



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Χημικών Μηχανικών
Τομέας II: Ανάλυσης, Σχεδιασμού &
Ανάπτυξης Διεργασιών και Συστημάτων

**Στοχαστικές προσομοιώσεις ετερογενών πληθυσμών *E. coli* με
ρυθμιστικό δίκτυο τύπου toggle switch**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Μέμμου Νικόλαου

Επιβλέπων Καθηγητής

Ανδρέας Γ. Μπουντουβής

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον Καθηγητή Ανδρέα Μπουντουβή για την ευκαιρία που μου έδωσε να εκπονήσω τη διπλωματική μου εργασία σε συνεργασία με τον ίδιο και την ερευνητική του ομάδα. Οι συμβουλές του ήταν καθοριστικές, τόσο σε επίπεδο διπλωματικής εργασίας, όσο και για τη μετέπειτα πορεία μου ως Χημικός Μηχανικός.

Επίσης, θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στους Δρ. Ιωάννη Αβιζιώτη και Μιχάλη Καβουσανάκη για την πολύτιμη βοήθεια τους. Οι συζητήσεις που είχαμε βοήθησαν σημαντικά στην εκπόνηση της διπλωματικής εργασίας.

Θα ήθελα να ευχαριστήσω ακόμη τους φίλους μου και ειδικότερα τον Νίκο Κάτση, τον Κωνσταντίνο Ρέβελα, τον Θοδωρή Μποτσάκο και τον Κωνσταντίνο Πανόπουλο για τα φοιτητικά χρόνια που περάσαμε, καθώς επίσης και την οικογένειά μου για την υποστήριξη που μου έδειξε.

Περίληψη

Τα ρυθμιστικά δίκτυα αποτελούν σειρές αντιδράσεων οι οποίες λαμβάνουν χώρα εντός του κυττάρου και επηρεάζουν το φαινότυπό τους, μέσω του ρυθμού έκφρασης των γονιδίων του δικτύου. Εξετάζονται πληθυσμοί κυττάρων *E. coli* τα οποία φέρουν το ρυθμιστικό δίκτυο toggle switch. Το δίκτυο αυτό περιλαμβάνει δύο παρεμποδιστές (repressors) με ανταγωνιστική δράση, παρεμποδίζοντας ο ένας την έκφραση του άλλου. Η δυναμική συμπεριφορά των συστημάτων αυτών περιγράφεται από συνήθεις διαφορικές εξισώσεις, οι οποίες σε συνθήκες μόνιμης κατάστασης παρουσιάζουν πολλαπλότητα λύσεων. Συγκεκριμένα, για ένα εύρος εξωκυτταρικών παραμέτρων, η μέση έκφραση του κυτταρικού πληθυσμού μπορεί να είναι είτε υψηλή είτε χαμηλή.

Η συνήθης πρακτική για τη μελέτη της δυναμικής των ρυθμιστικών δικτύων είναι η θεώρηση ομοιογενών κυτταρικών πληθυσμών ως προς το κυτταρικό τους περιεχόμενο. Μια τέτοια θεώρηση αν και οδηγεί σε απλά μαθηματικά μοντέλα (συνήθεις διαφορικές εξισώσεις) αγνοεί ένα βασικό χαρακτηριστικό των κυτταρικών πληθυσμών που είναι η ετερογένεια, τόσο ως προς το σχήμα και το μέγεθος των κυττάρων όσο και προς τον αριθμό των ρυθμιστικών μορίων που αυτά περιέχουν. Η μαθηματική προτυποποίηση της κυτταρικής ετερογένειας μπορεί να πραγματοποιηθεί με τη χρήση Ισοζυγίων Κυτταρικών Πληθυσμών (ΙΚΠ), τα οποία περιγράφονται από μερικές ολοκληρωτικές διαφορικές εξισώσεις (integro-differential equations).

Ωστόσο, η περιγραφή ετερογενών κυτταρικών πληθυσμών με ντετερμινιστικά πρότυπα δε λαμβάνει υπόψη την εγγενή στοχαστικότητα που παρατηρείται σε αυτά τα συστήματα. Για το λόγο αυτό γίνεται χρήση στοχαστικών αλγορίθμων τύπου kinetic Monte Carlo. Παρουσιάζεται η χρονική εξέλιξη της μέσης έκφρασης του ενδοκυτταρικού περιεχομένου πληθυσμών σε διαφορετικές εξωκυτταρικές συνθήκες και για διαφορετικά μεγέθη πληθυσμών. Επίσης, παρουσιάζονται στιγμιότυπα της κατανομής των πληθυσμών αυτών ως προς το ενδοκυτταρικό τους περιεχόμενο.

Τα αποτελέσματα της χρονικής εξέλιξης που προέκυψαν για διαφορετικούς κυτταρικούς πληθυσμούς, συγκρίνονται με τα αντίστοιχα του ντετερμινιστικού. Παρατηρείται συμφωνία του στοχαστικού με το ντετερμινιστικό μοντέλο, καθώς αυξάνεται το πλήθος των κυττάρων. Όσον αφορά τη χρονική εξέλιξη των κατανομών, παρατηρείται ότι από μια αρχική κατανομή Gauss, ο πληθυσμός μεταπίπτει σε πιο απότομες κατανομές μέχρι να σταθεροποιηθεί. Τέλος, από τα αποτελέσματα προέκυψε ότι ο παράγοντας ασυμμετρίας έχει σημαντική επίδραση στην κατανομή.

Stochastic simulations in heterogeneous populations of E.coli with the toggle switch model

Diploma Thesis

by

Nikolaos Memmos

Abstract

Gene-regulatory networks are synthesized by a combination of intracellular reactions occurring in cellular systems. The matter of investigation is populations consisting of E.coli cells, which carry the toggle switch genetic network. This network is composed of two repressors presenting mutual inhibition. Such systems promote multiplicity of the average gene expression within a range of extracellular parameters.

The most common method for studying the dynamic behavior of genetic networks is the assumption of homogeneous cell populations with respect to the average context of cells. Despite the fact that this assumption drives on simple mathematical models (ordinary differential equations), neglects the heterogeneity in the shape, size and of the concentration of regulatory molecules within the cells. The deterministic Cell Population Balance (CPB) models which are described by partial integro-differential equations are able to capture the effect of heterogeneity on cell populations.

However, deterministic models disregard the discrete nature of cell populations and for these reasons stochastic modeling is adopted, such as kinetic Monte Carlo algorithms. In this thesis, time development of average concentration of intracellular context for different number of cells and different extracellular conditions is presented. Furthermore, snapshots of the distribution of the intracellular context are presented.

The results of time evolution of average concentration of intracellular context for different cell populations are compared with the results of the deterministic model. It is observed that as the number of cells increases, the stochastic and deterministic model give almost identical

results. Furthermore, the distributions evolve from an initial Gaussian distribution to more sharp distributions. Finally, the results show that the asymmetric factor has a significant impact on the distribution.

Περιεχόμενα

Ευχαριστίες.....	3
Περίληψη	4
Abstract.....	6
Κεφάλαιο 1 Εισαγωγή	11
1.1 Κυτταρική ετερογένεια.....	11
1.2 Είδη κυτταρικής ετερογένειας.....	12
1.3 Ρυθμιστικά δίκτυα	12
1.4 Οργάνωση της εργασίας	13
Κεφάλαιο 2 Ρυθμιστικό δίκτυο toggle switch.....	14
2.1 Δομή του toggle switch.....	14
2.2 Το φαινόμενο της διπλοευστάθειας	16
2.3 Μαθηματικό μοντέλο	19
2.4 Εξαγωγή μαθηματικού μοντέλου.....	21
Κεφάλαιο 3 Το στοχαστικό μοντέλο.....	23
3.1 Ισοζύγια Κυτταρικών Πληθυσμών.....	23
3.2 Εισαγωγή στα στοχαστικά μοντέλα.....	27
3.2 Ανάπτυξη μοντέλου	28
3.3 Constant Number Monte Carlo (CNMC).....	33
Κεφάλαιο 4 Αποτελέσματα.....	35
4.1 Εισαγωγή	35
4.2 Χρονική εξέλιξη μέσω συγκεντρώσεων	37
4.3 Κατανομές κυτταρικού περιεχομένου	40
Κεφάλαιο 5 Συμπεράσματα-Μελλοντική Έρευνα.....	43
5.1 Συμπεράσματα.....	43
5.2 Μελλοντική έρευνα	44
Παράρτημα	46
Βιβλιογραφία	47

Κεφάλαιο 1

Εισαγωγή

Στο κεφάλαιο αυτό γίνεται μία εισαγωγή στην κυτταρική ετερογένεια, καθώς και στα είδη αυτής που υπάρχουν. Επιπλέον, γίνεται μία σύντομη αναφορά στα ρυθμιστικά δίκτυα. Τέλος, περιγράφεται ο στόχος της παρούσας εργασίας καθώς και η δομή της διπλωματικής.

1.1 Κυτταρική ετερογένεια

Οι εξελίξεις των τελευταίων χρόνων στη μοριακή βιολογία, στα genomics, τα transcriptomics και τα proteomics, μας παρέχουν τα απαραίτητα εργαλεία για την μελέτη των διεργασιών που πραγματοποιούνται σε κυτταρικό επίπεδο [1]. Επίσης, τεχνικές από τον κλάδο της υπολογιστικής βιολογίας χρησιμοποιούνται προκειμένου να διερευνηθούν φαινόμενα που πραγματοποιούνται στον ενδοκυτταρικό χώρο. Παρόλα αυτά ο φαινότυπος¹ ενός κυτταρικού πληθυσμού είναι αποτέλεσμα διαφόρων και πολύπλοκων διεργασιών που πραγματοποιούνται τόσο εντός του κυττάρου, όσο και σε επίπεδο κυτταρικού πληθυσμού. Παράμετροι όπως το σχήμα, το μέγεθος, οι αντιδράσεις εντός των κυττάρων και η ανομοιόμορφη κατανομή του κυτταρικού περιεχομένου κατά τη διαίρεση οδηγούν σε διακυμάνσεις του φαινοτύπου από κύτταρο σε κύτταρο. Το φαινόμενο αυτό απαντάται σε διάφορα βιολογικά συστήματα, όπως στα μεγέθη των ιών phage [2], σε καλλιέργειες αποτελούμενες από σπόρια του *B. subtilis* [3] και στους κυτταρικούς πληθυσμούς του *E. coli* [4].

Το πλήθος των βιοτεχνολογικών εφαρμογών που υπάρχουν σήμερα, έχουν ως στόχο τη μεγιστοποίηση της παραγωγής ουσιών από κυτταρικούς πληθυσμούς. Επιπλέον, οι περισσότερες ιατρικές εφαρμογές και προσεγγίσεις που εξετάζουν παθολογικές συνθήκες αντιμετωπίζουν τον κυτταρικό πληθυσμό ως σύνολο και όχι κάθε κύτταρο μεμονωμένα. Η κυτταρική ετερογένεια αποτελεί έναν σημαντικό παράγοντα ο οποίος μπορεί να παρεμποδίσει την αποδοτική παραγωγή

¹ Ως φαινότυπος ορίζεται το σύνολο των παρατηρούμενων χαρακτηριστικών ενός οργανισμού μια δεδομένη χρονική στιγμή. Δεν πρέπει να συγχέεται με τον γονότυπο, ο οποίος αποτελεί το σύνολο των γονιδίων που κληρονομούνται από έναν οργανισμό. Ο φαινότυπος επηρεάζεται τόσο από τον γονότυπο, όσο και από το περιβάλλον.

βιοτεχνολογικών προϊόντων [5, 6], αλλά και την προσαρμογή τους σε απότομες περιβαλλοντικές αλλαγές [7, 8, 9]. Σημειώνεται ότι ένα βασικό εμπόδιο για τη θεραπεία του καρκίνου αποτελεί η ετερογένεια των νεοπλασματικών αλλά και μεταστατικών κυττάρων [10, 11, 12, 13]. Συνεπώς, η σε βάθος κατανόηση της δυναμικής σχέσης μεταξύ των διεργασιών που λαμβάνουν χώρα σε επίπεδο ενός κυττάρου και σε επίπεδο ολόκληρου κυτταρικού πληθυσμού μπορεί να παρέχει σημαντικές πληροφορίες για την ανίχνευση, την πρόγνωση αλλά και τη θεραπεία του καρκίνου.

1.2 Είδη κυτταρικής ετερογένειας

Ο φαινότυπος κάθε κυττάρου εξαρτάται από το είδος και τον αριθμό των γονιδίων που μεταφράζονται (transcription) κάθε χρονική στιγμή, αλλά και από τις μεταβολικές-ενδοκυτταρικές αντιδράσεις. Οι εξωκυτταρικές περιβαλλοντικές συνθήκες επηρεάζουν άμεσα τις ενδοκυτταρικές διεργασίες, ιδίως σε ισογονιδιακούς πληθυσμούς, δηλαδή σε πληθυσμούς που τα κύτταρα φέρουν το ίδιο γενετικό δίκτυο. Σε αυτού του είδους τους πληθυσμούς έχει παρατηρηθεί ότι πολύ μικρές διακυμάνσεις των ποσοτήτων των ρυθμιστικών μορίων εντός του κυττάρου (regulatory molecules) [14] μπορούν να επηρεάσουν σημαντικά το φαινότυπό τους, συνεπώς είναι πιθανό να υπάρχουν ταυτόχρονα κύτταρα με περίπου τον ίδιο αριθμό ρυθμιστικών μορίων που να συμπεριφέρονται τελείως διαφορετικά. Αυτού του είδους η ετερογένεια ονομάζεται ενδογενής (intrinsic).

Επίσης, τα μητρικά κύτταρα, κατά τη διαίρεσή τους, κληροδοτούν στα θυγατρικά κύτταρα το κυτταρικό τους περιεχόμενο κατά ανισομερή τρόπο (εκτός του DNA). Οι διαφορές στις ποσότητες συγκεκριμένων ουσιών (π.χ RNA, πρωτεΐνες) και κυρίως στον αριθμό των ρυθμιστικών μορίων που κληροδοτούνται στα θυγατρικά κύτταρα έχει ως αποτέλεσμα την εκδήλωση διαφορετικών φαινοτύπων. Αυτού του τύπου η ετερογένεια καλείται εξωγενής (extrinsic) και είναι αυτή που θα μελετηθεί στην παρούσα εργασία.

1.3 Ρυθμιστικά δίκτυα

Από πειραματικές διαδικασίες έχει προκύψει ότι η ετερογένεια στους κυτταρικούς πληθυσμούς συνδέεται άρρηκτα με τις λειτουργικές ιδιότητες των ρυθμιστικών γονιδιακών δικτύων (gene regulatory networks) [15]. Ως ρυθμιστικά δίκτυα ορίζονται το σύνολο των τμημάτων του DNA σε ένα κύτταρο που αλληλεπιδρούν μεταξύ τους μέσω της μεταγραφής του DNA σε RNA και στη συνέχεια της μετάφρασης του τελευταίου σε mRNA, καθώς και με άλλες ουσίες στο κύτταρο. Ο φαινότυπος κάθε κυττάρου εξαρτάται από το είδος και τον αριθμό των

γονιδίων που εκφράζονται κάθε χρονική στιγμή, καθώς και από τις επακόλουθες ενδοκυτταρικές αντιδράσεις. Στην παρούσα εργασία μελετάται το ρυθμιστικό δίκτυο τύπου toggle switch, το οποίο περιγράφεται αναλυτικά στο κεφάλαιο 2.

1.4 Οργάνωση της εργασίας

Η εργασία αυτή είναι οργανωμένη σε έξι κεφάλαια. Στο παρόν κεφάλαιο γίνεται εισαγωγή στις έννοιες της κυτταρικής ετερογένειας και των ρυθμιστικών δικτύων. Στο Κεφάλαιο 2 γίνεται περιγραφή του γενετικού δικτύου toggle switch που εξετάζεται στην παρούσα διπλωματική εργασία. Στο Κεφάλαιο 3 περιγράφονται το ντετερμινιστικό και το στοχαστικό μοντέλο αντίστοιχα. Στο Κεφάλαιο 4 παρατίθενται τα αποτελέσματα του στοχαστικού μοντέλου, σε σύγκριση με αυτά τα του ντετερμινιστικού ενώ το Κεφάλαιο 5 περιέχει τα συμπεράσματα που εξήχθησαν από τα αποτελέσματα, καθώς και προτάσεις για μελλοντική έρευνα.

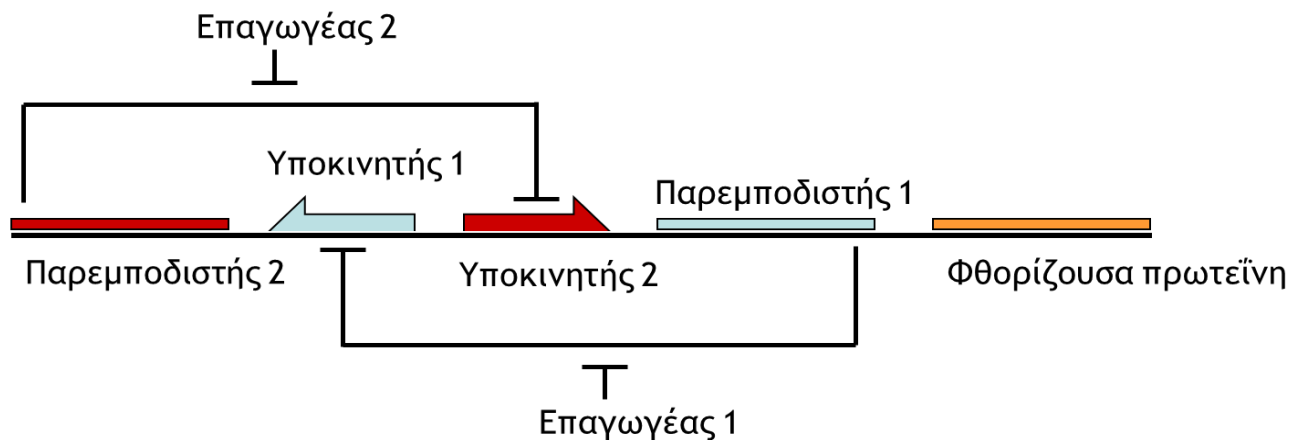
Κεφάλαιο 2

Ρυθμιστικό δίκτυο toggle switch

Στο κεφάλαιο αυτό παρουσιάζεται το ρυθμιστικό δίκτυο toggle switch, το οποίο μελετάται στην παρούσα εργασία. Αναλύεται η λειτουργία του και το μαθηματικό μοντέλο που το περιγράφει.

2.1 Δομή του toggle switch

Ένα από τα πρώτα ρυθμιστικά δίκτυα τα οποία κατασκευάστηκαν σε εργαστηριακό επίπεδο είναι το genetic toggle switch. Το toggle switch αποτελείται από δύο παρεμποδιστές (repressors) και δύο υποκινητές (promoters) [16] (βλ. Σχήμα 2.1).

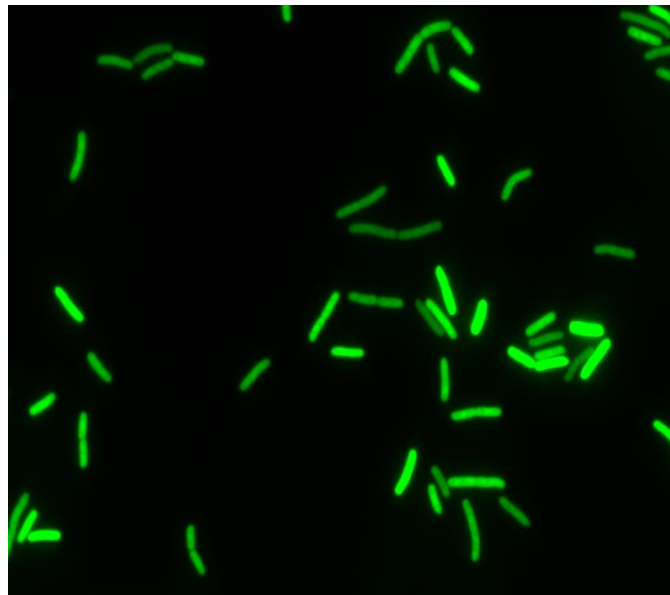


Σχήμα 2.1: Σχηματική αναπαράσταση του toggle switch.

Ο εκάστοτε υποκινητής είναι υπεύθυνος για τη μεταγραφή του αντίστοιχου παρεμποδιστή, δηλαδή ο υποκινητής 1 είναι το σημείο όπου η RNA πολυμεράση προσκολλάται για να ξεκινήσει η μεταγραφή του γονιδίου του παρεμποδιστή 2. Αντίστοιχα, ο υποκινητής 2 είναι υπεύθυνος για τη μεταγραφή του παρεμποδιστή 1 και μιας πράσινης φθορίζουσας πρωτεΐνης που σηματοδοτεί την έκφραση του γονιδίου αυτού σε εργαστηριακό επίπεδο (βλ. Σχήμα 2.2).

Πειραματικά, έχουν κατασκευαστεί δύο κατηγορίες γενετικών δικτύων toggle switch, η κατηγορία pTAK και η κατηγορία pIKE. Το κοινό των δύο αυτών κατηγοριών είναι ότι χρησιμοποιούν την λακτόζη ως παρεμποδιστή (IacI) σε συνδυασμό με τον υποκινητή P_{trc-2}. Το δεύτερο ζευγάρι παρεμποδιστή-υποκινητή είναι αυτό που διαφέρει. Συγκεκριμένα, στην κατηγορία pTAK χρησιμοποιείται ο υποκινητής P_{Ls1con} σε συνδυασμό με τον ευαίσθητο σε θερμοκρασιακές μεταβολές παρεμποδιστή clts, ενώ στην κατηγορία pIKE χρησιμοποιείται ο P_{LtetO-1} υποκινητής σε συνδυασμό με τον παρεμποδιστή Tet (tetR).

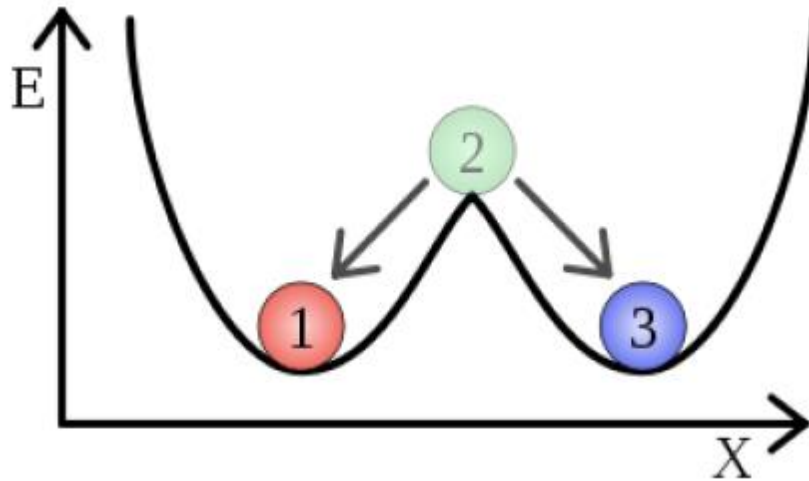
Το βασικότερο χαρακτηριστικό του συστήματος αυτού είναι ότι για διάφορες εξωκυτταρικές συνθήκες, όπως είναι η συγκέντρωση του κατάλληλου επαγωγέα, μπορεί να υπάρξει διαφορετικός φαινότυπος από κύτταρο σε κύτταρο. Η διαφορά αυτή στο φαινότυπο γίνεται αντιληπτή εργαστηριακά μέσω της φθορίζουσας πρωτεΐνης που αναφέρθηκε προηγουμένως. Το φαινόμενο αυτό είναι γνωστό ως διπλοευστάθεια λύσεων, η οποία χαρακτηρίζει τόσο κάθε κύτταρο ξεχωριστά, όσο και τη συμπεριφορά του κυτταρικού πληθυσμού σε μακροσκοπικό επίπεδο, και απαντάται σε πληθώρα βιολογικών συστημάτων.



Σχήμα 2.2: Κύτταρα E. coli που φέρουν το γενετικό δίκτυο toggle switch [17].

2.2 Το φαινόμενο της διπλοευστάθειας

Ως διπλοευστάθεια ορίζεται η ύπαρξη δύο ευσταθών σημείων ισορροπίας για ένα δυναμικό σύστημα (βλ. Σχήμα 2.3).



Σχήμα 2.3: Σχηματική αναπαράσταση διπλοευστάθειας. Τα σημεία 1 και 3 είναι ευσταθή, ενώ το 2 ασταθές.

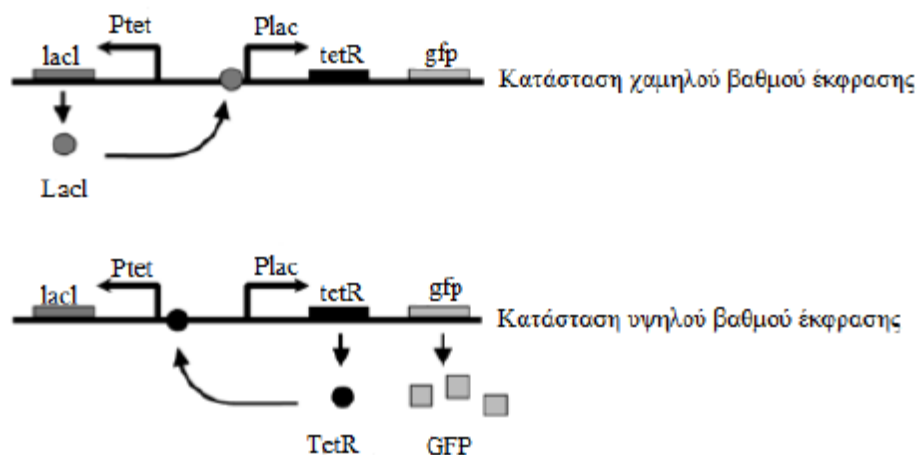
Το φαινόμενο της διπλοευστάθειας είναι ευρέως παρατηρούμενο σε βιολογικά συστήματα και η κατανόησή του είναι απαραίτητη για τα φαινόμενα που πραγματοποιούνται κατά τη φάση της κυτταρικής λειτουργίας, όπως είναι οι διαδικασίες κατά τη διάρκεια του κυτταρικού κύκλου, η κυτταρική διαφοροποίηση και η απόπτωση [18]. Επιπλέον, έχει παρατηρηθεί στην απώλεια της κυτταρικής ομοιόστασης [19], η οποία συνδέεται άρρηκτα με τον καρκίνο και τις ασθένειες της πρωτεΐνης prion [20]. Επίσης, φαινόμενα διπλοευστάθειας έχουν αναφερθεί ότι εμφανίζονται στην έκφραση του φαινότυπου σε πληθυσμούς κλώνων που είναι σημαντικοί για τη γέννηση νέων ειδών [21].

Η συμπεριφορά αυτή έχει παρατηρηθεί ακόμη σε βιολογικά συστήματα, όπως είναι το lac operon [22] σε βακτήρια *E. coli*, δηλαδή ένα σύνολο γονιδίων τα οποία καταστέλλονται παρουσία γλυκόζης, ενώ μεταγράφονται παρουσία λακτόζης και απουσία γλυκόζης. Ένα χαρακτηριστικό παράδειγμα ακόμη είναι ο ιός phage λ σε κύτταρα *E. coli*, που υπάρχει σε δύο καταστάσεις. Στην πρώτη κατάσταση ο ιός επιβιώνει ανεξάρτητα από τον φορέα (λυσογονική

κατάσταση), ενώ στη δεύτερη τα κύτταρα του *E. coli* που έχουν προσβληθεί από τον ιό οδηγούνται σε λύση (λυτική κατάσταση) [23].

Η διπλοευστάθεια στο toggle switch προκαλείται λόγω της αμοιβαίας αναίρεσης των δύο παρεμποδιστών. Απουσία των επαγωγέων από το σύστημα, δύο καταστάσεις ισορροπίας είναι πιθανές, η κατάσταση χαμηλού και η κατάσταση υψηλού βαθμού έκφρασης ως προς την πράσινη φθορίζουσα πρωτεΐνη (βλ. Σχήμα 2.4) :

- Ο παρεμποδιστής 1 μεταγράφεται, μαζί με την πράσινη φθορίζουσα πρωτεΐνη, ο οποίος με τη σειρά του παρεμποδίζει τη μεταγραφή του παρεμποδιστή 2 (Κατάσταση υψηλού βαθμού έκφρασης).
- Ο παρεμποδιστής 2 μεταγράφεται, ο οποίος με τη σειρά του παρεμποδίζει τη μεταγραφή του παρεμποδιστή 1 (Κατάσταση χαμηλού βαθμού έκφρασης).

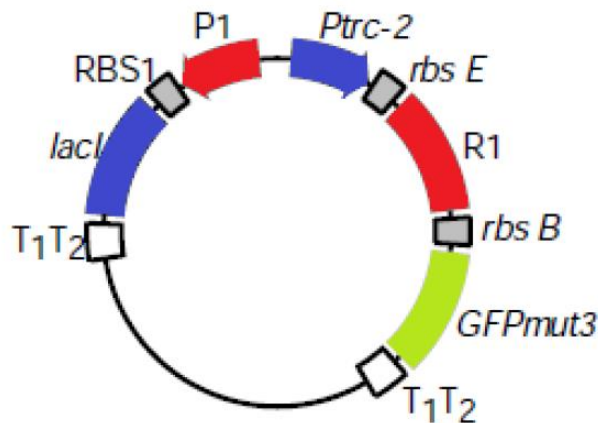


Σχήμα 2.4: Ύπαρξη των δύο καταστάσεων στην περίπτωση απουσίας επαγωγέων [24]. Στην κατάσταση χαμηλού βαθμού έκφρασης, ο *lacI* μεταγράφεται από τον *Ptet* υποκινητή. Ο *lacI* συνδέεται με τον *Plac* υποκινητή, αναστέλλοντας έτσι τη μεταγραφή του *TetR* και του *GFP*. Στην κατάσταση υψηλού βαθμού έκφρασης ο *TetR* και ο *GFP* μεταγράφονται από τον *Plac* υποκινητή. Ο *TetR* συνδέεται με τον *Ptet* υποκινητή, αναστέλλοντας τη μεταγραφή του *lacI*.

Η εναλλαγή από τη μία κατάσταση στην άλλη επιτυγχάνεται με τη σταδιακή εφαρμογή του επαγωγέα (*inducer*) που αντιστοιχεί στον παρεμποδιστή που βρίσκεται σε υψηλή συγκέντρωση

(δηλαδή ο υποκινητής που είναι υπεύθυνος για τη μεταγραφή του είναι ενεργός). Ο επαγωγέας προσδένεται πάνω στα μόρια του παρεμποδιστή που εκφράζεται, απελευθερώνοντας τον υποκινητή που ήταν προσωρινά ανενεργός. Αυτό με τη σειρά του οδηγεί στη μεταγραφή του αντίστοιχου παρεμποδιστή που θα παρεμποδίσει τη μεταγραφή του πρώτου. Για παράδειγμα, αν υποθεθεί ότι ο υποκινητής 1 είναι ενεργός, δηλαδή ο παρεμποδιστής 2 σε υψηλή συγκέντρωση, τότε εφαρμόζοντας τον επαγωγέα 1, θα μεταβούμε στην κατάσταση όπου ο υποκινητής 2 είναι ενεργός, άρα ο παρεμποδιστής 1 σε υψηλή συγκέντρωση. Σημειώνεται ότι το σύστημα αυτό έχει μνήμη, τούτέστιν αν διακοπεί η εφαρμογή του επαγωγέα, το σύστημα δεν επιστρέφει στην πρότερη του κατάσταση αλλά διατηρεί την υπάρχουσα.

Στην κατηγορία pTAK η ουσία isopropyl-b-D-thiogalactopyranoside (IPTG) ή η θέρμανση χρησιμοποιούνται ως επαγωγείς. Αντίστοιχα, στην κατηγορία pIKE ως επαγωγείς χρησιμοποιούνται η ουσίες IPTG και η anhydrotetracycline (aTc).



Σχήμα 2.5: Το toggle switch. Οι υποκινητές (P1 και Ptrc-2), οι παρεμποδιστές (R1 και lacI), ο reporter (GFPmut3), τα ριβοσώματα (RBS1 και rbs B) και οι τερματιστές (T1T2) [16].

Για τη μαθηματική μοντελοποίηση των βιολογικών συστημάτων υπάρχει πληθώρα μεθόδων. Μερικά από τα μαθηματικά μοντέλα που χρησιμοποιούνται είναι τα στοχαστικά [25, 26], τα διακριτά/λογικά μοντέλα, αυτά δηλαδή που χρησιμοποιούν άλγεβρα Boole και χρησιμοποιούνται για ποιοτικά αποτελέσματα περισσότερο [27, 28]. Στο πλαίσιο της παρούσας

εργασίας, η μοντελοποίηση του γενετικού δικτύου γίνεται μέσω συνήθων διαφορικών εξισώσεων.

2.3 Μαθηματικό μοντέλο

Για την περιγραφή του toggle switch σε επίπεδο ενός κυττάρου χρησιμοποιούνται οι παρακάτω συνήθεις διαφορικές εξισώσεις [16], r_1 και r_2 , όπου x , y οι αδιάστατες συγκεντρώσεις των παρεμποδιστών 1 και 2 αντίστοιχα και τ ο αδιάστατος χρόνος (βλ. Κεφάλαιο 3:

$$r_1 \equiv \frac{dx}{d\tau} = \frac{a_1}{1 + y^\beta} - \delta x \quad (2.1)$$

$$r_2 \equiv \frac{dy}{d\tau} = \frac{a_2}{1 + (\frac{x}{g})^\gamma} - \delta y \quad (2.2)$$

$$g = (1 + \frac{[IPTG]}{K})^n \quad (2.3)$$

όπου :

a_1 : αδιάστατος ρυθμός σύνθεσης του παρεμποδιστή 1

a_2 : αδιάστατος ρυθμός σύνθεσης του παρεμποδιστή 2

β : συνεργατικότητα της καταστολής του υποστηρικτή 2

γ : συνεργατικότητα της καταστολής του υποστηρικτή 1

δ : αδιάστατος ρυθμός της απώλειας του ενδοκυτταρικού περιεχομένου (υποβάθμιση)

K : σταθερά διάσπασης του IPTG από τη LacR

n : συνεργατικότητα του IPTG

$[IPTG]$: συγκέντρωση του εξωκυτταρικού επαγωγέα IPTG (M)

Η έκφραση του μοντέλου toggle switch μέσω των ανωτέρω διαφορικών εξισώσεων διατηρεί τις δύο βασικές πτυχές του συγκεκριμένου γενετικού δικτύου : τη συνεργατική παρεμπόδιση των ενεργά μεταγραφόμενων υποστηρικτών (ο 1^{ος} όρος της εξίσωσης) και την υποβάθμιση των παρεμποδιστών (2^{ος} όρος της εξίσωσης).

Στο πλαίσιο της παρούσας εργασίας χρησιμοποιήθηκαν οι ακόλουθες τιμές [16]:

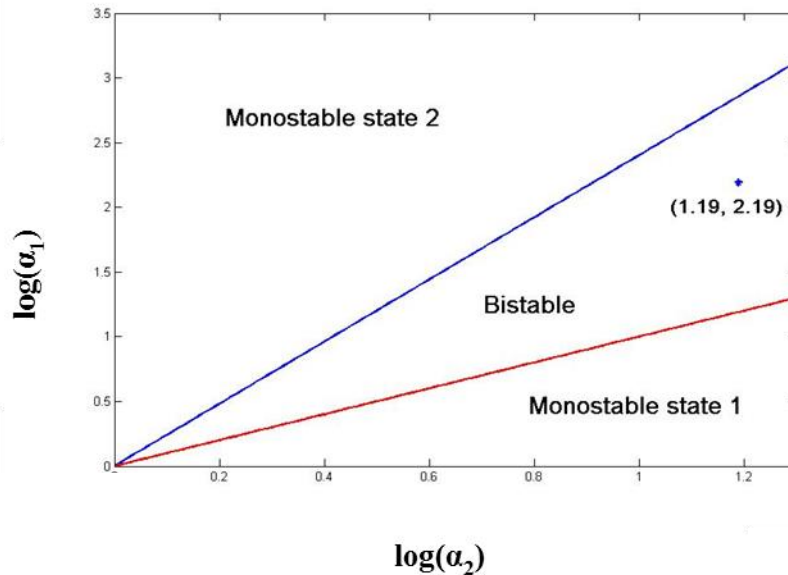
$$\alpha_1=156.25, \alpha_2=15.6, \beta=2.4, \gamma=1$$

$$n=2.0015, \delta=0.005, K=2.9618 \cdot 10^{-5}$$

Αποδεικνύεται μαθηματικά ότι τα όρια της περιοχής της διπλοευστάθειας δίνονται από τις εξισώσεις (βλ. Σχήμα 6) :

$$\log a_1 = \frac{1}{\gamma} \log a_2 \quad (2.4)$$

$$\log a_1 = \beta \log a_2 \quad (2.5)$$



Σχήμα 2.6: Όρια της περιοχής διπλοευστάθειας.

Παρατηρείται ότι οι τιμές των παραμέτρων α_1 και α_2 ανήκουν στην περιοχή διπλοευστάθειας, συνεπώς το σύστημα παρουσιάζει δύο ευσταθείς λύσεις.

2.4 Εξαγωγή μαθηματικού μοντέλου

Η εξαγωγή των εξισώσεων του toggle switch (2.1) και (2.2), προκύπτει από το παρακάτω δίκτυο αντιδράσεων :



Οι δύο πρώτες αντιδράσεις περιγράφουν την πρόσδεση των δύο παρεμποδιστών (R_1 και R_2) στους δύο υποκινητές (P_1 και P_2) προκειμένου να αναστείλουν τη μεταγραφή. Οι δύο τελευταίες αντιδράσεις δηλώνουν ότι τα μονομερή των δύο παρεμποδιστών πολυμερίζονται. Οι σταθερές πρόσδεσης K_1 και K_2 έχουν μονάδες M^{-1} , ενώ οι σταθερές πολυμερισμού, K_3 και K_4 , έχουν μονάδες $(M)^{1-\gamma}$ και $(M)^{1-\beta}$ αντίστοιχα. Οι ρυθμοί των αντιδράσεων είναι :

$$\frac{d[R_1]}{dt} = \frac{k_1[P^T]}{1 + K_1 K_4 [R_2]^\beta} - \varepsilon [R_1] \tag{2.7}$$

$$\frac{d[R_2]}{dt} = \frac{k_2[P^T]}{1 + K_2 K_3 [R_1]^\beta} - \varepsilon [R_2] \tag{2.8}$$

όπου k_1, k_2 οι σταθερές του ρυθμού σύνθεσης των δύο παρεμποδιστών, ε ο ρυθμός διάσπασης τους, $[P^T]$ η συνολική συγκέντρωση των υποκινητών και $[R_1], [R_2]$ οι συγκεντρώσεις των παρεμποδιστών 1 και 2 αντίστοιχα.

Εισάγεται ο αδιάστατος χρόνος $\tau = t / t^* = \mu t$, και οι εξισώσεις (2.7) και (2.8) γίνονται :

$$\frac{d[R_1]}{d\tau} = \frac{1}{\mu} \frac{k_1 [P^T]}{1 + K_1 K_4 [R_2]^\beta} - \delta [R_1] \tag{2.9}$$

$$\frac{d[R_2]}{d\tau} = \frac{1}{\mu} \frac{k_2[P^T]}{1 + K_2K_3[R_1]^\beta} - \delta[R_2] \quad (2.10)$$

όπου $\delta = \varepsilon/\mu$ ο αδιάστατος ρυθμός διάσπασης. Για την πλήρη αδιαστατοποίηση των δύο εξισώσεων, θεωρούνται οι αδιάστατες συγκεντρώσεις :

$$x = [R_1](K_2K_3)^{1/\gamma} \quad (2.11)$$

$$y = [R_2](K_1K_4)^{1/\beta} \quad (2.12)$$

και οι αδιάστατοι ρυθμοί σύνθεσης :

$$\alpha_1 = \frac{k_1[P^T](K_2K_3)^{1/\gamma}}{\mu} \quad (2.13)$$

$$\alpha_2 = \frac{k_2[P^T](K_1K_4)^{1/\beta}}{\mu} \quad (2.14)$$

Τέλος, δεδομένου ότι ο επαγωγέας IPTG προσδένεται πάνω στον παρεμποδιστή 1, απελευθερώνοντας έτσι τον υποκινητή 1, ενώ η ουσία IPTG διασπάται από την LacR, προκύπτουν οι εξισώσεις (2.1)-(2.3).

Για την περιγραφή της συμπεριφοράς ολόκληρου του κυτταρικού πληθυσμού χρησιμοποιούνται τα Ισοζύγια Κυτταρικών Πληθυσμών (ΙΚΠ), καθώς επίσης τα στοχαστικά μοντέλα, τα οποία θα αναλυθούν στο Κεφάλαιο 3.

Κεφάλαιο 3

Το στοχαστικό μοντέλο

Ένα από τα βασικά μοντέλα που χρησιμοποιούνται για τη μελέτη της επίδρασης της ετερογένειας στους κυτταρικούς πληθυσμούς είναι τα Ισοζύγια Κυτταρικών Πληθυσμών (Cell Population Balance Models). Παρόλα αυτά, η στοχαστικότητα που διέπει την κυτταρική διαίρεση έχει σημαντική επίδραση στους μικρούς κυτταρικούς πληθυσμούς και δεν μπορεί να ληφθεί υπόψη από τα ντετερμινιστικά μοντέλα, όπως τα Ισοζύγια Κυτταρικών Πληθυσμών (ΙΚΠ). Για το λόγο αυτό, χρησιμοποιούνται αλγόριθμοι kinetic Monte Carlo, όπως είναι ο Constant Number Monte Carlo (CNMC) που περιγράφεται στο παρόν κεφάλαιο, ενώ γίνεται και μια σύντομη περιγραφή του ΙΚΠ.

3.1 Ισοζύγια Κυτταρικών Πληθυσμών

Ένα από τα γνωστότερα μαθηματικά μοντέλα προκειμένου να μελετηθεί η κατανομή των φαινοτύπων σε έναν κυτταρικό πληθυσμό είναι τα Ισοζύγια Κυτταρικών Πληθυσμών [29, 30, 31], τα οποία αποτελούν μερικές ολοκληρωτικές διαφορικές εξισώσεις (integro-differential equations), από την επίλυση των οποίων προκύπτει ο αριθμός των κυττάρων ενός κυτταρικού πληθυσμού τη χρονική στιγμή t που έχουν ενδοκυτταρικό περιεχόμενο στα διαστήματα $(x, x+dx)$ και $(y, y+dy)$ (το x και το y εκφρασμένα σε μόρια ή moles) ανά μονάδα όγκου, $F(x, y, t) dx dy$. Συχνά όμως, αντί για την συνάρτηση $F(x, y, t)$, ο άγνωστος είναι η συνάρτηση «πυκνότητας πιθανότητας» των κυττάρων, $n(x, y, t)$ [31] :

$$n(x, y, t) = \frac{F(x, y, t)}{\int_0^{x_{max}} \int_0^{y_{max}} F(x, y, t) dx dy} \quad (3.1)$$

όπου x_{max} και y_{max} οι μέγιστες τιμές που μπορεί να λάβει το ενδοκυτταρικό περιεχόμενο. Η συνάρτηση $n(x, y, t)$ δηλώνει τον αριθμό των κυττάρων με περιεχόμενο x και y , διαιρεμένο με τον συνολικό αριθμό των κυττάρων. Είναι φανερό από την εξίσωση (3.1) ότι πρέπει να ισχύει η συνθήκη κανονικοποίησης :

$$\int_0^{x_{max}} \int_0^{y_{max}} n(x, y, t) dx dy = 1 \quad (3.2)$$

Για την πλήρη περιγραφή του κυτταρικού πληθυσμού απαιτούνται τριών ειδών συναρτήσεις που αναφέρονται στη διεθνή βιβλιογραφία ως Intrinsic Physiological State Functions (IPSF) [32]. Οι συναρτήσεις αυτές είναι :

- Οι ρυθμοί των ενδοκυτταρικών αντιδράσεων
- Ο ρυθμός διαίρεσης του κυττάρου
- Η συνάρτηση κατανομής του κυτταρικού περιεχομένου

Οι ρυθμοί των ενδοκυτταρικών αντιδράσεων στην παρούσα μελέτη είναι οι εξισώσεις (2.1) και (2.2). Ο ρυθμός της διαίρεσης του κυττάρου δίνεται από τον τύπο (power law) [33] :

$$\gamma(x, y) = \mu \left(\frac{x}{\langle x \rangle} \right)^{m_1} \left(\frac{y}{\langle y \rangle} \right)^{m_2} \quad (3.3)$$

όπου m_1, m_2 οι παράμετροι που ρυθμίζουν την ταχύτητα της κυτταρικής διαίρεσης και μ είναι το μέτρο που εκφράζει το ρυθμό ανάπτυξης του πληθυσμού και έχει μονάδες αντίστροφου χρόνου. Η παραπάνω σχέση δεν έχει χρησιμοποιηθεί μόνο σε υπολογιστικά μοντέλα, αλλά έχει βρεθεί να περιγράφει πειραματικά δεδομένα κυτταρικών πληθυσμών ζυμών (yeast). Μεγάλες τιμές των εκθετών m_1 και m_2 αντιστοιχούν σε πιο απότομη διαίρεση ή πιο συγκεκριμένα σε μεγαλύτερη διαφορά ανάμεσα στους ρυθμούς διαίρεσης κυττάρων με ενδοκυτταρικό περιεχόμενο μεγαλύτερο και μικρότερο από κάποιο «φράγμα» ενδοκυτταρικού περιεχομένου. Για την επίλυση του ΙΚΠ, θεωρείται ένας χρόνος αναφοράς, t^* , ο οποίος ορίζεται ως :

$$t^* = \frac{1}{\mu} \quad (3.4)$$

Με βάση με αυτόν το χρόνο αναφοράς γίνεται χρονική αδιαστατοποίηση, και η αδιάστατη πλέον εξίσωση (3.3) γράφεται :

$$\Gamma(x, y) = \left(\frac{x}{\langle x \rangle}\right)^{m_1} \left(\frac{y}{\langle y \rangle}\right)^{m_2} \quad (3.5)$$

με :

$$\tau = \frac{t}{t^*} \quad (3.6)$$

Για τη μελέτη της επίδρασης της ασυμμετρίας στην κατανομή του ενδοκυτταρικού περιεχομένου κατά την κυτταρική διαίρεση, θεωρήθηκε η απλούστερη δυνατή μορφή για τη συνάρτηση κατανομής του κυτταρικού περιεχομένου [34] :

$$P_1(x, x') = \frac{1}{2f} \delta(fx' - x) + \frac{1}{2(1-f)} \delta((1-f)x' - x) \quad (3.7)$$

$$P_2(y, y') = \frac{1}{2f} \delta(fy' - y) + \frac{1}{2(1-f)} \delta((1-f)y' - y) \quad (3.8)$$

όπου δ είναι η συνάρτηση Dirac και f το κλάσμα του ενδοκυτταρικού περιεχομένου που κληροδοτείται στο μικρότερο θυγατρικό κύτταρο και αντίστοιχα το κλάσμα $1-f$ στο άλλο θυγατρικό κύτταρο. Είναι φανερό ότι με τον ορισμό αυτό, η τιμή f κυμαίνεται μεταξύ του 0 και του 0.5.

Οι δύο παραπάνω συναρτήσεις αποτελούν συναρτήσεις πυκνότητας πιθανότητας οι οποίες περιγράφουν την πιθανότητα ένα κύτταρο με κυτταρικό περιεχόμενο x' και y' να διαιρεθεί σε δύο θυγατρικά κύτταρα με το πρώτο να έχει κυτταρικό περιεχόμενο x, y ενώ το δεύτερο $x'-x$ και $y'-y$. Λόγω της διατήρησης της μάζας πρέπει να ισχύει η παρακάτω συνθήκη για τις δύο συναρτήσεις :

$$P_1(x, x') = P_1(x' - x, x') \quad (3.9)$$

$$P_2(y, y') = P_2(y' - y, y') \quad (3.10)$$

Οι δύο παραπάνω συναρτήσεις κατανομής μπορούν να λάβουν υπόψη τους την ασύμμετρη κατανομή του κυτταρικού περιεχομένου κατά την κυτταρική διαίρεση. Προφανώς για $f=0.5$ έχουμε συμμετρική κατανομή του κυτταρικού περιεχομένου. Λαμβάνοντας υπόψη τα παραπάνω το Ισοζύγιο Κυτταρικού Πληθυσμού (ΙΚΠ) που περιγράφει το παρόν σύστημα είναι [35] :

$$\begin{aligned} & \frac{\partial n(x, y, \tau)}{\partial \tau} + \frac{\partial}{\partial x} (R_1(x, y)n(x, y, \tau)) + \frac{\partial}{\partial y} (R_2(x, y)n(x, y, \tau)) = \\ & = \frac{1}{f^2} \Gamma\left(\frac{x}{f}, \frac{y}{f}\right) n\left(\frac{x}{f}, \frac{y}{f}, \tau\right) + \frac{1}{(1-f)^2} \Gamma\left(\frac{x}{1-f}, \frac{y}{1-f}\right) n\left(\frac{x}{1-f}, \frac{y}{1-f}, \tau\right) \\ & - n(x, y, \tau) \int_0^{x_{max}} \int_0^{y_{max}} \Gamma(x, y)n(x, y, \tau) dx dy - \Gamma(x, y)n(x, y, \tau) \end{aligned} \quad (3.11)$$

Οι συνοριακές συνθήκες είναι :

$$n(0, y, \tau) = n(x_{max}, y, \tau) = n(x, 0, \tau) = n(x, y_{max}, \tau) = 0 \quad (3.12)$$

Σύμφωνα με τις συνοριακές συνθήκες του ΙΚΠ, δεν μπορεί να υπάρξει κύτταρο με ενδοκυτταρικό περιεχόμενο μεγαλύτερο του x_{max} και του y_{max} . Η αρχική συνθήκη για την εξίσωση (3.11) είναι μία κατανομή $n_0(x, y, 0)$ (πχ. μια μετατοπισμένη κατανομή Gauss με μέση τιμή $\mu_1 = \mu_2$ και τυπική απόκλιση $\sigma_1 = \sigma_2$).

Ένας από τους βασικούς περιορισμούς στην εφαρμογή των Ισοζυγίων Κυτταρικών Πληθυσμών για τη μελέτη της δυναμικής συμπεριφοράς των κυτταρικών πληθυσμών και της αλληλεπίδρασης τους με τα γενετικά δίκτυα, είναι η δυσκολία επίλυσης τους. Αναλυτικές λύσεις έχουν προσδιορισθεί μόνο για πολύ ειδικές περιπτώσεις. Στη γενική περίπτωση τα Ισοζύγια Κυτταρικών Πληθυσμών δεν μπορούν να λυθούν αναλυτικά, καθιστώντας έτσι την εφαρμογή αριθμητικών μεθόδων απαραίτητη.

Το εύρος των αριθμητικών μεθόδων που μπορούν να εφαρμοστούν για την επίλυση των ΙΚΠ είναι μεγάλο. Παρόλα αυτά, όπως αναφέρθηκε προηγουμένως, για την επίλυσή τους πρέπει να είναι γνωστό το μέγιστο ενδοκυτταρικό περιεχόμενο, κάτι που είναι πρακτικά αδύνατο, σε

αντίθεση με το μικρότερο ενδοκυτταρικό περιεχόμενο που μπορεί να θεωρηθεί γνωστό και ίσο με το μηδέν. Επομένως, στις πρακτικές εφαρμογές, η κατανομή των φαινοτύπων εξελίσσεται δυναμικά σε ένα άγνωστο χωρίο ενδοκυτταρικών περιεχομένων. Συνεπώς, είναι επιτακτική η ανάγκη διαμόρφωσης αλγορίθμων με κινητά ή ελεύθερα σύνορα (moving or free boundary algorithms) για την αντιμετώπιση της παραπάνω δυσκολίας.

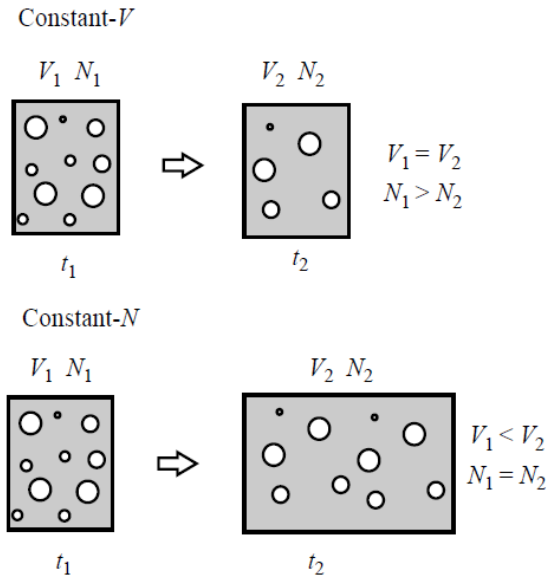
3.2 Εισαγωγή στα στοχαστικά μοντέλα

Τα ντετερμινιστικά μοντέλα παρότι δίνουν εξαιρετική πρόβλεψη της συμπεριφοράς μεγάλων κυτταρικών πληθυσμών, εν τούτοις για μικρούς κυτταρικούς πληθυσμούς αποκλίνουν λόγω της «τυχαιότητας» που διέπει την κυτταρική διαίρεση. Επίσης, τα ντετερμινιστικά μοντέλα δεν λαμβάνουν υπόψη τους τον ενδογενή θόρυβο, ο οποίος προέρχεται από τις αντιδράσεις που συμβαίνουν εντός του κυττάρου [36].

Για να ξεπεραστεί αυτή η «αστοχία» των ντετερμινιστών μοντέλων, χρησιμοποιούνται αλγόριθμοι τύπου kinetic Monte Carlo. Ο όρος Monte Carlo αναφέρεται σε μία ευρεία κατηγορία αλγορίθμων, οι οποίοι στοχεύουν στην επίλυση προβλημάτων με τη χρήση τυχαίων αριθμών. Ο πιο διάσημος αλγόριθμος που ανήκει στην κατηγορία αυτή είναι ο αλγόριθμος Metropolis [37] και αποτελεί έναν πολύ αποτελεσματικό τρόπο εξαγωγής κατάλληλων δειγμάτων από πολύπλοκες κατανομές πιθανοτήτων, ενώ χρησιμοποιείται επίσης σε πολλούς κλάδους και για μεγάλο εύρος εφαρμογών. Παρόλα αυτά, ο αλγόριθμος αυτός δεν είναι κατάλληλος για προσομοιώσεις σε μεγάλες χρονικές κλίμακες. Στην περίπτωση μεγάλων χρονικών κλιμάκων χρησιμοποιούνται αλγόριθμοι τύπου kinetic Monte Carlo [38]. Η μόνη αδυναμία των αλγορίθμων αυτών είναι ότι πρέπει να είναι γνωστοί οι ρυθμοί των κινητικών που χρησιμοποιούνται, οι οποίοι υπολογίζονται είτε πειραματικά είτε με υπολογιστικές μεθόδους. Οι αλγόριθμοι kinetic Monte Carlo αδυνατούν να υπολογίσουν τους ρυθμούς των κινητικών από μόνοι τους.

Δύο διαδεδομένοι τύποι αλγορίθμων kinetic Monte Carlo είναι ο Constant Number Monte Carlo (CNMC) και ο Constant Volume Monte Carlo (CVMC) (βλ. Σχήμα 4.1) [39, 40]. Όπως γίνεται κατανοητό από το όνομα των δύο αλγορίθμων, στο μεν CNMC το πλήθος του δείγματος διατηρείται σταθερό, ενώ στον αλγόριθμο CVMC ο όγκος. Στον αλγόριθμο CNMC ο αριθμός των σωματιδίων διατηρείται σταθερός, ανεξαρτήτως από το αν η πραγματική διαδικασία οδηγεί σε απώλεια ή αύξησή τους [41]. Σε πραγματικές διεργασίες, ο αριθμός συγκέντρωσης των σωματιδίων ενός δείγματος μπορεί να αυξάνεται, να μειώνεται ή να υφίσταται πιο πολύπλοκες

μεταβάσεις. Ανεξάρτητα όμως από τη συμπεριφορά που ακολουθούν τα σωματίδια, στον αλγόριθμο CVMC η αύξηση ή η συρρίκνωση του αριθμού των σωματιδίων του δείγματος είναι ευθέως ανάλογη της συγκέντρωσης των σωματιδίων του φυσικού συστήματος, για το οποίο εκτελείται η προσομοίωση [42].



Σχήμα 3.1: Σχηματική αναπαράσταση του CVMC (πάνω) και του CNMC (κάτω) σε ένα πρόβλημα συσσωμάτωσης σωματιδίων [42].

3.2 Ανάπτυξη μοντέλου

Οι αλγόριθμοι που αναφέρθηκαν προηγουμένως έχουν ως στόχο να προσεγγίσουν την εξίσωση master. Η εξίσωση master αποτελεί μια διαφορική εξίσωση, η επίλυση της οποίας δίνει την πιθανότητα το σύστημα μας να βρίσκεται σε μια δεδομένη κατάσταση. Συγκεκριμένα, ορίζουμε $\mathbf{X}(t) = \mathbf{x}$ το διάνυσμα κατάστασης του συστήματος τη χρονική στιγμή t και $\mathbf{X}(t_0) = \mathbf{x}_0$ την αρχική κατάσταση για χρόνο t_0 . Η πιθανότητα που βρίσκεται με την επίλυση της εξίσωσης master ορίζεται ως :

$$P(\mathbf{x}, t | \mathbf{x}_0, t_0) = \text{η πιθανότητα το σύστημα να βρίσκεται στην κατάσταση } \mathbf{x} \text{ τη χρονική στιγμή } t, \text{ για δεδομένη κατάσταση } \mathbf{x}_0 \text{ (} t \geq t_0 \text{)} \quad (3.13)$$

Παρόλα αυτά, η επίλυση της εξίσωσης master, είτε αναλυτικά είτε αριθμητικά, είναι πρακτικά αδύνατη. Για το λόγο αυτό, αναπτύχθηκε ένα ισοδύναμο μοντέλο από τον Gillespie το 1976, γνωστό ως Stochastic Simulation Algorithm (SSA) [43].

3.2.1 Reaction probability density function

Για τη μελέτη της χρονικής εξέλιξης του συστήματος δύο ερωτήσεις είναι αυτές που πρέπει να απαντηθούν προκειμένου να έχουμε πλήρη περιγραφή : πότε θα γίνει η επόμενη διαίρεση και ποιο κύτταρο θα διαιρεθεί; Για την απάντηση των δύο αυτών ερωτήσεων ορίζεται η παρακάτω συνάρτηση $P(\tau, j)$:

$P(\tau, j)d\tau \equiv$ η πιθανότητα, δεδομένης μιας κατάστασης $S_i = \{M_0(t)=N; \mathbf{X}(t)=\mathbf{x}_i, i=1,2,\dots,N\}$ τη χρονική στιγμή t , η επόμενη αντίδραση θα πραγματοποιηθεί σε χρονικό διάστημα $(t+\tau, t+\tau+d\tau)$ και θα είναι το κύτταρο j , όπου $M_0(t)$ ο αριθμός των κυττάρων τη χρονική στιγμή t και \mathbf{X} το διάνυσμα κατάστασης του συστήματος που περιλαμβάνει το ενδοκυτταρικό περιεχόμενο για κάθε κύτταρο

Η συνάρτηση $P(\tau, j)$ καλείται «reaction probability density function» και αποτελεί μία διμεταβλητή συνάρτηση πυκνότητας πιθανότητας, όπου η τ αποτελεί μία συνεχή τυχαία μεταβλητή ($0 \leq \tau < \infty$), ενώ η j μια διακριτή ($j=1,2,\dots,N$). Η εύρεση των δύο αυτών τιμών μας δίνει την απάντηση στα δύο ερωτήματα που τέθηκαν προηγουμένως. Η βασική (και μοναδική ταυτόχρονα) υπόθεση που βασίζεται το μοντέλο αυτό προκειμένου να γίνει η εύρεση τους είναι :

$$\Gamma(x_j(t), y_j(t))dt = \text{η πιθανότητα να διαιρεθεί το κύτταρο } j \text{ σε χρονικό διάστημα } dt \quad (3.15)$$

Για τον υπολογισμό της πιθανότητας στη σχέση (3.14), αρχικά ορίζεται η $P_0(\tau)$ που είναι η πιθανότητα να μη γίνει καμία κυτταρική διαίρεση το χρονικό διάστημα $(t, t+\tau)$. Στη συνέχεια η $P(\tau, j)$ γράφεται ως γινόμενο των πιθανοτήτων P_0 και της πιθανότητας της σχέσης (3.15), δηλαδή :

$$P(\tau, j)d\tau = P_0(\tau)\Gamma(x_j(\tau), y_j(\tau))d\tau \quad (3.16)$$

Η έκφραση για την πιθανότητα $P_0(\tau)$ δίνεται από την επίλυση της παρακάτω διαφορικής εξίσωσης :

$$P_0(\tau' + d\tau') = P_0(\tau')[1 - \sum_{j=1}^N \Gamma(x_j(\tau'), y_j(\tau')) d\tau'] \quad (3.17)$$

Ο όρος $[1 - \sum_{j=1}^N \Gamma(x_j(\tau'), y_j(\tau')) d\tau']$ είναι η πιθανότητα να μη γίνει καμία κυτταρική διαίρεση το χρονικό διάστημα $d\tau'$. Από την επίλυση της εξίσωσης (3.17) και με αρχική συνθήκη $P_0(0)=1$ προκύπτει :

$$P_0(\tau) = \exp\left(-\int_0^{\tau} \sum_{j=1}^N \Gamma(x_j(\tau'), y_j(\tau')) d\tau'\right) \quad (3.18)$$

Με αντικατάσταση της σχέσης (3.18) στην (3.16) προκύπτει :

$$P(\tau, j) = \begin{cases} \Gamma(x_j(\tau), y_j(\tau)) \exp\left(-\int_0^{\tau} \sum_{j=1}^N \Gamma(x_j(\tau'), y_j(\tau')) d\tau'\right), & j = 1, 2 \dots N \text{ και } 0 \leq \tau \leq \infty \\ 0, & \text{αλλιώς} \end{cases} \quad (3.19)$$

3.2.2 Stochastic Simulation Algorithm (SSA)

Από την εξίσωση (3.19) γίνεται εμφανές ότι πρέπει να βρεθεί το ζευγάρι (τ, j) που θα την επαληθεύει. Η $P(\tau, j)$ μπορεί να γραφεί ως γινόμενο δύο επιμέρους πιθανοτήτων (conditioning), $P_1(\tau)d\tau$ και $P_2(j|\tau)$, όπου η $P_1d\tau$ δηλώνει την πιθανότητα να γίνει μία οποιαδήποτε κυτταρική διαίρεση το χρονικό διάστημα $(t+\tau, t+\tau +d\tau)$ και P_2 να διαιρεθεί το κύτταρο j τη χρονική στιγμή $t+\tau$. Συνεπώς η $P(\tau, j)$ γράφεται :

$$P(\tau, j) = P_1(\tau)P_2(j|\tau) \quad (3.20)$$

Η πιθανότητα $P_j(\tau)$ υπολογίζεται βάσει του θεωρήματος ολικής πιθανότητας :

$$P_1(\tau) = \sum_{j=1}^N \Gamma(x_j(\tau), y_j(\tau)) \exp\left(-\int_0^{\tau} \sum_{j=1}^N \Gamma(x_j(\tau'), y_j(\tau')) d\tau'\right) \quad (3.21)$$

Αντικαθιστώντας τις σχέσεις (3.19) και (3.21) στην (3.20) και επιλύοντας ως προς $P_2(j|\tau)$, προκύπτει :

$$P_2(j|\tau) = \frac{\Gamma(x_j(\tau), y_j(\tau))}{\sum_{j=1}^N \Gamma(x_j(\tau), y_j(\tau))} \quad (3.22)$$

Από την εξίσωση (3.21) μπορεί να προκύψει με ολοκλήρωση η αθροιστική συνάρτηση πυκνότητας πιθανότητας $F(\tau)$, άρα :

$$F(\tau \leq T) = 1 - \exp\left(-\int_0^T \sum_{j=1}^N \Gamma(x_j(\tau'), y_j(\tau')) d\tau'\right) \quad (3.23)$$

Προκειμένου να βρεθεί ο χρόνος που μεσολαβεί μεταξύ δύο κυτταρικών διαιρέσεων, πρέπει να επιλυθεί η εξίσωση (3.23). Μια ισοδύναμη σχέση της (3.23) είναι :

$$Q(T) \equiv \frac{\int_0^T \sum_{j=1}^N \Gamma(x_j(\tau'), y_j(\tau')) d\tau'}{\ln(1 - rand_1)} + 1 = 0 \quad (3.24)$$

Σύμφωνα με την παραπάνω σχέση, μέσω της γέννησης ενός τυχαίου αριθμού, $rand_1$ από την ομοιόμορφη κατανομή $(0,1]$, ο οποίος αντιστοιχίζεται σε μία τιμή της συνάρτησης $F(\tau \leq T)$,

γίνεται εύρεση του αδιάστατου χρόνου, T , που μεσολαβεί μεταξύ δύο κυτταρικών διαιρέσεων μέσω της επίλυσης της (3.24). Στη συνέχεια πρέπει να βρεθεί το κύτταρο το οποίο θα διαιρεθεί.

3.2.3 Ευθεία Μέθοδος (Direct Method)

Για την επιλογή του γεγονότος που θα πραγματοποιηθεί σε κάθε βήμα της προσομοίωσης υπάρχουν διάφορες παραλλαγές του παραπάνω αλγορίθμου. Δύο πολύ διαδεδομένες παραλλαγές του είναι :

- Η μέθοδος πρώτης αντίδρασης (First Reaction Method)
- Η ευθεία μέθοδος (Direct Method)

Για τη μέθοδο πρώτης αντίδρασης παράγονται χρόνοι άφιξης για όλα τα δυνατά γεγονότα και πραγματοποιείται σε κάθε βήμα αυτό με το μικρότερο χρόνο άφιξης, δηλαδή επιλέγεται η εξίσωση (3.24) N φορές, μία για κάθε κύτταρο στην προκειμένη περίπτωση, και αυτό με τον μικρότερο χρόνο, T , επιλέγεται. Στην παρούσα εργασία χρησιμοποιήθηκε η ευθεία μέθοδος (Direct Method). Παρότι οι δύο μέθοδοι είναι ισοδύναμες, η μέθοδος πρώτης αντίδρασης είναι λιγότερο αποδοτική από άποψη υπολογιστικού κόστους, καθώς απαιτεί N τυχαίους αριθμούς προκειμένου να επιλυθεί η εξίσωση (3.24), ενώ όπως θα δειχθεί στη συνέχεια η ευθεία μέθοδος χρειάζεται μόνο δύο.

Για την ευθεία μέθοδο απαιτούνται δύο τυχαίοι αριθμοί. Ο πρώτος τυχαίος αριθμός ($rand_1$) χρησιμοποιείται προκειμένου να επιλυθεί η εξίσωση (3.24) όπως αναφέρθηκε προηγουμένως. Ο δεύτερος τυχαίος αριθμός ($rand_2$), που ανήκει στο διάστημα $(0,1]$, χρησιμοποιείται προκειμένου να βρεθεί το κύτταρο το οποίο θα διαιρεθεί σύμφωνα με την παρακάτω σχέση :

$$\frac{\sum_{i=1}^{j-1} \Gamma(x_i(\tau), y_i(\tau))}{\sum_{i=1}^N \Gamma(x_i(\tau), y_i(\tau))} < rand_2 \leq \frac{\sum_{i=1}^j \Gamma(x_i(\tau), y_i(\tau))}{\sum_{i=1}^N \Gamma(x_i(\tau), y_i(\tau))} \quad (3.25)$$

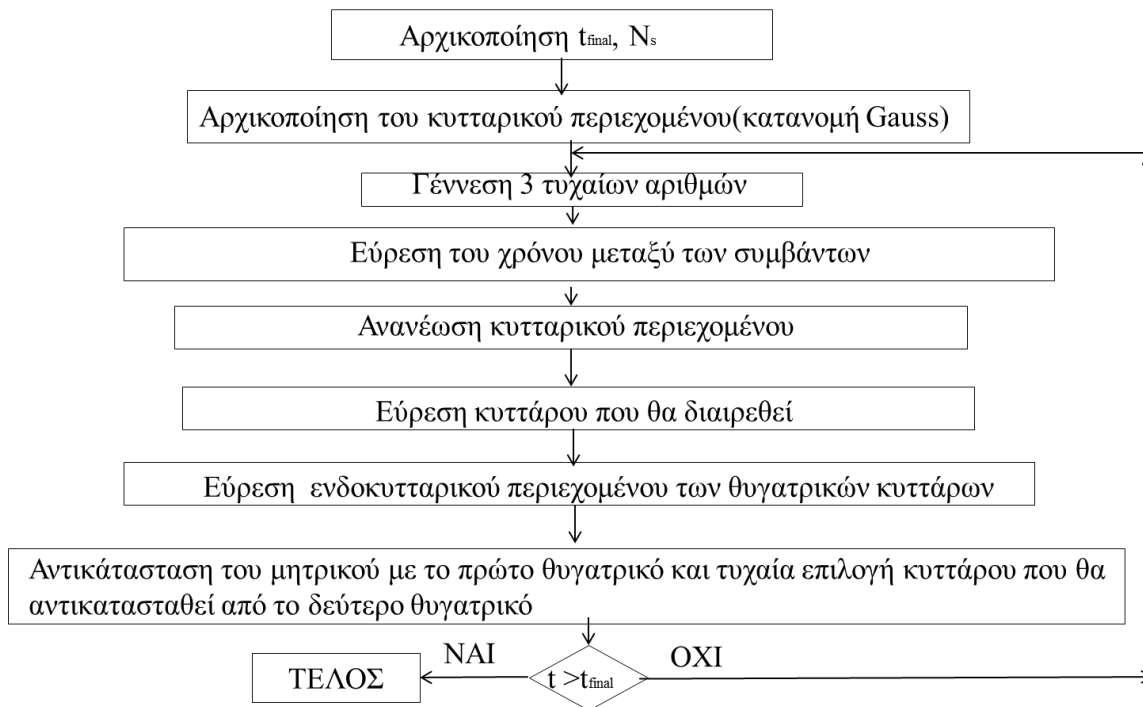
Σύμφωνα με την σχέση (3.25), αν ο δεύτερος τυχαίος αριθμός βρίσκεται μεταξύ αυτών των δύο τιμών, οι οποίες αντιστοιχούν σε τιμές μιας αθροιστικής συνάρτησης πυκνότητας πιθανότητας, τότε επιλέγεται το κύτταρο j προκειμένου να διαιρεθεί. Η μαθηματική εξαγωγή της παραπάνω σχέσης παρατίθεται στο παράρτημα.

Για την παραπάνω σχέση, όσο αφορά στο υπολογιστικό κομμάτι, αρχικά υπολογίζονται τα αθροίσματα και στη συνέχεια γεννάται τυχαίος αριθμός. Για τον υπολογισμό των παραπάνω αθροισμάτων, πρέπει οι τιμές της συνάρτησης $\Gamma(x, y)$ να ταξινομηθούν κατά αύξουσα σειρά, μαζί με τα αντίστοιχα x, y , μέσω μιας υπορουτίνας Quicksort, προκειμένου να υπολογιστούν οι τιμές της αθροιστικής συναρτήσεως πυκνότητας πιθανότητας. Στη συνέχεια γεννάται τυχαίος αριθμός και γίνεται αναζήτηση μεταξύ ποιών τιμών βρίσκεται.

3.3 Constant Number Monte Carlo (CNMC)

Ο αλγόριθμος CNMC στηρίζεται στον SSA που περιγράφηκε προηγουμένως. Τα βήματα που ακολουθούνται για την υλοποίησή του είναι τα ακόλουθα :

1. Αρχικοποίηση χρόνου προσομοίωσης, t_{final} , και αριθμού κυττάρων, N .
2. Αρχικοποίηση του κυτταρικού περιεχομένου, βάσει μιας κατανομής Gauss
3. Γέννεση τριών τυχαίων αριθμών από την ομοιόμορφη κατανομή
4. Εύρεση του χρόνου που μεσολαβεί μεταξύ δύο κυτταρικών διαιρέσεων μέσω της εξίσωσης (3.24)
5. Ανανέωση του κυτταρικού περιεχομένου μέσω των εξισώσεων (2.1) και (2.2)
6. Εύρεση του κυττάρου που θα διαιρεθεί βάσει της σχέσης (3.25)
7. Εύρεση του κυτταρικού περιεχομένου των δύο θυγατρικών κυττάρων βάσει των συναρτήσεων κατανομής, (3.7) και (3.8)
8. Αντικατάσταση του μητρικού κυττάρου με το πρώτο θυγατρικό
9. Αντικατάσταση ενός κυττάρου από τον κυτταρικό πληθυσμό με το δεύτερο θυγατρικό, βάσει του τρίτου τυχαίου αριθμού, $rand_3$.
10. Αν $t \leq t_{final}$ επέστρεψε στο βήμα 3, αλλιώς τέλος αλγορίθμου



Σχήμα 3.2: Αλγόριθμος Constant Number Monte Carlo.

Σημειώνεται ότι για την επίλυση των διαφορικών εξισώσεων που αντιστοιχούν στους ρυθμούς των ενδοκυτταρικών αντιδράσεων χρησιμοποιήθηκε η άμεση μέθοδος Euler, η οποία όχι μόνο δίνει αξιόπιστα αποτελέσματα στην συγκεκριμένη περίπτωση, αλλά είναι οικονομικότερη από άποψη υπολογιστικού κόστους, εν αντιθέσει με τις μεθόδους τύπου Runge-Kutta που είναι κοστοβόρες. Επίσης, ο υπολογισμός του ολοκληρώματος στην εξίσωση (3.24) έγινε με τον κανόνα του τραπεζίου, ενώ η επίλυσή της με την επαναληπτική μέθοδο Newton-Raphson. Η επίλυσή της είναι απλή, καθώς ο κανόνας του τραπεζίου οδηγεί το ολοκλήρωμα σε ένα πολώνυμο πρώτης τάξης ως προς T , οπότε η Ιακωβιανή που προκύπτει είναι σταθερός αριθμός. Στο τέλος του κώδικα ελέγχεται αν το περιεχόμενο σε κάθε κύτταρο είναι μεγαλύτερο ή ίσο του μηδενός. Αν είναι αρνητικό, τίθεται ίσο με το μηδέν.

Κεφάλαιο 4

Αποτελέσματα

Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα που προέκυψαν από το στοχαστικό μοντέλο. Οι κώδικες για την πραγματοποίηση της υπολογιστικής μελέτης είναι γραμμένοι σε γλώσσα προγραμματισμού C++ και εκτελούνται στο εμπορικό λογισμικό MATLAB. Οι υπολογισμοί εκτελέστηκαν στο cluster Andromeda. Αρχικά, παρουσιάζεται η χρονική εξέλιξη των μέσων συγκεντρώσεων για τους δύο παρεμποδιστές για διαφορετικούς πληθυσμούς, ενώ γίνεται σύγκριση με το ντετερμινιστικό μοντέλο. Στη συνέχεια παρουσιάζεται η χρονική εξέλιξη των κατανομών και παρατηρείται τότε το σύστημα οδηγείται σε μόνιμη κατάσταση.

4.1 Εισαγωγή

Όπως αναφέρθηκε στα προηγούμενα κεφάλαια, για τη μελέτη της συμπεριφοράς του κυτταρικού πληθυσμού που φέρει το γενετικό δίκτυο toggle switch εφαρμόστηκε στοχαστικός αλγόριθμος τύπου kinetic Monte Carlo και συγκεκριμένα ο Constant Number Monte Carlo. Τα αποτελέσματα που προέκυψαν από τον στοχαστικό προσομοιωτή, συγκρίθηκαν με αυτά του ντετερμινιστικού μοντέλου, το οποίο είχε μελετηθεί σε πρόσφατη εργασία, από την οποία προέκυψαν τα αποτελέσματα για το ντετερμινιστικό μοντέλο [44]. Συνοψίζονται οι αντιδράσεις που περιγράφουν το toggle switch:

$$r_1 \equiv \frac{dx}{d\tau} = \frac{a_1}{1 + y^\beta} - \delta x \quad (2.1)$$

$$r_2 \equiv \frac{dy}{d\tau} = \frac{a_2}{1 + (\frac{x}{g})^\gamma} - \delta y \quad (2.2)$$

$$g = (1 + \frac{[IPTG]}{K})^n \quad (2.3)$$

Ο ρυθμός κυτταρικής διαίρεσης των κυττάρων δίνεται από τη σχέση :

$$\Gamma(x, y) = \left(\frac{x}{\langle x \rangle}\right)^{m_1} \left(\frac{y}{\langle y \rangle}\right)^{m_2} \quad (3.5)$$

ενώ το ενδοκυτταρικό περιεχόμενο των θυγατρικών κυττάρων από τις συναρτήσεις πυκνότητας πιθανότητας

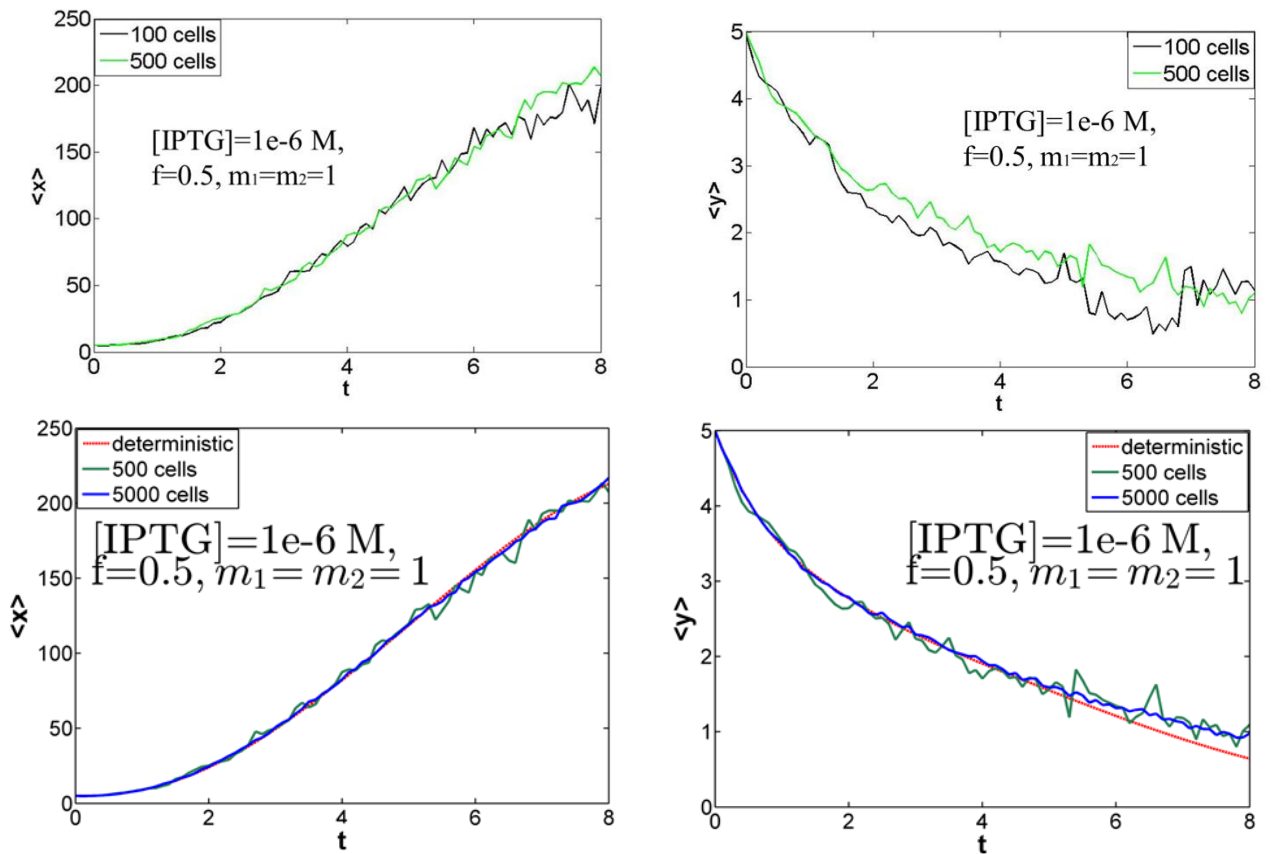
$$P_1(x, x') = \frac{1}{2f} \delta(fx' - x) + \frac{1}{2(1-f)} \delta((1-f)x' - x) \quad (3.7)$$

$$P_2(y, y') = \frac{1}{2f} \delta(fy' - y) + \frac{1}{2(1-f)} \delta((1-f)y' - y) \quad (3.8)$$

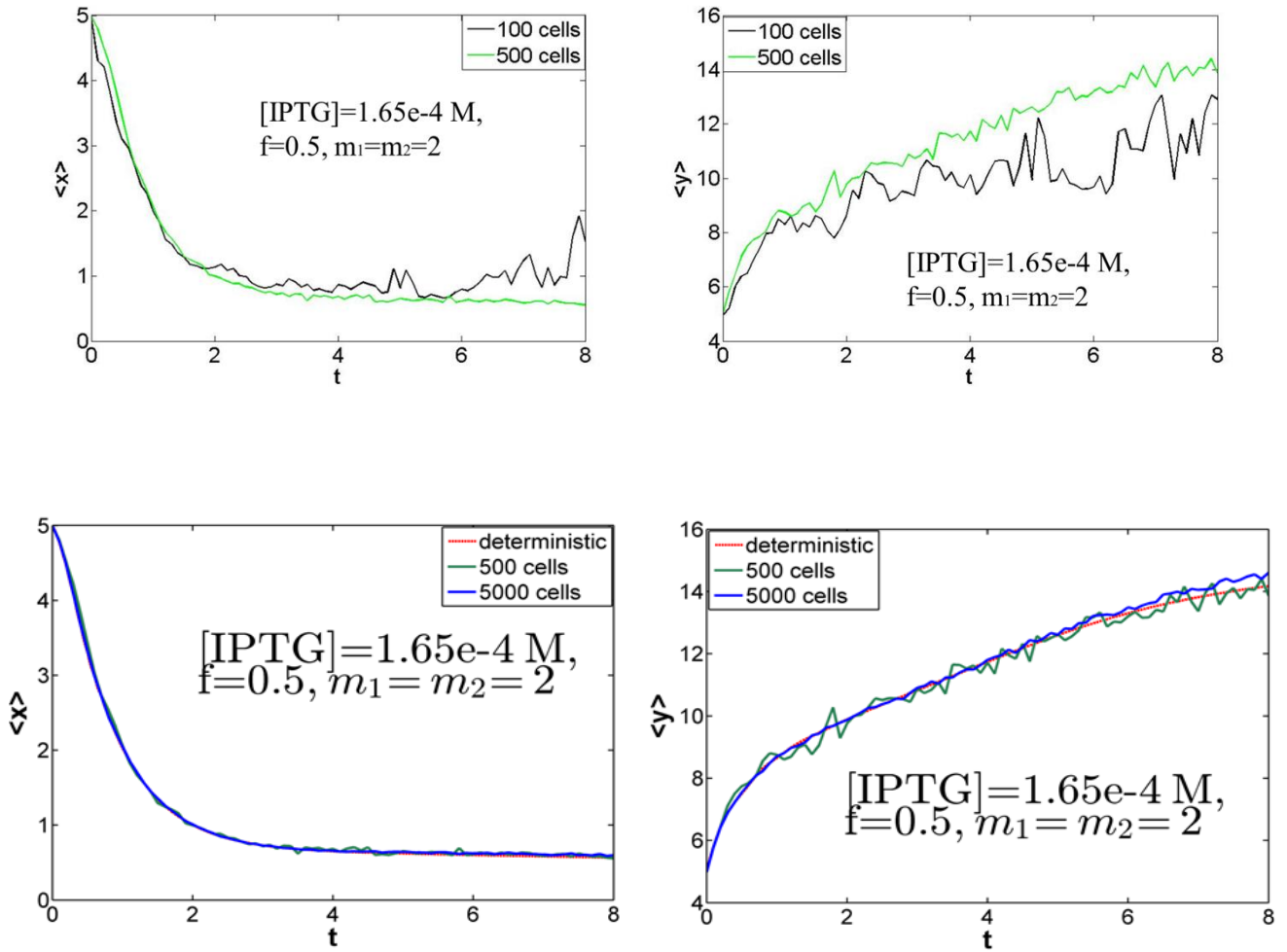
Στις προσομοιώσεις του αλγορίθμου kinetic Monte Carlo χρησιμοποιήθηκαν πέντε αντίγραφα προκειμένου να μειωθεί το στατιστικό σφάλμα. Συνεπώς, οι κατανομές που προέκυψαν είναι ο μέσος όρος μεταξύ πέντε αντιγράφων. Για τον υπολογισμό της χρονικής εξέλιξης των μέσων συγκεντρώσεων εκτελέστηκε ο στοχαστικός αλγόριθμος για διαφορετικούς, αδιάστατους χρονικούς ορίζοντες με βήμα 0.1. Σε κάθε προσομοίωση χρησιμοποιείται ως αρχική συνθήκη μια μετατοπισμένη κατανομή Gauss με μέσες τιμές $\mu_1 = \mu_2 = 5$ και τυπικές αποκλίσεις $\sigma_1 = \sigma_2 = 0.25$.

4.2 Χρονική εξέλιξη μέσω συγκεντρώσεων

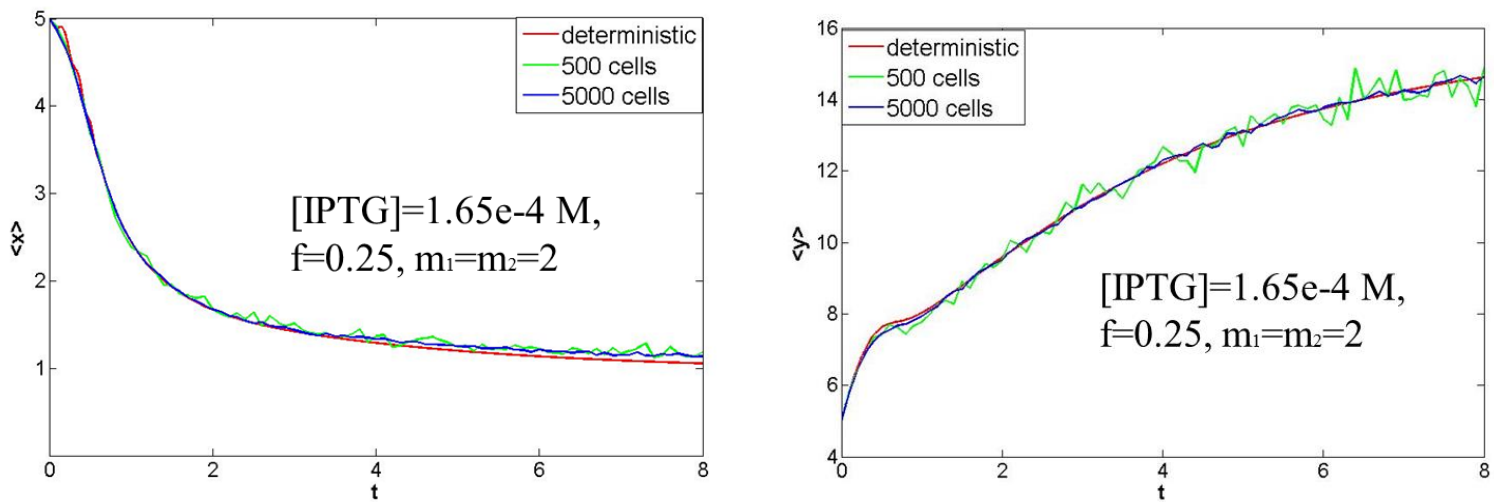
Η μέση συγκέντρωση των δύο παρεμποδιστών είναι το μακροσκοπικό εκείνο μέγεθος το οποίο εξάγεται από την προσομοίωση. Μελετήθηκε πώς ο παράγοντας ασυμμετρίας, f , καθώς και οι παράμετροι m_1 , m_2 επηρεάζουν τη χρονική εξέλιξη της μέσης συγκέντρωσης. Επίσης εξετάζεται πώς το εξωκυτταρικό περιβάλλον, δηλαδή η εξωκυτταρική συγκέντρωση του IPTG, επηρεάζει τη μέση έκφραση των παρεμποδιστών. Οι συγκεντρώσεις του [IPTG] επιλέχθηκαν σύμφωνα με τα αποτελέσματα που προέκυψαν από τη συστημική ανάλυση που έγινε με το ντετερμινιστικό μοντέλο, προκειμένου οι λύσεις να βρίσκονται σε ευσταθή κλάδο και όχι κοντά σε σημείο στροφής [46].



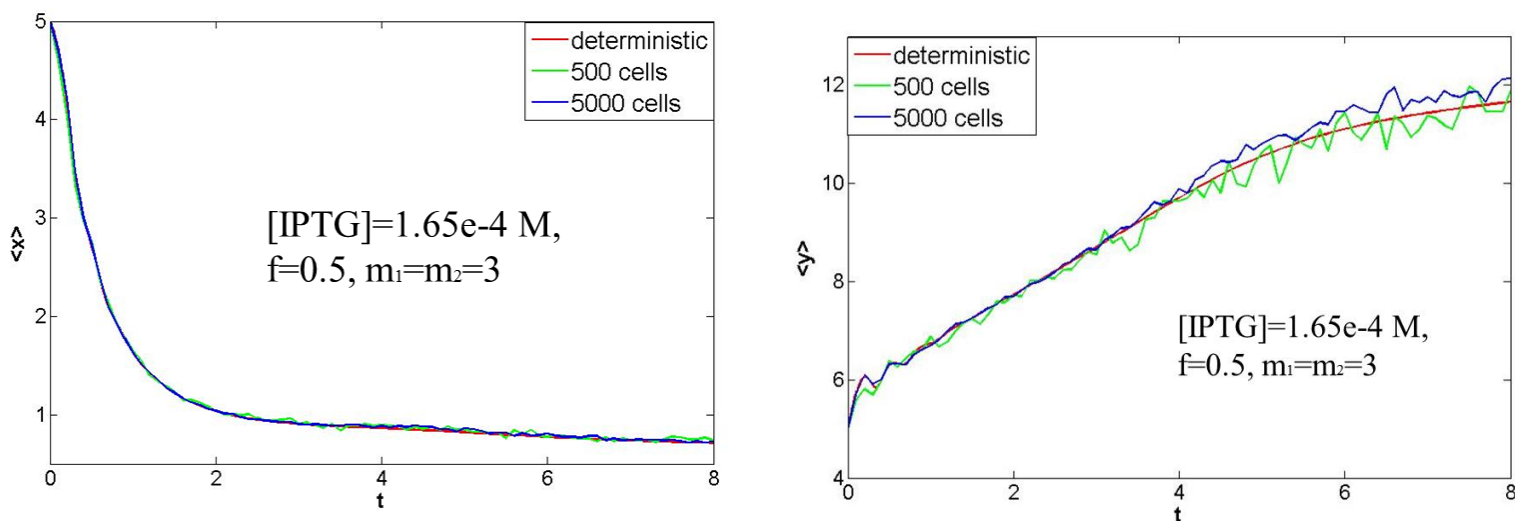
Σχήμα 5.1: Αποτελέσματα για διαφορετικούς κυτταρικούς πληθυσμούς με $[IPTG]=1e-6 M$, $f=0.5$, $m_1=m_2=1$.



Σχήμα 5.2: Αποτελέσματα για διαφορετικούς κυτταρικούς πληθυσμούς με $[IPTG]=1.65e-4 M$, $f=0.5$, $m_1=m_2=2$.



Σχήμα 5.3: Αποτελέσματα για διαφορετικούς κυτταρικούς πληθυσμούς με $[IPTG]=1.65e-4 M$, $f=0.25$, $m_1=m_2=2$.



Σχήμα 5.4: Αποτελέσματα για διαφορετικούς κυτταρικούς πληθυσμούς με $[IPTG]=1.65e-4 M$, $f=0.5$, $m_1=m_2=3$.

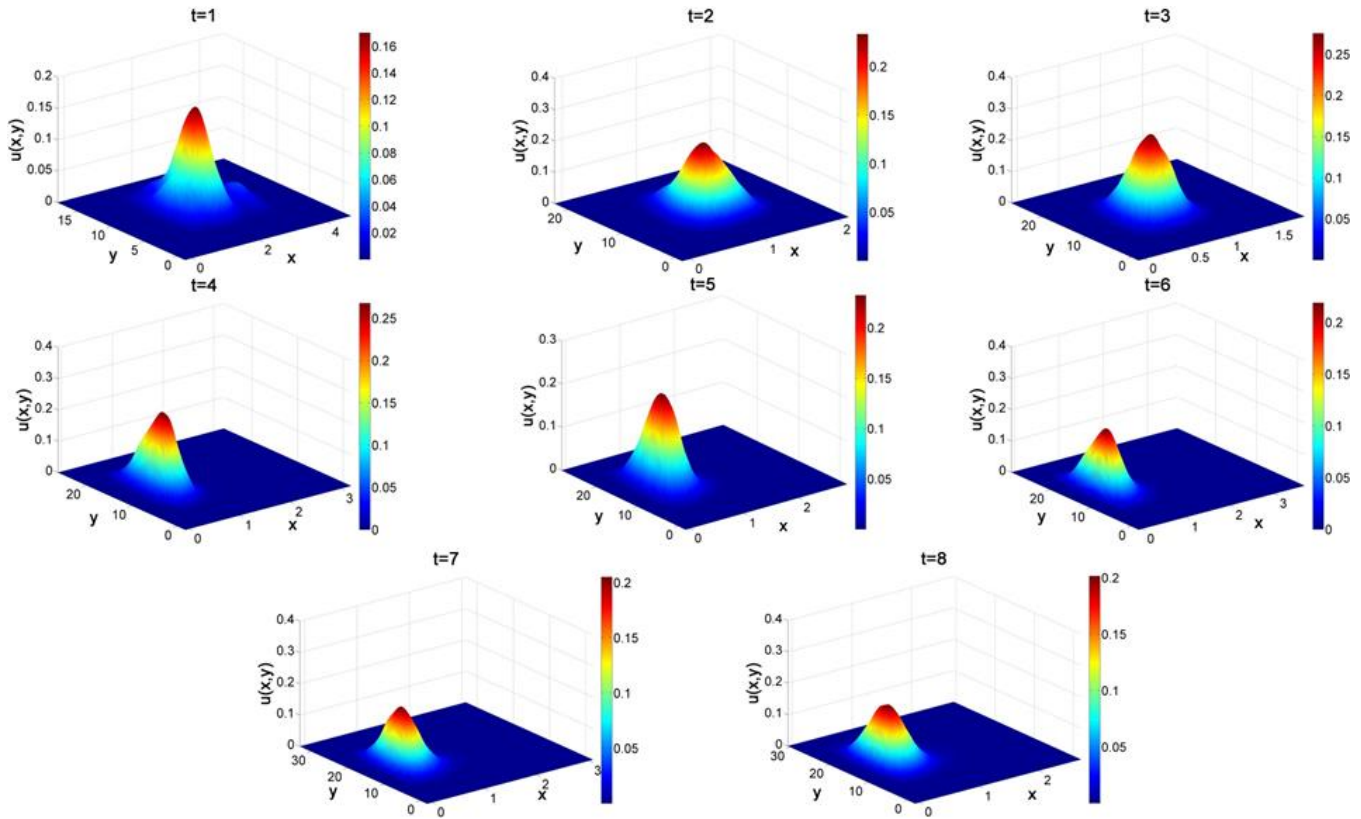
Όπως παρατηρείται από τα Σχήματα 5.1-5.4, το ντετερμινιστικό (κόκκινη γραμμή) με το στοχαστικό μοντέλο (πράσινη και μπλε γραμμή) δίνουν παραπλήσια χρονική εξέλιξη των μέσων συγκεντρώσεων για τους δύο παρεμποδιστές. Για μικρούς κυτταρικούς πληθυσμούς, στην προκείμενη περίπτωση για πεντακόσια (πράσινη γραμμή) και για εκατό κύτταρα (μαύρη γραμμή, βλ. Σχήματα 5.1-5.2), υπάρχει θόρυβος. Ο θόρυβος αυτός οφείλεται στη στοχαστικότητα που διέπει την κυτταρική διαίρεση. Με αύξηση του πλήθους των κυττάρων παρατηρείται σύγκλιση των δύο μοντέλων. Παρόλα αυτά, δεν εξαλείφεται πλήρως ο θόρυβος σε όλες τις περιπτώσεις, επί παραδείγματι στο Σχήμα 5.4 παρατηρείται ότι υπάρχει θόρυβος όσον αφορά τη μέση συγκέντρωση του παρεμποδιστή 2 ($\langle y \rangle$).

Επίσης, παρατηρείται ότι για διαφορετικές συγκεντρώσεις IPTG (βλ Σχήμα 5.1-5.2) προκύπτει διαφορετική έκφραση των δύο παρεμποδιστών. Συγκεκριμένα, για μικρές συγκεντρώσεις IPTG παρατηρείται υψηλή έκφραση του παρεμποδιστή 1, ενώ για υψηλές του παρεμποδιστή 2.

Τέλος, από τα παραπάνω αποτελέσματα εξάγεται το συμπέρασμα ότι οι παράμετροι f και m_1 , m_2 δεν έχουν σημαντική επίδραση στη χρονική εξέλιξη των μέσων συγκεντρώσεων.

4.3 Κατανομές κυτταρικού περιεχομένου

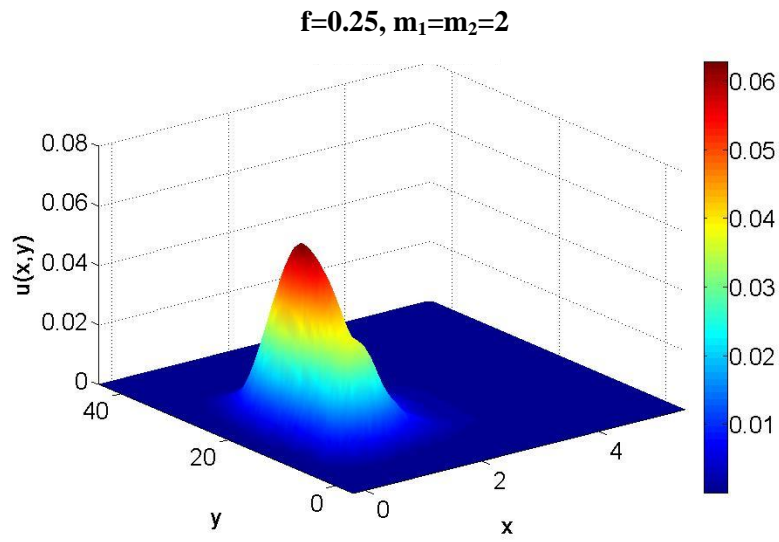
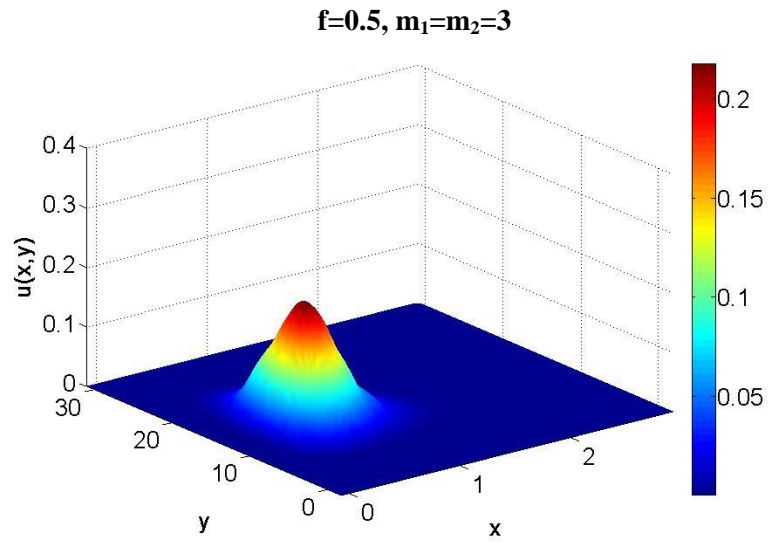
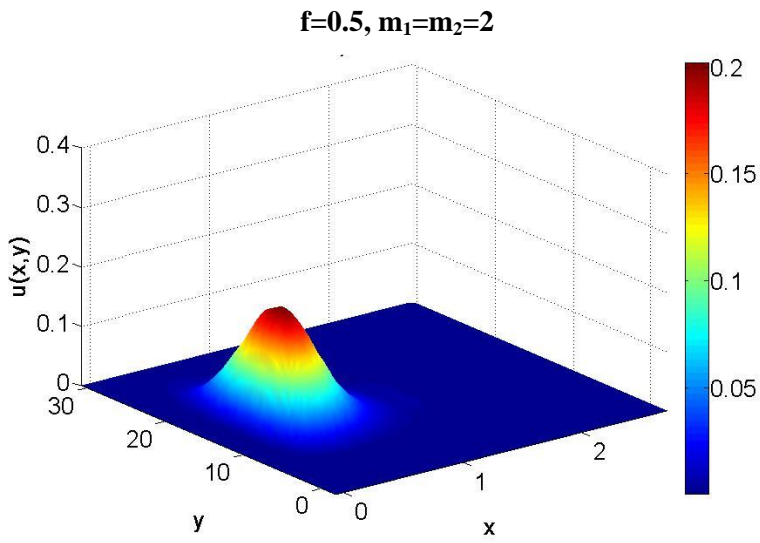
Στην παράγραφο αυτή απεικονίζονται η χρονική εξέλιξη των κατανομών, $u(x,y)$, για 5000 κύτταρα και για διάφορους αδιάστατους χρόνους. Ως παράμετροι χρησιμοποιήθηκαν $f=0.5$, $m_1=m_2=2$ και $[IPTG]=1.65e-4$.



Σχήμα 5.5: Χρονική εξέλιξη κατανομής για διάφορους αδιάστατους χρόνους με παραμέτρους $f=0.5$, $m_1=m_2=2$ και $[IPTG]=1.65e-4$ M.

Σχετικά με τη χρονική εξέλιξη των κατανομών, παρατηρείται ότι ξεκινώντας από μία μετατοπισμένη κατανομή Gauss για $t=0$, το σύστημα μεταπίπτει σε πιο απότομες κατανομές, ενώ για χρόνους $t > 7$, η κατανομή γίνεται ομαλότερη και σταθεροποιείται, κάτι που σημαίνει ότι το σύστημα έχει φτάσει σε μόνιμες συνθήκες.

Στη συνέχεια παρουσιάζονται οι κατανομές σε μόνιμες συνθήκες (για αδιάστατο χρόνο $t=15$) για διαφορετικές τιμές παραμέτρων f και m_1, m_2 .



Σχήμα 5.6: Κατανομές σε αδιάστατο χρόνο $t=15$ και $[IPTG]=1.65e-4$ M για διάφορες τιμές f και m_1, m_2 .

Παρατηρείται ότι για διαφορετικές τιμές m_1, m_2 η μορφή της κατανομής δεν επηρεάζεται ιδιαίτερα, πέρα από κάποια στένωση στην κατανομή όσο αυξάνονται οι τιμές αυτές. Σημαντική επίδραση, όμως, έχει ο παράγοντας ασυμμετρίας, f , κάτι που είναι λογικό αφού όσο μικρότερος είναι, οδηγεί ολοένα και σε πιο ασύμμετρη κατανομή του κυτταρικού περιεχομένου.

Το υπολογιστικό κόστος εξαρτάται τόσο από το μέγεθος του κυτταρικού πληθυσμού, όσο και από τον χρονικό ορίζοντα που ορίζεται στην προσομοίωση. Για παράδειγμα για 100 κύτταρα και $t_{final}=5$, χρειάζονται 0.09 δευτερόλεπτα, ενώ για 1000 κύτταρα 7.59 δευτερόλεπτα. Αντίστοιχα, για 1000 κύτταρα και $t_{final}=3$, χρειάζονται 4.8 δευτερόλεπτα.

Κεφάλαιο 5

Συμπεράσματα-Μελλοντική Έρευνα

Στο κεφάλαιο αυτό συνοψίζονται τα συμπεράσματα που προκύπτουν από την ανάλυση των αποτελεσμάτων. Επίσης, παρατίθενται προτάσεις για μελλοντική έρευνα.

5.1 Συμπεράσματα

Στην εργασία αυτή αναπτύχθηκε ένα μοντέλο περιγραφής κυτταρικών πληθυσμών που φέρουν το γενετικό δίκτυο toggle switch. Βασικός στόχος ήταν η μελέτη της επίδρασης της ετερογένειας που οφείλεται στη στοχαστικότητα που διέπει την κυτταρική διαίρεση, ενώ έγινε σύγκριση με το ντετερμινιστικό μοντέλο το οποίο βασίζεται στα Ισοζύγια Κυτταρικών Πληθυσμών. Επίσης, μελετήθηκε πώς παράμετροι, όπως το εξωκυτταρικό περιβάλλον, επιδρά στη έκφραση των γονιδίων. Για τη προσομοίωση του συστήματος χρησιμοποιήθηκε στοχαστικός αλγόριθμος τύπου kinetic Monte Carlo. Χρησιμοποιήθηκε ο αλγόριθμος Constant Number Monte Carlo.

Η πρώτη ιδιότητα που μελετήθηκε ήταν η χρονική εξέλιξη των μέσων συγκεντρώσεων των δύο παρεμποδιστών για διαφορετικούς πληθυσμούς και για διαφορετικές παραμέτρους, όπως είναι η εξωκυτταρική συγκέντρωση του IPTG, ο παράγοντας ασυμμετρίας, f , και οι παράμετροι m_1 , m_2 . Επίσης, έγινε σύγκριση των αποτελεσμάτων του στοχαστικού μοντέλου με το ντετερμινιστικό. Η μέση συγκέντρωση, όπως παρατηρήθηκε, επηρεάζεται από την εξωκυτταρική συγκέντρωση του IPTG, ενώ οι υπόλοιπες παράμετροι έχουν αμελητέα συνεισφορά. Συγκεκριμένα, για πολύ μικρές τιμές της συγκέντρωσης του IPTG υπάρχει υψηλή έκφραση του παρεμποδιστή 1 και χαμηλή του παρεμποδιστή 2. Το αντίθετο συμβαίνει για υψηλές συγκεντρώσεις. Συνεπώς, το εξωκυτταρικό περιβάλλον έχει βαρύνουσα σημασία για την έκφραση των γονιδίων.

Επιπλέον, για μικρούς κυτταρικούς πληθυσμούς παρατηρήθηκε η ύπαρξη «θορύβου», που οφείλεται στα τυχαία φαινόμενα και συγκεκριμένα στην κυτταρική διαίρεση. Το στοχαστικό μοντέλο αποτελεί ένα ακριβέστερο μοντέλο για μικρούς κυτταρικούς πληθυσμούς, σε σύγκριση με το ντετερμινιστικό, καθώς μπορεί να λάβει υπόψη του τη στοχαστικότητα της κυτταρικής διαίρεσης. Για μεγάλους κυτταρικούς πληθυσμούς τα δύο μοντέλα δίνουν παραπλήσια

αποτελέσματα. Αυτό είναι αναμενόμενο από μαθηματικής σκοπιάς, δεδομένου ότι το ΙΚΠ που περιγράφει το παρόν σύστημα αποτελεί έναν μέσο όρο της master equation, η οποία προσεγγίζεται από το στοχαστικό μοντέλο.

Στη συνέχεια μελετήθηκε η χρονική εξέλιξη των κατανομών για πληθυσμό 5000 κυττάρων και για 5 αντίγραφα. Ως παράμετροι χρησιμοποιήθηκαν $f=0.5$, $m_1=m_2=2$ και $[IPTG]=0.000165$ M. Παρατηρήθηκε ότι από μια αρχική κατανομή Gauss, το σύστημα προχωράει σε πιο απότομες κατανομές, μέχρι να σταθεροποιηθεί για $t>7$.

Τέλος, μελετήθηκε η επίδραση του παράγοντα ασυμμετρίας και των m_1 , m_2 . Παρατηρήθηκε ότι μεγαλύτερη επίδραση στην κατανομή έχει ο παράγοντας ασυμμετρίας, f , και αυτό διότι όσο μικρότερος είναι, τόσο πιο ανισομερής είναι η κατανομή του κυτταρικού περιεχομένου κατά την κυτταρική διαίρεση, άρα η κατανομή που προκύπτει είναι ολόενα και πιο ασύμμετρη.

5.2 Μελλοντική έρευνα

Οι προσομοιώσεις με τον αλγόριθμο CNMC, αν και ακριβείς, είναι δαπανηρές από άποψη υπολογιστικού χρόνου, καθώς όταν αυξάνεται ο κυτταρικός πληθυσμός κατά μία τάξη μεγέθους, για τον ίδιο χρονικό ορίζοντα, το υπολογιστικό κόστος εκατονταπλασιάζεται. Αντίστοιχα, όσο μεγαλύτερος είναι ο χρονικός ορίζοντας, τόσο πιο χρονοβόρα είναι η προσομοίωση. Συνεπώς, είναι πρακτικά αδύνατο να γίνει συστημική ανάλυση, έτσι ώστε να βρεθεί ολόκληρος ο χώρος λύσεων του κυτταρικού πληθυσμού, κάτι που είναι ιδιαίτερης σημασίας στο συγκεκριμένο σύστημα, καθώς παρουσιάζει το φαινόμενο της διπλοευστάθειας λύσεων, όπου για μία τιμή της εξωκυτταρικής συγκέντρωσης του IPTG μπορεί να υπάρχουν δύο ευσταθείς λύσεις.

Για το λόγο αυτό θα μπορούσε να χρησιμοποιηθεί η «ελευθέρα εξίσωση» [45,46] μεθοδολογία, η οποία χωρίς τη χρήση κάποιας εξίσωσης που να περιγράφει μακροσκοπικά το σύστημα, εξάγει αποτελέσματα για την μακροσκοπική συμπεριφορά του κυτταρικού πληθυσμού, σε ρεαλιστικό υπολογιστικό χρόνο, δίνοντας έτσι τη δυνατότητα να γίνει παραμετρική ανάλυση.

Παράρτημα

Η επιλογή ενός τυχαίου ακέραιου αριθμού i όπως αναφέρθηκε στο Κεφάλαιο 3, γίνεται μέσω της σχέσης (3.25). Για την απόδειξή της αρχικά θεωρούμε την συνάρτηση πυκνότητας πιθανότητας $P(i)$, η οποία δηλώνει την πιθανότητα ο τυχαίος ακέραιος αριθμός που θα επιλεγεί να είναι ο i . Η αθροιστική συνάρτηση πυκνότητας πιθανότητας είναι :

$$F(i) = \sum_{-\infty}^i P(i') \quad (\text{A1})$$

Η πιθανότητα $P(i')$ μπορεί να γραφεί και ως εξής :

$$F(i) - F(i - 1) = \sum_{-\infty}^i P(i') - \sum_{-\infty}^{i-1} P(i') = P(i') \quad (\text{A2})$$

Άρα, η ισοδύναμη σχέση για την επιλογή ενός τυχαίου ακέραιου αριθμού, i , βάσει ενός τυχαίου αριθμού από την ομοιόμορφη κατανομή, r , είναι :

$$F(i - 1) < r \leq F(i) \quad (\text{A3})$$

Στην παρούσα περίπτωση η (A3) αντιστοιχεί στη σχέση (3.25) και η $P(i)$ είναι :

$$P(i) = \frac{\Gamma(x_i(\tau), y_i(\tau))}{\sum_{i=1}^N \Gamma(x_i(\tau), y_i(\tau))} \quad (\text{A4})$$

Βιβλιογραφία

1. N. V. Mantzaris, "Stochastic and deterministic simulations of heterogenous cell population dynamics," *Journal of Theoretical Biology*, vol. 241, pp. 690-706, 2006.
2. M. Delbrück, "The burst size distribution in the growth of bacterial viruses (bacteriophages)," *Journal of Bacteriology*, vol. 50, pp. 131-135, 1945.
3. J. D. Chung and G. Stephanopoulos, "Studies of transcriptional state heterogeneity in sporulating cultures of *Bacillus subtilis*," *Biotechnology and Bioengineering*, vol. 47, pp. 234-242, 1995.
4. M. B. Elowitz, A. J. Levine, E. D. Siggia and P. S. Swain, "Stochastic gene expression in a single cell," *Science*, vol. 297, pp. 1183-1186, 2002.
5. F. Delvigne and P. Coffin, "Microbial heterogeneity affects bioprocess robustness: Dynamical single-cell analysis contributes to understanding of microbial population," *Biotechnology Journal*, vol. 9, pp. 67-72, 2013.
6. S. Alonso, M. Rendueles and M. Diaz, "Physiological Heterogeneity of *Pseudomonas taetrolens* during Lactobionic Acid Production," *Applied Microbiology and Biotechnology*, vol. 96, pp. 1465-1477, 2012.
7. B. R. Boles, M. Thoendel and P. K. Singh, "Self-generated diversity produces 'insurance effects' in biofilm communities," *Proceedings National Academy of Sciences of the United States of America*, vol. 101, pp. 16630-16635, 2004.
8. C. Goerke, M. Gressinger, K. Endler, C. Breitkopf, K. Wardecki, M. Stern, C. Wolz and B. C. Kahl, "High phenotypic diversity in infecting but not in colonizing *Staphylococcus aureus* populations," *Environmental Microbiology*, vol. 9, pp. 3134-3142, 2007.

9. B. R. Boles and P. K. Singh, "Endogenous oxidative stress produces diversity and adaptability in biofilm communities," *Proceedings National Academy of Sciences of the United States of America*, vol. 105, pp. 12503-12508, 2008.
10. J. F. Gainor and A. T. Shaw, "Emerging paradigms in the development of resistance to tyrosine kinase inhibitors in lung cancer," *Journal of Clinical Oncology*, vol. 31, p. 3987–3996, 2013.
11. B. Marte, "Tumour heterogeneity," *Nature*, vol. 501, p. 327, 2013.
12. R. A. Burrell, N. McGranahan, J. Bartek and C. Swanton, "The causes and consequences of genetic heterogeneity in cancer evolution," *Nature*, vol. 501, p. 338–345, 2013.
13. R. Govindan, L. Ding, M. Griffith and et al, "Genomic landscape of non-small cell lung cancer in smokers and never-smokers," *Cell*, vol. 150, p. 1121–1134, 2012.
14. A. Alberts, D. Bray, J. Lewis, M. Raff, K. Robers and J. D. Watson, *Molecular Biology of the Cell*. third edition, New York: Garland Publishing, 1994.
15. Μ. Καβουσανάκης, Αδρομερής Προσομοίωση Περίπλοκων Φυσικοχημικών και Βιολογικών Συστημάτων με Μεθόδους Ανάλυσης Πολλαπλών Κλιμάκων, Διακλάδωσης και Ομοιοτυπίας. Διδακτορική Διατριβή, Αθήνα: ΕΜΠ, 2007.
16. T. S. Gardner, C. R. Cantor and J. J. Collins, "Construction of a genetic toggle switch in *Escherichia coli*," *Nature*, 2000, vol. 403, p. 339-342, 2000.
17. K. Spetsieris, K. Zygourakis and N. V. Mantzaris, "A novel assay based on fluorescence microscopy and image processing for determining phenotypic distributions of rod-shaped bacteria," *Biotechnology and Bioengineering*, vol. 102, pp. 598-615, 2009.

18. T. Eissing, H. Conzelmann, E. D. Gilles, F. Allgoewer, E. Bullinger and P. Scheurich, "Bistability analyses of a caspase activation model for receptor-induced apoptosis," *The Journal of Biological Chemistry*, vol. 279, pp. 36892-36897, 2004.
19. D. Kim, O. Rath, W. Kolch and K. H. Cho, "A hidden oncogenic positive feedback loop caused by crosstalk between Wnt and ERK pathways," *Oncogene*, vol. 26, pp. 4571-4579, 2007.
20. N. Kellershohn and M. Laurent, "Prion Diseases: Dynamics of the Infection and Properties of the Bistable Transition," *Biophysical Journal*, vol. 81, pp. 2517-2529, 2001.
21. J. W. Veening, W. K. Smiths and O. P. Kuipers, "Bistability, epigenetics, and bet-hedging in bacteria," *Annual Review of Microbiology*, vol. 62, pp. 193-210., 2008.
22. N. V. Mantzaris, "From single-cell genetic architecture to cell population dynamics: quantitatively decomposing the effects of different population heterogeneity sources for a genetic network with positive feedback architecture," *Biophysical Journal*, vol. 92, p. 4271–4288, 2007.
23. T. Tian and K. Burrage, "Bistability and switching in the lysis/lysogeny genetic regulatory network of bacteriophage λ ," *Journal of Theoretical Biology*, vol. 227, p. 229–237, 2004.
24. E. M. Judd, M. T. Laub and H. H. McAdams, "Toggles and oscillators: new genetic circuit designs," *Bioessays*, vol. 22, pp. 507-509, 2000.
25. A. Arkin, J. Ross, and H. McAdams, "Stochastic kinetic analysis of developmental pathway bifurcation in phage λ -infected escherichia coli cells," *Genetics*, vol. 149, p. 1633–1648, 1998.
26. M. Khammash and H. El Samad, "Stochastic modeling and analysis of genetic networks," in *Proc. 44th IEEE Conf. Decision and Control*, Seville, Spain, 2005.
27. R. Thomas, "Boolean formalization of genetic control circuits," *Journal of Theoretical Biology*, vol. 42, pp. 563-585, 1973.

28. R. Albert and H. Othmer, "The topology of the regulatory interactions predicts the expression pattern of the drosophila segment polarity genes," *Journal of Theoretical Biology*, vol. 223, pp. 1-18, 2003.
29. J. M. Eakman, A. G. Fredrickson and H. M Tsuchiya, "Statistics and Dynamics of Microbial Cell Populations.," *Chemical Engineering Progress*, vol. 62, pp. 37-49, 1966.
30. H. M Tsuchiya, A. G. Fredrickson and R. Aris, "Dynamics of Microbial Cell Populations," *Advances in Chemical Engineering*, vol. 6, pp. 125-206, 1966.
31. A. G. Fredrickson, D. Ramkrishna and H. M Tsuchiya, "Statistics and Dynamics of Prokaryotic Cell Populations," *Mathematical Biosciences*, vol. 1, pp. 327-374, 1967.
32. J. J. Liou, F. Sreenc and A. G. Fredrickson, "Solutions of population balance models based on a successive generations approach," *Chemical Engineering Science*, vol. 52, p. 1529–1540, 1997.
33. B. S. Dien, Aspects of Cell Division Cycle Related Behavior of *Saccharomyces Cerevisiae* Growing in Batch and Continuous Culture: A Single-Cell Growth Analysis. Ph.D. Thesis, University of Minnesota, 1994.
34. N. V. Mantzars, "Transient and asymptotic behaviour of the binary," *Journal of Physics A: Mathematical and General*, vol. 38, p. 5111–5132, 2005.
35. M. E. Kavousanakis, N. V. Mantzaris and A. G Boudouvis, "A novel free boundary algorithm for the solution of cell population balance model," *Chemical Engineering Science*, vol. 64, pp. 4247-4261, 2009.
36. J. R. Beckwith and D. Zisper, *The Lactose Operon*, New York: Cold Spring Harbor, ed. C.S.H. Laboratory, 1970.
37. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth., A. H. Teller and E. Teller "Equation of state calculations by fast computing machines," *The Journal of Chemical Physics*, vol. 21, p. 1087, 1953.

38. D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *The Journal of Physical Chemistry*, vol. 81, pp. 2340-2361, 1977.
39. N. V. Mantzaris, "Stochastic and deterministic simulations of heterogeneous cell," *Journal of Theoretical Biology*, vol. 241, pp. 690-706, 2006.
40. N. V. Mantzaris, "From Single Cell Genetic Architecture to Cell Population Dynamics: Quantitatively Decomposing the Effects of Different Population Heterogeneity Sources for a Genetic Network with Positive Feedback Architecture," *Biophysical Journal*, vol. 92, pp. 4271-4288, 2007
41. M. Smith and T. Matsoukas, "Constant-number Monte Carlo simulation of population balances," *Chemical Engineering Science*, Vols. 1777-1786, p. 53, 1998.
42. Y. Lin, K. Lee and T. Masoukas, "Solution of the population balance equation using constant-number Monte Carlo," *Chemical Engineering Science*, vol. 57, pp. 2241-2252, 2002.
43. D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *Journal of Computational Physics*, vol. 22, pp. 403-434, 1976.
44. Π. Χρυσίνας, Αριθμητική επίλυση πολυμεταβλητών Ισοζυγίων Κυτταρικών Πληθυσμών: Διερεύνηση μη γραμμικών φαινομένων σε ετερογενείς πληθυσμούς E.coli με ρυθμιστικό δίκτυο τύπου toggle switch. Μεταπτυχιακή Διατριβή, Αθήνα: ΕΜΠ, 2015.
46. I. G. Gillespie, C. W. Gear, J. Hyman, P. Kevrekidis, O. Runborg and K. Theodoropoulos, "Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis". *Communications in Mathematical Science*, vol. 1, pp. 715-762, 2003.

47. I. G. Gillespie, C. W. Gear and G. Hummer, "Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis". *AICHE Journal*, vol. 50, pp. 1346-1354, 2004.

