Maximizing protein translation rate in the non-homogeneous ribosome flow model: a convex optimization approach

Gilad Poker1, Yoram Zarai1, Michael Margaliot1,2 and Tamir Tuller2,3

1School of EE-Systems, 2The Sagol School of Neuroscience, and 3Department of Biomedical Engineering, Tel Aviv University, Tel Aviv 69978, Israel

Translation is an important stage in gene expression. During this stage, macro-molecules called ribosomes travel along the mRNA strand linking amino acids together in a specific order to create a functioning protein. An important question, related to many biomedical disciplines, is how to maximize protein production. Indeed, translation is known to be one of the most energy-consuming processes in the cell, and it is natural to assume that evolution shaped this process so that it maximizes the protein production rate. If this is indeed so then one can estimate various parameters of the translation machinery by solving an appropriate mathematical optimization problem. The same problem also arises in the context of synthetic biology, namely, re-engineer heterologous genes in order to maximize their translation rate in a host organism. We consider the problem of maximizing the protein production rate using a computational model for translation–elongation called the ribosome flow model (RFM). This model describes the flow of the ribosomes along an mRNA chain of length \( n \) using a set of \( n \) first-order nonlinear ordinary differential equations. It also includes \( n + 1 \) positive parameters: the ribosomal initiation rate into the mRNA chain, and \( n \) elongation rates along the chain sites. We show that the steady-state translation rate in the RFM is a strictly concave function of its parameters. This means that the problem of maximizing the translation rate under a suitable constraint always admits a unique solution, and that this solution can be determined using highly efficient algorithms for solving convex optimization problems even for large values of \( n \). Furthermore, our analysis shows that the optimal translation rate can be computed based only on the optimal initiation rate and the elongation rate of the codons near the beginning of the ORF. We discuss some applications of the theoretical results to synthetic biology, molecular evolution, and functional genomics.

1. Introduction

Gene expression is the process by which the information encoded in the genes is used to synthesize proteins. The two major steps of gene expression are the transcription of the genetic information from DNA to messenger RNA (mRNA) by RNA polymerase, and the translation of the mRNA molecules to proteins. During gene translation, the genetic information is deciphered into proteins by molecular machines called ribosomes that move along the mRNA chain in a unidirectional manner from the 5' end to the 3' end [1]. Each triplet of the mRNA consecutive nucleotides, called a codon, is decoded by a ribosome into a corresponding amino acid. The rate in which proteins are produced during the translation step is referred to as the protein production rate or translation rate.

Translation occurs in all organisms, in all known cells and under almost all conditions. Thus, understanding translation has important implications in many scientific disciplines, including medicine, biotechnology, functional genomics, evolutionary biology and much more. The amount of biological findings related to translation increases at an exponential rate and this leads to considerable interest in computational models that can integrate and analyse these findings [2–11].
A fundamental challenge in biotechnology and synthetic biology is to control the expression of heterologous genes in a host organism in order to synthesize new proteins or to improve certain aspects of the host fitness [12–14]. Computational models of translation are also important in this context, as they allow one to simulate and analyse the effect of various manipulations of the genomic machinery on the translation process.

A conventional computational model of translation–elongation is the totally asymmetric simple exclusion process (TASEP) [15,16]. TASEP is a stochastic model that describes particles moving along a one-dimensional lattice of sites. The term totally asymmetric is used to indicate unidirectional motion along the chain. Each site can be either empty or occupied by a single particle. This captures interaction between the particles, as a particle in site \(i\) blocks the movement of a particle in site \(i - 1\). Hence, the term simple exclusion. At each time instant, the sites are scanned and provided that a site is occupied by a particle and the next site is empty, the particle hops to the next site with some probability. The two sides of the chain are connected to particle reservoirs, and particles can hop into the chain (if the first site is empty) and out of the chain (if the last site is full). TASEP is a fundamental model in non-equilibrium statistical mechanics that has been used to model numerous natural and artificial processes [17]. Analysis of TASEP is based on determining the probabilities of steady-state configurations using matrix products (see the excellent review paper [18]).

The ribosome flow model (RFM) [19] is a deterministic model for translation–elongation that can be obtained via a mean-field approximation of TASEP (see e.g. [17, §4.9.7] and [18, p. R345]). The RFM for a chain with \(n\) sites includes \(n\) first-order, nonlinear ordinary differential equations and \(n + 1\) positive parameters: the initiation rate \(R_0\) and elongation rates \(R_i\), \(i = 1, 2, \ldots, n\), between every two consecutive sites.

There are indications in some genes all the elongation rates along the mRNA chain are (approximately) equal [20]. This may be modelled by assuming constant elongation rates in the RFM. This yields the homogeneous ribosome flow model (HRFM) [21] that includes only two positive parameters: the initiation rate \(R_0\) and the constant elongation rate \(R_c\).

In a previous study [22], we have shown that the steady-state protein translation rate in the HRFM, denoted by \(R = R(R_0, R_c)\), is a concave function of the parameters \(R_0, R_c\). The proof of this result is based on analysing the Hessian matrix \(H\) of \(R\) in the HRFM. Note that \(H\) has dimensions \(2 \times 2\) for all \(n\). However, the assumption of equal elongation rates is often too strong. For example, it was shown that factors such as the adaptation of codons to the RNA pool [23–25], folding of the mRNA [6,25] and local amino acid charge [6,25,26] affect translation elongation speed. This induces variations between different elongation rates. In these cases, the HRFM is not a suitable model, and one must use the RFM. The steady-state translation rate in the RFM is a function of \(n + 1\) parameters, i.e. \(R = R(R_0, \ldots, R_n)\). In this paper, we show that \(R(R_0, \ldots, R_n)\) is a strictly concave function of its \(n + 1\) positive parameters. Here, the Hessian matrix has dimensions \((n + 1) \times (n + 1)\), and it seems that the approach applied in [22] cannot be extended to handle the RFM. The proof of our main result is thus based on an entirely new technique.

To explain the importance of the strict concavity of \(R\), consider figure 1 that depicts, for simplicity, a scalar strictly concave function \(y = f(x)\). Strict concavity in this case means the following. Given any two different values \(x_1\) and \(x_2\), with the corresponding function values \(y_1 = f(x_1)\) and \(y_2 = f(x_2)\), let \(l = l(x)\) denote the line that connects the points \((x_1, y_1)\) and \((x_2, y_2)\). Then \(f(x) > l(x)\), for all \(x \in (x_1, x_2)\). In other words, the graph of the function lies above the line \(l(x)\).

Concave functions have many useful and desirable properties. First, a concave function is differentiable almost everywhere. Second, recall that a point \(x_m\) is called a local maximum of a function \(f\) if the function values in some neighborhood of \(x_m\) are smaller than or equal to \(f(x_m)\). It is a global maximum if the function values in its entire domain of definition are smaller than or equal to \(f(x_m)\). For a concave function, any local maximum is also a global maximum. If the function is strictly concave then this maximum is unique.

Furthermore, strict concavity implies that a simple ‘hill climbing’ algorithm can be used to find the global maximum. In the depicted one-dimensional function, this can be explained as follows. Select an arbitrary point \(x_0\) in the domain of definition of \(f\) as a candidate for a maximum point. Next, determine two points \(x_1\) and \(x_2\) that are ‘close’ to \(x_0\) and satisfy \(x_0 < x_1 < x_2\). Denote \(y_0 = f(x_0)\), \(y_1 = f(x_1)\) and \(y_2 = f(x_2)\). If \(y_0 \geq y_1\) and \(y_0 \geq y_2\), then \(x_0\) is a local, and thus global, maximum of the function and the algorithm terminates. Otherwise, at least one of the two values \(y_0\) or \(y_2\) is larger than \(y_0\). The corresponding point, i.e. \(x_1\) or \(x_2\), becomes the new candidate for a maximum, and the algorithm is iterated. Under mild assumptions, this simple algorithm is guaranteed to converge to the global maximum of the concave function. More generally, there exist highly efficient algorithms for finding the global maximum of multi-dimensional concave functions [27].

A function \(g\) is called (strictly) convex if \(-g\) is (strictly) concave. Thus, the problem of finding the maximum value of a concave function is equivalent to the problem of finding the minimum value of a convex function. A famous quote by Rockafellar states that: ‘...the great watershed in optimization isn’t between linearity and nonlinearity, but convexity and non-convexity’ [28, p. 185]. We note in passing that a linear function \(y(x) = ax + b\) is both concave and convex.

Summarizing, our main result implies that the problem of maximizing the protein translation rate, under a simple constraint on the RFM parameter values, admits a unique
solution, and that this solution can be found numerically using highly efficient algorithms. It is important to note that many systems and processes have been modelled and analysed using TASEP. These include translation, traffic flow, molecular motors, surface growth, the movement of ants and more [17]. All these processes may also be modelled using the RFM, and the problem of maximizing $R$ seems to be of importance in all of them.

We now describe some possible applications of the main result in the context of translation. A recent work [29] studied the effect of the intracellular translation factor abundance on the protein production rate. Abundance of the encoded translation factor was experimentally manipulated to a sub-wild-type level [29] using the tet07 construct. The reported results suggest that the mapping from levels of translation factors to protein production rate is concave [29, fig. 1]. This may provide an experimental support to the results presented in this paper. Note that Firczuk et al. [29] used the model organism Saccharomyces cerevisiae, which is known to have non-constant elongation rates [6,23]. Thus, the RFM, and not the HRFM [22], is a better computational model for describing these experiments.

Translation is one of the major energy-consuming processes in the cell [1,30,31]. A reasonable assumption is that in organisms under strong evolutionary pressure the genomic machinery has evolved so that it optimizes the translation rate given the available resources. This assumption can be studied in the context of the RFM since the concavity of the translation rate implies that one can easily determine the optimal parameter values, and then compare them with biological findings. This may help in understanding the level of selection pressure acting on the genomes of various organisms and the evolutionary changes in various micro-organisms [32].

In synthetic biology, an important problem is to re-engineer a genetic system by manipulating the transcript sequence, and possibly other intra-cellular variables, in order to obtain an optimal translation rate. Using our results on the RFM can provide verifiable predictions on how this can be done efficiently. Another related problem is optimizing the translation efficiency and protein levels of heterologous genes in a new host [30,31,33,34]. These genes actually compete with endogenous genes for the available resources, e.g. initiation factors. Consuming too much resources by the heterologous gene may kill the host [30,31]. Thus, any optimization of the protein translation rate should not consume too many resources, as otherwise the fitness of the host may be significantly reduced. This seems to fit well with the constrained optimization problem that we pose here for the RFM.

The remainder of this paper is organized as follows. Section 2 briefly reviews the RFM. Section 3 presents the main results. Section 4 describes the implications of our results to systems biology, evolution and synthetic biology, and describes several possible directions for further research. To streamline the presentation, all the proofs are placed in appendix A.

2. Preliminaries

The RFM [19] is a deterministic mathematical model for translation–elongation. In the RFM, mRNA molecules are coarse-grained into a unidirectional chain of $n$ sites of codons. The RFM is a set of $n$ first-order nonlinear ordinary differential equations:

$$
\dot{x}_i = \lambda_0(1 - x_i) - \lambda_1 x_i (1 - x_{i+1}) + \lambda_2 x_i (1 - x_{i+2}) - \lambda_3 x_i x_3 (1 - x_4),
$$

(2.1)

Here, $x_i(t) \in [0, 1]$ is the occupancy level at site $i$ at time $t$, normalized so that $x_i(0) = 0$ [$x_i(1) = 1$] implies that site $i$ is completely empty (completely full) at all time $t$. The units of $x_i$ are thus 1/time. The parameter $\lambda_0 > 0$ is the initiation rate into the chain, and $\lambda_i > 0$, $i \in \{1, \ldots, n\}$, is a parameter that controls the flow from site $i$ to site $i + 1$. In particular, $\lambda_0$ controls the output rate at the end of the chain. These parameters have units of 1/time.  

The rate of ribosome flow into the system is $\lambda_0(1 - x_1(t))$. The rate of ribosome flow exiting the last site, i.e. the protein production rate, is $\lambda_n x_n(t)$. The rate of ribosome flow from site $i$ to site $i + 1$ is $\lambda_2 x_i (1 - x_{i+1}) (1 - x_{i+2})$ (figure 2). Note that this rate increases with $x_i(t)$ (i.e. when site $i$ is fuller) and decreases with $x_{i+1}(t)$ (i.e. when the consecutive site is becoming fuller).

In this way, the RFM, just like TASEP, takes into account the interaction between the ribosomes in consecutive sites.

We emphasize that in the RFM the state variables take values in the closed interval $[0, 1]$ and are not limited to the values $\{0, 1\}$. This is different from TASEP, where a site can be either empty or full. Indeed, the $x_i$’s in the RFM may be interpreted as time-averaged occupancy levels in TASEP, and this average takes values in $[0, 1]$.

Let $x(t, a)$ denote the solution of (2.1) at time $t \geq 0$ for the initial condition $x(0) = a$. Since the state variables correspond to normalized occupation levels, we always assume that $a$ belongs to the closed $n$-dimensional unit cube:

$$C^n := \{x \in \mathbb{R}^n : x_i \in [0, 1], \quad i = 1, \ldots, n\}.$$  

It is straightforward to verify that this implies that $x(t, a) \subseteq C^n$ for all $t \geq 0$. In other words, $C^n$ is an invariant set of the dynamics [35].
Let int(C^n) denote the interior of C^n. It was shown in [35] that the RFM is a monotone dynamical system [36] and that this implies that (2.1) admits a unique equilibrium point e ∈ int(C^n). Furthermore,
\[ \lim_{t \to \infty} x(t, a) = e \quad \text{for all } a \in C^n. \]
This means that all trajectories converge to the steady-state e.

We note in passing that monotone dynamical systems have recently found many applications in systems biology (see, e.g. [37–39] and references therein).

For \( x = e \), the left-hand side of all the equations in (2.1) is zero, so
\[ \lambda_0(1 - e_1) = \lambda_1e_1(1 - e_2) = \lambda_2e_2(1 - e_3) = \cdots = \lambda_{n-1}e_{n-1}(1 - e_n) = \lambda_ne_n. \]
Denoting the steady-state translation rate by
\[ R := \lambda_0e_n \]
yields
\[ R = \lambda_i e_i(1 - e_{i+1}), \quad i \in \{1, \ldots, n\}, \]
where we define \( e_{n+1} := 0 \). Also,
\[ e_n = \frac{R}{\lambda_i}, \quad e_{n-1} = \frac{R}{(\lambda_{n-1}(1 - e_n))}, \quad \cdots, \quad e_2 = \frac{R}{(\lambda_1(1 - e_3))}, \quad e_1 = \frac{R}{(\lambda_1(1 - e_2))}, \]
and
\[ e_1 = \frac{1 - R}{\lambda_0}. \]
Combining (2.5) and (2.6) provides a finite continued fraction [40] expression for R:
\[ 0 = 1 - \frac{R/\lambda_0}{1 - \frac{R/\lambda_1}{1 - \frac{R/\lambda_2}{1 - \frac{R/\lambda_{n-1}}{1 - \frac{R}{\lambda_n}}}}}. \]
Note that this equation has several solutions for R (and thus also several solutions for \( e_0 = R/\lambda_0 \)), however, we are interested only in the unique feasible solution, i.e. the solution corresponding to \( e \in int(C^n) \).

Equation (2.7) may be written as \( p(R) = 0 \), where \( p \) is a polynomial of degree \([(n + 1)/2]\) in R with coefficients that are algebraic functions of the \( \lambda_i \)'s. For example, for \( n = 3 \), (2.7) yields
\[ (\lambda_0\lambda_2 + \lambda_1\lambda_2 + \lambda_1\lambda_3)R^2 - (\lambda_0\lambda_1\lambda_2 + \lambda_0\lambda_1\lambda_3 + \lambda_0\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_3)R + \lambda_0\lambda_1\lambda_2\lambda_3 = 0. \]

Recent biological findings suggest that in some cases the transition rate along the mRNA chain is approximately constant [20]. This may also be the case for gene transcription [41]. To model this, Margaliot & Tuller [21] have considered the RFM in the special case where
\[ \lambda_1 = \lambda_2 = \cdots = \lambda_n := \lambda_c, \]
that is, the transition rates \( \lambda_i, i = 1, 2, \ldots, n \), are all equal, and \( \lambda_c \) denotes their common value. Since this HRFM includes only two parameters, \( \lambda_0 \) and \( \lambda_c \), the analysis is simplified. In particular, (2.7) becomes
\[ 0 = 1 - \frac{R/\lambda_0}{1 - \frac{R/\lambda_c}{1 - \frac{R}{\lambda_c}}}. \]
where \( \lambda_c \) appears a total of \( n \) times.

Several recent papers analysed the RFM or HRFM. To model ribosome recycling (see, e.g. [42] and references therein), Margaliot & Tuller [43] have considered a closed-loop RFM with a positive linear feedback from the output R to the input \( \lambda_0 \). It has been shown that the closed-loop system admits a unique globally asymptotically stable equilibrium point. In [44], it has been shown that the state variables (and thus the protein production rate) in the RFM entrain to periodically time-varying initiation and/or transition rates. This provides a computational framework for studying entrainment to a periodic excitation (e.g. the cell cycle) at the translation level. The HRFM with an infinitely long chain, (i.e. with \( n \to \infty \) ) was considered in [45]. There, a simple closed-form expression for \( e_\infty := \lim_{n \to \infty} e_n \) was derived, as well as explicit bounds for \( |e_n - e_\infty| \) for all \( n \geq 2 \).

In the RFM the steady-state production rate R is a function of the positive parameters \( \lambda_0, \lambda_1, \ldots, \lambda_n \). In this paper, we study the dependence of R on these parameters. Our results are based on a novel, linear-algebraic approach linking the protein translation rate to the maximum eigenvalue of a symmetric, non-negative tridiagonal matrix whose components are functions of the \( \lambda_i \)'s.

### 3 Main results

#### 3.1. Concavity

The next result is the main result in this section. Recall that all the proofs are placed in appendix A. Let \( R^{n+1}_+ := \{ x \in R^{n+1}_+: x_i \geq 0, \quad i = 1, \ldots, n+1 \} \).

**Theorem 3.1.** Consider the RFM with dimension n. The steady-state translation rate \( R = R(\lambda_0, \ldots, \lambda_n) \) is a strictly concave function on int(\( R^{n+1}_+ \)).

The next example demonstrates theorem 3.1.

**Example 3.2.** Consider the RFM with \( n = 1 \). In this case, (2.5) and (2.6) yield \( e_1 = R/\lambda_1 \) and \( e_1 = 1 - R/\lambda_0 \), so
\[ R(\lambda_0, \lambda_1) = \frac{\lambda_0\lambda_1}{\lambda_0 + \lambda_1}. \]

Figure 3 depicts \( R(\lambda_0, \lambda_1) \) as a function of its arguments. It may be seen that this is indeed a strictly concave function on int(\( R^2 \)).

Recall that a function \( f: R^p \to R \) is called positively homogeneous of degree \( m \) if \( f(cx) = c^mf(x) \) for all \( c > 0 \) and all
x \in \mathbb{R}_+^3$. For example, the function $f(x_1, x_2) := x_1x_2 + x_2^2$ is positively homogeneous of degree 2. The following result follows immediately from the fact that $R$ always appears in (2.7) only in terms of the form $R/\lambda_i$.

**Fact 3.3.** Consider the RFM with dimension $n$. The function $R = R(\lambda_0, \ldots, \lambda_n)$ is positively homogeneous of degree 1.

In other words,

$$R(c\lambda_0, \ldots, c\lambda_n) = cR(\lambda_0, \ldots, \lambda_n) \quad \text{for all } c > 0. \quad (3.2)$$

From a biophysical point of view, this means that multiplying the initiation rate and all the elongation rates by the same factor $c > 0$ increases of the steady-state production rate by a factor of $c$. This also means that the steady-state occupancy levels $c_{iv} i = 1, 2, \ldots, n$, remain unchanged with respect to such a multiplication.

**Example 3.4.** Consider the RFM with dimension $n = 2$. In this case, the feasible solution of (2.5) and (2.6) (i.e. the solution corresponding to a value $c \in \mathbb{C}$ for all $\lambda_0, \lambda_1, \lambda_2 > 0$) is

$$R(\lambda_0, \lambda_1, \lambda_2) = \frac{\lambda_0\lambda_1 + \lambda_0\lambda_2 + \lambda_1\lambda_2 - \sqrt{(\lambda_0\lambda_1 + \lambda_0\lambda_2 + \lambda_1\lambda_2)^2 - 4\lambda_0\lambda_1^2\lambda_2}}{2\lambda_1}, \quad (3.3)$$

and clearly this implies that $R(c\lambda_0, c\lambda_1, c\lambda_2) = cR(\lambda_0, \lambda_1, \lambda_2)$.

Recall that a function $f : \mathbb{R}_+^n \rightarrow \mathbb{R}$ is called superadditive if $f(x + y) \geq f(x) + f(y)$ for all $x, y \in \mathbb{R}_+^n$. It is well known that for a positively homogeneous function, concavity is equivalent to superadditivity [46]. Combining this with fact 3.3 and theorem 3.1 yields the following result.

**Corollary 3.5.** Consider the RFM with dimension $n$. The function $R = R(\lambda_0, \ldots, \lambda_n)$ is superadditive.

This means that

$$R(\lambda_0 + \lambda_0, \ldots, \lambda_n + \lambda_n) \geq R(\lambda_0, \ldots, \lambda_n) + R(\lambda_0, \ldots, \lambda_n),$$

for all $\lambda_0, \ldots, \lambda_n, \lambda_0, \ldots, \lambda_n > 0$. From a biophysical point of view this means the following. Consider two RFMs, one with initiation rate $\lambda_0$ and transition rates $\lambda_1, \lambda_2, \ldots, \lambda_n$, and the second with initiation rate $\lambda_0$ and transition rates $\lambda_1, \lambda_2, \ldots, \lambda_n$. The sum of the production rates of these two RFMs is smaller or equal to the production rate of a single RFM with initiation rate $\lambda_0 + \lambda_0$ and transition rates $\lambda_1 + \lambda_1, \lambda_2 + \lambda_2, \ldots, \lambda_n + \lambda_n$. In other words, a single RFM with rates $\lambda_i + \lambda_j$ is at least as efficient as the total of two separate RFMs, one with rates $\lambda_i$ and the second with rates $\lambda_j$, for $i = 0, 1, \ldots, n$.

### 3.2. Constrained maximization of the protein translation rate

Consider the problem of determining the parameter values $\lambda_0, \ldots, \lambda_n$ that maximize $R$ (or, equivalently, that minimize $-R$) in the RFM. Obviously, to make this problem meaningful we must constrain the possible parameter values. This leads to the following constrained optimization problem.

**Problem 3.6.** Given the parameters $w_0, w_1, \ldots, w_n$, $b > 0$, minimize $-R = -R(\lambda_0, \ldots, \lambda_n)$, with respect to its parameters $\lambda_0, \ldots, \lambda_n$, subject to the constraints:

$$\sum_{i=0}^{n} w_i \lambda_i \leq b \quad \text{and} \quad \lambda_0, \ldots, \lambda_n \geq 0.$$

In other words, given an affine constraint on the total rates, namely, the initiation rate $\lambda_0$ and the transition rates $\lambda_1, \ldots, \lambda_n$, maximize the protein translation rate. The values $w_i$, $i = 0, 1, \ldots, n$ can be used to provide different weighting to the different rates.

The motivation for the constraint (3.4) is that all the rates require common, and limited, cellular resources. For example, all tRNA molecules are transcribed by the same transcription factors (TFIIIB) and by RNA polymerase III. The constraints on amino acids, elongation/initiation factors, and aminoacyl tRNA synthetases are also similar (their concentrations affect many $\lambda_i$‘s, and they are limited, as generating them consumes significant amounts of cellular energy).

It is not difficult to show that the optimal solution $\lambda^*$ of problem 3.6 always satisfies $\lambda^* \in \text{int}(\mathbb{R}_+^{n+1})$. Theorem 3.1 implies that problem (3.6) is a convex optimization problem [27]. It thus enjoys many desirable properties.

The next result shows that increasing any of the $\lambda_i$‘s increases the translation rate.

**Proposition 3.7.** Consider the RFM with dimension $n$. Then $(\partial / \partial \lambda_i) R > 0$ for $i = 0, 1, \ldots, n$.

In other words, increasing either the initiation rate or the elongation rate at any site improves the production rate.

**Remark 3.8.** Suppose that the optimal solution $\lambda_0^*, \ldots, \lambda_n^*$ satisfies $\sum_{i=0}^{n} w_i \lambda_i < b$. Then we can increase, say, $\lambda_0^*$ a little so that the constraint (3.4) still holds. By proposition 3.7, this will increase $R$. This contradicts the assumption that the solution is optimal. Thus, an optimal solution always satisfies the constraint (3.4) with equality.

**Example 3.9.** Consider problem 3.6 for the RFM with dimension $n = 2$. In this case, $R$ is given by (3.3). Let $b = w_0 = w_1 = w_2 = 1$, i.e. the constraint is $\lambda_0 + \lambda_1 + \lambda_2 = 1$. 

![Figure 3. Steady-state translation rate $R(\lambda_0, \lambda_1)$ in the RFM with dimension $n = 1$.](http://rsif.royalsocietypublishing.org/Downloadedfrom)
Theorem 3.9. Consider problem 3.6 for the RFM with dimension $n$. Then

$$R^* = \frac{(\lambda_0)^2}{\lambda_0^2 + \frac{2}{n} \lambda_1^2} \quad (3.5)$$

In other words, the optimal translation rate $R^*$ can be computed given the optimal initiation rate and the first optimal elongation rate (and their corresponding weights in the affine constraint). This result holds regardless of the length of the transcript. It is interesting to note that several biological studies showed that various signals encoded in the 5'UTR of the transcript. It is interesting to note that several biological studies showed that various signals encoded in the 5'UTR and the beginning of the ORF can predict the protein levels of endogenous genes with relatively high accuracy [24,30,47–50].

It follows from the proof of theorem 3.9 (given in appendix A) that in order to determine $R^*$ one needs the ratio $w_i/w_{i+1}$ and the values $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}$ for some $i$. In particular, for $i = 0$ this yields the expression in (3.5).

Example 3.10. Consider again example 3.9. In this case, $w_1/w_0 = 1$, $\lambda_0^1 = 0.3008$ and $\lambda_1^1 = 0.3984$, so (3.5) yields

$$R^* = \frac{(0.3008)^2}{0.3008 + 0.3984} = 0.1294,$$

and this agrees with the result in example 3.9.

(i) Maximization with equal constraint weights

It is interesting to consider the specific case where all the weights $w_i$ in the constrained optimization problem are equal. Indeed, in this case the weights give equal preference to all the rates, so if the optimal solution satisfies $\lambda_i^* > \lambda_j^*$ for some $i, j$ then this may be interpreted as saying that, in the context of maximizing $R$, $\lambda_i$ is 'more important' than $\lambda_j$.

Figure 5 depicts the optimal values $\lambda_i$ for the case where $b = 1$ and $w_1 = 1$ for all $i$. In other words, the constraint is $\sum_{i=0}^n \lambda_i = 1$. Three cases are shown corresponding to $n = 30, 10$ and 4. The optimal values were found numerically using a simple search algorithm that is guaranteed to converge for convex optimization problems.

It may be observed that the optimal transition rates are symmetric with respect to the index $i = n/2$. In general, the transition rate $\lambda_{n/2}^*$ is larger than all other rates and the optimal values decrease as we move towards any edge of the chain. The difference between $\lambda_{n/2}^*$ and $\lambda_i^*$ (or $\lambda_j^*$) is always visible, but the difference between $\lambda_{n/2}^*$ and $\lambda_{(n/2)\pm i}$ with $i$ being small, becomes negligible as $n$ increases.

Intuitively, these results may be interpreted as follows. The importance of an elongation rate (or the corresponding site) depends on its 'centrality', or the mean distance of this site to other sites in the chain. Site $n/2$ is thus always the most 'important' site in the chain. As $n$ increases, the sites near the middle site almost have the same mean distance to the other sites, and thus become almost as important.

Figure 6 depicts the optimal translation rate $R^*$ as a function of $n$ for two different constraints: $\sum_{i=0}^n \lambda_i = n$ and $\sum_{i=0}^n \lambda_i = n^{1.03}$. The first case corresponds to the scenario where the total available resources increases linearly with $n$ (i.e. $b = n$). It may be observed that in this case the optimal translation rate $R^*$ decreases monotonically with $n$. On the other hand, increasing the total available resources by a rate which is slightly larger than a linear rate (i.e. $b = n^{1.03}$) changes the behaviour; $R^*$ in this case increases monotonically with $n$. This result suggests that in order to maintain the same optimal translation rate value as $n$ increases, the total allocated resources should increase at a rate that is slightly higher than a linear rate in $n$.

4. Discussion

The RFM is a deterministic mathematical model for translation–elongation. It can be derived via a mean-field approximation of a fundamental model from nonequilibrium statistical mechanics called TASEP. The RFM encapsulates both the simple exclusion and the total asymmetry properties of the stochastic TASEP model. The RFM is characterized by an order number $n$, corresponding to the number of sites along the mRNA strand, a positive initiation rate $\lambda_0$ and a set of positive elongation rates $\lambda_1, \ldots, \lambda_n$.

In this paper, we show that the steady-state protein translation rate $R = R(\lambda_0, \ldots, \lambda_n)$ in the RFM is a strictly concave

![Figure 4. Steady-state translation rate $R$ in the RFM with dimension $n = 2$ as a function of $\lambda_0$ and $\lambda_1$ under the constraint $\lambda_2 = 1 - \lambda_0 - \lambda_1$.](http://rsif.royalsocietypublishing.org/)
function of its (positive) parameters. This implies that: (i) a local maximum of $R$ is the global maximum (and this maximum is unique) and (ii) the problem of maximizing the steady-state protein translation rate under an affine constraint on the RFM parameters is a convex optimization problem. Such problems can be solved numerically using highly efficient algorithms. The constraint here aims to capture the limited biosynthetic budget of the cell.

We now describe the possible implications of these results in various disciplines including biology, synthetic

Figure 5. Optimal elongation rates $\lambda^*_i$ as a function of $i$ for the constraint $\sum_{i=0}^n \lambda_i = 1$. (a) $n = 30$, (b) $n = 10$ and (c) $n = 4$.

Figure 6. Optimal translation rate $R^*$ as a function of $n$ for the constraint $\sum_{i=0}^n \lambda_i = b$. (a) $b = n$ and (b) $b = n^{1.03}$.
biology, molecular evolution, and functional genomics. As mentioned above, the functional dependence of the translation rate on various variables can also be examined experimentally. A recent paper [29] studied the effect of the intracellular translation factor abundance on protein synthesis. Experiments based on a tet07 construct were used to manipulate the production of the encoded translation factor to a sub-wild-type level, and measure the translation rate (or protein levels) for each level of the translation factor(s). An analysis of fig. 1 in [29] suggests that the mapping from levels of translation factors to the translation rate is indeed concave. Our results thus provide the first mathematical support for the observed concavity in these experiments.

In synthetic biology, re-engineering gene expression is frequently used to synthesize proteins for medical and agricultural goals [12–14,51,52]. For example, in genetically modified crops new genes are introduced to the genome of the host in order to improve its resistance to certain pests/diseases or for improving the nutrient profile of the crop [52]. Another example is the commercial production of human proteins in recombinant microorganisms for therapeutic use [12–14,51]. This is sometimes based on the natural ability of certain bacteria to efficiently secrete properly folded human proteins (for example, insulin [51]). In this context, the fundamental problem is to maximize the translation rate of the heterologous gene (and thus the protein production rate) under the given constraints, e.g. the limited availability of intracellular components involved in translation. These constraints are also needed because very high initiation and elongation rates mean that the expression of the heterologous gene consumes too much of the resources of the translational machinery (e.g. ribosomes, tRNA molecules, etc.), thus significantly deteriorating the fitness of the host. In addition, very high levels of protein abundance may eventually contribute to aggregation of proteins [53,54], leading to a decrease in the yield of heterologous protein production. All these aspects are encapsulated in the convex optimization problem that is addressed here for the RFM. We believe that this mathematical problem may thus be used to provide verifiable predictions on how to efficiently manipulate the various biological factors.

There is a rich literature on using optimization theory, combined with evolutionary arguments, in biology (see, e.g. [55–57] and references therein). This approach has often been criticized, but it has undoubtedly provided insights into the process of adaptation under biological constraints, as well as helped to discriminate between alternative hypotheses for a suitable ‘fitness function’ in various biological mechanisms. Furthermore, experiments showed evolutionary adaptation of biological processes towards optimal operation levels. Examples include optimal metabolic fluxes in E. coli [58], and optimal protein expression levels from the lac operon [59]. We believe that the optimization problem posed here may lead to further progress in studying the evolution of the translation machinery.

The translation machinery is affected by mutations such as duplication/deletion of a tRNA gene or synonymous mutation affecting the codon bias usage. The concavity of the translation rate \( R \) may suggest that the selection of mutations that increase fitness indeed converges towards the optimal parameter values, as explained by the simple ‘hill climbing’ argument described above (figure 7).

Recent studies have shown that in various organisms the ribosomal density at the 5’ and 3’ ends of the ORF is higher than in the middle of the ORF [20,25,31,60]. In addition, the genomic ribosomal density is relatively constant in the middle of the ORF (usually more than 30 codons away from the two ORF ends). The elongation rate \( \lambda_i \) is negatively correlated with the ribosomal density (or the probability that a site is occupied) at site \( i \). Indeed, if \( \lambda_i \), which controls the elongation rate from site \( i \), is small, then there is a higher probability to see a ribosome in this site. Thus, these biological studies suggest that the elongation rates at the end of the chain are lower than in the middle of the chain, and that the rates near the middle are approximately equal. This agrees well with the optimal elongation rates derived based on our analysis in the case of equal weighting in the constraint (figure 5).

Our results in the case of equal weights in the constraint also show that if the total biosynthetic budget \( b(n) \) is a sub-linear or linear function of \( n \) then \( R^* \) decreases monotonically with \( n \); however, when \( b(n) \) grows faster than a linear function in \( n \) then \( R^* \) increases monotonically with \( n \). The relation between expression levels and gene length has been studied experimentally. It has been shown that in some organisms, such as humans and \( S. cerevisiae \) [3,61], expression levels tend to monotonically decrease with gene length (shorter genes have higher expression levels). However, in other organisms, such as plants [62], an opposite relation was reported (longer genes have higher expression levels). Our analysis may suggest that one should take into account not only the difference in gene length, but also the difference in the available resources of the translational machinery.

An interesting question for further research is whether the translation rate in other models of translation, including various versions of TASEP [6,8,9,31,43,63], is also a concave function of its parameters.

Another possible research direction is the design and implementation of biological experiments based on the analytical results described above. Such experiments should combine: (i) methods for manipulating the translation machinery and/or the transcript of certain gene(s) and (ii) online estimation of ribosomal density along the mRNA (e.g. using ribosome profiling [64]). The elongation rate of each codon can be estimated based on a method described in [23].
Manipulation of the translation machinery can include deletion of tRNA genes, and using the tetO7 construct to downregulate the initiation and elongation factors [29,65]. Techniques for local manipulation of a transcript include generating libraries of a certain heterologous non-functional gene (e.g. a GFP protein). In each of the variants a few mutations (relatively to the wild-type) are introduced either in the 5’UTR (corresponding to λ0) or the ORF (corresponding to λ1, λ2, . . . ), and the protein levels and ribosomal densities are measured [24,66]. The fact that the heterologous gene is non-functional to the host assures that the observed changes in translation efficiency are due to the introduced modifications.

As a specific example, one can measure the effect of small changes in the affinity constants on the translation rate of different codons (corresponding to λ1, λ2, . . . ) by introducing synonymous mutations in different parts of the ORF. We expect that a graph depicting the translation rate as a function of elongation rates will be concave (as in figure 4).

It is important to note that many times recombinant protein products are secreted. In this situation the rules for optimal translation could be fundamentally altered because of the disruption of ribosome flow following translation of the signal sequence. For example, recently it was suggested that secreted proteins undergo selection for slower translation of the N-terminal of their transcripts to enable accurate maturation and folding of these proteins [67]. This, and similar phenomena, may be addressed via the affine constraint described here by assigning larger p’s to regions that require a slower elongation rate.

Finally, the effect of single mutations in different parts of the transcript on translation rate is a fundamental question related to various biomedical disciplines. Specifically, it is known that codon substitutions in different parts of the coding sequence affect elongation and initiation rates (i.e. the λ’s) via various mechanisms (e.g. mRNA folding and adaptation to the tRNA pool [6,24,32,68]). Our result is based on linking R to the Perron root of a tridiagonal matrix that depends on the λ’s. This can serve as a starting point for a rigorous sensitivity analysis of R, i.e. analysing the effect of small changes in the λ’s on R. This topic is currently under study.

Acknowledgement. We thank the anonymous reviewers for their helpful comments.

Funding statement. This research is partially supported by research grants from the ISF and from the ElaKodesz Institute for Medical Engineering and Physical Sciences.

Appendix A. Proofs

Proof of theorem 3.1. The proof consists of the following steps:

1. Expressing the term on the right-hand side of (2.7) as a ratio between two polynomials p(R) and v(R).
2. Linking the numerator polynomial p(R) to the determinant of a symmetric, non-negative tridiagonal matrix A whose entries depend on the λ’s.
3. Proving that R−1/2 is the largest eigenvalue of the matrix A.
4. Using the properties of the largest eigenvalue of a symmetric, non-negative matrix to show that R is a strictly concave function of its parameters.

Step 1: Define

\[ g_{n+1}(z, \lambda_0, \ldots, \lambda_n) := 1 - \frac{z/\lambda_0}{1 - \frac{z/\lambda_1}{1 - \frac{z/\lambda_2}{\cdots - \frac{z/\lambda_{n-1}}{1 - z/\lambda_n}}}} \]  

(A 1)

Then we can rewrite (2.7) as

\[ g_{n+1}(R, \lambda_0, \ldots, \lambda_n) = 0. \]  

(A 2)

By the theory of convergents of continued fractions [40] it follows that

\[ g_{n+1}(z, \lambda_0, \ldots, \lambda_n) = \frac{p_{n+1}(z, \lambda_0, \ldots, \lambda_n)}{v_{n+1}(z, \lambda_1, \ldots, \lambda_n)}, \]  

(A 3)

where \( p_{n+1} \) and \( v_{n+1} \) are defined recursively by

\[
\begin{align*}
\left\{
\begin{array}{l}
p_{-1}(z) = 1, \\
& \quad p_0(z) = 1 \\
\end{array}
\right.
\quad \text{and} \quad
\left\{
\begin{array}{l}
p_k(z) = p_{k-1}(z) - z\lambda_{k-1}^{-1} p_{k-2}(z), & k \geq 1 \\
& \quad v_0(z) = 1 \\
\end{array}
\right.
\quad \text{and} \quad
\left\{
\begin{array}{l}
v_{-1}(z) = 0, \\
& \quad v_k(z) = v_{k-1}(z) - z\lambda_{k-1}^{-1} v_{k-2}(z), & k \geq 1.
\end{array}
\right.
\]

(A 4) (A 5)

For example, for \( n = 2 \), equation (A 1) yields

\[ g_3 = 1 - \frac{z/\lambda_0}{1 - \frac{z/\lambda_1}{1 - \frac{z/\lambda_2}{1 - z/\lambda_3}}}, \]

\[ = \frac{(\lambda_0^{-1} \lambda_1^{-2} z^2 - (\lambda_0^{-1} + \lambda_1^{-1} + \lambda_2^{-1})z + 1}{1 - (\lambda_1^{-1} + \lambda_2^{-1})z}, \]

whereas (A 4) and (A 5) yield

\[ p_3 = (\lambda_0^{-1} \lambda_2^{-1})z^2 - (\lambda_0^{-1} + \lambda_1^{-1} + \lambda_2^{-1})z + 1, \]

and

\[ v_3 = 1 - (\lambda_1^{-1} + \lambda_2^{-1})z. \]

Note that (A 4) and (A 5) imply that \( p_k = p_k(z, \lambda_0, \ldots, \lambda_{k-1}) \) and \( v_k = v_k(z, \lambda_1, \ldots, \lambda_{k-1}) \), and

\[ v_{k+1}(z, \lambda_1, \ldots, \lambda_n) = p_k(z, \lambda_0, \ldots, \lambda_{k-1}). \]  

(A 6)

Furthermore, (2.6) and (2.5) yield

\[ v_k = \frac{p_k(R, \lambda_0, \ldots, \lambda_{k-1})}{p_{k-1}(R, \lambda_0, \ldots, \lambda_{k-2})}, \quad k = 1, \ldots, n, \]  

(A 7)

so

\[ p_k = \frac{p_k}{p_{k-1}} \frac{p_{k-1}}{p_{k-2}} \cdots \frac{p_2}{p_1} \frac{p_1}{p_0}, \]

(A 8)

Since \( v_1 \in (0, 1) \) it follows that \( p_k(R) \in (0, 1) \) for all \( k = 1, 2, \ldots, n. \)

From (A 2) and (A 3), it follows that

\[ \frac{p_{n+1}(R, \lambda_0, \ldots, \lambda_n)}{v_{n+1}(R, \lambda_1, \ldots, \lambda_n)} = 0. \]  

(A 9)

Suppose for a moment that \( v_{n+1}(R, \lambda_1, \ldots, \lambda_n) = 0. \) Then (A 6) yields \( p_k(R, \lambda_0, \ldots, \lambda_{k-1}) = 0, \) and combining this with (A 7) yields \( v_0 = 0. \) This is a contradiction, as \( e \in \text{int}(C^0). \) We conclude that the denominator in (A 9) is not zero, so (A 9) is
where

To relate the polynomial $p_k$ to a tridiagonal matrix, define the polynomials

$$q_0(s) := s^{k+1} p_k(s^{-2}) \quad k = -1, 0, \ldots, n + 1,$$

where $s \in \mathbb{R} \setminus \{0\}$. Then (A 4) yields

$$q_0(s) = s^{k+2} - (\lambda_k^2 + 1) q_k(s),$$

and

$$q_{k+1}(s) = s q_k(s) - q_{k-1}(s), \quad k \geq 1.$$

Define a $(n + 2) \times (n + 2)$ Jacobi matrix $A = A(\lambda_0, \ldots, \lambda_n)$ by

$$A :=
\begin{bmatrix}
0 & \lambda_0^{-1/2} & 0 & \cdots & 0 \\
-\lambda_0^{-1/2} & 0 & \lambda_1^{-1/2} & \cdots & 0 \\
0 & -\lambda_1^{-1/2} & 0 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 \\
0 & 0 & \cdots & -\lambda_{n-1}^{-1/2} & 0 \\
0 & 0 & \cdots & 0 & -\lambda_n^{-1/2}
\end{bmatrix}.
$$

Let $I_{n+2}$ denote the $(n + 2) \times (n + 2)$ identity matrix. Then

$$sI_{n+2} - A$$

and it is straightforward to verify that the determinant of the $(l + 1) \times (l + 1)$ leading principal minor of $sI_{n+2} - A$ is $q_l(s)$. In particular, $q_{n+1}(s) = \det(sI_{n+2} - A)$. Combining (A 10) and (A 11) implies that $q_{n+1}(R^{-1/2}) = 0$, so $R^{-1/2}$ is an eigenvalue of the matrix $A$.

Step 3: Recall that the spectral radius of a square matrix is the maximum over the absolute values of its eigenvalue. The spectral radius of a non-negative matrix is an eigenvalue of the matrix called the Perron root [70]. The next result shows that $R^{-1/2}$ is the largest eigenvalue of the non-negative matrix $A$.

**Proposition A.1.** The Perron root of the matrix $A$ is $R^{-1/2}$. 

**Proof of proposition A.1.** It follows from known results on Jacobi matrices [70, ch. 0] that all the eigenvalues of the matrix $A$ are real and distinct, and that if we order them as

$$\alpha_1 < \alpha_2 \leq \cdots \leq \alpha_{n+2},$$

then the number of sign changes in the sequence

$$\{q_0(\alpha_1), \ldots, q_0(\alpha_{n+2})\}$$

is $n + 2 - j$. Let $i$ be the index such that $\alpha_i = R^{-1/2}$. By (A 11), $q_0(\alpha_i) = R^{-1/2} + \beta_i P(R)$, and (A 8) yields $q_k(\alpha_i) > 0$ for all $k = 0, 1, \ldots, n$. Thus, the number of sign changes in the sequence

$$\{q_0(\alpha_1), \ldots, q_0(\alpha_{n+2})\}$$

is zero, so $i = n + 2$, i.e. $\alpha_{n+2} = R^{-1/2}$. 

Step 4: Given a vector $x = [x_0, \ldots, x_n] \in \mathbb{R}^{n+1}$, let $T(x)$ denote the $(n + 2) \times (n + 2)$ tridiagonal matrix whose main diagonal is zero, and sub- and super-diagonals are the vector $x$. Note that this matrix is non-negative and irreducible. Let $s_i = s(T(x)), i = 1, \ldots, n + 2$ denote the eigenvalues of $T(x)$ ordered so that

$$s_1 < s_2 \leq \cdots < s_{n+2}.$$ 

We already know that $s_{n+2}(T(x)^{-1/2}) = R^{-1/2}(x)$. Note that the matrix $A$ in (A 13) can be written as $A = T(\lambda^{-1/2})$, where

$$\lambda^{-1/2} := [\lambda_0^{-1/2}, \ldots, \lambda_n^{-1/2}] = [\lambda^{-1/2}, \ldots, \lambda^{-1/2}].$$

Pick $x, y \in \mathbb{R}^{n+1}$, with $x \neq y$, and $k \in (0, 1)$. Let $u := kx + (1 - k)y$. Then

$$R(u) = s_{n+2}^2(\lambda^{-1/2}(kx^{-1/2}) + (1 - k)y^{-1/2}).$$

The function $u(w) := w^{-1/2}$ is strictly convex on $w \in (0, \infty)$, so

$$u^{-1/2} < k\lambda^{-1/2} + (1 - k)y^{-1/2},$$

where the inequality between the vectors should be interpreted component-wise. Since $T(u^{-1/2})$ is irreducible, this implies that [71, ch. 8]

$$s_{n+2}(T(u^{-1/2})) < s_{n+2}(Q),$$

where $Q := T(k\lambda^{-1/2} + (1 - k)y^{-1/2})$. Combining this with (A 14) yields

$$R(u) > s_{n+2}^2(Q).$$

Since $Q$ is non-negative and symmetric, its induced 2-norm is equal to its spectral radius $s_{n+2}(Q)$. Using the fact that a norm is always convex, it is straightforward to see that the norm-squared $s_{n+2}^2(Q)$ is also convex, so

$$s_{n+2}^2(Q) = s_{n+2}^2(kT(x^{-1/2}) + (1 - k)T(y^{-1/2}))$$

$$\leq k^2 s_{n+2}^2(T(x^{-1/2}) + (1 - k)s_{n+2}^2(T(y^{-1/2})).$$

Combining this with (A 15) yields

$$R(u) > \max \left\{ \frac{1}{k^2 s_{n+2}^2(T(x^{-1/2})), \frac{1}{s_{n+2}^2(T(y^{-1/2})}, \frac{1}{s_{n+2}^2(Q)} \right\}.$$ 

Now (A 14) implies that

$$R(kx + (1 - k)y) > \min \{R(x), R(y)\},$$

i.e. $R$ is strictly quasi-concave on int($\mathbb{R}^{n+1}$).

Pick $t \in (0, 1)$, and let $\mu := R(x)/R(y) = (1 - t)R(y)$. Then $\mu \in (0, 1)$, so (A 16) and the homogeneity of $R$ (see Fact 1 above) yield

$$R \left( \frac{\mu}{R(x)} \right) + (1 - \mu) \frac{y}{R(y)} > \min \left\{ \frac{x}{R(x)}, \frac{y}{R(y)} \right\} = 1.$$

Thus, 

$$R \left( \frac{tx + (1 - t)y}{R(x) + (1 - t)R(y)} \right) > 1.$$

Using the homogeneity of $R$ again gives

$$R(tx + (1 - t)y) > R(x + (1 - t)y),$$

and this completes the proof of theorem 3.1. 

**Proof of proposition 3.7.** Let $v = [v_1, \ldots, v_{n+2}] \in \mathbb{R}^{n+2}$ denote a Perron eigenvector of the symmetric matrix $A$, i.e. an eigenvector corresponding to the Perron root $R^{-1/2}$. It follows from
known results [72] that
\[
\frac{\partial}{\partial \lambda_i} R^{1/2} = \frac{v'}{v} R^{1/2} \left( \frac{\partial}{\partial \lambda_i} \log R \right),
\]
and combining this with (A 13) yields
\[
\frac{\partial}{\partial \lambda_i} R = \frac{2R^{3/2}}{\lambda_i^{1/2} v'} \sum_{j=1}^n v_j \lambda_{j/2}.
\] (A 17)
Since all the components of the Perron eigenvector are strictly positive [71], this implies that \(\frac{\partial}{\partial \lambda_i} R > 0\) for all \(i = 0, \ldots, n\).

Proof of theorem 3.9. The proof is based on formulating the Lagrangian function associated with problem 3.6 and determining the optimal parameter values by equating its derivatives to zero [27]. The Lagrangian is
\[
L(\lambda_0, \ldots, \lambda_n, \theta) := R(\lambda_0, \ldots, \lambda_n) + \theta \left( b - \sum_{i=0}^n \omega_i \lambda_i \right),
\]
where \(\theta \in \mathbb{R}\) is the Lagrange multiplier. Differentiating this with respect to \(\lambda_i\) and equating to zero yields
\[
\theta \omega_i = \left( \frac{\partial}{\partial \lambda_i} \log R \right) |_{\omega},
\]
where \(\omega\) means that the equation holds once the optimal values \(\lambda^*_i\) are substituted. This implies in particular that
\[
\frac{\omega_0}{\omega_1} = \left( \frac{\partial}{\partial \lambda_0} \log R \right) |_{\omega},
\]
and combining this with (A 17) yields
\[
\frac{\omega_0}{\omega_1} = \left( \frac{\lambda^*_0}{\lambda^*_1} \right)^{3/2} R^1 v^1_{i}.
\] (A 18)
where \(v^*\) is the unique (up to scaling) positive Perron eigenvector of the non-negative and irreducible matrix \(A^* := A(\lambda^*_0, \ldots, \lambda^*_n)\) [71, ch. 8]. The equation \(A^* v^* = (R^*)^{-1/2} v^*\) yields
\[
\begin{align*}
(\lambda^*_0)^{-1/2} v^*_1 + (\lambda^*_1)^{-1/2} v^*_2 &= (R^*)^{-1/2} v^*_1, \\
(\lambda^*_1)^{-1/2} v^*_2 + (\lambda^*_2)^{-1/2} v^*_3 &= (R^*)^{-1/2} v^*_2, \\
& \vdots \\
(\lambda^*_n)^{-1/2} v^*_{n+1} &= (R^*)^{-1/2} v^*_{n+2}.
\end{align*}
\] (A 19)
Thus
\[
(\lambda^*_0)^{-1/2} v^*_1 = ((R^*)^{-1} (\lambda^*_0)^{1/2} - (\lambda^*_0)^{-1/2}) v^*_1,
\]
and substituting this in (A 18) and simplifying yields (3.5).

Endnote

1 In previous papers on the RFM, the notation \(\lambda\) was used to denote the initiation rate. Here, we use \(\lambda_0\) as this leads to a more consistent notation.

References


