Coherent regulation in yeast’s cell-cycle network

Neşe Aral and Alkan Kabakçioğlu
Department of Physics, Koç University, Rumelifeneri Yolu Sarayı 34450, Istanbul, Turkey

1 Author to whom any correspondence should be addressed.
E-mail: akabakcioglu@ku.edu.tr

Keywords: coherent gene regulation, Boolean networks, yeast cell cycle, Saccharomyces cerevisiae

Abstract
We define a measure of coherent activity for gene regulatory networks, a property that reflects the unity of purpose between the regulatory agents with a common target. We propose that such harmonious regulatory action is desirable under a demand for energy efficiency and may be selected for under evolutionary pressures. We consider two recent models of the cell-cycle regulatory network of the yeast, Saccharomyces cerevisiae as a case study and calculate their degree of coherence. A comparison with random networks of similar size and composition reveals that the yeast’s cell-cycle regulation is wired to yield an exceptionally high level of coherent regulatory activity. We also investigate the mean degree of coherence as a function of the network size, connectivity and the fraction of repressory/activatory interactions.

List of abbreviations

| Abbreviation | Definition |
|--------------|------------|
| GRN          | Gene regulatory network |
| YCC          | Yeast cell cycle |
| TF           | Transcription factor |

1. Introduction

Despite vast amount of data on genetic regulatory interactions in a growing number of organisms [1–7], the interplay between structure and function in these systems is still unclear and an active field of research [8–10]. Much effort has been devoted to local structures and their role in promoting multistationarity, stability and robustness in a background of fluctuating environmental conditions [11–14]. Other studies argue that regulatory dynamics is ‘consistent’ in the sense that, a transcription factor’s (TF) influence on a target gene (activation/inhibition) is not context dependent [15]. Similarly, architectural features allowing better controllability have been of recent interest [16–18]. Such guiding principles provide valuable insight and a bird’s-eye perspective on networks of gene regulatory interactions, where a qualitative understanding of the underlying generative mechanisms is hindered by the prohibitive complexity of the cell and the shear amount of experimental data.

We here introduce the concept of ‘coherent regulation’ as a similar guideline for the evolutionary design of regulatory networks. Below, we give a mathematical definition for the degree of coherence, as a measure of the extent to which a given regulatory network structure and the equations of motion for the regulation dynamics are optimized towards minimum waste of material/energy resources available to the cell. We then calculate it for the yeast cell-cycle (YCC) network, a well-studied subnetwork responsible for cell-division in the yeast Saccharomyces cerevisiae [19], by means of two previously proposed models in the literature. Further analyses on random graphs yield a reference baseline as a function of the network’s structural parameters, by means of which we show that the yeast’s cell-cycle network is, in fact, optimally organized to maximize coherent regulatory action.

2. Model

2.1. Coherent regulation
Regulation of gene activity requires production of regulatory proteins which demands energy and raw material. The need for the economical use of these resources has been a constant and dominant evolutionary pressure in much of the last 4 billion years and
virtually for all species [20–22]. Therefore, it is natural to expect regulatory networks to have been optimized towards their optimal use.

A full-fetched analysis of such optimization would require a much deeper understanding than we currently have of a cell’s functions on the protein level. Here, we will discuss network efficiency only in terms of the choice of interactions that determine the time dependence of protein expression, and assert that a prerequisite for optimality is to avoid simultaneous production of regulatory elements that perform opposing tasks.

In particular, it is desirable that in a steady state, two regulatory elements where one is an activator and the other is a repressor of the same protein are not co-expressed. Exceptions are the relatively brief transition periods (induced by internal or environmental conditions) between steady states or between the states in cyclic attractors (such as circadian cycles [23, 24]), where changes in the expression levels are due to role switches between such antagonistic regulators.

We coin the concordance of simultaneously active regulatory elements sharing a common target as ‘coherent regulation’. More precisely, we say that a gene is coherently regulated if its expression state is determined unanimously by its active regulatory partners, as described in figure 1. Note that, coherence defined as above is different from ‘cooperativity’ in regulatory networks discussed in the literature [25–27], which refers to increase (or decrease) of the binding affinity of a TF when other TFs are already bound to the promoter region of the gene. It can also be noted that, a coherent network as defined above is expected to be more robust against fluctuations in the expression rates of the co-regulating elements. This is because, when the regulatory inputs received by a gene are compatible, fluctuations in their relative strengths (e.g., due to time delays) are less significant than when the regulatory agents act antagonistically, in which case the fluctuations may result in a sign change in the net regulatory message received by the target gene.

The degree of coherent regulation is clearly structure dependent, i.e., some networks are more supportive of such harmonious regulatory activity than others. Identifying the architectural aspects of these networks is an interesting problem in itself which shall be addressed in a future study. The efficiency acquired through coherence also has a truly dynamical aspect, since coherent action is reached through the time-dependent interplay between expression levels of communicating genes. Starting from the fact that the structure shapes regulatory dynamics in a phenotype, the question we ask here is whether evolutionary processes may have provided the means to feed information from the dynamics back to the structure, promoting modifications that maximize coherent regulatory action.

2.2. Quantifying coherence in regulatory networks

Given a fixed environment, the genes/proteins that participate in a genetic regulatory network spend most of their time in the attractors of the regulation dynamics. Let \( \tau(\sigma) \) be the appropriate discrete-time-evolution operator \( \tau \) that propagates the expression levels \( \{\sigma_j\} \) associated with the nodes \( j = 1 \ldots N \) of the network from time \( t \) to \( t + \delta t \), i.e., \( \sigma(t + \delta t) = \tau(\sigma(t)) \). Then the attractors are determined by the condition \( \sigma^* = \tau^q[\sigma^*] \). A point attractor corresponds to \( q = 1 \). Otherwise, the attractor is a cycle of period \( q \delta t \) with the cycle states \( \{\sigma^*, \tau[\sigma^*], \ldots, \tau^{q-1}[\sigma^*]\} \). We here consider Boolean models where \( \sigma \) is a binary vector with entries \( \epsilon \in \{0, 1\} \).

We can now define the ‘coefficient of coherence’ for an attractor \( i \) as \( \alpha_i^{(i)} \), the fraction of genes which are not subject to simultaneous and conflicting regulatory messages from their regulating partners active in state \( i \). Describing up-(down-) regulation of gene \( j \) by the associated regulators \( k_j \) by the associated regulators \( c_{\beta,j} = 1(-1) \), we can express \( \alpha_i^{(i)} \) in terms of (Boolean) expression levels as:

![Figure 1. Coherent versus incoherent regulation for a node in the regulatory network with two regulating partners.](image-url)
work is de
uniformly weighted averages, in order to demonstrate
here present our results with both basin-weighted and
and, therefore, is irrelevant in the biological sense. We
typically never appear throughout the life cycle of a cell
because a randomly constructed expression state will
weighted averaging should be employed with caution,
i

where $N$ is the number of nodes in the network, 11 in
our case. $i$ is the attractor index, hence for a network
with $n_a$ attractors $i$ runs from 1 to $n_a$. $\text{int}(\cdot)$ function
returns the integer part of its argument, which
effectively sets the coefficient of coherence to unity if
the currently expressed subset of regulators $\{k_j\}$ of
gene $j$ are all activators or all repressors, and to zero
otherwise. In other words, coherence of a node at a
given attractor is also described by a Boolean variable,
in the spirit of the models considered here. Possible
generalizations to continuous and stochastic models is
discussed at the end of the paper. Also note that, the
trivial fixed point $\sigma = 0$ (common to both networks)
is left out in our analysis so that the denominator in
equation (1) never vanishes.

For a cyclic attractor with period $q$, $\delta t$, $\alpha^{(j)}(i)$ is taken
as the arithmetic average over the cycle states $\sigma^{(j)}(k)$
($k = 1,\ldots,q$). The coefficient of coherence for a net-
work is defined to be the mean of $\alpha^{(j)}(i)$ over all the
attractors $i$, i.e.

\[ \alpha_c = \frac{1}{n_a} \sum_{i=1}^{n_a} \alpha^{(j)}(i) \]

where, again, $n_a$ is the number of attractors. Alternatively,
one could consider the average weighted by the basin sizes (the fraction of uniformly randomly picked
initial states that end up in a given attractor):

\[ \alpha_{\text{basin}} = \frac{1}{\sum_{i=1}^{n_a} b_i} \sum_{i=1}^{n_a} b_i \alpha^{(j)}(i), \]

where $b_i$ is the number of initial states reaching
attractor $i$ under the given dynamics. However, such
weighted averaging should be employed with caution,
because a randomly constructed expression state will
typically never appear throughout the life cycle of a cell
and, therefore, is irrelevant in the biological sense. We
here present our results with both basin-weighted and
uniformly weighted averages, in order to demonstrate
that our conclusions do not critically depend on this
choice.

### 3. Results

#### 3.1. Yeast’s cell-cycle: a case study

The budding yeast, *S. cerevisiae*, is a well-studied
single-cell eukaryote [2, 28–30]. Its 307 known TFs
out of ~6000 distinct proteins participate in (currently
estimated) ~2 × 10^5 regulatory associations [2]. Yet,
these are still not sufficiently well characterized to tell
whether an arbitrarily chosen TF up- or down-
regulates a target gene in a given context. Therefore,
the structural optimality (in the above sense) of the
yeast’s regulatory network as a whole is not possible
to assess with the currently available information.

Nevertheless, the small subnetwork responsible
for the cell-cycle in yeast (YCC) is very well studied
[19, 31–34]. A handful of regulatory proteins or com-
plexes that drive the cell division and the interactions
between them have been identified. Furthermore, sev-
eral models of regulatory dynamics involving these
proteins have been shown to yield the experimentally
observed expression profiles [19, 33, 34]. We will con-
sider two such models here, and demonstrate their
high degree of coherence as a proof of concept.

Once the cell reaches a critical size [35, 36], the cell
division is initiated. It starts from the $G_1$ phase of the
mother cell and ends in the $G_2$ phases of the mother
daoth cells. The two models we consider here
reproduce in proper order all the intermediate expres-
station stages that the cell traverses during division, via
the time-evolution equations given in table 1. They
feature the same set of key TFs/proteins/complexes
(vertices of the network) that drive DNA replication
and the accompanying cell division in yeast. These
models are depicted in figure 2, where blue nodes indi-
cate auto-repressor elements ($c_{ii} = -1$) and two-
colored (blue+magenta) nodes represent multi-pro-
tein complexes. The model proposed by Li *et al* [19] in
figure 2(a) is, to our knowledge, the first successful
implementation of Boolean dynamics to yeast’s cell-
and has been widely used in later studies. We will
refer to this model as M1 from now on. More recently,
Mangla *et al* improved M1 by introducing additional
interactions and varying interaction weights, while at
the same time extending the Boolean formalism to
allow some of the vertices to have three different states 
(0, 1, 2) [33]. This work was later refined by Hong et al [34] 
in order to capture crucial checkpoint conditions 
overlooked in M1. We will refer to the YCC model of 
Hong et al (shown in figure 2(b)) as M2 from now on.

Time-evolution equations for M1 and M2 are 
given in table 1. Here, \( \sigma_i(t) \) is the expression state 
which takes the value \( '1' \) when the node \( i \) is active and 
\( '0' \) otherwise, except for Swi5 and Clb2 which can also 
take the value of \( '2' \) (strong expression) in M2. The 
interactions considered in M1 (figure 2(a)) differ only 
by their sign, i.e., \( c_{ij} \in \{-1, 0, 1\} \), while for M2 
\( c_{ij} \in \{0, \pm 1/3, \pm 1, \pm 3\} \) (shown with different edge 
thicknesses in figure 2(b)). \( \theta_{1,1} \) and \( \theta_{1,2} \) are threshold 
values that trigger the state shifts \( 0 \leftrightarrow 1 \) and \( 1 \leftrightarrow 2 \), 
respectively. While reproducing the time-evolution of 
M2 given in [34], we concluded that the observed 
dynamics and basin sizes are reproducible only if an 
additional parameter, \( \delta \), that introduces a hysteresis to 
up/down regulation events is added to the evolution 
dynamics (see table 1). This crucial detail appears to be 
accidentally omitted in the reference article. We give 
the numerical values of the parameters above in table 1 
and refer the reader to the original references for fur-
ther details of the two models.

While a deterministic implementation of the 
above equations follows the experimentally resolved 
expression dynamics, a stochastic version will occasion-
ally end up in different attractors listed in table 2. 
The biological relevance of such relatively low prob-
ability states is unclear [19, 37]. Nevertheless, the 
degree of coherence supported by the structure of the 
regulatory network under the dynamical evolution 
given in table 1 should to be defined through all of its 
attractors. This also facilitates a fair comparison of the 
YCC network with randomized ensembles for which a 
classification of attractors according to their biological 
relevance is meaningless. On the other hand, we do 
present numerical results obtained through basin-size 
weighted averaging, which, in the case of YCC, corre-
sponds to essentially focusing on the \( G_1 \) attractor.

![Figure 2](image.png)

**Figure 2.** Two model systems for yeast cell cycle. (a) Model M1 by Li et al (b) model M2 by Hong et al. The weights for interactions in 
M2 are 1, 1/3 and 3 for normal, thin and dashed lines, respectively.

| F     | Clb3 | MBF | Cln1.2 | Cdc14 | Cln1 | Clb1 | Sic1 | Cln3 | DNA repl. |
|-------|------|-----|--------|-------|------|------|------|------|----------|
|       |      |     |        |       |      |      |      |      |          |
| F1    |      |     |        |       |      |      |      |      |          |
| F2    |      |     |        |       |      |      |      |      |          |
| F3    |      |     |        |       |      |      |      |      |          |
| F4    |      |     |        |       |      |      |      |      |          |
| F5    |      |     |        |       |      |      |      |      |          |
| F6    |      |     |        |       |      |      |      |      |          |
| F7    |      |     |        |       |      |      |      |      |          |
| F8    |      |     |        |       |      |      |      |      |          |

**Table 2.** YCC attractors under the dynamics given in table 1 and 
their coherence, \( \alpha_c \). Black boxes indicate the active genes and the 
first row corresponds to the experimentally verified \( G_1 \) phase. The 
trivially coherent \( \sigma = 0 \) fixed point is left out of the analysis.

| F     | Clb3 | MBF | Cln1.2 | Cdc14 | Cln1 | Clb1 | Sic1 | Cln3 | DNA repl. | \( \alpha_c \) |
|-------|------|-----|--------|-------|------|------|------|------|----------|-------------|
|       |      |     |        |       |      |      |      |      |          |             |
| F1    |      |     |        |       |      |      |      |      |          |             |
| F2    |      |     |        |       |      |      |      |      |          |             |
| F3    |      |     |        |       |      |      |      |      |          |             |
| F4    |      |     |        |       |      |      |      |      |          |             |
| F5    |      |     |        |       |      |      |      |      |          |             |
| F6    |      |     |        |       |      |      |      |      |          |             |
| F7    |      |     |        |       |      |      |      |      |          |             |
| F8    |      |     |        |       |      |      |      |      |          |             |

| F     | Clb3 | MBF | Cln1.2 | Cdc14 | Cln1 | Clb1 | Sic1 | Cln3 | DNA repl. | \( \alpha_c \) |
|-------|------|-----|--------|-------|------|------|------|------|----------|-------------|
|       |      |     |        |       |      |      |      |      |          |             |
| F1    |      |     |        |       |      |      |      |      |          |             |
| F2    |      |     |        |       |      |      |      |      |          |             |
| F3    |      |     |        |       |      |      |      |      |          |             |
| F4    |      |     |        |       |      |      |      |      |          |             |
| F5    |      |     |        |       |      |      |      |      |          |             |
| F6    |      |     |        |       |      |      |      |      |          |             |
| F7    |      |     |        |       |      |      |      |      |          |             |
| F8    |      |     |        |       |      |      |      |      |          |             |
Also given in table 2 are the coefficients of coherence for the attractors of M1 and M2, calculated using equation in table 1. For the whole network, we obtain $\alpha_1 = 0.97$ (M1) and 0.89 (M2) with uniformly weighted attractors and 0.99 (M1) and 0.98 (M2) with basin-size weighted averaging. We observe that, all the common attractors of M1 and M2 (including the biologically relevant $G_1$ phase) are either fully coherent or almost so.

Is the YCC network optimized towards maximal coherence? A meaningful assessment of the above numbers is possible only in the background of corresponding results on properly selected random ensembles (for a similar analysis to determine frequent network motifs in biological networks see, for example, [12, 38].) Therefore we next perform an analysis on directed graphs of the same size with similar characteristics. We do this in two different ways, as explained below.

3.2. Coherence: yeast versus random networks

Even a network composed of random associations, each gene up- or down-regulating an arbitrary subset of genes, will display a nonzero degree of coherence determined by the laws of statistics. Two obvious factors that contribute to $\alpha_1$ are: the number of regulating partners per node, $k$, and the fraction of up-regulating interactions in the network, $p$. (See section 3.3.) While we will investigate the dependence of $\alpha_1$ (average of $\alpha_1$ values over an ensemble of networks) on $k$ and $p$ below, it is sensible to compare yeast’s cell-cycle network against the appropriate random ensemble with $(k, p) = (k_{\text{YCC}}, p_{\text{YCC}})$. Excluding self loops, model M1 above has 14 inhibitory and 15 activatory interactions, that is, $(k_{\text{YCC}}, p_{\text{YCC}}) \approx (2.63, 0.52)$. Similarly, model M2 has 20 inhibitory and 15 activatory interactions, yielding $(k_{\text{YCC}}, p_{\text{YCC}}) \approx (3.18, 0.43)$.

In order to construct an ensemble of networks with fixed $(N, k, p)$, we first generated connected, undirected random graphs with $N$ regulatory elements (nodes), and $N k$ interactions (edges) between them ($k \geq 1$ is required for connectivity.) Then, each edge was assigned one of the two possible interaction types (activator or repressor), such that a fraction $p$ of them up-regulate, and the rest down-regulate their targets. Note that, a node may act as an activator for one target and as a repressor for another. This is in line with the behavior observed in YCC, as well as recent research which suggests that the TFs involved in both kinds of regulation may be more common than previously thought [39, 40]. We compared models M1 and M2 with the corresponding random ensembles R1 and R2 generated using network parameters $(N_{\text{YCC}}, k_{\text{YCC}}, p_{\text{YCC}})_{1,2,}$ respectively.

We also applied a more stringent test of coherence in the chosen YCC models by comparing them with more ‘yeast-like’ ensembles. These ensembles (Y1 and Y2) were generated by forcing each node to preserve its number of incoming and outgoing edges, separately for each type of interaction. This was achieved by a simple edge-shuffling procedure that switches the targets of two randomly selected edges in the network, subject to the condition that the edges are of the same type. In all the random ensembles considered, we also separately fixed the fraction of the auto-repressor nodes ($c_{ij} = -1$) to the corresponding values 5/11 (M1) and 3/11 (M2), for a fair comparison.

Overall, two sets of $10^4$ random networks (R1, Y1) and (R2, Y2) were generated for each YCC model. The probability distributions of $\alpha_1$ obtained from each ensemble are shown in figures 3 and 4 separately for M1 and M2, together with the value obtained from the actual models for comparison.

Next, we compared the $\alpha_1$ distributions above with those from the ‘yeast-like’ ensembles Y1 and Y2, which preserve the particular set of in-/out-degree pairs for the two interaction types at each node of the model networks M1 and M2. Surprisingly, Y1 (both with and without basin-size weighted averaging) and Y2 (without basin-size weighted averaging) yield a reduced degree of coherence when compared with the respective ensembles R1 and R2, as is evident from figure 3. The corresponding mean values along with standard deviations for $\alpha_1$ are listed in table 3. We found that the coherence values of the actual networks lie in the top 4% in all of the random ensembles R1, R2, Y1 and Y2. Based on these observations, we conclude that the exceptionally high level of coherent regulatory activity observed in the two YCC models is not solely due to the particular local decoration of the networks’ nodes.

3.3. Coherence versus network parameters

In order to investigate the dependence of mean degree of coherence on $k$ and $p$, we next considered (for simplicity of implementation) only the YCC model M1 above. We, again, generated $10^4$ distinct, random networks of size $N = 11$ (as in YCC), now with $k \in \{1.6, 5\}$ and $p = 0.5 \approx 0.5 p_{\text{YCC}}$ as well as with $k = k_{\text{YCC}} = 29/11$ and $p \in [0, 1]$. Possible isomorphism of the ensemble members was ruled out by ensuring that each network in the ensemble has a different eigenvalue spectrum. We also generated much larger networks with $N = 110$ nodes (corresponding to a 10 fold increase in size) for comparison and with both $k$ and $p$ values in the same intervals as above, in order to check the network size dependence of our results. Then the attractors of the dynamics given in table 1 were found for each network by an exhaustive search for $N = 11$ and by sampling $2^{11}$ initial states for $N = 110$. $\alpha_1$ were calculated separately for each network over corresponding attractors of the dynamical evolution described by equation (2). The dependence of the ensemble mean $\alpha_1$ on $p$ and $k$ is shown in figures 5(a)–(d) respectively.
\(\alpha_{\bar{c}}\) changes only slightly with the number of nodes \(N\) for a given \((k, p)\) pair: increasing the network size 10 fold yields a similar behavior, with \(<9.8\%\) difference in the worst case (figure 5). Therefore, despite the smallness of the model system considered here, our results on random ensembles may be expected to serve as a reasonable null-hypothesis for coherence in the global regulatory network of yeast and other organisms.

It may at first sight be surprising that \(\alpha_{\bar{c}}\) is asymmetric with respect to \(p = 1/2\), given that the number of genes with incoming edges that both up- and down-regulate the gene is maximized precisely at this point. The observed asymmetry is due to the fact that the fraction of active nodes at the fixed points is a monotonically increasing function of \(p\) (the insets of figure 5). Close to \(p = 0\) and \(p = 1\), most genes are regulated in the same direction by their regulating

---

**Table 3.** Mean and standard deviation values for \(\alpha_{c}\) of random ensembles along with the \(\alpha_{c}\) values of model networks M1 and M2.

|          | Uniformly weighted | Baseline weighted |
|----------|--------------------|-------------------|
| \(\alpha_{c} \pm \sigma\) | 0.85 ± 0.08        | 0.81 ± 0.09       |
|          | 0.82 ± 0.07        | 0.73 ± 0.10       |
|          | 0.97               | 0.99              |
|          | 0.79 ± 0.07        | 0.75 ± 0.11       |
|          | 0.75 ± 0.09        | 0.87 ± 0.06       |
|          | 0.89               | 0.98              |
partners, hence an increase in coherence is observed in these extremes. While the fraction of genes with conflicting incoming edges peaks at $p = p_12$, the minimum of $\alpha_c$ is shifted towards larger $p$ values (more up-regulation), where the chances are higher for antagonistic regulatory partners of a gene to be simultaneously expressed.

A simple analytical model based on the above reasoning can be shown to reproduce the observed behavior. Having shown the weak dependence of our results on the network size, let us consider the large-$N$ limit. For fixed $(k, p)$, define the average number of active regulating partners per node be $r_a(k, p)$. After the empirical observation that $r_a$ increases uniformly with $p$ for fixed $k$ (see figures 5(a), (b) inset), we fit the data for $N = 11$ and $N = 110$ with following analytical form for convenience:

$$ r_a(k, p) = r_{min} + p'(r_{max} - r_{min}) $$

(4)

where $r_{min,max}$ are $k$-dependent. Note that, a typical attractor of a network with all inhibitory interactions ($p = 0$) will still have a few nodes ‘on’ in absence of active down-regulating partners, hence $r_{min} > 0$. When $p = 1$, we expect most of the nodes to be turned on. A node is coherently regulated, if all its active regulating partners act unanimously. $\alpha_c$ measures essentially the probability for this event across the network’s attractors. Treating the states of neighbors regulating a given node as independent random variables, this probability can be modeled as that of observing two extreme outcomes in $r$ binomial trials where the binomial outcomes reflect the two possible assignments, up and down regulation with respective probabilities $p$ and $(1 - p)$, for each interaction between the target node and its $r$ active regulators. Substituting for $r$ its expected value $r_a$ with a mean-field approach, we obtain:

$$ \alpha_c(p, k) = p^r + (1 - p)^r. $$

(5)

Despite the empirical nature of our treatment, equations (4) and (5) yield a good fit for both $r_a$ and $\alpha_c$ with $(r_{min}, r_{max}, \gamma)$ values listed in table 4. The agreement between the basin-averaged $\alpha_c$ versus $p$ for $N = 110$ and the analytical form above is shown in figure 6. On the other hand, the parameter values yielding the best fit to $\alpha_c$ and $r_a$ differ. This discrepancy may be attributed to the nonuniformity of the in-degree distribution and the contribution of correlations, both of which are not captured by the simple treatment above.

The monotonic decrease of $\alpha_c$ with increasing $k$, the mean number of regulating partners per node, shown in figures 5(c), (d) is hardly surprising. It reflects the fact that the genes with fewer active
regulators are less likely to receive conflicting inputs. The number of active neighbors per node when \( p = 0.5 \approx p_{\text{YCC}} \) increases almost linearly with \( k \), suggesting that the overall number of simultaneously active nodes remains a constant, essentially determined by \( p \).

The results presented here are identical to the naked eye when non-point-attractors (with period \( > 1 \)) are excluded from the calculation of \( \alpha_c \) on random networks. Such cycles appear relatively infrequently with the given regulatory dynamics (see figure 6 inset). Therefore, the fact that the yeast’s networks in (figure 2) has only point attractors does not seem to be relevant to its extreme degree of coherence, either. Investigation of structural features of coherent networks (such as local motif statistics, directed cycles, etc) is planned as a future study.

4. Discussion

We defined a measure for the degree of coherence in gene regulatory networks (GRNs). Using the proposed measure, we showed that the cell-cycle network of the budding yeast displays exceptionally high degree of harmonious regulatory activity, which is in line with earlier observations on its robustness. The proposed measure of coherence sets the YCC network aside, within an ensemble of networks of the same size and composition. We found that, achieving coherent regulation is most difficult in systems when roughly 25% of the interactions are repressory. While this ratio is about 1/2 for the yeast’s cell-cycle network, YCC yields a degree of coherence which is significantly higher than those of random networks with identical network parameters such as size, average degree, and +/- interaction ratio. When basin basin sizes are ignored (uniform weighting), optimality of the YCC network is even more pronounced within the ensemble of yeast-like networks that preserve the in-/out-degrees of nodes separately for each interaction type. A deeper analysis of highly coherent random networks is necessary to pinpoint the structural determinants of these unique and possibly biologically relevant class of networks.

Although not considered in our work, the formalism here can be generalized to continuous models of gene expression by replacing the averages over states with time averages and employing an appropriate extension of the degree of coherence to continuous variables. For simple linearized models of the form

\[
\frac{dn_j(t)}{dt} = \sum_{k_j} f_{k_j} n_{k_j}(t),
\]

where \( n_j \) typically represent protein/RNA concentrations, a possible generalization could be to replace equation (1) with

\[
\alpha_c^{(i)} = \frac{1}{T_i N} \sum_{j=1}^N \int_0^{T_i} \left| \frac{1}{N} \sum_{k_j} f_{k_j} n_{k_j}(t) \right| dt,
\]

where \( T_i \) is the period of the \( i \)th dynamical attractor in consideration. Our formulation in section 2.2 can also be easily generalized to stochastic systems after substituting for \( r \) a stochastic transition matrix, and for \( \sigma \) a probability distribution on the ensemble of states \( \sigma \). While noise is a relevant determinant of cell activity [41–43], biological systems are typically robust to fluctuations in gene expression levels [44, 45]. Therefore, both stochastic and deterministic models of regulatory dynamics are frequently used in the literature [19, 46, 47].

Perhaps the most far-reaching question inspired by our findings is whether coherence in gene regulation is a prevalent motif across organisms in nature. In order to check this, the presented analysis needs to be extended to the cell-cycle regulation networks of other organisms (for example the fission yeast [48]), as well as GRNs responsible for different biological processes (examples in the literature include cell differentiation and segmentation networks [49–51]). Such an investigation is currently in progress and will shed light on the generality of our observations on the cell-cycle network of the budding yeast. An affirmative answer to the question above would point to the interesting possibility of a Hebbian-like selection mechanism for regulatory interactions. That is because, the coherence property defined in this work can be seen as a generalization of the Hebb’s rule for neural networks [52], which is usually stated as ‘neurons that fire together, wire together’. A mechanism of selection for coherent networks would similarly promote positive regulation (and demote negative regulation) among genes that are expressed simultaneously, and the vice versa for an ‘on’ gene regulating an ‘off’ gene. Rapid evolution of regulatory interactions as found in recent studies [53] resonates with this scenario, for it provides the fertile ground on which such a mechanism can effectively produce coherent networks in short evolutionary time scales.
Acknowledgments

A K is grateful for the hospitality of M Kardar during a visit to MIT and a stimulating discussion with the rest of the biophysics group there. We also thank D Yuret and C Dunn of Koç University for beneficial exchanges, and an anonymous referee for pointing out the references [33, 34].

References

[1] Yang T H, Wang C C, Wang Y C and Wu W S 2014 Database 2014 ba0114
[2] Teixeira M C et al 2014 Nucleic Acids Res. 42 161–6
[3] Guldener U et al 2005 Nucleic Acids Res. 33 364–8
[4] Münch R, Hiller K, Barg H, Heldt D, Linz S, Wingender E and Jahn D 2003 Nucleic Acids Res. 31 266–9
[5] Drysdale R A et al 2005 Nucleic Acids Res. 33 390–5
[6] Chatr-aryamontri A et al 2013 Nucleic Acids Res. 41 816–23
[7] Tsai H K, Lu H H S and Li W H 2005 Proc. Natl Acad. Sci. USA 102 13512–7
[8] Klemm K and Bornholdt S 2005 Proc. Natl Acad. Sci. USA 102 18414–9
[9] Mangan S and Alon U 2003 Proc. Natl Acad. Sci. USA 100 11980–5
[10] Thieffry D 2007 Briefings Bioinformatics 8 220–5
[11] Tyson J and Novak B 2010 Annu. Rev. Phys. Chem. 61 219
[12] Alon U 2006 An Introduction to Systems Biology: Design Principles of Biological Circuits ( Boca Raton, FL: CRC Press)
[13] Sevim V and Rikvold P A 2006 J. Theor. Biol. 253 323–32
[14] Cinquin O and Demongeot J 2002 C. R. Biologies 325 1085–95
[15] Sansom R 2008 Br. J. Phil. Sci. 59 169–200
[16] Yuan Z, Zhao C, Di Z, Wang W X and Lai Y C 2013 Nat. Commun. 4 2447
[17] Yan G, Ren J, Lai Y C, Lai C H and Li B 2012 Phys. Rev. Lett. 108 218703
[18] Wuchty S 2014 Proc. Natl Acad. Sci. 111 7156–60
[19] Li F, Long T, Lu Y, Ouyang Q and Tang C 2004 Proc. Natl Acad. Sci. USA 101 4781–6
[20] Zhang X, Jantama K, Moore J C, Jarobe L R, Sham mugam K T and Ingram I O 2009 Proc. Natl Acad. Sci. 106 20180–5
[21] Niven J E and Laughlin S B 2008 J. Exp. Biol. 211 1792–804
[22] Laughlin S B, van Steveninck R R d R and Anderson J C 1998 Nat. Neurosci. 1 36–41
[23] Eelderink-Chen Z, Mazzotta G, Sturre M, Bosman I, Roenneberg T and Merrow M 2010 Proc. Natl Acad. Sci. 107 2043–7
[24] Akman O E, Watterson S, Parton A, Binns N, Millar A J and Ghazal P 2012 J. R. Soc. Interface 9 2365–82
[25] Stefan M I and le Novère N 2013 PLoS Comput. Biol. 9 e1003106
[26] Hsu J T, Peng C H, Hsieh W P, Lan C Y and Tang C Y 2011 BMC Bioinformatics 12 281
[27] Banerjee N and Zhang M Q 2003 Nucleic Acids Res. 31 7024–31
[28] Kaizu K, Gho sh S, Matsuoka Y, Moriya H, Shimizu-Yoshida Y and Kitano H 2010 Mol. Syst. Biol. 6
[29] Hedges P E, Mc Kee A H, Davis B P, Payne W E and Garrels J I 1999 Nucleic Acids Res. 27 69–73
[30] Bader G D, Hilblut A, Andrews B, T yers M, Hughes T and Boone C 2003 Trends Cell Biol. 13 344–56
[31] Lee W B and Huang J Y 2009 FEBS Lett. 583 927–32
[32] Okabe Y and Sasai 2007 Biophys. J. 93 3451
[33] Mangla K, Dill D L and Horowitz M A 2010 PLoS One 5 e8906
[34] Hong C, Lee M, Kim D, Kim D, Cho K H and Shin I 2012 BMC Syst. Biol. 6 129
[35] Johnston G, Pringle J and Hartwell L 1977 Exp. Cell Res. 105 79–98
[36] Rupe l I 2002 Trends Genetics 18 479–85
[37] Albert R and Othmer H G 2003 J. Theor. Biol. 223 1–18
[38] Milo R, Shen- Orr S, Itzkovitz S, Kashtan N, Chklovskii D and Alon U 2002 Science 298 824
[39] Hickman M J and Winston F 2007 Mol. Cell. Biol. 27 7414–24
[40] Ptashne M 2011 Nat. Chem. Biol. 7 484–7
[41] Thattai M and van Oudenaarden A 2001 Proc. Natl Acad. Sci. 98 8614–9
[42] Swain P S, Elowitz M B and Siggia E D 2002 Proc. Natl Acad. Sci. 99 12798–809
[43] Paulsson J 2004 Nature 427 415–8
[44] Ingolia N T 2004 PLoS Biol. 2 e123
[45] Kaneko K 2007 PLoS One 2 e434
[46] Zhang Y, Qian M, Ouyang Q, Deng M, Li F and Tang C 2006 Physica D 219 35–39
[47] Fauré A, Naldi A, Chaouiya C and Thieffry D 2006 Bioinformatics 22 e124–e131
[48] Davidich M I and Bornholdt S 2008 PLoS One 3 e1672
[49] Remy E and Ruet P 2008 Bioinformatics 24 1220–1226
[50] Mendoza L, Thieffry D and Alvarez-Buylla E R 1999 Bioinformatics 15 593–606
[51] Krumkies J, Marr C, Schroeder T and Theis F J 2011 PLoS One 6 e22649
[52] Hebb D O 1949 The Organization of Behavior: An Neuropsychological Theory (New York: Wiley)
[53] Schmidt D et al 2010 Science 328 1036–40