

Università degli Studi dell'Insubria

DIPARTIMENTO DI SCIENZA ED ALTA TECNOLOGIA
Corso di laurea in Fisica (L-30)



Totally Asymmetric Simple Exclusion Process: un modello per lo studio della traduzione del mRNA

Relatore:

Prof. Ginelli Francesco

Laureando:

Arrigoni Stefano

Matricola: 737870

Anno accademico 2023–2024

"There's as many atoms in a single molecule of your DNA
as there are stars in the typical galaxy.
We are, each of us, a little universe."

- Neil deGrasse Tyson

Indice

	Page
1 Sinossi	3
2 Introduzione	5
2.1 Background biologico	5
2.2 TASEP model: soluzione di campo medio	7
2.2.1 Periodic Boundary	8
2.2.2 Open Boundary	10
3 Modello per la traduzione del mRNA	13
3.1 Modello a due stati	13
3.2 Soluzione di campo medio	14
3.2.1 Periodic boundary	14
3.2.2 Open boundary	15
3.3 Maximal Current Principle	17
4 TASEP inomogeneo	18
4.1 Definizione del modello	18
4.1.1 Master Equation	19
4.2 Power Series Approximation (PSA)	20
4.3 TASEPy	22
4.3.1 Risoluzione Iterativa del modello disomogeneo	22
4.3.2 TASEPy: python per velocizzare e semplificare i calcoli	23
5 Conclusioni	26
5.1 TASEP omogeneo: conclusioni sul modello a due fasi	26
5.2 TASEPy e risoluzione del TASEP disomogeneo: applicazioni e risultati	27

1 Sinossi

In questo elaborato di natura compilativa si approfondirà tramite l'Asymmetric Simple Exclusion Process (ASEP) o Totally Asymmetric Simple Exclusion Process (TASEP) la sintesi del polipeptide, sequenza di amminoacidi che formeranno in seguito ad altri processi chimici una proteina.

Il TASEP model permette lo studio della traduzione del mRNA definendo il ribosoma come una particella che si muove lungo un reticolo, il filamento di mRNA stesso.

Il modello TASEP, inizialmente introdotto alla fine degli anni 60 da Carolyn T. MacDonald and Julian H. Gibbs [1] [2], è tutt'oggi utilizzato per lo studio della dinamica di diversi fenomeni tra cui, appunto, la traduzione del mRNA, lo studio del traffico e l'evoluzione delle code.

Questa modellizzazione permette di evidenziare le caratteristiche principali del processo biologico.

Il processo di traduzione è un processo sequenziale che parte dall'accoppiamento del tRNA con il codone presente sul filamento di RNA messaggero fino all'abbandono del tRNA stesso dopo il rilascio dell'aminoacido tramite il processo di elongazione del polipeptide. Il modello TASEP descrive lo spostamento di particelle con una dinamica stocastica in una sola direzione lungo un reticolo unidimensionale, riprendendo appunto la sequenzialità del processo biologico nel quale un ribosoma scorre lungo il filamento di mRNA. Il vincolo di esclusione inoltre riprende l'impossibilità di più ribosomi di sovrapporsi sullo stesso codone. La parte introduttiva approfondirà le proprietà biologiche e descriverà i modelli originali degli anni '68 [1] e '69 [2]. È interessante notare come una semplice approssimazione di campo medio permette di descrivere le differenze di fase che possono verificarsi nel sistema, i punti critici, la densità media e locale delle particelle tra i vari siti e la corrente di spostamento delle particelle, trascurando la correlazione tra i diversi siti.

Nel capitolo successivo si metterà in evidenza l'importanza del processo di accoppiamento tra tRNA e codone, trascurato dai modelli originali che lo assumevano come istantaneo. I dati sperimentali indicano invece che il rate di accoppiamento corrisponde con il processo più lento limitante la cinetica, dunque Ciandrini et al (2010) [3] mettono in evidenza questa proprietà definendo un grado di libertà aggiuntivo. Lo stato interno aggiuntivo permette quindi di descrivere il momento in cui si trova il ribosoma durante il processo di traduzione.

La descrizione del processo biologico viene riassunto in due stati principali: lo stato di accoppiamento (primo stato) e lo stato di traslocazione (secondo stato). L'assunzione di ambiente omogeneo, ovvero la presenza omogenea dei tRNA nell'ambiente citoplasmatico, permette di trascurare le fluttuazioni di densità e l'approssimazione di campo medio permette di trascurare la correlazione tra i diversi siti.

Questa nuova modellizzazione consente di approfondire lo studio della corrente di spostamento dei ribosomi sul reticolo discreto e permette di indagare più a fondo il comportamento della densità locale e media associando a ciascuno stato interno una propria densità, migliorando la comprensione del processo biologico.

Nell'ultimo capitolo si tratterà il modello TASEP inomogeneo che si differenzia dai modelli visti in precedenza poiché l'hopping rate (probabilità istantanea di salto) tra un sito e l'altro varia a seconda del sito considerato.

Il TASEP inomogeneo descrive la variazione di velocità di scorrimento del ribosoma a causa della diversa disponibilità dei tRNA portatori degli amminoacidi correlati ai codoni. Si è dimostrato che gli amminoacidi sono decodificati da 2-6 codoni distribuiti non in maniera casuale [4]; infatti in ambito biotecnologico è stato possibile aumentare la produzione delle proteine di circa mille volte modificando la codificazione degli amminoacidi con codoni meno frequenti.

Questa inomogeneità dello scorrimento può essere quindi vista come una diversa probabilità di salto da un sito ad un altro.

Il modello inomogeneo non ha ancora una soluzione analitica esatta a differenza del modello omogeneo. Le soluzioni principali sono state trovate tramite l'approssimazione di campo medio nel caso in cui i rate di salto variano lentamente nel reticolo e il reticolo stesso è molto esteso, ma questa prima approssimazione trascura inoltre la correlazione tra i diversi siti.

Una migliore soluzione è possibile studiarla tramite l'approssimazione in serie di potenze (PSA) [4]. Il parametro di espansione utilizzato per la PSA è il rate di inizializzazione α che regola l'immissione di nuove particelle sulla catena. Tipicamente infatti quest'ultimo è il processo più lento e dunque α è il rate più piccolo. L'approssimazione è dominata dai primi termini, dunque studiando i primi k termini dell'espansione e ricavando i primi coefficienti è possibile ottenere il calcolo di densità e corrente con buona approssima-

zione.

Il numero di coefficienti k da calcolare per avere una buona approssimazione come si può intuire è tuttavia enorme. Per questo motivo Ciandrini et al (2023) [5] hanno sviluppato un pacchetto in codice python, TASEPy, che permette la risoluzione in PSA in modo efficiente e rapido. TASEPy infatti permette la risoluzione in serie di potenze e il calcolo dei coefficienti nel giro di pochi minuti tramite l'utilizzo di: 4 funzioni personalizzate, la definizione della matrice degli hopping rate e la definizione della lunghezza (ℓ) del ribosoma.

Di seguito è rappresentato schematicamente l'intero processo biologico di sintesi di una proteina che consta del processo di trascrizione, nel quale il DNA viene copiato in RNA messaggero (mRNA) grazie all'enzima RNA polimerasi che traduce la sequenza del DNA nella sequenza nucleotidica che compone il filamento di mRNA, e del processo di traduzione del mRNA in filamento amminoacilico (il polipeptide) che in seguito ad altri processi chimici diventerà la proteina stessa; l'elaborato di tesi tratterà solo il fenomeno della traduzione del mRNA.

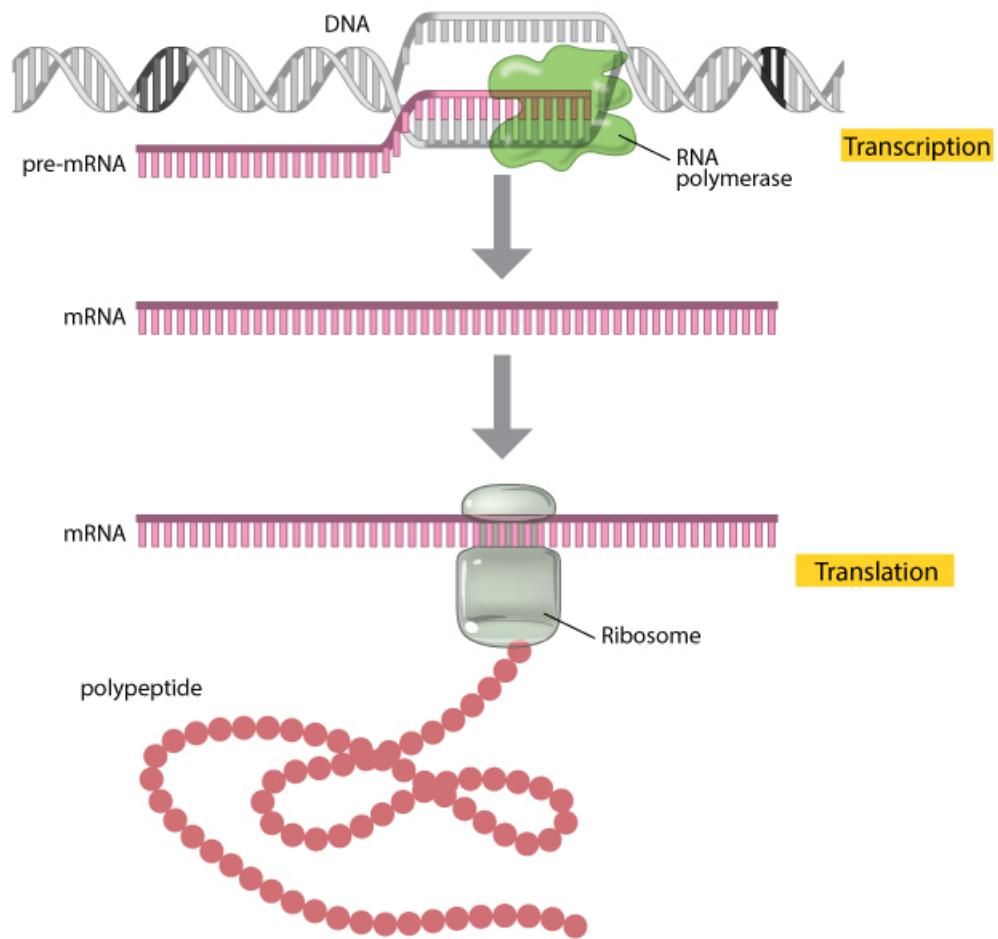


Figura 1: Trascrizione e traduzione: dal DNA fino alla sintesi del polipeptide. Nel presente elaborato si andrà ad analizzare il fenomeno della traduzione. [6]

2 Introduzione

2.1 Background biologico

L'introduzione di una nomenclatura biologica di base è imprescindibile per poter capire come il TASEP model venga associato al fenomeno biologico della traduzione:

- *acido ribonucleico messaggero* (mRNA): è una molecola di RNA a filamento singolo generata durante il processo di trascrizione del DNA e corrisponde alla codifica dei geni presenti nel DNA. È formato da una sequenza di nucleotidi (unità cellulare formata da una base azotata, uno zucchero e un gruppo fosfato);
- *ribosoma*: complesso macromolecolare senza membrana e formato da 4 molecole di rRNA e 80 proteine, è responsabile per il processo di traduzione. Durante tale processo il ribosoma si lega al filamento di mRNA e può essere suddiviso in 3 sub-unità funzionali: il sito amminoacilico (A) in cui un tRNA che trasporta uno specifico amminoacido lega il suo anticodone al codone del mRNA, il sito polipeptidico (P), ovvero il sito in cui l'amminoacido trasportato dal tRNA viene rilasciato e legato all'amminoacido precedente a formare il polipeptide, e il sito di espulsione (E) nel quale il tRNA ormai privo del suo amminoacido viene rilasciato nel citoplasma;
- *RNA di trasferimento* (tRNA): filamenti singoli di RNA che diffondono liberamente nel citoplasma e trasportano gli amminoacidi. Sono la chiave del processo di accoppiamento tra codone e anticodone durante la traduzione del mRNA;
- *codone*: unione di 3 basi azotate presenti sul mRNA
- *anticodone*: unione di 3 basi azotate presenti sul tRNA che andrà a legarsi col rispettivo codone (da cui il nome)
- *Amminoacido*: unità elementare di una proteina; codificato da un codone nel mRNA e dal rispettivo anticodone nel tRNA.

La traduzione del mRNA è di fondamentale importanza per la sintesi degli amminoacidi e dunque delle proteine. L'intero processo coinvolge mRNA, Ribosomi e tRNA e si divide in tre fasi principali: *inizializzazione, elongazione e terminazione*.

1. **Inizializzazione:** una sequenza di inizializzazione (circa 170 amminoacidi) definisce il legame tra un mRNA ed un ribosoma. Il legame instaurato determina l'inizio della fase di codifica dopo che un tRNA che trasporta un amminoacido chiamato *metionina* si lega, tramite il suo anticodone, al primo codone della sequenza di codifica. La metionina verrà scartata e segnerà semplicemente l'inizio della traduzione del mRNA. Questo processo "inizializza" la catena polipeptidica del sito P;
2. **elongazione:** è la fase in cui il ribosoma scorre lungo il filamento di mRNA e lega tra loro gli amminoacidi trasportati dal tRNA per costruire la proteina. Il processo di appaiamento codone-anticodone avviene nel sito A. Una volta che il codone del mRNA e l'anticodone del tRNA si sono legati correttamente sul sito A del ribosoma, la catena polipeptidica passa da P ad A dove si lega all'amminoacido trasportato dal tRNA. Dopo la formazione del legame peptidico il ribosoma avanza di un codone e il tRNA si sposta sul sito E dove poi verrà espulso, lasciando il sito A libero di poter ripetere il processo di appaiamento e di allungamento della catena polipeptidica. Questo processo si ripeterà fino alla fine della sequenza di codifica;
3. **terminazione:** I tre codoni finali non hanno alcun anticodone associato, questo permette al ribosoma di svuotare tutti e tre i siti e facilità il rilascio del complesso ribosomale completando il processo di traduzione. Questa sequenza di codoni è chiamata *fattori di rilascio*.

È importante notare che durante il processo di traduzione possono esserci più ribosomi che scorrono sullo stesso filamento di RNA messaggero. Il processo di traduzione è però un processo sequenziale e dunque non è possibile che due o più codoni si sovrappongano sullo stesso codone. Questa proprietà è ripresa dal modello TASEP, che è approfondito nei capitoli successivi ed è chiamata proprietà di esclusione, poiché un sito o codone nel caso biologico può essere occupato da un solo ribosoma alla volta.

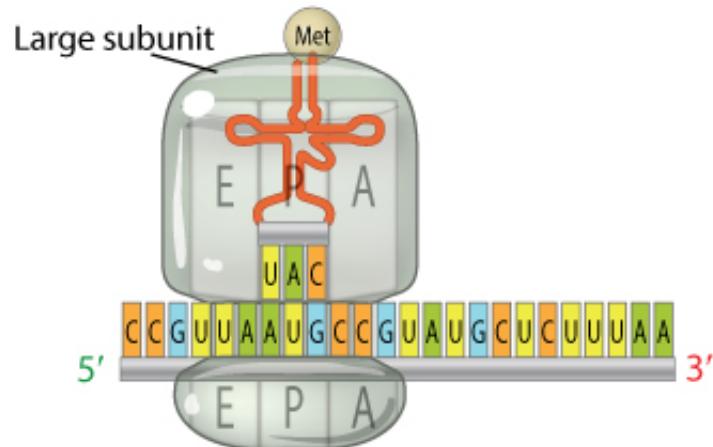


Figura 2: Legato il primo anticodone il ribosoma trasla e libera il sito A per iniziare il processo di elongazione [6]

2.2 TASEP model: soluzione di campo medio

In prima approssimazione il processo di traduzione del mRNA può essere riassunto nel movimento del ribosoma sul mRNA stesso. È infatti governato principalmente dallo scorrimento del ribosoma sul filamento di RNA messaggero. La traduzione avviene in modo sequenziale, il ribosoma può scorrere solo in una direzione e i processi chimici che permettono la sintesi di un amminoacido e l'allungamento del polipeptide devono avvenire, come descritto, in modo sequenziale. Questo permette di racchiudere il processo di traduzione nello scorrimento del ribosoma sul filamento di mRNA rendendo semplice una prima associazione del modello TASEP con quello della traduzione. Il TASEP model infatti è governato da un processo stocastico i cui elementi fondamentali si possono identificare in una o più particelle, ovvero uno o più ribosomi, e un reticolo discreto unidimensionale, il filamento di mRNA i cui siti sono rappresentati dai codoni.

Le proprietà di base del processo stocastico sono ben descritte dall'acronimo TASEP:

- **Totally Asymmetric (TA):** la particella ha la possibilità di spostarsi solo da sinistra verso destra, come se una forza $F \rightarrow \infty$ favorisse lo spostamento verso destra e impedisse quello verso sinistra;
- **Simple Exclusion (SE):** ogni sito può essere occupato solamente da una particella alla volta.

Il processo di esclusione impone una condizione necessaria allo spostamento della particella, infatti presa una particella sul sito j la possibilità di salto della particella al sito $j+1$ è condizionata dal fatto che il sito $j+1$ sia vuoto. Questa proprietà riflette a pieno il processo descritto dal ribosoma sul reticolo composto di codoni; infatti il processo sequenziale messo in atto dai ribosomi dipende dalla dinamica del ribosoma successivo che può, se eccessivamente lento, bloccare l'avanzamento di uno o più ribosomi alle sue spalle. L'evoluzione dinamica è dunque definita come segue: partendo da una configurazione qualsiasi con N particelle, si sceglie randomicamente un sito k ; se il sito k è pieno e il sito $k+1$ è vuoto allora la particella si sposterà da k a $k+1$. Definendo s la configurazione iniziale e s' la configurazione dopo che la particella si è spostata sul sito $k+1$ il rate di salto da s a s' è:

$$w_{ss'} = \begin{cases} 1 & \text{se transizione permessa} \\ 0 & \text{se transizione proibita} \end{cases}$$

È intuitivo immaginare che dopo un tempo transiente il sistema si troverà in uno stato stazionario, ovvero uno stato le cui proprietà statistiche sono indipendenti dal tempo.

L'introduzione del modello TASEP è dunque una prima approssimazione del processo traslatorio del ribosoma e ne riprende i tratti principali: il processo sequenziale del ribosoma che include l'accoppiamento codone-tRNA, la produzione dell'amminoacido e l'abbandono del tRNA nella sub-unità E del ribosoma, riassumendo quindi le reazioni chimiche nella semplice presenza o assenza del ribosoma su di un particolare sito, approssimando come istantanei i processi chimici intermedi e dando rilevanza solo al processo di accoppiamento (salto sul sito i -esimo) e di uscita del tRNA (salto sul sito $(i+1)$ -esimo e uscita dal sito i -esimo). Il modello TASEP permette quindi lo studio della densità e della corrente di spostamento tramite la dinamica delle transizioni permesse, includendo così anche il processo di esclusione che impedisce a due o più ribosomi di sostare su di uno stesso codone contemporaneamente.

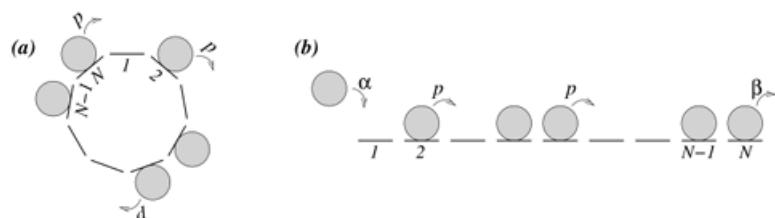


Figura 3: Nella figura è possibile vedere una rappresentazione schematica del modello TASEP. (a) coincide con la condizione di periodic boundary, (b) rappresenta la condizione di open boundary. [2]

2.2.1 Periodic Boundary

La condizione al contorno più semplice da poter studiare è quella di condizioni al contorno periodiche (o **periodic boundary conditions**): preso un reticolo ad N siti, la particella che salta a destra del N -esimo sito si sposta sul primo sito, immaginando così il filamento di mRNA come un cerchio chiuso su se stesso (fig.2.2(a)).

Definendo s come la configurazione iniziale è possibile introdurre la master equation: l'equazione che governa la probabilità di salto da una configurazione della catena all'altra.

La master equation descrive la derivata nel tempo della probabilità di trovarsi in una certa configurazione s come la somma delle probabilità di essere in una configurazione s' per il rate di transizione da s' a s meno la somma delle probabilità di essere nella configurazione s per il rate di transizione da s a s' . Dunque definisce la variazione della probabilità nel tempo come il flusso entrante in una data configurazione meno il flusso uscente da quella configurazione.

$$\frac{dp(s)}{dt} = \sum_{s'} p(s') w_{s' \rightarrow s} - \sum_s p(s) w_{s \rightarrow s'}$$

La condizione di stato stazionario implica che la probabilità di trovarsi in una data configurazione è invariante nel tempo ($p^{st}(s)$) dunque la master equation può essere riscritta come:

$$\frac{dp(s)}{dt} = 0 = \sum_{s'} p(s') w_{s' \rightarrow s} - \sum_s p(s) w_{s \rightarrow s'} \quad (1)$$

Inoltre il numero di configurazioni da cui si può entrare in s ($N_{in}(s)$) è uguale al numero delle configurazioni da cui si può uscire da s ($N_{out}(s)$). Questo è facilmente intuitivo grazie alla definizione del modello, infatti le configurazioni si differenziano semplicemente dallo spostamento verso destra (configurazione verso s) o verso sinistra (configurazione uscente da s) di una particella.

Tramite la definizione dei rate $w_{s \rightarrow s'}$ e $w_{s' \rightarrow s}$, dove gli unici valori ammessi sono 1 se la transizione è permessa e 0 se la transizione è proibita, unita alla condizione per cui $N_{out}(s) = N_{in}(s)$ è possibile scrivere la seguente uguaglianza:

$$N_{in}(s) = \sum_{s'} w_{ss'} = \sum_{s'} w_{s's} = N_{out}(s)$$

Utilizzando la master equation (1) e l'equazione sovrastante è possibile ricavare che nella condizione di stato stazionario $p^{st}(s) = \text{cost}$, ovvero che ogni configurazione degli stati microscopici è equiprobabile. Quindi la probabilità di essere in una data configurazione s in una condizione di stato stazionario è equiprobabile per qualsiasi sia la configurazione s considerata. Inoltre l'equiprobabilità può anche essere vista come una conseguenza dell'invarianza traslazionale di una catena periodica, visualizzabile come un anello.

È importante sottolineare che solo il numero di stati coincide, la particolare configurazione è differente tra gli stati di partenza e quelli di arrivo poiché $w_{ss'} = 1$ implica $w_{s's} = 0$.

Nella situazione di stato stazionario le due quantità principali che possono essere studiate sono:

- \mathbf{p}_i : probabilità che il sito i sia occupato;
- $\mathbf{J}_{i,i+1}$: la corrente di particelle tra il sito i e $i+1$. $J_{i,i+1}$ è diversa da zero se e solo se il sito i è occupato e il sito $i+1$ è vuoto.

L'equiprobabilità degli stati microscopici implica assenza di correlazione tra i diversi siti, ovvero la probabilità che il sito i sia occupato è indipendentemente dallo stato del sito j -esimo. Questa proprietà dello stato stazionario ci permette di utilizzare l'approssimazione di campo medio con esattezza e dunque possiamo calcolare:

$$\begin{aligned} p_i &= \langle n_i \rangle \\ J_{i,i+1} &= \langle n_i \times (1 - n_{i+1}) \rangle \end{aligned}$$

Dove $\langle \rangle$ indica il valor medio e $n_i = 0$ se il sito è vuoto o $n_i = 1$ se è pieno.

Grazie all'invarianza traslazionale dello stato stazionario sia $J_{i,i+1}$ che p_i sono indipendenti dal particolare sito:

- $p_i = p = \frac{N}{L}$ con $N \equiv \#$ particelle presenti nel reticolo = $cost$, $L \equiv \#$ totale di siti;
- $J_{i,i+1}$ partendo dal correlatore $\langle n_i n_{i+1} \rangle = \frac{\sum_s p^{st}(s) \times n_i(s) \times n_{i+1}(s)}{\sum_s p^{st}(s)}$ e applicando l'equiprobabilità degli stati microscopici (ovvero $p^{st}(s) = cost$) si può ridurre l'equazione a

$$\langle n_i n_{i+1} \rangle = \frac{\# \text{ stati in cui } i \text{ e } i+1 \text{ sono occupati}}{\# \text{ totale delle configurazioni}} = \frac{N(N-1)}{L(L-1)}$$

Dove l'unica correlazione è dovuta alla limitatezza del nostro sistema: una volta che un sito è occupato con probabilità $\frac{N}{L}$ avremo $N-1$ particelle e $L-1$ siti e dunque la probabilità sarà $\frac{N-1}{L-1}$. Prendendo il limite $N, L \gg 1$ la correlazione diventerà infinitesimale e dunque

$$\langle n_i n_{i+1} \rangle = \langle n_i \rangle \times \langle n_{i+1} \rangle \quad \text{che implica} \quad J_{i,i+1} = p(1-p)$$

Ed è facile intuire che la corrente massima si avrà per $p = \frac{1}{2}$ e si avrà $J_{i,i+1} = 0$ per $p = 0; 1$

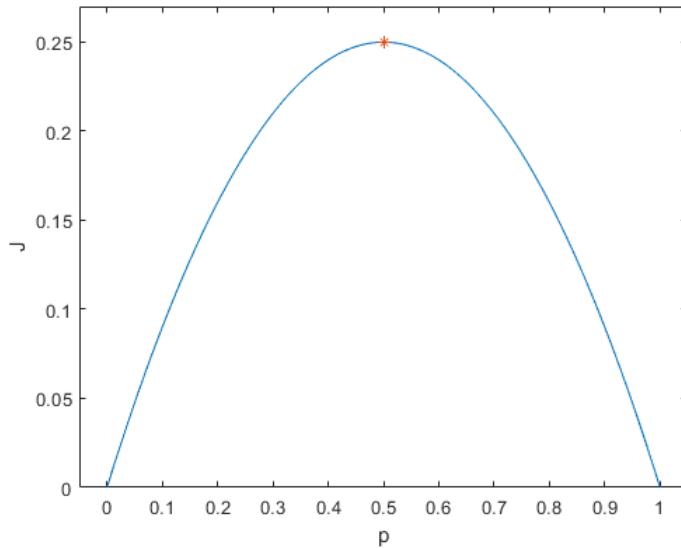


Figura 4: Grafico della corrente al variare della probabilità.

2.2.2 Open Boundary

La seconda condizione al contorno che si può studiare è quella di condizioni al contorno aperte (o **open boundary conditions**) (fig. 2.2(b)), situazione che riprende il processo biologico in modo più reale. Infatti il processo di traduzione avviene all'interno del citoplasma in cui i ribosomi si muovono e si accoppiano ad un'estremità del filamento di mRNA o vengono rilasciati all'interno del citoplasma stesso. La condizione di limiti aperti dunque emula l'ambiente citoplasmatico della particella in cui i ribosomi sono liberi di muoversi e non sono intrappolati nel loop periodico del filamento di RNA messaggero che non è circolare, ma più simile ad un segmento retto con un inizio ed una fine.

Le particelle, ovvero i ribosomi, possono essere introdotte nel reticolo con un rate α nel primo sito ($i = 1$) e essere rimosse con un rate β dall'ultimo sito ($i = L$).

La nuova dinamica seguirà queste regole:

- Si sceglie un sito k randomicamente (con $k \in [0, L]$)
- Se $k = 0$ e $n_1 = 0$ viene introdotta una nuova particella nel sistema con un rate α
- Se $k \in (0, L)$ con $n_k = 1$ e $n_{k+1} = 0$ la particella presente in k viene mossa in $k + 1$
- Se $k = L$ e $n_L = 1$ viene rimossa una particella con rate β

Dalla dinamica è facile intuire che l'invarianza traslazionale non è più valida e che quindi l'approssimazione di campo medio perde di efficacia, ma rimane comunque un approccio valido in prima istanza.

Una configurazione stazionaria è caratterizzata da una media indipendente dal tempo e dunque, nonostante $p_i = \langle n_i \rangle$ può essere dipendente dal sito i , la corrente $J_{i,i+1}$ deve essere indipendente da i , poiché uno sbilanciamento tra $J_{i-1,i}$ e $J_{i,i+1}$ implicherebbe una variazione temporale di p_i , quindi

$$J_{i,i+1} = J_{i-1,i} = J$$

$$\frac{d(\langle n_i \rangle)}{dt} = \frac{d(p_i)}{dt} = \langle n_{i-1}(1 - n_i) \rangle - \langle n_i(1 - n_{i+1}) \rangle$$

Ed applicando un'approssimazione di campo medio:

$$\frac{dp_i}{dt} = p_{i-1}(1 - p_i) - p_i(1 - p_{i+1})$$

Si ottengono quindi $L + 1$ equazioni:

$$\begin{aligned} \alpha(1 - p_i) &= J \\ p_i(1 - p_{i+1}) &= J \quad \text{con } i = 1, 2, \dots, L-1 \\ \beta p_L &= J \end{aligned}$$

Introducendo due siti fintizi $i = 0$ e $i = L + 1$ con $p_0 = \alpha$ e $p_{L+1} = (1 - \beta)$ è possibile utilizzare solamente l'equazione:

$$p_i(1 - p_{i+1}) = J \text{ con } i = 0, 1, 2, \dots, L$$

Studio della stabilità

L'equazione ricorsiva di p_i ci permette di studiarne la stabilità e i punti fissi utilizzando le conoscenze generali sulle mappe. Una mappa è un'equazione ricorsiva definita da una funzione $x_{n+1} = F(x_n)$ e un suo punto fisso è un punto x^* tale che $x^* = F(x^*)$ e se $x_0 = x^*$ allora $x_n = x^* \forall n$. È possibile eseguire un'analisi lineare della stabilità quando $x_0 - x^* \ll 1$. Definito $x_n = x^* + \epsilon_n$ si ha $\epsilon_{n+1} = F'(x^*)\epsilon_n$ che ha come soluzione $\epsilon_n = (F'(x^*))^n\epsilon_0$. La stabilità dipende quindi dal valore di $|F'(x^*)|$:

- se $|F'(x^*)| < 1$ allora $\epsilon_n \rightarrow 0$ e il punto fisso è stabile
- se $|F'(x^*)| > 1$ allora $\epsilon_n \rightarrow \infty$ esponenzialmente e il punto fisso è instabile
- se $|F'(x^*)| = 1$ si rende necessario uno studio al secondo ordine

È possibile dunque riscrivere l'equazione come:

$$p_{i+1} = 1 - \frac{J}{p_i} = F(p_i) \quad \text{con } i = 0, 1, 2, \dots, L$$

I punti fissi sono le soluzioni dell'equazione di secondo grado in p : $p_a = \frac{1}{2} - \sqrt{\frac{1}{4} - J}$ e $p_b = 1 - p_a$.

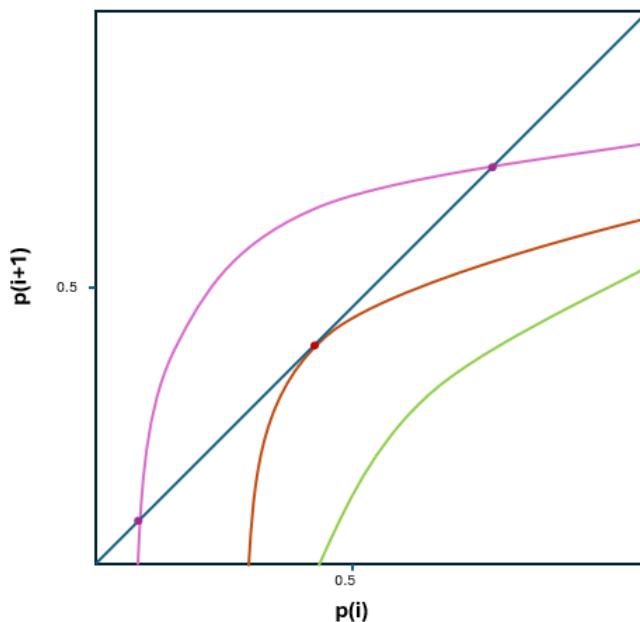


Figura 5: Punti fissi al variare di J

Come è possibile vedere nel grafico 5 al variare di J si possono avere uno, due o nessun punto fisso. Tenendo a mente che per le condizioni al contorno di limiti aperti si ha $p_0 = \alpha$ e $p_{L+1} = 1 - \beta$, si può entrare nel dettaglio delle soluzioni ottenute.

Nel caso in cui $J < \frac{1}{4}$ si avrà un punto fisso stabile in $p = 1 - p_a$ con bacino di attrazione per $p > p_a$: un qualsiasi $p_0 = \alpha$ che ricade in questa casistica evolverà verso di esso secondo la relazione

$$\lim_{i \rightarrow \infty} p_i = 1 - p_a = 1 - \beta \quad \text{con} \quad p_a < \frac{1}{2} \implies \beta < \frac{1}{2} \text{ e } \alpha > \beta.$$

Dalla relazione soprastante si può dedurre che la densità è maggiore di $\frac{1}{2}$ quasi ovunque, ovvero le particelle che vengono iniettate nel sistema sono maggiori di quelle che vengono espulse con una densità di particelle che diventa costante avvicinandosi all'uscita. Questa regione dello spazio della dinamica è chiamata **High Density Phase (HD)** ed è possibile calcolare il valore della corrente come $J = \beta(1 - \beta)$.

La dinamica delle particelle che si muovono verso destra può essere compresa, ribaltando il punto di vista, come la dinamica dei siti vuoti che si muovono verso sinistra. β può essere dunque visto come il rate di inserimento di un sito vuoto all'estrema destra e α il rate della rimozione di un sito vuoto all'estrema sinistra. Poiché la densità dei siti vuoti è uno meno la densità delle particelle si può intuire che al profilo di densità HD delle particelle corrisponde una situazione di Low density per i siti vuoti.

Nel caso in cui $\alpha < \frac{1}{2}$ e $\beta > \alpha$ (situazione simmetrica alla precedente) si avrà una condizione di **Low Density Phase (LD)** per le particelle e il punto fisso stabile a cui tende la dinamica sarà $p = p_a$ con corrente $J = \alpha(1 - \alpha)$.

È possibile applicare lo stesso ragionamento a $J = \frac{1}{4}$. Valore per cui si ha $\alpha = \frac{1}{2}$ $\beta > \frac{1}{2}$ e $\beta = \frac{1}{2}$ $\alpha > \frac{1}{2}$.

Per $\alpha, \beta > \frac{1}{2}$ si ha la fase denominata di **Maximal Current (MC)** in cui per $0 < J - \frac{1}{4} \ll 1$ la mappa è stabile intorno a $p = \frac{1}{2}$ per un tempo che cresce esponenzialmente tanto più $J - \frac{1}{4} \rightarrow 0$. Questa fase è chiamata di corrente massima proprio perché in questa fase la corrente raggiunge il suo valore massimo $J = \frac{1}{4}$.

Infine, per completezza, è utile far notare che per $J \gg \frac{1}{4}$ in pochi step la mappa raggiunge il punto [0;1] che è fuori dal dominio fisico.

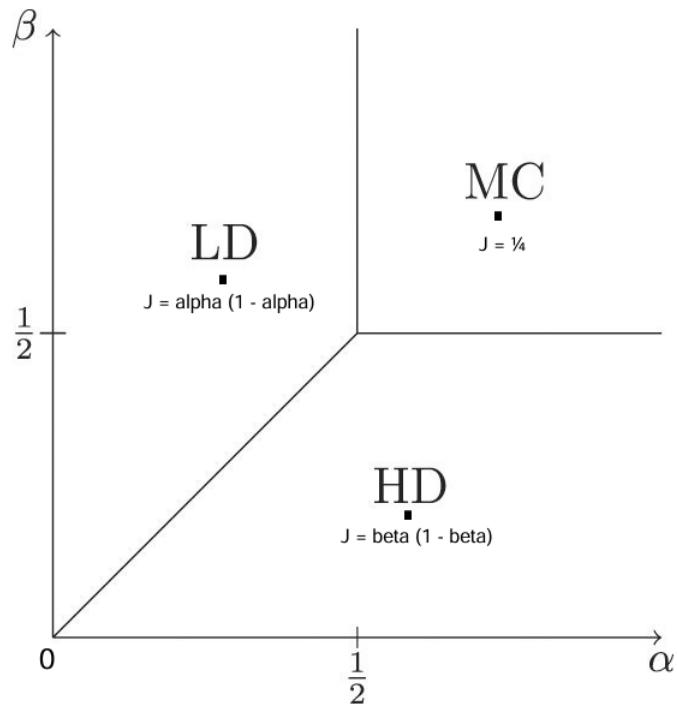


Figura 6: Diagramma di fase del modello TASEP originale [7]

3 Modello per la traduzione del mRNA

3.1 Modello a due stati

I dati sperimentali indicano che l'accoppiamento tra tRNA e codone è lo step limitante del processo di traduzione. È possibile definire un grado di libertà interno della particella (o stato interno) legato al momento del processo in cui si trova il particolare ribosoma sul sito i-esimo. Ciandrini et al. (2010) [3] definiscono un processo a due stati prendendo in considerazione l'effetto limitante del processo di accoppiamento: il passaggio dal primo stato al secondo stato corrisponde al processo di accoppiamento; il processo di traslocazione da un sito all'altro avviene poi solo a partire da una particella nel secondo stato. Per semplicità di calcolo si assume che l'ambiente citoplasmatico in cui mRNA e ribosomi sono immersi sia omogeneo, ovvero la concentrazione dei tRNA è omogenea e l'ambiente grande abbastanza permette di trascurare le fluttuazioni rendendo costante nel tempo i rate di transizione e indipendente dal particolare codone il rate di accoppiamento.

Come già accennato il filamento di mRNA è rappresentato come un reticolo i cui siti coincidono coi codoni e i ribosomi sono le particelle che si muovono tra di essi. Il grado di libertà interno della particella è definito tramite il *numero di occupazione* n_i :

- $n_i = 0$ coincide col sito vuoto
- $n_i = 1; 2$ rappresenta rispettivamente il primo e il secondo stato, ovvero il ribosoma prima e dopo l'accoppiamento codone-anticodone

La configurazione del sistema è rappresentata da $\eta = n_1, n_2, \dots, n_L$ dove L è la lunghezza del reticolo. La dinamica è caratterizzata, come descritto nel capitolo precedente, da delle regole definite sia dal modello che dalla sequenzialità dei processi del ribosoma stesso. Le transizioni dinamiche ammesse sono solamente due:

$$\begin{cases} 1 \rightarrow 2 & \text{con rate } k_i \\ 20 \rightarrow 01 & \text{con rate } \gamma_i \end{cases}$$

La *prima transizione* si riferisce semplicemente al passaggio dal primo al secondo stato. Il rate di transizione (k_i) in questo caso è, in generale, dipendente dal particolare sito; tuttavia, seguendo il lavoro di Ciandrini e colleghi [3], verrà studiato il caso in cui $k_i = k = \text{cost}$.

La *seconda transizione* dinamica è una rappresentazione di ciò che succede quando il ribosoma, uscito dal secondo stato, lascia un codone e atterra su quello successivo. In questa seconda operazione, come visto, il ribosoma torna ricettivo a nuovi tRNA e quindi passa nuovamente al primo stato.

Densità media

È possibile suddividere la densità media delle particelle sul sito i (ρ_i) in due densità medie che rappresentano i due stati in cui si può trovare il ribosoma.

λ_i è la densità media delle particelle sul sito i mentre si trovano nello stato 1; viceversa σ_i rappresenta la densità media delle particelle sul sito i nel secondo stato.

Entrambe le densità posso essere scritte in funzione del numero di occupazione:

$$\begin{aligned} \lambda_i &= \langle n_i(2 - n_i) \rangle \\ \sigma_i &= \left\langle \frac{n_i(n_i - 1)}{2} \right\rangle \end{aligned}$$

Dove $\langle \rangle$ indica una media temporale.

La densità media totale si può facilmente calcolare come $\rho = \lambda_i + \sigma_i$.

L'assunzione di assenza di correlazione fra i vari siti, ovvero che $\langle n_i n_j \rangle \sim \langle n_i \rangle \langle n_j \rangle$ permette di semplificare l'analisi e di riscrivere le equazioni di campo medio che governano l'evoluzione della densità media nei vari siti:

$$\frac{d\lambda_i}{dt} = \sigma_{i-1}(1 - \lambda_i - \sigma_i)\gamma - k\lambda_i \quad (2)$$

$$\frac{d\sigma_i}{dt} = k\lambda_i - \sigma_i(1 - \lambda_{i+1} - \sigma_{i+1})\gamma \quad (3)$$

Corrente

Definizione

1. Il numero di particelle che entrano nel sito i per unità di tempo è la corrente J_+^i .
2. Il numero di particelle che escono dal sito i per unità di tempo è la corrente J_-^i .

Partendo dalle equazioni della densità media è possibile rappresentare le correnti J_\pm^i come segue:

$$J_+^i = \sigma_{i-1}(1 - \lambda_i - \sigma_i)\gamma \quad (4)$$

$$J_-^i = \sigma_i(1 - \lambda_{i+1} - \sigma_{i+1})\gamma \quad (5)$$

In ultimo ricordiamo che la condizione di "steady state" o stato stazionario è rappresentata da $\frac{d\lambda_i}{dt} = \frac{d\sigma_i}{dt} = 0$; dove la corrente è costante lungo tutto il reticolo $J_+^i = J_-^i = J \quad \forall i$.

3.2 Soluzione di campo medio

3.2.1 Periodic boundary

L'analisi comincia dalla condizione di *periodic boundary* come per i modelli trattati nel capitolo precedente. Come accennato in precedenza l'analisi viene eseguita nel caso speciale in cui $k_i = k$ e dunque è possibile considerare ogni sito identico ed eliminare l'indicizzazione.

Le equazioni (2,3,4,5) posso essere riscritte notando che il numero medio di particelle nel primo stato è uguale alla densità locale λ e in ugual modo il numero medio di particelle nel secondo stato alla densità locale σ :

$$\lambda = \frac{\langle \sum_{i=1}^L n_i \delta_{n_i,1} \rangle}{L} \quad (6)$$

$$\sigma = \frac{\langle \sum_{i=1}^L n_i \delta_{n_i,2} \rangle}{2L} \quad (7)$$

e le equazioni precedenti possono essere scritte come:

$$J = (1 - \lambda - \sigma)\gamma \quad (8)$$

$$\lambda = \frac{J}{k} \quad (9)$$

$$\sigma = \rho - \lambda = \rho - \frac{J}{k} \quad (10)$$

Utilizzando ρ come parametro di controllo è possibile riscrivere le equazioni soprastanti come:

$$J = \frac{\rho(1 - \rho)k}{\frac{k}{\gamma} + (1 - \rho)} \quad (11)$$

$$\lambda = \frac{\rho(1 - \rho)}{\frac{k}{\gamma} + (1 - \rho)} \quad (12)$$

$$\sigma = \frac{\frac{k}{\gamma}\rho}{\frac{k}{\gamma} + (1 - \rho)} \quad (13)$$

Fissato un valore arbitrario per γ è possibile ricavare il valore di ρ per cui la corrente è massima:

$$\rho^* = 1 + \frac{k}{\gamma} - \frac{k}{\gamma} \sqrt{1 + \frac{\gamma}{k}} = 1 - \chi \quad (14)$$

Da cui si possono ricavare i valori massimali di λ e σ :

$$\lambda^* = 1 + \frac{2k}{\gamma} \left(1 - \sqrt{1 + \frac{\gamma}{k}}\right) = 1 - 2\chi \quad (15)$$

$$\sigma^* = \rho^* - \lambda^* = \chi \quad (16)$$

Dalle equazioni (11 - 16) è evidente la dipendenza di λ e σ dal rapporto $\frac{k}{\gamma}$, dipendenza che è confermata dalle simulazioni e dalle soluzioni analitiche del sistema.

Il sistema sovrastima la corrente J solamente quando $\frac{k}{\gamma} \ll 1$, poiché la correlazione non può più essere trascurata.

Dall'equazione (14) è possibile notare che con l'incremento del rapporto $\frac{k}{\gamma}$ la densità tende al valore massimo di 0.5.

Il caso limite $\frac{k}{\gamma} \rightarrow \infty$ coinciderebbe con il considerare la transizione da stato uno a stato due istantaneo, ricollegando il modello al caso classico di TASEP con un singolo rate di salto γ .

Infine si può notare come le curve di λ e σ al variare di ρ si incrociano, nel caso in cui $k < \gamma$, nel punto $\rho_d = 1 - \frac{k}{\gamma}$. Si possono dedurre quindi due diversi regimi:

- $\rho < \rho_d$ identifica un regime per cui $\lambda > \sigma$, ovvero il reticolo è popolato da ribosomi nella fase iniziale del legame con il codone;
- $\rho > \rho_d$ rappresenta la situazione opposta in cui il reticolo è maggiormente popolato di ribosomi pronti a saltare sul sito successivo;

Per $k \geq \gamma$ l'unico regime possibile è $\lambda < \sigma$.

3.2.2 Open boundary

Nella condizione di *open boundary* α è il rate con cui una particella viene inserita nel reticolo e β è il rate con cui lo abbandona.

La dinamica all'interno del reticolo segue le stesse regole che definivano la dinamica nelle condizioni periodiche al contorno.

All'interno del reticolo quindi la dinamica e le equazioni che la governano sono le stesse studiate nella sezione precedente.

Nella fase di iniezione, ovvero la fase in cui le particelle vengono introdotte nel reticolo, la dinamica sarà come segue:

$$\frac{d\lambda_1}{dt} = \alpha(1 - \lambda_1 - \sigma_1) - k\lambda_1 \quad (17)$$

$$\frac{d\sigma_1}{dt} = k\lambda_1 - \sigma_1(1 - \lambda_2 - \sigma_2)\gamma \quad (18)$$

Analogamente nella fase in cui la particella abbandona il reticolo l'evoluzione dinamica sarà:

$$\frac{d\lambda_L}{dt} = \sigma_{L-1}(1 - \lambda_L - \sigma_L)\gamma - k\lambda_L \quad (19)$$

$$\frac{d\sigma_L}{dt} = k\lambda_L - \beta\sigma_L \quad (20)$$

Stato stazionario

Nello stato stazionario tramite le equazioni (2, 3) è possibile ricavare una mappa iterativa per la densità σ e i rispettivi punti fissi:

$$\sigma_{i+1} = 1 - J\left(\frac{1}{k} + \frac{1}{\gamma\sigma_i}\right) \quad (21)$$

$$\sigma_{\pm} = \frac{1}{2}\left(1 - \frac{J}{k}\right) \pm \sqrt{\left(1 - \frac{J}{k}\right)^2 - \frac{4J}{\gamma}} \quad (22)$$

È possibile riscrivere l'equazione ricorsiva per ricavare il generico valore σ_i , utilizzando il valore di σ_1 :

$$\sigma_i = \frac{-\sigma_- \sigma_+ (\sigma_+^{i-1} - \sigma_-^{i-1}) + \sigma_1 (\sigma_+^i - \sigma_-^i)}{-\sigma_- \sigma_+ (\sigma_+^{i-2} - \sigma_-^{i-2}) + \sigma_1 (\sigma_+^{i-1} - \sigma_-^{i-1})} \quad (23)$$

Studio delle fasi

Come per il modello originale ad uno stato il sistema mostra tre differenti regimi: *Low Density Phase*, *High Density Phase*, *Maximal Current Phase*.

- **Low Density Phase** – $\sigma_1 \simeq \sigma_-$, $\sigma_L < \sigma_+$:

Le condizioni da dover soddisfare per trovarsi in questo regime sono $\alpha \leq \alpha_{crit} = \gamma\chi$ e $\beta > \alpha$.

Partendo ad iterare la mappa vicino al punto instabile, il valore di σ_i rimane inizialmente vicino a questo valore per poi allontanarsi.

Risolvendo le equazioni precedenti si ottengono le soluzioni:

$$\sigma_1 = \frac{\alpha}{\gamma} \quad (24)$$

$$\sigma_L = \frac{\alpha k (\gamma - \alpha)}{\beta \gamma (k + \alpha)} \quad (25)$$

$$J = \frac{\alpha k (\gamma - \alpha)}{\gamma (k + \alpha)} \quad (26)$$

- **High density phase** – $\sigma_1 > \sigma_-$, $\sigma_L \simeq \sigma_+$:

La fase di alta densità si ottiene solo nel caso in cui $\beta \leq \beta_{crit} = \gamma\chi$, $\beta < \alpha$.

A differenza di prima lo starting point si trova nel bacino di attrazione del punto stabile (σ_+), dunque dopo l'iterazione della mappa, il punto σ_i sarà arbitrariamente vicino a σ_+ .

La soluzione che si ricava è la seguente:

$$\sigma_1 = \frac{\alpha k \gamma - \beta k \gamma + \alpha \beta^2 + k \beta^2}{\alpha \gamma (\beta + k)} \quad (27)$$

$$\sigma_L = \frac{k (\gamma - \beta)}{\gamma (k + \beta)} \quad (28)$$

$$J = \frac{\beta k (\gamma - \beta)}{\gamma (k + \beta)} \quad (29)$$

- **Maximal Current Phase** – $\sigma_1 \geq \frac{1}{2} > (1 - \frac{J}{k})$, $\sigma_1 \leq \frac{1}{2}(1 - \frac{J}{k})$:

Questa regime si raggiunge quando i due punti fissi σ_{\pm} collassano in uno e il reticolo ha la massima corrente permessa. Questa condizione si ha solo quando $\alpha \geq \alpha_{crit}$, $\beta \geq \beta_{crit}$.

È possibile calcolare dunque:

$$\sigma_1 = 1 - \frac{J}{k} - \frac{J}{\alpha} \quad (30)$$

3.3 Maximal Current Principle

Il *Maximal Current Principle* (MCP), sviluppato inizialmente da Krug negli anni ottanta e portato avanti nei primi anni duemila, permette di ottenere gli stessi risultati visti in precedenza.

Questo principio sostituisce le condizioni al contorno aperte con delle riserve di particelle agli estremi del reticolo; la dinamica presente all'interno del reticolo è la stessa tra le riserve di particelle e il reticolo stesso.

La corrente J nella fase di Maximal Current di un reticolo "open boudary" è vista secondo il principio MCP come:

$$J = \max_{\rho \in [\rho_{L+1}, \rho_0]} J(\rho)$$

Dove ρ_0 e ρ_{L+1} sono rispettivamente le densità delle riserve alla sinistra e alla destra del reticolo. Per poter ottenere gli stessi risultati basta prendere dei valori fissi per le densità delle riserve di particelle e calcolare di conseguenza i rate di iniezione e abbandono (α, β).

Immaginando di avere un sito extra per $i = 0$, che corrisponde ad una riserva di particelle alla sinistra del reticolo, con una densità $\rho_0 = \lambda_0 + \sigma_0$ si può riscrivere α come $\alpha = \gamma P(n_0 = 2)$, poiché $P(n_0 = 2) = \sigma_0$. Quindi si può riscrivere il parametro di iniezione $\alpha = \sigma_0 \gamma$. Analogamente si può ripetere il ragionamento per il sito $i = L + 1$ (la riserva alla destra del reticolo) e ottenere il parametro $\beta = (1 - \rho_{L+1}) \gamma$.

Ricavato l'andamento dei valori di α e β per poter ricavare i valori critici per cui si hanno le transizioni di fase è possibile definire:

$$\sigma_0 = \sigma^* = \rho^* + \lambda^* \quad (31)$$

$$\rho_{L+1} = \rho^* \quad (32)$$

Ottenendo gli stessi valori critici ottenuti in precedenza α_{crit} e β_{crit} e i rispettivi valori di corrente e densità per le diverse fasi possibili:

$$J = \begin{cases} \frac{\alpha k(\gamma - \alpha)}{\gamma(k + \alpha)} & \text{per } \alpha < \beta < \gamma\sigma^* \quad (\textbf{LD}) \\ \frac{\beta k(\gamma - \beta)}{\gamma(k + \beta)} & \text{per } \beta < \alpha < \gamma\sigma^* \quad (\textbf{HD}) \\ k(1 - 2\sigma^*) & \text{per } \alpha, \beta \geq \gamma\sigma^* \quad (\textbf{MC}) \end{cases}$$

$$\rho = \begin{cases} \rho_0 & \text{per } \alpha < \beta < \gamma\sigma^* \quad (\textbf{LD}) \\ \rho_{L+1} & \text{per } \beta < \alpha < \gamma\sigma^* \quad (\textbf{HD}) \\ 1 - \sigma^* & \text{per } \alpha, \beta \geq \gamma\sigma^* \quad (\textbf{MC}) \end{cases}$$

Grazie ai risultati ottenuti si può notare che le tre regioni che determinano la dinamica (LD, HD, MC) sono le stesse previste dai primi modelli; ciò che è possibile approfondire col modello introdotto in questo capitolo è la transizione tra le diverse fasi: la transizione verso la regione di Maximal Current è continua; la transizione tra LD e HD è, invece, una discontinua ed è definita da una linea retta sul piano $\alpha - \beta$ e coincide con la condizione di $\alpha = \beta$, infatti i due parametri critici hanno la stessa dipendenza da k e γ .

In ultimo, considerando il caso limite $k \rightarrow \infty$, i valori critici si approssimano ai valori ottenuti nel modello TASEP standard, poiché corrisponde al considerare la transizione $n_i = 1 \rightarrow n_i = 2$ istantanea.

4 TASEP inomogeneo

4.1 Definizione del modello

I ribosomi possono muoversi con diverse velocità sul filamento del RNA messaggero. La variazione nella velocità è stata dimostrata essere legata alla diversa disponibilità dei tRNA che portano il corretto amminoacido al ribosoma, infatti è stato osservato che la maggior parte degli amminoacidi è codificata da 2-6 codoni con una frequenza non randomica.

Ciandrini et al. (2023) [4] considerano questa diversa velocità di traduzione attraverso il TASEP inomogeneo, ovvero un modello TASEP in cui i rate di salto variano in dipendenza del sito considerati. Questa modellizzazione a differenza del modello omogeneo non ha una risoluzione analitica esatta. In prima istanza è possibile utilizzare un'approssimazione di campo medio, trascurando la correlazione tra i diversi siti, giustificata solamente nel caso in cui i rate di salto variano lentamente lungo il reticolo e il numero di siti è grande. Utilizzando l'approssimazione di campo medio e la soluzione analitica del modello omogeneo è possibile ricavare un'equazione per la densità media.

In alternativa è possibile ricavare una soluzione più accurata tramite un'espansione in serie di potenza, o Power Series Approximation (PSA), tenendo conto della correlazione tra i diversi siti. L'espansione in serie di potenze è eseguita rispetto al rate di inizializzazione delle particelle, ovvero il rate con cui il ribosoma si lega al codone iniziale. Essendo il rate di inizializzazione il rate limitante la dinamica, ovvero il valore minore dell'intero processo (e quindi il processo più lento), solo i primi termini sono quelli dominanti.

A differenza del modello precedente, considerando un reticolo con L siti, il ribosoma è ora un oggetto esteso definito da due caratteristiche principali:

- ℓ rappresenta la lunghezza del ribosoma e indica il numero di siti che saranno occupati da parte del ribosoma stesso;
- il **tracking site** è un sito del ribosoma scelto in maniera arbitraria che ci permette di individuare la posizione del ribosoma stesso, in questa trattazione sarà il sito centrale

Come in precedenza il rate α è il rate per cui un ribosoma entra nel reticolo, ciò che realmente cambia è che un ribosoma entra nel reticolo solo se i primi ℓ siti del reticolo sono liberi, dunque per poter introdurre un nuovo ribosoma il tracking site del ribosoma successivo presente sul reticolo dovrà trovarsi sul sito $\ell + 1$. β è il rate di uscita dal reticolo dall'ultimo sito $i = L$. Il ribosoma abbandona il reticolo quando lo fa col suo tracking site.

ω_i è il rate di passaggio dal sito i al sito $i+1$ e come per l'iniezione è necessario che il tracking site successivo si trovi a $i + \ell$.

Definendo τ_i come il numero di occupazione del sito i tale che:

$$\tau_i = \begin{cases} 1 & \text{se il tracking site si trova sul sito } i \\ 0 & \text{altrimenti} \end{cases}$$

È possibile definire la configurazione C del sistema $[\tau_1, \tau_2, \dots, \tau_L]$.

La dinamica può essere dunque riassunta attraverso il numero di occupazione:

Inizializzazione $\tau_1 = 0 \xrightarrow{\alpha} 1$ con $\tau_1 = \dots = \tau_\ell = 0$

Elongazione $[\tau_i, \tau_{i+1}] = [1, 0] \xrightarrow{\omega_i} [0, 1]$ con $\tau_{i+\ell} = 0$ per $i = 1, \dots, L - \ell$

Terminazione $\tau_L = 1 \xrightarrow{\beta} 0$

Esattamente come per i modelli precedenti l'idea di comprendere la dinamica del modello e approfondirne la comprensione passa dallo studio della corrente e della densità media delle particelle nello stato stazionario.

Trascurando la correlazione tra le particelle e i vari siti la corrente è conservata lungo tutto il reticolo e si può scrivere come:

$$J = \alpha \langle \prod_{i=1}^{\ell} (1 - \tau_i) \rangle = \beta \langle \tau_L \rangle \quad (33)$$

L'equazione della corrente può essere vista come la corrente di iniezione nel reticolo, ovvero il rate α per la probabilità che i primi ℓ siti siano vuoti, e come la corrente di uscita per la probabilità che l'ultimo sito sia pieno.

A questo punto possiamo definire la densità locale (ρ_i), ovvero la probabilità di trovare un tracking site nel sito i , e la densità media (ρ):

$$\rho_i = \langle \tau_i \rangle \quad (34)$$

$$\rho = \frac{1}{L} \sum_{i=1}^L \rho_i \quad (35)$$

4.1.1 Master Equation

Prima di approfondire la master equation rispettata dalla probabilità di stato stazionario $\mathbf{P}(\mathbf{C})$ ridefiniamo la configurazione C come l'insieme delle coordinate x_m che indicano la posizione del tracking site di ogni ribosoma:

$$C = [x_1, \dots, x_i, \dots, x_N] \text{ con } N = \sum_{i=1}^L \tau_i = \# \text{ di ribosomi presenti nel reticolo}$$

Per poter rispettare il principio di esclusione è necessario inoltre avere che:

$$x_m + \ell \leq x_{m+1} \text{ con } m = 1, \dots, L-1$$

È possibile scrivere la master equation:

$$0 = \sum_{C'} W(C' \rightarrow C) P(C') - \sum_{C'} W(C \rightarrow C') P(C) \quad (36)$$

Dove è facile intuire che $W(C' \rightarrow C)$ è il rate di transizione dalla configurazione C' alla configurazione C . Esplicitando i termini si ricava:

$$e(C) P([x_1, \dots, x_N]) = \sum_{m=1}^N \omega_{x_{m-1}} \mathbf{E}^m P([x_1, \dots, x_N]) + \omega_L P([x_1, \dots, x_N, L]) \mathbf{I}_{X_N \leq L-\ell} \quad (37)$$

Con:

$$\begin{cases} \omega_0 = \alpha \text{ e } \omega_L = \beta \\ e(C) = \text{exit rate da } C = \sum_{C'} W(C \rightarrow C') = \omega_0 \mathbf{I}_{x_1 > \ell} + \sum_{m=1}^{N-1} \omega_{x_m} \mathbf{I}_{x_{m+1} - x_m > \ell} + \omega_{x_N} \\ \mathbf{I} = \text{matrice identità} \end{cases}$$

\mathbf{E}^m è l'operatore scala che muove la particella m -esima una posizione a sinistra. Si scrive come:

$$\begin{aligned} \mathbf{E}^m P([x_1, \dots, x_N]) &= \begin{cases} P([x_1, \dots, x_{m-1}, x_m, \dots, x_N]) \text{ per } x_m - x_{m-1} > \ell \\ 0 \text{ per } x_m - x_{m-1} = \ell \end{cases} \\ \mathbf{E}^1 P([1, \dots, x_N]) &= \begin{cases} P([2, \dots, x_N]) \text{ per } N > 1 \\ P(\emptyset) \text{ per } N = 1 \end{cases} \end{aligned}$$

Ricapitolando, quello che è necessario per calcolare la master equation per una configurazione C è calcolare l'exit rate da quella stessa configurazione e prendere in considerazione tutte le possibili configurazioni C' che hanno una qualsiasi particella una posizione più a sinistra rispetto a C (sempre rispettando le transizioni permesse). In questo caso le configurazioni C' rappresentano tutte quelle configurazioni da cui C può essersi generato.

4.2 Power Series Approximation (PSA)

La generazione della master equation vista in precedenza permette di utilizzare la *Power Series Approximation* (PSA), ovvero approssimare la probabilità $P(C)$ con un'espansione in serie di potenze. Il rate utilizzato per l'espansione è il rate iniziale α , poiché si è dimostrato essere molto più piccolo di qualunque altro rate [4]

$$P(C) = \sum_{n=0}^{\infty} c_n(C) \alpha^n \quad (38)$$

Inoltre, poiché la sommatoria su C di tutte le probabilità $P(C)$ è uguale a 1 abbiamo che:

$$\sum_C c_n(C) = \begin{cases} 1 & \text{per } n = 0 \\ 0 & \text{per } n \geq 1 \end{cases} \quad (39)$$

Consideriamo ora $W(C \rightarrow C')$.

Tramite l'operatore $I_{C,C'} = \begin{cases} 1 & \text{se } C \rightarrow C' \text{ corrisponde con l'inizializzazione} \\ 0 & \text{Altrimenti} \end{cases}$

W può essere riscritto come:

$$\begin{cases} W(C \rightarrow C') = \alpha I_{C,C'} + W_0(C \rightarrow C') \\ W_0(C \rightarrow C') = (1 - I_{C,C'})W(C \rightarrow C') \end{cases} \quad (40)$$

E riportando tutto nella master equation (36):

$$\begin{aligned} & \sum_{C'} [I_{C',C} \sum_{n=1}^{\infty} c_{n-1}(C') \alpha^n + W_0(C' \rightarrow C) \sum_{n=0}^{\infty} c_n(C') \alpha^n] \\ &= \sum_{C'} [I_{C,C'} \sum_{n=1}^{\infty} c_{n-1}(C) \alpha^n + W_0(C \rightarrow C') \sum_{n=0}^{\infty} c_n(C) \alpha^n] \end{aligned}$$

Per ottenere una forma per il calcolo dei coefficienti $c_n(C)$ è necessario raccogliere tutti i termini in α^n e porre la loro somma a zero (grazie alla definizione iniziale della master equation):

- $C \neq \emptyset$ e $n \geq 1$:

$$c_n(C) = \frac{1}{e_0(C)} [\sum_{C'} W_0(C' \rightarrow C) c_n(C') + \sum_{C'} c_{n-1}(C') I_{C',C} - c_{n-1}(C) \sum_{C'} I_{C,C'}] \quad (41)$$

$$e_0(C) = \sum_{C'} W_0(C \rightarrow C') = \text{exit rate totale da } C \text{ escludendo l'iniziazione}$$

- $C \neq \emptyset$ e $n = 0$:

$$e_0(C) c_0(C) = \sum_{C'} W_0(C' \rightarrow C) c_0(C') \quad (42)$$

- $C = \emptyset$ e $\forall n$:

$$c_n(\emptyset) = \delta_{n,0} - \sum_{C' \neq \emptyset} c_n(C') \text{ (grazie all'equazione 39)} \quad (43)$$

Dall'equazione (42), ricordando che coincide con la master equation per $\alpha = 0$, è possibile ricavare:

$$c_0(C) = \begin{cases} 1 & \text{per } C = \emptyset \\ 0 & \text{altrimenti} \end{cases} \quad (44)$$

Da cui segue che tutti i coefficienti c_n di ordine n minore di $N(C)$ (# particelle presenti nella configurazione C) è nullo:

$$c_n = 0 \quad \forall n < N(C) \quad (45)$$

Questo permette di semplificare i calcoli nell'approssimazione in serie di potenze rendendola utilizzabile e pratica. Inoltre, grazie a quest'ultima precisazione è possibile vedere l'espansione in serie di potenze come un forma di perturbazione, dove ogni evento di inizializzazione corrisponde ad un ordine della teoria perturbativa, partendo dall'ordine zero che corrisponde con la configurazione vuota.

È possibile finalmente inserire la PSA nell'equazione della corrente, della densità media e della densità locale (equazioni: 33, 34, 35) per ottenere:

$$J = \sum_{n=0}^{\infty} J_n \alpha^{n+1}, \quad J_0 = 1, \quad J_n = c_{n-1}(\emptyset) + \sum_{\substack{C \\ x_1 \geq \ell+1}} c_{n-1}(C) \quad \forall n \geq 1 \quad (46)$$

$$\rho_i = \sum_{n=0}^{\infty} \rho_{i,n} \alpha^n, \quad \rho_{i,0} = 0, \quad \rho_{i,n} = \sum_{\substack{C \\ \tau_i=1}} c_n(C), \quad i = 1, \dots, L \quad \forall n \geq 1 \quad (47)$$

$$\rho = \frac{1}{L} \sum_{i=1}^L \rho_i = \sum_{n=0}^{\infty} \rho_n \alpha^n, \quad \rho_0 = 0, \quad \rho_n = \frac{1}{L} \sum_{i=1}^L \sum_{\substack{C \\ \tau_i=1}} c_n(C) \quad \forall n \geq 1 \quad (48)$$

Dove n, J_n , $\rho_{i,n}$, ρ_n corrispondono all'ordine di espansione della serie di potenze.

4.3 TASEPy

4.3.1 Risoluzione Iterativa del modello disomogeneo

In questa ultima sezione verrà studiata la risoluzione iterativa dei coefficienti c_n e verrà approfondito il pacchetto sviluppato da Ciandrini e colleghi [5] per permettere il calcolo dei coefficienti e di conseguenza della corrente e delle densità fondamentali.

Innanzitutto si può riscrivere l'equazione (41) esplicitando la dipendenza $C = [x_1, \dots, x_N]$, ricordando che $1 < N < n$:

$$c_n([x_1, \dots, x_N]) = \frac{1}{e_0([x_1, \dots, x_N])} \left(\underbrace{c_{n-1}([x_2, \dots, x_N]) \mathbf{I}_{x_1=1}}_a + \underbrace{\omega_{x_1-1} \mathbf{E}^1 c_n([x_1, \dots, x_N]) \mathbf{I}_{x_1>1}}_{b_1} + \right. \\ \left. + \underbrace{\sum_{m=2}^N \omega_{x_m-1} \mathbf{E}^m c_n([x_1, \dots, x_N])}_{b_2} + \underbrace{\omega_L c_n([x_1, \dots, x_N, L]) \mathbf{I}_{x_N \leq L-\ell} \mathbf{I}_{n \geq N+1}}_c + \right. \\ \left. - \underbrace{c_{n-1}([x_1, \dots, x_N]) \mathbf{I}_{x_1>\ell} \mathbf{I}_{n-1 \geq N}}_d \right) \quad (49)$$

Dove:

$$e_0([x_1, \dots, x_N]) = \sum_{m=1}^{N-1} \omega_{x_m} \mathbf{I}_{x_{m+1}-x_m > \ell} + \omega_{x_N}$$

Come è possibile vedere dall'equazione sovrastante la dipendenza dei coefficienti è iterativa e ogni termine sul lato destro dell'equazione dipende da una configurazione precedente a quella considerata per il calcolo del coefficiente c_n :

- **Termine a:** contributo che deriva dall'ordine $n - 1$ a condizione che la prima particella si trovi in posizione $x_1 = 1$;
- **Termini b:** rappresenta il contributo derivato da ogni configurazione da cui può derivare la configurazione sotto analisi (C) muovendo una particella di un sito verso destra. In particolare il termine b_1 si riferisce alla particella più a sinistra possibile, mentre il termine b_2 a tutte le restanti particelle;
- **Termine c:** include nel calcolo tutte quelle configurazione che hanno una particella in più nel reticolo nell'ultimo sito L ;
- **Termine d:** elimina il contributo della configurazione C stessa ad un ordine inferiore;

Quindi per poter risolvere l'equazione (49) e calcolare il coefficiente c_n l'iter è semplice:

Assumendo di aver calcolato tutti i coefficienti c_{n-1} e partendo da una configurazione *stacked*, ovvero una configurazione in cui n particelle (minori o uguali al numero massimo di particelle che il reticolo può ospitare $N_{max} = \frac{L}{\ell} + 1$) sono nelle posizioni $x_i = 1 + (i - 1)\ell$. Calcolato il primo coefficiente $c_n([1, 1 + \ell, \dots, 1 + (n - 1)\ell])$ è possibile calcolare il secondo muovendo la n -esima particella una posizione più a destra, il terzo spostando di un'ulteriore posizione più a destra l' n -esima particella e così via fino ad arrivare al sito L e rimuovere la particella n -esima.

L'iter a questo punto viene ripetuto per la $(n-1)$ -esima particella per ogni posizione ammessa della particella n -esima. E così via fino ad arrivare alla configurazione vuota.

È dunque possibile definire: S_n come insieme delle configurazioni esplorate da questa procedura iterativa che corrisponde a $\{C | N(C) \leq n\}$.

4.3.2 TASEPy: python per velocizzare e semplificare i calcoli

Ciandrini et al [5] hanno sviluppato un pacchetto in codice python che permette di automatizzare questi calcoli, introducendo 4 diverse funzioni che replicano il processo iterativo descritto nel capito precedente e permette di risolvere le formule (46, 47, 48):

- **psa_compute:** utilizzando la matrice delle transizioni di salto, l'ordine massimo dell'espansione in serie (K) e la lunghezza della particella (ℓ), permette di calcolare i coefficienti J_n e $\rho_{i,n}$. Per migliorare le performance di calcolo non vengono salvati di default tutti i coefficienti c_n calcolati, ma se necessari possono essere estrapolati semplicemente settando il parametro `save_coeff` a "True"; i coefficienti verranno salvati in un file .csv.

La funzione `psa_compute` genera la configurazione stacked tenendo semplicemente in considerazione l'ordine dell'espansione in serie e confrontandola con il valore massimo di particelle che possono essere contenute nel filamento definito. Il valore massimo di particelle (N_{max}) viene calcolato confrontando il parametro ℓ che rappresenta la lunghezza del ribosoma e il numero di siti presenti sul reticolo (L). Calcolato N_{max} se l'ordine dell'approssimazione in serie ($npsa$) è minore di N_{max} allora, come descritto nelle equazioni (44, 45), si avranno $npsa$ particelle che occuperanno appunto i primi $npsa$ siti. Se, invece, $npsa > N_{max}$ allora si avranno N_{max} particelle che riempiranno tutti i siti del reticolo.

Definita la configurazione stacked la funzionalità `next_config()` di python permette di scorrere attraverso tutte le S_n configurazioni possibili. Calcolato il coefficiente c_n per la configurazione stacked viene salvato e verrà utilizzato per il calcolo dei coefficienti della densità per ogni sito i e per la corrente. L'iter viene ripetuto per ogni ordine $n = 0, \dots, K$.

Come spiegato in precedenza analizzando l'equazione 49 poiché ogni coefficiente necessita solo gli ordini n e $n-1$ per il suo calcolo, il sistema tiene in memoria solo i coefficienti strettamente necessari.

- **local_density:** grazie ai coefficienti $\rho_{i,n}$ calcolati tramite la funzione `psa_compute` e grazie ad un rate α definito dall'utente la funzione `local_density` permette di calcolare la densità locale per ogni sito per ogni ordine dell'espansione in serie ($n = 0, \dots, K$).

$$\rho_i^{(n)} = \sum_{k=0}^n \rho_{i,k} \alpha^k$$

- **mean_density:** permette il calcolo della densità media ad ogni ordine di espansione in serie grazie all'equazione sottostante.

$$\rho^{(n)} = \frac{1}{L} \sum_{i=1}^L \rho_i^{(n)}$$

- **current:** come per la densità locale, la funzione `current` permette di calcolare la corrente partendo dai coefficienti J_n e il rate α .

$$J^{(n)} = \sum_{k=0}^n J_k \alpha^{k+1}$$

Esempio di uso del codice:

Durante lo sviluppo del pacchetto Ciandrini et al (2023) [4] hanno confrontato i calcoli ottenuti tramite il codice TASEPy e la risoluzione diretta della master equation (36).

Un primo confronto è stato eseguito per un reticolo con solo 4 siti e per diverse lunghezze del ribosoma ($\ell = 1, 2, 3$). I rate di salto sono stati scelti in modo randomico ([1.88, 1.52, 1.09, 1.38]) e l'ordine di espansione è stato troncato a 5. I risultati ottenuti nel calcolo dei coefficienti ρ_n e J_n sono stati gli stessi sia tramite la risoluzione diretta della master equation sia tramite l'uso del pacchetto TASEPy.

Un secondo controllo è stato eseguito confrontando i profili di densità ottenuti tramite la simulazione stocastica del modello TASEP attraverso l'uso dell'algoritmo di Gillespie e la risoluzione del modello con l'utilizzo del pacchetto TASEPy. I parametri utilizzati sono stati $L = 50$, $\ell = 5$, $K = 4$ e $\alpha = 0.2$. Come è possibile vedere nella figura 7 i profili di densità tra le due metodologie sono statisticamente simili.

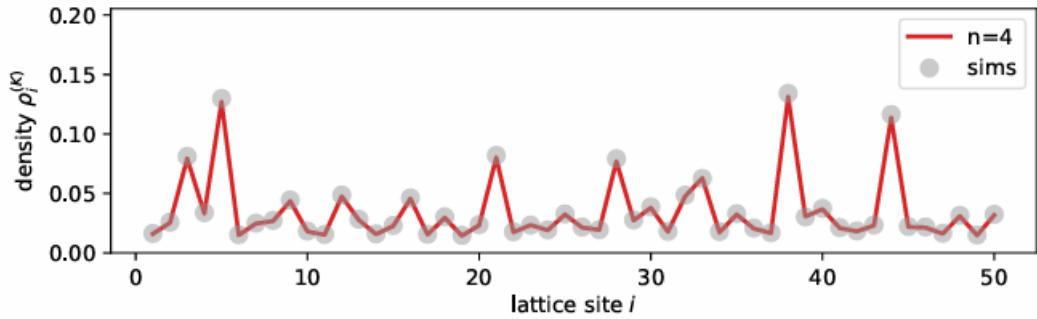


Figura 7: Confronto dei risultati ottenuti tramite le simulazioni stocastiche attraverso l'algoritmo di Gillespie (marcatori in grigio) e il calcolo dello spettro di densità media tramite l'uso del pacchetto TASEPy al quarto ordine dell'espansione in serie [4].

Inoltre, definito un errore percentuale $100 \times \frac{|J_{TASEPy}^n - J_{sims}|}{J_{sims}}$ dove $J_{sims} = 0.14262$ corrisponde alla corrente ottenuta tramite le simulazioni con l'algoritmo di Gillespie e n all'ordine di espansione in serie di potenze. I risultati ottenuti mostrano che al quarto ordine di espansione l'errore percentuale è del 1%.

Di seguito i risultati ottenuti settando i parametri a $L = 100$, $\ell = 1$, $\alpha = 0.2$ e $K = 4$ i risultati ottenuti sono verosimili e statisticamente accettabili rispetto ai dati sperimentali [4] come quelli precedentemente visti.

	Ordine 0	Ordine 1	Ordine 2	Ordine 3	Ordine 4
ρ	0.0	0.0571	0.0615	0.0624	0.0626
J	0.2000	0.1950	0.1948	0.1947	0.1947

Tabella 1: Dati estrapolati utilizzando il codice TASEPy con l'approssimazione PSA. Dove $L=100$, $\ell = 1$, $\alpha = 0.2$ e la lista dei rate è scelta in modo randomico tra 1 e 10

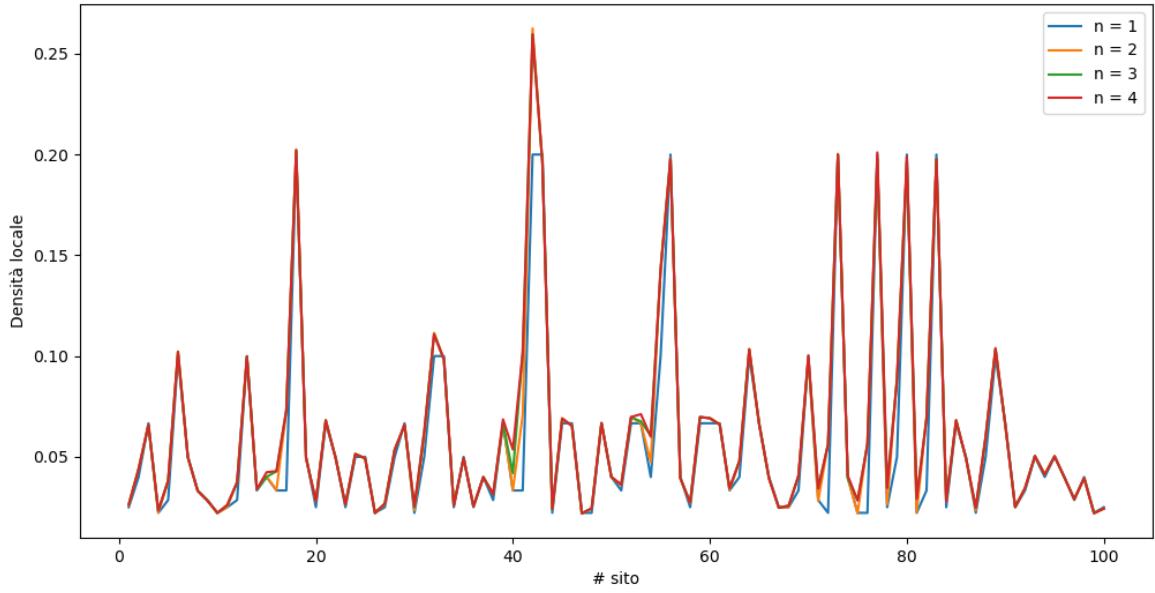


Figura 8: Spettro della densità locale in funzione dei siti e del grado di espansione in serie

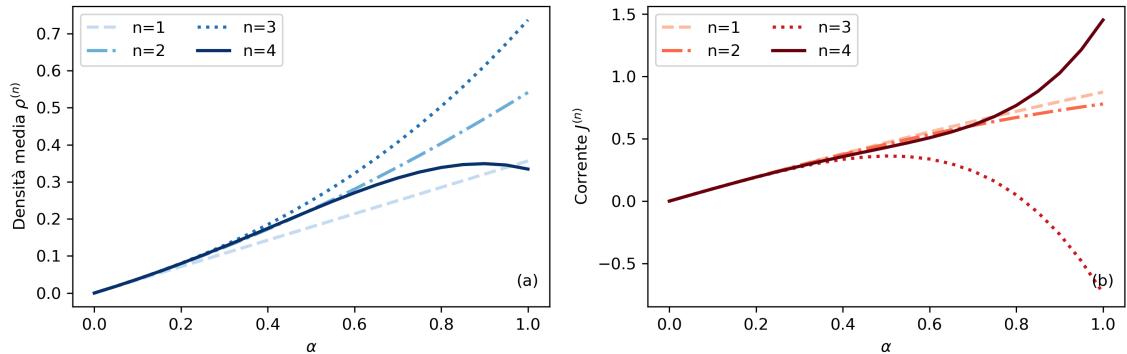


Figura 9: Valori medi calcolati tramite il pacchetto TASEPy della corrente e della densità per diversi valori di α e n . Dove $L = 100$, $\ell = 1$, $\alpha = 0.2$ e la lista dei rate è scelta in modo randomico tra 1 e 10

Nella figura 9 è possibile notare come al crescere di α l'approssimazione in serie di potenze non è più una buona approssimazione in quanto gli ordini più grandi della serie dominano sui primi facendo divergere i valori di ρ e J .

5 Conclusioni

5.1 TASEP omogeneo: conclusioni sul modello a due fasi

Il modello a due stati introdotto nel 2010 [4] ha dimostrato la capacità di fare previsioni molte accurate. Ha messo in evidenza che i punti critici in cui avvengono le transizioni di fase sono diversi da quelli predetti dai modelli precedenti, dimostrando che i punti critici e le variazioni di fase dipendono dal rapporto tra k e γ .

Il rapporto tra k e γ rappresenta la differenza di rapidità tra il passaggio da primo a secondo stato, ovvero tra accoppiamento tRNA-codone, e processo di traslocazione da un sito ad un altro. Nel terzo capitolo è stato approfondito come il processo di accoppiamento, definito dal rate k , è il passaggio limitante del processo biologico. Osservando il diagramma di fase rappresentato in figura 10 è evidente come la fase di massima corrente diminuisca all'aumentare del rapporto fra k e γ . A valori di k maggiori la MC phase si riduce fino alla condizione per cui $\frac{k}{\gamma} \rightarrow \infty$ per cui si ottengono i valori descritti dai modelli originali [1], [2]. La condizione limite coincide col rappresentare il passaggio da primo stato a secondo istantaneo. Essendo il rate di accoppiamento il passaggio limitante del processo biologico una trattazione più dettagliata della traduzione del mRNA richiede chiaramente la considerazione esplicita del processo di accoppiamento tra tRNA e codone.

Al diminuire del rapporto $\frac{k}{\gamma}$ viene messo in evidenza il contributo del processo di traslocazione che aumenta drasticamente la dimensione della fase di massima corrente come si può vedere in figura 10. Dai dati sperimentali emerge che il valore reale di γ è dell'ordine di 35 s^{-1} (costante per ogni codone) e il rapporto $\frac{k}{\gamma}$ si trova nel range $[0.05-3.38]$ in dipendenza del codone scelto. Tuttavia, per la grande maggioranza dei codoni (87%) si ha, dai dati sperimentali, che $\frac{k}{\gamma} < 1$.

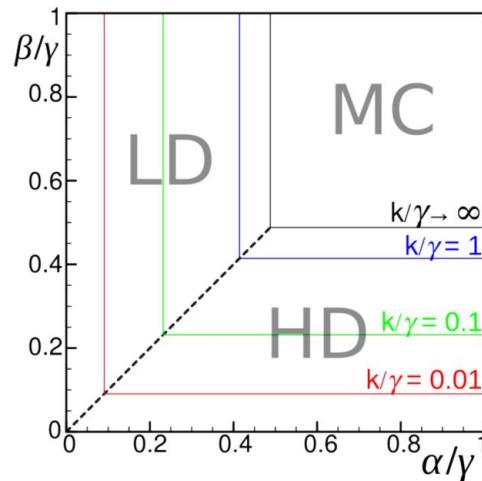


Figura 10: Diagramma di fase dal lavoro di Ciandrini et al. (2010) [3]

Il risultato principale del modello a due stati è dimostrare che il processo di traduzione giace quasi costantemente in una fase di corrente massima, in cui la sintesi delle proteine è ottimizzata. Questa affermazione è sostenuta dal fatto che la fase di massima corrente è la zona preponderante nel diagramma di fase definito tramite i dati ottenuti sperimentalmente e, inoltre, il profilo della densità estrapolato da alcuni esperimenti si può approssimare con quello legato alla MC phase [8].

A sostegno di ciò che il diagramma di fase permette di evincere si è dimostrato che un ribosoma appena uscito dall'ultimo sito può ripartire da capo come in un processo ricorrente, ottimizzando ulteriormente il rate di inizializzazione α ed avvicinando così questo parametro all'ordine del valore critico α_{crit} che permette di giacere nella fase di massima corrente [9].

5.2 TASEPy e risoluzione del TASEP disomogeneo: applicazioni e risultati

Il modello di calcolo descritto dal pacchetto TASEPy permette il calcolo dei coefficienti legati al PSA per il TASEP inomogeneo in modo rapido ed efficace. Inoltre TASEPy può essere impiegato per calcoli inversi dei sistemi di traduzione del mRNA.

Un primo esempio è il calcolo del rate iniziale alpha a partire da una data densità sperimentale; questo approccio è stato usato da Ciandrini e colleghi in un ulteriore articolo [10] in cui hanno dimostrato di poter risalire al valore di α per la sintesi proteica del *S. cerevisiae* (il lievito di birra) partendo dal solo valore della densità media. Prima dell'introduzione del pacchetto in codice python quello che il team di fisici avrebbe dovuto fare era calcolare tutti i coefficienti e tutte le densità ipotizzando dei valori per α da aggiustare ed ottimizzare dopo il calcolo, invece con l'uso del pacchetto TASEPy a partire dal calcolo dei coefficienti con `psa_compute` è possibile automatizzare il processo utilizzando i tool di ottimizzazione che esplorano i possibili valori del parametro restituendo quello che minimizza la funzione $f(\alpha) = \rho_\alpha - \rho_{exp}$.

Un secondo impiego di TASEPy permette di calcolare il rapporto tra α e i vari ω_i a partire dalla densità locale. Definito $k_i = \frac{\omega_i}{\alpha}$, si possono calcolare i coefficienti con `psa_compute` e minimizzando la deviazione quadratica media

$$RM(k_1, \dots, k_L) = \sqrt{\frac{\sum_{i=1}^L [\rho_i(k_1, \dots, k_L) - \rho_i^{(exp)}]^2}{L}}$$

si possono ricavare i rapporti dei rate di salto ricercati. I valori calcolati con TASEPy risultano molto accurati come è possibile vedere nella figura sottostante.

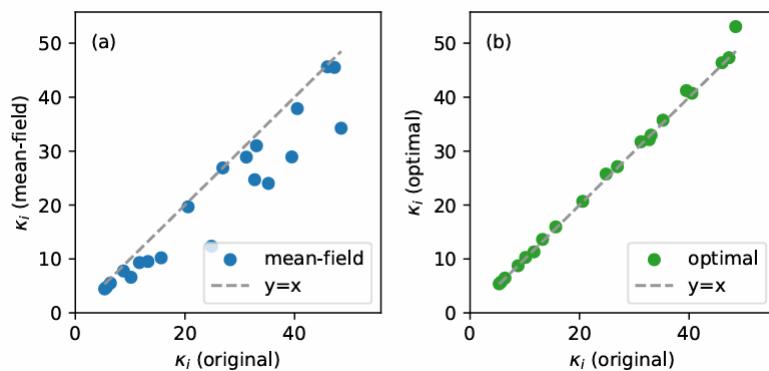


Figura 11: Sono stati confrontati i valori reali con i valori calcolati con due metodologie diverse dei rate di salto. Ogni punto corrisponde ad un sito specifico. Nella figura (a) è stata utilizzata l'approssimazione di campo medio per il calcolo dei rapporti k e come è possibile vedere i valori si distanziano maggiormente dalla bisettrice $y=x$ e quindi dai valori sperimentali rispetto alla figura (b) in cui è stata utilizzata l'approssimazione in serie di potenze e il calcolo è stato eseguito tramite TASEPy. [4]

L'utilizzo di TASEPy rimane comunque vincolato all'approssimazione in serie di potenze. Infatti nonostante sia conosciuto e studiato sin dagli anni 80/90 il modello inomogeneo non ha ancora una risoluzione corretta, ma dipende appunto dall'approssimare PSA rispetto ad rate limitante, nel nostro caso α . Per un rate di inizializzazione grande l'approssimazione perde le sue fondamenta restituendo risultati non coerenti, queste limitazioni corrispondono a delle condizioni che J e ρ devono soddisfare per poter applicare con successo la PSA:

- | | |
|---|---|
| 1. $0 \leq \rho_i \leq 1$
2. $\frac{d\rho_i}{d\alpha_{11}} \geq 0$ | 3. $0 \leq J \leq \alpha$
4. $0 \leq \frac{dJ}{d\alpha_{11}} \leq 1$ |
|---|---|

TASEPy rappresenta dunque un progresso verso la risoluzione del modello inomogeneo e dunque verso una migliore comprensione del processo biologico reale.

Riferimenti bibliografici

- [1] Carolyn T. MacDonald, Julian H. Gibbs, and Allen C. Pipkin. Kinetics of biopolymerization on nucleic acid templates. *Biopolymers*, Jan 1968.
- [2] Carolyn T. MacDonald and Julian H. Gibbs. Concerning the kinetics of polypeptide synthesis on polyribosomes. *Biopolymers*, Feb 1969.
- [3] L. Ciandrini, I. Stansfield, and M. C. Romano. Role of the particle's stepping cycle in an asymmetric exclusion process: A model of mRNA translation. *Phys. Rev. E*, May 2010.
- [4] Luca Ciandrini, Richmond L. Crisostomo, and Juraj Szavits-Nossan. Tasepy: A python-based package to iteratively solve the inhomogeneous exclusion process. *SciPost Phys. Codebases*, Dec 2023.
- [5] Luca Ciandrini, Richmond L. Crisostomo, and Juraj Szavits-Nossan. Codebase release 1.1 for tasepy. *SciPost Phys. Codebases*, Dec 2023.
- [6] S. Clancy and W. Brown. Translation: Dna to mRNA to protein. *Nature Educating*, page 10, Aug 2008.
- [7] Roberto Livi and Paolo Politi. *Nonequilibrium statistical physics: a modern perspective*. Cambridge University Press, Jan 2017.
- [8] Nicholas T. Ingolia, Sina Ghaemmaghami, John R. S. Newman, and Jonathan S. Weissman. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *science*, Apr 2009.
- [9] Tom Chou. Ribosome recycling, diffusion, and mRNA loop formation in translational regulation. *Biophysical Journal*, Feb 2003.
- [10] Luca Ciandrini, Ian Stansfield, and M. Carmen Romano. Ribosome traffic on mRNAs maps to gene ontology: Genome-wide quantification of translation initiation rates and polysome size regulation. *PLOS Computational Biology*, Jan 2013.
- [11] R. A. Blythe and M. R. Evans. Nonequilibrium steady states of matrix-product form: a solver's guide. *Journal of Physics A: Mathematical and Theoretical*, oct 2007.
- [12] Aakash Basu and Debashish Chowdhury. Traffic of interacting ribosomes: Effects of single-machine mechanochemistry on protein synthesis. *Phys. Rev. E*, Jun 2007.
- [13] Hermioni Zouridis and Vassily Hatzimanikatis. A model for protein translation: polysome self-organization leads to maximum protein synthesis rates. *Biophysical journal*, June 2007.
- [14] Ashok Garai, Debanjan Chowdhury, Debashish Chowdhury, and TV Ramakrishnan. Stochastic kinetics of ribosomes: single motor properties and collective behavior. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, July 2009.
- [15] MACKAY VL. Gene expression analyzed by high-resolution state array analysis and quantitative proteomics. *Mol Cell Proteomics*, Feb 2004.

Ringraziamenti

Non sono mai stato il migliore degli studenti, quello coi voti più alti o quello che rispettava i tempi di laurea al minuto, ma sono contento della mia crescita in questi anni, sono contento del mio percorso.

Iniziai questo corso di laurea con la consapevolezza inconscia che non sarei mai riuscito a finirlo, mi dicevo "troppo difficile", "non sei all'altezza, fisica è impossibile" per fortuna non sono mai stato solo. Per questo motivo vorrei ringraziare tutte le persone che mi hanno aiutato a crescere e migliorare.

La prima persona che vorrei ringraziare è la mia professoressa di fisica delle superiori, senza di lei non mi sarei mai appassionato alla fisica, non avrei mai preso in considerazione questo percorso. Mi ha passato la curiosità e la voglia di scoprire il mondo attraverso questa incredibile materia.

Vorrei anche ringraziare il professor Ginelli che mi ha seguito e aiutato durante la stesura di questo elaborato adattando spesso il suo calendario ai miei problemi dovuti all'orario del lavoro. Lo ringrazio perché mi ha mostrato come anche in un ambiente competitivo ed elevato, come quello accademico, ci siano persone comprensive e pronte ad aiutarti.

Il mio percorso in questi anni è passato da tanti alti e bassi, per tanti anni ho fatto fatica a capirmi e essere in pace con me stesso. Per questo motivo vorrei ringraziare Chiara, la persona che mi ha mostrato una via per accettarmi, dandomi i consigli di cui avevo bisogno per apprezzarmi, capirmi e perdonarmi.

Negli anni ho incontrato tantissime persone che mi hanno accompagnato nel mio percorso. Tante persone, di ogni età, che mi hanno mostrato nuovi punti di vista di me e del mondo. Giulio, Tommaso, Luca, Giulia, Matilde, Andrea, Marco, William, Matteo e tutte le altre persone che ora mi sfuggono. Vi ringrazio, ognuno di voi, anche se con alcuni ho perso i contatti, vi ringrazio per aver contribuito a ciò che sono oggi.

Ringrazio Ilaria per il tempo passato insieme, per mostrarmi ogni giorno un mondo diverso da quello che vedo io, per sopportarmi ogni giorno e starmi accanto con la dolcezza che solo una persona come lei può avere. Che questo sia il punto di inizio per tutti i nostri progetti insieme.

In ultimo, ma non per importanza, vorrei ringraziare la mia famiglia. Non sono il figlio più facile da avere, sono un po' strano, ho tante cose per la testa e spesso sono un po' testardo. Vi ringrazio per starmi accanto, accettarmi e sostenermi in qualsiasi cosa io faccia o scelga. Vi ringrazio per darmi un posto sicuro in cui stare tutte le volte che ne ho bisogno. Grazie. Davvero.