Topology of biological networks and reliability of information processing

Konstantin Klemm†‡ and Stefan Bornholdt†§

†Interdisciplinary Center for Bioinformatics, University of Leipzig, Kreuzstr. 7b, D-04103 Leipzig, Germany, ‡Bioinformatics Group, Dept. of Computer Science, University of Leipzig, Kreuzstr. 7b, D-04103 Leipzig, Germany, §Institute for Theoretical Physics, University of Bremen, Otto-Hahn-Allee, D-28359 Bremen, Germany

(Dated: February 9, 2008)

Biological systems rely on robust internal information processing: Survival depends on highly reproducible dynamics of regulatory processes. Biological information processing elements, however, are intrinsically noisy (genetic switches, neurons, etc.). Such noise poses severe stability problems to system behavior as it tends to desynchronize system dynamics (e.g. via fluctuating response or transmission time of the elements). Synchronicity in parallel information processing is not readily sustained in the absence of a central clock. Here we analyze the influence of topology on synchronicity in networks of autonomous noisy elements. In numerical and analytical studies we find a clear distinction between non-reliable and reliable dynamical attractors, depending on the topology of the circuit. In the reliable cases, synchronicity is sustained, while in the unreliable scenario, fluctuating responses of single elements can gradually desynchronize the system, leading to non-reproducible behavior. We find that the fraction of reliable dynamical attractors strongly correlates with the underlying circuitry. Our model suggests that the observed motif structure of biological signaling networks is shaped by the biological requirement for reproducibility of attractors.

I. INTRODUCTION

The processes of Life in living cells and organisms rely on highly reproducible information processing. It has been a long-standing question how this is accomplished even though it involves elements with non-reproducible, noisy dynamics [1]. In nerve cells, for instance, firing of spikes is not fully determined by synaptic input [2]. Similarly in gene regulation, protein concentration evolves in a relatively irreproducible manner under given promoter levels [2]. In a steady state, fluctuations can be dampened by the properties of the single elements as, e.g., autoregulation attenuates noise in genetic transcription [4]. In a dynamical scenario with rising and falling levels of activation, intrinsic noise of the interacting elements causes fluctuations of switching delays [3]. These cause stability problems in larger systems of switching elements, as timing and coordination may fail. In order to generate reproducible dynamics, the non-deterministic responses of single elements have to be compensated for by a suitable circuit design.

During recent years, biochemical interactions involved in the information processing within or between cells of an organism have been characterized systematically. For many systems of sizes up to a few hundred constituents a significantly large fraction of interactions have been identified such that the network-like structure of the system becomes visible. Such networks have been obtained for gene regulation [1, 3], signal transduction (See for instance http://stke.sciencemag.org), and neuronal information processing [8]. Despite their different functional origins, these systems show strong similarity in their graph representation as nodes and interactions. In [3], local wiring structure has been analyzed for several network examples in terms of abundance of connected subgraphs. Comparing with randomly wired graphs, real networks typically exhibit significant abundance of some subgraphs (called motifs), while others are strongly suppressed [10]. Networks of signal transduction and transcription and neural networks share similar statistics of local wiring patterns [2]. A vastly differing statistics of subgraphs is observed in non-signaling networks from different areas, such as the World Wide Web, social acquaintance webs and the graph of word-adjacency in various languages.

In this paper, we argue that biological signaling networks (as gene networks or neural networks) are shaped by the common requirement of robust signal processing. We study reliability of dynamics in small networks of information processing units with fluctuating response times. Starting with the simultaneous oscillation of two mutually coupled nodes we already observe that there are two distinct classes of dynamics. In reliable dynamics the nodes regain synchronization after a perturbation (retarded response) while for unreliable dynamics the system does not have a tendency to self-synchronize. This distinction appears in descriptions at different levels of abstraction. The differential equations for continuous variables (macromolecule concentrations or firing rates, respectively) yield the same synchronization properties as a model with discrete binary state variables. Turning to networks with three nodes we find that the occurrence of reliable dynamics is highly correlated with the underlying topology. One observes that reliable dynamics is more likely to appear on those triads that have been found as building blocks (motifs) in real biological networks. Further insight into the relationship between topology and reliability is gained by the analysis of cyclic behavior (attractors). The dynamics are reliable only if all switching events are connected by a single causal chain. An espe-
cially instructive case are isolated feedback loops. Here the fraction of initial conditions with reliable dynamics is obtained easily. In the following, let us start with the simplest example of two coupled nodes.

II. FEEDBACK LOOPS OF TWO NODES

The distinction between reproducible and non-reproducible dynamics becomes obvious already in the setting of two mutually coupled nodes. Two basic scenarios are to be distinguished: (i) Nodes influence each other with the same sign. For instance node 2 represses node 1 while node 1 activates node 2. (ii) Both nodes have the same coupling, i.e., either both couplings are activating or both are repressing.

Starting with scenario (i), let us describe the dynamics of the system by rate equations for normalized concentrations (firing rates) \( u, v \in [0, 1] \) of the two nodes

\[
\dot{u}(t) = \frac{\alpha_1 v_h(t - \tau_1)}{k_1^h + v_h(t - \tau_1)} - \beta_1 u(t) \quad (1)
\]

\[
\dot{v}(t) = \frac{\alpha_2 k_2^h}{k_2^h + u_h(t - \tau_2)} - \beta_2 v(t) . \quad (2)
\]

The time constants \( \alpha_i \) and \( \beta_i \) are the rates of production and degradation of messenger substance, respectively. In the neural system they represent the time constants of the varying firing rate. The non-linear production term with Hill exponent \( h \) describes collective chemical effects or a non-linear neural response function, respectively. Note that the rate equations involve explicit delay times \( \tau_1 \) and \( \tau_2 \) for signal transmission. By randomly varying these delay variables in time, an uncertainty in response times of the signal processing elements is implemented.

Figure 1 shows time series for the above system under varying transmission delay \( \tau_1 \). The variables \( u(t) \) and \( v(t) \) oscillate with a time lag that depends on the current transmission delay \( \tau_1 \). Note that (a) the dynamics remains qualitatively the same independently of \( \tau_1 \), as \( v(t) \) always lags behind \( u(t) \) by about 1/4 of an oscillation period; (b) the current phase lag does not depend on the history of the system. The system retains synchronization irrespective of the changes in the transmission delay.

Approximating the sigmoid response curves by step functions and considering the limit of fast production and degradation \( \alpha_i, \beta_i \to \infty \) in Eqs. (1) and (2) we obtain the simplified dynamics

\[
u(t) = v(t - \tau_1) \quad (3)
\]

\[
v(t) = 1 - u(t - \tau_2) \quad (4)
\]

where now we have binary state variables \( u, v \in \{0, 1\} \).

Fig. 1 shows that this approximate description has qualitatively the same behavior as the original, continuous rate equations. Let us now turn to case (ii) above. Consider two mutually coupled nodes with the same type of coupling in both directions. Here we take both couplings to be inhibitory. Again we describe the system in terms of rate equations

\[
\dot{u}(t) = \frac{\alpha_1 k_1^h}{k_1^h + v_h(t - \tau_1)} - \beta_1 u(t) \quad (5)
\]

\[
\dot{v}(t) = \frac{\alpha_2 k_2^h}{k_2^h + u_h(t - \tau_2)} - \beta_2 v(t) . \quad (6)
\]

The approximation by binary state variables reads

\[
u(t) = 1 - v(t - \tau) \quad (7)
\]

\[
v(t) = 1 - u(t - \tau) . \quad (8)
\]

The time series for both the continuous and the binary system are plotted in Fig. 2. Initially, the variables \( u(t) \) and \( v(t) \) are perfectly synchronized, due to identical initial values. Changing the transmission delay \( \tau_1 \) causes a non-zero phase lag, that is retained even when the transmission delay is reset to the original value. In fact, varying \( \tau_1 \) inside a small interval causes accumulation of phase lag between \( u(t) \) and \( v(t) \). The system does not self-synchronize and may therefore be driven out of phase by slight fluctuations of transmission delay.

Consequently, system (i) with antisymmetric interaction and system (ii) with symmetric interactions differ with respect to reproducibility, see Fig. 3. With transmission delays varying randomly with time and across

FIG. 1: Dynamics of the feedback loop with antisymmetric coupling. (a) Fluctuations of transmission delay \( \tau_1 \). (b) Evolution of phase lag between \( u(t) \) and \( v(t) \) for the continuous case (closed circles) and the discrete approximation (open triangles). Cf. Eqs. (1), (2) and (3), (4), respectively. (c) The phase lag is defined as the time difference between subsequent passages of variables \( u \) (dashed curve) and \( v \) (solid curve) across the value 1/2 from below, as indicated by the black double arrows in the lower panel. Parameters in Eqs. (1) and (2) have values \( \tau_2 = 0.5, \alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 10, k_1 = k_2 = 10^{1/2}, h = 2 \). Broad variations of these parameters give qualitatively the same results.
several runs, system (i) shows reproducible behavior. The phase lag practically remains constant, as the oscillation with a given phase lag is the stable (attractive) mode. In contrast, the behavior of system (ii) varies across runs. Synchronized oscillation is a marginally stable mode of this system such that the fluctuations drive the system away from this mode.

III. THREE NODES

In the following we shall see that the clear distinction between reproducible dynamics with intrinsically stable synchronicity and non-reproducible dynamics sensitive to fluctuations extends beyond simple oscillations in systems of two nodes. Let us study dynamics on 3-node circuits as shown in the top row of Fig. 4. To this end, we first need to define the dynamics of nodes with more than one input. Consider a node $i$ that is directly influenced by nodes $j$ and $k$. Restricting ourselves to the binary approximation of states from now, node $i$ switches according to

$$x_i(t) = f_i[x_j(t - \tau_i), x_k(t - \tau_i)]$$  \hspace{1cm} (9)

where we define $\tau_i$ as the (time-dependent) transmission delay of node $i$. The Boolean function $f_i$ maps the 4 pairs of binary states $(x_j, x_k)$ to the set $\{0,1\}$. We choose $f_i$ from the set of canalizing functions $\mathbb{1}$, i.e. we do not use the exclusive-or function and its negation. From the 14 canalizing functions we further exclude those that are constant with respect to one or both of the inputs. We are then left with 8 Boolean functions of two inputs (these are OR, AND and their variants generated by negation of one or both inputs). For nodes with one input, we allow only the non-constant Boolean functions Identity (output = input) and Negation (output $\neq$ input).

As subgraph reliability we define the probability of obtaining reliable dynamics when preparing a random initial condition and a random assignment of Boolean functions. The strict subgraph reliability is the probability that all initial conditions yield reliable dynamics for a random assignment of Boolean functions. The small system size allows us to obtain the exact values of these quantities by full enumeration of all combinations of initial conditions and function assignments. See "Methods" for details on the simulation procedure.

Viewing the results in Fig. 4 (bottom), we note first that feed-forward wiring (subgraphs 1, 2, 3 and 7) always gives reliable dynamics because from any initial condition the system reaches a fixed point after a short time. The least reproducible dynamics is obtained for the pure feedback loop (subgraph 8) with strict reliability zero. For an explanation see the section on “Feedback loops of arbitrary length”. Interestingly, the feedback loop with an additional link in the “opposite” direction (subgraph 11) has a larger strict reliability 1/2.

Now let us compare the dynamical reliability of the subgraphs with their abundance in biological signaling networks. Consider first subgraphs 7–13, which contain a closed triad. Among these, subgraph 7 (the feed-forward loop) and subgraph 9 (two bidirectionally coupled nodes receiving from a common third node) have particularly large reliability. These subgraphs are also observed to be highly abundant in the signaling networks. Furthermore, subgraph 12 has larger reliability than subgraphs 11 and 13. This superiority is also reflected by abundance measure in the networks. The large abundance (quantified...
FIG. 4: Abundance and dynamical reliability of subgraphs. Top: All directed connected graphs of three nodes. Center: Triad significance profile [9] in various signaling networks. Bottom: Graph reliability (squares) and strict graph reliability (circles) as defined in the main text. Triad significance profiles in the center panel are given for the signal-transduction in mammalian cells (squares), genetic transcription for development in the fruit fly (dots) and the sea-urchin (triangles), synaptic connections in C. elegans (circles). Triad significance is the difference between numbers of occurrence in the given network and in an ensemble of randomly rewired surrogate networks, divided by the standard deviation. The triad significance profile is the normalized vector of triad significances. See reference [9] for details.

by the Z-score measure as defined in [9] of subgraph 10, however, cannot be predicted from the dynamical reliability. For the remaining subgraphs 1–6, wiring diagrams without closed triads, there is a clear correlation between the empirical Z score and our measure of dynamical reliability.

The presence or absence of a given subgraph cannot be fully explained by the reliability measure presented here. First, the Z score tends to increase with the number of connections in the subgraph. The bias towards densely (but not fully) connected motifs is a consequence of the networks’ modular structure with functional clusters of nodes [12]. This property of the networks is not directly related to robust dynamics and therefore is not reflected in our reliability measure. Second, we have neglected the network environment of the motifs, that may greatly change the dynamics. For instance, the feed-forward subgraph 3 alone yields perfectly reliable dynamics. The 4-node feedback loop 1 → 2 → 3 → 4 → 1, however, does not give reliable dynamics (as shown below), even though subgraph 3 is its only motif. Third, we have defined reliability of a subgraph by averaging over all assignments of plausible functions. In reality, however, the nodes’ functions may be correlated with the wiring diagram. For instance, subgraph 10 gives perfectly reliable dynamics if one assumes antisymmetric influence (one promoter, one repressor) between the bidirectionally linked nodes.

Despite these limitations in the present analysis, the correlation between abundance and reliability of subgraphs in Fig. 4 is striking. In the following we gain more detailed insight into the mechanism leading to reliability by considering attractors in feedback loops.

IV. ATTRACTORS AND CAUSALITY

Even though in biological systems transmission delays fluctuate, it is instructive to regard constant transmission delays as a reference case. Setting \( \tau_i = 1 \) for all nodes \( i \) in Eq. (9), we recover the time-discrete synchronous update mode often employed in Boolean network models. In this idealized picture, all signal transmissions require exactly the same time and nodes flip synchronously as if driven by a central clock. The deterministic dynamics eventually reaches a periodic attractor — an indefinitely repeated sequence of network states as illustrated in Fig. 5.

FIG. 5: Attractor of a Boolean network. Left: A Boolean network with \( N = 3 \) nodes with possible states 0 (off) and 1 (on). Each node has a lookup table to determine its state given the input from other nodes. Here the two light nodes simply copy their single input, while the dark node performs the Boolean function NOR on its inputs. Right: In the deterministic case of constant transmission delays, the system always reaches a periodic reproducible sequence of states.

Reliability under fluctuating transmission delays affects the stability of this perfectly synchronized mode. When slightly perturbed by one retarded signal transmission, does the system autonomously re-establish synchronicity? Or does the system “remember” the perturbed timing? In the latter case, just as in Fig. 2, a series of perturbations may drive the system away from the predictable periodic behavior. The key to answering the question about stability is the causal structure of the attractor. Let us call a “flipping event” a pair \((i, t)\), given that node \( i \) changes state at time \( t \). For the attractor in
Fig. 6 we draw arrows from each flipping event to all flipping events it causes to happen in the next time step. In the resulting plot, Fig. 6(a), there is a single closed chain of flipping events. Retardation of one event by a time $t_{ret}$ simply retards all subsequent events by the same amount of time. Event $(i, t)$ in the unperturbed scenario becomes event $(i, t + t_{ret})$, but the sequence of states encountered by the system remains the same. This built-in compensation of fluctuations renders the dynamics reliable.

An example of an unreliable dynamical attractor is shown in Fig. 6(b). In this case there are two separate chains of flipping events, one connecting the on-events and the other connecting the off-events. Retarding an event in one of the chains does not influence the timing of the events in the other chain. By repeatedly retarded on-events the time span a node is on is gradually reduced and eventually reaches zero. Then the system encounters a fixed point and does not follow the attractor any longer. In general, the dynamics is reliable if and only if the attractor contains exactly one causal chain.

![FIG. 6: Causal structure of attractors. (a) The attractor shown in Fig. 5 has one causal chain of flipping events triggering each other. (b) A different attractor (on the same wiring diagram) with two independent causal chains of flipping events. The attractor in (b) is obtained after replacing the NOR in Fig. 5 with the Boolean function that gives 1 if and only if it receives 0 from the top node and 1 from the left node.](image)

V. FEEDBACK LOOPS OF ARBITRARY LENGTH

The relation between topology and reproducibility is particularly obvious in isolated feedback loops. Consider $N \geq 2$ nodes connected in a directed cycle, i.e., each node receives input only from its (clockwise) predecessor. If node $i = 1$ changes state (“flips”) at time $t$, node $i$ will change state at time $t + \tau_i$, node $i + 1$ will change state at time $t + \tau_i + \tau_{i+1}$ and so forth. The dynamics can be interpreted as “wave fronts” of flipping events traveling at constant speed on a ring. The nodes are located on this ring at distances given by the transmission delays $\tau_i$, as illustrated in Fig. 8. For constant transmission delays, the dynamics is periodic (with period $\tau_1 + \ldots + \tau_N$). However, when a transmission delay $\tau_i$ fluctuates, consecutive passages of wave fronts from node $i = 1$ to node $i$ take different times. Eventually one wave front may catch up with the other. Wave fronts annihilate upon encounter: Flipping from 0 to 1 and back to 0 at the same time results in no flipping at all. Annihilations of wave fronts happen stochastically as they are driven by the random fluctuations of transmission delays.

Consequently, the dynamics is reproducible only if annihilation of wave fronts is excluded. When all nodes perform the function Identity, as in Fig. 7, the number of wave fronts is even. Two or more wave fronts can annihilate, eventually leading to irreproducible dynamics. The only reproducible dynamics in this case is a system that stays on a fixed point, corresponding to zero wave fronts.

However, if one of the node performs the function Negation then this node acts as a resting wave front, because states on the two sides are always different. The total number of wave fronts (including the resting one) is still even, but now the number of traveling wave fronts is odd. Initial conditions exist such that there is a single traveling wave front, giving reproducible dynamics.

The two cases generalize easily. For an even number of inhibitory couplings (i.e., an even number of nodes performing Inversion) the dynamics is reliable if and only if one of the two fixed points is chosen as initial condition. Then the fraction of initial conditions with reproducible dynamics is

$$q_{even} = 2/2^N = 2^{N-1}. \quad (10)$$

for a feedback loop of $N$ nodes. Analogously we find

$$q_{odd} = 2N/2^N = N2^{N-1} \quad (11)$$

because in a feedback loop with an odd number of Negations, there are $2N$ initial conditions that generate a single wave front. Note that the only feedback loop that yields reproducible dynamics for all initial conditions has $N = 2$ nodes, one performing Inversion and the other Identity. This is the system studied as case (i) given by Eqs. 1 and 2.

![FIG. 7: The “wave front” picture of dynamics in a feedback loop with $N = 4$ nodes all performing the Boolean function Identity.](image)

VI. DISCUSSION AND CONCLUSIONS

Biological information processing systems are constrained by the need for reproducible output from circuits of intrinsically noisy elements. Compensation for
the effects of fluctuations on the system level can be achieved through the selection of a suitable circuit design. In the present work we have studied the influence of network topology on the reliability of information processing by networks of switches with fluctuating response times. Analytical considerations and simulation studies have shown that the dynamical scenarios occurring in the deterministic case fall into two classes in the presence of fluctuations. In reliable scenarios, the elements cooperatively suppress the fluctuations and tend to synchronize their operations. Unreliable scenarios, in contrast, do not favor synchronization and show irreproducible, desynchronized behavior when response times fluctuate.

In general, reliable and unreliable dynamical scenarios may be distinguished by considering the causal chains (cascades) of switching events in the network. A dynamics that can be separated into causal chains without a common checkpoint is unreliable. Even though they might be involving the same nodes, flipping events on different causal chains cannot synchronize.

The occurrence of the two dynamical classes is strongly biased by the topology. Whether or not the system shows reliable dynamics can to a large degree be deduced from the unlabeled wiring diagram without information on the type of couplings and functions of switches. To capture this connection between topology and dynamical robustness quantitatively we have defined graph reliability as the probability to obtain reliable dynamics on a given directed graph.

We have obtained the graph reliability for each of the 13 directed networks of three switching elements. Comparing with empirical networks of genetic transcription, signal transduction, and the nervous system we find that the statistics of the local wiring structure is closely related to the reliability measure. Reliable triads tend to occur significantly more frequently in natural networks compared to the randomized versions of the networks, while unreliable triads are typically suppressed when comparing empirical nets with their randomized counterparts.

The fact that without knowledge of functionality prominent features of the local wiring follow solely from the constraint of reliability, supports the hypothesis that noise effects are suppressed at the system level by selection of topology. Reversely, our approach might be used to predict function as the most reliable dynamics given an observed wiring pattern. For instance, we may infer that a feedback loop should contain an odd number of suppressing elements because then the dynamics is reliable more often than in the case of even number of suppressors.

Several extensions of this work are conceivable. In this first dynamical study of motifs we have neglected the network context by assuming all external input to be constant. Further work should drop this assumption and study larger groups of nodes containing a given subgraph to be tested. Another interesting outlook is a study of the dynamics on an empirical network and to compare its reliability with rewired counterparts. We expect that dynamical studies on biological network topologies will teach us about origin and function of these systems, even when still lacking full knowledge of network properties and dynamics.

VII. METHODS

Differential equations are integrated by first order Euler method using a time increment Δt = 10^{-5}. For systems with Boolean variables (Eqs. (3), (4), (7), (8), and (9)) integration is performed with continuous time t (exact up to machine precision). In the simulations in section III transmission delays τ_i are varied as follows: Whenever node i changes state, a new target delay time τ_i^* is drawn from the homogeneous distribution on the interval [0.9:1]. Then the delay τ_i approaches the target value τ_i^* according to τ_i = Δt \text{sgn}(τ_i^* − τ_i) (if τ_i were set to a new value abruptly, temporal order of signal transmission would not be conserved).

Criterion for reliability: For a given dynamical scenario (choice of Boolean functions and initial condition) 100 independent simulation runs of duration T = 1000 are performed and the sequence of sustained states is recorded. A sustained state is a state (x_1, ..., x_N) that is assumed by the system for time at least t_{sus} = 1/2. The given dynamical scenario on the given subgraph is called reproducible if the series of sustained states of all 100 runs are identical. Varying the number of runs and their duration by one order of magnitude does not change the set of reliable scenarios.

VIII. ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft DFG.

[1] Rao C.V., Wolf D.M., & Arkin A.P. (2002) Nature 420, 231.
[2] Allen C., Stevens C.F. (1994) Proc. Natl. Acad. Sci. USA 91, 10380.
[3] McAdams H.H., Arkin A. (1997) PNAS 94, 814-819.
[4] Becskei A., Serrano L. (2000), Nature 405, 590.
[5] Beierholm U., Nielsen C.D., Ryge J., Aelstrom P., Kiehn O. (2001), J. Neurophysiol. 86, 1858-1868.
[6] Thieffry D., Huerta A.M., P´erez-Rueda E., & Collado-Vides J. (1998), BioEssays 20, 433-440.
[7] Costanzo M.C., Crawford M.E., Hirschman J.E., Kranz J.E., Olsen P., Robertson L.S., Skrzypek M.S., Braun
B.R., Hopkins K.L., et al. (2001), Nucleic Acids Res. **29**, 75-79.

[8] White J.G., Southgate E., Thompson J.N., Brenner S. (1986), Phil. Trans. R. Soc. Lond., B **314** 1-340.

[9] Milo R., Itzkovitz S., Kashtan N., Levitt R., Shen-Orr S., Ayzenshtat I., Sheffer M., Alon R. (2004) Science **303**, 1538-1542 (2004).

[10] Milo R., Shen-Orr S., Itzkovitz S., Kashtan N., Chklovskii D., Alon U. (2002), Science **298**, 824-827.

[11] Kauffman S.A., Physica D **10**, 145 (1984)

[12] Milo R., Itzkovitz S., Kashtan N., Levitt R., Alon R. (2004) Science **305**, 1107d.

[13] Kauffman S., Peterson C., Samuelsson B., Troein C. (2003) PNAS **100**, 14796-14799.