Probabilistic Boolean Network Modelling and Analysis Framework for mRNA Translation

Yun-Bo Zhao and J. Krishnan

Abstract—mRNA translation is a complex process involving the progression of ribosomes on the mRNA, resulting in the synthesis of proteins, and is subject to multiple layers of regulation. This process has been modelled using different formalisms, both stochastic and deterministic. Recently, we introduced a Probabilistic Boolean modelling framework for mRNA translation, which possesses the advantage of tools for numerically exact computation of steady state probability distribution, without requiring simulation. Here, we extend this model to incorporate both random sequential and parallel update rules, and demonstrate its effectiveness in various settings, including its flexibility in accommodating additional static and dynamic biological complexities and its role in parameter sensitivity analysis. In these applications, the results from the model analysis match those of TASEP model simulations. Importantly, the proposed modelling framework maintains the stochastic aspects of mRNA translation and provides a way to exactly calculate probability distributions, providing additional tools of analysis in this context. Finally, the proposed modelling methodology provides an alternative approach to the understanding of the mRNA translation process, by bridging the gap between existing approaches, providing new analysis tools, and contributing to a more robust platform for modelling and understanding translation.

Index Terms—mRNA translation, probabilistic Boolean networks, parallel update, recycle, regulation, sensitivity

1 Introduction

mRNA translation is a crucial step of protein synthesis in cells. This process involves the production of the protein from its corresponding mRNA, and is usually conceptualized and described in terms of ribosome movement along the mRNA. This movement consists of three main stages: first, the ribosome attaches to the mRNA to start the translation (initiation), then the ribosome reads the codons (triplet of nucleotides) on the mRNA to build the nascent peptide chain (elongation) and finally the ribosome leaves the mRNA releasing the completed peptide chain (termination) [1], [2]. Overlaid upon this basic process are other layers of regulatory complexity. The combination of stochastic multi-ribosome transport, regulatory complexity and other factors such as resource limitation contribute to the intrinsic complexity of this process.

mRNA translation can be mathematically described by various models. These models describe translation at different levels of abstraction and can be either deterministic or stochastic [3], [4], [5]. At one extreme models of translation may describe this process as a single ODE. At the other extreme, the various individual movement steps of the ribosome as well as various biochemical (and even mechanochemical) steps can be described explicitly [6]. Depending on the goals of modelling, different stages may be described in a very coarse-grained fashion, and the resulting models can still capture essential aspects of the underlying process, and make predictions which are mirrored experimentally [7], [8], [9]. In this paper we will focus on codon-based models which aim at a description of ribosomal progression along the mRNA, incorporating codon-based effects: this may be regarded as a basic but representative common-denominator for mRNA translation, incorporating its basic ingredients. In the codon-based models as depicted in Fig. 1, the mRNA is represented by a one-dimensional lattice with each site being one codon on the mRNA, and the dynamics of the process conceptually consists of three different types of events, corresponding to the three stages in mRNA translation. These events represent the entry of the ribosome from the leftmost end of the mRNA (initiation), the hops of the ribosome one codon at a time to the right (elongation) and the exit of the ribosome from the rightmost end of the mRNA (termination), where the occurrence of each event is associated with a stochastic rate [10], [11]. For a ribosome to attach the mRNA, the first r codons (which is the number of codons that a ribosome covers as it moves along the mRNA) must be empty, and for the ribosome to exit, its head must be at the last codon of the mRNA. Three distinct features are observed from this model. First, each codon on the mRNA can be occupied by no more than one ribosome. Second, the ribosome can hop in only one direction. Third, multiple non-overlapping ribosomes can be on the mRNA simultaneously [12]. This unidirectional, non-overlapping ribosome movement ("traffic") of the codon-based model can be mathematically described by ODEs, Totally Asymmetric Simple Exclusion Process (TASEP) [13], [14], [15], Petri net [16] or other possible approaches. In particular, the TASEP model has been validated extensively with experimental data, for *S. cerevisiae* [17], *E. coli* and

- *Y.-B. Zhao is with the Department of Chemical Engineering, Centre for Process Systems Engineering, Imperial College London, South Kensington, London SW7 2AZ, UK, and the Department of Automation, Zhejiang University of Technology, Hangzhou 310023, China. E-mail: dr.y.zhao@ieee.org.*
- *J. Krishnan is with the Department of Chemical Engineering, and Centre for Process Systems Engineering, and Institute for Systems and Synthetic Biology, Imperial College London, South Kensinton, London SW7 2AZ, UK. E-mail: J.Krishnan@imperial.ac.uk.*

Manuscript received 5 Sept. 2014; revised 2 Feb. 2015; accepted 20 May 2015. Date of publication 14 Sept. 2015; date of current version 4 Aug. 2016.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TCBB.2015.2478477

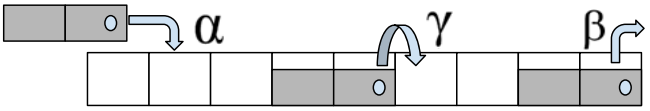


Fig. 1. The codon-based lattice model for mRNA translation, where the ribosome covers two codons on the mRNA. The entry and exit of the ribosome are assumed to be instantaneous. α, β, and γ are initiation, termination, and elongation rate constants, respectively.

B. subtilis [18], and so forth. These various models aim to understand such features as steady state translation rates, ribosomal distributions, average coverage and their dependence on different parameters.

Each of these modelling formalisms can be used effectively to make predictions in individual contexts. Furthermore, each formalism possesses certain advantages and limitations (discussed in detail in [19]). For instance, ODE models allow for the use of dynamical systems tools and bifurcation analysis, but do not impose strict exclusion, stochastic models enforce strict exclusion but are generally less tractable to effectively dissect. This motivates the need for modelling and analysis tools which can retain the features of the stochastic models but dissect them better.

Within the codon-based modelling framework, we recently proposed a Probabilistic Boolean Network (PBN) modelling and analysis framework [19]. The PBN model shares the same model structure as TASEP (as illustrated in Fig. 1) but treats the model structure from a different perspective by modelling the codon status as Boolean variables. It is a codon-based, stochastic model, allowing for multi-codon coverage (e.g., a ribosome covers two codons in Fig. 1). Importantly the PBN framework allows for the use of tools of analysis of Boolean networks to obtain exact numerical solutions of the steady state distribution of the mRNA, without simulation.

Recent experimental methods such as the advanced ribosomal profiling technique, allow for a more detailed and comprehensive study of mRNA translation. Such techniques can provide information about ribosomal presence and dynamic progression along the mRNA [20], [21], [22], [23]. In addition many other aspects about the progression of ribosomes are being uncovered, including codon dependent elongation rates, and different regulatory mechanisms. In such a setting, tools which can allow for a systematic analysis of different aspects of ribosomal progression and their effect of translation become increasingly important. Current tools include comprehensive, and coarse-grained stochastic simulations as well as analytical and numerical methods of dissecting mean-field models. However, there is currently the need to analyse different aspects of ribosomal progression in translation, which requires understanding how different parameters and factors affect translation. There is thus the need to go well beyond just steady state codon coverage density and translation rate, and at the same time understand how the stochastic aspects of translation are sensitive to parameter changes. Thus tools which allow for dissection of stochastic behaviour and dependence on parameters are valuable and the PBN model provides new tools in this regard.

In fact, even for the calculation of the steady state codon coverage density and translation rate, the PBN model and associated analysis methods provides an alternative/additional tool. These steady state profiles can usually be calculated using stochastic simulations and mean-field TASEP (analytical or numerical). On the one hand, the essential stochastic effects in mRNA translation necessitate a large number of sample simulations to be run for the steady state analysis, which can be very time consuming, while the statistical average from sample simulations does not have absolute guarantee of accuracy [24], [25], and assessing how the results depend on parameters and auxiliary factors in this manner is also time consuming. On the other, analytical solutions to the mean-field TASEP are usually difficult, and either approximate analytical or numerical solution is required. Different from either time-consuming stochastic simulations or approximate analytical solutions, the PBN model provides a numerical but exact solution (without resorting to simulations) to the codon-based stochastic model for mRNA translation. Then, as an complementary analysis tool, the PBN model provides an alternative capability of analysing the mRNA translation process. Some of this analysis can be difficult for other approaches, as will be discussed later.

In order to interpret experimental data, a coarse-grained TASEP-like model, called the ribosome flow model (RFM), was proposed [26], which can either be stochastically simulated or analytically treated using the mean-field method. Since both the PBN model and the RFM model share the same model structure, the PBN model can be used in a similar coarse-grained way as well. Again, due to the stochastic nature and the exact numeral calculation of the stochastic distribution, the PBN model provides a useful complementary tool to the RFM model at the coarse-grained level, just like its complementary role to simulations and mean-field approaches in analysing TASEP models.

Overall the PBN model provides an alternative approach to the understanding of the complex mRNA translation process. As illustrated in Fig. 2, the modelling and analysis range from simple lumped ODE models, to comprehensive detailed models incorporating biochemical reactions, ribosomal progression and other factors. Two approaches exist for the comprehensive understanding of the complexities of mRNA translation. One involves stochastic simulations of the coarse-grained RFM or the codon-based TASEP, and the other is deterministic analysis of such models using mean-field or other kinds of approximation. The PBN based approach allows for the bridging of the gap between these two approaches. Combined with the other approaches, it provides a more unified approach for dissecting the complexity of translation.

In the context of the codon-based models for mRNA translation, one has to specify the movement order, referred to as the update rules, for the multiple ribosomes on the mRNA. This is an important factor, since different update rules can result in important differences in the resulting steady state of the mRNA translation process [16], [19], [27]. The proposed PBN model has been developed for only one often-used rule called the random-sequential update rule. It has been realized that another rule, referred to as the parallel update rule, is also reasonable to mimic the mRNA translation process, and can be even more realistic in certain respects [16], [19]. This fact justifies one of our purposes in this work to extend the PBN model to the parallel update rule. With this expanded PBN modelling framework, we

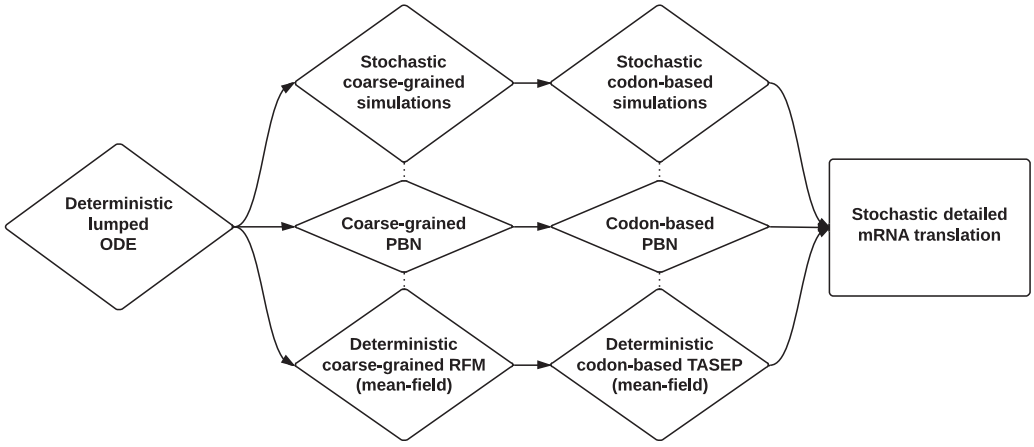


Fig. 2. A spectrum of models of translation ranging from simple lumped ODE to detailed stochastic models accounting for many factors is depicted. The coarse-grained ribosome flow model and the codon-based TASEP provide an approach to understanding the stochastic, complex mRNA translation process, based on deterministic analysis (mean-field approximation) or stochastic simulations. The PBN model can also be used at both the coarse-grained and codon-based levels, thus providing a complementary approach, which is stochastic, along with tools for direct computation of steady states.

then demonstrate the aforementioned advantages of the PBN model by various example applications. Finally, the extension of the PBN model to allow for the parallel update rule completes the general framework of the PBN modelling methodology.

The remainder of the paper is organized as follows. For completeness we first introduce the basics of the PBN model for the random-sequential update rule in Section 2. The PBN model for the parallel update rule follows in Section 3. Then, Section 4 discusses various contexts building on the basic translation process, where the PBN model provides valuable information, not easily obtained from other approaches, and Section 5 concludes the paper.

Notation. The total codons on the mRNA and the number of the codons that a ribosome covers are represented by n and r, respectively. The state of the n codons on the mRNA, or the "mRNA state", is denoted by a vector $x = [x_1 \ \dots \ x_n]$ where $x_i = 1$ means the ith codon is occupied by a ribosome and $x_i = 0$ means the ith codon is empty. An event is identified by the position of the head of the ribosome when the event occurs. The set of the possible events is then $\mathbb{E} := \{e_i : i \in \mathbb{I}_e\}$ with the index set being $\mathbb{I}_e := \{0, r, r+1, \dots, n\}$, i.e., e_0 the ribosome entry, e_n the ribosome exit, and $e_i, i = r, r+1, \dots, n-1$ the ribosome hopping from codon i to $i+1$. Each event is associated with a rate, denoted by α for the entry, β for the exit and $\gamma_i, i = r, \dots, n-1$ for the hops, respectively. We also use γ_0 for α, γ_n for β, $\Gamma := \{\gamma_i : i \in \mathbb{I}_e\}$ being the set of the event rates, for simplicity of notation.

2 Preliminaries on PBN with Random-Sequential Update

At the outset, we mention that the basic mRNA translation process can be described in a Boolean setting. This is quite natural as the coverage status of a codon is a Boolean variable. The translation process can then be cast in the framework of a Boolean network, with Boolean variables, rules which define the transition between states and parameters which define various stochastic transition rates. This has been discussed in [19] and the main details for representing the basic translation process in this setting are described in the Appendix, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TCBB.2015.2478477. With this as a backdrop, we now discuss how the mRNA translation process cast in this form is simulated. This involves describing how this Boolean network model is updated.

2.1 Simulating mRNA Translation with the Random-Sequential Update Rule

The central aspects of the codon-based model are the event rates, which are interpreted as follows. Within a sufficiently small time interval dt, the probability of event e_i to occur is $\gamma_i dt$ if the mRNA state at the time allows such an event to occur; otherwise the probability is 0. Therefore, for mRNA state x, the actual occurrence rate of event e_i, denoted by $p_{e_i}(x)$, is dependent not only on the event rate γ_i, but the event occurrence probability, denoted by $\psi_i(x)$, thus leading to the following rate law,

$$p_{e_i}(x) = \psi_i(x)\gamma_i, \tag{1}$$

where the event occurrence probability is determined by the mRNA state x, as follows,

$$\psi_i(x) = \begin{cases} P\{x_j = 0, j = 1, 2, \dots, r\}, & i = 0 \\ P\{x_n = 1\}, & i = n \\ P\{x_i = 1, x_{i+1} = 0\}, & i \neq 0, n. \end{cases} \tag{2}$$

The rate law itself does not completely define the codon-based model. Indeed, the model is updated in discrete time steps and within one time step, more than one update event can be allowed since multiple ribosomes are on the mRNA simultaneously. This fact implies that an order of the update events, termed as the "update rule", has to be specified.

The popular random-sequential update rule assumes no particular order of the update events, that is, at any time step, the update event is chosen randomly with equal probability from all the update events in $\mathbb{E}$. Its simulation algorithm is shown in Algorithm 1 where $p_i^a := \gamma_i / \sum_{i \in \mathbb{I}_e} \gamma_i$ and $dt_a := 1 / \sum_{i \in \mathbb{I}_e} \gamma_i$. For more details of the algorithm and its alternatives, the reader is referred to [19].

Algorithm 1. Simulating mRNA Translation with the Random-Sequential Update Rule

1) Given n, r and the event rates in Γ: entry rate α, exit rate β, and hopping rates $\gamma_i, i = r, \ldots, n-1$.
2) At time t, simulate a random number according to the discrete distribution with the probability being $\{p_i^a\}$, and determine the next event index i.
   a) If $i = 0$. A new ribosome enters and occupies codons from 1 to r if they are empty; otherwise do nothing;
   b) $i = n$. The last r codons become empty if they are occupied; otherwise do nothing;
   c) $i \neq 0, n$. The ribosome hops from i to $i+1$ if codon i is occupied and $i+1$ empty; otherwise do nothing.
3) Let $t = t + dt_a$ and repeat Step 2.

2.2 PBN as a Model for mRNA Translation with the Random-Sequential Update Rule

Since the state of each codon on the mRNA is a Boolean variable which can be either 1 (occupied) or 0 (empty), the mRNA state is a Boolean vector, whose dynamics is governed by the update events, i.e., entry, hopping and exit of the ribosome. Notice that each ribosome occupies consecutively r codons, thus making the possible mRNA states (termed as the "effective mRNA states") much less than 2^n which is the number of the states of a Boolean network with n nodes. For example, for $n = 3$, $r = 2$, only three mRNA sates are possible out of the eight Boolean states, i.e., [1 1 0], [0 1 1] and [0 0 0], where we have assumed that the entry and exit of the ribosome are instantaneous and this assumption is held throughout the paper. Therefore, our discussion in what follows applies only to the effective mRNA states, denoted by $\mathbb{M}(n, r)$ and $m_{n,r} := |\mathbb{M}(n, r)|$ is the number of the effective mRNA states. We use $\mathbb{M}$ and m for simplification of notation.

Using a tool called the semi-tensor product for Boolean networks [28], [29], [30], the update events can be fully and uniquely described by a linear system (see Appendix, available in the online supplemental material), as follows,

$$x(t+1) = L_t x(t), x(t) \in \mathbb{M}, \tag{3}$$

where $x(t)$ represents the mRNA state at time t and the matrix L_t, referred to as the structure matrix, determines the mRNA state update according to the update event. Since there are $n - r + 2$ possible update events in total, L_t is therefore chosen probabilistically as follows, according to Algorithm 1,

$$P\{L_t = L_i\} = p_i^a, i \in \mathbb{I}_e. \tag{4}$$

Take the expectation of (3), we have $Ex(t+1) = L_S Ex(t), x(t) \in \mathbb{M}$, where L_S, referred to as the transition matrix, is defined by

$$L_S := \sum_{i \in \mathbb{I}_e} p_i^a L_i. \tag{5}$$

The transition matrix L_S can be reduced to only the effective mRNA states, thus leading to the following simplified version,

$$M_S = L_S|_{\mathrm{Row}(L_S) \in \mathbb{M},\ \mathrm{Col}(L_S) \in \mathbb{M}}, \tag{6}$$

where M_S is obtained by deleting from L_S all the columns and rows that do not belong to $\mathbb{M}$. The transition matrix, L_S or M_S then fully describes the mRNA translation dynamics and thus defines the PBN model for the random-sequential update rule.

For more details of the PBN model for the random-sequential update rule, the reader is referred to Appendix, available in the online supplemental material, and [19].

3 PBN with Parallel Update

3.1 The Parallel Update Rule

Unlike the random-sequential update rule where the update event is chosen from all the possible update events leading to events which are not completely independent from each other, with the parallel update rule, the ribosomes move independently from each other. Therefore, within one time step, the random-sequential update rule updates only one chosen event, while with the parallel update rule, all the allowed events are updated simultaneously by their individual update probabilities.

mRNA translation with the parallel update rule can be simulated as in Algorithm 2, where the allowed events are updated sequentially from the left to the right within a single time step, since the parallel update rule is equivalent to this ordered-sequential update rule from the left [27]. For mRNA state x, the probabilities of the allowed event e_i being updated is given as $p_i^b(x) = \frac{\gamma_i}{\max_{i \in \mathbb{I}_e} \gamma_i}$. Then the rate law implies that $\gamma_i dt_b = p_i^b(x)$, which further leads to the time step of $dt_b = \frac{1}{\max_{i \in \mathbb{I}_e} \gamma_i}$.

The above definitions of $p_i^b(x)$ and dt_b ensure that for given mRNA state x, the events that are allowed and not allowed to be updated are actually updated with rates $p_i^b(x)/dt_b = \gamma_i$ and 0, respectively, and therefore this update rule is still consistent with the rate law in (1), though it is clearly different from the random-sequential update rule.

Algorithm 2. Simulating mRNA Translation with the Parallel Update Rule

1) Given n, r and the event rates in Γ: entry rate α, exit rate β, and hopping rates $\gamma_i, i = r, \ldots, n-1$.
2) At time t,
   a) Check the current state $x(t)$ to obtain all the possible update events, $\mathbb{I}_e(x)$;
   b) From the left to the right, update event $e_i \in \mathbb{I}_e(x)$ with probability $p_i^b(x)$.
3) Let $t = t + dt_b$ and repeat Step 2.

3.2 PBN as a Model for mRNA Translation with the Parallel Update Rule

An important feature of the PBN model with the random-sequential update rule is that the choice of the structure matrix L_t (the update event) is independent from the mRNA state, as shown in (3) and (4). At each time step, the update event is chosen with fixed probability, and whether it is actually updated or not is determined by the structure matrix and the mRNA state. This algorithm structure gives the

constant transition matrices $L_i, i \in \mathbb{I}_e$, and then enables the mean dynamics with L_S and M_S being defined in (5) and (6) in a relatively straightforward way.

With the parallel update rule, however, the update varies with the mRNA state. Indeed, for mRNA state x, denote by $\mathbb{I}_e^0(x)$ the index set of the allowed events, and then $\mathbb{I}_e^0(x)$ will be dependent on specific x and may contain multiple update events. With the parallel update rule, any update event in $\mathbb{I}_e^0(x)$ is possibly (or not) updated. Therefore, the actual update events within the time step can be any arbitrary combination of the $k_x := |\mathbb{I}_e^0(x)|$ possible update events, thus giving 2^{k_x} possible mRNA states at the next time step, with no fixed structure matrix analogous to L_t in the random-sequential case being defined.

To define the PBN model for the parallel update rule, we employ the underlying mathematical structure of the PBN model as a Markov chain. Indeed, the PBN model for the random-sequential update rule relies on the definition of the transition matrix M_S in (6). M_S defines the transition probability between any two effective mRNA states in $\mathbb{M}$, and consequently defines the Markov chain for the random-sequential update rule. For the parallel update rule, we notice that the effective mRNA states remain unchanged. Therefore, the PBN model for the parallel update rule can be defined provided a similar transition matrix analogous to M_S can be given. In what follows, we define the transition matrix directly from the underlying Markov chain, that is, define the transition probability between any two effective mRNA states for the parallel update rule, in the absence of the linear system description in (3) and (4) for the random-sequential update rule. This is done, for any given effective mRNA state, by finding all the possible mRNA states transferable by one step of parallel update and their corresponding probabilities.

To that end, let $\Phi(x)$ be the set of all the k_x-dimensional vector with each element being only 0 or 1. Then any $\phi := [\phi_1 \dots \phi_{k_x}] \in \Phi$ defines an event for the parallel update rule, where $\phi_i = 1$ means the ith possible event is actually updated and $\phi_i = 0$ the ith possible event being not chosen. Similar to the update equation for the random-sequential update rule as in (3), the update equation for the parallel update rule can be written as follows,

$$x(t+1) = L_x^\phi x(t), x(t) \in \mathbb{M}, \phi \in \Phi, \tag{7}$$

where L_x^ϕ is defined by the product of all the transition matrices whose element in ϕ is 1, that is,

$$L_x^\phi := \prod_{i=1}^{k_x} (L_{j_i})^{\phi_i}, \tag{8}$$

where j_i is the ith element in $\mathbb{I}_e^0(x)$. The matrix L_x^ϕ represents the composite transition matrix which results from a parallel update in one time step.

Since the ribosome movements are independent, the probability of the update event being described by L_x^ϕ is then given by

$$p_x^\phi := P\{L_x^p = L_x^\phi\} = \prod_{i=1}^{k_x} (p_{j_i}^b)^{\phi_i} (1 - p_{j_i}^b)^{1-\phi_i}. \tag{9}$$

Unlike the random-sequential update rule in (3), (4) and (5), for the parallel update rule described in (7), (8) and (9), no similar transition matrix L_S as in (5) can be directly given, since L_x^ϕ and p_x^ϕ are dependent on x for the parallel update rule.

However, from the perspective of a Markov chain, (7), (8) and (9) mean that for any effective mRNA state x, at the next step it switches to $L_x^\phi x$ with probability p_x^ϕ. With this, we are able to construct the following transition matrix, similar as M_S in (6) for the random-sequential update rule,

$$M_P := (p_{ij})_{2^m \times 2^m}, \tag{10}$$

where $p_{ij} := P\{j = L_x^\phi x | i = x\} = p_x^\phi, x \in \mathbb{M}$ denotes the probability that an mRNA state of index i is transformed into one of index j.

The transition matrix M_P then defines the PBN model for the parallel update rule, which is described by PBN, and essentially based on a time-homogeneous Markov chain governed by the transition matrix M_P.

3.3 The Exact Steady State Distribution Derived from the PBN Model

With the codon-based models for the mRNA translation process, we are interested in steady state properties such as the translation rate (the number of translated proteins per unit time; or in the codon-based model the number of termination events per unit time), the codon coverage density (the probability of the codon being covered by a ribosome, the same as the term "density" in TASEP models), and the state density (the probability of the mRNA state being a specific one among $\mathbb{M}$). Noticing the fact in the mRNA translation process that any two effective mRNA states are transferable and that any effective mRNA state can stay unchanged with a certain positive probability, we are then able to say that the underlying Markov chains for both the random-sequential and the parallel update rules are irreducible and aperiodic. Therefore, limit (stationary) distributions, which is also the state density at the steady state of the PBN models, exist for both update rules. Denote the limiting distribution for the parallel update rule by $\pi_P := [\pi_P^1 \ \pi_P^2 \ \dots \ \pi_P^m]^T$, where π_P^i is the probability of the mRNA state being the ith in $\mathbb{M}$ (denoted by χ_i), then π_P can be calculated by either of the following two equations,

$$M_P \pi_P = \pi_P, \tag{11a}$$

$$\lim_{i \to \infty} M_P^i = [\pi_P \ \pi_P \ \dots \ \pi_P]. \tag{11b}$$

Similar conclusions apply to the random-sequential update rule as well.

With the state density π_P, the codon coverage density ρ and the translation rate c at the steady state can be calculated as follows,

$$\rho = \sum_{i=1}^{m} \pi_i \chi_i, \tag{12a}$$

$$c = \beta \rho_n. \tag{12b}$$

The calculation procedure for the parallel update rule is organized in Algorithm 3.

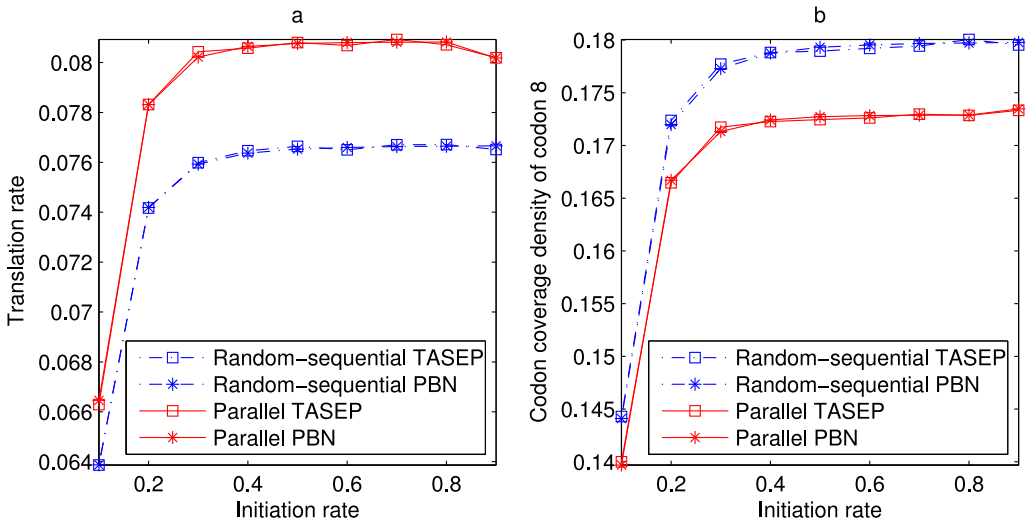


Fig. 3. The PBN models and TASEP simulations produce the same steady state translation rate (Fig. 3a) and codon coverage density (Fig. 3b) profiles for both the random-sequential and parallel updates.

Algorithm 3. Calculating the Steady State Profiles from the PBN Model with the Parallel Update Rule

1) Given n, r, and the event rates in Γ: entry rate α, exit rate β, and hopping rates γ_i.
2) Define the set of the effective mRNA states $\mathbb{M}$.
3) For all $x \in \mathbb{M}$, define the transition probabilities as in (9), thus giving the transition matrix M_P in (10).
4) Determine the limit distribution by (11a) or (11b).
5) Calculate the codon coverage density and translation rate profiles by (12a) and (12b).

Note that π_P contains the stochastic distribution information of the mRNA being at any state, meaning that the codon coverage density ρ and the translation rate c are but two applications of π_P. We will show by various examples in the next section, how π_P can provide valuable information which aids in the dissection of the translation process.

Remark 1. The number of the non-zero elements in M_P for the parallel update rule is $\sum_{x\in\mathbb{M}} 2^{k_x}$, while in M_S for the random-sequential rule this number becomes $\sum_{x\in\mathbb{M}} k_x$. In both cases the transition matrices M_P and M_S are sparse and the random-sequential update rule yields an even more sparse transition matrix, which can be understood from the different natures of the update rules. The sparse nature of the transition matrices have been used to improve the algorithm efficiency.

In the implementation of Algorithm 3, one particular difficulty is to determine the set of the effective mRNA states $\mathbb{M}$, since decisions have to be made on 2^n mRNA states whether they are effective or not. In the current implementation, a recursive algorithm is used, which is one of the main sources of the feasible size constraint of n in Algorithm 3.

4 Use of PBN Modelling Framework for Dissecting mRNA Translation

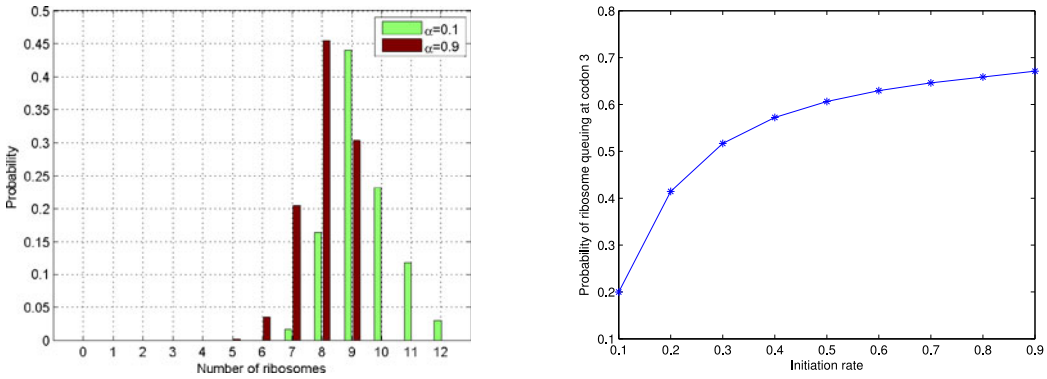
In this section we consider several scenarios involving the basic translation process. We start by demonstrating the correctness of the PBN model by comparing it with TASEP simulations. We then examine mRNA translation with added static and dynamic biological complexities to show the applicability and flexibility of the PBN modelling framework. In these examples, we show that the PBN model can be an alternative tool of analysis and also highlight the fact that the stochastic distribution derived from the PBN model can be useful in a systematic understanding of the intrinsic complexity of mRNA translation, aspects of which are being revealed by recent experimental approaches. With this analysis, we demonstrate that the PBN model can be used as a general modelling framework for mRNA translation, complementing existing tools. The parameters used in the following figures presented in this section are listed in Appendix, available in the online supplemental material.

4.1 Exact Steady State Stochastic Distribution for Basic mRNA Translation Process by the PBN Model

As discussed earlier, the PBN models for both the random-sequential and parallel update rules are derived based on the same assumptions made for corresponding TASEP simulations. Therefore, they should result in the same steady state profiles. These profiles are calculated numerically but exactly by the PBN models (without simulations), and are statistically averaged from sample simulations of the TASEP model.

To verify the correctness of the PBN models in Fig. 3 we calculate the translation rate and the codon densities at the steady state with varying initiation rates. A good match is found and the minor differences between the PBN models and TASEP simulations is likely due to the stochastic effects in the TASEP simulations. This demonstrates the correctness of the PBN model, and provides a secure platform allowing us to use the PBN models for mRNA translation in other contexts.

The solution of the PBN model provides the stochastic distribution of the mRNA being at any state, which contains more information than just the translation rate and the codon densities. For example, with the PBN model, we are able to show the probabilities of the mRNA being occupied by any number of ribosomes (Fig. 4a). We see that in the case where a slow codon may be present, despite the faster



(a) Small initiation rate leads to more ribosomes on the mRNA, revealed by the detailed probabilities of the mRNA being occupied by any number of ribosomes.

(b) The probability of the ribosome queuing at slow codon 3 as a function of the initiation rate. The queuing probability is calculated as the occurrence frequency of the pattern that codons 1-3 are occupied at the same time.

Fig. 4. Applications of the stochastic distribution information obtained from the PBN model.

initiation rate constant of $\alpha = 0.9$, the translation process is in a certain sense more resource optimized for the much slower initiation rate $\alpha = 0.1$, since the latter case sees more average ribosomes on the mRNA. This observation is further explained by the probabilities of ribosome queuing at slow codon 3 (Fig. 4b) since the increased initiation rate, though increasing the translation rate, results in higher probability of ribosome queuing at the slow codon 3 which decreases the translation efficiency.

In fact, we are able to show that for any $\alpha \geq 0.1$, the mRNA state $[1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$ which contributes to the ribosome queuing at slow codon 3 occurs at the highest probability, e.g., 19.23 percent for $\alpha = 0.9$, among all the 4,096 possible mRNA states. The unusually high probability of the mRNA being at this specific state helps our understanding of the role of the slow codon, which can also be useful if we are to engineer or design circuits involving mRNA translation with slow codons using synthetic or other means.

4.2 Modelling Additional Static Complexities: Ribosome Recycle

The PBN model can accommodate additional biological complexity. It has been suggested that in certain situations a ribosome may reattach to the mRNA rather than enter into the free pool when the beginning of the mRNA is not occupied by another ribosome, with the help of a so-called ribosome recycling factor [17], [31], [32]. This happens because of the interactions between the 3′ end binding proteins (like PABP) and the translational machinery due to the circularization of mRNA during translation. Such a recycling process improves the effective use of ribosomes and consequently increases the translation efficiency. Suppose the released ribosome is recycled with probability λ if the entry sites allows it to. Then, a ribosome at the last codon will first leave the mRNA, and in the case that the first r codons on the mRNA are empty, it can reattach the beginning of the mRNA with probability λ. Note that there are slightly different variants which can be used in modelling recycle, but that is tangential to the focus of this paper.

This extra complexity can be easily incorporated in the Boolean framework. As described in [19] the various events: initiation, elongation and termination can be described as appropriate Boolean functions, which facilitates the use of analysis tools. A brief summary of relevant points is included in the Appendix, available in the online supplemental material. In a similar manner, a recycle update event can also be written as a Boolean function, as follows,

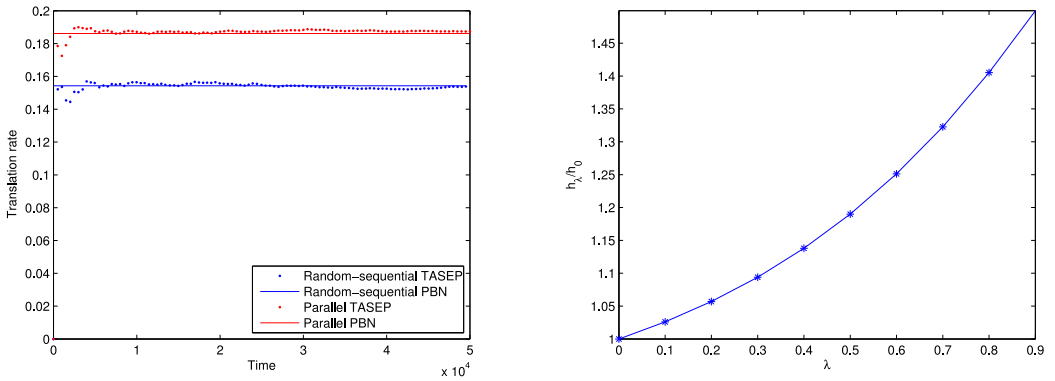
$$\vec{f}_{n^r}: \begin{cases} x_1(t+1) & = (x_n(t) \wedge \neg x_r(t)) \vee x_1(t) \\ & \vdots \\ x_r(t+1) & = (x_n(t) \wedge \neg x_r(t)) \vee x_r(t) \\ x_{r+1}(t+1) & = x_{r+1}(t) \\ & \vdots \\ x_{n-r}(t+1) & = x_{n-r}(t) \\ x_{n-r+1}(t+1) & = (\neg x_n(t)) \vee x_r(t)) \wedge x_{n-r+1}(t) \\ & \vdots \\ x_n(t+1) & = (\neg x_n(t)) \vee x_r(t)) \wedge x_n(t), \end{cases} \tag{13}$$

where the logic operators used are as follows: $\wedge$ for conjunction, or logic AND; $\vee$ for disjunction, or logic OR and $\neg$ for negation, or logic NOT. It is readily seen that the Boolean function in (13) ensures that the released ribosome is recycled provided the mRNA is ready to accept new ribosomes (that is, the first r codons on the mRNA are empty).

Adding this new update event with its probability to the original PBN model leads to the PBN model for mRNA translation with ribosome recycle. No particular difficulty is caused by the inclusion of this biological complexity, for both update rules.

We show the trajectory of the translation rate averaged over time by TASEP simulations and compare it to the exact value obtained by the PBN model in Fig. 5a, where a good match is observed for both update rules. This exact solution of the PBN model can be a great advantage, which will be further discussed in Section 4.4.

On the other hand, due to ribosome recycle, the mRNA state in which the mRNA is ready to accept new ribosomes (the first r codons are empty) and release a ribosome (the last r codons are occupied) at the same time, may lead to the mRNA state that a new ribosome just enters (the first r codons are occupied) and no ribosome is at the end of the mRNA (the last r codons are empty). Denote the former by $[0\ \cdots\ 1]$ and the latter by $[1\ \cdots\ 0]$, and their occurrence probability by $p_{01}(\lambda)$ and $p_{10}(\lambda)$, respectively, where λ is



(a) The PBN model matches the TASEP simulations in the presence of ribosome recycle.

(b) The dependence of the mRNA state $(1 \cdots 0)$ on $(0 \cdots 1)$ increases with the ribosome recycle rate constant. Parallel update is shown.

Fig. 5. The use of the PBN model in analyzing translation with ribosome recycle.

used to indicate the fact that these probabilities are dependent on the ribosome recycle probability λ. We may expect that the larger the recycle probability is, the more the state $[1 \cdots 0]$ can be correlated with $[0 \cdots 1]$. Therefore, the scaled ratio $\frac{h_\lambda}{h_0}$ where $h_\lambda := \frac{p_{10}(\lambda)}{p_{01}(\lambda)}$ may be expected to be an increasing function of λ,. This is indeed found to be the case, as demonstrated in Fig. 5b. This finding means that from the ribosome profiling data of any translation process, the unusually high correlation between the two states $[0 \cdots 1]$ and $[1 \cdots 0]$ may indicate the existence and strength of ribosome recycle. This is useful knowledge which may not be as straightforwardly obtained by other existing tools.

4.3 Modelling Added Dynamic Complexities: Translation in the Presence of Feedback Regulation

Section 4.2 shows that the PBN model can be readily modified to accommodate added biological complexities. The added complexity shown there is however only static. By "static" we mean the added complexity does not make the event rates vary, although the added recycle probability does complicate the system structure. Adding additional constant rates (associated with events) to the PBN model will not change the fundamental structure of the PBN model, since the transition matrix for the underlying Markov chain can still be obtained as before. In these situations, the construction and analysis of the PBN model is not essentially different from the basic case.

On the other hand, other types of biological complexities exist, which make the event rates vary. Feedback regulation is one typical example, where the rates (either initiation, elongation or termination) are not constant, as result of feedback [33]. We understand that the PBN model is based on an underlying Markov chain which is essentially time-homogeneous, and the exact calculation of the steady state profiles benefit from this characteristic. By breaking the time-homogeneous characteristic of the PBN model, the inclusion of dynamic complexities makes the exact calculation of the steady state profiles difficult.

We show here that, despite the absence of the exact calculation, the steady state profiles can be approximated iteratively by the PBN model in the presence of feedback regulation, and this approximation can be obtained with arbitrary accuracy.

To that end, we consider the negative autoregulation of the initiation stage observed in the expression of the poly(A)-binding protein (PABP), and borrow the model from [19], where the initiation rate is assumed to evolve as follows,

$$\alpha = \frac{1}{1 + k_I \rho_I} \alpha_I, \tag{14}$$

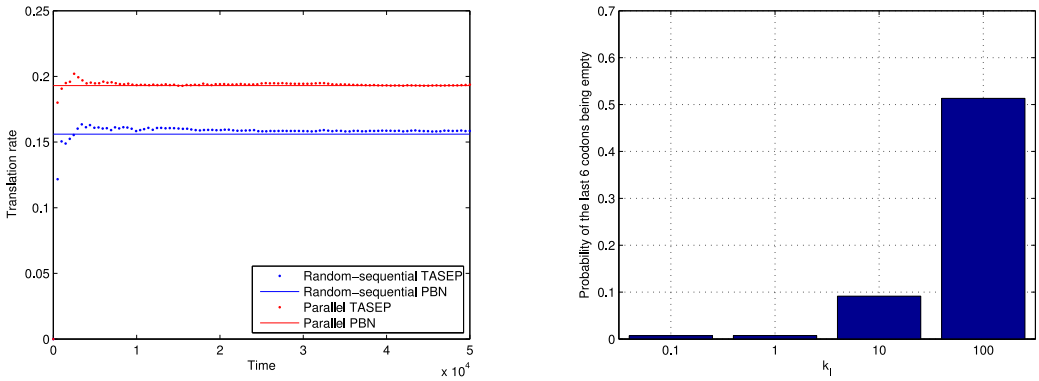
where α_I is interpreted as the maximum initiation rate, ρ_I is the protein concentration of PABP, and $k_I > 0$ as a constant controls the autoregulation strength.

We consider protein synthesis from translation, as described above, with a constant protein degradation rate d_I. Therefore the protein concentration at steady state is given by $\rho_I = c/d_I$, where c is the translation rate. Assuming this steady state is stable, this is the steady state which is reached.

For any fixed α, the PBN model can give rise to the exact steady state. However, due to the existence of the initiation regulation, α is varying with ρ_I, and therefore, the steady state of the above system cannot be calculated directly by the PBN model.

The equation for ρ_I at steady state can be used to calculate the steady state of the PBN model. There are multiple ways of doing this. One way is to use a simple iterative method: we start from some initial ρ_I (and thus α) and calculate the corresponding translation rate c, use this to compute the new ρ_I and iterate. By setting different tolerance bounds, in principle the steady state can be calculated with arbitrary accuracy. Other more sophisticated iterative methods are also possible, which may be facilitated by the direct calculation of parametric sensitivities.

We show the translation rate at the steady state calculated by the aforementioned PBN-based method and compare it with TASEP simulations in Fig. 6a. A satisfactory match is observed for both the random-sequential and parallel update rules. We also show the probabilities of a stretch of the mRNA being empty change with varying feedback strength in Fig. 6b. This information can be exactly calculated by the PBN model (or approximated by simulations) and helps our understanding of how feedback behaves and how it contributes to ribosomal progression



(a) The PBN model can be used to calculate the steady state by iteration in the presence of dynamic feedback. $d_I = 0.5$.

(b) The probability of the last 6 codons being empty with varying feedback strength k_I: calculated by the PBN model with parallel update.

Fig. 6. The use of the PBN model in the presence of dynamic feedback.

and mRNA coverage. What is more, the empty sections on the mRNA may contribute to the instability of the mRNA through other regulatory mechanism (not modelled herein). For instance, it is reported that nonsense mediated decay may be activated if a stretch of the mRNA is left uncovered for sufficiently long time [34]. Therefore, Fig. 6b may then indicate that large k_I is not acceptable in practice.

In synthetic biology, the mRNA translation process is often manipulated at the initiation stage, in order to rapidly control protein synthesis at a low cost. In our model in (16), this means that the maximum initiation rate α_I can be manipulated, and this manipulation can be done in various ways, as the initiation rate is in fact determined by many factors, which can be controlled by upstream signals. From a control engineering perspective, the manipulation of α_I then introduces a feedforward control mechanism to the existing feedback loop. Therefore, in such a case the varying initiation rate α is controlled by both the intrinsic feedback regulation and the extrinsic feedforward manipulation.

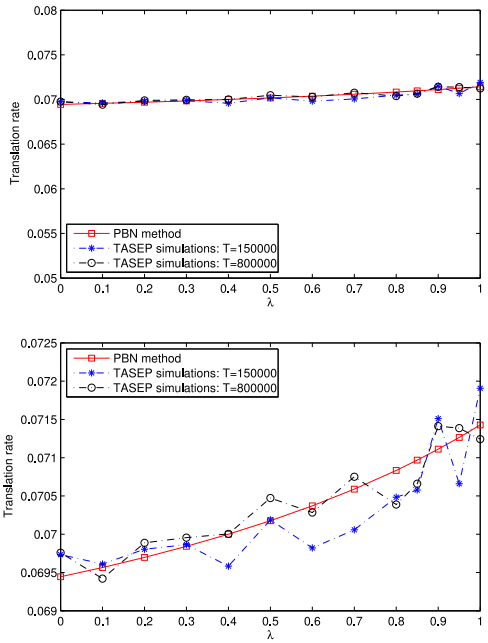


Fig. 7. PBN based exact numerical solution helps the sensitivity analysis of translation rate on λ.

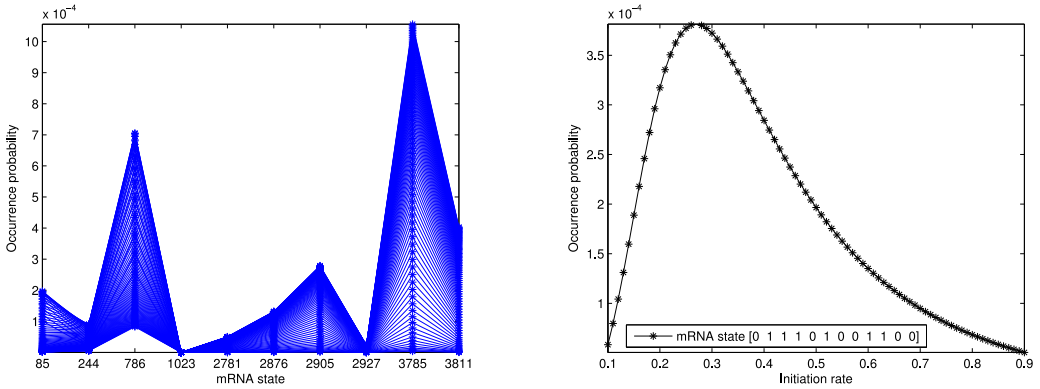
Basic control circuits are often designed based on very simple descriptions of mRNA translation. It is worth pointing out that any such circuit (and for that matter any circuit involving transcription as well) involves the process of translation. In such cases, codon-based features, queueing and stochasticity in translation can all play vital roles and affect the behaviour of the circuit. Therefore systematically accounting for these aspects at the design stage is important. Simulations alone are not a strong enough tool for practical design guidance, especially when an interplay of the above factors comes to the fore. In such a case the PBN method can provide a valuable computational tool in synthetic biology.

We can build on this case, and examine the case of more complex synthetic genetic circuits, which contain translational control mechanisms (whether feedforward, feedback or both). Again, while one way to model these circuits is through ODE models, it is often desirable and necessary to understand other aspects such as the effect of noise, or the interplay between the synthetic circuits and host cells. In these cases, it is necessary to model translation at multiple levels of granularity to understand these issues, and the tools presented here can be extremely useful.

4.4 Exact Numerical Solution of PBN for Parameter Sensitivity Analysis

The exact numerical solution of the PBN model can be useful for the sensitivity analysis of translation with respect to parameters. To show this we consider again the ribosome recycle example in Section 4.2, and analyse the sensitivity of the translation rate to the recycle probability λ for given rates.

To that end, we calculate the translation rates for various λ. This can be done by either TASEP simulations or the PBN model. However, with a large n, the system is controlled by many parameters and in some cases, the variation of λ itself does not lead to dramatic variations of the translation rate. The exact calculation in this case is especially useful. We show a sample analysis in Fig. 7. where we use the simulation times of $T = 150{,}000$ and $T = 800{,}000$, respectively. Notice that both Figs. 7a and 7b are plotted using the same data and the difference lies only in the different scales of the y-axis. No significant difference is observed from Fig. 7a, since for the two sample simulations, the maximum



(a) Different mRNA states have different sensitivity levels to varying initiation rates.

(b) The occurrence probability of the mRNA state may not change monotonically with the initiation rate.

Fig. 8. Using PBN model for sensitivity analysis where the sensitivity is captured by the height (difference between the maximum and minimum values) of the vertical lines corresponding to each mRNA state in Fig. 8a.

variations of the translation rate compared with the PBN model are only around 0.8 percent for $T = 150{,}000$ and 0.6 percent for $T = 800{,}000$, respectively, and the maximum difference between the two sample simulations is around 1 percent, hardly observable on the plot. However, when we zoom in further in Fig. 7b we clearly see the smooth increasing trend of the translation rate with λ by the PBN model, while for TASEP simulations, an inaccurate calculation distorts the observable pattern.

By increasing further the simulation time, one may observe similar patterns from TASEP simulations as well, but the point is, increasing the simulation time will dramatically increase the computational resources requirement and more importantly, whatever the simulation time is, we may never be absolutely confident with the results obtained in this manner (for instance if the accuracy in a certain parameter regime requires greater time integration). In this sense the exact numerical calculation provided by the PBN model is extremely valuable.

In this context it is worth pointing out that the sensitivities of the stationary distribution with respect to parameters can be obtained directly, without necessarily performing any finite difference operation. Indeed, it is possible to directly differentiate the equation determining the stationary distribution (with also the fact that the sum of sensitivities of different state densities with respect to parameters is zero), to obtain the relevant sensitivities using matrix vector products.

The exact stochastic distribution obtained by the PBN model makes it possible for us to investigate the sensitivity of specific mRNA states. In Fig. 8a, we plot the occurrence probabilities of some randomly chosen mRNA states (in the basic case of translation without recycle) with varying initiation rates from $\alpha = 0.1$ to $\alpha = 0.9$ with the increasing step being 0.01, where the mRNA states labelled for the x axis are the decimal value of the binary expression of the mRNA states, e.g., "85" represents mRNA state $[0\ 0\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1]$. Hence, the height of the vertical line for each state (composed of all the occurrence probabilities produced with the initiation rates between $\alpha = 0.1$ to $\alpha = 0.9$ with the increasing step 0.01) is then an indication of the sensitivity of that mRNA state to varying initiation rates. In this regard different mRNA states have dramatically different sensitivity levels. On the other hand, the occurrence probabilities of more than $2/3$ mRNA states are found not to change monotonically with the initiation rate, as illustrated in Fig. 8b for the mRNA state $[0\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 0]$. This calculations is based on the initiation rate range of 0.1-0.9, hence more non-monotonic states may be expected if this range is further extended. This finding again points to the fact that understanding translation at the mRNA state level, as required by the ribosome profiling data, necessitates stochastic codon-based models and appropriate tools of analysis for its effective dissection.

5 Discussion and Conclusion

The TASEP and PBN models describe the same process of mRNA translation but are developed and analysed from very different perspectives. The TASEP model regards the mRNA translation process as a particle transportation system (and a related many-body system) while the PBN model treats the process as a Boolean network with its state evolution being governed by an underlying Markov chain. Hence the PBN model provides a useful complementary perspective for modelling mRNA translation. As mentioned earlier, methods of analysis of such PBN models served to bridge the gap between mean field analysis of TASEP models and stochastic simulations. In addition, PBN models of moderate size themselves bridge the gap between simple lumped ODE models and detailed codon-based stochastic models. Therefore, such models create bridges between different models and different methods of analysis contributing to a consolidated and secure platform for modelling translation in systems and synthetic biology.

We have developed and extended the PBN modelling framework to incorporate both random-sequential and parallel update rules. The parallel update rule is an alternative reasonable (perhaps, in some respects even more preferable) possibility for modelling mRNA translation, while relevant studies are rare [16]. This is probably because the TASEP models with the parallel update rule are more difficult to handle. In contrast, except for the possible requirement of more computational resources (the resulting transition matrix of the parallel update rule is less sparse compared to the random-sequential case), the PBN model with the parallel update rule is not much more difficult to construct and analyse compared to such a model with the random-sequential

update rule. This thus makes the PBN model particularly useful in the case of the parallel update rule.

The computational complexity of the PBN model increases exponentially with the increase of the system size, as the number of the effective mRNA states increases exponentially with n (but decreases with r). This fact means that the PBN model may not be able to directly model and analyse translation at single codon resolution a lot of mRNAs. However, as mentioned earlier, the PBN model can be used in a coarse-grained manner for this purpose. Indeed, a coarse-grained approach called the "ribosome flow model" has already been reported in other settings [26], [35], [36]. This model coarse-grains the codons on the mRNA into segments, resulting in the same model structure as the PBN model as illustrated in Fig. 1 with single coverage ($r = 1$). In [26], the authors developed coarse grained models incorporating available data on tRNA concentrations and activation energies of each step. This was done for cellular systems ranging from bacteria to human: *S. cerevisiae*, *S. pombe*, *E. coli*, and *Human* liver. This predicted successfully the fundamental features of the translation process, including the translation rate, ribosome densities, etc. and the relations between all the variables. The ribosome flow model outperformed other often used predictors like the *tAI* [37]. This makes the ribosome flow model, and consequently similar coarse-grained models such as the PBN model, a very useful tool in analysing mRNA translation and connecting with experimental data. It is reported that the size of the segments being between 25-35 codons results in the best correlation between the coarse-grained model and the biological data, which corresponds to the system size of about 15-18 coarse grained segments for most mRNAs. In such cases, most mRNAs are within the computational limit of the PBN model, since this can be implemented on a personal computer without optimizing the algorithm efficiency. Hence, the fact that the PBN model shares the same model structure as the ribosome flow model and that the size of the coarse-grained model can be easily handled by the PBN model, then enable the PBN model to be well positioned to interpret biological data. On the other hand, the PBN model has its own value even independent of experimental data, that is, the exact calculation of the PBN model provides us with distinct advantages compared to the stochastic simulations, as based on exact calculations, we are able to infer general properties and trends of the mRNA translation process with more confidence from sample systems with moderate sizes. For that purpose, the limitations on the computational efficiency is not an essential issue.

The emerging ribosome profiling technique provides an additional context for using the stochastic PBN tool. To interpret genome scale data, we need analysis tools that can provide more than just averaged properties such as the translation rate and codon coverage density, and the exact stochastic distribution of the mRNA states derived by the PBN approach can be useful for this purpose. In addition, for modelling and understanding such a complicated system, it is important to understand how different aspects of the stationary distribution vary with parameters and the PBN model provides value here too.

mRNA translation is a highly complex process regulated at multiple levels, enabling other complex behaviour. For example, if the initiation rate in (14) is regulated via positive feedback, the translation process is then found to possess bistability. Mean-field TASEP models are convenient tools to analyse such behaviour due to the dynamical systems framework. However, the stochastic steady state distribution(s) is not available. Stochastic simulations are a useful, but far from sufficient tool to investigate bistability due to the lack of clear mapping from parameter variations to the steady states using simulations. In this case the PBN model provides a valuable tool, which can numerically calculate all the steady states (not mean-field approximated) with also the corresponding detailed stochastic distributions. Further, we may be able to use mean-field TASEP and the PBN tools in conjunction to investigate such complex dynamics: mean-field TASEP conveniently allows us to understand the deterministic dynamics and get estimates of relevant parameter regimes, based on which stochastic distributions can be calculated by the PBN tool. Finally, it is worth pointing out that the stochastic distribution of the unstable steady state can be also calculated. This cannot be obtained either from stochastic simulations or mean-field approaches. This can allow for an understanding of stochastic evolution from this state (paralleling similar studies in deterministic dynamical systems). Further the ability to compute distributions of unstable steady states provides further information if one wishes to monitor or manipulate both aspects of the steady state distribution and stability through synthetic or other means.

Overall, we see that the PBN model and analysis tools allow us to probe aspects of translation and discern clear trends through approaches which complement simulation. Thus the stationary distribution, and in fact sensitivities of the stationary distribution can be obtained for given parameter values directly. There are a number of situations where such capabilities may be especially useful.

In systems biology, many regulatory steps in translation are being uncovered. These regulatory steps are at multiple levels and related to other parts of the cellular system, and therefore systematic modelling frameworks are required for understanding the translation process accounting for this [2], [33], [38]. In synthetic biology, individual steps of translation are manipulated through multiple approaches, including feedforward and feedback regulation. Synthetic biology is now engineering riboswitches, ribozymes, small RNAs [39], [40], and other possible regulatory molecules to regulate protein synthesis, suggesting that sophisticated dynamic regulation of protein synthesis may be possible in the future. Developing more complex circuits which include such regulation naturally thus requires the use of such tools. Even circuits which are built through transcriptional regulation, involve translation and protein synthesis as part of the circuit. While ODE models may be partially useful in some of these cases, if one requires a systematic understanding of system or circuit behaviour, determining the effect of stochasticity and examining the interaction between the translation system and its surroundings, such tools become very important. The broader need for synthetic biology to confront the complexity of biology is being acknowledged [41], and likewise sophisticated engineering of translation needs to account for the complexity of translation. Such tools are also useful if it is desired to optimize certain aspects of the translation process and protein synthesis in biotechnology.

It is worth emphasizing that one combine detailed stochastic simulations, mean-field and ODE approaches, and the PBN approach as a broad toolkit for dissecting and manipulating translation and protein synthesis, since each tool presents unique advantages. Stochastic simulations can be performed on relatively large size systems, incorporating a lot of detail, and including stochastic effects. Mean-field and ODE approaches present effective tools of analysis, including the dynamical systems framework. The PBN model provides stochastic information and a direct tool of analysis without requiring repeated simulations, allowing for the dissection of the interplay between system features, parameters and stochastic aspects. Finally it is worth pointing out that the elucidation of other molecular traffic processes such as transport by motors (which are not necessarily unidirectional) can be also benefit from such tools [42], [43].

All our discussions and analysis show that, the flexibility, the exact numerical solution and the role as a bridging tool, make the PBN model a very valuable tool for understanding the complexity of mRNA translation.

ACKNOWLEDGMENTS

This work was supported by BBSRC through grant BBI04254/1.

REFERENCES

[1] B. Lewin, *Genes IX*. Sudbury, MA, USA: Jones & Bartlett, 2007.
[2] S. Gokhale, D. Nyayanit, and C. Gadgil, "A systems view of the protein expression process," *Syst. Synthetic Biol.*, vol. 5, no. 3-4, pp. 139–150, 2011.
[3] R. Heinrich and T. A. Rapoport, "Mathematical modelling of translation of mRNA in eucaryotes; steady states, time-dependent processes and application to reticulocytes," *J. Theor. Biol.*, vol. 86, no. 2, pp. 279–313, 1980.
[4] T. Chou and G. Lakatos, "Clustered bottlenecks in mRNA translation and protein synthesis," *Phys. Rev. Lett.*, vol. 93, p. 198101, Nov 2004.
[5] A. Mehra and V. Hatzimanikatis, "An algorithmic framework for genome-wide modeling and analysis of translation networks," *Biophys. J.*, vol. 90, no. 4, pp. 1136–1146, 2006.
[6] P. Shah, Y. Ding, M. Niemczyk, G. Kudla, and J. B. Plotkin, "Rate-limiting steps in yeast protein translation," *Cell*, vol. 153, no. 7, pp. 1589–1601, 2013.
[7] T. von der Haar, "Mathematical and computational modelling of ribosomal movement and protein synthesis: An overview," *Comput. Struct. Biotechnol. J.*, vol. 1, no. 1, p. e201204002, 2012.
[8] E. de Silva, J. Krishnan, R. Betney, and I. Stansfield, "A mathematical modelling framework for elucidating the role of feedback control in translation termination," *J. Theor. Biol.*, vol. 264, no. 3, pp. 808–821, 2010.
[9] N. Skjndal-Bar and D. R. Morris, "Dynamic model of the process of protein synthesis in eukaryotic cells," *Bull. Math. Biol.*, vol. 69, no. 1, pp. 361–393, 2007.
[10] M. Kozak, "The scanning model for translation: An update," *J. Cell. Biol.*, vol. 108, no. 2, pp. 229–241, 1989.
[11] A. Basu and D. Chowdhury, "Modeling protein synthesis from a physicist's perspective: A toy model," *Am. J. Phys.*, vol. 75, no. 10, pp. 931–937, 2007.
[12] R. Zia, J. Dong, and B. Schmittmann, "Modeling translation in protein synthesis with TASEP: A tutorial and recent developments," *J. Statist. Phys.*, vol. 144, no. 2, pp. 405–428, 2011.
[13] K. E. P. Sugden, "Nonequilibrium statistical physics applied to biophysical cellular processes," Ph.D. dissertation, department of Physics, The Univ. of Edinburgh, Edinburgh, UK, 2009.
[14] J. Dong, B. Schmittmann, and R. Zia, " Towards a model for protein production rates," *J. Statist. Phys.*, vol. 128, no. 1-2, pp. 21–34, 2007.
[15] L. B. Shaw, R. K. P. Zia, and K. H. Lee, "Totally asymmetric exclusion process with extended objects: A model for protein synthesis," *Phys. Rev. E*, vol. 68, p. 021910, 2003.
[16] C. A. Brackley, D. S. Broomhead, M. C. Romano, and M. Thiel, "A max-plus model of ribosome dynamics during mRNA translation," *J. Theor. Biol.*, vol. 303, pp. 128–140, 2012.
[17] E. Marshall, I. Stansfield, and M. Romano, "Ribosome recycling induces optimal translation rate at low ribosomal availability," *J. Roy. Soc. Interface*, vol. 11, no. 98, p. 20140589, 2014.
[18] A. Dana and T. Tuller, "The effect of trna levels on decoding times of mrna codons," *Nucleic Acids Res.*, vol. 42, no. 14, pp. 9171–9181, 2014.
[19] Y.-B. Zhao and J. Krishnan, "mRNA translation and protein synthesis: an analysis of different modelling methodologies and a new PBN based approach," *BMC Syst. Biol.*, vol. 8, no. 1, p. 25, 2014.
[20] N. T. Ingolia, L. F. Lareau, and J. S. Weissman, "Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes," *Cell*, vol. 147, no. 4, pp. 789–802, 2011.
[21] M. V. Gerashchenko, A. V. Lobanov, and V. N. Gladyshev, "Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 43, pp. 17 394–17 399, 2012.
[22] M. Guttman, P. Russell, N. T. Ingolia, J. S. Weissman, and E. S. Lander, "Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins," *Cell*, vol. 154, no. 1, pp. 240–251, 2013.
[23] N. T. Ingolia, "Ribosome profiling: New views of translation, from single codons to genome scale," *Nat. Rev. Genetics*, vol. 15, no. 3, pp. 205–213, 2014.
[24] X. Cai and X. Wang, "Stochastic modeling and simulation of gene networks-a review of the state-of-the-art research on stochastic simulations," *IEEE Signal Process. Mag.*, vol. 24, no. 1, pp. 27–36, Jan. 2007.
[25] D. T. Gillespie, "Stochastic simulation of chemical kinetics," *Annu. Rev. Phys. Chem.*, vol. 58, no. 1, pp. 35–55, 2007.
[26] 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.*, vol. 7, no. 9, p. e1002127, 2011.
[27] N. Rajewsky, L. Santen, A. Schadschneider, and M. Schreckenberg, "The asymmetric exclusion process: Comparison of update procedures," *J. Statist. Phys.*, vol. 92, no. 1-2, pp. 151–194, 1998.
[28] D. Cheng, H. Qi, and Y. Zhao, *An Introduction to Semi-Tensor Product of Matrices and Its Applications*. Singapore, World Scientific, 2012.
[29] D. Cheng, H. Qi, and Z. Li, *Analysis and Control of Boolean Networks: A Semi-tensor Product Approach*. New York, NY, USA: Springer, 2011.
[30] D. Cheng and H. Qi, "A linear representation of dynamics of Boolean networks," *IEEE Trans. Autom. Control*, vol. 55, no. 10, pp. 2251–2258, Oct. 2010.
[31] M. A. Gilchrist and A. Wagner, "A model of protein translation including codon bias, nonsense errors, and ribosome recycling," *J. Theor. Biol.*, vol. 239, no. 4, pp. 417–434, 2006.
[32] L. Janosi, I. Shimizu, and A. Kaji, "Ribosome recycling factor (ribosome releasing factor) is essential for bacterial growth," *Proc. Nat. Acad. Sci. USA*, vol. 91, no. 10, pp. 4249–4253, 1994.
[33] R. Betney, E. de Silva, C. Mertens, Y. Knox, J. Krishnan, and I. Stansfield, "Regulation of release factor expression using a translational negative feedback loop: A systems analysis," *RNA*, vol. 18, no. 12, pp. 2320–2334, 2012.
[34] S. Brogna and J. Wen, "Nonsense-mediated mRNA decay (NMD) mechanisms," *Nat. Struct. Mol. Biol.*, vol. 16, no. 2, pp. 107–113, 2009.
[35] M. Margaliot and T. Tuller, "On the steady-state distribution in the homogeneous ribosome flow model," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 9, no. 6, pp. 1724–1736, Nov.-Dec. 2012.
[36] Y. Zarai, M. Margaliot, and T. Tuller, "Explicit expression for the steady-state translation rate in the infinite-dimensional homogeneous ribosome flow model," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 10, no. 5, pp. 1322–1328, Sep. 2013.
[37] M. dos Reis, R. Savva, and L. Wernisch, "Solving the riddle of codon usage preferences: A test for translational selection," *Nucleic Acids Res.*, vol. 32, no. 17, pp. 5036–5044, 2004.
[38] C. Rato, S. R. Amirova, D. G. Bates, I. Stansfield, and H. M. Wallace, "Translational recoding as a feedback controller: Systems approaches reveal polyamine-specific effects on the antizyme ribosomal frameshift," *Nucleic Acids Res.*, vol. 39, no. 11, pp. 4587–4597, 2011.

[39] J. Chappell, M. K. Takahashi, S. Meyer, D. Loughrey, K. E. Watters, and J. Lucks, "The centrality of RNA for engineering gene expression," *Biophys. J.*, vol. 8, no. 12, pp. 1379–1395, 2013.
[40] F. J. Isaacs, D. J. Dwyer, and J. J. Collins, "RNA synthetic biology," *Nat. Biotechnol.*, vol. 24, no. 5, pp. 545–554, 2006.
[41] J. J. Collins, "Synthetic biology: How best to build a cell," *Nature*, vol. 509, pp. 155–157, 2014.
[42] R. Lipowsky, Y. Chai, S. Klumpp, S. Liepelt, and M. J. Muller, "Molecular motor traffic: From biological nanomachines to macroscopic transport," *Physica A*, vol. 372, no. 1, pp. 34–51, 2006.
[43] Y. Chai, S. Klumpp, M. J. I. Müller, and R. Lipowsky, "Traffic by multiple species of molecular motors," *Phys. Rev. E*, vol. 80, p. 041928, 2009.

Yun-Bo Zhao received the BSc degree in mathematics from Shandong University, Shandong, China, in 2003, the MSc degree in systems theory from the Institute of Systems Science, Chinese Academy of Sciences, Beijing, China, in 2007, and the PhD degree from the University of Glamorgan (now University of South Wales), Pontypridd, United Kingdom, in 2008. He is currently a professor with the Zhejiang University of Technology, China. His research interests include systems biology, networked control systems, and Boolean networks.

J. Krishnan did his undergraduate studies in chemical engineering at IIT-Madras and received the PhD degree from Princeton University, also in chemical engineering. His research work focussed on nonlinear dynamics and pattern formation on catalyst surfaces. He subsequently was an associate research scientist in the Department of Electrical Engineering at the Johns Hopkins University, which was his entree into the world of systems biology. He subsequently moved to Imperial College London, where he is currently a senior lecturer in the Department of Chemical Engineering and the Centre of Process Systems Engineering, with affiliate appointments in the Institute of Systems and Synthetic Biology and the Centre for Bioinformatics. His work focuses on the elucidation of functioning of cellular processes in terms of the underlying signalling and gene regulatory networks. This involves mathematical modelling of concrete problems of basic and applied interest, in collaboration with cell biologists, biomedical scientists, and synthetic biologists. This work is complemented by theoretical and systems efforts, involving dynamical systems, systems engineering, control engineering, and networks, which provide frameworks and tools for the efficient elucidation, and manipulation of cellular information processing.