Effective interaction graphs arising from resource limitations in gene networks

Yili Qian and Domitilla Del Vecchio

Abstract—Protein production in gene networks relies on the availability of resources necessary for transcription and translation, which are found in cells in limited amounts. As various genes in a network compete for a common pool of resources, a hidden layer of interactions among genes arises. Such interactions are neglected by standard Hill-function-based models. In this work, we develop a model with the same dimension as standard Hill-function-based models to account for the sharing of limited amounts of RNA polymerase and ribosomes in gene networks. We provide effective interaction graphs to capture the hidden interactions and find that the additional interactions can dramatically change network behavior. In particular, we demonstrate that, as a result of resource limitations, a cascade of activators can behave like an effective repressor or a biphasic system, and that a repression cascade can become bistable.

I. INTRODUCTION

Context dependence, the unintended interactions among genetic circuits and host factors, is a current challenge in the analysis and design of biomolecular networks [1]. Such unintended interactions hinder our ability to predict design outcomes, which often leads to lengthy and ad hoc design processes. Therefore, much research has sought to better understand and mitigate context dependence [1], [2]. In this paper, we are concerned with the context dependence problem arising from the limitations of cellular resources. In particular, we study gene transcription networks, where genes are transcribed by RNA polymerase (RNAP) into mRNA, and mRNA is translated by ribosomes into proteins. Proteins can be transcription factors (TFs) that regulate each other by binding to the promoter site of a gene, which would either activate or repress its ability to recruit RNAP for transcription. The total amount of RNAP and ribosomes is limited and all genes simultaneously compete for these resources [3]. This limitation has been largely neglected so far, due to the small scale and simplicity of circuits considered. In larger circuits, however, the competition for limited resources has been shown to introduce interactions in gene expression levels even in the absence of explicit regulatory links [4].

In this paper, we consider general gene transcription networks and develop a modeling framework to predict the effective interactions arising from limitations in RNAP and ribosomes availability. Related theoretical works have recently appeared that study resource sharing problems in biomolecular networks. De Vos et al. analyze the response of network flux toward changes in total competitors (mRNAs) and common targets (ribosomes) [5]. Yeung et al. illustrate, using tools from dynamical systems, that resource sharing leads to non-minimum phase zeros in the transfer function of a linearized genetic cascade circuit [6]. Gyorgy et al. develop the notion of realizable region for steady state gene expression to account for the limitations in the availability of RNAP and ribosomes [7]. Hamadeh et al. analyze and compare different feedback architectures to mitigate resource competition [8].

Our work focuses on the idea of effective interactions to help illustrate how sharing of RNAP and ribosomes alters the dynamics of a general gene transcription network. For example, when a TF activates the production of protein $x_1$, more RNAP is recruited to produce a larger number of mRNA $m_1$. Increased $m_1$ further increases the demand for ribosomes to produce $x_1$. Both effects decrease the amount of resources available to produce other protein species (for example, protein $x_2$) in the network. This waterbed effect creates an effective inhibition of protein $x_2$ and can be incorporated into an interaction graph, which is commonly used to describe transcriptional regulation interactions (activation/repression) among TFs.

Here, we propose a general model based on deterministic reaction rate equations and ODEs in a resource limiting environment. The model is able to account for resource limitations while maintaining the same dimension as the standard Hill-function-based models [2], [9]. Employing this model, we provide simple rules to identify the hidden interactions due to resource limitations, and the resulting effective interactions in the network. We apply our results to two-stage activation and repression cascades and illustrate how the hidden interactions can dramatically change system’s behavior. In an activation cascade, resource sharing can completely invert the desired steady state I/O response or lead to biphasic behavior, while in a two-stage repression cascade, resource limitations can lead to bistability.

This paper is organized as follows. In Section II, we give a motivating example. In Section III, we introduce our general modeling framework. In Section IV, we illustrate the effective interaction graph of a general gene network. The activation and repression cascade examples are in Section V. We discuss the limitations of our approach and provide directions for future investigation in Section VI.

II. A MOTIVATING EXAMPLE

Cascade circuits are one of the most common network motifs in both natural and synthetic gene networks due to their ability to amplify signals and achieve “switch-like” behavior [9]. In Fig. 1, we consider a simple two-stage activation cascade composed of gene 1 and gene 2. Protein $u$ is the input TF that binds with promoter $p_1$ to activate the production of protein $x_1$. Protein $x_1$ is an activator for the output protein ($x_2$). The structure of this motif can be represented by the interaction graph as $u \rightarrow x_1 \rightarrow x_2$. The
dynamics of binding reactions and mRNA dynamics are often neglected because they are much faster than protein dynamics [2, 9]. We use \( u, x_1 \) and \( x_2 \) to represent the concentration of \( u, x_1 \) and \( x_2 \), respectively. In a standard model, we use Hill functions to describe gene activation, thus we have:

\[
\dot{x}_1 = \frac{\alpha_0 + \alpha \left( \frac{x_1}{k_1} \right)^n}{1 + \left( \frac{x_1}{k_1} \right)^n} - \gamma_1 x_1, \quad \dot{x}_2 = \frac{\beta_0 + \beta \left( \frac{x_2}{k_2} \right)^n}{1 + \left( \frac{x_2}{k_2} \right)^n} - \gamma_2 x_2,
\]

where \( \alpha_0 \) and \( \beta_0 \) are the basal production rate constants; \( \alpha \) and \( \beta \) are the production rate constants with activation; \( k_1 \) and \( k_2 \) are the dissociation constants of activators \( u \) and \( x_1 \) binding with their respective promoters, \( \gamma_1 \) and \( \gamma_2 \) are the dilution/degradation rate of the proteins, and \( n \) and \( m \) are the cooperativity coefficients. Solving for the steady state of equation (1) gives a monotonically increasing I/O response (Fig. 2A).

To examine whether the standard model in (1) is a good representation of system response under resource limitations, we simulate the system with a mechanistic model that explicitly accounts for the usage of RNAP and ribosomes, and for their conservation law (listed in Section III). Surprisingly, simulation of this mechanistic model reveals that the steady state I/O response can be biphasic (Fig. 2B).

With reference to Fig. 2A, decrease of steady state expression of \( x_2 \) with \( u \) at high input level in Fig. 2B can be explained by the following resource sharing mechanism. When promoter \( p_1 \) and mRNA \( m_1 \) have much stronger ability to sequester resources than promoter \( p_2 \) and mRNA \( m_2 \), as we increase \( u \), the production of protein \( x_1 \) sequesters resources from the production of protein \( x_2 \), decreasing the amount of free resources available to produce \( x_2 \). When this effective repression is stronger than the activation \( x_1 \rightarrow x_2 \), \( x_2 \) decreases with \( u \).

This paper is aimed to obtain an explicit model, with the same dimension as the standard model in (1), that predicts such effective interactions due to resource limitations.

III. GENERAL MODELING FRAMEWORK

A. Gene Expression in a Transcriptional Component

We consider a transcriptional component as a node in the gene network [10]. A transcriptional component takes a number of TFs to bind with its gene promoter \( p_i \) and triggers a series of chemical reactions to produce a TF \( x_i \) as output. The input TFs can either activate or repress the expression of gene \( i \) by changing the binding strength of \( p_i \) with RNAP. Since most gene promoters take at most two input TFs [2, 9], we consider a node \( i \) taking two input TFs \( (x_1 \) and \( x_2 \)) that form complexes with \( p_i \). The reactions are:

\[
p_i + n_1 \cdot x_1 \rightleftharpoons c_i^1, \quad p_i + n_2 \cdot x_2 \rightleftharpoons c_i^2 \equiv c_i^2, \quad c_i^1 + n_2 \cdot x_2 \rightleftharpoons c_i^{12} \equiv c_i^{12}, \quad c_i^1 + n_2 \cdot x_2 \rightleftharpoons c_i^{12} \equiv c_i^{12},
\]

where \( n_1 \) and \( n_2 \) are the cooperativities of \( x_1 \) and \( x_2 \) binding with \( p_i \), respectively. The promoter \( p_i \) and the promoter/TF complexes \( \{c_i^1, c_i^2, c_i^{12}\} \) recruit free RNAP \( (y) \) to form an open complex for transcription. The reactions are given by:

\[
p_i + y \rightleftharpoons C_i, \quad c_i^j + y \rightleftharpoons C_i^j \quad (j = 1, 2, 12).
\]

These transcriptionally active complexes can then be transcribed into mRNA \( (m_i) \), with reactions given by:

\[
m_i + z \rightleftharpoons M_i, \quad M_i \rightarrow m_i + z + x_i, \quad m_i \rightarrow \emptyset, \quad M_i \rightarrow \omega_i, \quad x_i \rightarrow \gamma_i \rightarrow \emptyset.
\]

Consequently, the concentration of each species (italic) in node \( i \) follows the following ODEs:

\[
\begin{align*}
\dot{c}_i^1 &= k_{c_1}^+ p_i x_1^n - k_{c_1}^- c_i^1 - a_j y c_i^1 + d_j C_i^1 + \alpha_j C_i^1, \\
\dot{c}_i^{12} &= k_{c_1}^+ c_i^1 x_2^n - k_{c_2}^- c_i^{12} + k_{c_12}^+ c_i^1 x_1^n - k_{c_21}^- c_i^{12}, \\
\dot{c}_i^1 &= a_j p_i y - d_i C_i - \omega_0 C_i, \\
\dot{C}_i &= a_j p_i y - d_i C_i - \omega_0 C_i, \\
\dot{m}_i &= \alpha_0 C_i + \alpha_1 C_i^1 + \alpha_2 C_i^2 + \alpha_{12} C_i^{12}, \\
\dot{M}_i &= \kappa^+ m_i z - \kappa^- M_i - \omega_i M_i - \theta_i M_i, \\
\dot{x}_i &= \delta_i M_i - \gamma_i x_i.
\end{align*}
\]
where indices \( j = 1, 2 \) and \( k = 1, 2, 12 \). Since DNA concentration is conserved [9], we have
\[
p_{i,T} = p_i + C_i + \sum_{j=1,2,12} (c_j^1 + c_j^2),
\]
where \( p_{i,T} \) is the total concentration of gene \( i \). Given that the binding reactions and mRNA dynamics are much faster than protein production and degradation [9], we can set (2) to (7) to quasi-steady state (QSS) to simplify our analysis. We first obtain the QSS concentration of complexes formed with \( p_i \):
\[
c_i^1 = \frac{p_i x^{n_1}_i}{k_i^1}, \quad c_i^2 = \frac{p_i x^{n_2}_i}{k_i^2}, \quad c_i^1 = \frac{p_i x^{n_1}_i x^{n_2}_i}{k_i^1 k_i^2}, \quad c_i^2 = \frac{p_i x^{n_1}_i x^{n_2}_i}{k_i^1 k_i^2}, \quad C_i = \frac{p_i y}{K_i^j}, \quad C_j^i = \frac{c_j y}{K_i^j} (j = 1, 2, 12),
\]
where dissociation constants are defined as:
\[
K_i^j = \frac{a_j^j + \alpha_j}{\delta_i^j}, \quad K_i^j = \frac{a_j^j + \alpha_j}{\delta_i^j}, \quad k_i^j = \frac{k_j}{k_j} (j = 1, 2, 12).
\]
Here, \( K_i^j \) is the basal dissociation constant of promoter \( p_i \) with RNAP \( y \), \( K_i^j \) is the dissociation constant of promoter/TF complex \( c_j^1 \) with \( y \), and \( K_i^j \) is the dissociation constant of TF \( x_i \) binding with \( p_i \). A smaller dissociation constant indicates stronger binding. When node 1 takes only one input, for simplicity, we write \( K_i \) for \( K_i^1 \) and \( k_i \) for \( k_i^1 \). To obtain the QSS concentration of mRNA complexes, we further assume that the transcription rates are independent of how transcriptions are initiated and thus \( \alpha_0 = \alpha_1 = \cdots = \alpha_i \). We can then substitute (10) into the QSS of ODEs (6) and (7) and obtain
\[
M_i = \frac{\alpha_i z}{\delta_i \kappa_i} (C_i + \sum_j C_j^i) = \frac{\alpha_i p_i}{\delta_i} \frac{z}{\kappa_i} F_i(u_i), \tag{11}
\]
where vector \( u_i = [x_1, x_2]^T \) and index \( j = 1, 2, 12 \), \( \kappa_i = (\kappa^- + \theta_i + \omega_i)/\kappa^+ \) is the dissociation constant of \( u_i \) binding with ribosomes \( z \). A smaller \( \kappa_i \) indicates stronger RBS strength. \( F_i(u_i) : \mathbb{R}^2 \to \mathbb{R} \) is the Hill function derived by substituting (10) into the conservation law in (9) and solving for \( c_i^1 + c_i^2 + c_i^3 + c_i^4 \). Assuming that the free amount of RNAP and ribosomes are limited, in particular,
\[
y \ll K_i, K_i^j \quad \text{and} \quad z \ll \kappa_i, \tag{12}
\]
\( F_i(u_i) \) can be written as:
\[
F_i(u_i) = 1 + \frac{a_i^1 x_1^{n_1} + a_i^2 x_2^{n_2} + a_i^3 x_1^{n_1} x_2^{n_2}}{1 + b_i^1 x_1^{n_1} + b_i^2 x_2^{n_2} + b_i^3 x_1^{n_1} x_2^{n_2}}, \tag{13}
\]
where
\[
a_i^1 = \frac{K_i^j}{K_i^1 k_i}, \quad a_i^2 = \frac{K_i^j}{K_i^2 k_i}, \quad a_i^3 = \frac{K_i^j}{K_i^3 k_i} \left( \frac{1}{k_i^1 k_i^2} + \frac{1}{k_i^2 k_i^3} \right), \quad b_i^1 = \frac{1}{k_i}, \quad b_i^2 = \frac{1}{k_i}, \quad b_i^3 = \frac{1}{k_i^1 k_i^2} + \frac{1}{k_i^2 k_i^3}. \tag{14}
\]
Situations in (12), where resources are limited, are described in the Appendix. Finally, we combine equation (11) and (8) to obtain the dynamics of \( x_i \):
\[
\dot{x}_i = \frac{\alpha_i \theta_i p_i}{\delta_i} \frac{y}{K_i} \frac{z}{\kappa_i} F_i(u_i) - \gamma_i \cdot x_i. \tag{15}
\]
Since \( y \) and \( z \) are shared among all nodes in the network, their free concentrations \( y, z \) need to be determined from the network context. This is the aim of the next subsection.

B. Resource Sharing in Gene Networks

A gene network \( \mathcal{N} \) is composed of \( N \) nodes and \( M \) external TF inputs \((v_1, \ldots, v_M)\). The concentration of the external inputs can be represented by \( \mathbf{v} = [v_1, \ldots, v_M]^T \) and the state of the network is represented by the concentrations of output proteins of each node \( \mathbf{x} = [x_1, \ldots, x_N]^T \). The set of all TFs in the network is \( \mathcal{X} = \{x_1, \ldots, x_N, v_1, \ldots, v_M\} \), and we use \( \mathbf{z} = [x^T, v^T]^T \) to represent the vector of their concentrations. Nodes can be connected by transcriptional regulation interactions where protein \( x_i \) can either activate or repress the production of \( x_j \) by binding to its promoter. We call \( x_i \) as a target of \( x_j \) and \( x_j \) as a parent of \( x_i \). We denote by \( U_i \subseteq \mathcal{X} \) the set of all parents of \( x_i \). Their concentrations are given by a vector \( u_i = Q_i \cdot \mathbf{z} \), where elements in \( Q_i \) are defined as:
\[
q_{jk} = \begin{cases} 1, & \text{if } \xi_k \text{ is the } j\text{th input to node } i, \\ 0, & \text{otherwise}. \end{cases}
\tag{16}
\]
Fig. 3 illustrates an example gene network. To determine the effect of RNAP and ribosome limitations on the gene network, we account for the fact that the total amount of resources available to network \( \mathcal{N} \) is constant [3]:
\[
y_T = y + \sum_{i=1}^N y_i, \quad z_T = z + \sum_{i=1}^N z_i, \tag{17}
\]
where \( y_T \) and \( z_T \) represent the total amount of RNAP and ribosomes, respectively. We let \( y_i \) and \( z_i \) denote the RNAP and ribosomes bound to (used by) node \( i \), thus \( y_i = C_i^1 + C_i^2 + C_i^3 + C_i^4 \) and \( z_i = M_i \). According to (11), we have:
\[
y_i = p_i \cdot y \frac{y}{K_i} F_i(u_i), \quad z_i = \frac{\alpha_i p_i}{\delta_i} \frac{y}{K_i} \frac{z}{\kappa_i} F_i(u_i). \tag{18}
\]
Combining equation (17) and (18), we obtain:
\[
y = \frac{y_T}{1 + \sum_{i=1}^N \left[ \frac{p_i}{K_i} F_i(u_i) \right]}, \quad z = \frac{z_T}{1 + y \sum_{i=1}^N \left[ \frac{\alpha_i p_i}{\delta_i K_i} F_i(u_i) \right]}. \tag{19}
\]
Hence,
\[ y \cdot z = \frac{y_T \cdot z_T}{1 + \sum_{i=1}^{N} \frac{p_i \cdot T}{K_i} \cdot (1 + \frac{\alpha_i}{\kappa_i \delta_i} \cdot y_T) \cdot F_i(u_i)} \cdot \frac{y_T \cdot z_T}{1 + \sum_{k=1}^{N} J_k F_k(u_k)}. \tag{19} \]

Substituting (19) into (15), the dynamics of \( x_i \) are given by:
\[ \dot{x}_i = \frac{T_i F_i(u_i)}{1 + \sum_{k=1}^{N} J_k F_k(u_k)} - \gamma_i x_i, \tag{20} \]

where \( J_i \) and \( T_i \) are lumped parameters defined as:
\[ J_i := \frac{p_i \cdot T}{K_i} \cdot (1 + \frac{\alpha_i}{\kappa_i \delta_i} \cdot y_T), \quad T_i := y_T \cdot z_T \cdot p_i \cdot T \cdot \frac{\theta_i \alpha_i}{K_i \kappa_i \delta_i}. \tag{21} \]

\( F_i(u_i) \) is the only element in equation (20) that reflects transcriptional regulations on node \( i \). According to equation (15), the form of \( F_i(u_i) \) is the same as those of the standard Hill functions described in [2] and [9]. Note that \( F_i(u_i) \equiv 1 \) when \( u_i = 0 \), hence, according to equation (15), \( T_i \) represents the “baseline” gene expression of node \( i \), because \( T_i \) quantifies production rate of \( x_i \) when \( u_i = 0, y = y_T \) and \( z = z_T \).

C. \( J_i \) as a Measure of Resource Usage by Node \( i \)

\( J_i \) is a constant for node \( i \) that defines its “baseline” resource usage when \( u_i = 0 \). We take \( J_i \) as a measure of resource usage by node \( i \) because the expression in equation (19) implies the “conservation law” for \( y \cdot z \):
\[ y_T \cdot z_T = \sum_{i=1}^{N} (J_i \cdot F_i(u_i) \cdot y \cdot z) \]

Furthermore, the only difference between our modified model in equation (20) and the standard no-resource-sharing model in [2] and [9] is the denominator term \( D = 1 + \sum_{k=1}^{N} J_k F_k(u_k) \). The following claim shows that when resources used by every node in \( N \) are negligible, the resource usage measure \( J_i \ll 1 \).

**Claim 1:** For every \( u_i \), if \( y_i \ll y \) and \( z_i \ll z \) for all \( i = 1, \cdots, N \), then \( J_i \ll 1 \) for all \( i = 1, \cdots, N \).

**Proof:** Using equation (18), \( y_i \ll y \) for every \( u_i \) is equivalent to \( p_i \cdot T(K_i) / K_i \ll 1 \) for every \( u_i \). Thus, we must have \( p_i \cdot T(K_i) / K_i \ll 1 \). Similarly, \( z_i \ll z \) for every \( u_i \) requires \( \frac{\alpha_i}{\kappa_i \delta_i} y_T < 1 \). Since \( y_i \ll y \) for all \( i \), \( y \approx y_T \).

Therefore, \( \frac{\alpha_i}{\kappa_i \delta_i} y_T \ll 1 \) and \( J_i \ll 1 \) for all \( i \).

This claim shows that when resource usage is negligible in the network, \( 0 < J_i \ll 1 \) and the modified model reduces back to the standard model in [2] and [9]:
\[ \dot{x}_i = T_i F_i(u_i) - \gamma_i x_i, \tag{23} \]

which has the same form as equation (1).

Equation (21) indicates that a node \( i \) is a strong resource sink when \( u_i = 0 \) if its (i) copy number is large; (ii) basal RNAP sequestering capability is strong (small \( K_i \)); (iii) transcription rate constant is large; (iv) ribosome sequestering capability is strong (small \( \kappa_i \)); (v) mRNA degradation rate is low and (vi) the total amount of RNAP is large. Conditions (i) and (ii) are associated with the \( p_i \cdot T(K_i) / \kappa_i \) term in equation (21), and describe the node’s capability to sequester RNAP. Conditions (iii) to (vi) are the contributions from the \( \alpha_i y_T / (\kappa_i \delta_i) \) term and characterize the node’s capability to sequester free ribosomes.

IV. EFFECTIVE INTERACTIONS DUE TO RESOURCE LIMITATIONS

Directed edges, such as those in Fig. 3, have been used to represent transcriptional regulation interactions, where one TF binds with the promoters of its targets to regulate the target’s production [9]. Here, we mathematically define the standard to draw interaction graphs and illustrate that resource limitations lead to effective interactions in gene networks that do not rely on TF regulation.

**Definition 1:** Let the dynamics of \( x_i \) be given by \( \dot{x}_i = G_i(\xi) - \gamma_i \cdot x_i \). We draw the interaction graph from \( \xi \) to \( x_i \) based on the following rules:

- If \( \frac{\partial G_i}{\partial \xi} = 0 \) for all \( \xi \in \mathbb{R}^+ \), then there is no interaction from \( \xi \) to \( x_i \);
- If \( \frac{\partial G_i}{\partial \xi} \geq 0 \) for all \( \xi \in \mathbb{R}^+ \) and \( \frac{\partial G_i}{\partial \xi} \neq 0 \) for some \( \xi \), then \( \xi \) activates \( x_i \) and we draw \( \xi \rightarrow x_i \);
- If \( \frac{\partial G_i}{\partial \xi} \leq 0 \) for all \( \xi \in \mathbb{R}^+ \) and \( \frac{\partial G_i}{\partial \xi} \neq 0 \) for some \( \xi \), then \( \xi \) represses \( x_i \) and we draw \( \xi \leftarrow x_i \);
- If \( \frac{\partial G_i}{\partial \xi} \geq 0 \) for some \( \xi \in \mathbb{R}^+ \) and \( \frac{\partial G_i}{\partial \xi} < 0 \) for some other \( \xi \), then the regulation of \( \xi \) on \( x_i \) is undetermined and we draw \( \xi \leftrightarrow x_i \);

Based on Definition 1, for the standard model in equation (23), \( G_i(\xi) = T_i F_i(Q_i, \xi) = T_i F_i(u_i) \), and therefore there is a link from \( \xi \) to \( x_i \) if and only if \( \xi \in U_i \). In our modified model in equation (20), instead we have
\[ G_i(\xi) = \frac{T_i F_i(Q_i, \xi)}{1 + \sum_{k=1}^{N} J_k F_k(Q_k, \xi)} = \frac{T_i F_i(u_i)}{1 + \sum_{k=1}^{N} J_k F_k(u_k)}, \]

which implies that the dynamics of \( x_i \) may be influenced by TFs that do not belong to its parents \( U_i \).

In what follows, we discuss the effective interactions from \( \xi \in \chi \) to protein \( x_i \) when (i) \( x_i \) is the only target of \( \xi \), (ii) \( x_i \) is one of the multiple targets of \( \xi \), and (iii) \( x_i \) is not a target of \( \xi \). We do not require \( x_i \neq \xi \) and assume that a TF cannot be both an activator and a repressor. When \( x_i \) is the only target of \( \xi \), the following claim shows that resource limitations do not alter the activation/repression of \( x_i \) by \( \xi \) in the interaction graph.

**Claim 2:** If \( \xi \in U_i \) and \( \xi \notin U_q \) for all \( (q \neq i) \). Then we have \( \text{sign} \left[ \frac{\partial G_i(\xi)}{\partial \xi} \right] = \text{sign} \left[ \frac{\partial F_i(Q_i, \xi)}{\partial \xi} \right] \).

**Proof:** According to equation (20),
\[
\frac{\partial G_i(\xi)}{\partial \xi} = \frac{\partial G_i}{\partial F_i} \cdot \frac{\partial F_i(Q_i, \xi)}{\partial \xi} \Rightarrow \text{sign} \left( \frac{\partial G_i}{\partial \xi} \right) = \text{sign} \left( \frac{\partial F_i}{\partial \xi} \right). 
\]
### Remark 1: In the case where $\xi_j \in \mathcal{U}_k, \ldots, \mathcal{U}_k \ (k \geq 2)$, the effective interactions from $\xi_j$ to its targets are undetermined. For example, if $\xi_j$ represses $x_1$ and $x_2$ simultaneously, the effective interaction from $\xi_j$ to $x_1$ is given by:

$$\frac{\partial G_1}{\partial \xi_j} = \frac{\partial G_1}{\partial \xi_j} + \frac{\partial F_1(Q_1 \xi_j)}{\partial \xi_j} + \frac{\partial F_2(Q_2 \xi_j)}{\partial \xi_j}.$$

As $\text{sign}(\partial G_1/\partial \xi_j)$ cannot be determined, the effective interaction from $\xi_j$ to $x_1$ is undetermined.

When $\xi_j$ is not a parent of $x_i$, the following claim shows $\xi_j$ is an effective repressor for $x_i$ if $\xi_j$ is an activator. Conversely, $\xi_j$ is an effective activator for $x_i$ if $\xi_j$ is a repressor.

### Claim 3: If $\xi_j \notin \mathcal{U}_i$, then $\frac{\partial G_i}{\partial \xi_j} = -\frac{\partial F_k(Q_k \xi_j)}{\partial \xi_j}$.

**Proof:** Since $\xi_j \notin \mathcal{U}_i$, $G_i \partial F_k < 0$ for all $k$.

$$\frac{\partial G_i}{\partial \xi_j} = \sum_k \frac{\partial G_i}{\partial F_k} \frac{\partial F_k}{\partial \xi_j}.$$

Therefore, $\text{sign}(\partial G_i/\partial \xi_j) = -\text{sign}(\partial F_k/\partial \xi_j)$.

The effective interactions for the above three cases are summarized in Table I, with illustrative examples given in each case. For any index $i, j \in \{1, \ldots, N\}$, a black solid line from node $j$ to node $i$ represents $\partial F_i(Q_i \xi_j)/\partial \xi_j$, the interaction due to transcriptional regulation, while a red dashed line represents any hidden (additional) interactions arising from $\partial G_i(Q_i \xi_j)/\partial \xi_j$.

### V. APPLICATION TO ACTIVATION AND REPRESSION CASCADES

#### A. Two-stage Activation Cascade

We first revisit the motivating example in Section II. u is the input and $x_1$ and $x_2$ are the two TFs cascaded by transcriptional regulation interactions (Fig. 4A). According to (20), the dynamics of the system can be written as:

$$\begin{align*}
\dot{x}_1 &= \frac{T_1 F_1(u)}{1 + J_1 F_1(u) + J_2 F_2(x_1)} - \gamma_1 x_1, \\
\dot{x}_2 &= \frac{G_1(u,x_1)}{1 + J_1 F_1(u) + J_2 F_2(x_1)} - \gamma_2 x_2.
\end{align*}$$

From Claim 3, since $u$ is an activator, there is a hidden repression from $u$ to $x_2$. Similarly, there is a hidden negative auto-regulation on $x_1$. These hidden interactions are represented by dashed lines in Fig. 4B. From Claim 2, since $u$ and $x_1$ both have only one target, we draw $u \rightarrow x_1$ and $x_1 \rightarrow x_2$ in Fig. 4B. The effective interaction graph of the activation cascade becomes that of an incoherent feed-forward loop (IFFL) [9]. The steady state I/O response of an IFFL can, depending on parameters, be qualitatively characterized by monotonically increasing, monotonically decreasing or biphasic functions [9] [11]. We can predict which of these function classes the steady state I/O response of the activation cascade falls into by linearizing the model in equations (24) and (25), as shown by the following claim.

### Claim 4: Consider a monostable nonlinear time-invariant SISO system: $\tilde{x} = f(\tilde{x}, u), \ y = g(\tilde{x}, u)$, where $f(\tilde{x}, u)$ and $g(\tilde{x}, u)$ are analytic functions with respect to their arguments. Let the linearized system at input $\tilde{u}$ and corresponding locally asymptotically stable equilibrium $\tilde{x}$ be $\tilde{x} = A \tilde{x} + B \tilde{u}, \ \tilde{y} = C \tilde{x} + D \tilde{u}$. Let the steady state output of the nonlinear system be $\tilde{y} = g(\tilde{x}, \tilde{u})$, then $\frac{dy}{du} \bigg|_{\tilde{x}, \tilde{u}} = H = -CA^{-1}B + D$.

**Proof:** In the linearized model, $A = \frac{\partial f}{\partial \tilde{x}} \bigg|_{\tilde{x}, \tilde{u}}, \ B = \frac{\partial f}{\partial u} \bigg|_{\tilde{x}, \tilde{u}}, \ C = \frac{\partial g}{\partial \tilde{x}} \bigg|_{\tilde{x}, \tilde{u}}$ and $D = \frac{\partial g}{\partial u} \bigg|_{\tilde{x}, \tilde{u}}$. Therefore,

$$\frac{dy}{du} \bigg|_{\tilde{x}, \tilde{u}} = \frac{\partial g(\tilde{x}, \tilde{u})}{\partial \tilde{x}} \cdot \frac{d\tilde{x}}{du} + \frac{\partial g(\tilde{x}, \tilde{u})}{\partial u} \cdot \frac{d\tilde{u}}{du}.$$  

Since $\tilde{x}$ satisfies $f(\tilde{x}, \tilde{u}) = 0$, using the implicit function theorem, we have,

$$\frac{d\tilde{x}}{du} = -\left( \frac{\partial f(\tilde{x}, \tilde{u})}{\partial \tilde{x}} \right)^{-1} \cdot \frac{\partial f(\tilde{x}, \tilde{u})}{\partial u} = -A^{-1}B.$$  

Matrix $A$ is invertible because the equilibrium $\tilde{x}$ is asymptotically stable. Combining equations (26) and (27), we obtain $\frac{dy}{du} \bigg|_{\tilde{x}, \tilde{u}} = H = -CA^{-1}B + D$. 

---

**Table I. Effective Interactions with Resource Limitations**

<table>
<thead>
<tr>
<th>Remark 1</th>
<th>Claim 2</th>
<th>Claim 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi_j$ to multiple targets</td>
<td>$\xi_j$ to its only target</td>
<td>$\xi_j$ to nodes that are NOT its target</td>
</tr>
<tr>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
<td><img src="image3.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

**Fig. 4.** In addition to the regulatory transcriptional activations (solid lines) captured by the ideal model in equation (1) (A), resource limitations introduce two hidden repressions (dashed lines) into the system: repression from input $u$ to output $x_2$ and negative auto-regulation of $x_1$. (B). Both nodes take a single input: $U_1 = u$ and $U_2 = x_1$. Thus, equations (14) yield $a_1 = \frac{K_1}{K_1 + x_1}, b_1 = \frac{1}{x_1}, a_2 = \frac{K_2}{K_2 + x_2}$ and $b_2 = \frac{1}{x_2}$, with the remaining parameters in (14) being 0. Denote by $u$ and $m$ the cooperativity coefficients for $u$ and $x_1$ binding with genes 1 and 2, respectively. From (13) we have:

$$F_1(u) = \frac{1 + a_1 u^n}{1 + b_1 u^n}, \ F_2(x_1) = \frac{1 + a_2 x_1^m}{1 + b_2 x_1^m}.$$
Applying Claim 4 to the two-stage activation cascade, we first linearize (24) and (25) at input $\bar{u}$ and equilibrium $\bar{x}$ to obtain:

$$A = \begin{bmatrix} \frac{\partial G_1}{\partial x_1} - \gamma_1 - \gamma_2 \end{bmatrix}, B = \begin{bmatrix} \frac{\partial G_1}{\partial u} \\ \frac{\partial G_2}{\partial u} \end{bmatrix}, C = [0 \ 1], D = 0. \tag{28}$$

Substituting (28) into $H$ from Claim 4, we can numerically find the slope of the steady state response ($\frac{dx_2}{du}$) for all linearization points and parameter conditions that admit an increasing/decreasing steady state response. Fig. 5 shows that a decreasing steady state response occurs when (a) resources are limited (large DNA copy number) and (b) $x_1$ has a stronger resource sequestering capability than $x_2$ (stronger RBS). The numerical result is in agreement with the following analytical results providing sufficient conditions for different steady state I/O responses.

**Claim 5:** If node 1 and 2 have the same DNA copy numbers $p_{1,T} = p_{2,T} = p_T$, and transcription rate constants $\alpha_1 = \alpha_2 = \alpha$, then in a two-stage activation cascade the slope of the steady state I/O response $\frac{dx_2}{du}$ satisfies:

1. $\frac{dx_2}{du} > 0$ for all $u > 0$ if (a) $K_1 \gg p_T$ and (b) $\kappa_1 \cdot \delta_1 \gg \alpha \cdot y_T$;
2. $\frac{dx_2}{du} < 0$ for all $u > 0$ if (a) $p_T \gg K'_2 \gg K_2 \gg K_1 \gg K_1$ and (b) $\alpha \cdot y_T \gg \delta_2 \cdot \kappa_2 \gg \delta_1 \cdot \kappa_1$;
3. $\frac{dx_2}{du} > 0$ when $u \to 0$ and $\frac{dx_2}{du} < 0$ when $u \to \infty$ if (a) $K'_1 \gg p_T \gg K_2 \gg K_1$ and (b) $\kappa_2 \cdot \delta_2 \gg \kappa_1 \cdot \delta_1 \gg \alpha \cdot y_T$.

**Proof:** Using equation (25), the steady state concentration of the output ($x_2$) can be written as:

$$\bar{x}_1(u, \bar{x}_1) = \frac{1}{\gamma_1} \cdot \frac{T_1 F_1(u)}{1 + J_1 F_1(u) + J_2 F_2(\bar{x}_1)}, \tag{29}$$

$$\bar{x}_2(u, \bar{x}_1) = \frac{1}{\gamma_2} \cdot \frac{T_2 F_2(\bar{x}_1)}{1 + J_1 F_1(u) + J_2 F_2(\bar{x}_1)}. \tag{30}$$

From Claim 2, $\bar{x}_1$ increases monotonically with $u$. When conditions in 1) are satisfied, we have $J_1 F_1(u) < \frac{\kappa_1}{\kappa_2} (1 + \alpha y_T) < 1$. Equation (30) becomes $\bar{x}_2(u, \bar{x}_1) = \frac{1}{\gamma_2} \cdot \frac{T_2 F_2(\bar{x}_1)}{1 + J_2 F_2(\bar{x}_1)}$, and therefore $\bar{x}_2$ increases with $u$. When conditions in 2) are satisfied, we have $J_1 F_1(u) \gg J_2 F_2(\bar{x}_1) \gg 1$.

Combining with (29), we have $\bar{x}_1 = T_1 / J_1 = \text{constant}$, and (30) becomes a single variable decreasing function of $u$. To prove 3), note that when $u \to 0$, $F_1(u) \to 1$ and when $u \to \infty$, $F_1(u) \to K_1 / K_1 \gg 1$. The conditions give $J_1 F_1(u) = \frac{1}{\gamma_2} < 1$ when $u \to 0$, and $J_1 F_1(u) \gg J_2 F_2(\bar{x}_2) \gg 1$ when $u \to \infty$. The rest of the proof follows from case 1) and 2).

**B. Two-stage Repression Cascade**

A two-stage repression cascade consists of two repressors: TF $u$ is the repressor for protein $x_1$, and $x_1$ is a repressor for output protein $x_2$ (Fig. 6A). A repressor inhibits the production of its target by binding with its promoter region, thus inhibiting RNAP ($y$) recruitment. A repression cascade is expected to have a unique steady state and a monotone I/O response [9]. Here, we apply our modified model to investigate its behavior under resource limitations. In our model, the inputs to the two nodes are $u_1 = u$ and $u_2 = x_1$, respectively. Using the results in (20), the two-stage repression cascade can be modeled as:

$$\dot{x}_1 = \frac{J_1 F_1(u)}{1 + J_1 F_1(u) + J_2 F_2(x_1)} - \gamma_1 x_1, \tag{31}$$

$$\dot{x}_2 = \frac{G_1(u,x_1)}{1 + J_1 F_1(u) + J_2 F_2(x_1)} - \gamma_2 x_2. \tag{32}$$

For simplicity, we assume that the repressors are not leaky such that when $u$ or $x_1$ are bound to the promoters of their targets, $y$ can not bind with the promoters. Using (14), we have $b_1 = \frac{1}{\kappa_1}$ and $b_2 = \frac{1}{\kappa_2}$ with the remaining parameters being 0. We denote $n$ and $m$ as the cooperativity coefficients of $u$ and $x_1$ binding with $p_1$ and $p_2$. We obtain from (13):

$$F_1(u) = \frac{1}{1 + \frac{1}{\kappa_1} u^n}, \quad F_2(x_1) = \frac{1}{1 + \frac{1}{\kappa_2} x_1^m}. \tag{33}$$

Using Claim 2, we find that resource limitations do not affect $u \neq x_1$ and $x_1 \neq x_2$ in the interaction graph. From Claim 3, we find that there is a hidden activation of $x_2$ by $u$ and a hidden positive auto-regulation on $x_1$. Positive feedback loops like the one in Fig. 6B have been closely related to bistable behaviors theoretically [12], and bimodal reporter gene distributions experimentally [13]. In order to determine whether the repression cascade can display bistability because of this positive auto-regulation, we perform nullcline analysis. The two nullcline equations of the nonlinear system in equations (31) and (32) at equilibrium $\bar{x} = [\bar{x}_1, \bar{x}_2]^T$ and
constant input \( \bar{u} \) (and thus, constant \( F_1(\bar{u}) \)) are given by:

\[
\begin{align*}
\frac{T_1 F_1(\bar{u})}{1 + J_1 F_1(\bar{u}) + J_2 F_2(\bar{x}_1)} - \gamma_1 \bar{x}_1 &= 0, \\
\frac{T_2 F_2(\bar{x}_1)}{1 + J_1 F_1(\bar{u}) + J_2 F_2(\bar{x}_1)} - \gamma_2 \bar{x}_2 &= 0.
\end{align*}
\]

(33) and (34)

Equation (33) is a single variable equation of \( \bar{x}_1 \), and equation (34) defines a unique \( \bar{x}_2 \) for every \( \bar{x}_1 \). Therefore, the number of equilibria of this nonlinear system is solely determined by equation (33) which can be re-written as:

\[
h_1(\bar{x}_1) = \frac{T_1 F_1(\bar{u})}{1 + J_1 F_1(\bar{u}) + J_2 F_2(\bar{x}_1)} = h_2(\bar{x}_1) = \gamma_1 \bar{x}_1.
\]

(35)

Since \( h_1(\bar{x}_1) \) is an increasing Hill function and \( h_2(\bar{x}_1) \) is an increasing linear function, they can have either 1 or 3 intersections when the cooperativity \( m > 1 \). Particularly, when there exists \( \bar{x}_1^1 < \bar{x}_1^2 < \bar{x}_1^3 \) satisfying \( h_1(\bar{x}_1^k) = h_2(\bar{x}_1^k) \)\((k = 1, 2, 3)\), \( \bar{x}_1^1 \) and \( \bar{x}_1^3 \) are locally stable nodes and \( \bar{x}_1^2 \) is a saddle point.

Now we seek to obtain parameter conditions that give rise to a bistable repression cascade. To do this, we utilize the following claim showing that the nonlinear repression cascade is bistable if and only if its linearized system is unstable at some equilibrium.

**Claim 6**: For a given input \( u^* \), let \( x^* \) be one of the corresponding equilibria. The nonlinear system in equation (31) and (32) is bistable if and only if \( -\gamma_2 + \frac{\partial G_1}{\partial x_1} x^* u^* > 0 \) for some \( (x^*, u^*) \).

**Proof**: (sketch) Note that \( \lambda_1 = -\gamma_2 + \frac{\partial G_1}{\partial x_1} x^* u^* \) and \( \lambda_2 = -\gamma_1 < 0 \) are the two eigenvalues of the linearization of nonlinear system in (31) and (32) at \((x^*, u^*)\). The linearized system is unstable if and only if \( \lambda_1 > 0 \).

(\( \Rightarrow \)) When the nonlinear system is bistable at input \( u^* \), according to our nullcline analysis, there are 3 equilibria: 2 stable nodes and a saddle point. Linearizing the system around the saddle point yields an unstable linearized system.

(\( \Leftarrow \)) We let \( H(x_1, u) = G_1(x_1, u) - \gamma_1 x_1 \), at fixed \( u = u^* \), with abuse of notation, we write \( H(x_1) = H(x_1, u^*) \). \( H(x_1) \) is continuously differentiable and solution to \( H(x_1) = 0 \) entirely determines the number of equilibria. When \( \lambda_2(x_1^*, u^*) = H'(x_1^*) > 0 \), since \( H(x_1) = 0 \), by continuity, there exists \( \epsilon > 0 \) such that \( H(x_1 - \epsilon) < 0 \) and \( H(x_1 + \epsilon) > 0 \). Also, when \( x_1 = 0 \), \( H(0) = G_1(0) > 0 \), and when \( x_1 \to \infty, H(x_1) \to -\infty \). According to the intermediate value theorem, there exist a \( x_1^- \) such that \( 0 < x_1^- < x_1 - \epsilon \) and satisfies \( H(x_1^-) = 0 \). Similarly, there exists a \( x_1^+ \) such that \( x_1^+ < x_1 + \epsilon \) and satisfies \( H(x_1^+) = 0 \). Since there are at most three zeros to the equation \( H(x_1) = 0 \), \( H'(x_1^-) \) and \( H'(x_1^+) \) are negative, and thus they are stable.

**Remark 2**: To obtain a bistable cascade, we need

\[
\lambda_1 = -\gamma_2 + \frac{\partial G_1}{\partial x_1} = -\gamma_2 - \frac{T_1 J_2 F_1(u^*) \frac{\partial F_2(x_1^*)}{\partial x_1}}{1 + J_1 F_1(x_1^*) + J_2 F_2(x_1^*)} > 0.
\]

(36)

Partial differentiation of \( \lambda_1 \) with respect to \( J_2 \) shows that \( \lambda_1 \) monotonically increases with \( J_2 \) when \( J_2 F_2(x_1^*) > 1 + J_1 F_1(u^*) \). Therefore, we can observe a bistable repression cascade if we increase the resource sequestering capability of node 2 \((J_2 F_2(x_1^*))\) and decrease that of node 1 \((J_1 F_1(u^*))\). Physically, these conditions increase the amount of resources released by node 2 upon repression from \( x_1 \), which effectively “activates” the production of \( x_1 \), promoting the hidden positive auto-regulation. Simulation results in Fig. 7A using (35) is consistent with these analysis. Full mechanistic model simulation using ODEs and resource conservations in Section III confirms that this deterministic system is bistable in some parameter and input ranges (Fig. 7B). Conversely, from (36), we can remove bistability by adding a sufficiently strong negative auto-regulation to node 1 such that \( \partial G_1/\partial x_1 < \gamma_2 \), which ensures monostability.

Resource-limitation-induced bistability can potentially explain the experimental results in [14] and [15]. Both works observed bimodal distribution of protein concentrations at the output of a repression cascade, which disappears when negative auto-regulation is added to the cascade. However, bimodal distribution can stem from a number of other sources in addition to deterministic bistability, such as transcriptional and translational bursts [16]. Further theoretical and experimental work is required to verify the source of bimodality in these experiments.

VI. DISCUSSION AND CONCLUSION

The proper function of a gene network requires cellular resources (here, we focus on RNAP and ribosomes) for transcription and translation. Inevitably, all nodes in the network are forced to compete for a limited amount of resources, introducing hidden interactions in addition to the transcriptional regulation interactions. In this work, we have developed a general modeling framework to describe the dynamics of gene networks in a resource-limited environment. The model reveals a hidden layer of interactions among nodes in the network, which have been largely neglected so far but will become more relevant when resources are limited. Such hidden interactions can alter the steady state I/O response or stability of a network, as we have demonstrated in the examples of activation and repression cascades.

One important limitation of our model is the oversimplification of cell metabolism by assuming constant total amount of resources. A cell system has a number of additional complications. Firstly, recent evidence suggests that resources are not distributed evenly in cells [17]. How spatial distribution of resources changes our current
results need to be investigated. Secondly, when exogenous circuits are overly activated, living cells tend to reduce the production of ribosomal proteins and produce heat shock proteins [17]. The redistribution of cellular resources under these conditions involves some regulation mechanisms that are still unknown and not accounted for in this resource conservation model. Moreover, although the key limiting factors appeared to be RNAP and ribosome [18], [19], resource sharing occurs at all levels of protein production. For instance, it is well known that RNAP compete for σ factors appeared to be RNAP and ribosome [18], [19], resource sharing occurs at all levels of protein production. The dissociation constant of T7 RNAP binding with promoter is $K = 220$[nM] [21]. T7 RNAP has stronger binding with promoters than other RNAP species [22], therefore, $K \gg 220$[nM]. Furthermore, since $y < y_T \approx 100$[nM] [3], we can assume $y \ll K$. Physically, this corresponds to the fact that promoters are rarely occupied by RNAP, which is common in experiments. For instance, Chrchward et al. find that DNA template is in excess of free RNAP in constitutively expressing lac genes [19]. The free amount of ribosome in E. coli is estimated to be $z < z_T \approx 1000$[nM] [3] at low growth rate of 1 doubling/hr, and a typical value of RBS dissociation constant is $\kappa \approx 5000$[nM] [23], which suggests $z \ll \kappa$. These assumptions are closer to reality when the network is larger in scale, and thus resources become more scarce.

ACKNOWLEDGMENT

We thank Eduardo D. Sontag, Abdullah Hamadeh, Hsin-Ho Huang and Andras Gyorgy for helpful discussions and suggestions.

REFERENCES


