Review

Discovering design principles for biological functionalities: Perspectives from systems biology

Priyan Bhattacharya[1,2,3] , Karthik Raman[2,3,4]* and
Arun K Tangirala[1,2,3]*

[1]*Department of Chemical Engineering, Indian Institute of Technology (IIT) Madras, Chennai 600 036, India*

[2]*Robert Bosch Centre for Data Science and Artificial Intelligence (RBCDSAI), IIT Madras, Chennai 600 036, India*

[3]*Initiative for Biological Systems Engineering (IBSE), IIT Madras, Chennai 600 036, India*

[4]*Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, IIT Madras, Chennai 600 036, India*

*Corresponding authors (Emails, kraman@iitm.ac.in; arunkt@iitm.ac.in)

Network architecture plays a crucial role in governing the dynamics of any biological network. Further, network structures have been shown to remain conserved across organisms for a given phenotype. Therefore, the mapping between network structures and the output functionality not only aids in understanding of biological systems but also finds application in synthetic biology and therapeutics. Based on the approaches involved, most of the efforts hitherto invested in this field can be classified into three broad categories, namely, computational efforts, rule-based methods and systems-theoretic approaches. The present review provides a qualitative and quantitative study of all three approaches in the light of three well-researched biological phenotypes, namely, oscillation, toggle switching, and adaptation. We also discuss the advantages, limitations, and future research scope for all three approaches along with their possible applications to other emergent properties of biological relevance.

Keywords. Adaptation; biological networks; bioswitches; computational approach; oscillator; rule-based; systems biology; systems theory

1. Introduction

Understanding biological systems requires a post-scriptum approach, which makes a significant part of the science essentially a reverse-engineering process. In this quest to find out the 'how' of 'parts leading to the whole', biologists often adopt a reductionist approach in which specific properties of a biological system under study are attributed to specific components, just

as Lazebnik ([2002](#)) suggested the role of MIC (most important component) in the regular operation of a radio transistor. Reducing every property of the system to one of its components leads to the conjecture that every module or constituent behaves identically in isolated and connected conditions. This is clearly not true for biological systems. Studies by Del Vecchio ([2013](#)) on loss of modularity and context-dependence showed that clock-generating, inter-cellular modules behave differently from their isolated versions when connected to a downstream system. Therefore, it is only natural to consider the manner of interconnections along with particular knowledge of the constituent

This article is part of the Topical Collection: Emergent dynamics of biological networks.

modules in order to understand (hence, design) biological networks (Hespanha and Sivakumar 2013). This necessitates a different treatment of biological entities that aims to understand the emergence of biological properties in a manner that provides equal importance to both the particular constituents and the interconnections (Voit 2012).

Systems biology explains any biological functionality as a response emanating from a biological system composed of several basic units such as proteins, metabolites, etc. (Konopka 2006). This formalism has inspired the systems-biology community to discover how the basic network units are connected for a specific biological functionality (Milo *et al.* 2002). Further, Ma *et al.* (2009) and Milo *et al.* (2002) hypothesized that the connection patterns for a given biological functionality remain conserved across the organism space. For instance, adaptation involved in chemotaxis of *E. coli* requires negative feedback between two proteins CheB and CheR (Bernardo and Tu 2003) just as the adaptation of Ca^{2+} homeostasis in mammalian cells employs negative feedback between the parathyroid hormone and vitamin D (Königs *et al.*

2020; Khammash 2021). This hypothesis has been central to establishing a system of mapping between the repertoire of emergent biological properties and the type of interconnections between the constituents of the underlying networks (figure 1). This can be used to arrive at the possible (feasible) constituents by knowing the required interconnections for a given biological functionality. From a network reconstruction perspective, knowledge of both the constituents and the interconnections between them enables us to design engineering systems that can mimic actual biological networks, i.e. utilizing the potential of synthetic biology in order to construct a model for different purposes such as therapeutic and drug design (Cameron *et al.* 2014).

The particular knowledge about the network in consideration aids in the construction of a mathematical description of the system, thereby rendering the problem of discovering design principles a project of qualitative systems identification (figure 2). This has enabled non-biologists to appreciate the problems of biological systems from a domain-agnostic perspective and made possible the translations of specific problems
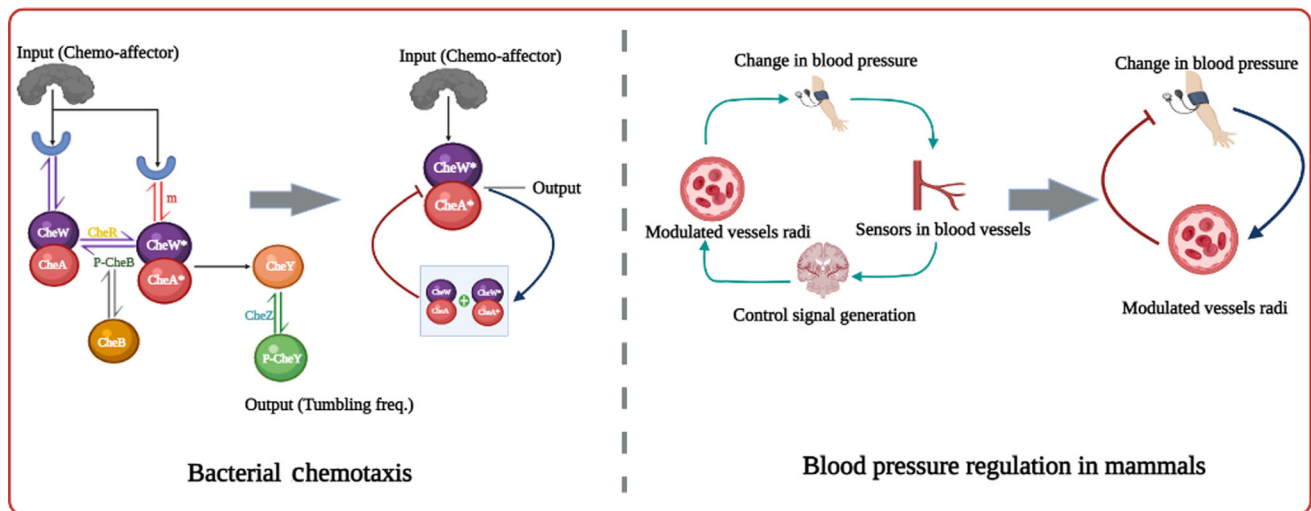

Figure 1. Conservation of design principles. The left panel shows the involvement of five proteins, namely, CheW, ChwA, CheB, CheY, and CheR. The methylated compound CheA*-CheW* affects the tumbling frequency of the flagella motors in presence of a chemo-attractant/repellent. The colours represent different proteins. The superscript '∗' refers to the methylated state of a particular protein. The red edge with blunt end and the blue edge with the arrow end refer to the activation and inhibition reactions, respectively, whereas the green arrow stands for the phosphorylation reaction. The methylation reaction is denoted by a subscript 'm' on the reaction link. Further, the phosphorylated species is denoted by a prefix P followed by the name. The entire network can be abstracted as a two-node negative feedback between CheA*-CheW* and the total concentration of CheA and CheW, with the former considered as the output. The right panel shows the same for blood pressure regulation in mammals. An increase in blood pressure is first sensed by the receptors in the blood vessels, which then communicate this information to the brain cells to increase the radius of blood vessels so that the blood pressure on the walls on the arteries decreases significantly. The resultant network can be translated to a negative feedback between the blood pressure and the diameter of the blood vessels. Therefore, in both the cases, although the network under construction is completely distinct, the resultant structure remains the same given the functionality of regulation.
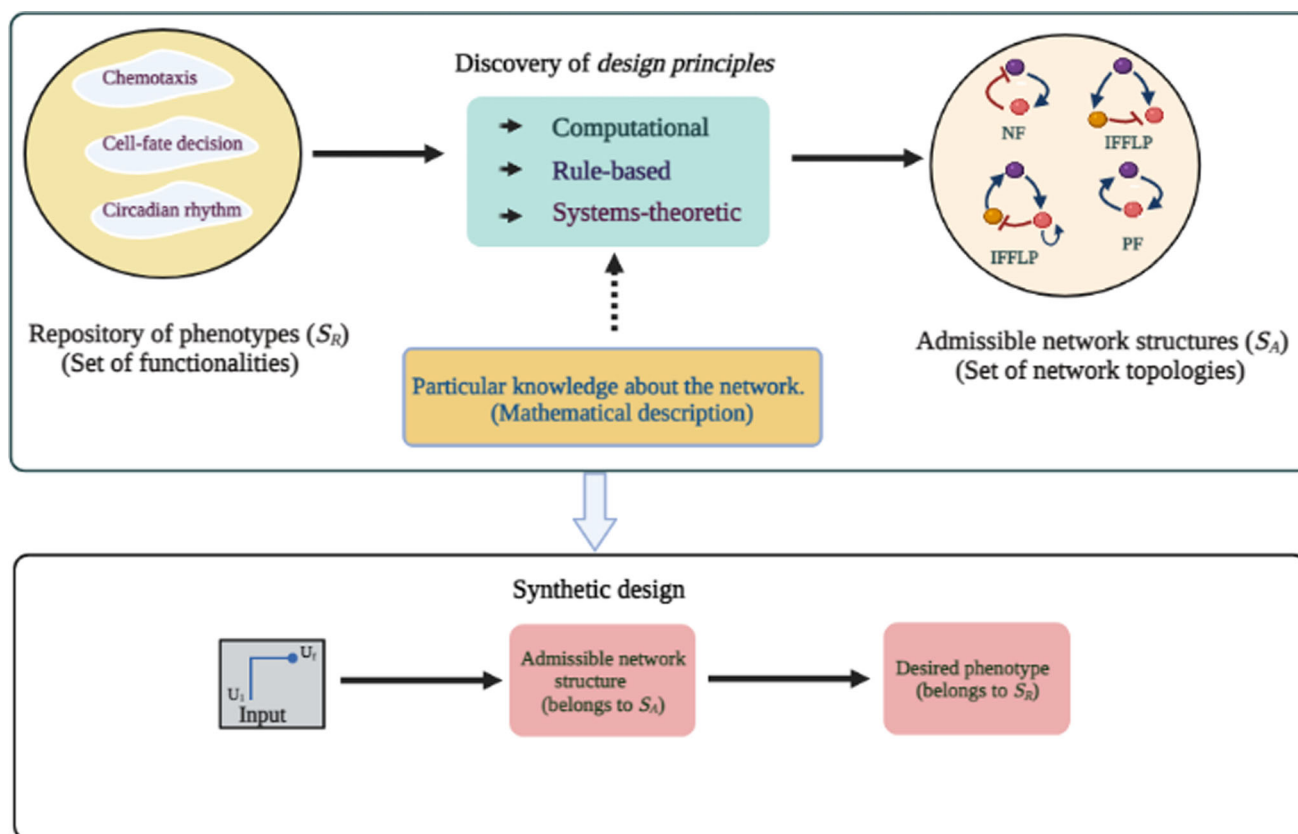
Figure 2. Reflective of the relationship between 'understanding' and 'designing' biological networks with desired properties.

pertaining to biology to a language that appeals to the broader community of science (Kulkarni *et al.* 2014). Disciplines ranging from mathematical systems theory to computational graph theory have been applied to deduce the design principles for useful biological functionalities (George 2002; Liu *et al.* 2011; Wang *et al.* 2016). Apart from the applications in mathematical modeling of biological networks, systems theory in particular has been instrumental in understanding biological networks. George (2002) developed a competent chemotherapy schedule for cancer using the principles of optimal control. Graph-theoretic methods and graph-based control theory have been applied in large biological networks (Liu *et al.* 2011). The concept of structural controllability has been employed to analyse and justify a number of network structures that emerged out of natural selection (Torday 2015). Del Vecchio (2013) and Hespanha and Sivakumar (2013) investigated the loss of modularity in biological networks from a control-theoretic viewpoint and suggested a number of solutions that were in agreement with experimental results. Hinczewski and Thirumalai (2014) argued that a phosphorylation unit of a cellular signal transduction network

adopts a Wiener–Kolmogorov filtering strategy to extract maximum amount of information from the upstream module.

The task of establishing a mapping between the network structures and phenotype can be facilitated primarily by two ways: first, an exhaustive search of the entire set of all possible network structures of all sizes – this approach can further be divided on two categories namely, computational screening and systems-theoretic methodologies, depending on the required prior knowledge about the underlying rate kinetics; second, if the functionality has a common occurrence in engineering systems, then by using the existing repertoire of design strategies in the domain of engineering design, it may be possible to pinpoint potential network motifs for the given functionality – this is denoted as the rule-based methodology.

The search-oriented approaches can further be divided in two categories based on the search techniques and assumed prior knowledge. In a computational screening method, the response of the dynamical system constituted by the chemical reactions characteristic to a particular network structure is compared against certain predefined performance parameters. Continuing

the exercise for the ensemble of structural possibilities and the associated parameters (rate constants), it is possible to arrive at the admissible network structures. Therefore, in a simulation/computational environment, the explicit knowledge of the rate kinetics along with the rate constants remains mandatory for assessing the response of a particular network structure with respect to the given phenotype.

On the other hand, systems-theoretic approaches start with introducing a few performance parameters (p) in such a way that there exists a well-defined mapping between p and the standard systems-theoretic parameters (such as poles, zeros, gain, system matrix, etc.) pertaining to any dynamical system. Thus, the performance parameters evaluated at the optimal scenario can serve as the precise mathematical conditions expressed in terms of systems-theoretic parameters for the given functionality. Subsequently, these conditions are mapped back to the general structural requirements of the network with the help of combinatorial mathematics. The central assumption in this framework is that there exists a *bijective* mapping between the network structure and the system matrix of the underlying dynamical network. As pointed out by a number of studies (Angeli *et al.* 2004; Sontag 2007; Ma'ayan *et al.* 2008; Fangzhou *et al.* 2021), this assumption holds true for most biochemical networks. Therefore, the systems-theoretic approaches do not require explicit knowledge on the rate kinetics of the network as long as the mapping holds true.

Design-oriented approaches refers to the rule-based methodology in which the prevailing design rules (in the domain of engineering) for a given functionality are applied towards identifying the admissible network structures that can serve as a subset of the admissible set. Further, the rule-based methodologies identify a subset of the admissible network topologies.

From a synthetic design perspective, as seen in figure 2, the three aforementioned methodologies can be used for discovering (and hence, understanding) the essential network structures for a given functionality. Subsequently, the admissible network structures (obtained by the said methodologies) can be used as a base for designing engineering systems that can mimic biological systems.

For the purpose of demonstrating the efficacy and advancement of the three approaches mentioned above, we selected three well-defined phenomena with great biological relevance: (i) bioswitches, (ii) simple oscillation, and (iii) adaptation. It is to be noted that all three of these functionalities are ubiquitous to every living organism. Therefore, it can be stated from the hypothesis of 'conservation of design principles' that there exist a few unified network structures for each of these biological functionalities that are common to most living organisms. Intuitively, it can be said that these admissible network structures constitute the core that governs a large network known to provide a particular biological functionality (Bhattacharya *et al.* 2018), thereby making it possible to predict the actual response of a large network by a thorough study of its constitutive motifs (Fiedler *et al.* 2013; Jorge *et al.* 2017).

Each living cell requires 'switch-like' devices to make certain decisions that become extremely crucial for different phases in the cell cycle. Evidently, these switches are constructed through the associated regulatory control in the cell (Chin *et al.* 2008). For instance, in the case of mammalian cell fate, depending on the control signal, a cell can either subject itself to differentiation or can be transformed to stem cells. Synthesis of these two possibilities (differentiation and transformation into a stem cell) is accomplished through a regulatory network with dynamics containing exactly two stable attractors (Waddington 1957; Macarthur *et al.* 2009). Apart from cell-fate decisions, controls with two stable attractors are also prevalent in prokaryotes (Ozbudak *et al.* 2004; Dubnau and Losick 2006; Guantes and Poyatos 2008). Interestingly, there exists a rich amount of literature that proves and illustrates with experimental and theoretical observations the requirement of at least one positive feedback loop in a network (irrespective of organisms and particularities of the kinetics) for it to be able to attain multistability (Laurent and Nicolas 1999; Wen and Ferrell 2003; Angeli *et al.* 2004; Macarthur *et al.* 2009; Hat *et al.* 2016; Leon *et al.* 2016). The discovery of positive feedback as an admissible motif for multistability changed the attitude of researchers toward positive feedback that existed till the early 1980s and opened up numerous structural possibilities. Contrary to the proposition of Jacob and Monod (1961), it has been proved that the metaphor of 'valley-rift-valley' in the famous Waddington diagram cannot be accomplished by a network built solely with negative feedback loops and feed-forward paths (Waddington 1957; El-Samad 2021). Further, due to their modular nature, biochemical switches have been one of the early functionalities to be designed in a synthetic environment (Gardner *et al.* 2000).

Along with multistate switches, every cell requires a synchronization mechanism or a 'clock-like device' to arrange its numerous activities ranging from

locomotion, intra-cellular communication, to regulation of important cellular behaviors in a desired order (Roenneberg and Merrow 2003). Further, these clock devices play an instrumental role in synchronous Boolean networks that arise in gene levels (Otero-Muras and Banga 2016). Several experiments have discovered the existence of regulatory sub-networks that control the sleep–wake cycle with endogenous clock pulses that try to mimic the oscillation pattern in the natural day–night cycle (Griffith 1968). Generally, there exists a variety of oscillator modules with a diversity of purposes (Tyson *et al.* 2003) ranging from neural to epidemiological oscillations (Goldbeter 1996). In this work, we shall limit our discussions to biochemical and circadian oscillations. The biochemical oscillator has a typical time period ranging from 1 to 20 mins. Each organism consists of multiple oscillator modules with a diverse set of time periods, out of which the circadian oscillator merits a special mention for its interesting properties. The circadian oscillator generates periodic responses with a time period of nearly 24 h (*circa diem*), i.e., in constant environmental conditions (darkness or light), as observed by the French astronomer Jean Jacques d'Ortous de Mairan in 1729 through his study of the unfolding and folding of the leaves of the *Mimosa* plant kept in a dark room. When an organism is subjected to a change in environmental conditions, for instance, in the case of travel that consists of a significant time difference, i.e., difference in the daylight pattern, the circadian rhythm also changes its time patterns according to the duration of the sunlight in the destination. This necessitates the existence of a 'master–slave' clock, where the downstream slave module follows the oscillatory properties of the upstream master (Joshi *et al.* 2020). Further, it has been shown through the seminal studies carried out by Hardin *et al.* (1992) and separately by Saez and Young (1996) that a network of multiple genes (including *per1* and *cry*) with negative feedback can be understood as the underlying mechanism of a circadian clock. These findings strengthened the long-anticipated transcriptional–translational feedback loop model involving the negative feedback between the Period (PER) and Cryptochrome (CRY) proteins (Fahrenkrug *et al.* 2006). Recent work by Pett *et al.* (2018) has also indicated the existence of a repressilator as the necessary condition for sustained endogenous oscillation in mammalian cells. Interestingly, although the specific proteins/genes involved in providing the oscillation vary across organisms, the existence of at least one negative feedback remains a commonality across all the oscillatory biochemical networks hitherto observed.

Unlike oscillation, certain activities in every living cells need to be invariant to the steep changes in the surroundings, for instance, regulation of the body temperature in different climate conditions or maintaining a desirable blood pressure in the presence of environmental fluctuations. These abilities in their totality enable a living organism to survive in the presence of the sudden changes in the surroundings. Apart from homeostasis, bacterial chemotaxis, the phenomena of a bacteria reacting to the presence of a chemo-attractant or chemo-repellent whose concentration is prone to sudden change also requires the sensory module to send a constant signal (evidently, through chemical reactions) to the flagella motors in spite of rapid fluctuations in the chemo-affector concentration (Bernardo and Tu 2003). Therefore, adaptation is an essential property of every living organism that consists of sensing environmental change and subsequent return to the pre-disturbance desired state. Further, experimental studies have discovered that apart from biochemical networks, there exist a class of networks such as voltage-gated sodium channels that can also provide adaptive behavior in the presence of a step-like disturbance (Ferrell 2016). These network structures provide adaptation response in the presence of a single step-type disturbance but fail to provide the same in subsequent changes in the environment (Friedlander and Brenner 2009).

This review elucidates all these three approaches aimed at discovering design principles in the light of the existing literature. To this end, we adopted three biological functionalities of utmost importance: (i) oscillation, (ii) toggle switches, and (iii) adaptation. Moreover, as depicted in figure 2, traversing the path of understanding biological network structures through these three approaches also paves way for the synthetic design of sophisticated biological networks. Sections 2, 3 and 4 illustrate the work done through computational screening, rule-based and systems-theoretic approaches with respect to the three aforementioned functionalities. Section 5 presents a thorough discussion and the potential future scope of these three methods.

2. Computational screening

Every biological network can be characterized by the underlying dynamical systems constituted by the rate reactions and rate constants associated with the unique stoichiometry of the network. The computational approach scans through all possible network structures by simulating the associated dynamical system.
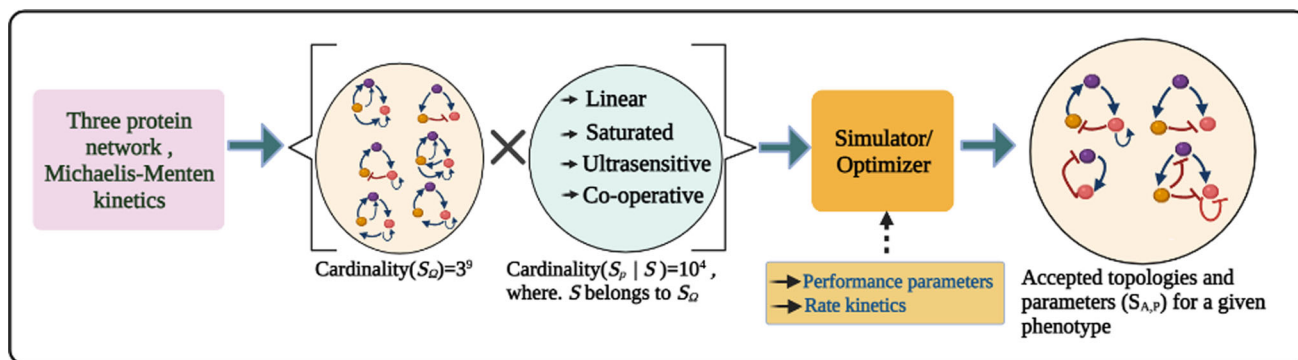
Figure 3. A schematic for computational approaches. With the prior information about the complete mathematical description (rate kinetics, parameters) of the network, all possible network structures combined with the set of biologically feasible parameter sets are subjected to simulation. The response for each pair of network structures and parameter set is assessed with respect to the predefined performance indices. In the case of an optimization-based approach, the objective function is evaluated for the entire topology–parameter space.

Further, in order to make the search exhaustive, each network structure is examined for multiple sets of rate constants (figure 3). The performance of a specific topology–parameter combination is assessed through certain performance parameters defined with respect to the reference (desired) functionality. The computational screening approaches have used the Q-ratio to measure robustness of a given network structure. It is to note here that Ma *et al.* (2009) defined robustness as the quality (pertaining to a given network structure) of keeping the adaptive response unaltered in the presence of parameter fluctuations. Given the desired response, the Q-ratio measures the number of times a particular network structure produces a satisfactory performance with respect to the performance parameters characteristic to the desired functionality to the total number of samples drawn from the parameter space. Therefore, the topology that produces a satisfactory performance for most combinations of the parameters (rate constants) is considered as the most robust network structure for a given functionality.

2.1 *Multiple steady states for two and three protein networks*

Since it has been found that biological switches can be constructed through gene regulatory networks, computational approaches in the existing literature have aimed to find the particular connection patterns between the relevant genes that can exhibit a switch-like behavior. To this purpose, Leon *et al.* (2016) started with a three-gene network. Each protein can either repress or activate the process of the synthesis for itself or other proteins. Considering the presence, types,

and directionality of the edges, there are 3^9 possible network structures for a three-node network. It is to be noted that each network structure is characterized by the underlying dynamics emanating from the chemical reactions in the pathway. Each chemical reaction in turn is characterized by its associated rate constants. Therefore, the stoichiometry with the concentrations of the genes as state variables, rate constants as the parameters, and the conservation rules together constitute the underlying differential algebraic equation system for a given network (Leon *et al.* 2016).

Leon *et al.* (2016) and Diegmiller *et al.* (2021) analysed multiple network structures for networks with at least two genes. Each network structure was examined for different parameter sets in order to assess the robustness of the network structure in consideration. The central idea is, given a particular network structure, to construct a prior for the parameter vector, and thereafter, with the prior distribution and the explicit knowledge of the rate reactions, obtain the log-likelihood in order to arrive at the posterior parameter range best suited for the desired response. Since the rate reactions in most of the cases are highly nonlinear, it becomes very difficult to obtain a closed form expression of the likelihood function. To circumvent this problem, Leon *et al.* (2016) used an approximate Bayesian computation with the sequential Monte Carlo method, wherein the first set of parameters (p_1) is drawn from the prior distribution. The model was simulated with p_1 for different initial conditions obtained from an extensive Latin hypercube sampling. The existence of stable steady states reflects as clusters in the state space. The Stabilityfinder tool, developed for this purpose, uses the K-means clustering algorithm followed by a Gap statistic to find the number of

clusters present in the phase space. A distance metric (d) was introduced to compute the distance between N_c and the desired number of steady states. The sample parameter set was considered acceptable if the associated d was less than an initial threshold (ε_0). The following parameter set was obtained by considering a weighted combination of previously accepted parameter sets. Further, with each iteration, the threshold ε_t is reduced until it reaches a user-specific value. Thus, for a given network structure, the posterior parameter distribution admissible for multistability of desired order can be obtained with the histogram of the ensemble of accepted parameter sets in the algorithm.

This algorithm was applied to the well-known two-node bistable network – also known as the Gardner switch – characterized by two mutually repressing nodes (genes, in this case). With Michaelis–Menten rate kinetics, the Stabilityfinder tool was implemented to deduce the posterior for four parameters: the rate of synthesis for each gene and the co-operativity indices for each repression. The degradation for each repressor has been assumed to be linear, with degradation rates being unity. The computationally burdensome simulations have been parallelized using Graphic Processing Units. The results from this study reconfirmed Gardner's hypothesis that, in order for a two-gene positive feedback network (with an even number of repressing edges) to exhibit bistability, the synthesis rates should be balanced and the cooperative indices should be greater than unity.

The study revealed that the previously existing conditions for a deterministic Gardner switch to provide a bistable behavior stands true in the stochastic scenario as well (Losick and Desplan 2008). The more generalized version (in terms of the form of dynamics) of the Gardner switch (Gardner *et al.* 2000) obtained by Lu *et al.* (2006) has also been investigated. Interestingly, it has been found that a pair of additional positive self-loops along with a mutual repression between the two proteins can induce tristable behavior. Although it has been conjectured that presence of four or more stable states can only be obtained through a network with at least three proteins, establishing a mapping between the number of stable steady states and the resultant network structure still remains an open challenge.

2.2 *All possible three-node oscillators*

Investigating network motifs through computational means for the sole purpose of finding the necessary (and hopefully sufficient) structural conditions behind oscillations was attempted first by Pett *et al.* (2018). It has been well established in the literature that a two-node network with a negative feedback can provide an oscillatory response (Bela Novak and Tyson 2008). Later, three- and higher-order networks consisting of at least a wholesome negative feedback have also been shown to portray oscillatory responses (Rössler 1972; Frank 1974; Novak and Tyson 1993; Pomerening *et al.* 2005). In order to circumvent the computational burden in examining the entire topology–parameter space and, at the same time, to draw reliable conclusions, Pett *et al.* (2018) started with a well-known, experimentally validated circadian oscillator network that involves *Bmall*, *Per2*, *Cryl*, *Dbp,* and the nuclear receptor *Rev-erb-α* (Pett *et al.* 2018). Experimental studies have shown that the network contains 17 edges with 34 parameters. From the given network, elimination of random nodes (both one at a time or multiple deletion) have been performed and it has been found that a three-protein network (eliminating *Bmal* or *Dbp*) can provide robust oscillations with both negative and positive feedbacks. Subsequent studies on gene knockouts from the basic network validate the theoretical findings by Thomas (1981) that a complete graph requires at least one negative feedback to provide oscillations. It is to be noted, that the basic assumption in all these studies has been that the elements of the system matrix obtained by linearizing the nonlinear system of rate equations do not change signs in the entire state space.

Another computational effort by Li *et al.* (2017) examined protein networks to find the structures that produce robust oscillations. To begin with, Li *et al.* (2017) examined all possible structural combinations of a two-protein network. Each network structure was simulated for multiple sets of parameters and initial conditions. This methodology has also been extended to three-node networks. Interestingly, it has been found that a negative feedback with an incoherency in the respective nodes yields more robust oscillations than a simple negative feedback. It is to be noted that the requirement of incoherency refers to a specific pattern of interconnections. For instance, in the case of an incoherent two-node network with a negative feedback between two proteins A B, if B represses A, then the condition of incoherency binds A to have a positive self-activation. Notably, these network structures have been validated experimentally (Higgins 1964; Goldbeter and Lefever 1972).

Apart from direct computational screening, there exist other optimization-based approaches that rely on mixed-integer programming to deduce the network

structure along with the parameters. Otero-Muras and Banga (2016) developed an approach wherein the desired functionality is represented by a standard function. Subsequently, the difference between the desired and obtained responses is minimized, subject to the rate constants and structural variables, which are represented as a signed stoichiometry matrix. This leads to mixed-integer nonlinear programming which produces the admissible topology–parameter combinations given the desired functionality. The study revealed that the existence of a negative feedback loop together with incoherent self-loops is admissible for oscillations in a delay-free system.

2.3 *Only two admissible structures for adaptation in a three-node network*

Similar to toggle switching and oscillation, computational screening methods have also been applied in the context of adaptation with great success. Ma *et al.* (2009), in their seminal studies, adopted a biochemical networks with three nodes. As a first step, the quantum of adaptation attained by a given network has been characterized by two performance parameters, namely, sensitivity and precision. Sensitivity is computed as the ratio of the relative difference between the peak and the initial levels of the output to the relative change in the input, whereas precision is defined as the ratio of the relative changes between the initial and final output levels to the changes in the input. The bare minimum values of sensitivity and precision that the response of a particular network structure should contain in order to be enlisted as the admissible motif have been chosen as 1 and 12 respectively. The study investigated $16,038$ possible network structures. Each structure was examined for $10,000$ different sets of parameters (rate constants) totaling $\approx 1.6 \times 10^8$ number of simulations. A numerical measure of robustness is defined as the ratio between the number of parameter sets for which a particular network structure showed satisfactory adaptive performance to the total number of parameter combinations considered. The study revealed that only 395 topologies were able to provide a sensitivity and precision value greater than the threshold. Also, all these 395 topologies have been shown to either contain a negative feedback loop with a buffer node (NFBLB) or at least two mutually opposing (incoherent) feedforward paths from the input to the output node (IFFLP). The buffer node in NFBLB refers to the specific protein in the negative feedback loop whose dynamics need to be independent of its own

concentration in order for the network structure to provide perfect adaptation. Upon robustness analysis, it was found that IFFLP is comparatively less robust than the NFBLB topology. Further, addition of negative feedbacks to both NFBLB and IFFLP has been shown to increase robustness (Ma *et al.* 2009).

Another important intervention by Qiao *et al.* (2019) began with a two-node biochemical network. Similar to Ma *et al.* (2009), a variant of Michaelis–Menten kinetics was adopted but the input signal was considered to be stemming from the scalar Langevin equation, i.e. the stochasticity in the network was introduced through the input. Each possible network structure for two- and three-node setups was investigated for multiple parameter sets.

The ratio between the output and the input signal-to-noise power ratio (SNR) was measured for each network structure, in addition to sensitivity and precision. The variance required to compute the ratio of output and input SNR was derived by using the Palson dissipative theorem (PDT). Subsequently, a correlation study was done by pairing the adaptation and noise attenuation credentials. As a result, it has been found that although sensitivity has a positive correlation with SNR, the correlation between precision and SNR was found to be negative for all the network structures admissible for adaptation. This suggests an inherent incompatibility of a three-node network to provide perfect adaptation and simultaneously minimize output variance. Therefore, the only way to achieve this goal was to connect an adaptation module with a distinct noise-filtering module. Further, it has been shown through computational study that an adaptation module followed by a noise filtering downstream can yield more robust response than the other way round.

Computational approaches have been able to generate highly reliable predictions for small-scale networks. Given particular kinetics, the structural conditions produced from these approaches have often been used as the benchmark to assess the correctness of the predictions yielded by the other two methods. Further, the study of robustness, i.e. sensitivity of the response with respect to the parameters, becomes quite straightforward in this method, which is not necessarily the case for the following methods (Ma *et al.* 2009).

3. Rule-based methods

Unlike search-oriented brute force or mixed-integer optimization, the rule-based method adopts a design-oriented approach. The repertoire of design tools in the
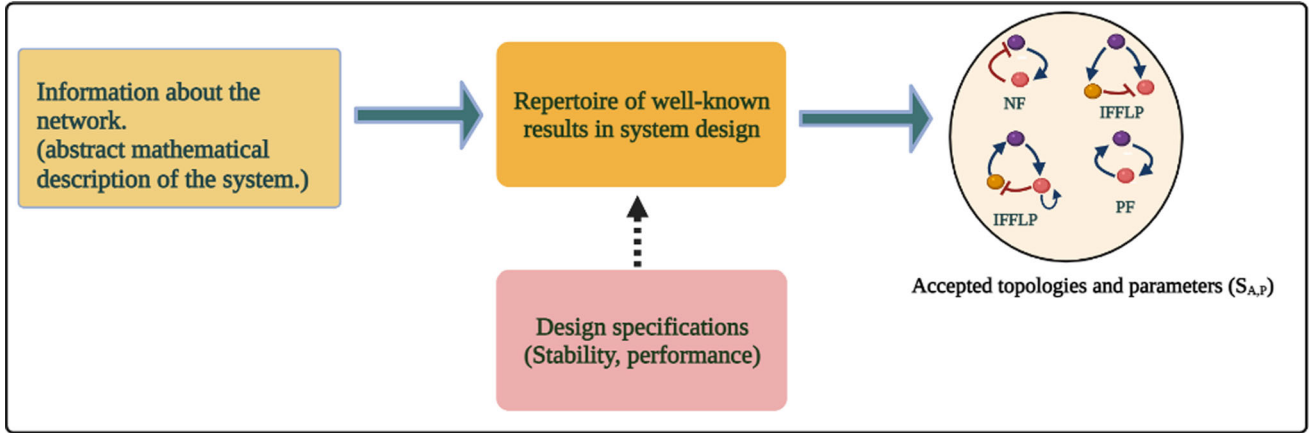
Figure 4. A schematic of rule-based methodologies at work. With the qualitative information about the mathematical description (form of rate dynamics, I/O nodes) of the network, well-known engineering principles for circuit synthesis are applied to construct the topology that satisfies the specified design requirements. The design specifications are obtained from the performance parameters.

existing literature is utilized to construct a biological network that can provide the desired functionality (figure 4). This approach has been instrumental in constructing the building blocks (basic modules) of large biological networks. It is also worth mentioning that this approach has an implicit assumption that biological networks are essentially modular, i.e. the functionality of a particular system remains unaltered if it is connected with a downstream node. Also, the fact that a biological network requires to adhere to specific rules to produce a particular response necessitates a translation of biological rules into a more generic (context-free) language of design specifications.

3.1 *Biological switches: admissible topologies and their stability properties*

The first successful application of rule-based design of biological switch dates back to the year 2000, when Gardner *et al.* (2000) synthetically constructed a genetic toggle switch in *Escherichia coli*. In this paper, the authors proposed to design a biochemical switch exhibiting bistability in *E. coli*. A two-protein network with positive feedback accomplished by mutual inhibition was considered as the ideal network structure for producing switch-like behavior. The dynamics of the resultant network were assumed to be

$$\dot{x} = \frac{K_x}{1 + y^{\theta_{xy}}} - x \tag{1}$$

$$\dot{y} = \frac{K_y}{1 + x^{\theta_{yx}}} - y \tag{2}$$

where x and y are the concentrations of the repressor proteins and θ_{ij} is the cooperativity index of gene j in the synthesis of gene j. Notably, the parameter region in which the system defined in equations 1 and 2 produces bistable behavior is proportional to the cooperativity indices $(\theta_{xy}, \theta_{yx})$ of the network. More importantly, it has been concluded that for guaranteeing bistability, at least one of the co-operativity indices has to be greater than unity (figure 5).

Apart from the design principles obtained via synthetic design, Angeli *et al.* (2004) designed a strategy to assess the stability characteristics of a positive feedback system (Angeli *et al.* 2004). Instead of analysing the entire N-node network, this method first divides the network into process and control modules in such a way that the steady states of the output state for the open-loop system (y^*) can be written as the function of the control input (u) in a particular way. Let us denote that at the steady state as

$$y^* = f(u) \tag{3}$$

Angeli *et al.* (2004) developed a graphical approach wherein $f(u)$ is plotted with respect to u along with the straight line $g(u) = u$. The intersection points $(u^*, f(u^*))$ are noted. It can be shown that $(f(u^*), u^*)$ is a stable equilibrium if the following condition is satisfied

$$f(u) < u|_{u^{*+}} \quad \text{and} \quad f(u) > u|_{u^{*-}} \tag{4}$$

Further, Angeli *et al.* (2004) showed that if the open-loop system is monotone and the resultant network does not contain any negative loop, then each
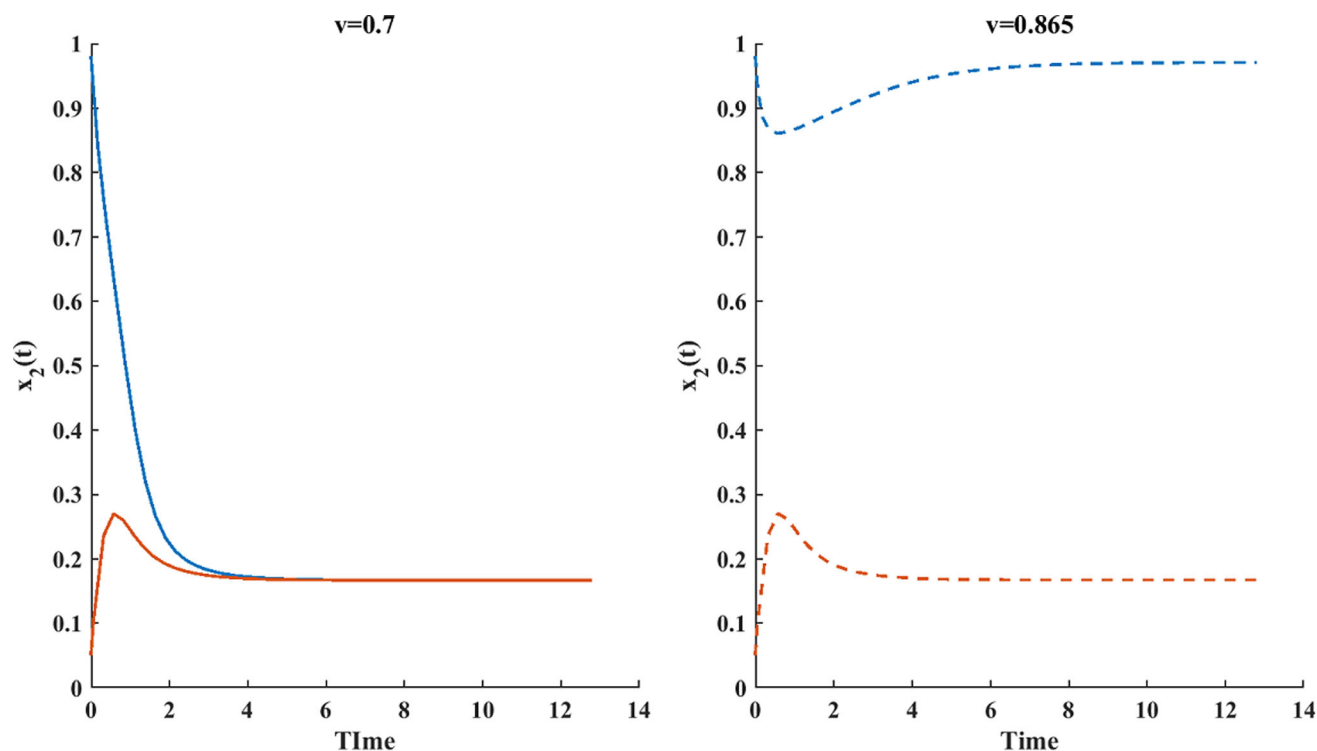
Figure 5. Example of bistable behavior. The left figure is simulated with $v = 0.7$ and the same for the figure at right is $v = 0.865$. The dynamical system in equations 1 and 2 possesses a pitch-fork bifurcation at $v = 0.865$ resulting in two stable steady states. The parameter set considered for the purpose of simulation is $\alpha_1 = 0.9$, $\alpha_2 = 1.05$, $\beta_1 = 200$, $\beta_2 = 9.89$, $\gamma_1 = \gamma_2 = 3.98$, $K_1 = 29.87$, $K_2 = 1.008$.

interaction point satisfying equation 5 provides an injective mapping with a stable steady state of the closed-loop system. Therefore, the conditions used in this methodology provide the reason why a positive feedback network with monotone dynamics can produce switch-like behavior.

3.2 Negative feedback with delay promotes oscillations: Studies on small-scale networks

As opposed to biological switches, oscillation, from the perspective of dynamical systems theory, requires the underlying dynamical system to possess at least one stable limit circle. An oscillatory behavior can be characterized by its amplitude and time period. Although the design of a circadian oscillator has been prioritized in the existing literature, the design strategies adopted in most cases can be easily extended to the construction of oscillators with arbitrary frequency. For instance, Albert Goldbeter (1996) observed that similar to circadian oscillators, design principles for the CAMP oscillator in mammals contain a negative feedback with delay. What differentiates the biochemical networks from other prevalent ones in biology is

that the system matrix for the associated linearized network dynamics can be thought of as a variant of the digraph matrix for the same biochemical network. This opens up the scope for the application of certain instrumental results in the field of combinatorial matrix theory (Maybee *et al.* 1989).

Biological oscillation is known to involve feedback loops. Initial applications of rule-based methods in designing biological networks involved the well-known engineering principle of 'negative feedback with delay' for designing oscillators (Mackey and Glass 1997). The basic idea is to introduce delay in the feedback loop along with integral control. The integral of the error signal provides a finite phase margin which can be met by the additional delay in the feedback loop. In the case of a linear system, a delay slightly greater than the required phase margin can lead to unstable behavior, leading to failure of oscillation. Further, even if the delay accurately compensates for the delay margin, contrary to the biological oscillators, the resultant oscillation produces a sustained tonal oscillation, with the amplitude being a function of the initial condition. Therefore, nonlinearity either in the process or the feedback loop is required to produce stable oscillatory responses, with amplitude independent of the initial conditions.

Mackey and Glass (1997) studied a single state model with explicit delay wherein the rate of the protein synthesis at any particular instance t is a function of the protein concentration at $(t - \tau)$. A self-repressive loop with sigmoidal kinetics can potentially generate oscillatory responses for certain admissible sets of rate constants. Further, it has been shown that a large value of co-operativity increases the nonlinearity, thereby increasing the chances of exhibiting a stable oscillatory response.

Instead of the explicit use of delay in the network dynamics, the introduction of multiple indirect paths or additional positive feedbacks can also serve the purpose. It can also be shown that a single protein system without delay cannot provide oscillations. Therefore, the absence of explicit delay has to be traded off by increasing the order of the network, i.e., introducing at least one more protein in the network. In the case of a two-protein network, Novak and Tyson (2008) suggested that negative feedback between two proteins with repression of the degradation of at least one protein can provide oscillations. Given two proteins A and B with concentrations $x_1(t)$ and $x_2(t)$, the proposed network dynamics can be written as

$$\dot{x}_1 = \frac{\alpha_{12}}{\beta_{12} + x_2^{n_1}} - \gamma_{11}x_1 \qquad (5)$$

$$\dot{x}_2 = \alpha_{21}x_1 - \gamma_{22}x_2 - \frac{\alpha_{22}x_2}{ax_2^2 + bx_2 + c} \qquad (6)$$

It can be seen that for a particular design choice of $(\gamma_{22}, \alpha_{22}, a, b, c)$, the expression $\frac{\partial \dot{x}_2}{\partial x_2}$ can become positive, indicating an effective self-loop for the node B. The resultant positive self-loop acts as a potential memory device which makes the current value of x_2 a function of its past value.

Apart from these two topologies, negative feedbacks with incoherent amplification are also shown to be able to provide oscillation. Pomerening *et al.* (2005) and Ananthasubramaniam and Herzel (2014) suggested at least one positive feedback loop along with the customary negative feedback to attain incoherent amplification of at least one node in the network. The experimental studies performed on the mitotic oscillator present in sea urchin embryos and *Xenopus* eggs (Goldbeter 1996) provided greater insight into the possible network structures for oscillations. Further, the models proposed by Goldbeter (2002) showed that a negative feedback loop with three proteins accompanied by a positive feedback produces robust oscillations and can reliably explain the behavior of the mitotic oscillators.

Moreover, as demonstrated in figure 6, rule-based methods have deduced three types of design principles for oscillation, namely, (1) negative feedback with delay, (2) negative feedback with positive feedback, and (3) negative feedback with incoherent amplification. It has been conjectured by Novak and Tyson (2008) that negative feedback is absolutely necessary for a biological network of any size to provide an oscillatory response. Further, it is to be noted that the class of oscillators which facilitates incoherent amplification with the presence of positive feedbacks over and above the mandatory negative feedback offers a robust performance of oscillations in the presence of parametric variations.

3.3 *Integral control facilitates adaptation*

The rule-based attempts on adaptation, an essential property of every living organism to regulate the system with respect to the desired state, have been based on the well-known engineering principle of using negative feedback for the purpose of regulation. The basic idea is to use a pure or proportional-integral control such that the output of the response follows the desired signal perfectly.

To illustrate this, let us consider a closed-loop linear, time-invariant system $G_{Cl}(s)$. The corresponding open-loop and control transfer functions are denoted as $G_P(s)$ and $G_C(s)$, respectively. Further, the disturbance $(D(s))$ is assumed to be added to the output of the open-loop plant. The controller is provided with the desired set points $(R(s))$. Considering negative feedback for stability purposes, the output $(Y(s))$ can be written as

$$Y(s) = \frac{G_P(s)G_C(s)}{1 + G_P(s)G_C(s)}R(s) + \frac{G_P(s)}{1 + G_P(s)G_C(s)}D(s) \qquad (7)$$

For the system to reject a step-type disturbance $\left(D(s) = \frac{1}{s}\right)$ and follow a constant reference, the following should hold:

$$\lim_{t \to \infty} y(t) = \lim_{s \to 0} sY(s) = R(s) \text{ (final value theorem)} \qquad (8)$$

$$\Rightarrow \lim_{s \to 0} s\left[\frac{G_P(s)G_C(s)}{G_P(s)G_C(s)}R(s) + \frac{G_P(s)}{s(1 + G_P(s)G_C(s))}\right]$$
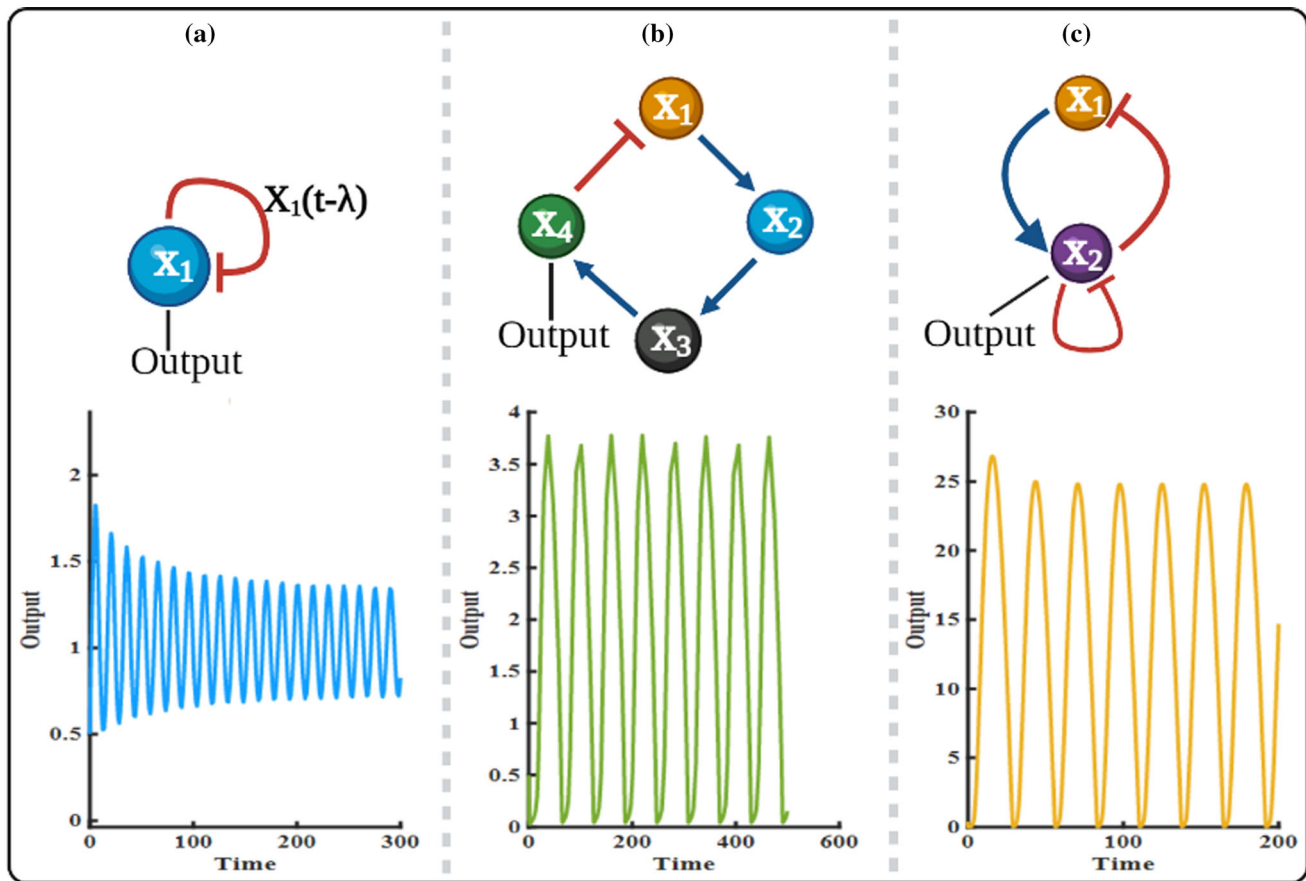$$= R(s) \qquad (9)$$

Figure 6. Example of oscillatory behavior. (**a**) Simulation to a single protein negative feedback motif with delay. As long as there exists sufficient nonlinearity and long delay it can produce sustained oscillation. (**b**) The response of a network involving two proteins and two intermediates connected in a negative feedback fashion. The multi-stepped, long negative feedback can act as potential delay necessary for sustained oscillation. The associated rate equations along with the rate constants both for (**a**) and (**b**) have been provided in the Appendix. (**c**) Response of a two-protein network with wholesome negative feedback and incoherency in the output node. The self-inhibition of the degradation of the output protein acts as an equivalent of positive feedback. The necessary parameters used here $\alpha_{12} = 4$, $\beta_{12} = 1$, $\gamma_{11} = 0.05$, $\alpha_{21} = 0.32$, $\gamma_{22} = 0.053$, $\alpha_{22} = 1$, $a = 0.1$, $b = 1$ and $c = 2.5$.

Since the reference $r(t)$ is a constant signal, its Laplace counterpart can always be expressed as $(s) = \frac{k_0}{s}$, where k_0 is a constant. Therefore, equation 9 can be satisfied if and only if $G_c(s)$ can be expressed as $\frac{\Phi(s)}{s}$, where $\Phi(s)$ and s are co-primes. This indicates the presence of an integral feedback control as a sufficient condition for perfect adaptation (Astrom and Richard 2010).

From the perspective of bifurcation analysis, the condition for perfect adaptation can be cast as the invariance of the output steady state (concentration of CheP in bacterial chemotaxis) with respect to a step-change in the chemoaffector. Interestingly, it is seen that in the two-state model built by Barkai and Stan (1997) for bacterial chemotaxis, the steady-state concentration of the output state CheP remains invariant to

the inputs for all possible biologically feasible values of total receptor concentration. Tau-Mu *et al.* (2000) showed that the Barkai–Leibler model employs a pure integral controller to render a zero sensitivity of the output steady state to the external input and other parameters through the internal model principle.

Subsequently, Briat *et al.* (2016) used another rule-based approach to deduce the design principles for adaptation in the presence of stochastic variations. In the scenario of low numbers of reactant molecules, it becomes infeasible to treat the reaction systems in the continuous-time deterministic framework. Further, considering the parameters as random variables results in a stochastic dynamical system. Although use of the Fokker–Planck equations (Kolmogorov's forward equation) seems a tempting option here, the

nonlinearity (non-Gaussianity of the evolving joint PDF of the states) of the chemical reactions can result in an infinite dimensional deterministic dynamical system in terms of the statistical moments – this is known as the moment closure problem in nonlinear stochastic dynamical systems. To circumvent this problem, instead of the distribution, Briat *et al.* (2016) focused on the chemical reactants directly and designed the control strategy so that the population average of the output species remains unaltered in the presence of stochastic variations. To this effect, a set of reference, sensor, and control reactions was proposed, and using certain important results on Markov processes, they showed that the negative feedback integral control strategy provides robust tracking of the desired set point in the presence of noise. One of the major shortcomings of this design is the increased variance of the output state. Subsequently, Briat *et al.* (2018) prescribed an additional negative feedback loop in order to reduce the variance of the output state. The general expression of variance has been obtained through solving a Lyapunov-like equation stemming from the linearized dynamics. Subsequently, it has been shown that negative feedback along with the existing antithetic integral control can serve the purpose.

From the design perspective, rule-based methodologies can serve as a great starting point because of their prominent lineages with well-tested engineering systems. Further, as beautifully summed up by Oberortner *et al.* (2015), the design rules can be classified into five distinct categories, namely, counting, pairing, positioning, orientation, and interactions. Counting rules refer to the maximum (minimum) number of genetic species that can be used for a given design. For instance, according to well-known engineering principles, any delay-free dynamical system needs to be of at least second-order so as to produce oscillatory response. This translates to the requirement of a two-gene oscillator system as shown by Tyson (1975). The pairing rules denote the number of appearances of a particular pair of genetic elements, thereby referring to the process of choosing appropriate biochemical species for a particular phenotype. In genetic design, ordering involves identifying the ideal chronology of the DNA sequence via spatial separation to timing a particular interaction with reference to the design process. This enables realizing the functional delay in the system. It has been shown by Mackey and Glass (1997) and Börsch and Schaber (2016) that introduction of delay can significantly reduce the minimum requirements on the order of the dynamical systems to exhibit a given behavior. The design specification of orientation requires knowledge of the orientation pattern of a particular genetic substance in order to meet the desired specifications. Finally, the design rule of interaction captures all the desired and observed interactions possible with the chosen genetic substances of the given order and orientation.

The aforementioned rules particularly aid in the synthetic design of the biological networks. Design principles obtained from the wisdom pertaining to the domain of engineering remain starved of the actual biological context. Therefore, selecting the appropriate biochemical species and engineering the appropriate promoters are instrumental in implementing engineering designs to biological networks.

4. Systems-theoretic approaches

Apart from the two formalisms mentioned above, systems-theoretic approaches have been at the forefront in unraveling the design principles for important biological functionalities. Mathematical systems theory can be applied to any phenomena that can be cast into the formalism of 'input → system/model → output'. In this sense, the confluence of systems theory and systems biology refers to the application of systems-theoretic principles in analysing the behavior of complex biological networks. Therefore, the task of identifying design principles given the desired response and the input disturbance pertains to a problem of qualitative systems identification. To this purpose, similar to the brute force approach, several hyper-parameters are introduced to characterize the reference response. These parameters are mapped as some conditions in terms of certain well-defined qualities such as stability, controllability, gain, etc*.,* of systems theory. These conditions, along with the application of combinatorial matrix theory, can provide the generic design principles for a given biological functionality (figure 7).

As noted by several authors (Angeli *et al.* 2004; Sontag 2007; Ma'ayan *et al.* 2008), the system matrix of the dynamical system linearized around an operating point can serve as a variant of the digraph matrix for almost every biochemical network. This observation has been crucial for the systems-theoretic methodologies since this renders the systems-theoretic approaches agnostic to the particularities of rate kinetics. The analysis of the digraph matrix has enabled one to use the wealth of combinatorial matrix theory to unravel the necessary structural requirements for any given functionality.

Let us consider a biochemical network containing N interacting biochemical species (say, proteins, for
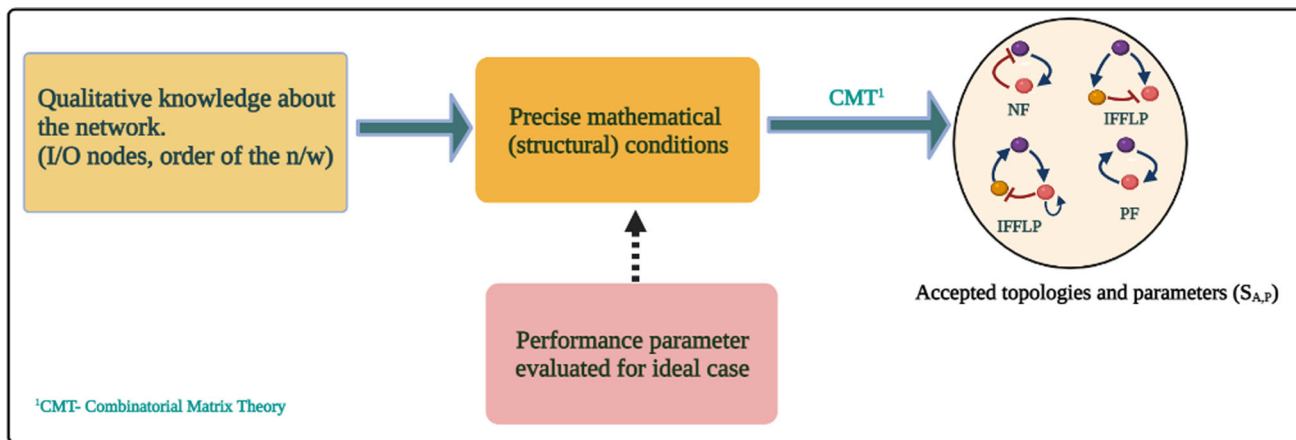
Figure 7. A schematic for systems-theoretic approaches to deduce the complete set of design principles against a specific functionality. First, the performance parameters are analysed and mapped to certain entities of systems theory. These conditions are then applied on the underlying dynamical system which is constructed with a qualitative knowledge of (specification of the I/O nodes), the network, in order to obtain precise mathematical conditions on the digraph matrix of the network. Subsequently, the complete set of admissible topologies is obtained through the application of combinatorial matrix theory.

instance) $[x_1, x_2, \cdots x_N]$ with the associated rate constants $[p_1, p_2, \cdots p_R]$. Further, assume that none of these N proteins entertains any linear/nonlinear conservation law. The resultant dynamical system can be expressed as

$$\dot{x} = f(x,p), y(t) = h(x,p) \tag{10}$$

where $x \in R^N$ and $p \in R^P$ are the states (concentrations) and the parameters (rate constants) of the dynamical system. It can be seen that the presence of any linear algebraic constraint stemming from certain conservation principles results in the reduction of the order of the dynamical system.

It is to be noted here that equation 10 is a symbolic representation of the dynamical system that lies beneath any biochemical reaction. The systems-theoretic approach works on this symbolic representation of the dynamics without assuming any particular rate law ($f(x)$) barring the central assumption that the elements of the matrix $\left[\frac{\partial f}{\partial x}\right]_{i,j}$ do not change signs in the entire state space.

Intuitively, any switch-like device requires toggling between at least two steady states. Therefore, given a network, finding design principles for biological switches requires deriving structural conditions for multistable dynamics.

As established by Sontag (2007), most of the biochemical networks maintain a set of properties that, in most of the cases, make a systems-theoretic intervention in drawing insights on the network structures from

the underlying dynamics fruitful. These properties are described as follows:

1. The system of differential equations in equation 10 constitutes a well-posed dynamical system. This ensures the existence and uniqueness of the concentration trajectories.

2. The $(i,j)^{th}$ element in the Jacobian (A) of $f(x,p)$ with respect to x refers to the edge from the jth node to the ith node. Further, the sign of $A_{i,j}$ dictates the type of interaction. In the case of repression, $A_{i,j}$ is negative in the manifold, whereas the opposite is true for activation.

4.1   *Positive feedback: Essential motif for bioswitches of any size*

Previously, Thomas (1981) conjectured that the existence of an odd number of inhibitory interactions in a loop (negative feedback) and an even number of inhibitory interactions (positive feedback) could potentially serve as the necessary condition for sustained oscillations and biological switches. The paper adopted both, two- and three-gene networks. A gene network can be constructed by considering the respective genes as nodes and the interactions as the edges. In the case of a Boolean network, each gene can possess only two states, ON (1) and OFF (0). The status of a particular gene A is determined by the generation process associated with A. Further, this generation process for A is

obtained by considering the cumulative or multiplicative effects exerted on A by its neighboring genes. Therefore, the present state of the process for gene A depends on the past values A and its neighboring genes. If the process is at the ON state, then the associated gene becomes activated after a certain amount of delay.

Using the above framework of Boolean networks, Thomas (1981) showed that a three-gene network requires at least one loop consisting of an even number of inhibitory regulations for any network structure to provide multiple steady states. In the case of oscillation, the network structures having at least one loop with an odd number of inhibitory interactions serve as the necessary condition for oscillation. This result played the central role behind the conjecture made by Thomas (1981) that the presence of positive feedback for any network irrespective of the number of nodes and edges is essential for switch-like behavior. Similarly, for oscillation, the existence of at least a negative feedback loop can be the potential necessary condition.

Later, Plahte *et al.* (1995) and Snoussi (1998) proved Thomas's conjecture for biological switches. The work adopted a continuous-time system with time-invariant parameters as described in equation 10. The concentration dynamics was assumed of the form

$$\dot{x} = f(x, p) - x \qquad (11)$$

Where the second term in equation 11 refers to the degradation process of each protein. It can be inferred from equation 11 that the corresponding Jacobian matrix of $f(x)$ with respect to x can act as the variant of the digraph matrix provided condition 2 is satisfied by $f(x)$.

Snoussi (1998) first showed that if the dynamical system in equation 11 has two stable steady states x_1^* and x_2^* such that $x_1^* < x_2^*$, i.e. $x_{1_i}^* < x_{2_i}^* \ \forall i = 1(i)N$, then the corresponding digraph generated by the Jacobian (A) of $f(x)$ with respect to x contains at least one positive feedback loop involving all activation interactions.

A more general case where there exists no particular relationship between x_1^* and x_2^* was resolved by first transforming the system in equation 11 by multiplication with a suitable diagonal matrix $P \in R^{N \times N}$ constructed by the following process:

$$P_{ii} = \begin{cases} 1 & x_{1_i}^* > x_{2_i}^* \\ -1 & otherwise \end{cases}$$

It can be verified that P exhibits projection matrix-like properties such as $P^2 = I$. Further, the transformed state space where $x' = Px$ has steady states Px_1^* and Px_2^*.

On applying the previous result on the transformed system, it can be concluded that the transformed system matrix must contain at least one positive feedback with all positive regulation in the network. This proves the fact that A should also consist of at least one positive feedback loop.

As has been argued before, due to sensitivity to initial conditions and parameter variations, a linear treatment to finding design principles for oscillation can be ruled out. Further, as we know that a first-order, nonlinear dynamical system without any delay always produces a monotone/quasi-monotone response, the chances of the same producing an oscillatory response can be eliminated. Therefore, we turn to nonlinear systems with orders more than unity. The existence of oscillatory behavior in any dynamical system can be traced from a stable limit circle in the phase space. A stable limit circle is called locally stable if the neighborhood trajectories converge to the limit circle. This implies that the vector field inside a stable limit circle is locally outwards and points towards the limit circle outside of it.

4.2 *Negative feedback loops: are they necessary for oscillation?*

Similar to biological switches, Snoussi (1998) provided the necessary condition for sustained oscillation in the same work. Given a network with the underlying dynamics mentioned in equation 10, Snoussi (1998) argued that if the Jacobian of $f(x)$ with respect to x represents a complete graph, i.e., each node is reachable from every other node in the network and the field $f(x)$ satisfies condition 2 (repression→negative, activation→positive), then the network requires at least one negative feedback involving more than one node to provide sustained oscillation. This was proved by contradiction, i.e., it was supposed that a complete graph with no negative feedback could provide oscillation. It then follows that in the case of a complete graph with all the loops being positive, all the paths from the ith to the jth node are of the same sign, which is constructed by multiplication of individual signs (-1 for activation and $+1$ for repression) of every edge involved in a given path. As the next step, a transformation similar to **P** defined for the case of multistability was carried out in the following manner:

$$\dot{x^i} = g(x') \qquad (12)$$

where the Jacobian matrix (A) associated with $f(x)$ and the same (A') for $g(x)$ are related as

$$A' = P^{-1}AP \tag{13}$$

Snoussi (1998) proved that if the network structure induced by A does not contain any negative feedback and all the paths are (activating) positive, then the off-diagonal elements of matrix A' are positive – this satisfies Kamke's theorem on monotone systems. Therefore, the associated flow of the dynamical system can be expressed as a monotonically increasing vector-valued function of the initial condition. Subsequently, Snoussi (1998) showed that systems satisfying Kamke's theorem could not contain a stable limit cycle, thereby eliminating the possibility of a complete network with no negative feedback loop to provide oscillation.

Although the above result on the design principles for the oscillatory network is revealing, it is to be remembered that these results on oscillation are only applicable to a complete network containing an arbitrary number of nodes. Apart from the relaxation of the completeness condition, future scopes include deducing the conditions for oscillation in a dynamical system of arbitrary order.

4.3 *Biological adaptation: all possible network topologies*

Unlike oscillation, there has been a flurry of approaches dedicated to discovering the design principles for adaptation. To begin with, Sontag argued that adaptation, in essence, is a disturbance rejection problem. Therefore, an internal model principle (IMP) should be the operating mechanism behind the networks capable of adaptation. The IMP states that in order to reject a disturbance $D(s)$, either the plant or the control must contain a copy of $U(s)$ within. According to IMP, adaptation to a step-type disturbance requires the presence of a step-type component ($\frac{k}{s}$) in the network structure. Subsequent works provided the necessary mathematical condition for adaptation for a linear time-invariant dynamical system (Drengstig *et al.* 2008, 2011; Waldherr *et al.* 2012; Bhattacharya *et al.* 2018). Given the dynamics underlying a biochemical network in equation 10 with N nodes, the linearized system can be written as

$$\dot{\delta}\mathbf{x} = \mathbf{A}\delta\mathbf{x} + \mathbf{B}\delta\mathbf{u} \tag{14}$$

$$\dot{\delta}y = \mathbf{C}\delta\mathbf{x} + \mathbf{D}\delta u \tag{15}$$

where $\delta x \in R^N$ is the deviation variables computed by taking the difference between $x(t)$ and the steady state x^*, and $A \in R^{N \times N}$ and $B \in R^{N \times 1}$ are the Jacobian matrices of $f(x)$ in equation 10 evaluated at x^*, u^*, with respect to x and u, respectively. Similarly, C and D are obtained by evaluating the Jacobian of $h(x)$ at x^*, u^* with respect to the states and the inputs respectively.

Waldherr *et al.* (2012) showed that for the system in equations 14–15 to provide adaptation, the Schur complement of A corresponding to the matrix $H = [ABCD]$ has to be zero. Evidently, this can be evaluated from the internal model principle argument. Further, Bhattacharya *et al.* (2018) used a similar approach to deduce the conditions for perfect adaptation for a three-protein network to the requirement of zero-gain dynamical system – this condition when mapped back to the realm of network structures refers to the requirement of a negative feedback with a buffer node or incoherent feed-forward structure.

Subsequently, Araujo and Lance (2018) and Wang *et al.* (2021) extended these results in to networks of arbitrary size. Both the works used the conditions for adaptation as a requirement of zero final gain along with a weak form of the stability condition $sgn(det(A)) = -1^N$ to deduce that there exist only two kinds of network structures that can provide adaptation – the balancer and the opposer modules. The balancer module contains at least one feedback loop that facilitates an integral control action to the output node. Araujo and Lance (2018) conjectured that for stability purpose, it is necessary for the balancer module to contain at least one negative feedback loop. On the other hand, the opposer module achieves perfect adaptation through multiple forward paths from the input-receiving to the output node with mutually opposite effects (signs). Recently, Bhattacharya *et al.* (2022) proved the Araujo conjecture for networks of any size to establish the necessity of negative feedback in balancer modules for perfect adaptation (figure 8).

Although the structure plays a determining role, attainment of perfect adaptation is only ensured by the right choice of the parameter sets.

For both opposer and balancer modules, it requires the associated rate constants to satisfy specific equality constraints to produce the desired response for adaptation. Therefore, the actual nonlinear system with parameter uncertainty often leads to a condition where the response contains a high but finite precision value. This phenomenon is also known as imperfect adaptation. Bhattacharya *et al.* (2021) showed that for a two-and three-node network, an adaptation-like response
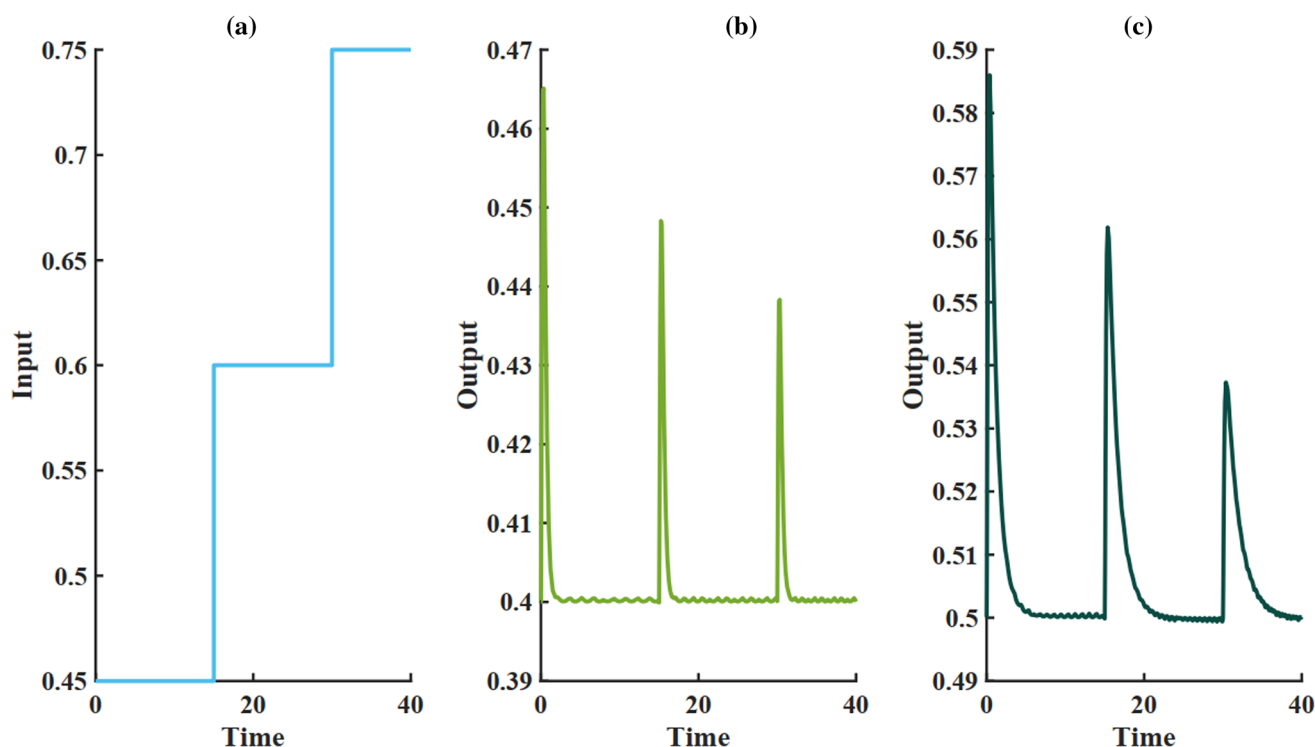
Figure 8. Example of adaptive behavior. (**a**) The staircase-like input. The step changes occurred at t = 0, 15, and 30 min. (**b**) The response of a three-node network involving two mutually opposing forward path from the input node A to the output node C. It is noteworthy that the response of an IFFLP is always non-oscillatory for the hyperbolic nature of the system matrix A. (**c**) The response of another three-protein network with negative feedback between nodes A and B. The associated rate equations along with the rate constants both for (**a**) and (**b**) have been provided in the Appendix.

could be obtained if and only if the zero of the underlying dynamical system is placed before the pole of the same with respect to the origin. The structural networks can be seen to remain unaltered for at least smaller (two- and three-node) networks.

The systems-theoretic approach, despite its limited success in the context of oscillations, has been the most suitable methodology to understand the governing network structure behind the emergence of a property. Due to the absence of both the computational burden and the need to find sufficient structural conditions, this method can be used for deducing all possible structural conditions behind a particular response. Further, certain assumptions, as stressed by Sontag (2007), empower the user to adopt principles for understanding larger networks and make reliable predictions on the mapping between structure and functionality (Sontag 2007).

5. Discussion and future scope

The three methods discussed in light of the three functionalities, namely, toggle switches, oscillation, and adaptation, have been very useful in unraveling the network structures. In spite of this, they have a number of necessary yet limiting assumptions. First, modeling, the stepping stone for any real-world system analysis, requires assumptions that are not always satisfied in reality. For instance, it is well known that the Michaelis–Menten kinetics, which have been used as the model rate kinetics for most of the computational approaches, assume that the binding reaction occurs at a rate faster than the synthesis reaction of the product. Therefore, the scope of any particular approach can be assessed through the assumptions made beforehand.

Second, most of the models aimed at predicting the concentration profile of different species are limited by an inherent assumption that the reactants are in close proximity with each other and there is no spatial gradient at play. This reduces the complex partial differential equations to a set of nonlinear ordinary differential equations, thereby increasing the possibility of a well-defined and bounded solution of the concentrations of each node species of the network.

It is to be noted that the very problem of discovering network structure associated with a particular functionality has an inherent assumption that the particular topology is exclusively tuned for performing that

particular functionality. On the contrary, the network structure, in reality, is only a part of an extensive biological network. Therefore, considering a network structure in isolation can result in unreliable prediction despite being a necessary and useful abstraction. A specific problem of this kind has been addressed by studying retroactivity and context-dependence, where Del Vecchio (2013) showed that a two-protein network with negative feedback changes from oscillation to an exponentially stable response when connected with the downstream system.

In summary, it is necessary to consider all of the above limitations in order to design biological networks. Since all of the afore-explained methodologies rely on certain important aspects of the modeling, the inherent assumptions also seep into the entire process. This merits a comparative evaluation of the three methodologies based on scalability, generalizability, and exclusivity as we set out to gauge a particular methodology's suitability for a given situation.

Given the reaction kinetics, scalability investigates how efficiently a particular methodology performs with increasing network size. As mentioned earlier, computational screening scans through the entire network topology space to find the network structures that satisfy a pre-defined performance criterion. Typically, for an N-node protein network, the number of simulations (N_s) required to be performed in the computational screening method can be expressed as

$$N_s = 3^{N^2} \times N_p$$

where N_p is the number of samples drawn from the parameter space in order to draw an assessment of the robustness of a particular network structure. Clearly, it becomes extremely burdensome to scale up the method for networks of large size.

On the other hand, the rule-based method starts with a preconceived design idea for the associated functionality. The challenge for rule-based methods when operated on networks of large size lies in demonstrating how the large network structures imbibe the predefined design strategy for the given functionality instead of answering the question of what are the possible network structures that can produce the desired response. This is why a scaling up of the network size does not become a worrisome issue for this methodology, given the design strategy is already in place for the functionality in question.

The systems-theoretic methodologies employ a search-oriented approach wherein the performance parameters that characterize the given functionality are first evaluated in the ideal scenario. Followed by this, essentially, these parameters are translated to some system requirements. Subsequently, the structural conditions behind the given functionality are deduced for networks of large size. It is important to note that the assumption that the dynamical systems underlying biochemical networks are quasi-monotonous plays an instrumental role in the last step of this methodology. Therefore, as long as the mapping between the performance parameters and the properties of the dynamical system is well-established, a search-oriented approach can yield reliable results for networks of any size.

The quality of generalizability assesses whether the design principles predicted through a particular methodology work well for different rate kinetics. According to the famous hypothesis on conservation of design principles, the network structure plays a governing role in the nature of the system response irrespective of the organisms and the different levels of the central dogma (Ma *et al.* 2009). Therefore, predictions on admissible network structures produced by a particular methodology should not depend on the particularity of the rate kinetics. This leads us to examine the above three methodologies in light of generalizability.

As discussed earlier, the computational screening methods involve simulating the topology parameter space, which requires explicit knowledge of the rate dynamics to generate the response. In the case of the methods that rely on mixed-integer nonlinear programming, the objective function, in reality, is a function of the rate dynamics. This makes the computational methods dependent on the particularities of the rate dynamics. On the other hand, due to the knowledge of the rate dynamics up front, it becomes very easy to evaluate the robustness to parameter changes in these methods (figure 9). In fact, the Q-measure provided by Ma *et al.* (2009) in the context of adaptation is the widely used, standard measure of robustness for any particular functionality.

Since the rule-based methods aim at a network that can provide the desired functionality with the pre-conceived design strategies, it does not depend on the particularities of the rate kinetics. For instance, the design strategy of negative feedback with delay and nonlinearity that inspired a number of novel works on designing networks for biological oscillation is applicable to almost all time-invariant nonlinear systems. Conversely, this also makes the calculation of robustness of a given network structure with respect to parameter uncertainty difficult to compute.

The systems-theoretic approaches hitherto used for biochemical networks suppose that the Jacobian matrix
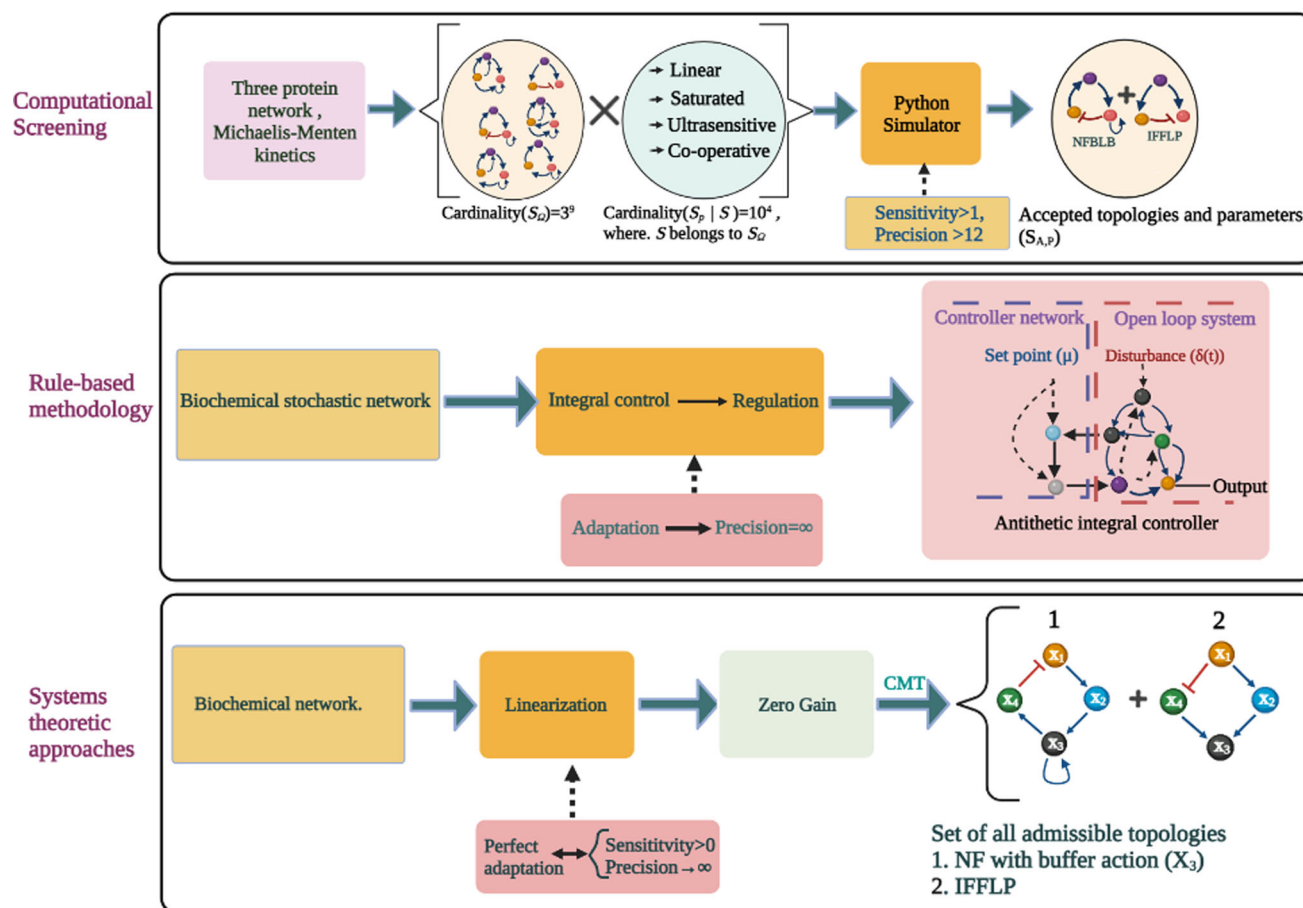
Figure 9. An illustration of three methodologies at work. The functionality considered here is adaptation. The illustration for computational screening approach is inspired from Ma *et al.* (2009), wherein the entire topology–parameter space was searched exhaustively with the dynamics being Michaelis–Menten in order to obtain the all possible network structures – negative feedback (NF) and incoherent feed-forward structure (IFFLP) for three-node network. The schematic shown for rule-based methodology is inspired from the work by Briat *et al.* (2016), which designed a negative feedback induced integral control strategy to achieve perfect regulation in a stochastic environment. Finally, the illustration on systems-theoretic approaches demonstrates the work by Bhattacharya *et al.* (2018), which discovered the necessary structural conditions for a network of any size to provide perfect adaptation. The topology 'NF' refers to negative feedback.

of the actual nonlinear vector field acts as the digraph matrix of the underlying network structure. This assumption has been shown to be satisfied for most biochemical networks. Therefore, the need for explicit knowledge of the rate dynamics can be circumvented, and the resultant network structures are truly generalizable in terms of different rate kinetics (figure 9). The network structures for adaptation produced by Bhattacharya *et al.* (2018) using an LTI systems approach stands true both in the case of Hill or Michaelis–Menten kinetics.

Last but not the least, exhaustivity determines whether a particular methodology can detect the entire set of structural conditions admissible to the functionality of interest. For the purpose of designing, obtaining an efficient design principle that can deliver the requirements is sufficient, but in order to understand how the biology works and also the evolution of a particular network pattern, it is important to find all the possible network structures admissible for the desired functionality.

Given the reaction kinetics of the network, the computational screening examines the entire possibility space of the topology. Similarly, in the case of the optimization-based approach, the set of all possible structural possibilities are captured through different combinations of the *N*-bit number. This *ipso facto* encompasses the entire possibility space through evaluating the objective function at all possible combinations of the *N*-bit binary sequence. Therefore, in principle, computational screening provides a reliable prediction in terms of the exhaustivity of the network structures.

Table 1. A consolidated representation of three different methods in discussion

Methodology	Scalability	Generalizability	Exhaustivity
Computational screening	✗ Highest size till date: four nodes	✗ Requires specific knowledge of the kinetics.	✓ Network size and the specific underlying rate kinetics should be known ➢ Generates exhaustive predictions.
Rule based	✓ ➢ Ideally, can be used for networks of any size. ➢ In synthetic design, the maximum number of nodes used is four.	✓ The dynamical system abides by certain broad assumptions (cite) ➢ This method can be applied to any dynamical system .	✗ ➢ For a given network, the structural predictions proposed by rule based methods are often non-exhaustive.
Systems theoretic	✓ There exists a well-defined mathematical relation between the performance parameters and the properties of the dynamical system. ➢ It is always possible to scale up this method for networks of any size.	✓ The Jacobian of the dynamical system *wrt* the states serve as the digraph matrix of the actual network. ➢ This method does not possess any explicit dependence on the particularities of the reaction kinetics.	✓ A well-defined mapping between the performance parameters and the mathematical properties of the system exists. ➢ This method yields all-encompassing set of admissible topologies.

The ticks symbolizes the ability of a methodology to satisfy a particular property, whereas the crosses represent the inability of the same. In all the cases, the underlying rate kinetics has been assumed to constitute a well-posed dynamical system.

The rule-based method, on the other hand, relies on a particular design strategy, thereby focusing on the efficient design principle rather than the entire possibility set. For instance, early rule-based initiatives on adaptation included only the network structures with negative feedback that facilitate integral control. Therefore, it is unlikely that a rule-based design strategy can unravel all possible admissible structural conditions for any given functionality.

The system-theoretic efforts start with the basic characterization of the functionality at hand and then build a bottoms-up approach for obtaining the full set of design principles. This becomes possible if a finite number of mappings exist between the performance parameters and the properties of the underlying dynamical system. Therefore, if all these conditions are satisfied, it is possible to arrive at an exhaustive set of design principles using this approach. Unlike the brute force methods, first, the admissible networks are obtained from the possibility space. Therefore, the optimization problem is solely focused on the parameters (rate constants), not the structure, thereby reducing the curse of dimensionality to a great extent (figure 9).

Although it might seem, from table 1, that the systems-theoretic approaches are the most sophisticated and effective among all the three methodologies, they contain a number of limiting assumptions. The first among these stems from the assumption of well-posedness of the rate kinetics, which, in reality, may not hold true for every biological network. Second, there exist no necessity theorems till now that guarantee a mapping between the performance and the systems-theoretic parameters. Therefore, the set of functionalities for which this method yields reliable structural predictions is still small in size compared to the computational approaches. On the other hand, computational screening is simple and extremely effective for small-scale networks. As mentioned earlier, the rule-based methods are best suited for synthetic design. Therefore, clearly no best methods exist. In fact, it can be observed that our knowledge of the design

principles of these important functions has always progressed by facilitating a cross-talk between these three methodologies. Typically, after the experimental confirmations of the occurrence, the similarity between the response in the engineering system and these functionalities are evaluated. This inspires a bit of rule-based efforts to design the network structure synthetically with the assistance of the engineering principle. Subsequently, the limitations discovered in the course of rule-based design and further experimental probings sometimes fetch an entirely different class of admissible network structures for a given functionality, thereby necessitating a computational screening of the entire topology–parameter space of small-scale networks. The results obtained from the computational screening and experiments serve as the Rosetta stone for subsequent works on the lines of systems theory, which then strive to obtain the necessary structural conditions for the functionality in a scalable, generic, and exhaustive manner.

Appendix

Equation for single node oscillator with delay:

$p = 2$, $k_m = 1$, $K_d = 1$, $S = 1$, $K_1 = 1$, $K_2 = 1$, $E_t = 1$, $\tau = 10$

$$\dot{x} = K_1 S K_d^p / \left(K_d^p + x(t - \tau)^p \right) - K_2 E_t x / (K_m + x);$$

Equation for four-node oscillator without explicit delay:

$p = 2$, $S = 1000/p$, $K_{dx1} = 10$, $K_{exp} = 0.2$, $K_{dx2} = 0.2$, $\varepsilon = 1$, $K_{dx4} = 8/p$, $K_{dx3} = 0.1$, $K_m = 0.1$, $K_{imp} = 0.1$;

$$\dot{x_1} = K_{dx1}\left(S/\left(1 + x_4^p\right) - x_1 - K_{exp}x_1\right)$$
$$\dot{x_2} = \varepsilon * K_{exp} * x_1 - K_{dx2}x_2$$
$$\dot{x_3} = K_{dx3} * (x_2 - x_3) - \varepsilon K_{imp} * x_3$$
$$\dot{x_4} = K_{imp} * x_4 - K_{dx4}x_4 / (K_m + x_4)$$

Equation for three-node IFFLP:

$$\dot{x_1} = I \frac{1 - x_1}{(0.0001 + (1 - x_1))} - 2x_1$$
$$\dot{x_2} = 10x_1 \frac{1 - x_2}{(0.0001 + (1 - x_2))} - 200x_2x_3$$
$$\dot{x_3} = 0.1x_2(1 - x_3)/(1.001 - x_3) - x_3$$

Equation for three-node NFBLB:

$$\dot{x_1} = I \frac{1 - x_1}{(0.0001 + [1 - x_1])} - 2x_1$$
$$\dot{x_2} = 350x_1(1 - x_2) - 350x_2x_3$$
$$\dot{x_3} = 2x_2 \frac{(1 - x_3)}{(1.001 - x_3)} - x_3/(0.001 + x_3)$$

References

Ananthasubramaniam B and Hanspeter H 2014 Positive feedback promotes oscillations in negative feedback loops. *PLoS One* **9** e104761

Angeli D, Ferrell Jr and Sontag Ed 2004 Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proc. Natl. Acad. Sci. USA* **101** 1822–1827

Araujo R and Lance L 2018 The topological requirements for robust perfect adaptation in networks of any size. *Nat. Commun.* **9** 1757–1769

Astrom K and Richard M 2010 *Feedback systems: An introduction for scientists and engineers* (Princeton University Press)

Barkai N and Stan L 1997 Robustness in simple biochemical networks. *Nature* **387** 913–921

Bernardo M and Tu Y 2003 Perfect and near-perfect adaptation in a model of bacterial chemotaxis. *Biophys. J.* **84** 2943–2956

Bhattacharya P, Karthik R and Tangirala A 2018 A systems-theoretic approach towards designing biological networks for perfect adaptation. *IFAC-PapersOnLine* **51** 307–312

Bhattacharya P, Karthik R and Tangirala A 2021 Systems-theoretic approaches to design biological networks with desired functionalities. *Methods Mol. Biol.* **2189** 133–155

Bhattacharya P, Karthik R and Tangirala A 2022 Discovering adaptation-capable biological network structures using control-theoretic approaches. *PLOS Comp. Biol.* **18** e1009769

Börsch A and Jörg S 2016 How time delay and network design shape response patterns in biochemical negative feedback systems. *BMC Syst. Biol.* **10** 82–90

Briat C, Gupta A and Khammash M 2016 Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Syst.* **2** 15–26

Briat C, Gupta A and Khammash M 2018 Antithetic proportional-integral feedback for reduced variance and improved control performance of stochastic reaction networks. *J. R. Soc. Interface* **15** 20180079

Cameron E, Bashor C and Collins J 2014 A brief history of synthetic biology. *Nat. Rev. Microbiol.* **12** 381–390

Chin C, Chubukov V, Jolly E, *et al*. 2008 Dynamics and design principles of a basic regulatory architecture controlling metabolic pathways. *PLoS Biol.* **6** 1343–1356

Del D 2013 A control theoretic framework for modular analysis and design of bio-molecular networks. *Annu. Rev. Control* **37** 333–345

Diegmiller R, Zhang L, Marcio G, *et al*. 2021 Mapping parameter spaces of biological switches. *PLoS Comp. Biol.* **17** e1008711

Drengstig T, Ueda H and Ruoff P 2008 Predicting perfect adaptation motifs in reaction kinetic networks. *J. Phys. Chem. B* **112** 16752–16758

Drengstig T, Kjosmoen T and Ruoff P 2011 On the relationship between sensitivity coefficients and transfer functions of reaction. *J. Phys. Chem. B* **115** 6272–6278

Dubnau D and Losick R 2006 Bistability in bacteria. *Mol. Microbiol.* **61** 564–572

El-Samad H 2021 Biological feedback control: respect the loops. *Cell Syst.* **12** 477–487

Fahrenkrug J, Birgitte G, Hannibal J, *et al*. 2006 Diurnal rhythmicity of the clock genes *Per1* and *Per2* in the rat ovary. *Endocrinology* **147** 3769–3776

Fangzhou X, Khammash M and Doyle J 2021 Stability and control of biomolecular circuits through structure. *2021 American Control Conference* pp 476–483

Ferrell Jr 2016 Perfect and near-perfect adaptation in cell signaling. *Cell Syst.* **2** 62–67

Fiedler B, Mochizuki A, Kurosawa G, *et al*. 2013 Dynamics and control at feedback vertex sets. I: Informative and determining nodes in regulatory networks. *J. Dyn. Diff. Equat.* **25** 563–604

Frank F 1974 Kinetic feedback processes in physicochemical oscillatory systems. *Faraday Symp. Chem. Soc.* **9** 137–149

Friedlander T and Brenner N 2009 Adaptive response by state-dependent inactivation. *Proc. Natl. Acad. Sci. USA* **106** 22558–22563

Gardner S, Cantor C and Collins J 2000 Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403** 339–342

George S 2002 Role of optimal control theory in cancer chemotherapy. *Math. Biosci.* **101** 237–284

Goldbeter A 2002 Computational approaches to cellular rhythms. *Nature* **420** 238–245

Goldbeter A and Lefever R 1972 Dissipative structures for an allosteric model. *Application to Glycolytic Oscillations. Biophys. J.* **12** 1302–1315

Goldbeter A 1996 *Biochemical oscillations and cellular rhythms: The molecular bases of periodic and chaotic behavior* (Cambridge University Press)

Griffith S 1968 Mathematics of cellular control processes. i. Negative feedback to one gene. *J. Theor. Biol.* **20** 202–208

Guantes R and Poyatos J 2008 Multistable decision switches for flexible control of epigenetic differentiation. *PLoS Comp. Biol.* **4** e1000235

Hardin E, Hall J and Rosbash M 1992 Circadian oscillations in period gene mRNA levels are transcriptionally regulated. *Proc. Natl. Acad. Sci. USA* **89** 11711–11715

Hat B, Kochan C, Bogdal M, *et al*. 2016 Feedbacks, bifurcations and cell fate decision-making in the P53 system. *PLoS Comp. Biol.* **12** e1004787

Hespanha J and Sivakumar H 2013 Towards modularity in biological networks while avoiding retroactivity. *Proceedings of the American Control Conference* pp 4550–4556

Higgins J 1964 A chemical mechanism for oscillation of glycolytic intermediates in yeast cells. *Proc. Natl. Acad. Sci. USA* **51** 989–994

Hinczewski M and Thirumalai D 2014 Cellular signaling networks function as generalized Wiener-Kolmogorov filters to suppress noise. *Phys. Rev. X* **4** 04101701–04101715

Jacob F and Monod J 1961 Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3** 318–356

Jorge Z, Gang Y and Albert R 2017 Structure-based control of complex networks with nonlinear dynamics. *Proc. Natl. Acad. Sci. USA* **114** 7234–7239

Joshi S, Sen S and Kar N 2020 Synchronization of master-slave oscillators: analysis and experimental results. *IFAC-PapersOnLine* **53** 226–231

Khammash M 2021 Perfect adaptation in biology. *Cell Syst.* **12** 509–521

Königs V, Camila O, Freitas M, *et al*. 2020 Srsf7 maintains its homeostasis through the expression of Split-ORFs and nuclear body assembly. *Nat. Struc. Mol. Biol.* **27** 260–273

Konopka K 2006 *Systems biology: Principles, methods, and concepts* (CRC Press)

Kulkarni V, Stan GB and Raman K 2014 *A systems theoretic approach to systems and synthetic biology I: Models and system characterizations* (Springer)

Laurent M and Nicolas K 1999 Multistability: A major means of differentiation and evolution in biological systems. *Trends Biochem. Sci.* **24** 418–422

Lazebnik Y 2002 Can a biologist fix a radio?—or, What I learned while studying apoptosis. *Cancer Cell* **2** 179–182

Leon M, Mae W, Alex F *et al*. 2016 A computational method for the investigation of multistable systems and its application to genetic switches. *BMC Syst. Biol.* **10** PMC5142341

Li Z, Liu S and Yang Q 2017 Incoherent inputs enhance the robustness of biological oscillators. *Cell Syst.* **5** 72–81

Liu Y, Slotine J and Barabási A 2011 Controllability of complex networks. *Nature* **473** 167–173

Losick R and Desplan C 2008 Stochasticity and cell fate. *Science* **320** 65–68

Lu J, Sherman D, Devor M, *et al*. 2006 A putative flip–flop switch for control of REM sleep. *Nature* **441** 589–594

Ma W, Trusina A, Hana S, *et al*. 2009 Defining Network topologies that can achieve biochemical adaptation. *Cell Syst.* **138** 760–773

Ma'ayan A, Lipshtat A, Iyengar R, *et al*. 2008 Proximity of intracellular regulatory networks to monotone. *IET Syst. Biol.* **2** 103–112

Macarthur D, Ma'ayan A and Lemischka R, 2009 Systems biology of stem cell fate and cellular reprogramming. *Nat. Rev. Mol. Cell Biol.* **10** 672–681

Mackey C and Glass L 1997 Oscillation and chaos in physiological control systems. *Science* **197** 287–289

Maybee J, Driessche P, Olesky D, *et al*. 1989 Matrices, digraphs, and determinants. *SIAM* **10** 500–519

Milo R, Shai O, Shalev I, *et al*. 2002 Network motifs: simple building blocks of complex networks. *Science* **298** 824–827

Novak B and Tyson J 1993 Numerical analysis of a comprehensive model of m-phase control in Xenopus oocyte extracts and intact embryos. *Cell Sci.* **106** 153–168

Novak B and Tyson J 2008 Design principles of biochemical oscillators. *Nat. Rev. Mol. Cell Biol.* **9** 981–991

Oberortner E, Bhatia S, Erik L, *et al*. 2015 A rule-based design specification language for synthetic biology. *J. Emerg. Technol. Comput. Syst.*. *https://doi.org/10.1145/2641571*

Otero-Muras I and Banga J 2016 Design principles of biological oscillators through optimization: forward and reverse analysis. *PLoS One* **11** e0166867

Otero-Muras I and Banga J 2019 Distilling robust design principles of biocircuits using mixed integer dynamic optimization. *Processes* **7** e7020092

Ozbudak M, Thattai M, Han L, *et al*. 2004 Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **427** 737–740

Pett J, Kondoff M, Bordyugov G, *et al*. 2018 Co-existing feedback loops generate tissue-specific circadian rhythms. *Life Sci. Alliance* **1** e201800078

Plahte E, Mestl T and Omholt J 1995 Feedback circuits, stability and multistationarity in dynamical systems. *J. Biol. Syst.* **3** 409–413

Pomerening R, Kim Y and Jr Ferrell 2005 Systems-level dissection of the cell-cycle oscillator: bypassing positive feedback produces damped oscillations. *Cell* **122** 565–578

Qiao L, Zhao W, Tang C, *et al*. 2019 Network topologies that can achieve dual function of adaptation and noise attenuation. *Cell Syst.* **9** 271–285

Roenneberg T and Merrow M 2003 The network of time: understanding the molecular circadian system. *Curr. Biol.* **13** 198–207

Rössler E 1972 A principle for chemical multivibration. *J. Theor. Biol.* **36** 413–417

Saez L and Young M 1996 Regulation of nuclear entry of the *Drosophila* clock proteins period and timeless. *Neuron* **17** 911–920

Snoussi E 1998 Necessary conditions for multistationarity and stable periodicity. *J. Biol. Syst.* **6** 3–9

Sontag Ed 2007 Monotone and near-monotone biochemical networks. *Syst. Synth. Biol*. **1** 59–87

Tau-Mu Y, Huang Y, Simon M, *et al*. 2000 Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl. Acad. Sci. USA* **97** 4649–4653

Thomas R 1981 On the relation between the logical structure of systems and their ability to generate multiple steady states or sustained oscillations. In: Della Dora J, Demongeot J, Lacolle B (eds) Numerical methods in the study of critical phenomena. Springer Series in Synergetics, vol 9. (Springer, Berlin, Heidelberg). *https://doi.org/10.1007/978-3-642-81703-8_24*

Torday J 2015 Homeostasis as the mechanism of evolution. *Biology* **3** 573–590

Tyson J 1975 On the existence of oscillatory solutions in negative feedback cellular control processes. *J. Math. Biol.* **1** 311–315

Tyson J, Katherine C and Novak B 2003 Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.* **15** 221–231

Voit E 2012 *A first course in systems biology* (Garland Science)

Waddington C 1957 *The strategy of genes* (Routledge)

Waldherr S, Streif S and Allgöwer F 2012 Design of biomolecular network modifications to achieve adaptation. *IET Syst. Biol.* **6** 223–231

Wang Z, Qi S, Gang H, *et al*. 2016 A geometrical approach to control and controllability of nonlinear dynamical networks. *Nat. Commun*. **7** ncomms11323

Wang Y, Huang Z, Antoneli F, *et al*. 2021 The structure of infinitesimal homeostasis in input–output networks. *J. Math. Biol*. **82** PMC8139887

Wen X and Ferrell J 2003 A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* **426** 460–465