Synchrony Patterns in Gene Regulatory Networks

Manuela A. D. Aguiar

Manuela Aguiar, Faculdade de Economia, Centro de Matemática, Universidade do Porto, Rua Dr Roberto Frias, 4200-464 Porto, Portugal

Ana P. S. Dias

Ana Dias, Departamento de Matemática, Centro de Matemática, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

Haibo Ruan

Haibo Ruan, Institute of Mathematics, Technical University of Hamburg, Am Schwarzenberg-Campus 3, D-21073 Hamburg, Germany

Abstract

Motivated by studying synchronization mechanisms in gene regulatory networks (GRNs) and their relation to evolutionary events such as genetic duplication and genetic redundancy, we consider two mathematical dynamical models of GRNs. We obtain results on robust synchronization on these dynamical models inspired by the existing theoretical results in the coupled cell network formalisms. We also explore the concepts of quotient networks and network lifting in the context of GRNs which are related to the process of gene duplication and the phenomenon of subfunctionalization as an outcome of functional divergence.

Keywords: Gene regulatory network, gene dynamical model, synchronization, gene duplication

2010 MSC: Primary: 92B20, 92B25, 34C14, Secondary: 34C15, 05C90

Email addresses: maguiar@fep.up.pt (Manuela A. D. Aguiar), apdias@fc.up.pt (Ana P. S. Dias), Corresponding author: haibo.ruan@tuhh.de (Haibo Ruan)

$\mathrm{act}(\cdot)$	activation regulatory function
$\mathrm{rep}(\cdot)$	repression regulatory function
m_i	concentration of mRNA in gene i
p_i	concentration of protein in gene i
$x_i = (m_i, p_i)^T$	dynamical variable of gene i
A_i	2×2-matrix for internal dynamics of gene i
$I_i^{\pm}$	the set of indices j where gene j activates resp. represses gene i
$W^{\pm} = (w_{ij}^{\pm})$	weight adjacency matrix for activation resp. repression with "$+$" resp. "$-$"
$w_{ij}^{\pm}$	non-negative weight in activation resp. repression from gene j to gene i
$M^{\pm} = (m_{ij}^{\pm})$	multiplicity adjacency matrix for activation resp. repression
C	the set of all gene indices
P	a partition of C
C_k	the k-th equivalence class under P
$[i]$	the equivalence class of i under P
$Q^{\pm} = (q_{ij}^{\pm})$	quotient weight adjacency matrix for activation resp. repression
$N^{\pm} = (n_{ij}^{\pm})$	quotient multiplicity adjacency matrix for activation resp. repression

Table 1: List of main notations

1. Introduction and Motivation

Gene regulatory networks (GRNs) are biochemically interacting systems that are composed of molecular regulators in a cell which interact with each other directly or indirectly through their RNA and protein expression products ([42, 58]). They underlie genetic regulatory mechanisms that determine and control cellular functions such as cell operations, cell-cycle progression and responses to environmental signals ([44]). With the continuous progress in genome sequencing technology, an enormous amount of experimental data on gene expression and regulation have been made available, which may provide the first step towards understanding how cells survive, reproduce and adapt ([51]). However, given the structural and dynamical interconnectedness of genomes, many cellular processes that control the development and adaptation of multicellular life forms still remain widely undetermined ([12, 52]). One such example is the role of gene duplication, genetic redundancy and functional divergence in network evolution ([31, 54]).

Mathematical models that offer tractable analysis of dynamical processes in GRNs and their collective behaviour are sought after for understanding, predicting and controlling regulatory mechanisms, see for example [6, 36, 37]. Among others, the theory of coupled cell networks and their associated dynamical systems, proposed by Golubitsky and Stewart [27], and by Field [23], have provided mathematical models for understanding collective dynamics on general coupled networks such as robust synchronization and synchrony-related bifurcations [53, 23, 29, 27]. Coupled cell systems are dynamical systems (*cells*) that are coupled together through mutual interactions, abstracted by the associated network, which exert influences on the temporal evolution of each other. A key advantage of applying these formalisms to analyse GRNs is that they allow theoretical deduction of collective dynamics based only on the underlying network structure, without referring to the specifics of each gene. This leads to theoretical results that are relatively independent of modelling specifics of individual genes and are more importantly, naturally compatible with structure-related dynamical processes in GRNs.

For example, the existence of different robust *patterns of synchrony* for a GRN where each pattern has one or more *clusters of synchronized genes*. Here, two genes are said synchronized if their state coincides for all time. In terms of the dynamics on the phase space, this means that a solution where certain genes are synchronized at one time remain synchronized for all time. Thus, each pattern of synchony corresponds to a *synchrony subspace*, that is, to a subspace described in terms of equalities of gene coordinates that is flow-invariant under the dynamics of every coupled cell system - any solution that starts in that space remains there for all time. The synchronization of the genes in the same cluster (class) is independent of the their internal dynamics which can range from a steady-state to chaotic dynamics. Moreover, it is robust for the dynamical models (all the admissible vector fields).

Morone, Leifera and Maksea [46] introduce the use of network fibration symmetries in the analysis of the transcriptional regulatory network of bacterium *Escherichia coli* where nodes are genes and a directed link represents a transcriptional regulation. They formalize the process of communication between genes with the notion of *input tree* of the gene where the input tree of a gene formalizes the information arriving to a gene containing the entire history transmitted through all pathways that reach that gene. Moreover, genes with isomorphic input trees form clusters of genes and the set of all input tree isomorphisms defines the *symmetry fibrations* of the network. They emphasize that the fibration symmetries in GRNs may identify *clusters of synchronized genes*

(genes with isomorphic input trees) which can be collapsed, each group into a single gene, as symmetry fibrations are transformations that preserve the dynamics of information flow in the network. See Figure 1.

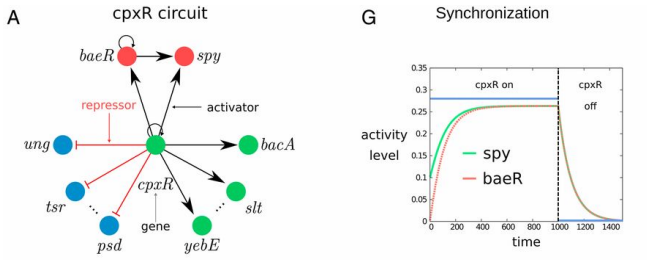

Figure 1: Part of [46, Figure 1] (reproduced with permission from [46]). Left (A) The *Escherichia coli* subcircuit controlled by the *cpxR* gene which regulates its own expression (via an autoregulation activator loop) and also regulates other genes grouped by the different colors. The circuit regulates more genes represented by the dotted lines which are not displayed for simplicity. Gene *cpxR* is not regulated by any other transcription factor in the network. Right (G) Symmetric genes *baeR* and *spy* when activated by its regulator *cpxR* can synchronize their activity to produce the same activity levels. See [46] for details including the models used for simulation.

In [18], DeVille and Lerman show that robust synchrony is a consequence of the existence of graph fibrations and, in particular, that surjective graph fibrations lead to synchrony subspaces. Moreover, as expected, network symmetries imply network synchronies. In fact, if a network has permutation symmetry Γ (by permuting the cells) and Σ is a subgroup of Γ, then the fixed point subspace of Σ is a synchrony subspace. Nevertheless, as pointed out by Antoneli and Stewart [9], for general networks there can be synchrony subspaces that are not forced by the symmetries of the network, if any. In particular, this applies to GRNs. The collapse of each cluster of genes into a single gene corresponds to the notion of quotient network in the coupled cell network formalism. See, for example Aguiar, Dias, Golubitsky and Leite [5] and Aguiar and Dias [3]. The quotient network associated to a pattern of synchrony is the gene network for the coupled systems in the restriction to the corresponding synchrony subspace.

In this paper, we consider two mathematical dynamical models of GRNs and identify their robust patterns of synchrony inspired by theoretical results from the coupled cell network formalism.

In the context of gene regulations, gene synchronization can be a result of gene duplication. Gene duplication is the primary source of new genes. Dewey and Galas [19] discuss a network growth model based on *gene duplication* pointing out evidence through a variety of genoma analysis that new genes are almost always created by duplication. After duplication, the two gene copies are generally assumed to be entirely redundant, that is, to perform the same function and consequently, the inactivation of one of the genes has little or no effect on the biological phenotype ([47]). As shown in [56, 57], genetic redundancy can be actively maintained by natural selection and it is not just a transient consequence of recent gene duplication. Nevertheless, the fate of most of the duplicated genes that are not lost is to undergo functionalization. Theory suggests three alternative outcomes in the evolution of duplicate genes that are preserved: (i) the conservation of all functions in both duplicates (gene conservation); (ii) the evolution of a new function in one of the duplicates (neofunctionalization); or (iii) the division of ancestral functions among duplicates (subfunctionalization) ([43, 32]). In the duplication-degeneration-complementation model of subfunctionalization the duplicated genes are preserved because each copy of the gene loses some, but not all, of its functions through degenerating mutations. This results in the degenerating mutations in one gene being complemented by the other and vice versa. Thus, the two degenerated genes complement each other and together carry out the same two functions as the preduplication gene. See [24] and [35], for several examples supporting the duplication-degeneration-complementation model for gene duplication, including analysis of a new engrailed gene in zebrafish and yeast genes. In [30], Gout and Lynch propose a refined version of subfunctionalization, quantitative subfunctionalization, where the ancestral gene functions are simply partitioned between the two duplicated genes such that expression levels necessary to produce enough of the final product are retained; no change in biochemical functions or expression patterns (i.e., when/where the genes and their products are expressed) is necessary. For a very recent work on whole-genome duplication see Kuzmin *et al.* [39].

The process of gene duplication for GNRs has its analog in the process of *lifting* in the theory of coupled cell networks. A lift of a coupled cell network is a new network where some of the cells give rise to copies of themselves and where the interactions linking the new cells in the bigger network guarantee that the dynamics of the bigger systems contains the dynamics of the smaller systems. See Aguiar, Dias, Golubitsky and Leite [5] and its extensions to weighted networks in Aguiar and Dias [3] and Ashwin, Aguiar, Dias and Field [1]. A trivial observation is that, in general, there is no uniqueness in this lifting process. That is, fixing a k-cell network, if $n > k$

then there are many n-cell networks which are lifts (or inflations) of the k-cell network. Moreover, there is a method of enumerating the lifts if the k-cell network is fixed. See Theorem 2.5 of Aguiar, Dias, Golubitsky and Leite [5] valid for coupled cell systems following the formalisms of Stewart, Golubitsky and Pivato [53], Golubitsky, Stewart and Török [29] and Field [23], where the edges in the graph network structure have assigned nonnegative integer numbers. See also the extension of this result to weighted networks, where the connections have attached strengths that can be any real number and the coupled cell systems have additive input structure, in Theorem 2.13 of Aguiar and Dias [3]. For both setups, the enumeration method goes through the characterization of the network adjacency matrices of the larger networks determined by the smaller network adjacency matrices. A nice observation in the enumeration method, is that, fixing n, the number of n-cell lifts is finite for nonnegative integer matrices and nonfinite in the weighted setup.

The lifting of a network is the reverse process of quotient. As mentioned earlier, an important aspect of the identification of patterns of synchrony is that, associated with each synchrony pattern, there is a smaller (quotient) network of the GRN, whose dynamics is related with the total GRN when restricted to the corresponding synchrony subspace. This is in accordance with a common technique used in science, from both theoretical and experimental points of view: to first investigate small networks either from dynamical or statistical point of view and then to expand the results to larger networks that can be related to the smaller ones. There are different approaches as to how smaller networks can be related to larger ones. In this paper, we are following the dynamical perspective and the process of expanding a (smaller) quotient network to a bigger network which ensures that the dynamics associated with the smaller network also occurs at the bigger network. See Aguiar, Dias, Golubitsky and Leite [5].

We exemplify next the processes of duplication and synchronization of genes taking a particular GRN.

1.1. Example

Consider the three transcriptional repressor system used by Elowitz and Leibler [22] to build a repressilator network, in *Escherichia coli*. Note that one of the results obtained in [22] is that, depending on the values of several parameters, such as, the dependence of transcription rate on repressor concentration, the translation rate, and the decay rates of the protein and messenger RNA, at least two types of solutions are possible: the system may converge towards a stable steady

state, or the steady state may become unstable, leading to sustained limit-cycle oscillations. See also [15]. In the equations model used in [22], each variable $x_i = (m_i, p_i)$, for $i = 1, 2, 3$, describes the concentrations of two gene products, the mRNAs and proteins, which vary in continuous time, and their time derivatives are expressed as functions of the variables. Graphically, each node in the network represents a gene. Also, an interaction between two nodes is represented by an edge. Moreover, in this example there is only a repression type of interaction with specific positive real weight α. See Figure 2. The *repression weight adjacency matrix* is

$$W^- = \left[w_{ij}^-\right] = \left(\begin{array}{c|c|c} 0 & 0 & \alpha \\ \hline \alpha & 0 & 0 \\ \hline 0 & \alpha & 0 \end{array}\right),$$

where if there is an interaction from gene j to gene i, the entry w_{ij}^- is nonzero and it denotes the positive weight of the repression type interaction. The three-gene network equations used in [22] are:

$$\left\{\begin{array}{l} \dot{x}_1 = A_\beta x_1 + [\alpha \mathrm{rep}(p_3) + \alpha_0] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_2 = A_\beta x_2 + [\alpha \mathrm{rep}(p_1) + \alpha_0] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_\beta x_3 + [\alpha \mathrm{rep}(p_2) + \alpha_0] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} \dot{x}_1 = A_\beta x_1 + \left(\frac{\alpha}{1+p_3^2} + \alpha_0\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_2 = A_\beta x_2 + \left(\frac{\alpha}{1+p_1^2} + \alpha_0\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_\beta x_3 + \left(\frac{\alpha}{1+p_2^2} + \alpha_0\right) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{array}\right., \tag{1.1}$$

where

$$A_\beta = \begin{pmatrix} -1 & 0 \\ 1 & -\beta \end{pmatrix} \quad \text{and} \quad \mathrm{rep}(p) = 1 - \frac{p^2}{1+p^2} = \frac{1}{1+p^2}\,.$$

Thus all the three genes have the same internal (linear) dynamics depending on β and all the weights are equal to α. For each x_i gene equation, only the m_i mRNA equation depends directly on the interaction (with weight α) from other gene j through the p_j protein concentration. Thus the system (1.1) has cyclic permutation symmetry $\mathbf{Z}_3$ on the three nodes which justifies nicely the oscillatory behaviour found in [22], see for example Theorem XVII 8.2 of Golubitsky, Stewart and Schaeffer [28] and Chapter 4 of Golubitsky and Stewart [26]. Also note that the cyclic permutation symmetry $\mathbf{Z}_3$ of the three-gene network in Figure 2 leads to the synchrony subspace $\{x : x_1 = x_2 = x_3\}$.

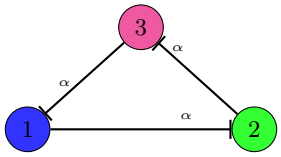

Figure 2: A GRN with three genes. It corresponds to the repressilator composed by a cyclic negative-feedback loop of three repressor genes presented in [22].

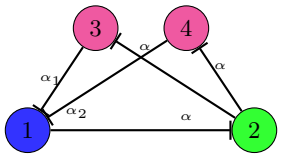

Figure 3: A GRN with four genes that can be interpreted from the three gene GRN of Figure 2 where gene 3 was duplicated giving rise to genes $3, 4$.

Gene duplication example

Consider any 4×4 repression weighted adjacency matrix of the form

$$\left(\begin{array}{c|c|cc} 0 & 0 & \alpha_1 & \alpha_2 \\ \hline \alpha & 0 & 0 & 0 \\ \hline 0 & \alpha & 0 & 0 \\ 0 & \alpha & 0 & 0 \end{array}\right)$$

where $\alpha_1, \alpha_2 > 0$ and consider the four gene GRN in Figure 3. Take the following model equations for the four gene GRN:

$$\left\{\begin{array}{l} \dot{x}_1 = A_\beta x_1 + [\alpha_1 \mathrm{rep}(p_3) + \alpha_2 \mathrm{rep}(p_4) + \alpha_0] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_2 = A_\beta x_2 + [\alpha \mathrm{rep}(p_1) + \alpha_0] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_\beta x_3 + [\alpha \mathrm{rep}(p_2) + \alpha_0] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_4 = A_\beta x_4 + [\alpha \mathrm{rep}(p_2) + \alpha_0] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{array}\right. \tag{1.2}$$

where, as above $\mathrm{rep}(p) = 1/(1+p^2)$.

We can interpret the four-gene GRN in Figure 3 obtained from the three-gene GRN of Figure 2 by *duplication* of gene 3 into genes $3, 4$. The four-gene GRN is said to be a *lift* network of the three-gene GRN of Figure 2 and $\{x : x_3 = x_4\}$ is a synchrony subspace.

If we assume $\alpha_1+\alpha_2 = \alpha$ then we are assuming subfunctionalization as the outcome of functional divergence for gene duplication. Restricting equations (1.2) to the synchrony subspace $\{x : x_3 = x_4\}$ we obtain equations (1.1). In particular, an oscillatory solution of the system (1.1) is in correspondence to an oscillatory solution of system (1.2) where genes $3, 4$ quantities are synchronized for all time. Systems with $\mathbf{Z}_3$ symmetry can support a discrete rotating wave by Hopf bifurcation yielding a traveling wave solution for equations whose restriction is $\mathbf{Z}_3$-symmetric. See the example in [53, pages 613-615]. We have then that the three-gene GRN of Figure 2 is obtained from the four-gene GRN in Figure 3 by *collapsing* the $3, 4$ gene cluster into a single gene 3. This cluster exists due to the fibration symmetries of the GRN of Figure 3. Equivalently, this is due to the facts that both genes $3, 4$ receive a repression type interaction with the same weight from (the cluster formed by the single) gene 2, and the sum of weights of the repression type interactions from the cells in the cluster of $3, 4$ to gene 1 is equal to the weight of the repression type interaction to gene 1. We use the terminology that the three-gene GRN of Figure 2 is a *quotient* of the four-gene GRN of Figure 3 by the synchrony space $\{x : x_3 = x_4\}$. We note that in the subfunctionalization model the synchrony subspaces for the quotient network are 'mantained' for the lift network. In this example, the synchrony subspace $\{x : x_1 = x_2 = x_3\}$ for the three-gene GRN of Figure 2 lifts to the synchrony subspace $\{x : x_1 = x_2 = x_3 = x_4\}$ for the four-gene GRN of Figure 3.

If we consider gene conservation for the evolution of duplicate genes, that is, if we consider entire gene redundancy, then $\alpha_1 = \alpha_2 = \alpha$. In this case the lift four-gene GRN also has $\{x : x_3 = x_4\}$ as a synchrony subspace but $\{x : x_1 = x_2 = x_3 = x_4\}$ is not a synchrony subspace. Moreover, the quotient network in the restriction to the synchrony space $\{x : x_3 = x_4\}$ is no longer the three-gene GRN of Figure 2.

In this work we aim to connect and adapt some of the existent results concerning network robust synchronization to GRNs. We address mainly two issues. One concerns the existence of synchrony spaces in GRNs and their description; the other is related to robustness of such synchrony patterns. That is, does the answer to the first issue depends on the model equations approach? We address the robustness question to SUM and MULT models (cf. (2.8)-(2.9). From our results we conclude that, in general, the synchronization patterns can be quite distinct from the SUM and MULT

model. Moreover, we see that if the activation and repression functions involved at the GRNs models imply structural relation, unexpected synchrony patterns may occur that are arising due to the specificities of the activation and repression functions. We also address the lifting and quotient processes in each of the two models, considering subfunctionalization as the outcome of functional divergence for gene duplication.

The paper is organized in the following way. In Section 2 we provide some basics of gene regulatory networks that we follow in this paper. We present two dynamical models, the SUM and MULT models, which differ in the way the GRN regulatory function operates over the gene inputs. In Section 3, we introduce the concept of synchrony pattern (synchrony subspace) for the two types of GRN models equations and characterize the synchrony patterns for GRNs. See Propositions 3.4-3.6 and Propositions 3.12-3.14. When the regulatory functions are Hill-like (cf. (2.6)-(2.7)), a corollary of these results is that the GRN synchrony patterns are completely determined by the GRN structural matrices. In Section 4, we combine Theorem 2.13 of Aguiar and Dias [3] with Propositions 3.4-3.6 and Propositions 3.12-3.14, obtaining a characterization method of the n-gene GRNs that are lifts of a fixed k-gene GRN, for the SUM using the k-gene GRN activation and repression weighed matrices and for MULT models, using the k-gene GRN activation and repression weighed matrices, and the multiplicities matrices. See Theorems 4.1-4.2. In Section 5, we remark that for models where the activation and repression regulatory functions are structurally related, then there may be other synchrony patterns forced by those regulator functions, which we call regulatory dependent synchrony patterns. We give an example of such regulatory-dependent synchrony patterns in the context of circadian rhythms clock models in Section 2.3. Finally, in Section 6, we summarize the work presented and point out directions for future work.

2. Preliminaries

2.1. Gene Regulatory Networks (GRNs)

Gene regulatory networks can be modelled by a large variety of mathematical formalisms. Here we consider the differential equation model, where the variables describe the concentrations of gene products such as mRNAs and proteins as continuous values, and their time derivative is expressed as a function of the variables themselves (cf. [16]).

The activity of a gene is regulated by other genes through the concentrations of their gene products, which function as transcription factors. Regulation can be quantified by the "response

characteristics" given by the level of gene expression as a function of transcription factors (cf. [41]). Formally, a GRN can be described by a set of ordinary differential equations, for $i = 1, 2, \ldots, n$, of the form

$$\begin{cases} \dot{m}_i(t) = -a_i m_i(t) + b_i(p_1(t), p_2(t), \ldots, p_n(t)) \\ \dot{p}_i(t) = -c_i p_i(t) + d_i m_i(t) \end{cases}, \tag{2.3}$$

where $m_i, p_i \in \mathbb{R}$ are the concentrations of mRNA and protein, respectively, of the i-th node, $a_i, c_i > 0$ are degradation rates of mRNA and protein, respectively, and $d_i > 0$ is a constant.

Graphically, the set of equations (2.3), for $i = 1, 2, \ldots, n$, translates a GRN as a network of n nodes, where each node $x_i := (m_i, p_i)$ is connected with another node if there is an interaction between them via the regulatory function b_i.

2.2. Models of GRNs and Regulatory Functions

The regulatory function b_i in (2.3) plays a key role in the dynamical modelling of GRNs. It is generally a nonlinear function of its variables $p_1(t), p_2(t), \ldots, p_n(t)$ and may assume monoticity in each variable in simplified cases (cf. [55], more references from [41]). Depending on all biochemical reactions involved, it may have a complicated form which is, in practice, determined heuristically.

There are typically two regulatory logics behind b_i, depending on whether each transcription factor acts *additively* or *multiplicatively* to regulate the i-th gene. It is called *SUM logic*, if

$$b_i = \sum_j b_{ij}(p_j(t))$$

and it is called *MULT* logic, if

$$b_i = \Pi_j b_{ij}(p_j(t)).$$

See for example Chesi [16], Chesi and Hung [17] and [38, 48, 49].

There are two types of regulatory functions: those that *activate* and those that *repress* the target gene expression in the i-th node, described by a monotonically *increasing* and *decreasing* function b_{ij}, respectively.

One of the most commonly used type of activation regulatory function is described by (cf. [21, 16] for example) the *Hill function*

$$\mathrm{act}(p) = \frac{p^n}{\beta^n + p^n}, \tag{2.4}$$

where $n \in \mathbf{N}$ characterizes the steepness of regulation and β marks the mid-value of maximal reachable value of act. See Remark 2.4 for details. The repression is frequently modelled by

$$\mathrm{rep}(p) = 1 - \mathrm{act}(p) = \frac{\beta^n}{\beta^n + p^n}\,. \tag{2.5}$$

For our purpose, to allow more general modelling possibilities, we will only assume that act and rep are strictly monotonic between 0 and 1, satisfy

$$\mathrm{act}(p) \to 0 \ \text{ and } \ \mathrm{rep}(p) \to 1 \ \text{ when } \ p \to 0 \tag{2.6}$$

$$\mathrm{act}(p) \to 1 \ \text{ and } \ \mathrm{rep}(p) \to 0 \ \text{ when } \ p \to \infty, \tag{2.7}$$

and call them the *Hill-like* regulatory functions.

Consider a GRN of n nodes (genes), $x_i = (m_i, p_i) \in \mathbb{R}^+_0$ for $i = 1, \ldots, n$, where the concentration m_i of mRNA and p_i of protein are measured for each node. Take the internal dynamics function of the ith node as the 2×2 matrix

$$A_i = \begin{pmatrix} -a_i & 0 \\ d_i & -c_i \end{pmatrix},$$

where a_i, c_i, d_i are positive real constants. That is, we assume that the internal dynamics of the genes is linear. In particular, it follows that each A_i is invertible. Given i, define

$$I_i^+ = \{j \in \{1, \ldots, n\} : \ \text{gene } j \text{ activates gene } i\},$$
$$I_i^- = \{j \in \{1, \ldots, n\} : \ \text{gene } j \text{ represses gene } i\}\,.$$

Remark 2.1. *We allow $I_i^+ \cap I_i^- \neq \emptyset$. That is, a gene can activate and repress a same gene.* $\diamond$

Denote the two $n \times n$ network adjacency matrices by $W^+ = [w_{ij}^+]$ and $W^- = [w_{ij}^-]$. If there is activation (resp. repression) from gene j to gene i, we have that $w_{ij}^+ > 0$ (resp. $w_{ij}^- > 0$) stands for the maximal achievable level of the activation (resp. repression) from gene j to gene i. Otherwise, $w_{ij}^+ = 0$ (resp. $w_{ij}^- = 0$). We call W^+ (resp. W^-) the network *weighted activation* (resp. *repression) adjacency matrix*.

The SUM model

In the SUM model, the regulatory function is operating by *addition*:

$$\dot{x}_i = A_i x_i + \sum_{j \in I_i^-} w_{ij}^- \mathrm{rep}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_{j \in I_i^+} w_{ij}^+ \mathrm{act}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (i = 1, \ldots, n)\,. \tag{2.8}$$

The MULT model

In the MULT model, the regulatory function is operating by *multiplication*:

$$\dot{x}_i = A_i x_i + \prod_{j\in I_i^-} w_{ij}^- \mathrm{rep}^{m_{ij}^-}(p_j) \prod_{j\in I_i^+} w_{ij}^+ \mathrm{act}^{m_{ij}^+}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (i = 1, \ldots, n)\,, \tag{2.9}$$

where $m_{i,j}^{\pm}$ indicates the multiplicative regulation of the gene i from the gene j. Graphically, these multiplicities are usually represented by multiple arrows. However, in the case of $0-1$ adjacency matrices we frequently indicate the multiplicity of the arrows with a number next to each arrow. In the general case, if we have both weights $w_{ij}^{\pm}$ and exponents $m_{ij}^{\pm}$ in the model, then we reserve the number next to the arrow for $w_{ij}^{\pm}$ and use multiple arrows to indicate $m_{ij}^{\pm}$. Let $M^{\pm} = (m_{ij}^{\pm})$ be the *activation and repression multiplicity matrices* with the assumption that each entry of W^+ (resp. W^-) is nonzero if and only if each entry of M^+ (resp. M^-) is nonzero. The weights $w_{ij}^{\pm}$, on the other hand, can be combined into one weight:

$$w_i = \prod_{j\in I_i^-} w_{ij}^- \prod_{j\in I_i^+} w_{ij}^+ \qquad (i = 1, \ldots, n)\ .$$

Thus, (2.9) is effectively

$$\dot{x}_i = A_i x_i + w_i \prod_{j\in I_i^-} \mathrm{rep}^{m_{ij}^-}(p_j) \prod_{j\in I_i^+} \mathrm{act}^{m_{ij}^+}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (i = 1, \ldots, n)\ . \tag{2.10}$$

Example 2.2. *Consider the three-cell GRN which appears in Figure 4 and have repression and activation adjacency and multiplicities matrices given by*

$$W^- = \begin{pmatrix} 0 & 0 & w_{13} \\ 0 & 0 & 3 \\ 2 & 0 & 0 \end{pmatrix},\ W^+ = \begin{pmatrix} 0 & w_{12} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ M^- = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},\ M^+ = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where $w_{12}w_{13} = 4$*, respectively. Note that gene* 1 *represses gene* 3*, gene* 2 *activates gene* 1*, and gene* 3 *represses genes* $1, 2$*. Moreover, all activations and repressions have multiplicity one. The MULT model equations in Example 2 of [16] correspond to this three-cell GRN where the activation function is the Hill function* $\mathrm{act}(p) = p^2/(1+p^2)$*, the repression function is* $\mathrm{rep}(p) = 1/(1+p^2)$

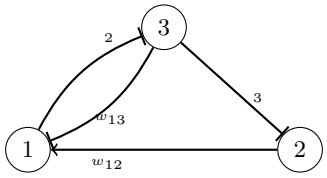

Figure 4: A GRN with three genes.

and the internal gene dynamics is linear and determined by certain 2×2 *matrices* A_1, A_2, A_3*:*

$$\left\{\begin{array}{l} \dot{x}_1 = A_1 x_1 + w_{12}\mathrm{act}(p_2) w_{13}\mathrm{rep}(p_3) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_2 = A_2 x_2 + 3\mathrm{rep}(p_3) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_3 x_3 + 2\mathrm{rep}(p_1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} \dot{x}_1 = A_1 x_1 + \frac{4p_2^2}{(1+p_2^2)(1+p_3^2)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_2 = A_2 x_2 + \frac{3}{1+p_3^2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_3 x_3 + \frac{2}{1+p_1^2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{array}\right. . \quad (2.11)$$

◇

Remark 2.3. *We remark that in some literature, for example [16], PROD models are used, which correspond to the particular case of the MULT models where* $m_{ij}^{\pm} = 1$*, for all* i, j*. That is, the PROD model equations are given by:*

$$\dot{x}_i = A_i x_i + \prod_{j \in I_i^-} w_{ij}^- \mathrm{rep}(p_j) \prod_{j \in I_i^+} w_{ij}^+ \mathrm{act}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (i = 1, \ldots, n) \ . \quad (2.12)$$

The PROD model equations in Example 2 of [16] correspond to the MULT model equations that appear in Example 2.2 above. ◇

The SUM and MULT differential equation models considered here are simplistic models for gene networks. These artificial gene networks models mimic as much as possible the characteristics of real gene networks. The study of how topological and kinetic properties of artificial gene networks may influence their behavior seems to be a good strategy to discover the underlying principles of biochemical regulation at the level of transcription, Mendes *et al.* [45] . See Zhang and Li [59] for the use of a SUM model for the gene regulatory network of the Hematopoietic stem cells differentiation and Elahi and Hasan [20] for the use of a MULT model for a network of interconnected nine genes SOS response in *Escherichia coli.*

2.3. An example of mammalian circadian rhythm

Following a statistical mechanical framework proposed in [13, 14], a concise GRN model was proposed in [38, 49], which consists of only 5 gene variables *Bmal1, Rev-erb-α, Per2, Cry1, Dbp* in studying the complex gene regulatory dynamics of the mammalian circadian oscillator. It was shown to describe the known phase relations, amplitudes and wave forms of clock gene expression profiles. Based on the same 5-gene model, different feedback loops (as sub-networks of the 5-gene network) have been identified to fit circadian gene expression profiles for different mammalian tissues ([48]).

The regulator functions are modelled by (cf. [38], Supplement)

$$\mathrm{rep}(x,b) = \frac{1}{1+bx} \tag{2.13}$$

$$\mathrm{act}(x,a) = \frac{1+ax}{1+x}, \tag{2.14}$$

where $a \geq 1$ is a parameter for fold activation and $b \geq 0$ is a repression parameter. We compare them with the classic Hill regulatory functions in regulation modelling parameters.

Remark 2.4. *Many gene regulation functions have been proposed heuristically. Besides the monotonicity with finite asymptotic values, the key features are usually characterized by steepness of regulation, basal transcription or transcription initiation (detected in the absence of transcription factors), the required transcription factor level and the maximal achievable level.*

We explain using activators as an example. In case of the regulation given by the Hill functions (cf. (2.4)-(2.5)), the transcription fold change is described by

$$H(x) = \frac{ax^n}{b^n + x^n},$$

where $a > 1$ is the maximal achievable transcription factor level, b is where the mid-value $a/2$ is attained and $n \in \mathbf{N}$ signifies the steepness of the curve measured by the slope at the mid-point

$$H'(b) = \frac{a}{4b}n. \tag{2.15}$$

In case of the activating function used for circadian clock models, the transcription fold change is described by

$$C(x) = \left(\frac{1 + a\cdot\frac{x}{b}}{1+\frac{x}{b}}\right)^n = \left(\frac{b+ax}{b+x}\right)^n,$$

where x/b is the normalized concentration, $a^n > 1$ is the maximal achievable transcription factor level, b is where the "mid-value" $\big((a+1)/2\big)^n$ is attained and $n \in \mathbf{N}$ signifies the steepness of the

curve measured by the slope at the mid-point

$$C'(b) = \frac{a-1}{4b}\left(\frac{a+1}{2}\right)^{n-1} n\,. \tag{2.16}$$

See Figure 5.

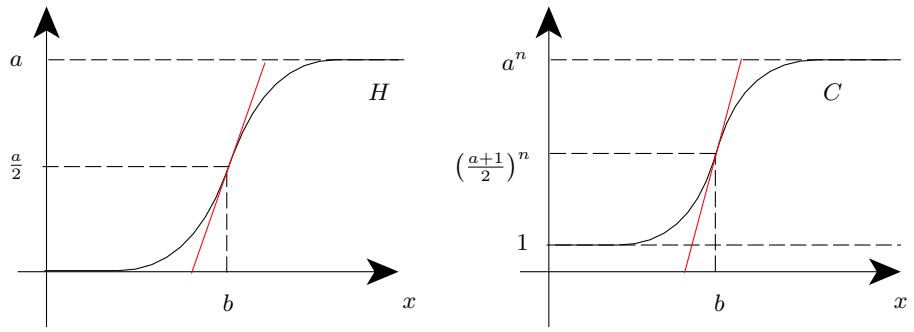

Figure 5: Two gene regulation functions for activating transcription factors. The Hill function $H(x) = (ax^n)/(b^n+x^n)$ (left) and the activating function $C(x) = \big((b+ax)/(b+x)\big)^n$ used in circadian rhythm model (right). The steepness is measured at the mid-point (cf. (2.15)-(2.16)).

These two functions give very similar qualitative descriptions of gene regulation activities including monotonicity, maximal achievable levels given by $a > 1$ and the steepness factor n.

One crucial difference, however, is the basal transcription given by the asymptotic value at $x = 0$, 235 *which corresponds to how much transcription can be detected in the absence of any activators. The Hill function H assumes no transcription activities if no activators are present, the function C models with a non-zero constant basal transcription, which in case of circadian models, is modulated by circadian transcription factors and input functions (cf. [38, 48, 48]).* ◇

Example 2.5. The circadian core 5-gene clock model is composed of genes *Bmal1, Rev-erb-α, Per2, Cry1,Dbp*, which interact through transcriptional feedback loops of negative and positive regulations. It describes gene regulatory dynamics of the mammalian circadian oscillator (cf. [38, 49]). See Figure 6. The equations employed to describe the dynamics are (cf. [38])

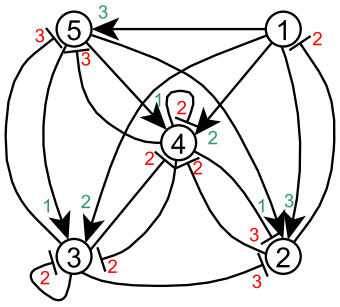

Figure 6: The circadian core 5-gene clock model, where cells $1, 2, 3, 4, 5$ correspond to the concentration of genes *Bmal1, Rev-erb-α, Per2, Cry1,Dbp*, respectively. The number next to the input arrow indicates the number of binding sites in the target gene (the red for negative, the green for positive feedback).

$$\begin{cases} \dot{x}_1 &= \mathrm{rep}^{\boxed{2}}(x_2, a_{12}) - d_1 x_1 \\ \dot{x}_2 &= \mathrm{act}^{\boxed{3}}(x_1, a_{21})\mathrm{act}(x_5, a_{25})\mathrm{rep}^{\boxed{3}}(x_3, 1)\mathrm{rep}^{\boxed{3}}(x_4, 1) - d_2 x_2 \\ \dot{x}_3 &= \mathrm{act}^{\boxed{2}}(x_1, a_{31})\mathrm{act}(x_5, a_{35})\mathrm{rep}^{\boxed{2}}(x_3, 1)\mathrm{rep}^{\boxed{2}}(x_4, 1) - d_3 x_3 \\ \dot{x}_4 &= \mathrm{act}^{\boxed{2}}(x_1, a_{41})\mathrm{act}(x_5, a_{45})\mathrm{rep}^{\boxed{2}}(x_2, 1)\mathrm{rep}^{\boxed{2}}(x_3, 1)\mathrm{rep}^{2}(x_4, 1) - d_4 x_4 \\ \dot{x}_5 &= \mathrm{act}^{\boxed{3}}(x_1, a_{51})\mathrm{rep}^{\boxed{3}}(x_3, 1)\mathrm{rep}^{\boxed{3}}(x_4, 1) - d_5 x_5, \end{cases}$$

where rep, act are given by (2.13)-(2.14) and the exponents are given by the number of binding sites available in the target gene. The boxed numbers for indicate that these genes x_j for $j \in I_i$ share the same number of binding sites available in the target gene x_i. See the numbers in Figure 6. ◇

Remark 2.6. *As mentioned above, in [38], the authors used (2.17) to model circadian clocks, where the boxed numbers are kept equal for genes sharing the same number of binding sites (that is available sites are always occupied) in the target gene. In [49] as an extension of work from [38], the authors used the same model to run through all oscillating networks and found statistical significance of a sub-network of* $\mathbf{Z}_3$*-symmetry composed of Rev-erb-α, Per2, Cry1 (corresponding to* x_2, x_3, x_4*) that works as a repressilator.* ◇

3. Synchrony in gene regulatory networks

Motivated by the formalisms of Stewart, Golubitsky and Pivato [53], Golubitsky, Stewart and Török [29] and Field [23] on coupled cell networks and robust network synchronization for the associated coupled cell systems, we define synchronization partitions for GRN model equations.

Definition 3.1. The *gene equivalence partition* of a GRN is the partition of the genes of the GRN where each part is formed by all the genes with the same internal dynamics function. Thus if $A_{i_1} = A_{i_2}$ (in (2.8) or (2.10)), we have that i_1, i_2 belong to the same part of the GRN gene equivalence partition. $\diamond$

Consider a GRN and an associated dynamical equations model as, for example, the SUM or MULT equations models, (2.8) and (2.10), respectively.

Definition 3.2. Given a partition P of the gene set of the GRN, define the space Δ_P to be the polydiagonal where gene products concentrations corresponding to the same class of genes in P are identified. We say that P is a *synchronization partition* of the GRN gene set, for the particular model considered, when Δ_P is flow-invariant under the equations of that model, for any given Hill-like regulator functions (cf. (2.6)-(2.7)). In that case, we call Δ_P a *synchrony pattern* or a *synchrony space* of the GRN. $\diamond$

In what follows, given a gene partition P of a GRN, we denote by $[i]$ the part of P that contains gene i. Moreover, we assume that a necessary condition for genes to synchronize in a robust way is that the genes have the same internal dynamics. That is, a synchronization partition P has to refine the gene equivalence partition.

Definition 3.3. Consider a GRN and an associated dynamical equations model. Given a synchronization partition P of the gene set of the GRN for the fixed model, if the restriction of the model equations to the synchrony subspace Δ_P are equations of the same model type then they are associated to a *quotient* GRN. The genes of the quotient GRN correspond to the parts in the synchronization partition P; the activation (repression) interactions between two genes in the quotient are the projection of the activation (repression) interactions between the genes in the corresponding parts of P in the original GRN. We also say that the original GRN is a *lift* of the quotient GRN. $\diamond$

We address now the issue of synchronization for gene regulatory networks, assuming the models equations are the SUM or the MULT models. More precisely, we characterize the synchrony patterns

for GRNs, for the SUM and MULT models, using the activation and repression adjacency and multiplicity matrices. From that, it follows that for the same GRN, distinct patterns of synchrony can occur for the two models, which in particular, lead to distinct dynamical properties for the corresponding dynamical systems. As already remarked, the synchrony subspaces that are forced by the symmetries of a symmetric GRN graph, the fixed-point subspaces, occur for both SUM and MULT GRN models considered here, as the associated GRNs equations for both models inherit the symmetries of the GRN. But, there can be synchrony subspaces that are not forced by the symmetries of the graph, if any. See, Antoneli and Stewart [9].

3.1. Synchrony for the SUM model

Proposition 3.4. *Take an n-gene GRN with adjacency matrices W^+ and W^- and a partition P of the gene set into classes $C_1, \dots, C_m$ refining the gene equivalence class. The partition P corresponds to a synchrony pattern Δ_P for the SUM equations model (2.8) if and only if for each part C, we have that for $k = 1, \dots, m$,*

$$\sum_{j \in I_i^+ \cap C_k} w_{ij}^+ \text{ is constant for } i \in C \tag{3.17}$$

and

$$\sum_{j \in I_i^- \cap C_k} w_{ij}^- \text{ is constant for } i \in C\,. \tag{3.18}$$

Proof. Assume that (3.17)-(3.18) hold for $k = 1, \dots, m$ on any equivalence class C under partition P. Then, we have for any $i \in C$ that

$$\begin{aligned}
\dot{x}_i &= Ax_i + \sum_{j \in I_i^-} w_{ij}^- \mathrm{rep}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_{j \in I_i^+} w_{ij}^+ \mathrm{act}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
&= Ax_i + \sum_{k=1}^{m} \underbrace{\sum_{j \in I_i^- \cap C_k} w_{ij}^-}_{w_{i,k}^-} \mathrm{rep}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_{k=1}^{m} \underbrace{\sum_{j \in I_i^+ \cap C_k} w_{ij}^+}_{w_{i,k}^+} \mathrm{act}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix},
\end{aligned}$$

where the sums $w_{i,k}^{\pm}$ by (3.17)-(3.18), remain the same in the equivalence class of i. Thus, for any $i \in C$, the dynamics of $x_{[i]} \in \Delta_P$ is governed by

$$\dot{x}_{[i]} = Ax_{[i]} + \sum_k w_{[i],k}^- \mathrm{rep}(p_{[j]}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_k w_{[i],k}^+ \mathrm{act}(p_{[j]}) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3.19}$$

where $[j] := C_k$ if $j \in C_k$ and k is such that $I_i^{\pm} \cap C_k \neq \emptyset$. Notice that by (3.17)-(3.18), $I_i^{\pm} \cap C_k \neq \emptyset$ if and only if $I_l^{\pm} \cap C_k \neq \emptyset$ for any $l \in [i]$. It follows that Δ_P is flow-invariant and consequently, a synchrony space for the SUM equations model (2.8).

Assume that Δ_P is a synchrony space for (2.8). Then, Δ_P is flow-invariant. In particular, we have $I_i^{\pm} \cap C_k \neq \emptyset$ if and only if $I_l^{\pm} \cap C_k \neq \emptyset$ for any i, l in the same part of P. Moreover, the flow restricted to Δ_P is of form (3.19) and for i, l in the same part of P, we have

$$\sum_{k=1}^{m} w_{i,k}^{-}\mathrm{rep}(p_{[j]}) + \sum_{k=1}^{m} w_{i,k}^{+}\mathrm{act}(p_{[j]}) = \sum_{k=1}^{m} w_{l,k}^{-}\mathrm{rep}(p_{[j]}) + \sum_{k=1}^{m} w_{l,k}^{+}\mathrm{act}(p_{[j]}).$$

For $k \in \{1, \ldots, m\}$, let $p_{[j]} = 0$ for all $[j] \neq C_k$. For $[j] = C_k$, if we let $p_{[j]} \to 0$, then $\mathrm{rep}(p_{[j]}) \to 1$ and $\mathrm{act}(p_{[j]}) \to 0$ hold, which implies that $w_{i,k}^{-} = w_{l,k}^{-}$ (cf. (2.6)). Similarly, if we let $p_{[j]} \to \infty$ for $[j] = C_k$, then $w_{i,k}^{+} = w_{l,k}^{+}$ (cf. (2.7)). □

Remark 3.5. It follows from Proposition 3.4 that, given a partition P of the gene set of a GRN, the associated polydiagonal susbspace Δ_P is a synchrony subspace for the SUM model of the GRN if and only if it is left invariant by the weighted adjacency matrices W^+ and W^- of the GRN. We have then, using the work of Aguiar and Dias in [2] and [3], that the set of the synchrony subspaces (synchrony patterns) for the SUM model of a GRN can be computed using the algorithm in Section 6 of [2]. The algorithm can be executed to find the set of polydiagonal subspaces that are left invariant by one of the adjacency matrices W^+ or W^- and then the set of synchrony subspaces of the GRN is the subset of those polydiagonals that are also left invariant by the other adjacency matrix. ◇

Proposition 3.6. *Take an n-gene GRN with adjacency matrices W^+ and W^-. Let P be a synchronization partition of the gene set with classes $C_1, \ldots, C_m$ refining the gene equivalence partition. Consider the $m \times m$ matrices $Q^+ = \left[q_{ik}^+\right]$ and $Q^- = \left[q_{ik}^-\right]$ where*

$$\sum_{j \in I_i^+ \cap C_k} w_{ij}^+ = q_{ik}^+ \quad \text{and} \quad \sum_{j \in I_i^- \cap C_k} w_{ij}^- = q_{ik}^- \quad (i, k = 1, \ldots, m) \ .$$

The restriction of SUM equations model (2.8) to Δ_P is a SUM equations model consistent with the quotient GRN of m genes with activation and repression weighted adjacency matrices Q^+ and Q^-, respectively.

Proof. Follows trivially from the conditions in Proposition 3.4 and its proof. □

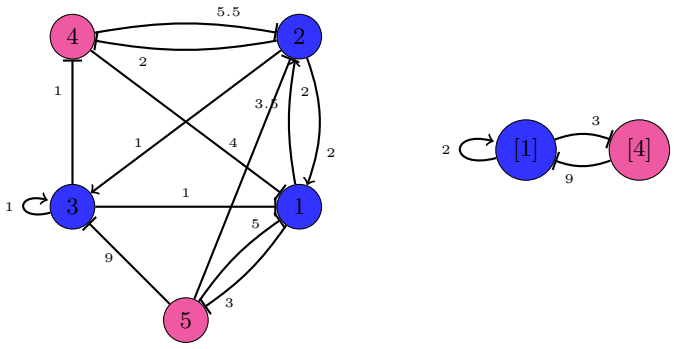

Figure 7: The GRNs of Example 3.7: (left) a five gene GRN and (right) its quotient network corresponding to the partition $P = \{[1] = \{1, 2, 3\}, [4] = \{4, 5\}\}$.

Example 3.7. *Take the five gene GRN on the left of Figure 7 with the following* 5×5 *activation and repression weighted adjacency matrices:*

$$W^+ = \left(\begin{array}{ccc|cc} 0 & 2 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right), \quad W^- = \left(\begin{array}{ccc|cc} 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 5.5 & 3.5 \\ 0 & 0 & 0 & 0 & 9 \\ \hline 0 & 2 & 1 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \end{array}\right).$$

By Proposition 3.4, the partition $P = \{[1] = \{1, 2, 3\}, [4] = \{4, 5\}\}$ *is a synchronization partition and corresponds to the synchrony space* $\Delta_P = \{x : x_1 = x_2 = x_3, x_4 = x_5\}$ *for the SUM model equations if* P *refines the gene equivalence partition. By Proposition 3.6, the corresponding quotient network is the two-gene GRN on the right of Figure 7 with the following activation and repression weighted adjacency matrices:*

$$Q^+ = \left(\begin{array}{c|c} 2 & 0 \\ \hline 0 & 0 \end{array}\right), \quad Q^- = \left(\begin{array}{c|c} 0 & 9 \\ \hline 3 & 0 \end{array}\right). \tag{3.20}$$

Note that Δ_P *is not a synchrony space for the MULT model equations.* ◇

Remark 3.8. *For any five gene GRN with* 5×5 *activation and repression weighted adjacency*

matrices given by

$$W^+ = \left(\begin{array}{ccc|cc} 0 & 2 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right), \quad W^- = \left(\begin{array}{ccc|cc} 0 & 0 & 0 & w_{14}^- & w_{15}^- \\ 0 & 0 & 0 & w_{24}^- & w_{25}^- \\ 0 & 0 & 0 & 0 & 9 \\ \hline 0 & 2 & 1 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \end{array}\right)$$

where the weights satisfy $w_{14}^- + w_{15}^- = w_{24}^- + w_{25}^- = 9$*, if the gene set partition* P *with classes* $\{1,2,3\}$, $\{4,5\}$ *refines the gene equivalence partition, then* P *is a synchronization partition corresponding to the synchrony space* $\Delta_P = \{x : x_1 = x_2 = x_3, x_4 = x_5\}$ *for the SUM model equations.* *The corresponding quotient network of two-gene GRN has activation and repression weighted adjacency matrices given in (3.20). This follows from Propositions 3.4 and 3.6. This example illustrates that there is an infinite number of five gene GRNs admitting the synchronization partition* P *and leading to the same two-gene quotient GRN for the SUM models.* ◇

Remark 3.9. Under the assumptions of Proposition 3.4 and recall equation (3.19), if $w_{i,k}^+ = w_{i,k}^-$, then as $\mathrm{rep}(p_k) = 1 - \mathrm{act}(p_k)$, the influence of the k-th cluster C_k to the evolution of the i-th gene reduces to $w_{i,k}^+\mathrm{act}(p_k) + w_{i,k}^-\mathrm{rep}(p_k) = w_{i,k}^- = w_{i,k}^+$, which becomes independent on the variable p_k. See Example 3.11. For general weighted networks, Aguiar, Dias and Ferreira [4] point out a similar phenomenon and in that case Δ_P is called a *spurious* synchrony pattern. See Definition 2.9 of [4].

◇

Remark 3.10. *There are GRNs such that* $I_i^+ \cap I_i^- = \emptyset$*, for all genes* i*, and admitting a synchronization partition such that, in the quotient, the same gene activates and represses another gene, as the following example illustrates.* ◇

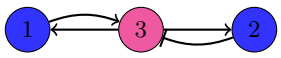

Figure 8: A GRN with three genes. Assuming genes 1 and 2 are equivalent then $P = \{C_1 = \{1,2\}, C_2 = \{3\}\}$ is a synchronization partition for the GRN SUM equations models with the two-gene quotient GRN pictured in Figue 9.

Example 3.11. *Consider the three-gene GRN in Figure 8 with* 3×3 *activation and repression*

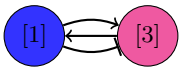

Figure 9: GRN with two genes where gene [1] activates and supresses gene [3] with the same strenght.

weighted adjacency matrices given by

$$W^+ = \left(\begin{array}{cc|c} 0 & 0 & 1 \\ 0 & 0 & 1 \\ \hline 1 & 0 & 0 \end{array}\right), \quad W^- = \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 1 & 0 \end{array}\right).$$

By Proposition 3.4, assuming genes 1 *and* 2 *are equivalent, we have that* $P = \{C_1 = \{1,2\}, C_2 = \{3\}\}$ *is a synchronization partition for the SUM equations models. Now, the two-gene quotient GRN as defined in Proposition 3.6, with* 2×2 *activation and repression weighted adjacency matrices given by*

$$Q^+ = \left(\begin{array}{c|c} 0 & 1 \\ \hline 1 & 0 \end{array}\right), \quad Q^- = \left(\begin{array}{c|c} 0 & 0 \\ \hline 1 & 0 \end{array}\right),$$

and satisfies $I^+_{[3]} \cap I^-_{[3]} = \{[1]\} \neq \emptyset$*. That is, in the quotient, gene* [1] *activates and represses gene* [3]*, see Figure 9. This follows from the fact that* $I^+_3 = \{1\}$*,* $I^-_3 = \{2\}$ *and* $1, 2$ *form the part* C_1 *of the synchronization partition* P*. Equations for the SUM model*

$$\begin{cases} \dot{x}_1 = A_1 x_1 + \mathrm{act}(p_3) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_2 = A_1 x_2 + \mathrm{act}(p_3) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_3 x_3 + [\mathrm{act}(p_1) + \mathrm{rep}(p_2)] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{cases}$$

restricted to $\Delta_P = \{x : x_1 = x_2\}$ *are given by:*

$$\begin{cases} \dot{x}_1 = A_1 x_1 + \mathrm{act}(p_3) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_3 x_3 + [\mathrm{act}(p_1) + \mathrm{rep}(p_1)] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{cases}.$$

With the assumption $\mathrm{act}(p_1) + \mathrm{rep}(p_1) \equiv 1$*, these equations simplify to*

$$\begin{cases} \dot{x}_1 = A_1 x_1 + \mathrm{act}(p_3) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_3 x_3 + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{cases} .$$

Thus, the gene equation for [3] *becomes independent of gene* [1]*. This example also illustrates Remark 3.9.* ◇

3.2. Synchrony for the MULT model

Let $W^{\pm} = (w_{ij}^{\pm})$ and $M^{\pm} = (m_{ij}^{\pm})$ be, respectively, the weighted matrices and the multiplicity matrices of the MULT model (2.10).

Proposition 3.12. *Take an n-gene GRN with adjacency matrices* $W^{\pm}$ *and* $M^{\pm}$ *and a partition* P *of the gene set into classes* $C_1, \ldots, C_m$ *refining the gene equivalence class. The partition* P *corresponds to a synchrony pattern for the MULT equations model (2.10) if and only if for each part* C *of* P*, we have that:*
(i) for $k = 1, \ldots, m$*,*

$$\sum_{j \in I_i^- \cap C_k} m_{ij}^- \quad \textit{is constant for } i \in C$$

and

$$\sum_{j \in I_i^+ \cap C_k} m_{ij}^+ \quad \textit{is constant for } i \in C$$

(ii)

$$\prod_{j \in I_i^+} w_{ij}^+ \prod_{j \in I_i^-} w_{ij}^- \textit{ is constant for } i \in C. \tag{3.21}$$

Proof. Assume that (i)-(ii) hold on any equivalence class C under partition P. Denote by $\mathbf{m}_{i,k}^{\pm}$ the sum of the multiplicities $m_{ij}^{\pm}$ in $I_i^{\pm} \cap C_k$, that is, $\mathbf{m}_{i,k}^- = \sum_{j \in I_i^- \cap C_k} m_{ij}^-$ and $\mathbf{m}_{i,k}^+ = \sum_{j \in I_i^+ \cap C_k} m_{ij}^+$.

Then, we have for any $i \in C$ that

$$\dot{x}_i = Ax_i + \prod_{j\in I_i^-} w_{ij}^- \mathrm{rep}^{m_{ij}^-}(p_j) \prod_{j\in I_i^+} w_{ij}^+ \mathrm{act}^{m_{ij}^+}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$= Ax_i + \underbrace{\Big(\prod_{j\in I_i^-} w_{ij}^- \prod_{j\in I_i^+} w_{ij}^+\Big)}_{w_i} \prod_k \prod_{j\in I_i^-\cap C_k} \mathrm{rep}^{m_{ij}^-}(p_j) \prod_k \prod_{j\in I_i^+\cap C_k} \mathrm{act}^{m_{ij}^+}(p_j) \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

where the product w_i, by (ii), remains the same in the equivalence class of i. In fact, w_i is the product of the $w_{ij}^\pm$ where j runs through $I_i^\pm$. Moreover, by (i), the sums $\mathbf{m}_{i,k}^\pm$ are also constant in the equivalence class of i. Thus, for any $i \in C$, the dynamics of $x_{[i]} \in \Delta_P$ is governed by

$$\dot{x}_{[i]} = Ax_{[i]} + w_{[i]} \prod_k \mathrm{rep}^{\mathbf{m}_{[i],k}^-}(p_{[j]}) \prod_k \mathrm{act}^{\mathbf{m}_{[i],k}^+}(p_{[j]}) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3.22}$$

where $[j] := C_k$ for $j \in C_k$ and k is such that $I_i^\pm \cap C_k \neq \emptyset$. By (ii), $I_i^\pm \cap C_k \neq \emptyset$ if and only if $I_l^\pm \cap C_k \neq \emptyset$ for any $l \in [i]$. Thus, Δ_P is flow-invariant and a synchrony space for the MULT equations model (2.10).

Assume that Δ_P is a synchrony space for (2.10). Then, Δ_P is flow-invariant. In particular, $I_i^\pm \cap C_k \neq \emptyset$ if and only if $I_l^\pm \cap C_k \neq \emptyset$ for any i, l in the same part of P. Moreover, the flow restricted to Δ_P is of shape (3.22) and for any i, l in the same part of P, we have

$$w_i \prod_k \mathrm{rep}^{\mathbf{m}_{[i],k}^-}(p_{[j]}) \prod_k \mathrm{act}^{\mathbf{m}_{[i],k}^+}(p_{[j]}) = w_l \prod_k \mathrm{rep}^{\mathbf{m}_{[l],k}^-}(p_{[j]}) \prod_k \mathrm{act}^{\mathbf{m}_{[l],k}^+}(p_{[j]}). \tag{3.23}$$

Assume to the contrary of (i) that $\mathbf{m}_{i,k}^- \neq \mathbf{m}_{l,k}^-$ for some k. Then, by letting $p_{[j]} = 0$ for all $[j]$ except when $[j] = C_k$ in (3.23), we have $\mathrm{r}ep^M(p)\mathrm{a}ct^N(p) \equiv$ constant for $p = p_{[j]}$ and some positive integers M, N. But since $\mathrm{r}ep^M(p)\mathrm{a}ct^N(p) \to 0$ both when $p \to 0$ and $p \to \infty$ by (2.6)-(2.7), we must have $\mathrm{r}ep^M(p)\mathrm{a}ct^N(p) \equiv 0$, which contradicts to the fact that rep and act are non-zero functions. Therefore, $\mathbf{m}_{i,k}^\pm = \mathbf{m}_{l,k}^\pm$ and (i) follows. By (3.23), we conclude that $w_i = w_l$ and (ii) follows. □

Remark 3.13. Given a GRN with weighted adjacency matrices W^+ and W^-, consider the diagonal matrix W^* where the ii entry is given by the product of the non-zero entries of the i-th row of both matrices W^+ and W^-,or zero otherwise. From Proposition 3.12, given a partition P of the gene set of the GRN, the associated polydiagonal susbspace Δ_P is a synchrony subspace for the MULT model of the GRN if and only if it is left invariant by the multiplicity matrices M^+ and M^-

of the GRN and by the matrix W^*. We have then, using the work of Aguiar and Dias in [2] and [3], that the set of the synchrony patterns for the MULT model of a GRN can be computed using the algorithm in Section 6 of [2]. The algorithm can be executed to find the set of polydiagonal subspaces that are left invariant by one of the multiplicity matrices M^+ or M^- and then the set of synchrony subspaces of the GRN is the subset of those polydiagonals that are also left invariant by the other multiplicity matrix and by the diagonal matrix W^*. ◇

Proposition 3.14. *Take an n-gene GRN with multiplicity matrices $M^\pm$ and weight matrices $W^\pm$ and a synchronized partition P of the gene set into classes $C_1, \dots, C_m$ refining the gene equivalence class for MULT equations model (2.9). The restriction of (2.9) to the synchrony space Δ_P is a MULT equations model consistent with any quotient m-gene GRN with the $m \times m$ multiplicity matrices $N^+ = \left[n_{ik}^+\right]$ and $N^- = \left[n_{ik}^-\right]$ given by*

$$n_{ik}^+ = \sum_{j \in I_i^+ \cap C_k} m_{ij}^+, \quad n_{ik}^- = \sum_{j \in I_i^- \cap C_k} m_{ij}^-, \quad (i,k = 1, \dots, m)$$

and $m \times m$ activation and repression weighted adjacency matrices $Q^+ = \left[q_{ik}^+\right]$ and $Q^- = \left[q_{ik}^-\right]$ satisfying

$$\prod_{k=1}^{m} q_{ik}^+ \, q_{ik}^- = \prod_{k=1}^{m} \left(\prod_{j \in I_i^+ \cap C_k} w_{ij}^+ \right) \left(\prod_{j \in I_i^- \cap C_k} w_{ij}^- \right) \quad (i = 1, \dots, m) \ . \tag{3.24}$$

Here, each product $\prod_{j \in I_i^+ \cap C_k} w_{ij}^+$ is considered only for k such that $I_i^+ \cap C_k \neq \emptyset$. In that case, $\prod_{j \in I_i^+ \cap C_k} w_{ij}^+$ is positive. Similarly, each product $\prod_{j \in I_i^- \cap C_k} w_{ij}^-$ is taken for k such that $I_i^- \cap C_k \neq \emptyset$.

Proof. This follows from Proposition 3.12 and the definition of MULT equations model (2.9). □

Remark 3.15. *Note that, in general, there is no uniqueness on the quotient GRN associated with the MULT model equations restricted to a synchrony space. This lack of uniqueness is due to the dependence of the product conditions (3.24), defining the activation and repression weighted adjacency matrices. See Example 3.16 below. Thus, the choice of a particular quotient GRN can be made taking into account the specifics of the problem under analysis.* ◇

Example 3.16. *Take the 4-gene GRN in Figure 10 with the following 4×4 activation and repression*

weighted adjacency matrices,

$$W^+ = \left(\begin{array}{cc|c|c} 2 & 0.5 & 0 & 0 \\ 1 & 9 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right),\ W^- = \left(\begin{array}{cc|c|c} 0 & 0 & 3 & 3 \\ 0 & 0 & 1 & 1 \\ \hline 0 & 2 & 0 & 0 \\ \hline 3 & 0 & 0 & 0 \end{array}\right),$$

and multiplicity matrices,

$$M^+ = \left(\begin{array}{cc|c|c} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right),\ M^- = \left(\begin{array}{cc|c|c} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ \hline 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \end{array}\right).$$

Assume that genes $1, 2$ *have the same internal dynamics. Consider the gene set partition* $P = \{C_1 = \{1, 2\}, C_2 = \{3\}, C_3 = \{4\}\}$. *By Proposition 3.12, we have that* P *is a synchronization partition. Note that* $2 \times 0.5 \times 3 \times 3 = 1 \times 9 \times 1 \times 1 = 9$ *and so condition (3.21) is satisfied. The synchronization partition* P *corresponds to the synchrony space* $\Delta_P = \{x : x_1 = x_2\}$ *for the MULT model equations:*

$$\begin{cases} \dot{x}_1 = A_1 x_1 + [2\text{act}(p_1)\, 0.5\text{act}(p_2)\, 3\text{rep}(p_3)\, 3\text{rep}(p_4)] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_2 = A_1 x_2 + [\text{act}(p_1)\, 9\text{act}(p_2)\, \text{rep}(p_3)\, \text{rep}(p_4)] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_3 x_3 + 2\text{rep}(p_2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_4 = A_4 x_4 + 3\text{rep}(p_1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{cases} \tag{3.25}$$

The restriction of equations (3.25) to $\Delta_P = \{x : x_1 = x_2\}$ *is:*

$$\begin{cases} \dot{x}_1 = A_1 x_1 + \left[9\left(\text{act}(p_1)\right)^2 \text{rep}(p_3)\, \text{rep}(p_4)\right] \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_3 = A_3 x_3 + 2\text{rep}(p_1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \dot{x}_4 = A_4 x_4 + 3\text{rep}(p_1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{cases}. \tag{3.26}$$

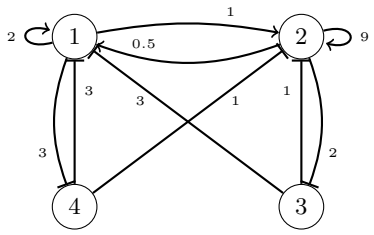

Figure 10: The four gene GRN of Example 3.16.

These equations are MULT model equations for any 3-gene GRN with the following 3×3 *multiplicity matrices,*

$$N^+ = \left(\begin{array}{c|c|c} 2 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right),\ N^- = \left(\begin{array}{c|c|c} 0 & 1 & 1 \\ \hline 1 & 0 & 0 \\ \hline 1 & 0 & 0 \end{array}\right)$$

and activation and repression weighted adjacency matrices,

$$Q^+ = \left(\begin{array}{c|c|c} q_{11}^+ & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right),\ Q^- = \left(\begin{array}{c|c|c} 0 & q_{12}^- & q_{13}^- \\ \hline 2 & 0 & 0 \\ \hline 3 & 0 & 0 \end{array}\right),$$

such that $q_{11}^+ q_{12}^- q_{13}^- = 9$. *Two particular choices of activation and repression weighted adjacency matrices are given by, respectively,*

$$Q^+ = \left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right),\ Q^- = \left(\begin{array}{c|c|c} 0 & 3 & 3 \\ \hline 2 & 0 & 0 \\ \hline 3 & 0 & 0 \end{array}\right).$$

and

$$Q^+ = \left(\begin{array}{c|c|c} 9 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right),\ Q^- = \left(\begin{array}{c|c|c} 0 & 1 & 1 \\ \hline 2 & 0 & 0 \\ \hline 3 & 0 & 0 \end{array}\right).$$

355 ◇

Note that in Example 3.16 above, the MULT model equations (3.25) fit the PROD model (recall Remark 2.3) but its restriction to the synchrony space $\Delta_P = \{x : x_1 = x_2\}$ providing equations (3.26) fit the MULT model but do not fit the PROD model. Thus the PROD model

shows some weaknesses in the context of synchrony spaces and their quotient networks, which are indispensable in describing robust synchrony patterns accommodated by GRN network structure. This disadvantage ceases to exist in the more general MULT model.

Example 3.17. *Motivated by the Case Study presented in Elahi and Hasan [20, Section 4], we consider now the network of nine genes SOS response in Escherichia coli and* 43 *interconnections in Figure 3.2 taken from [10, 25]. The nine genes named lexA, recA, recF, rpoD, rpoS, dinI, umuDC, rpoH and ssB are considered to be directly participating in the SOS response. As mentioned in [20], SOS response in Escherichia coli is an inducible DNA repair system that enables bacteria to survive under severe DNA damage. The associated MULT model considered in Elahi and Hasan [20] has the form*

$$\dot{x}_i = f_i(X, \Theta) \quad (i = 1, \ldots, 9)$$

where $X = [x_1, \ldots, x_9] = [lexA, recA, recF, rpoS, rpoD, umuD, dinI, ssB, rpoH]$ *denotes the state gene variables and* Θ *is the set of* 104 *unknown parameters which have to be estimated from the experimental genome data. See Section 4 of [20] for more details on the equations form.*

We remark that, depending on the model parameters, the gene network can have the following synchrony subspaces: $\{lexA = ssB\}$, $\{recA = dinI\}$, $\{umuDC = dinI\}$ *and* $\{rpoS = rpoH\}$. *Moreover, the intersection on any number of these synchrony subspaces is also a synchrony subspace for the gene network.* ◇

4. Network lifting enumeration and gene duplication

As stated in the introductory Section 1, network lifting, as defined in the theory of coupled cell networks, corresponds to gene duplication in GRNs followed by subfunctionalization as divergence process given that the sum of the weights of the edges from the duplicate genes that target a given gene corresponds to the weight of the edge from the ancestral gene targeting that gene. It was also remarked that, for general coupled cell networks and weighted networks, fixing a k-cell (quotient) network, if $n > k$ then there are many n-cell networks which are lifts (or inflations) of the k-cell network, that is, that admit the fixed network as a quotient. Moreover, there is a method of enumerating the lifts of a fixed k-cell network. See Theorem 2.5 of Aguiar, Dias, Golubitsky and Leite [5] valid for coupled cell systems following the formalisms of Stewart, Golubitsky and

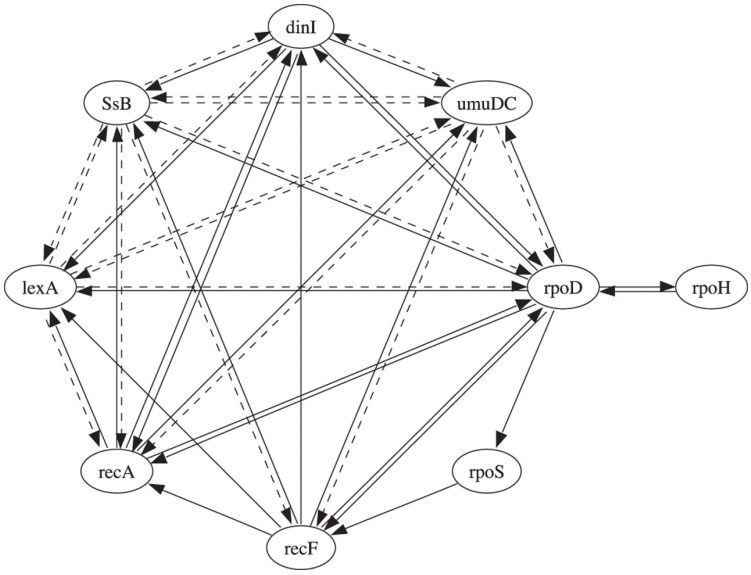

Figure 11: Figure taken from [10] of a GRN network between nine genes SOS response in *Escherichia coli* and 43 interconnections. Solid lines denote activating interactions, dotted lines denote inhibiting interactions and arrow heads show the direction of effect. Reproduced with permission from [10].

Pivato [53], Golubitsky, Stewart and Török [29] and Field [23], and its extension to weighted networks for coupled cell systems with additive input structure in Theorem 2.13 of Aguiar and Dias [3]. The enumeration method relies upon the characterization of the network adjacency matrices of the larger networks determined by the smaller network adjacency matrices. Also, fixing n, the number of n-cell lifts is finite for nonnegative integer matrices and nonfinite in the weighted setup.

In this section, we combine Theorem 2.13 of Aguiar and Dias [3] with Propositions 3.4, 3.6, and Propositions 3.12, 3.14, obtaining a characterization method of the n-gene GRNs that are lifts of a fixed k-gene GRN, for both the SUM and MULT models, using the activation and repression weighed matrices (and the multiplicities matrices for the MULT model quotients).

Theorem 4.1. *Consider Q an m-gene GRN for the SUM equations model with activation and repression weighted adjacency matrices $Q^+ = \left[q_{ik}^+\right]$ and $Q^- = \left[q_{ik}^-\right]$. An n-gene GRN for the SUM equations model with set of genes $\{1,\dots,n\}$, where $n > m$, is a lift of Q if and only if there is a partition of $\{1,\dots,n\}$ into m classes, $C_1,\dots,C_m$, refining the gene equivalence partition such that, after renumbering the genes if necessary, the activation and repression weighted adjacency matrices*

W^+ and W^- have the following block structures:

$$W^+ = \begin{pmatrix} Q^+_{11} & \cdots & Q^+_{1m} \\ \vdots & \cdots & \vdots \\ Q^+_{m1} & \cdots & Q^+_{mm} \end{pmatrix}, \qquad W^- = \begin{pmatrix} Q^-_{11} & \cdots & Q^-_{1m} \\ \vdots & \cdots & \vdots \\ Q^-_{m1} & \cdots & Q^-_{mm} \end{pmatrix} \tag{4.27}$$

where each $Q^\pm_{ik}$ is an $\#C_i \times \#C_k$-matrix with nonnegative real entries whose row sum is $q^\pm_{ik}$.

Proof. Direct application of Proposition 3.4 and Theorem 2.13 in [3]. □

Theorem 4.2. *Consider Q an m-gene GRN for the MULT equations model with $m \times m$ multiplicity matrices $N^+ = \left[n^+_{ik}\right]$, $N^- = \left[n^-_{ik}\right]$ and activation and repression weighted adjacency matrices $Q^+ = \left[q^+_{ik}\right]$ and $Q^- = \left[q^-_{ik}\right]$. An n-gene GRN for the MULT equations model with set of genes $\{1, \ldots, n\}$, where $n > m$, is a lift of Q if and only if there is a partition of $\{1, \ldots, n\}$ into m classes, $C_1, \ldots, C_m$, refining the gene equivalence partition such that, after renumbering the genes if necessary:*

(i) The $n \times n$ multiplicity matrices M^+ and M^- have the following block structures:

$$M^+ = \begin{pmatrix} N^+_{11} & \cdots & N^+_{1m} \\ \vdots & \cdots & \vdots \\ N^+_{m1} & \cdots & N^+_{mm} \end{pmatrix}, \qquad M^- = \begin{pmatrix} N^-_{11} & \cdots & N^-_{1m} \\ \vdots & \cdots & \vdots \\ N^-_{m1} & \cdots & N^-_{mm} \end{pmatrix} \tag{4.28}$$

where each $N^\pm_{ik}$ is an $\#C_i \times \#C_k$-matrix with nonnegative integer entries whose row sum is $n^\pm_{ik}$.

(ii) The activation and repression matrices W^+ and W^- have block structures

$$W^+ = \begin{pmatrix} Q^+_{11} & \cdots & Q^+_{1m} \\ \vdots & \cdots & \vdots \\ Q^+_{m1} & \cdots & Q^+_{mm} \end{pmatrix}, \qquad W^- = \begin{pmatrix} Q^-_{11} & \cdots & Q^-_{1m} \\ \vdots & \cdots & \vdots \\ Q^-_{m1} & \cdots & Q^-_{mm} \end{pmatrix}$$

where the matrices Q^+_{ij}, Q^-_{ij} have nonnegative real entries satisfying the following: for $i = 1, \ldots, m$, the product of the nonzero entries of each row of $\left(Q^+_{i1} \ldots Q^+_{im} Q^-_{i1} \ldots Q^-_{im}\right)$ equals the product of the nonzero entries of $\left(q^+_{i1} \ldots q^+_{im}\, q^-_{i1} \ldots q^-_{im}\right)$.

Proof. Direct application of Proposition 3.12 and Theorem 2.13 in [3]. □

Example 4.3. *Let Q be the 3-gene GRN for the MULT model equation (considered in Example 3.16) with the 3×3 multiplicity matrices:*

$$N^+ = \left(\begin{array}{c|c|c} 2 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right), \; N^- = \left(\begin{array}{c|c|c} 0 & 1 & 1 \\ \hline 1 & 0 & 0 \\ \hline 1 & 0 & 0 \end{array}\right)$$

and activation and repression weighted adjacency matrices

$$Q^+ = \left(\begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right), \; Q^- = \left(\begin{array}{c|c|c} 0 & 3 & 3 \\ \hline 2 & 0 & 0 \\ \hline 3 & 0 & 0 \end{array}\right).$$

Considering the 4-gene GRNs that are lifts of Q for the MULT model, we have that any 4-gene GRN with 4×4 multiplicity matrices

$$M^+ = \left(\begin{array}{cc|c|c} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right), \quad M^- = \left(\begin{array}{cc|c|c} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ \hline 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \end{array}\right)$$

and activation and repression weighted adjacency matrices

$$W^+ = \left(\begin{array}{cc|c|c} w_{11}^+ & w_{12}^+ & 0 & 0 \\ w_{21}^+ & w_{22}^+ & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right), \quad W^- = \left(\begin{array}{cc|c|c} 0 & 0 & w_{13}^- & w_{14}^- \\ 0 & 0 & w_{23}^- & w_{24}^- \\ \hline 0 & 2 & 0 & 0 \\ \hline 3 & 0 & 0 & 0 \end{array}\right),$$

where $w_{11}^+ w_{12}^+ w_{13}^- w_{14}^- = w_{21}^+ w_{22}^+ w_{23}^- w_{24}^- = 9$, is a lift of the 3-gene GRN. By Proposition 3.12, assuming that genes $1, 2$ have the same internal dynamics, any such lift has the synchronization partition P with the parts $\{1,2\}$, $\{3\}$, $\{4\}$, $\{5\}$ and the restriction of MULT model equations to the synchrony space $\Delta_P = \{x : x_1 = x_2\}$ gives rise to the MULT model equations (3.26). ◇

5. Regulatory-dependent synchrony spaces

Gene regulation functions provide an unique characteristic of GRNs relating concentrations of transcription factors such as activators or repressors to the promoter activities. They have

been investigated in two ways ([55]). Classical molecular biology explores mechanistic details of transcription and translation activities ([34, 50]), whereas the emerging field of system biology quantifies gene expressions on a larger scale in a statistical mechanical framework without requiring physical details of macromolecular interactions ([12, 52, 13, 14]).

So far, assuming that the regulator functions act and rep are Hill-like according to (2.6)-(2.7), we have shown, for the GRN equations models SUM and MULT, that the synchrony spaces are completely determined by their structural matrices (adjacency matrices $W^{\pm}$ in case of SUM models; multiplicity matrices $M^{\pm}$ and adjacency matrices $W^{\pm}$ in case of MULT models). See Propositions 3.4 and 3.12.

For regulator functions that are not Hill-like this may not be true: there may be other synchrony subspaces forced by those regulator functions, which we call *regulatory-dependent synchrony spaces.*

We illustrate this phenomena with the next example.

Example 5.1. *Consider the repressor and activation functions given by (2.13)-(2.14), in Section 2.3 of the GRN model for the mammalian circadian oscillator, and the GRN MULT model equations*

$$\begin{cases} \dot{x}_1 = \mathrm{act}(x_2, r)\mathrm{rep}(x_3, r) - x_1 \\ \dot{x}_2 = \mathrm{act}(x_1, a) - x_2 \\ \dot{x}_3 = \mathrm{act}(x_1, a) - x_3 \\ \dot{x}_4 = \mathrm{rep}(x_3, 1) - x_4 \end{cases} \tag{5.29}$$

for any $a \geq 1$, $r \geq 0$ which can be represented by a multi-arrowed graph. See Figure 12(left). We have that

$$S_o = \{x_2 = x_3\}, \qquad S = \{x_1 = x_4, x_2 = x_3\}$$

are synchrony subspaces of (5.29). More precisely, S_o is a synchrony subspace, since $P = \{\{1\}, \{2, 3\}, \{4\}\}$ is a synchronization partition for the MULT model equations (5.29). On the other hand, S is flow-invariant for the equations (5.29) due to the fact that $\mathrm{act}(x, r)\mathrm{rep}(x, r) = \mathrm{rep}(x, 1)$ holds for all $r, x \geq 0$. Thus, S is a regulatory dependent synchrony pattern for equations (5.29). Note that $I_1^+ = \{2\}$ while $I_4^+ = \emptyset$. The synchrony subspace $S_o = \{x_2 = x_3\}$ can be recognized on the graph, since cells $2, 3$ have identical positive input sets. However, $S = \{x_1 = x_4, x_2 = x_3\}$ may not be immediately identified looking at the graph.

One way to observe that S is a synchrony subspace, is to start with $S_o = \{x_2 = x_3\}$ and consider its quotient network, which can be represented by two different graphs. Indeed, when restricted to S_o, (5.29) becomes

$$\begin{cases} \dot{x}_1 = \mathrm{act}(x_{2,3}, r)\mathrm{rep}(x_{2,3}, r) - x_1 \\ \dot{x}_{2,3} = \mathrm{act}(x_1, a) - x_{2,3} \\ \dot{x}_4 = \mathrm{rep}(x_{2,3}, 1) - x_4 \end{cases} \tag{5.30}$$

or equivalently,

$$\begin{cases} \dot{x}_1 = \mathrm{rep}(x_{2,3}, 1) - x_1 \\ \dot{x}_{2,3} = \mathrm{act}(x_1, a) - x_{2,3} \\ \dot{x}_4 = \mathrm{rep}(x_{2,3}, 1) - x_4 \end{cases} \tag{5.31}$$

due to the relation $\mathrm{act}(x, r)\mathrm{rep}(x, r) = \mathrm{rep}(x, 1)$. *They correspond to Figure 12 middle and right, respectively.* ◇

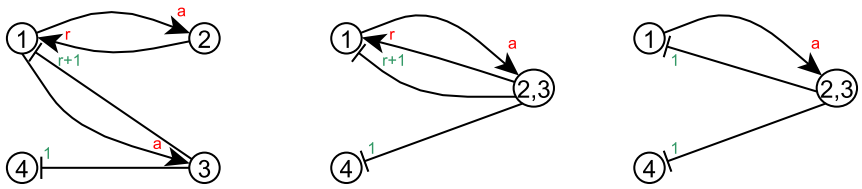

Figure 12: Left: The graph representation of the 4-gene MULT model GRN equations (5.29). When restricted to $S_o = \{x_2 = x_3\}$, the system (5.29) has two different graph representations: one corresponds to (5.30) (middle) and the other corresponds to (5.31) (right). This is due to the particular relation between the activation and repression functions that are considered in equations (5.29).

As shown by Example 5.1, regulatory dependent synchrony spaces (which are not inferred by the network structure) can exist, which are directly related to the modelling of the network itself including the choice of regulator functions and choice of models. Also, as shown in Example 3.11, hidden relations between act and rep functions can cause cancellations of regulator functions, which results in the "de-coupling" from regulating genes.

6. Conclusions and Outlook

This work is a first contribution to show how the theory of coupled cell systems can play a relevant role in the study of GRNs. We considered two dynamical models of GRNs, depending on whether the gene regulatory is additive (the SUM model) or multiplicative (the MULT model). Using theoretical results from coupled cell networks, we analyse the robust patterns of synchrony supported by these gene regulatory models and found out that the gene synchronization patterns can be quite different for the SUM and MULT models in general. Moreover, we have shown that other unexpected synchrony patterns may occur when the activation and repression functions satisfy specific relations.

Motivated by the study of the process of gene duplication and the divergence phenomena of subfunctionalization, we explore the concept of quotient networks and network lifting in both SUM and MULT models. From our results, it follows in particular that, if a SUM or MULT model of equations for a small GRN presents, for example, oscillatory behaviour, then we can enumerate bigger GRNs that admit this small GRN as a quotient network while preserving the initial oscillatory behaviour with some of the genes being synchronized. That is, each of these bigger GRNs will be a lifting of the smaller one with the guarantee that the bigger GRN has a synchrony space on which the associated quotient SUM or MULT equations restricted to the synchrony space are precisely the SUM or MULT model equations of the smaller GRN presenting the oscillatory behaviour. Furthermore, we described a method for constructing such network lifts. It was pointed out that there is no uniqueness in this process and so the choice of the lifts can be adapted to the particular types of applications under consideration.

Although our results are theoretical and by no means indicate how they stand in relation to empirical evidence or experimental data in GRNs, we wish to remark that coupled cell network formalisms and GRNs do share common features in structural dynamical properties. Therefore, it could be worthwhile incorporating theoretical considerations using coupled cell networks in the discussion of dynamical processes that have been addressed in GRNs.

In a future work, motivated by the work presented in [38] (see Remark 2.6), we aim to explore other potential approaches of embedding an oscillatory gene network with fewer genes into a bigger gene network, without using synchrony subspaces and where all genes may oscillate in a non-synchronized fashion.

Acknowledgments

The authors thank the reviewers for the suggestions that have improved the readability of the paper. MA and AD were partially supported by CMUP (UID/MAT/00144/2013), which is funded by FCT (Portugal) with national (MEC) and European structural funds (FEDER), under the partnership agreement PT2020. HR would like to thank Dr. Sabine Le Borne at Technical University of Hamburg, Germany for supporting her independent research and wish to express her gratitude to Hefei University, China for dynamic research exchanges and communications. HR thanks CMUP and the Mathematics Department of the University of Porto for the partial finantial support and the hospitality during her short to Porto in July of 2020.

References

[1] M.Aguiar, P.Ashwin, A.Dias and M.Field. Dynamics of coupled cell networks: synchrony, heteroclinic cycles and inflation. *J. Nonlinear Sci.* **21** (2011) (2) 271–323.

[2] M.A.D.Aguiar and A.P.S.Dias. The Lattice of Synchrony Subspaces of a Coupled Cell Network: Characterization and Computation Algorithm. *J. Nonlinear Sci.* **24** (2014) (6) 949–996.

[3] M.A.D.Aguiar and A.P.S.Dias. Synchronization and Equitable Partitions in Weighted Networks. *Chaos* **28** (2018) (7) 073105.

[4] M.A.D.Aguiar, A.P.S.Dias and F.Ferreira. Patterns of synchrony for feed-forward and auto-regulation feed-forward neural networks. *Chaos* **27** (2017) 013103.

[5] M.A.D.Aguiar, A.P.S.Dias, M.Golubitsky and M.C.A.Leite. Bifurcations from regular quotient networks: A first insight. *Physica D: Nonlinear Phenomena* **238** (2009) 137–155.

[6] U.Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC Press, Boca Raton, FL, 2006.

[7] F.Antoneli, A.P.S.Dias and R.C. Paiva. Hopf Bifurcation in Coupled Cell Networks with Interior Symmetries. *SIAM Journal on Applied Dynamical Systems* **7** (2008) (1) 220–248.

[8] F.Antoneli, M.Golubitsky and I.Stewart. Homeostasis in a feed forward loop gene regulatory motif. *J. Theoret. Biol.* **445** (2018) 103–109.

[9] F.Antoneli and I.Stewart. Symmetry and Synchrony in Coupled Cell Networks 1: Fixed-Point Spaces. *Int. Jour. of Bif. and Chaos* **16** (2006) (3) 559–577.

[10] M.Bansal, G.D.Gatta, D.di Bernardo. Inference of gene regulatory networks and compound mode of action from time course gene expression profiles. *Bioinformatics* **22** (2006) 815–822.

[11] M.Baumann, J.Pontiller and W.Ernst. Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: An overview. *Mol. Biotechnol.* **45** (2010) 241–247.

[12] M.A.Bee, and S.Tavazoie. Predicting gene expression from sequence. *Cell* **117** (2004) (2) 185–198.

[13] L.Bintu, N.E.Buchler, H.G.Garcia, U.Gerland, T.Hwa, J.Kondev, R.Phillips. Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev* **15** (2005) (2) 116–124.

[14] L.Bintu, N.E.Buchler, H.G.Garcia, U.Gerland, T.Hwa, J.Kondev, T.Kuhlman, R.Phillips. Transcriptional regulation by the numbers: applications. *Curr. Opin. Genet. Dev.* **15** (2005) (2) 125–135.

[15] C.A.Buzzi and J.Llibre. Hopf bifurcation in the full repressilator equations. *Math. Meth. Appl. Sci.* **38** (2015) 1428–1436.

[16] G.Chesi. Polynomial relaxation-based conditions for global asymptotic stability of equilibrium points of genetic regulatory networks. *International Journal of Systems Science* **41** (2010) (1) 65–72, DOI: 10.1080/00207720903078867

[17] G.Chesi and Y.S.Hung. Stability Analysis of Uncertain Genetic SUM Regulatory Networks. *Automatica* **44** (2008) (9) 2298–2305.

[18] L. DeVille and E. Lerman. Modular dynamical systems on networks. *J. Eur. Math. Soc.* **17** (2015) 2977–3013.

[19] G.T.Dewey and D.J.Galas. Gene Regulatory Networks. In: *Madame Curie Bioscience Database* [Internet]. Austin (TX): Landes Bioscience; 2000-2013. Available from: https://www.ncbi.nlm.nih.gov/books/NBK6167/

[20] F.E.Elahi and A.Hasan. A method for estimating Hill function-based dynamic models of gene regulatory networks. *R. Soc. open sci.* **5** (2018) (2) 171226.

[21] M.B.Elowitz. *Transport, Assembly, and Dynamics in Systems of Interacting Proteins.* Thesis, Princeton Univ. Princeton, 1999.

[22] M.B.Elowitz and S.Leibler. A Synthetic Oscillatory Network of Transcriptional Regulators. *Nature* **403** (2000) 335–338.

[23] M.Field. Combinatorial dynamics. *Dynamical Systems* **19** (2004) (3) 217–243.

[24] A.Force, M.Lynch, F.B.Pickett, A.Amores, Y.Yan and J.Postlethwait. Preservation of Duplicate Genes by Complementary, Degenerative Mutations. *Genetics* **151** (1999) (4) 1531–1545.

[25] T.S.Gardner, D.di Bernardo, D.Lorenz, J.J.Collins. Inferring genetic networks and identifying compound mode of action via expression profiling. *Science* 301 (2003) 102–105.

[26] M.Golubitsky and I.Stewart. *The Symmetry Perspective: From Equilibrium to Chaos in Phase Space and Physical Space*. Birkhauser, 2002.

[27] M.Golubitsky and I.Stewart. Nonlinear dynamics of networks: the groupoid formalism, *Bull. Amer. Math. Soc.* **43** (2006) (3) 305–364.

[28] M.Golubitsky, I.N.Stewart and D.G.Schaeffer. *Singularities and Groups in Bifurcation Theory: Vol. II.* Applied Mathematical Sciences, Springer-Verlag **69**, 1988.

[29] M.Golubitsky, I.Stewart, and A.Török. Patterns of synchrony in coupled cell networks with multiple arrows. *SIAM J. Appl. Dyn. Syst.* **4** (2005) (1) 78–100.

[30] J.-F.Gout and M.Lynch Maintenance and Loss of Duplicated Genes by Dosage Subfunctionalization. *Mol. Biol. Evol.* **32** (2015) (8) 2141–2148.

[31] Z.Guo, W.Jiang, N.Lages, W.Borcherds and D.Wang. Relationship between gene duplicability and diversifiability in the topology of biochemical networks. *BMC Genomics* (2014) 15:577.

[32] M.W.Hahn. Distinguishing Among Evolutionary Models for the Maintenance of Gene Duplicates. *Journal of Heredity* **100** (2009) (5) 605–617.

[33] T.F.Hansen. *Is modularity necessary for evolvability?* Remarks on the relationship between pleiotropy and evolvability. *Biosystems* **69** (2003) 83–94.

[34] P.H.von Hippel. From "simple" DNA-protein interactions to the macromolecular machines of gene expression. *Annu. Rev. Biophys. Biomol. Struct.* **36** (2007) 79–105.

[35] A.van Hoof. Conserved Functions of Yeast Genes Support the Duplication, Degeneration and Complementation Model for Gene Duplication. *Genetics* **171** (2005) 1455–1461.

[36] G.Karlebach and R.Shamir. Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.* **9** (2008) 770–780.

[37] E.Klipp, W.Liebermeister, C.Wierling, A.Kowald. *Systems Biology.* Wiley-VCH, Weinheim, Germany, 2016.

[38] A.Korenčič, R.Kosir, G.Bordyugov, R.Lehmann, D.Rozman and H.Herzel. Timing of circadian genes in mammalian tissues. *Scientific Reports* **4** (2014) 5782.

[39] E.Kuzmin, B.VanderSluis, A.N.Nguyen Ba, W.Wang, E.N.Koch, M.Usaj, A.Khmelinskii, M.Mattiazzi Usaj, J.van Leeuwen, O.Kraus, A.Tresenrider, M.Pryszlak, M.-C.Hu, B.Varriano, M.Costanzo, M.Knop, A.Moses, C. L.Myers, B. J.Andrews, C.Boone. Exploring whole-genome duplicate gene retention with complex genetic interaction analysis. *Science* **368** (2020), Issue 6498.

[40] J.Lawrence. Selfish operons: the evolutionary impact of gene clustering in prokaryotes and eukaryotes. *Curr. Opin. Genet. Dev.* **9** (1999) 642–648.

[41] C.Li, L.Chen and K.Aihara. Stability of genetic networks With SUM regulatory logic: Lur'e system and LMI approach. *IEEE Transactions on Circuits and Systems I: Regular Papers* **53** (2006) 2451–2458.

[42] E.Liu, L. Li and L.Cheng. Gene regulatory network review. In: S.Ranganathan, K.Nakai, C.Sch´onbach C. and Gribskov, M. (eds.), *Encyclopedia of Bioinformatics and Computational Biology* **2** (2019) pp. 155–164, Oxford: Elsevier.

[43] M.Lynch and J.Conery. The evolutionary fate and consequences of duplicate genes. *Science*, 290 5494 (2000) 1151–1155.

[44] H.McAdams, B.Srinivasan and A.Arkin. The evolution of genetic regulatory systems in bacteria. *Nat. Rev. Genet.* **5** (2004) 169–178.

[45] P.Mendes, W.Sha and K.Ye. Artificial gene networks for objective comparison of analysis algorithms. *Bioinformatics* **19** (2005) (2) ii122–ii129.

[46] F.Morone, I.Leifera and H.A.Maksea. Fibration symmetries uncover the building blocks of biological networks. *PANS* **117** (2020) (15) 8306–8314.

[47] M.A.Nowak, M.C.Boerlijst, J.Cooke and J.M.Smith. Evolution of genetic redundancy. *Nature* **388** (1997) 167–171.

[48] J.P.Pett, M.Kondoff, G.Bordyugov, A.Kramer, H.Herzel. Co-existing feedback loops generate tissue-specific circadian rhythms. *Life Sci. Alliance* **1** (2018) (3) e201800078. https://doi.org/10.26508/lsa.201800078

[49] J.P.Pett, A.Korenčič, F.Wesener, A.Kramer, H.Herzel. Feedback loops of the mammalian circadian clock constitute repressilator. *PLoS Comput Biol.* **12** (2016) (12) e1005266. Doi:10.1371/journal.pcbi.1005266

[50] M.Ptashne. *Genetic Switch: Phage Lambda Revisited*. Cold Spring Harbor Laboratory Press, New York, 2004.

[51] R.Röttger, U.Rückert, J.Taubert and J.Baumbach. How little do we actually know? On the size of gene regulatory networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **9** (2012) (5) 1293–1300. Doi: 10.1109/TCBB.2012.71.

[52] E.Segal, T.Raveh-Sadka and U.Gaul. Predicting expression patterns from regulatory sequence in Drosophila segmentation. *Nature* (2008) 451:535–540.

[53] I.Stewart, M.Golubitsky and M.Pivato. Symmetry groupoids and patterns of synchrony in coupled cell networks. *SIAM J. Appl. Dyn. Syst.* **2** (2003) (4) 609–646.

[54] S.A.Teichmann and M.M.Babu. Gene regulatory network growth by duplication. *Nat. Genet.* **36** (2004) (5) 492–496.

[55] V.Teif. Predicting Gene-regulation functions: Lessons from temperate bacteriophages. *Biophys J.* **98** (2010) (7) 1247–1256.

[56] J.Tischler, B.Lehner, N.Chen and A.G.Fraser. Combinatorial RNA interference in Caenorhabditis elegans reveals that redundancy between gene duplicates can be maintained for more than 80 million years of evolution. *Genome Biol.* **7** (2006) (8):R69.

[57] T.Vavouri, J.I.Semple and B.Lehner. Widespread conservation of genetic redundancy during a billion years of eukaryotic evolution. *Trends Genet.* **24** (2008) (10) 485–488.

[58] N.Vijesh, S.K.Chakrabarti and J.Sreekumar. Modeling of gene regulatory networks: A review. *Journal of Biomedical Science and Engineering* **6** (2013) (02) 223–231.

[59] C.Zhang and C.Li. Revealing the mechanism of lymphoid and myeloid cell differentiation and transdifferentiation through landscape quantification. *Phys. Rev. Research* **3** (2021) 013186.