2010 International Symposium on Nonlinear Theory and its Applications
NOLTA2010, Krakow, Poland, September 5-8, 2010

Generalized modeling of heterogeneous nonlinear networks

Thilo Gross, Dirk Stiefs, Lars Rudolf, and Martin Zumsande

Max-Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, 01187 Dresden, Germany.

Abstract—Generalized modeling (GM) can be used to explore the dynamics of large, complex networks of nonlinear dynamical elements. Importantly, the interaction terms in the corresponding dynamical systems do not need to be fixed but can be expressed as general (i.e., unspecified) functions, describing a large class of different conventional models. A normalization procedure is then used to derive the Jacobian matrix governing the dynamics close to all steady states in the whole class of models. By GM the Jacobian matrices can be expressed as functions of parameters that have a well defined interpretation in the context of the respective application. By an ecological example, we show how GM can be used to obtain meaningful information on the dynamics on networks comprising 50 variables and thousands of parameters.

1. Introduction

Many systems in nature, ranging from the biochemical reactions inside cells to societies and ecosystems can be described as networks of interacting factors [1, 2]. Today the structure of many of these networks is known or purported, offering an opportunity to gain detailed insights in their functioning and failure. However, the functioning of a complex regulatory networks is often linked more closely to the network's dynamics than to the structure [3]. Extracting the dynamics supported by a network with given structure is therefore a central goal for theory [4–9].

Here we focus on the approach of generalized modeling (GM) by which certain features of the dynamics of a system can be extracted efficiently and robustly. GM was first applied in [10] and was subsequently proposed as a general method for the analysis of nonlinear systems in [11]. Subsequently, it was applied to a wide range of different topics including cell signaling [12–14], metabolism [15–18], ecology [11, 12, 19–27], laser physics[11], epidemiology [28] and history [11, 12]. In systems of ordinary differential equations [11] and partial differential equations [25], GM can determine the stability of steady states, detect the local bifurcations in which the stability is lost, identify parameter regions where complex dynamics are likely [21], and also has potential applications in model reduction [27].

2. Modeling approaches: a motivation

Investigation of the dynamics of regulatory networks faces three major obstacles: The networks of interest are typically large and heterogeneous, with network nodes corresponding, for instance, to metabolites with very different chemical properties. Furthermore, the mathematical functions describing processes in the network are often strongly nonlinear and cause dynamics on many different time scales. Finally, there is typically a high degree of uncertainty, which is reflected in most models by a (potentially large) number of unknown parameters.

For the analysis of biological networks, methods from network science and spectral graph theory have recently received much attention. These methods consider the network as an abstract graph and use statistical properties such as the number of connections, the occurrence of certain network motives, and the length and path of shortest connections between pairs of nodes [1]. Although providing certain general insights into the dynamics of the network, this approach cannot easily make use of specific biological insights, such as the heterogeneity of nodes.

For incorporating specific knowledge and formulating detailed predictions and hypotheses, networks are typically modeled as a dynamical systems, where each network node corresponds to a dynamical variable. The time evolution of these variables is then governed by a system of ordinary differential equations (ODEs). The task of analyzing the dynamics of a complex network is thus mapped to the analysis of a large nonlinear dynamical system.

By far the most common approach to model analysis is simulation. However, for the investigation of the long term behavior this is fundamentally inefficient because detailed information on the transient dynamics, which is later discarded, is obtained at a high computational cost.

Because of the high degree of uncertainty involved in most models, one would ideally wish for an analytical method that can reveal insights without requiring the researcher to fix the parameters to specific values. Indeed, applying the tools of dynamical systems theory [29] can reveal the bifurcation points, i.e., the critical points in parameter space, where qualitative changes in the long-term dynamics occur. This *conventional approach* to model analysis starts typically with com-

putation the steady states of the model, where all variables remain constant in time. After a small perturbation from the steady state, a system may either return asymptotically or depart entirely to approach a different attractor. In the former (latter) case the steady state is called asymptotically stable (unstable). The stability of a given steady state is analyzed by computing the Jacobian matrix, which constitutes a local linearization of the dynamical system. The steady state is asymptotically stable if all eigenvalues of the corresponding Jacobian have negative real parts. If changes in parameters cause one or more eigenvalues to acquire positive real parts the stability of the steady state is lost in a local bifurcation. Investigating the Jacobian matrices corresponding to the steady states of the model can therefore analytically reveal the critical parameter values at which the system departs from stationary behavior.

Beyond the basic steps described above further tools of dynamical systems theory may reveal boundaries of more complex dynamics, such as bifurcations of limit cycles and tori. However, even the basic steps outlined above can present a considerable challenge. In particular the first step, the computation of steady states can be prohibitively difficult if the system contains more than three or four dynamical equations. Even numerical extensions of the analytical procedure frequently fail in the analysis of heterogeneous dynamical networks with more than, say 20, nodes.

The mathematical difficulties, of the conventional approach, which mainly arise from the computation of steady states, are circumvented in *random matrix models* [30]. The central idea of this approach is that the Jacobian of a sufficiently complex dynamical system can be modeled as a random matrix. Because of their simplicity random matrix models can often be investigated analytically even for large systems. Moreover, a random matrix model does not require the researcher to restrict the underlying processes to a specific functional forms. On the one hand, formulating a random matrix model thus requires less assumptions and can thereby provide more robust insights. On the other hand, the abstract nature of random matrix models makes answering specific questions often very difficult.

3. Generalized Modeling

GM offers an intermediate way between conventional and random matrix models. Specifically, GM comes close to the generality and efficiency of random matrix models, while offering interpretability comparable to conventional models. Below we illustrate the approach of GM by discussing the key ingredients of a general food web model that has recently been investigated [11, 24].

As a first step consider just a single biological population X, changing in time due to biological reproduction S and mortality M, leading to the dynamical system

$$\frac{\mathrm{d}}{\mathrm{dt}}X = S(X) - M(X). \tag{1}$$

Because our focus is on systems where little information is available, we avoid restricting S and M to specific functional forms. The aim of our analysis is to determine the conditions under which a steady state in the system is stable. Although we cannot compute steady states at the desired level of generality, we can formally denote a steady state under consideration as X^*. We further denote the rates of the two processes in X^* as $M^* = S^*$. We then normalize the system by introducing $x = X/X^*$, $m(x) = M(X)/M^*$, and $s(x) = S(X)/S^*$. Using the normalized variables and functions the model can be written as

$$\frac{\mathrm{d}}{\mathrm{dt}}x = \alpha(s(x) - m(x)), \tag{2}$$

where $\alpha = S^*/X^*$. In the normalized variables the steady state under consideration is at $x^* = 1$, we can therefore write the corresponding Jacobian as

$$\mathbf{J} = \alpha(s_\mathrm{x} - m_\mathrm{x}), \tag{3}$$

where $s_\mathrm{x} = \partial s(x)/\partial x|_1$ and $m_\mathrm{x} = \partial m(x)/\partial x|_1$ are coefficients from the linearization.

So far we have succeeded in writing the Jacobian corresponding to an arbitrary steady state of the system as a function of the constants s_x, m_x, and α. The central insight of GM is that these constants can be treated as unknown parameters and have in general a well-defined meaning in the context of the model: The parameter α denotes the per-capita growth and mortality rate in the steady state. It is therefore simply the inverse of the life expectancy of an individual in the population. The parameters s_x and m_x are so-called elasticities, logarithmic derivatives of the original functions in the steady state. For instance $s_\mathrm{x} = \partial \log S/\partial \log X|_*$. This implies that if S is any power-law AX^p, then the corresponding parameter is $s_\mathrm{x} = p$. In contrast to conventional parameters, e.g. half-saturation constants, elasticities can be directly measured in data observed in the steady state and do not require reference to an artificial state (e.g. the half-saturation point).

In the one-dimensional example considered here one can directly read off the single eigenvalue, $\lambda = \alpha(s_\mathrm{x} - m_\mathrm{x})$, of the Jacobian. We can therefore conclude that in every system of the form of Eq. 1 every given steady state is stable if the elasticity of the mortality in the steady state exceeds the elasticity of the reproduction.

Essentially the same procedure applied above can also be used to study much larger sytems. In the remainder of this paper we discuss three complications

that may arise in such larger systems. The most harmless of these is encountered when equations of motion contain more than two terms. Consider for instance the example

$$\frac{\mathrm{d}}{\mathrm{dt}}X = S(X) - M(X) - F(X, W), \tag{4}$$

which is analogous to Eq. 1, except that we have included an additional loss term describing predation by a predator W. Typically, W will follow its own equation of motion, which we ignore here. Applying essentially the same procedure as above, we find the normalized system

$$\frac{\mathrm{d}}{\mathrm{dt}}x = \alpha(s(x) - \beta m(x) - \bar{\beta} f(x, w)) \tag{5}$$

where $\alpha = S^*/X^* = (M^* + F^*)/X^*$, $\beta = M^*/(M^* + F^*)$, and $\bar{\beta} = F^*/(M^* + F^*)$. Here, we had to introduce the new parameter β and its complement $\bar{\beta} = 1 - \beta$ weighting the different types of losses. In words, $\bar{\beta}$ is the probability that an individual will eventually be eaten and β the probability that it will live until its natural death.

In general, the normalization of a dynamical equation containing N terms will require introducing $N-1$ parameters. It is generally advantageous to introduce one parameter (α) describing the per-capita turnover rate and $N-2$ parameters (β) weighting the contributions to the total gain and loss rates, respectively.

A second more subtle complication is for instance encountered in a model where X is a predator feeding on two prey populations Y and Z such that

$$\frac{\mathrm{d}}{\mathrm{dt}}X = S(X, Y, Z) - M(X) \tag{6}$$

When we carry out the normalization procedure, we end up with a Jacobian containing the elasticities $s_\mathrm{y} = \partial s/\partial y$ and $s_\mathrm{z} = \partial s/\partial z$ describing the sensitivity of the predation rate to the size of the prey populations. However, ecological knowledge may tell us that the two elasticities are not unrelated. For instance if one of the two species were very abundant then the sensitivity of the predation to the size of the rarer prey population is greatly reduced because the predator is likely to be saturated from a encounters with the abundant one.

Insights as the one described above can be integrated into a previously derived GM by an iterative refinement procedure. Suppose for instance that beyond what is stated in Eq. (6) we know predation to depend only on the sum of the sizes of the two prey populations, i.e. $S(X, Y, Z) = S(X, T)$, where $T = Y + Z$. The algebraic equation for the total amount of prey, T, can be normalized like the differential equation, yielding $t = \gamma y + \bar{\gamma} z$, where the parameters $\gamma = Y^*/(Y^* + Z^*)$, $\bar{\gamma} = 1 - \gamma$ measure the relative contributions of the two prey species. Using this new relation we can now write

$$g_\mathrm{y} = \frac{\partial s}{\partial y} = \frac{\partial t}{\partial y}\frac{\partial s}{\partial t} = \gamma g_\mathrm{t} \tag{7}$$

and analogously $g_\mathrm{z} = \bar{\gamma} g_\mathrm{t}$, where g_t is now the elasticity of the predation rate with respect to T. By substituting the equations for g_y and g_z we can rewrite a previously derived Jacobian in terms of the new parameters incorporating the additional insight on the dependence of g_y and g_z. Although this refinement does not generally reduce the number of parameters in a GM it often facilitates the interpretation of results.

The final and perhaps most obvious complication encountered in larger models is having more dynamical variables and hence more equations of motion and larger Jacobians. In this case the normalization procedure is applied to all of the equations of motion. For large systems containing tens or even hundreds of equations, the manual work can be reduced by using the matrix formalism proposed in [17]. When dealing with GMs of large systems the main challenge is therefore to extract information from large Jacobians. Using the method described in [10] analytical computation of the local bifurcations is feasible for systems of up to 10 dynamical variables. Furthermore, a numerical procedure that is applicable to larger system is described in [12].

In large systems identifying the decisive parameters having a strong impact on stability can be a challenging task. In GMs this is often accomplished by a Monte Carlo sampling of the parameter space: We create a large ensemble of parameter sets, where each parameter in each set is randomly chosen. In the second step we compute the stability of the steady states corresponding to the random parameter sets by substituting one set at a time into the Jacobian and numerically computing the leading eigenvalue of the matrix that is obtained. To each parameter set we assign a stability value q_i, which is 1=stable if the leading eigenvalue is negative and 0=unstable otherwise. We can then estimate the impact of a given parameter, say m_x, on stability by computing the correlation between the values of m_x and q_i in the ensemble. More detailed insights can be gained for instance by plotting a histogram of the fraction of stable states found, over one of the parameters [17, 24].

The analysis described above profits greatly from the efficiency of GMs. Because the computation of the leading eigenvalue of a matrix is much faster than, for instance, the simulation of the corresponding dynamical system very large ensembles can be studied. For instance in [24] approx. 10^{11} sample parameter sets were used for analyzing a 50-dimensional systems containing several thousand parameters.

References

[1] A.-L. Barabasi and Z. Oltvai. "Network biology: Understanding the cell's functional organization". *Nat. Rev. Gen.*, vol. 5, pp. 101, 2004.

[2] P. Turchin. *Historical dynamics: why states rise and fall.* Princeton University Press, 2003.

[3] J. J. Tyson. "Bringing cartoons to life". *Nature*, vol. 445, p. 823, 2007.

[4] K. Schallau and B. H. Junker. "Simulating Plant Metabolic Pathways with Enzyme-Kinetic Models". *Plant Phys.*, vol. 152, pp. 1763–1771, 2010.

[5] A. Rodriguez and D. Infante. "Network models in the study of metabolism". *Electr. J. Biotech.*, vol. 12, pp. 11-12, 2009.

[6] N. Jamshidi and B. O. Palsson. "Formulating genome-scale kinetic models in the post-genome era". *Mol. Sys. Biol.*, vol. 4, pp. 171, 2008.

[7] R. Steuer and B. H. Junker. "Computational models of metabolism". *Adv. Chem. Phys.*, vol. 142 pp. 105–251, 2009.

[8] L. J. Sweetlove, D. Fell and A. R. Fernie. "Getting to grips with the plant metabolic network". *Biochem. J.*, vol. 409, pp. 27–41, 2008.

[9] R. Steuer. "Computational approaches to the topology, stability and dynamics of metabolic networks". *Phytochemistry*, vol. 68, pp. 2139, 2007.

[10] T. Gross and U. Feudel. "Analytical search for bifurcation surfaces in parameter space". *Physica D*, vol. 195, pp. 292–302, 2004.

[11] T. Gross and U. Feudel. "Generalized modeling as a universal approach to nonlinear dynamical systems". *Phys. Rev. E*, vol. 73, 016205, 2006.

[12] D. Stiefs, T. Gross, R. Steuer and U. Feudel. "Computation and visualization of bifurcation surfaces". *Int. J. Bif. Chaos*, vol. 18, pp. 2191, 2008.

[13] M. Zumsande and T. Gross. "Bifurcations and chaos in the mapk signaling cascade". to appear in *J. Theor. Biol.*, 2010. doi: 10.1016/j.jtbi.2010.04.025.

[14] A. de la Fuente, G. Fotia, F. Maggio, G. Mancosu and E. Pieroni. "Insights into biological information processing". *J. Phys. A*, vol. 41, 224013, 2008.

[15] S. Bulik, S. Grimbs, C. Huthmacher, J. Selbig and H. G. Holzhuetter. "Kinetic hybrid models composed of mechanistic and simplified enzymatic rate laws". *FEBS*, vol. 276, pp. 410–424, 2009.

[16] R. Steuer, A. N. Nesi, A. R. Fernie, T. Gross, B. Blasius and J. Selbig. "From structure to dynamics of metabolic pathways". *Bioinformatics*, vol. 23, pp. 1378–1385, 2007.

[17] R. Steuer, T. Gross, J. Selbig and B. Blasius. "Structural kinetic modeling of metabolic networks". *PNAS*, vol. 103, pp. 11868–11873, 2006.

[18] S. Grimbs, J. Selbig, S. Bulik, H.-G. Holzhuetter and R. Steuer. "The stability and robustness of metabolic states". *Mol. Sys. Biol.*, vol. 3, pp. 146, 2007.

[19] T. Gross. *Handbook of Biological Networks*, pp. 417–436. World Scientific, Singapore, 2009.

[20] T. Gross, W. Ebenhöh and U. Feudel. "Enrichment and foodchain stability". *J. Theor. Biol.*, vol. 227, pp. 349–358, 2004.

[21] T. Gross, W. Ebenhöh and U. Feudel. "Long food chains are in general chaotic". *Oikos*, vol. 109, pp. 135–155, 2005.

[22] U. Feudel, M. Baurmann and T. Gross. *Irreversible Prozesse und Selbstorganization*, pp. 303–311. Logos, Berlin, 2006.

[23] G.A.K. Van Voorn, D. Stiefs, T. Gross, B.W. Kooi, U. Feudel and S.A.L.M. Kooijman. "Stabilization due to predator interference". *Math. Biosci. Eng.*, vol. 5, pp. 567–583, 2008.

[24] T. Gross, L. Rudolf, S. A. Levin and U. Dieckmann. "Generalized models Reveal Stabilizing Factors in Food Webs". *Science*, vol. 325, pp. 747–750, 2009.

[25] M. Baurmann, T. Gross and U. Feudel. "Instabilities in spatially extended predator-prey systems". *J. Theor. Biol.*, vol. 245, pp. 220–229, 2007.

[26] D. Stiefs, G. V. Voorn, B. Kooi, U. Feudel and T. Gross. "Food quality in producer-grazer models". to appear in *Am. Nat.* , 2010.

[27] T. Gross and U. Feudel. "Local dynamical equivalence of certain food webs". *Ocean Dynamics*, vol. 59, pp. 417–427, 2009.

[28] D. Stiefs, E. Venturino and U. Feudel. "Evidence of chaos in eco-epidemic models". *Math. Biosci. Eng.*, vol. 6, pp. 855–871, 2009.

[29] Y. A. Kuznetsov. *Elements of Applied Bifurcation Theory.* Springer-Verlag, New York, 2004.

[30] R. M. May. "Will a Large Complex System be Stable". *Nature*, vol. 238, pp. 413–&, 1972.