2021 American Control Conference (ACC)
New Orleans, USA, May 25-28, 2021

Mediating Ribosomal Competition by Splitting Pools

Jared Miller, M. Ali Al-Radhawi, and Eduardo D. Sontag

Abstract— **Synthetic biology constructs often rely upon the introduction of "circuit" genes into host cells, in order to express novel proteins and thus endow the host with a desired behavior. The expression of these new genes "consumes" existing resources in the cell, such as ATP, RNA polymerase, amino acids, and ribosomes. Ribosomal competition among strands of mRNA may be described by a system of nonlinear ODEs called the Ribosomal Flow Model (RFM). The competition for resources between host and circuit genes can be ameliorated by splitting the ribosome pool by use of orthogonal ribosomes, where the circuit genes are exclusively translated by mutated ribosomes. In this work, the RFM system is extended to include orthogonal ribosome competition. This Orthogonal Ribosomal Flow Model (ORFM) is proven to be stable through the use of Robust Lyapunov Functions. The optimization problem of maximizing the weighted protein translation rate by adjusting allocation of ribosomal species is formulated.**

I. Introduction

The process of protein expression is mediated by ribosomes. After a gene in DNA has been transcribed into messenger RNA (mRNA), a ribosome binds to the mRNA and begins protein translation. The mRNA is divided into a set of 3-nucleotide segments called codons, and each codon corresponds to an amino acid or a stop instruction. The ribosome attracts a tRNA carrying an amino acid which matches the currently read codon, and appends it to a growing polypeptide chain. Once the ribosome hits a stop codon, it falls off the mRNA and releases the amino acid chain as a polypeptide, which is subsequently post-translationally processed into a final protein product. The mRNAs remain in the cell until they are degraded or destroyed (such as by small interfering RNA). Intact mRNAs continually attract ribosomes and produce protein, and all mRNAs in the cell compete for a finite pool of resources which includes ribosomes. A substantial literature exists on the problem of gene expression burden including ribosome competition; see for example the many references reviewed in [1]. When new (circuit) genes are implanted into a cell, the circuit genes compete for ribosomes with the original (host) genes. One way to decouple host and circuit genes is to split the pool of resources via orthogonal ribosomes [2], [3], [4]. Orthogonal ribosomes are mutated ribosomes which only translate specifically modified circuit genes. One such construction replaces the 16S-rRNA (component of small ribosomal subunit) in *E. coli* with a synthetic version. The circuit's mRNAs binding sites are designed so that only mutated ribosomes will translate circuit mRNAs; host ribosomes are not attracted to the circuit mRNAs [5]. Orthogonal ribosomes may be able to decrease competition and increase protein throughput by splitting the pool.

We present the Orthogonal Ribosomal Flow Model (ORFM) extending existing RFM network models with ribosome competition. Global asymptotic stability of the ORFM is certified by means of Robust Lyapunov Functions. A bisection algorithm is detailed to compute the unique ORFM steady state. The protein throughput of the system can be changed by adjusting the production rate of ribosomal species.

The structure of the paper is as follows: Section II reviews existing models for protein translation and ribosome competition. Section III introduces the ORFM. Section IV presents an optimization problem to maximize protein throughput. Section V adds a feedback controller to regulate production of ribosomes. Section VI concludes the paper. The stability of the ORFM is proved in the Appendix.

II. Ribosomal Flow Models (RFMs)

A. Single-Strand Translation Models

The RFM [6] is a deterministic mean-field approximation to lattice models of steady-state ribosome distribution on mRNAs [7], [8]. Each codon on the length-n mRNA has a normalized ribosomal density $x_j(t) \in [0,1]$ for $j = 1, \ldots, n$, which one may think of as the probability that a ribosome is present on codon j at time t. Transition rates between codons are denoted here as λ_j. The initiation rate λ_0 is the rate at which the mRNA attracts the ribosome to begin translation, while the $\lambda_{j<n}$ are elongation rates and represent the amount of time required for the ribosome to attract the codon's respective tRNA. Finally, λ_n is the rate at which ribosomes separate from the mRNA and release the completed polypeptide chain. The quantity $y = \lambda_n x_n$ is the rate (protein/time) at which protein is produced by the mRNA. An RFMIO (input/output) is a tridiagonal polynomial single-input single-output system

$$\begin{aligned} \dot{x}_1 &= \lambda_0 u(1-x_1) - \lambda_1 x_1(1-x_2) \\ \dot{x}_j &= \lambda_{j-1} x_{j-1}(1-x_j) - \lambda_j x_j(1-x_{j+1}), 1 < j < n \\ \dot{x}_n &= \lambda_{n-1} x_{n-1}(1-x_n) - \lambda_n x_n \\ y &= \lambda_n x_n. \end{aligned} \tag{1}$$

The output y is the translation rate, and the input u is the rate at which new ribosomes become available. The probability that codon j is empty is x_j, and the probability that codon $j+1$ is empty is $1 - x_{j+1}$. The likelihood that ribosomes flow from codons j to $j+1$ is proportional to $x_j(1-x_{j+1})$, which is the density relation times the elongation rate in (1).

The authors are with the Department of Electrical & Computer Engineering, Northeastern University, Boston, MA. E.D. Sontag, is also with the Department of Bioengineering, Northeastern University, and the Laboratory of Systems Pharmacology, Harvard Medical School, Boston, MA. Emails: {miller.jare,malirdwi,e.sontag}@northeastern.edu.

2021 American Control Conference (ACC) | 978-1-6654-4197-1/21/$31.00 ©2021 IEEE | DOI: 10.23919/ACC50511.2021.9483415

Authorized licensed use limited to: Northeastern University. Downloaded on February 07,2022 at 14:13:37 UTC from IEEE Xplore. Restrictions apply.

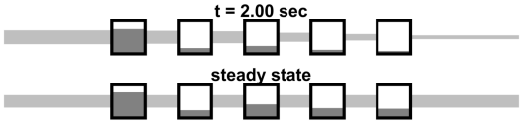


Fig. 1: A pictorial representation of the inflow, outflow, and densities of a 5-codon RFMIO at times $t = 2$ and ∞.

For a constant input $u > 0$, there is a unique steady-state in $[0, 1]^n$ for system (1), which we denote as e. Figure 1 shows the steady state of an 5-codon RFM. Each codon is a black box, and x_j is the filled proportion of each codon. Ribosomes flow from left to right, and the bars between codons show the flux rates $\lambda_j x_j(1-x_{j+1})$ (which must be all equal, as is clear from the equations). The steady state codon occupancies e_j and output $y = \lambda_n e_n$ may be computed by solving a finite continued fraction, which results in a polynomial equation of degree $(n + 1)/2$ [6]. A spectral formulation for the constant u was presented by Poker *et. al.*, and is reviewed in Algorithm 1. The operator **tridiag**(α, β) for $\alpha \in \mathbb{R}^b, \beta \in \mathbb{R}^{b-1}$ produces a symmetric $b \times b$ tridiagonal matrix with main diagonal α and 1-off-diagonals β. For a square symmetric matrix M, $\sigma_{max}(M)$ and $v_{max}(M)$ are its largest eigenvalue and respective eigenvectors. The vector ζ is the Perron-Frobenius eigenvector of the Jacobi matrix formed by the rates λ (dominant eigenvector with nonnegative entries).

Algorithm 1: RFMIO Steady State (from [9])

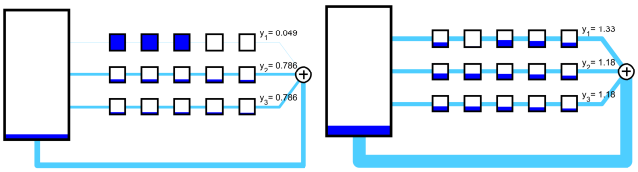
input : Rates λ, Constant Input u
output: Codon Steady States e_j
$\mu_j = 1/\sqrt{\lambda_j}$ for $j = 0, \ldots, n$
$J =$ **tridiag**$(\mathbf{0}_{n+2}, \mu)$
$\sigma = \sigma_{max}(J),\ \zeta = v_{max}(J)$
$e_j = \frac{\mu_j}{\sigma}\frac{\zeta_{j+2}}{\zeta_{j+1}}$ for $j = 1 \ldots n$

The RFM system is a monotone control system. The set of admissible controls $u \in \mathcal{U}$ is the set of bounded and measurable functions $u \in \mathbb{R}_+ \to \mathbb{R}_+$. The RFM is state and output-controllable, and desired translation rates and patterns can be achieved by proper choice of u and λ [10].

B. Ribosomal Competition

In real genetic systems, all mRNA's in the cell compete for a finite (and possibly time-varying) number of ribosomes. Particularly slow transition rates λ on codons can lead to strands of mRNA hoarding ribosomes, reducing the availability of ribosomes in the pool for all other mRNA's and leading to a globally depressed translation rate.

Raveh *et. al.* introduced a ribosomal flow model network with a pool (RFMNP) to abstractly describe the impact of ribosomal competition with multiple strands of mRNA [11]. Each of the s strands of mRNA is modeled by a RFMIO (x_j^i for codon j of mRNA i), and all RFMIO are connected to a common pool z. The total number of ribosomes in the system $N = z(t) + \sum_{i,j} x_j^i(t)$ is conserved. The input u of each mRNA is an increasing saturation function $G^i(z)$ (commonly



(a) Strand 1 is slow (b) Strand 1 is fast

Fig. 2: RFM competition at steady state

z or $\tanh(z)$), which describes the likelihood that a ribosome from the pool will attach itself to strand i. The translation rate $y^i = \lambda_n^i x_n^i$ is the output, and ribosomes leaving x_i return to the pool. The pool dynamics are therefore

$$\dot{z} = \textstyle\sum_i \lambda_n^i x_n^i - \sum_i \lambda_0^i G^i(z)(1 - x_1^i). \tag{2}$$

The RFMNP is a closed loop system. The number of ribosomes N defines a stoichiometric class, and RFMNP has a globally asymptotic equilbrium point with all $e_j^i \in (0, 1)$ and $z \in (0, N)$ for each N. Stability can be proven by contraction, where the weighted L_1 norm between trajectories is non-expanding over time [11].

Competition effects can be observed by perturbing parameters λ_j^i and analyzing resultant steady states. If λ_j^i changes for a specific strand of mRNA x^i, then the protein translation rate y^i will change in that same direction (increasing λ_j^i will increase y_j). The translation rate of all other mRNA in the network will change uniformly: an increase in λ_j^i will raise y^i and will either raise or lower all $y^{k \neq i}$. Further discussion on competition effects can be found in Section 4.2 of [11].

Fig. 2 shows two examples of competition effects in RFMNPs where each of the three strands of mRNA have five codons each. When all strands have transition rates of $\lambda_j^i = 5$ and saturation functions $G^i(z) = \tanh(z)$, each mRNA has a protein translation rate of $y_{ss}^{1:3} = 1.15$ with $N = 5$ ribosomes. In Fig. 2a, ribosomes cannot easily pass from codons 4 to 5 because $\lambda_3^1 = 0.05$. This bottleneck drops the translation efficiency of strand 1 to $y_{ss}^1 = 0.049$ and the others to $y_{ss}^{2:3} = 0.789$. Fig. 2b modifies strand 1 to $\lambda_2^1 = 40$. Ribosomes quickly exit strand 1 and enter the pool again for use in further translation. This improves the translation efficiency of all strands, raising $y_{ss}^1 = 1.33$ and $y_{ss}^{2:3} = 1.18$.

III. Orthogonal Ribosomal Flow Model

Orthogonal ribosomes can be added to the RFM scheme (ORFM) by splitting the ribosome pool z. Code for ORFM simulation and visualization is publicly available online at `https://gitlab.com/jarmill/Ribosomes`.

A. ORFM Formulation

The proposed model of ribosome translation has M species of ribosomes. Each ribosome species has a pool z_p for $p = 1, \ldots, M$ of available ribosomes. Strands of mRNA that use ribosome type p form an RFMNP with pool z_p.

Ribosomes of type p are formed by the combination of a 16S-rRNA of type p and the remaining ribosome components. The protein backbone and large ribosomal subunit are treated as an 'Empty' ribosome E, which has no translation capacity on its own. Empty ribosomes bind with rRNA at the rate

Authorized licensed use limited to: Northeastern University. Downloaded on February 07,2022 at 14:13:37 UTC from IEEE Xplore. Restrictions apply.

Algorithm 2: ORFM Steady State

input : $\lambda^{pi},\ N,\ G^{pi},\ K_p,\ \epsilon$
output: Steady States z_E, z_p, e_j^{pi}
$\bar{N}_{min} = 0,\ \bar{N}_{max} = N$
repeat
  $\bar{N}_{mid} = (\bar{N}_{max} + \bar{N}_{min})/2$
  $z_E = \frac{\bar{N}_{mid}}{1+\sum_p K_p} \qquad z_p = \frac{K_p \bar{N}_{mid}}{1+\sum_p K_p}$
  $e^{pi} = \text{RFMIO SS}(\lambda^{pi}, G^{pi}(z_p))$ (Alg. 1 $\forall p, i$)
  $N_{curr} = \bar{N}_{mid} + \sum_{p,i,j} e_j^{pi}$
  if $N_{curr} \leq N$ **then**
    $\bar{N}_{min} \to \bar{N}_{mid}$
  else
    $\bar{N}_{max} \to \bar{N}_{mid}$
  end
until $|N_{curr} - N| \leq \epsilon$;

k_p^+, and the ribosomal complex dissociates at the rate k_p^-, assuming rRNA's abundance. These kinetics are inspired by Darlington *et. al.*'s mass-action model of protein translation and cell metabolism [12]. The M pools of translating ribosomes z_p are each connected to the pool of empty ribosomes z_E. Each of the s_p strands of mRNA that obtain ribosomes from pool z_p will have a corresponding length n^{pi}. Copies of mRNA with identical initial conditions are represented as a single strand with multiplicity m^{pi}. The codon at location $j \in 0,\ldots,n^{pi}-1$ of RNA strand $i \in 1,\ldots,s_p$ coming from pool p is x_j^{pi}. The constants of this system are the N ribosomes, binding rates k_p^+ and k_p^-, and transition rates λ_j^{pi}. The total number of ribosomes N is a conserved quantity

$$N = z_E + \sum_{p=1}^{M} z_p + \sum_{p=1}^{M} \sum_{i=1}^{s_p} \left(m^{pi} \sum_{j=1}^{n^{pi}} x_j^{pi} \right). \quad (3)$$

The translation dynamics are

$$\dot{z}_E = \sum_p k_p^- z_p - \sum_p k_p^+ z_E \quad (4a)$$

$$\dot{z}_p = k_p^+ z_E - k_p^- z_p + \sum_i m^{pi} \lambda_{n^{pi}}^{pi} x_{n^{pi}}^{pi} - \sum_i m^{pi} \lambda_0^{pi} (1 - x_1^{pi}) G^{pi}(z_p) \quad (4b)$$

$$\dot{x}_1^{pi} = \lambda_0^{pi}(1 - x_1^{pi}) G^{pi}(z_p) - \lambda_1^{pi}(1 - x_2^{pi}) x_1^{pi} \quad (4c)$$

$$\dot{x}_j^{pi} = \lambda_{j-1}^{pi}(1 - x_j^{pi}) x_{j-1}^{pi} - \lambda_j^{pi}(1 - x_{j+1}^{pi}) x_j^{pi} \quad (4d)$$

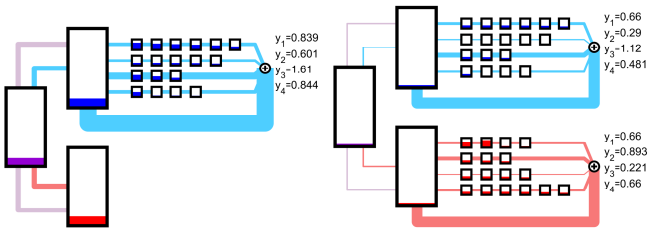
$$\dot{x}_{n^{pi}}^{pi} = \lambda_{n^{pi}-1}^{pi}(1 - x_{n^{pi}-1}^{pi}) x_{n^{pi}}^{pi} - \lambda_{n^{pi}}^{pi} x_{n^{pi}}^{pi}. \quad (4e)$$

Let $\bar{N} = z_E + \sum_p z_p$ be the number of the pool ribosomes at steady state. For a fixed $\bar{N}$, the steady state pool occupancies $z_p,\ z_E$ can be obtained by solving the linear system $0 = \dot{z}_p + \sum_{ij} m^{pi} \dot{x}_j^{pi},\ p = 1, .., N_p$ where $K_p = k_p^+ / k_p^-$

$$z_p = K_p\ z_E, \qquad \bar{N} = z_E + \sum_p z_p, \quad (5)$$

Algorithm 2 computes an approximation to the steady-state of the general model by bisection on $\bar{N}$. It converges faster than numerical integration, which is a significant difference in the optimization routine. A built-in MATLAB solver (`fzero`) may find a point that is outside the valid region.

Fig. 3 shows examples of mRNA competing for $N = 10$ ribosomes, where each mRNA may take ribosomes from either the host pool ($p = 1$, blue) or the circuit pool ($p = 2$, red). The 'empty' ribosomes are displayed in purple in the third pool. In this network, $k_p^+ = k_p^- = 1$, so at steady state,



(a) No circuit demand (b) High circuit demand

Fig. 3: Orthogonal RFM visualizations

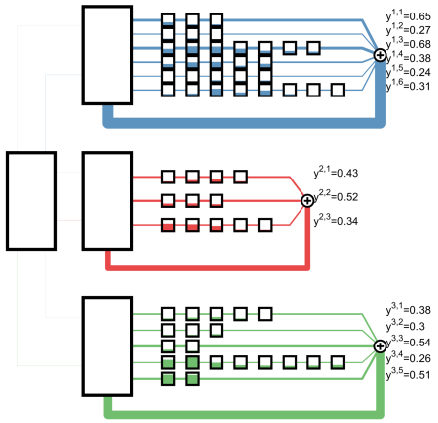


Fig. 4: ORFM with 3 ribosomal subspecies

the pool quantities $z_1 = z_2 = z_E$. In Figure 3a no mRNA takes ribosomes from the circuit pool z_2, so the ribosomes in z_2 induce deadweight loss for the system.

Fig. 4 illustrates a general model orthogonal ribosomal system with pool across three ribosomal subspecies and N = 20. Now $k_p^+ = k_p^- = 0.1,\ p = 1, 2, 3$, so all pools will have an equal number of ribosomes at equilibrium (z=0.286).

B. Stability of the ORFM

The RFM has recently been studied by constructing explicit Robust Lyapunov Functions (RLFs) [13] via writing it as a Chemical Reaction Network (CRN) and utilizing relevant methods [14], [15]. Such techniques provide explicit formulae of Lyapunov functions for general kinetics (not limited to Mass-Action), and have an easy-to-use software package. In this subsection, we derive the stability of the ORFM model (4a)-(4e) via an RLF. Let $x =: [x_j^{pi}]_{i,j,p} \in [0,1]^{N_c}$ be the vector of all codon occupancies, and N_c be the total number of codons. Let $z := [z_1, .., z_M, z_E]^T \in [0, N]^{M+1}$ be the vector of all pool occupancies. We therefore have:

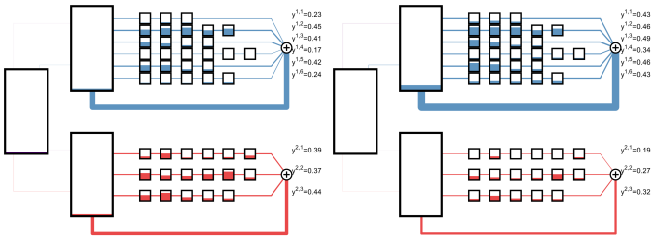
Theorem 1: Consider the system (4a)-(4e). Then,
(1) The function:

$$V(x, z) = \sum_{p,i} m^{pi} \sum_j |\dot{x}_j^{pi}| + \sum_p |\dot{z}_p| + |\dot{z}_E|,$$

is a (non-strict) Lyapunov function for any choice of $\{\lambda_j^{p,i}\} > 0$ and monotone functions G^{pi}, and
(2) For any fixed total number of ribosomes in the system $N_r > 0$, there exists a unique positive globally asymptotically stable steady-state (x_e, z_e) for (4a)-(4e).

A proof of Theorem 1 is given in the Appendix. Alternatively, the system (4a)-(4e) can be transformed into a CRN,

Authorized licensed use limited to: Northeastern University. Downloaded on February 07,2022 at 14:13:37 UTC from IEEE Xplore. Restrictions apply.



(a) Parity: $y_W = 3.11$ (b) Optimal: $y_W = 3.39$

Fig. 5: Optimize the total protein output $(w = \mathbf{1})$

and Theorem 1 can be verified via the software package LEARN [16] for a *fixed* number of pools, strands, and codons.

IV. Optimization

This section poses an optimization problem to maximize the weighted sum of protein production by mRNA. Optimization of protein throughput in a single strand of mRNA has previously been treated by Poker *et. al.* in [9]. Given an n-codon mRNA with allowable choices of λ_j in a convex set, finding rates λ^* that maximize the throughput $y = \lambda_n e_n$ is a concave optimization problem. This section adjusts rates K with constant λ to maximize a weighted-sum objective is $y_w = \sum_{p,i} w^{pi} m^{pi} y^{pi}$. Appropriate selection of the weights w can specify particular proteins as desirable. As a problem setting, consider an ORFM with M species and N ribosomes. There will be $M+1$ pools: z_E for the empty ribosomes and z_p for the translating species. If there exists sufficient flexibility to adjust $k_p^+ \in [k_p^{+,min}, k_p^{+,max}] > 0$ and $k_p^- \in [k_p^{-,min}, k_p^{-,max}] > 0$, then $K_p \in [K_p^{min}, K_p^{max}] = \left[\frac{k_p^{+,min}}{k_p^{-,max}}, \frac{k_p^{+,max}}{k_p^{-,min}}\right]$. It is assumed that $0 < K_p^{min} \leq K_p^{max} < \infty$ for each species p. For each point $K = \{K_p\}_{p=1}^M$, the weighted protein output $y_w(K)$ can be obtained by finding the ORFM steady state with respect to K using Algorithm 2 and then evaluating $y_w = \sum w^{pi} m^{pi} y^{pi}$. An optimization problem to maximize y_w at steady state can therefore be formulated:

$$y_w^* = \max_K \sum_{p,i} w^{pi} m^{pi} y^{pi}, \tag{6}$$

$$\textit{subject to}: K_p \in [K_p^{min}, K_p^{max}] \quad \forall p = 1, \ldots, M$$

Steady state of dynamics in (4a)-(4e).

The search variables of problem (6) are K_p, and the steady states (z_E, z_p, e_j^{pi}) are derived quantities of K_p. Fig 5 shows an ORFM with $s_1 = 6$ and $s_2 = 3$ strands with an objective of maximizing total protein output $(w = \mathbf{1})$. If all mRNA were connected to a common pool of ribosomes (RFMNP), the resultant output is $y_w = 3.18$. When $K = [5, 5]$, then $y_2 = 3.11$ as shown in Fig. 5a. Solving problem (6) with a grid search for $K^{min} = \frac{1}{5}$ and $K^{max} = 5$ results in the optimum rates $K^* = [5, 0.734]$ and protein output $y_w^* = 3.38$ in Fig. 5b. The optimization landscape of y_w vs. $(\log_{10}(K_1), \log_{10}(K_2))$ is displayed in Fig. 6.

A. Parameter Effects

This subsection analyzes the effects on properties of the ORFM steady state by incrementally changing a rate K_p. For a single strand x^{pi} taking ribosomes from pool z_p, the

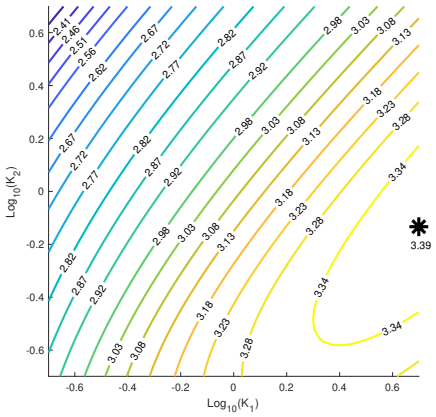


Fig. 6: Contours of total protein output (protein/sec)

effective intake rate from Eq. (1) is $\lambda_0 G(z)$. Define the throughput $y^{pi}(z_p)$ as the protein output in strand pi given pool occupancy z_p, and $N^{pi}(z_p) = \sum_{ij} e_j^{pi}$ as the ribosome occupancy of strand pi given z_p. Since $G(z)$ is monotonically increasing and through results in [9], the functions y^{pi} and N^{pi} are both nonnegative increasing functions of z_p. Let the weighted output $y_{wp}(z_p) = \sum_i w^{pi} m^{pi} y^{pi}(z_p)$, and the number of ribosomes of species p be $N_p(z_p) = z_p + \sum_i m^{pi} N^{pi}(z_p)$ across all strands of species p. The total weighted output is $y_w = \sum_p y_{wp}(z_p)$, and the occupancy is

$$N = z_E + \sum_p z_p + \sum_{p,i,j} m^{pi} e_j^{ij} = z_E + \sum_p N_p(z_p). \tag{7}$$

As a shorthand let $\partial = \partial_{K_p}$ for a chosen species p, $N_p' = dN_p'/dz$, and $y_{wp}' = dy_{wp}/dz$. From Eqs. (5) and (7)

$$0 = \partial(K_p z_E - z_p) = z_E + K_p \partial z_E - \partial z_p \tag{8a}$$

$$0 = \partial(K_{p'} z_E - z_{p'}) = K_{p'} \partial z_E - \partial z_{p'}, \qquad \forall p' \neq p \tag{8b}$$

$$0 = \partial N = \partial z_E + \sum_{p'=1}^M \partial N_{p'}(z_{p'}). \tag{8c}$$

By the chain rule, $\partial N_p(z_p) = N_p'(z_p) \partial z_p$. The quantity ∂z_E can be found from Eqs. (8a), (8b), and (8c)

$$0 = \partial z_E + \sum_{p'=1}^M \partial z_p N_p'(z_p) \tag{9a}$$

$$= \partial z_E + z_E N_p'(z_p) + \sum_{p'=1}^M K_{p'} N_{p'}'(z_{p'}) \partial z_E \tag{9b}$$

$$\partial z_E = \frac{-z_E N_p'(z_p)}{1 + \sum_{p'=1}^M K_{p'} N_{p'}'(z_{p'})} < 0. \tag{9c}$$

The change $\partial z_E < 0$ because all values $z_E, K_{p'}, N_p'(z_p) > 0$ over the valid range of K. Likewise $\partial z_{p'} = K_{p'} \partial z_E < 0$ for $p \neq p'$. In contrast,

$$\partial z_p = z_E \left(1 - \frac{K_p N_p'(z_p)}{1 + \sum_{p'=1}^M K_{p'} N_{p'}'(z_{p'})}\right) > 0. \tag{10}$$

Plugging back into the chain rule, $\partial N_p(z_p) = N_p'(z_p) \partial z_{p'} > 0$ and $\partial N_{p'}(z_{p'}) = N_{p'}'(z_{p'}) \partial z_{p'} < 0$ for $p' \neq p$. This is an intuitive conclusion: increasing the production rate of species p will increase the number of ribosomes of type p at the expense of all other species (including z_E).

The change in objective y_w by increasing K_p is

$$\partial y_w = z_E y_{wp}'(z_{p'}) + \partial z_E \sum_{p'=1}^M \partial K_{p'} y_{wp}'(z_{p'}). \tag{11}$$

The z_E term of (11) is > 0 while the ∂z_E term is < 0. The sign of ∂y_w may change over the valid region of K. Problem (6) is generically a non-concave problem in terms of K, and

Authorized licensed use limited to: Northeastern University. Downloaded on February 07,2022 at 14:13:37 UTC from IEEE Xplore. Restrictions apply.

may feature more than one local maximum depending on the choice of w. Standard approaches of global optimization such as Grid Search, Bayesian Optimization, and Basin Hopping can be used to approximate K^*, where the cost function $y_w(K)$ at each K is evaluated by Algorithm 2.

V. Self-Inhibiting Feedback Controller

The introduced optimization framework adjusts the ribosomal production rates K to maximize the weighted protein output y_w (6) assuming perfect knowledge of systems and parameters. However, if the transition rates and the number of mRNA strands per pool are subject to changes or are not known a priori, we propose a controller inspired by [12] to regulate the expression of ribosomes by self-inhibiting feedback. For instance, the system in Fig. 3b with high circuit demand can be optimized by (6) to maximize the total amount of protein ($w = \mathbf{1}$). If the circuit demand drops as shown in Fig. 3a, circuit ribosomes will be maintained without being used in translating protein (high K_2). Changing the system demand would require re-optimization of K.

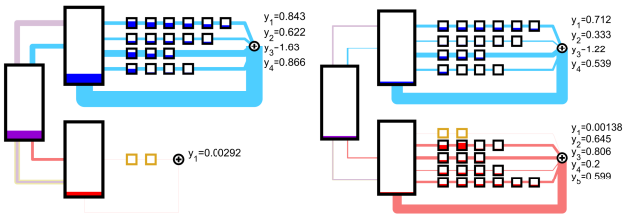
Alternatively, a feedback controller can be introduced to dynamically adjust the previously constant K, only producing ribosomes of species p when there exists corresponding demand. This can be accomplished by creating one distinguished mRNA with codon occupancies x^{pF} per species p. The new mRNA is translated into a protein F_p. The dynamics of the protein F_p with a degradation rate δ_p is

$$\dot{F}_p = \lambda_n^{pF} x_n^{pF} - \delta_p F_p. \tag{12}$$

A Hill-like inhibition term can be used to suppress the creation of new ribosomes (based on [12, S1-(18)]). For a nominal ribosome creation rate k_p^{+0}, exponent γ_p, and constant F_p^0, the suppressed k_p^+ is: $k_p^+ = k_p^{+0}/(1+F_p/F_p^0)^{\gamma_p}$. In summary, z_p activates F_p while F_p inhibits z_p. The controller is deemed to perform as desired if the number of circuit ribosomes adjusts according to the circuit demand. Removing the circuit from the system would initially raise F_p, which represses k_p^+ strongly and hence reduces the number of circuit ribosomes. If circuit genes are introduced, the new genes will compete with the inhibitor mRNA and the number of circuit ribosomes will increase.

Fig. 7 shows an example of an ORFM with 10 ribosomes, where the circuit pool has a feedback controller. The inhibitor protein F_2 is translated from the golden two-codon mRNA with occupancies x^{2F} and rates $\lambda^{2F} = [0.005, 5, 5]$. The low initiation rate $\lambda_0^F = 0.005$ allows other mRNA if present to take ribosomes from z_2 first. The inhibitor parameters are $F_2^0 = 1.5, \gamma_2 = 4$, and $\delta_2 = 0.01$. The golden sheath in Fig. 7 between z_2 and z_E represents the nominal rate $k_2^{+0} z_E$, compared to the purple core's inhibited $k_2^+ z_E$. With no circuit demand, the circuit ribosome creation rate in Fig 7a falls from 1.393 to 0.669 as desired.

Let $N_1 = z_1 + \sum_{i,j} x_j^{1i}$ be the number of host ribosomes (species 1), and y_w be the total non-inhibitory protein output (excluding F_2). With no inhibitor or circuit mRNA, the system in Fig. 3a has an expected $N_1 = 7.63$ and $y_w = 3.89$. With the inhibitor in Fig. 7a, the expected total



(a) No circuit demand (b) High circuit demand

Fig. 7: ORFM with inhibitor protein F_2, at $t = 5000$.

number of host ribosomes rises to $N_1 = 7.97$, $y_w = 3.96$. Fig. 3b has a high circuit demand and no inhibitor, and $N_1 = 3.94$, $y_w = 4.98$. With an inhibitor in Fig. 7b, $N_1 = 4.41$, $y_w = 5.06$. At no circuit demand the steady state $F_2 = 0.292$ units of inhibiting protein, and at high circuit demand $F_2 = 0.138$ units.

VI. Conclusion

Competition for finite resources are inevitable in protein translation. Orthogonal ribosomes have been developed to boost protein throughput by decoupling circuit genes from the host pool of ribosomes. We extended the existing RFM to orthogonal ribosomes, and generalized the system to an arbitrary number of ribosomal species. Stability results through RLFs and a simple algorithm to compute steady states were presented. Maximizing the weighted sum of protein throughput can be formulated as an optimization problem. A self inhibiting feedback controller can adjust ribosomal production as needed. Future work includes matching results with lab experiments and further development of the feedback controller. The problem of cross-talk in translation is discussed in an extended version of this paper [17].

Appendix: Proof of Theorem 1

For simplicity, assume that $s_p = s, n^{p,i} = n$, for all p, i. Also, assume $m^{p,i} = 1$, for all p, i. The argument for the general case will be the similar. Denote $\sigma(x) := \text{sgn}(x)$.

To use the techniques in [15], [13], we lift (4a)-(4e) to a higher-dimensional space by defining the *vacancy* $w_j^{p,i} := 1 - x_j^{pi}$, for all j, i, p. Hence, terms of the form $(1 - x_{j-1}^{pi}) x_j^{pi}$ take the familiar Mass-Action form: $w_{j-1}^{pi} x_j^{pi}$. We will generalize this further by considering arbitrary *monotone* functions of the form $R(w_{j-1}^{pi}, x_j^{pi})$. Hence, we write (4a)-(4e) as:

$$\begin{aligned}
\dot{z}_E &= \textstyle\sum_p R_p^-(z_p) - \sum_p R_p^+(z_E) \\
\dot{z}_p &= R_p^+(z_E) - R_p^-(z_p) + \textstyle\sum_i \left(R_n^{pi}(x_n^{pi}) - R_0^{pi}(w_1^{pi}, z_p) \right) \\
\dot{x}_1^{pi} &= R_0^{pi}(w_1^{pi}, z_p) - R_1^{pi}(w_2^{pi}, x_1^{pi}) \\
\dot{x}_j^{pi} &= R_{j-1}^{pi}(w_j^{pi}, x_{j-1}^{pi}) - R_j^{pi}(w_{j+1}^{pi}, x_j^{pi}) \\
\dot{x}_n^{pi} &= R_{n-1}^{pi}(w_n^{pi}, x_{n-1}^{pi}) - R_n^{pi}(x_n^{pi}).
\end{aligned} \tag{13}$$

We only assume that the rates $R_j^{p,i}, R_p^{\pm}$ are monotone w.r.t their reactants, see [13] for the full assumptions.

Consider $V(x, z) = \sum_{i,p,j} |\dot{x}_j^{pi}| + \sum_p |\dot{z}_p| + |\dot{z}_E|$. We show that V is non-increasing. Using (13), note that the space can be partitioned into regions, and V is *linear* in *rates* on each region. Fix an *open* region $\mathcal{W}$ and write $V = \sum_{p,i,j} \alpha_j^{pi} R_j^{pi}(x, z) + \sum_p \beta_p (R_p^-(z_p) - R_p^+(z_E))$ on $\mathcal{W}$.

Authorized licensed use limited to: Northeastern University. Downloaded on February 07,2022 at 14:13:37 UTC from IEEE Xplore. Restrictions apply.

Note V is differentiable on $\mathcal{W}$, and the signs of the currents $\dot{x}_j^{pi}, \dot{z}_p, \dot{z}_E$ are constant on $\mathcal{W}$. We claim that the derivative of each term in V is non-positive. We show first that $\alpha_j^{pi}\dot{R}_j^{pi} \leq 0$ for all p, i, j. We study three cases: $j = 0, n$, $0 < j < n$. If $j \neq 0, n$, then R_j^{pi} appears in $\dot{x}_j^{pi}, \dot{x}_{j+1}^{pi}$. If both have the same sign on $\mathcal{W}$, then (13) implies $\alpha_j^{pi} = 0$. Therefore, $\alpha_j^{pi} \neq 0$ implies $\sigma(\alpha_j^{pi}) = -\sigma(\dot{x}_j^{pi}) = \sigma(\dot{x}_{j+1}^{pi}) = -\sigma(\dot{w}_{j+1}^{pi})$. Hence, $\alpha_j^{pi}\dot{R}_j^{pi} = \alpha_j^{pi}(\frac{\partial R_j^{pi}}{\partial x_j^{pi}}\dot{x}_j^{pi} + \frac{\partial R_j^{pi}}{\partial w_{j+1}^{pi}}\dot{w}_{j+1}^{pi}) \leq 0$ as claimed, where non-negativity of the partial derivatives follows from monotonicity. Next, consider the case $j = 0$ where R_0^{pi} appears in two currents $\dot{x}_1^{pi}, \dot{z}_p$. Similar to the previous case, we conclude that if $\alpha_0^{pi} \neq 0$ then $\sigma(\alpha_0^{pi}) = -\sigma(z_p) = -\sigma(\dot{w}_1^{pi})$. Since R_0^{pi} is a monotone function of w_1^{pi}, z_1^{pi}, then $\alpha_0^{pi}\dot{R}_0^{pi} \leq 0$ as claimed. Finally, if $j = n$, similar analysis can be repeated to conclude that if $\alpha_n^{pi} \neq 0$ then $\sigma(\alpha_n) = -\sigma(x_n^{pi})$ and $\alpha_n\dot{R}_n^{pi} \leq 0$ as claimed. Next, we want to show that $\beta_p\dot{R}_p^- \leq 0$. Fix p. Note that R_p^- appears in two currents: $\dot{z}_p, \dot{z}_E$. If they have the same sign then $\beta_p = 0$. If they have different signs then (13) implies that $\sigma(\beta_p) = -\sigma(\dot{z}_p) = \sigma(\dot{z}_E)$. Since R_p^- is a function of z_p, then $\beta_p\dot{R}_p^- \leq 0$. A similar argument can be replicated to show that $-\beta_p\dot{R}_p^+ \leq 0$. Since $\mathcal{W}$ and p, i, j were arbitrary, we conclude that $\dot{V} \leq 0$ over the interior of all regions. The boundaries between the regions can be dealt with via the definition of the upper Dini's derivative, see [15]. This concludes the proof of the first statement of Theorem 1.

We proceed to prove part 2. The RLF utilized is non-strict, hence it cannot yield Global Asymptotic Stability (GAS) directly. The LaSalle argument proposed in [15] can be used, but it is long. Alternatively, we will show that the Jacobian of (13) (reduced to the invariant space for a fixed N_r) is robustly non-degenerate. This, coupled with the RLF, implies the GAS of any steady-state [18], [13].

To prove non-degeneracy, it has been shown in [14], [13] that the Jacobian J of (13) at any point in the space can be written as $J = \sum_{\ell=1}^{L}\rho_\ell Q_\ell$ for some $\rho_\ell \geq 0$, rank-one matrices Q_ℓ and some L. Here, $L = M(2+s(2n+1))$. The coefficients $\{\rho_\ell\}$ correspond to the partial derivatives of the rates $\{R_j^{pi}\}, \{R_p^{\pm}\}$, while $\{Q_\ell\}$ correspond to the network structure and are independent of the rates. Furthermore, it has been shown also [13] that the existence of an RLF implies that it sufficient to find *one positive point* $\rho^* = (\rho_1^*, .., \rho_L^*)$ for which the (reduced) Jacobian is non-degenerate to show that is non-degenerate for all $\rho_\ell > 0$. We will find that point next by studying the structure of the Jacobian.

Similar to a single pool [11, SI], it can be easily seen that J has non-negative off-diagonals and strictly negative diagonals (i.e, J is Metzler). In addition, conservation of the number of ribosomes implies that $\mathbf{1}^T J = 0$. By the definition of the Jacobian, all the entries in each column contain only the partial derivatives with respect to the state variable associated to the column. Hence, we can choose the corresponding ρ_ℓ's such the diagonal entry in each column is scaled to -1. Therefore, we consider the Jacobian evaluated at the chosen point ρ^* such that $J^* = P - I$, where I is the identity matrix and P a nonnegative irreducible column-stochastic matrix. By Perron-Frobenius Theorem, P has a maximal eigenvalue 1 with algebraic multiplicity 1. Therefore, J^* has a single eigenvalue at 0 and the remaining eigenvalues have strictly negative real-parts. Hence, the reduced Jacobian at ρ^* is non-degenerate. Robust non-degeneracy and GAS follows.

The existence of a steady-state follows from Brouwer's fixed point theorem since (13) evolves in a compact space (for a fixed $N_r > 0$), and uniqueness follows from non-degeneracy and GAS. The positivity of the steady-state follows from persistence of the ORFM which can be shown graphically by the absence of critical siphons [19]. ■

Acknowledgements: This research was partially funded by NSF grants 1849588 and 1716623. Jared Miller thanks Prof. Mario Sznaier for his support and discussions.

REFERENCES

[1] T. Frei, F. Cella, F. Tedeschi, J. Gutierrez, G. B. Stan, M. Khammash, and V. Siciliano. Characterization and mitigation of gene expression burden in mammalian cells. *Nat Commun*, 11(1):4641, 09 2020.

[2] O. Rackham and J.W. Chin. A network of orthogonal ribosome·mRNA pairs. *Nature Chemical Biology*, 1(3):159–166, 2005.

[3] Oliver PT Barrett and Jason W Chin. Evolved orthogonal ribosome purification for in vitro characterization. *Nucleic acids research*, 38(8):2682–2691, 2010.

[4] A. Costello and A.H. Badran. Synthetic biological circuits within an orthogonal central dogma. *Trends in Biotechnology*, In press, 2020.

[5] L.M. Chubiz and C.V. Rao. Computational design of orthogonal ribosomes. *Nucleic acids research*, 36(12):4038–4046, 2008.

[6] S. Reuveni, I. Meilijson, M. Kupiec, E. Ruppin, and T. Tuller. Genome-scale analysis of translation elongation with a ribosome flow model. *PLoS Comput Biol*, 7(9):e1002127, 2011.

[7] C.T. MacDonald, J.H. Gibbs, and A.C. Pipkin. Kinetics of biopolymerization on nucleic acid templates. *Biopolymers: Original Research on Biomolecules*, 6(1):1–25, 1968.

[8] C.T. MacDonald and J.H. Gibbs. Concerning the kinetics of polypeptide synthesis on polyribosomes. *Biopolymers: Original Research on Biomolecules*, 7(5):707–725, 1969.

[9] G. Poker, Y. Zarai, M. Margaliot, and T. Tuller. Maximizing protein translation rate in the non-homogeneous ribosome flow model: a convex optimization approach. *Journal of The Royal Society Interface*, 11(100):20140713, 2014.

[10] Y. Zarai, M. Margaliot, E.D. Sontag, and T. Tuller. Controllability analysis and control synthesis for the ribosome flow model. *IEEE/ACM Trans. Comput. Biol. Bioinf.*, 15(4):1351–1364, 2017.

[11] A. Raveh, M. Margaliot, E.D. Sontag, and T. Tuller. A model for competition for ribosomes in the cell. *Journal of The Royal Society Interface*, 13(116):20151062, 2016.

[12] A.P.S. Darlington, J. Kim, J.I. Jiménez, and D.G. Bates. Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes. *Nature Communications*, 9(1):695, 2018.

[13] M. Ali Al-Radhawi, D. Angeli, and E.D. Sontag. A computational framework for a Lyapunov-enabled analysis of biochemical reaction networks. *PLoS Computational Biology*, 16(2):e1007681, 2020.

[14] M. Ali Al-Radhawi and D. Angeli. Robust Lyapunov functions for complex reaction networks: An uncertain system framework. In *53rd IEEE CDC*, pages 3101–3106, Dec 2014.

[15] M. Ali Al-Radhawi and D. Angeli. New approach to the stability of chemical reaction networks: Piecewise linear in rates Lyapunov functions. *IEEE Trans. on Automatic Control*, 61(1):76–89, 2016.

[16] M. Ali Al-Radhawi. Lyapunov-enabled analysis of reaction networks (LEARN) http://github.com/malirdwi/learn, 2020.

[17] J. Miller, M. Ali Al-Radhawi, and E. D. Sontag. Mediating ribosomal competition by splitting pools. *arXiv preprint arXiv:2009.00539*, 2020.

[18] F. Blanchini and G. Giordano. Polyhedral Lyapunov functions structurally ensure global asymptotic stability of dynamical networks iff the Jacobian is non-singular. *Automatica*, 86:183–191, 2017.

[19] D. Angeli, P. De Leenheer, and E.D. Sontag. A Petri net approach to the study of persistence in chemical reaction networks. *Mathematical Biosciences*, 210(2):598–618, 2007.

Authorized licensed use limited to: Northeastern University. Downloaded on February 07,2022 at 14:13:37 UTC from IEEE Xplore. Restrictions apply.