Modelling, simulation, and inferring regulatory networks
Ashyraliyev, M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Summary

The research described in this thesis is devoted to different aspects of modelling, simulation, and inferring regulatory networks. We have considered here only deterministic models given by systems of differential equations that are capable of quantitatively reproducing spatio-temporal gene expression patterns.

In the first half of the thesis we have applied a ‘connectionist’ model given by a system of nonlinear Ordinary Differential Equations for simulating the gap gene network in the early development of the fruit fly *Drosophila melanogaster*. Using this model, the gap gene network has been extensively studied in the literature. However, in all previous studies the main attention has been focused on the estimation of model parameters, among which are most important the regulatory weights representing a regulatory influence of one gene on another, and subsequently on the analysis of the functioning of the gap gene system based on the values of estimated regulatory parameters. The identifiability analysis of inferred parameters has been missing and therefore, the reliability of all previous findings has remained unclear. We have tried to fill this gap by applying a posteriori identifiability analysis to assess the quality of the obtained parameters and studying its implications for conclusions deduced from the values of parameter estimates.

Similar to previous studies of the gap gene system, in Chapter 2 we have considered a 6-gene network consisting of gap genes *hunchback* (*hb*), *Krüppel* (*Kr*), *knirps* (*kni*), and *giant* (*gt*), terminal gap gene *tailless* (*tll*), and maternal coordinate gene *caudal* (*cad*), while the other maternal gene *bicoid* (*bcd*) has been implemented as an external input constant in time. The identifiability analysis of inferred regulatory weights has shown that none of them can be determined quantitatively with reasonable accuracy. Although we have been able to draw reliable qualitative conclusions for some of the regulatory weights, it has been found that many other interactions cannot be determined even qualitatively. So, the regulatory topology of the gap gene network deduced by only considering the values of parameter estimates has been confirmed only partially with the parameter determinability analysis. We have illustrated that an overall poor determinability of regulatory parameters is due to the presence of correlations between them. We have shown that these correlations are a property of the model, rather than being originated from the data.

In Chapter 3, we have considered a 4-gene network including gap genes *hb*,
In contrast to the 6-gene network, we have implemented $bcd$, $cad$, and $tll$ as time-variable external inputs. Moreover, we have supplemented the gap gene network with the terminal gap gene $huckebein$ ($hkb$) also implemented as time-variable external input. The results obtained with this reformulated gap gene network have provided a number of improvements in comparison with the results for the 6-gene network. Firstly, it refers to a correct regulation of the posterior $hb$ domain. We have shown that the posterior boundary of this domain is set up correctly and its anterior shift in time is reproduced by model outputs. Secondly, the identifiability analysis has revealed a significant improvement in the qualitative determinability of the regulatory weights.

In Chapter 3, we have also demonstrated that with available gap gene data the Weighted Least Squares sum with appropriately chosen weights is a more suitable measure for data fitting than the Ordinary Least Squares sum which has been used in all previous studies. This has been confirmed by a better fit of the boundaries of the gap gene expression domains, an absence of patterning defects in the model outputs, more reliable qualitative conclusions for a number of regulatory weights, and correct prediction of gap gene expression in $tll$ and $hkb$ mutants.

The cell-based model for simulating regulatory networks is given by a reaction-diffusion system with singular reaction source terms. Each source term is defined by a Dirac delta function expression on a lower dimensional surface. In Chapter 4, we have numerically studied this type of problems. Due to singularities, their solutions are not differentiable and this lack of smoothness causes order reduction when standard spatial discretization schemes are used on the uniform grid. We have used the finite volume approach based on the integral form such that the numerical treatment of the singular source terms is mathematically clear. We have demonstrated the reduction from order two to order one in the maximum norm when the standard second-order spatial discretization scheme on the uniform grid is applied for a number of 1D and 2D problems with singular source terms. To overcome this reduction, we have examined the discretization on a number of special locally refined grids, in 1D analytically and in 2D experimentally. We have shown that by an appropriate locally refined grid the maximum norm second-order convergence can be regained.

The model of regulatory networks incorporating a delay in the protein production is given by a system of Delay Differential Equations with a right-hand side being discontinuous in time and time-lag parameters. In Chapter 5, we have studied this type of problems. Solutions of such problems have a lack of smoothness in parameters, notably the derivatives with respect to parameters (gradients) can be discontinuous. As a consequence, the correct application of gradient-based optimization methods for parameter estimation as well as the validity of parameter determinability analysis applied on the parameter estimates can be questionable. In order to overcome these difficulties, we have examined a standard regularization technique to make the right-hand side of the model continuous at an $\epsilon$-neighborhood of the discontinuity points. We have proven analytically the
convergence of the solution of the regularized model to the solution of the original problem as $\epsilon \to 0$. Moreover, we have derived the rate of convergence. This result implies that the parameter estimates inferred from the regularized model converge to the corresponding estimates of the original problem as $\epsilon \to 0$. Additionally, we have shown that the convergence results do not depend explicitly on the way the model is regularized. We have supported our findings with numerical illustrations for simple test problems.