



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΥΠΟΛΟΓΙΣΤΙΚΗ ΜΗΧΑΝΙΚΗ»

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
του **ΙΩΑΝΝΗ Γ. ΑΒΙΖΙΩΤΗ**

**ΑΔΡΟΜΕΡΕΙΣ ΥΠΟΛΟΓΙΣΜΟΙ
ΙΣΟΖΥΓΙΩΝ ΚΥΤΤΑΡΙΚΩΝ ΠΛΗΘΥΣΜΩΝ**

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ
ΑΝΔΡΕΑΣ Γ. ΜΠΟΥΝΤΟΥΒΗΣ

ΑΘΗΝΑ 2013

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον Καθηγητή Ανδρέα Μπουντουβή, επιβλέποντα της μεταπτυχιακής αυτής εργασίας, για τη δυνατότητα που μου έδωσε να εργαστώ πάνω σε ένα τόσο ενδιαφέρον θέμα και για την επιστημονική του καθοδήγηση σε όλη τη διάρκεια της εργασίας επί του θέματος αυτού. Με τις παρατηρήσεις, συμβουλές και οδηγίες του, πάντα επίκαιρες και εύστοχες, κατέστη δυνατή η επιτυχής ολοκλήρωση της εργασίας. Τον ευχαριστώ, επίσης, θερμά γιατί η υποστήριξη και η συμπαράσταση που μου προσέφερε, διαμόρφωσαν ένα ιδανικό πλαίσιο εργασίας και έρευνας.

Επίσης, θα ήθελα να ευχαριστήσω τα μέλη της ερευνητικής ομάδας μας και ιδιαίτερος τον Δρ. Αντώνη Σπυρόπουλο για τη βοήθειά του σε θέματα υπολογισμών που διεξήχθησαν στα «clusters» που διαθέτει το Υπολογιστικό Κέντρο της σχολής Χημικών Μηχανικών και τον υποψήφιο Διδάκτορα Νίκο Χαμάκο για τις συζητήσεις που είχαμε.

Όμως, πάνω απ'όλα ένα μεγάλο και θερμό ευχαριστώ στον Δρ. Μιχάλη Καβουσανάκη για την υπομονή και την επιμονή που επέδειξε και για την καθοδήγησή του καθ'όλη τη διάρκεια της μεταπτυχιακής εργασίας. Με την εμπειρία και τις γνώσεις του προσπάθησε να με μυήσει στο θέμα των αδρομερών υπολογισμών και να μου μεταλαμπαδεύσει αυτές τις γνώσεις. Μιχάλη σ' ευχαριστώ για όλα.

Περίληψη

Βασικός στόχος της παρούσας εργασίας είναι η ανάπτυξη μιας υπολογιστικής μεθοδολογίας που επιτρέπει την αποτελεσματική μελέτη της μακροσκοπικής συμπεριφοράς συστημάτων, η περιγραφή των οποίων είναι διαθέσιμη σε επίπεδο μικρο-κλίμακας, χωρίς την ανάγκη εξαγωγής ενός μοντέλου διαφορικών εξισώσεων που να περιγράφει τη δυναμική συμπεριφορά σε επίπεδο μακρο-κλίμακας.

Αντιμετώπιστηκαν προβλήματα ισοζυγίων κυτταρικών πληθυσμών, στα οποία διερευνάται η επίδραση της κυτταρικής ετερογένειας στο φαινότυπο κυττάρων που φέρουν το ίδιο ρυθμιστικό γενετικό δίκτυο. Η μελέτη αυτή αναδεικνύει ένα δρόμο για τη γεφύρωση του χάσματος μεταξύ κλιμάκων που διαφέρουν κατά τάξεις μεγέθους (κύτταρο – φαινότυπος κυτταρικού πληθυσμού).

Στο πρώτο στάδιο της εργασίας κατασκευάστηκε ένας στοχαστικός αλγόριθμος kinetic Monte Carlo ο οποίος περιγράφει τη συμπεριφορά κυτταρικών πληθυσμών. Πραγματοποιήθηκαν προσομοιώσεις με τη χρήση του αλγορίθμου, οι οποίες οδήγησαν στη χρονική εξέλιξη του εξεταζόμενου συστήματος μέχρι αυτό να φτάσει στη μόνιμη κατάσταση.

Στη συνέχεια, αναπτύχθηκε υπολογιστικός κώδικας για την αδρομερή προσομοίωση του συστήματος. Κατά τους αδρομερείς υπολογισμούς ενδιαφερόμαστε για τα στατιστικά μεγέθη του πληθυσμού (μέση τιμή, τυπική απόκλιση-ροπές χαμηλότερης τάξης), τα οποία αποτελούν τη μακροσκοπική πληροφορία που εξάγεται από τις στοχαστικές προσομοιώσεις που έχουν διεξαχθεί σε επίπεδο μικρο-κλίμακας. Οι αδρομερείς υπολογισμοί απαιτούν σημαντικά λιγότερο υπολογιστικό κόστος από την ευθεία εφαρμογή του στοχαστικού αλγορίθμου. Μέσω αυτών εξετάζεται το κατά πόσο οι προκύπτουσες κατανομές είναι ακριβείς στην περιγραφή της συμπεριφοράς του πληθυσμού.

Μετά την ανάπτυξη του αδρομερούς προσομοιωτή, πραγματοποιήθηκε παραμετρική ανάλυση, με σκοπό να διαπιστωθεί αν για κάποιες τιμές παραμέτρων ο φαινότυπος του κυτταρικού πληθυσμού εμφανίζει πολλαπλότητα λύσεων (συγκεκριμένα διπλο-ευστάθεια). Σε πρώτη φάση έχει επιτευχθεί μόνο ο υπολογισμός ευσταθών καταστάσεων, με τη χρήση της επαναληπτικής μεθόδου Newton-Raphson που διαμορφώνεται κατάλληλα για να ενσωματώσει την αδρομερή περιγραφή του συστήματος. Για τον υπολογισμό καταστάσεων που ανήκουν στον ασταθή κλάδο είναι απαραίτητη η μελλοντική εφαρμογή της μεθόδου μήκους-τόξου (arc-length

continuation), η οποία, επίσης, πρέπει να διαμορφωθεί λαμβάνοντας υπόψη τις αδρομερείς προσομοιώσεις.

POSTGRADUATE DIPLOMA THESIS

By I.G. AVIZIOTIS

Coarse-grained computations for cell population balances

Abstract

The main objective of this thesis was the development of a computational methodology that allows the efficient analysis of the macroscopic behavior of systems, the description of which is available only at the micro-scale level, without the need to extract a set of differential equations describing the dynamic behavior at the macro-scale level.

For this purpose, cell populations were considered, where the effect of the cellular heterogeneity in cell phenotype carrying the same regulatory network, was investigated. The analysis of cell population balances reveals a path for bridging the gap between scales, varying over few orders of magnitude (cell-phenotype of the cell population).

At first, a kinetic Monte Carlo algorithm was developed that describes the behavior of the cell population. Simulations using this algorithm were performed to observe the time evolution of the investigated system until it reaches steady state. Then, an algorithm for coarse-grained simulations was built. In these simulations we are interested for the statistics of the cell population (mean value, standard deviation-generally for low-order moments), which are the macroscopic information extracted from the micro-scale through stochastic simulations. Coarse-grained computations require significantly less computer time compared to the direct implementation of the detailed algorithm and result to distributions which are accurate in describing the dynamics of the cell population.

In the last part of this thesis, a bifurcation analysis was attempted with the purpose to determine the existence of bistability regions in the phenotype of the cell population, for a range of parameter values. In this work, we were able to compute only stable branches by means of Newton-Raphson iteration appropriately designed as to incorporate the coarse-grained description of the system. Along the same lines, an arc-length continuation algorithm is necessary to enable the calculation of phenotypes on unstable branches.

Περιεχόμενα

Ευχαριστίες.....	1
Περίληψη.....	2
Abstract.....	4
Κεφάλαιο 1: Το φυσικό πρόβλημα.....	7
1.1. Κυτταρική ετερογένεια.....	7
1.2. Ισοζύγια κυτταρικών πληθυσμών.....	9
1.3. Γενετική αρχιτεκτονική με θετική ανάδραση.....	11
Κεφάλαιο 2: Το στοχαστικό μοντέλο.....	17
2.1. Εισαγωγή.....	17
2.2. Υπολογισμός του χρόνου που μεσολαβεί μεταξύ συμβάντων διαίρεσης.....	19
2.3. Καθορισμός του μητρικού κυττάρου που διαιρείται.....	22
2.4. Καθορισμός του ενδοκυτταρικού περιεχομένου των δύο θυγατρικών κυττάρων.....	23
2.5. «Αποκατάσταση» του κυτταρικού πληθυσμού.....	23
Κεφάλαιο 3: Αδρομερείς υπολογισμοί δυναμικών συστημάτων.....	26
3.1. Εισαγωγή.....	26
3.2. Αδρομερής χρονικός ολοκληρωτής.....	27
3.3. Άρση με τη μέθοδο της μέγιστης εντροπίας (lifting with the maximum entropy method).....	30
3.4. Υπολογισμοί αδρομερών μόνιμων καταστάσεων.....	33
Κεφάλαιο 4: Αποτελέσματα υπολογιστικής ανάλυσης.....	38
4.1. Εισαγωγή.....	38

4.2. Αποτελέσματα του στοχαστικού μοντέλου.....	38
4.3. Αποτελέσματα της άρσης με τη μέθοδο της μέγιστης εντροπίας.....	45
4.4. Επιλογή του κατάλληλου αριθμού ροπών.....	47
4.5. Παραμετρική ανάλυση.....	51
Κεφάλαιο 5: Συμπεράσματα.....	53
Βιβλιογραφία.....	56

Κεφάλαιο 1: Το φυσικό πρόβλημα

1.1. Κυτταρική ετερογένεια

Οι πρόσφατες εξελίξεις στη μοριακή βιολογία, στα genomics, τα transcriptomics και τα proteomics, μας παρέχουν ισχυρά εργαλεία για τη μελέτη διεργασιών τόσο σε μοριακό όσο και σε κυτταρικό επίπεδο. Τεχνικές συστημικής βιολογίας έχουν εφαρμοστεί για την κατανόηση πολύπλοκων διαδικασιών στο επίπεδο του κυττάρου [16]. Η φαινοτυπική συμπεριφορά ενός κυτταρικού πληθυσμού ωστόσο δεν είναι αποτέλεσμα μόνο των αλληλεπιδράσεων μεταξύ των πολυάριθμων συστατικών εντός των μεμονωμένων κυττάρων, αλλά εξαρτάται ισχυρά και από την πολυπλοκότητα, τις άμεσες και τις έμμεσες αλληλεπιδράσεις ανάμεσα στα κύτταρα ενός πληθυσμού. Οι διακυτταρικές και οι ενδοκυτταρικές αλληλεπιδράσεις οδηγούν σε σημαντικές διακυμάνσεις του φαινοτύπου από κύτταρο σε κύτταρο και το βιολογικό αυτό φαινόμενο χαρακτηρίζεται ως ετερογένεια κυτταρικών πληθυσμών. Τέτοια φαινόμενα έχουν παρατηρηθεί σε πολλά συστήματα όπως στα μεγέθη των ιών phage, σε καλλιέργειες αποτελούμενες από σπόρια του *Bacillus subtilis* που περιέχουν συγχωνεύσεις μεταξύ γονιδίων παραγωγής σπόρων και της *lacZ* και σε ποικίλα ισογονικά συστήματα της *E Coli* [16].

Η πληθώρα των βιοτεχνολογικών εφαρμογών στοχεύει στη μεγιστοποίηση της παραγωγής ουσιών που παράγονται από ένα πληθυσμό κυττάρων. Επιπροσθέτως, κύριος στόχος των περισσότερων ιατρικών εφαρμογών και των προσεγγίσεων που εξετάζουν παθολογικές συνθήκες είναι η μεταχείριση του πληθυσμού των κυττάρων ως σύνολο και όχι η αντιμετώπιση κάθε κυττάρου ξεχωριστά. Ακόμα και οι πειραματικές τεχνικές, οι οποίες είναι διαθέσιμες σήμερα (π.χ. ακολουθίες DNA, υγρή χρωματογραφία, φασματομετρία μάζας), συλλέγουν δεδομένα από ολόκληρο τον πληθυσμό και όχι από κάθε κύτταρο ξεχωριστά. Η ετερογένεια των κυτταρικών πληθυσμών μπορεί να οδηγήσει σε παρεμπόδιση της αποδοτικής παραγωγής βιοτεχνολογικών προϊόντων και σε μειωμένη ικανότητα προσαρμογής τους σε απότομες περιβαλλοντικές αλλαγές. Συνεπώς, η κατανόηση της δυναμικής σχέσης μεταξύ των φαινομένων που λαμβάνουν χώρα στο επίπεδο ενός κυττάρου και στο επίπεδο ολόκληρου του κυτταρικού πληθυσμού μπορεί να δώσει σημαντικές πληροφορίες για την ανίχνευση, την πρόγνωση αλλά και τη θεραπεία διάφορων ασθενειών όπως του καρκίνου.

Η παρατηρούμενη ετερογένεια ενός ισογονιδιακού πληθυσμού κυττάρων, δηλαδή ενός πληθυσμού που χαρακτηρίζεται από πανομοιότυπη γενετική σύνθεση, σε ένα ομοιόμορφο εξωκυτταρικό περιβάλλον προέρχεται από δύο διαφορετικές βασικές πηγές. Το σύνολο των ενδοκυτταρικών συστατικών του μητρικού κυττάρου, με εξαίρεση το DNA, κληροδοτείται ανισομερώς στα θυγατρικά κύτταρα. Οι διακυμάνσεις στο περιεχόμενο, μόρια ή moles μιας δεδομένης ουσίας όπως RNA ή πρωτεΐνες, των θυγατρικών κυττάρων και κυρίως στον αριθμό των ρυθμιστικών μορίων (regulatory molecules) οδηγούν σε διαφορετικούς φαινοτύπους. Λόγω της λειτουργίας του κυτταρικού κύκλου, αυτό το φαινόμενο ανατροφοδοτείται οδηγώντας έτσι σε περαιτέρω διακυμάνσεις. Η ετερογένεια που προκύπτει από αυτή την πηγή χαρακτηρίζεται ως εξωγενής (extrinsic). Τα ρυθμιστικά μόρια, που καθορίζουν σε μεγάλο βαθμό τον κυτταρικό φαινότυπο, βρίσκονται σε μικρές συγκεντρώσεις. Έτσι, οι ρυθμοί των ενδοκυτταρικών αντιδράσεων, οι οποίες ρυθμίζονται από αυτά τα μόρια, χαρακτηρίζονται από τυχαίες διακυμάνσεις. Γι' αυτό το λόγο σε καθορισμένες χρονικές στιγμές ακόμα και κύτταρα με τον ίδιο αριθμό ρυθμιστικών μορίων μπορεί να συμπεριφέρονται με τελείως διαφορετικό τρόπο. Η ετερογένεια που προέρχεται από τέτοιου είδους στοχαστικά ενδοκυτταρικά γεγονότα χαρακτηρίζεται ως ενδογενής (intrinsic). Σημειώνεται ότι κάθε κυτταρικό συστατικό που δρα στοχαστικά οδηγεί σε ποικίλες κυτταρικές καταστάσεις με αποτέλεσμα να αποτελεί πηγή εξωγενούς ετερογένειας. Έτσι, οι δύο τύποι ετερογένειας είναι συζευγμένες.

Έχει διαπιστωθεί ότι η ετερογένεια των κυτταρικών πληθυσμών εξαρτάται ισχυρά από τις λειτουργικές ιδιότητες των ρυθμιστικών δικτύων (gene regulatory networks) [27]. Τα ρυθμιστικά δίκτυα είναι ένα σύνολο αλληλουχιών DNA σε ένα κύτταρο, οι οποίες αλληλεπιδρούν μεταξύ τους (μέσω του περιεχόμενου RNA και των προϊόντων των εκφράσεων των πρωτεϊνών) αλλά και με άλλες ουσίες στο κύτταρο. Συνεπώς τα ρυθμιστικά δίκτυα είναι αυτά που καθορίζουν τους ρυθμούς με τους οποίους μεταφράζονται τα γονίδια σε mRNA [27]. Ο φαινότυπος κάθε κυττάρου εξαρτάται από το είδος και τον αριθμό των γονιδίων που εκφράζονται σε κάθε χρονική στιγμή, καθώς και από τις επακόλουθες ενδοκυτταρικές αντιδράσεις [27]. Οι ενδοκυτταρικές διεργασίες επηρεάζονται σημαντικά και από τις εξωκυτταρικές περιβαλλοντικές συνθήκες, κυρίως, όμως, από τη δράση των ενδοκυτταρικών ρυθμιστικών μορίων.

Με βάση τα όσα ειπώθηκαν παραπάνω, το βασικό ερώτημα που γεννάται είναι πώς ένα συγκεκριμένο ρυθμιστικό δίκτυο που δρα στο επίπεδο του ενός

κυττάρου επηρεάζει την κατανομή των φαινοτύπων σε επίπεδο κυτταρικού πληθυσμού, ο οποίος απαρτίζεται από κύτταρα που περιέχουν το εν λόγω ρυθμιστικό γενετικό δίκτυο.

1.2. Ισοζύγια κυτταρικών πληθυσμών

Για να δοθεί απάντηση στο παραπάνω ερώτημα είναι απαραίτητο να κατασκευαστεί ένα μαθηματικό μοντέλο το οποίο θα μπορεί να περιγράψει τις αλληλεπιδράσεις των γονιδίων σε επίπεδο κυττάρου. Το μοντέλο θα πρέπει, επίσης, να είναι σε θέση να περιγράψει τη δυναμική ενός πληθυσμού κυττάρων, και για το λόγο αυτό είναι απαραίτητο να λαμβάνει υπόψη του την εγγενή κυτταρική ετερογένεια του πληθυσμού και να περιλαμβάνει τη μαθηματική περιγραφή των ενδοκυτταρικών διεργασιών, που χαρακτηρίζουν το ρυθμιστικό γενετικό δίκτυο κάθε κυττάρου.

Τα μοντέλα αυτά εντάσσονται σε μια γενική κατηγορία μαθηματικών μοντέλων που ονομάζονται Ισοζύγια Κυτταρικών Πληθυσμών (ΙΚΠ, Cell Population Balance Models) και έχουν παρουσιαστεί και μελετηθεί διεξοδικά [3,6,16,27]. Πρόκειται για χρονικά μεταβαλλόμενα ισοζύγια, βασικός άγνωστος των οποίων είναι ο αριθμός των κυττάρων ανά μονάδα όγκου $F(x,t)dx$ που τη χρονική στιγμή t έχει ενδοκυτταρικό περιεχόμενο μεταξύ x και dx . Τότε η δυναμική της κατανομής $F(x,t)$, αγνοώντας τις επιδράσεις των στοχαστικών διεργασιών, δίνεται από το ακόλουθο ντετερμινιστικό μοντέλο ΙΚΠ [20]:

$$\frac{\partial F(x,t)}{\partial t} + \frac{\partial}{\partial y} [r(y)F(x,t)] + \gamma(x)F(x,t) = 2 \int_x^{x_{\max}} \gamma(x')P(x,x')F(x',t)dx'. \quad (1.1)$$

Μια γενικότερη διατύπωση ΙΚΠ εκφράζει τη δυναμική της συνάρτησης πυκνότητας αριθμών $n(x,t)$ (number density function) που περιγράφει τον αριθμό των κυττάρων με περιεχόμενο x κατά τη χρονική στιγμή t διαιρεμένο με το συνολικό αριθμό κυττάρων στην ίδια χρονική στιγμή [27]:

$$\begin{aligned} \frac{\partial n(x,t)}{\partial t} + \frac{\partial}{\partial x} [R(x)n(x,t)] + \Gamma(x)n(x,t) = \\ = 2 \int_x^{x_{\max}} \Gamma(x')P(x,x')n(x',t)dx' - n(x,t) \int_0^{x_{\max}} \Gamma(x)n(x,t)dx. \end{aligned} \quad (1.2)$$

Ο πρώτος όρος του αριστερού μέλους περιγράφει τη συσσώρευση των κυττάρων με ενδοκυτταρικό περιεχόμενο x μέσα στον όγκο ελέγχου. Ο δεύτερος όρος αποτελεί το ρυθμό, με τον οποίο κύτταρα με ενδοκυτταρικό περιεχόμενο x χάνονται από τον κυτταρικό πληθυσμό λόγω των ενδοκυτταρικών αντιδράσεων οι οποίες αντιστοιχούν στο υπό μελέτη ρυθμιστικό γενετικό δίκτυο. Ο τρίτος όρος αντιστοιχεί στο ρυθμό με τον οποίο κύτταρα με ενδοκυτταρικό περιεχόμενο x χάνονται από τον πληθυσμό λόγω της κυτταρικής διαίρεσης που οδηγεί σε δημιουργία κυττάρων με μικρότερο ενδοκυτταρικό περιεχόμενο.

Ο πρώτος όρος του δευτέρου μέλους της (1.2) αντιστοιχεί στο ρυθμό γέννησης κυττάρων περιεχομένου x που έχουν προέλθει από τη διαίρεση κυττάρων με περιεχόμενο μεγαλύτερο από x . Ο αριθμός 2 που πολλαπλασιάζει αυτό τον όρο οφείλεται στο γεγονός ότι από κάθε κυτταρική διαίρεση προκύπτουν 2 νέα κύτταρα. Ο δεύτερος όρος περιγράφει τη μείωση της συνολικής συγκέντρωσης των κυττάρων λόγω της διαίρεσης των κυττάρων του πληθυσμού (dilution effect).

Σύμφωνα με την εξίσωση (1.2) η δυναμική συμπεριφορά ενός πληθυσμού κυττάρων ορίζεται πλήρως από 3 συναρτήσεις που στη βιβλιογραφία είναι γνωστές ως Intrinsic Physiological State Functions (IPSF) και περιγράφουν διεργασίες που συμβαίνουν στο επίπεδο του ενός κυττάρου. Πιο συγκεκριμένα $R(x)$ είναι ένα δίκτυο αντιδράσεων που περιγράφει το ρυθμό μεταβολής του ενδοκυτταρικού περιεχομένου ανά μονάδα χρόνου. $\Gamma(x)$ είναι ο ρυθμός με τον οποίο ένα κύτταρο με περιεχόμενο x διαιρείται προς δύο κύτταρα μικρότερου ενδοκυτταρικού περιεχομένου και έχει μονάδες αντίστροφου χρόνου. Η συνάρτηση $P(x,x')$ είναι η συνάρτηση πυκνότητας πιθανότητας που περιγράφει την πιθανότητα με την οποία ένα διαιρούμενο μητρικό κύτταρο περιεχομένου x' , γεννάει ένα θυγατρικό κύτταρο περιεχομένου x και ένα δεύτερο περιεχομένου $x' - x$. Οι συναρτήσεις που περιγράφηκαν αποτελούν το μοντέλο ενός κυττάρου (single-cell model).

Οι συνοριακές συνθήκες που εφαρμόζονται για την επίλυση της (1.2) αποτελούν τις λεγόμενες συνθήκες “containment” ή συνοριακές συνθήκες κανονικότητας και περιγράφουν την απαίτηση τα κύτταρα του πληθυσμού να μην έχουν περιεχόμενο εκτός των ορίων του διαστήματος $[0, x_{\max}]$ κατά τη χρονική στιγμή t :

$$n(0, t) = n(x_{\max}, t) = 0. \quad (1.3)$$

Η άγνωστη συνάρτηση της (1.2) είναι η συνάρτηση πυκνότητας πιθανότητας γι’ αυτό θα πρέπει να ικανοποιεί τη σχέση:

$$\int_0^{x_{\max}} n(x, t) dx = 1. \quad (1.4)$$

Στην παρούσα εργασία μελετάται ένα ρυθιστικό δίκτυο με αρχιτεκτονική θετικής ανάδρασης, τα οποία έχουν μελετηθεί σε μεγάλη έκταση, με χαρακτηριστικότερο παράδειγμα αυτό του lac operon για την παραγωγή λακτόζης [19].

1.3. Γενετική αρχιτεκτονική με θετική ανάδραση

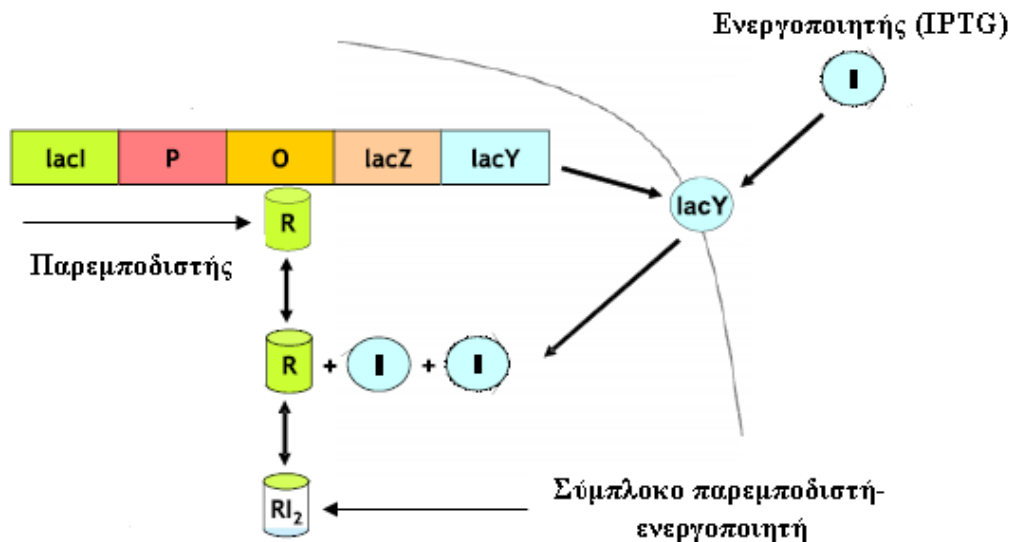
Το πρότυπο ρυθμιστικό γενετικό δίκτυο που μελετάται είναι το lac operon [1,19] (σχήμα 1.1.). Αποτελείται από τον υποστηρικτή lacP (promoter), από το χειριστή lacO (operator) και τρία γονίδια που κωδικοποιούν τις πρωτεΐνες οι οποίες είναι απαραίτητες για το μεταβολισμό της λακτόζης. Το ένα γονίδιο lacY, κωδικοποιεί τη lac permease, η οποία συμβάλλει στη μεταφορά της λακτόζης ή κάποιου αναλόγου της, όπως TMG (Methyl-1-Thio-β-D-Glactoside) ή IPTG (Isopropyl β-D-1-Thiogalactopyranoside) στο εσωτερικό του κυττάρου. Το γονίδιο lacZ κωδικοποιεί το ένζυμο β-galactosidase και το lacA την transacetylase [8,27].

Απουσία λακτόζης ή κάποιου αναλόγου της παρεμποδίζεται η έκφραση των τριών γονιδίων από την ανασχετική δράση της πρωτεΐνης lacI, η οποία εκφράζεται συνεχώς. Η lacI ενώνεται με το τμήμα DNA που προηγείται των τριών γονιδίων του lac operon (operator sequence) και εμποδίζει την RNA πολυμεράση να

πραγματοποιήσει τη μεταγραφή του DNA των τριών γονιδίων στο αντίστοιχο mRNA [5,8,27].

Παρουσία λακτόζης, αυτή μεταφέρεται στο κύτταρο στο οποίο ένα κλάσμα της υδρολύεται προς γαλακτόζη και γλυκόζη, που ενεργοποιούν τη γλυκόλυση και τον κύκλο TCA (Ticarboxylic Acid Cycle). Το υπόλοιπο κλάσμα της λακτόζης μετασχηματίζεται σε αλλολακτόζη από τη β-γαλακτοσιδάση. Η αλλολακτόζη προσδένεται στον καταστολέα lacI του lac operon, περιορίζοντας την ανασχετική του δράση. Αυξάνονται, έτσι, οι ελεύθερες θέσεις του lac operon. Ταυτόχρονα, το συσσωρευμένο σύμπλοκο cAMP-CRP (Cyclic Adenosine Monophosphate-Catabolic Repressor Protein) προσδένεται στο γονίδιο lacP, ενισχύοντας τη δράση της RNAP (RNA polymerase holoenzyme). Αυτό έχει σαν αποτέλεσμα την έναρξη της μεταγραφής των τριών γονιδίων του lac operon [15,17].

Η έκφραση του γονιδίου lacY διευκολύνει την περαιτέρω μεταφορά του ενεργοποιητή IPTG και αυτός με τη σειρά του, χρησιμοποιώντας τον ίδιο μηχανισμό, οδηγεί σε περαιτέρω έκφραση των γονιδίων του lac operon. Επομένως σε αυτό το ρυθμιστικό γενετικό δίκτυο, η έκφραση του lacY ενισχύει την περαιτέρω έκφρασή του, καθιστώντας το lac operon ένα αυτοκαταλυτικό σύστημα ή σύστημα με θετική ανάδραση (positive feedback loop).



Σχήμα 1.1.: Απεικόνιση του γενετικού ρυθμιστικού δικτύου lac operon, με θετική ανάδραση και ενεργοποίηση με IPTG [8].

Τέτοιες αρχιτεκτονικές παρουσιάζουν πολλαπλότητα λύσεων σε μόνιμη κατάσταση στο επίπεδο του ενός κυττάρου και θέλουμε να μελετήσουμε αν αυτή η

ιδιότητα επεκτείνεται και στο επίπεδο του κυτταρικού πληθυσμού [17]. Αρχικά καταστρώνεται ένα μαθηματικό μοντέλο, τέτοιο ώστε να μπορεί να περιγράψει τα βασικά χαρακτηριστικά της αρχιτεκτονικής με θετική ανάδραση [10,15]. Το μοντέλο αυτό περιλαμβάνει το ακόλουθο σύστημα ενδοκυτταρικών αντιδράσεων:



Ο χειριστής O μπορεί να είναι ενωμένος ή μη με τον παρεμποδιστή (IacI). Ο ρυθμός έκφρασης του γονιδίου στην κατάσταση κατά την οποία ο χειριστής είναι ενωμένος με τον παρεμποδιστή (k_1) είναι σημαντικά υψηλότερος σε σχέση με αυτόν στην κατάσταση στην οποία ο χειριστής δεν ενώνεται με τον παρεμποδιστή (k_o). Ως αποτέλεσμα αυτού του γεγονότος, η έκφραση του γονιδίου ενισχύει την περαιτέρω έκφραση του ίδιου γονιδίου, κάτι που είναι χαρακτηριστικό των ρυθμιστικών δικτύων με θετική ανάδραση. Στο παραπάνω σύστημα αντιδράσεων τα O_o, O_1 αντίστοιχα, σχετίζονται με την ελεύθερη και την κατειλημμένη κατάσταση του χειριστή, ενώ το Y και το Z αντίστοιχα, με τον αριθμό των μονομερών και των διμερών μορίων. Υποθέτοντας ότι οι ρυθμοί παραγωγής είναι ανάλογοι με τα κλάσματα της ελεύθερης και της κατειλημμένης κατάστασης του χειριστή και ο ρυθμός αποσύνθεσης είναι μια γραμμική συνάρτηση του ενδοκυτταρικού περιεχομένου, η δυναμική συμπεριφορά σε κάθε κύτταρο για την περίπτωση μιας μεταβλητής περιγράφεται από τη σχέση:

$$\frac{dx}{dt} = k_o O_o + k_1 O_1 - \lambda x. \tag{1.6}$$

όπου, το λ είναι η σταθερά αποσύνθεσης. Λόγω της διατήρησης των δύο καταστάσεων του χειριστή ισχύει ότι:

$$O_o + O_1 = 1. \tag{1.7}$$

Υποθέτουμε, επίσης, ότι οι δύο καταστάσεις του χειριστή βρίσκονται σε ισορροπία μεταξύ τους:

$$O_o z = \alpha O_1. \quad (1.8)$$

και το ίδιο ισχύει και για την αντίδραση διμερισμού:

$$x^2 = \beta z. \quad (1.9)$$

όπου α και β αντίστοιχα, είναι οι σταθερές ισορροπίας των αντιδράσεων μετάβασης και διμερισμού του χειριστή. Αντικαθιστώντας τις εξισώσεις (1.7-1.9) στην εξίσωση (1.6) οδηγούμαστε στη σχέση:

$$r(x) \equiv \frac{dx}{dt} = \frac{k_o \alpha \beta + k_1 x^2}{\alpha \beta + x^2} - \lambda x. \quad (1.10)$$

Αδιαστατοποιώντας το ρυθμό αντίδρασης $r(x)$ καταλήγουμε στη μορφή:

$$R(x) = \frac{\pi \rho + x^2}{\rho + x^2} - \delta x. \quad (1.11)$$

η οποία είναι μια εξίσωση με τρεις αδιάστατες παραμέτρους:

$$\pi = \frac{k_o}{k_1}, \quad (1.12)$$

$$\rho = \alpha \beta \left(\frac{\mu}{k_1} \right)^{2/L+1}, \quad (1.13)$$

$$\delta = \frac{\lambda}{k_1} \left(\frac{k_1}{\mu} \right)^{1/L+1}. \quad (1.14)$$

όπου το μ είναι ένα μέτρο του ρυθμού ανάπτυξης του πληθυσμού των κυττάρων και έχει μονάδες αντίστροφου χρόνου, ενώ ο εκθέτης L σχετίζεται με τον τρόπο που γίνεται η διαίρεση. Έτσι, μεγάλες τιμές αυτού του εκθέτη αντιστοιχούν σε πιο απότομη διαίρεση και πιο συγκεκριμένα σε μεγαλύτερη διαφορά μεταξύ των ρυθμών διαίρεσης κυττάρων με ενδοκυτταρικό περιεχόμενο μεγαλύτερο και μικρότερο από κάποιο φράγμα ενδοκυτταρικού περιεχομένου (βλ. εξ. 1.15) [8,27].

Με βάση τα όσα έχουν αναφερθεί παραπάνω σχετικά με τις δύο πιθανές καταστάσεις που μπορεί να βρίσκεται ο χειριστής, ισχύει ότι $\pi \ll 1$. Το ρ είναι αντιστρόφως ανάλογο του τετραγώνου της εξωκυτταρικής συγκέντρωσης του ενεργοποιητή IPTG και τέλος το δ είναι ο αδιάστατος ρυθμός αποσύνθεσης της συγκέντρωσης του γονιδίου lacY.

Για την πληρότητα του μαθηματικού μοντέλου καθορίζονται επίσης και ο ρυθμός διαίρεσης ως συνάρτηση του ενδοκυτταρικού περιεχομένου ($\Gamma(x)$), καθώς και η συνάρτηση πυκνότητας πιθανότητας ($P(x, x')$), η οποία περιγράφει την κατανομή του ενδοκυτταρικού περιεχομένου κατά την κυτταρική διαίρεση στα δύο θυγατρικά κύτταρα.

Για το ρυθμό διαίρεσης χρησιμοποιείται ο παρακάτω εκθετικός νόμος:

$$\Gamma(x) = \left(\frac{x}{\langle x \rangle} \right)^L. \quad (1.15)$$

Αυτή η έκφραση για το ρυθμό διαίρεσης του ενδοκυτταρικού περιεχομένου έχει εφαρμοστεί σε πολλές θεωρητικές και υπολογιστικές προσεγγίσεις του προβλήματος των ΙΚΠ [8,16,27]. Ο βασικός λόγος για την επιλογή της είναι ότι έχει βρεθεί να περιγράφει πειραματικά δεδομένα κυτταρικών πληθυσμών ζυμών (yeast) [27].

Για τη μελέτη του πώς κατανέμεται το ενδοκυτταρικό περιεχόμενο στα δύο θυγατρικά κύτταρα κατά την κυτταρική διαίρεση, χρησιμοποιήθηκε η διακριτή και ασύμμετρη συνάρτηση πυκνότητας πιθανότητας, που αποτελεί την απλούστερη δυνατή μορφή:

$$P(x, x') = \frac{1}{2f} \delta(fx' - x) + \frac{1}{2(1-f)} \delta((1-f)x' - x), \quad (1.16)$$

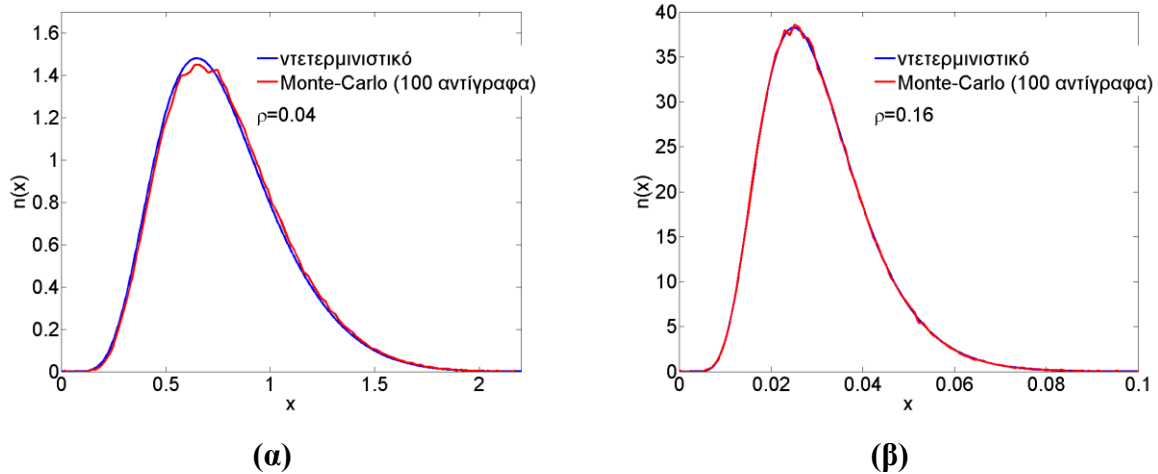
όπου δ είναι η συνάρτηση Dirac και f το κλάσμα ενδοκυτταρικού περιεχομένου που κληροδοτείται από το μητρικό κύτταρο στο μικρότερο θυγατρικό κύτταρο. Το δεύτερο κύτταρο κληρονομεί ενδοκυτταρικό περιεχόμενο $1-f$. Με τον ορισμό αυτόν γίνεται φανερό ότι η τιμή του f μπορεί να κυμαίνεται μεταξύ 0 και 0.5. Στην περίπτωση που $f=0.5$ βρισκόμαστε στην περίπτωση διακριτής και συμμετρικής συνάρτησης πυκνότητας πιθανότητας, ενώ οι μικρότερες τιμές του f αναλογούν σε μεγαλύτερο βαθμό ασυμμετρίας κατά την κυτταρική διαίρεση [16].

Το μαθηματικό μοντέλο όπως διατυπώθηκε αποτελεί μια προσέγγιση του προβλήματος των ΙΚΠ, η οποία όμως δε λαμβάνει υπόψη της την «τυχαιότητα» που χαρακτηρίζει μικρούς πληθυσμούς κυττάρων. Για τη μελέτη και της τυχαίας φύσης κυτταρικών πληθυσμών αναπτύχθηκε ένα στοχαστικό μοντέλο, το οποίο βασίζεται σε μεγάλο βαθμό στο μαθηματικό πρόβλημα που περιγράφηκε παραπάνω και υλοποιείται με την κατασκευή αλγορίθμου Monte Carlo.

Κεφάλαιο 2: Το στοχαστικό μοντέλο

2.1. Εισαγωγή

Το μαθηματικό μοντέλο που περιγράφηκε στο προηγούμενο κεφάλαιο είναι ντετερμινιστικό και δε λαμβάνει υπόψη του την «τυχειότητα» που διέπει μικρούς πληθυσμούς κυττάρων, καθώς επίσης και τη στοχαστική φύση των κυτταρικών διαιρέσεων, η οποία επιδρά σημαντικά στη δυναμική συμπεριφορά κυτταρικών πληθυσμών σε στάδια ανάπτυξης, όταν η κυτταρική πυκνότητα είναι χαμηλή. Ένα άλλο μειονέκτημά του είναι ότι δεν μπορεί να συμπεριλάβει τον ενδογενή θόρυβο στο επίπεδο του ενός κυττάρου, που είναι σημαντικός για την κυτταρική ετερογένεια [1]. Ένα τέτοιο μοντέλο προσφέρει μερική εκτίμηση της κυτταρικής ετερογένειας και των επιδράσεων της στη συμπεριφορά του συστήματος. Για να ξεπεραστούν οι ελλείψεις αυτές εφαρμόστηκε ένα στοχαστικό μοντέλο, το οποίο περιέχει τις ίδιες πληροφορίες με το ντετερμινιστικό (ρυθμός αντίδρασης και διαίρεσης στο επίπεδο του ενός κυττάρου και συνάρτηση πυκνότητας πιθανότητας για την κατανομή του ενδοκυτταρικού περιεχομένου) και ταυτόχρονα λαμβάνει υπόψη του και τη στοχαστικότητα των κυτταρικών διαιρέσεων, όπως και το στοχαστικό θόρυβο που εισάγεται από τις ενδοκυτταρικές αντιδράσεις. Οι διαφορές του στοχαστικού μοντέλου από το ντετερμινιστικό φαίνονται στα σχήματα 2.1. (α) και 2.1. (β), όπου απεικονίζεται η κατανομή των κυττάρων $n(x)$ συναρτήσει του ενδοκυτταρικού περιεχομένου x , σε μόνιμη κατάσταση, για δύο τιμές της αδιάστατης παραμέτρου εξωκυτταρικής συγκέντρωσης, ρ .



Σχήμα 2.1.: Σύγκριση του στοχαστικού μοντέλου με το ντετερμινιστικό για τιμές παραμέτρου (α) $\rho = 0.04$ και (β) $\rho = 0.16$.

Στο σχήμα 2.1. παρατηρούμε ότι η συμφωνία μεταξύ των αποτελεσμάτων του στοχαστικού και του ντετερμινιστικού μοντέλου είναι πολύ καλή και για τις δύο τιμές παραμέτρου ρ που χρησιμοποιήθηκαν. Οι όποιες διαφορές υπάρχουν στις κατανομές οφείλονται στη στοχαστικότητα και στον ενδογενή θόρυβο που είναι χαρακτηριστικά των κυτταρικών δράσεων (διαιρέσεων και ενδοκυτταρικών αντιδράσεων) και περιλαμβάνονται στο στοχαστικό μοντέλο, όχι όμως και στο ντετερμινιστικό.

Το στοχαστικό μοντέλο αναπτύχθηκε με τη βοήθεια αλγορίθμου kinetic Monte Carlo που δρα μεταξύ σταθερού όγκου και σταθερού αριθμού (constant volume-constant number Monte Carlo, CVMC-CNMC) [16,17]. Στον αλγόριθμο Monte Carlo σταθερού αριθμού ο αριθμός των σωματιδίων διατηρείται σταθερός ανεξάρτητα από το αν η πραγματική διαδικασία οδηγεί σε απώλεια σωματιδίων ή σε αύξησή τους [23]. Σε πραγματικές διεργασίες ο αριθμός συγκέντρωσης των σωματιδίων ενός δείγματος μπορεί να αυξάνεται, να μειώνεται, ή να υφίσταται πιο πολύπλοκες μεταβάσεις. Ανεξάρτητα από τη συμπεριφορά αυτή όμως, στον αλγόριθμο Monte Carlo σταθερού όγκου 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 [13]. Το διάγραμμα ροής του αλγορίθμου παρουσιάζεται στο σχήμα 2.2.

Αρχικά υπολογίζεται ο χρόνος που μεσολαβεί μεταξύ δύο συμβάντων διαίρεσης T . Αφού αυξηθεί ο συνολικός χρόνος της προσομοίωσης κατά T και βρεθεί το ενδοκυτταρικό περιεχόμενο κάθε κυττάρου τη χρονική στιγμή $\tau + T$, εντοπίζεται το κύτταρο που έχει υποστεί διαίρεση και καθορίζεται το ενδοκυτταρικό

περιεχόμενο των δύο θυγατρικών κυττάρων που προέκυψαν από τη διαίρεση. Τέλος, επειδή ο κυτταρικός πληθυσμός δεν μπορεί να υπερβεί μια μέγιστη τιμή, γίνεται έλεγχός του και επιλέγεται τυχαία ένα κύτταρο να απομακρυνθεί από αυτόν προς αποκατάστασή του.

2.2. Υπολογισμός του χρόνου που μεσολαβεί μεταξύ συμβάντων διαίρεσης

Σε μια δεδομένη στιγμή του αδιαστατοποιημένου χρόνου τ η τυχαία κατάσταση $(S(\tau))$ στην οποία βρίσκεται ο πληθυσμός των κυττάρων καθορίζεται από το συνολικό αριθμό των κυττάρων $(M_o(\tau))$ και από το ενδοκυτταρικό περιεχόμενο που περιέχεται σε αυτά τα κύτταρα $(X_i(\tau))$ στο χρόνο τ :

$$S_\tau \equiv \{M_o(\tau) = N_s; X_i(\tau) = x_i, i = 1, 2, \dots, N_s\}. \quad (2.1)$$

Τα κύτταρα διαιρούνται με ρυθμό $\Gamma(x)$ (εξ. 1.15) και το ενδοκυτταρικό περιεχόμενο των νέων κυττάρων καθορίζεται από τη συνάρτηση πυκνότητας πιθανότητας $P(x, x')$ (εξ. 1.16). Μεταξύ των κυτταρικών διαιρέσεων το περιεχόμενο κάθε κυττάρου μεταβάλλεται σύμφωνα με τη διαφορική εξίσωση:

$$\frac{dx_i}{dt} = R(x_i), \quad i = 1, 2, \dots, N_s. \quad (2.2)$$

όπου $R(x_i)$ είναι το δίκτυο των ενδοκυτταρικών αντιδράσεων που περιγράφεται από την εξίσωση 1.11.

Ένα πολύ σημαντικό μέγεθος στις στοχαστικές προσομοιώσεις κυτταρικών πληθυσμών είναι ο χρόνος T , μεταξύ δύο διαδοχικών κυτταρικών διαιρέσεων. Ο χρόνος αυτός είναι μια τυχαία μεταβλητή, που εξαρτάται από την κατάσταση $(S(\tau))$ στην οποία βρίσκεται ο πληθυσμός των κυττάρων τη χρονική στιγμή τ , με συνάρτηση κατανομής πιθανότητας (cumulative distribution function-CDF) [19]:

$$F_T(z|\tau) \equiv \Pr\{T \leq z | S_\tau\} = 1 - \Pr\{T > z | S_\tau\} = P_T(z|\tau). \quad (2.3)$$

Στηριζόμενοι στο γεγονός ότι μόνο ένα συμβάν διαίρεσης μπορεί να μεταβάλλει τον πληθυσμό των κυττάρων, μπορούμε να συσχετίσουμε τις πιθανότητες $P_T(z + dz | \tau)$ και $P_T(z | \tau)$ ως εξής:

$$P_T(z + dz | \tau) = P_T(z | \tau) \left[1 - \sum_{i=1}^{N_s} \Gamma(x_i(\tau + z)) \right]. \quad (2.4)$$

Παίρνοντας το όριο του $dz \rightarrow 0$ και διαιρώντας και τα δύο μέλη της εξίσωσης 2.4 με το dz , οδηγούμαστε στη διαφορική εξίσωση για την πιθανότητα $P_T(z | \tau)$, που έχει την ακόλουθη μορφή:

$$\frac{dP_T}{dz} = -P_T \sum_{i=1}^{N_s} \Gamma(x_i(\tau + z)). \quad (2.5)$$

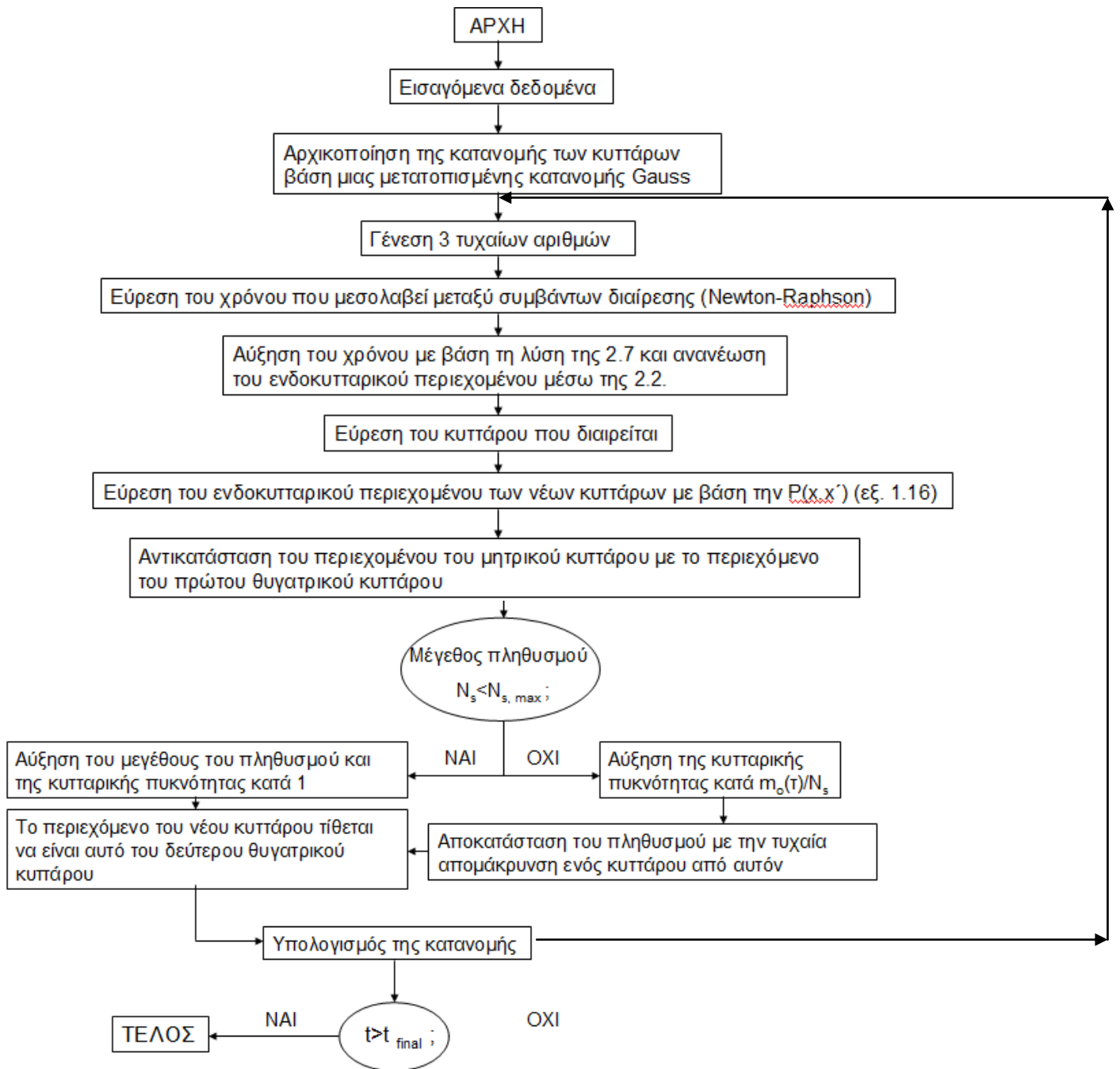
Με βάση τον ορισμό της πιθανότητας $P_T(z | \tau)$, η αρχική συνθήκη είναι η $P_T(0 | \tau) = 0$. Λύνοντας το πρόβλημα αρχικής τιμής (εξ. 2.5), με αρχική συνθήκη την $P_T(0 | \tau) = 0$ και αντικαθιστώντας στην εξίσωση 2.3, βρίσκουμε ότι:

$$F_T(z | \tau) = 1 - \exp \left[- \int_0^z \sum_{i=1}^{N_s} \Gamma(x_i(\tau + z')) dz' \right]. \quad (2.6)$$

Τα κύτταρα συνεχίζουν να αντιδρούν κατά τη διάρκεια του χρόνου που μεσολαβεί μεταξύ συμβάντων κυτταρικών διαιρέσεων, ενώ ο πληθυσμός τους παραμένει αμετάβλητος. Γι' αυτό, το ενδοκυτταρικό περιεχόμενο που εμφανίζεται στο ολοκλήρωμα της εξίσωσης 2.6 υπολογίζεται από τη χρονική ολοκλήρωση της εξίσωσης 2.2. Έτσι, για να υπολογιστεί η τιμή του χρόνου T που μεσολαβεί μεταξύ των κυτταρικών διαιρέσεων, από την εξίσωση κατανομής πιθανότητας 2.6 γεννάται ένας τυχαίος αριθμός από την ομοιόμορφη κατανομή στο διάστημα $(0,1)$, εξισώνεται με την εξίσωση 2.6 και επιλύουμε ως προς T . Η προκύπτουσα εξίσωση εξαρτάται από τη μορφή που έχει η αντίδραση $R(x)$ και είναι μη-γραμμική [16]:

$$Q(T) \equiv \frac{\sum_{i=1}^{N_s} \int_0^T \Gamma(x_i(\tau+z)) dz}{\ln(1-rand)} + 1 = 0. \quad (2.7)$$

Η εξίσωση 2.7 επιλύεται επαναληπτικά με γνωστές μεθόδους (Newton-Raphson). Για τον υπολογισμό του ολοκληρώματος εφαρμόστηκε ο κανόνας του τραπεζίου. Για να βρεθεί το ενδοκυτταρικό περιεχόμενο κάθε κυττάρου τη χρονική στιγμή $\tau + T$, που απαιτείται για την ολοκλήρωση με τον κανόνα του τραπεζίου, επιλύεται η διαφορική εξίσωση 2.2 με τη ρητή μέθοδο Euler.



Σχήμα 2.2.: Σχηματική απεικόνιση του διαγράμματος ροής του αλγορίθμου kinetic Monte Carlo [16].

2.3. Καθορισμός του μητρικού κυττάρου που διαιρείται

Αφού υπολογιστεί ο χρόνος που μεσολαβεί μεταξύ των κυτταρικών διαρρέσεων και αυξηθεί ο συνολικός χρόνος της προσομοίωσης κατά T , καθορίζεται το κύτταρο που έχει υποστεί διαίρεση. Ορίζεται ένας δείκτης k , τέτοιος ώστε να ισχύει $k = j$ όταν ένα κύτταρο με περιεχόμενο x_j διαιρείται. Η συνάρτηση κατανομής πιθανότητας των κυττάρων που παριστάνονται με αυτόν τον δείκτη,

λαμβάνοντας υπόψη την κατάσταση στην οποία βρίσκεται ο συνολικός πληθυσμός σε αυτή τη χρονική στιγμή, δίνεται από τη σχέση:

$$\Pr\{k = j | S_{\tau+T}\} = \frac{\Gamma(x_j(\tau+T))}{\sum_{i=1}^{N_s} \Gamma(x_i(\tau+T))}. \quad (2.8)$$

Με τη γένεση ενός δεύτερου τυχαίου αριθμού από την ομοιόμορφη κατανομή στο διάστημα $(0,1)$, υπολογίζεται η τιμή του j που υπακούει στη συνάρτηση κατανομής πιθανότητας 2.8 και, έτσι, ανιχνεύεται το κύτταρο το οποίο έχει διαιρεθεί.

2.4. Καθορισμός του ενδοκυτταρικού περιεχομένου των δύο θυγατρικών κυττάρων

Μετά την ανίχνευση του μητρικού κυττάρου που διαιρείται και τον υπολογισμό του ενδοκυτταρικού του περιεχομένου, καθορίζεται το περιεχόμενο των δύο νέων κυττάρων τα οποία προκύπτουν από τη διαίρεση. Ο καθορισμός γίνεται βάση της διακριτής και ασύμμετρης συνάρτησης πυκνότητας πιθανότητας $P(x, x')$ (εξ. 1.16), η οποία χρησιμοποιείται σε αυτή την περίπτωση. Το περιεχόμενο του πρώτου θυγατρικού κυττάρου είναι ίσο με ένα κλάσμα f του περιεχομένου του μητρικού κυττάρου. Το ενδοκυτταρικό περιεχόμενο του δεύτερου θυγατρικού κυττάρου προκύπτει από την αφαίρεση του περιεχομένου του πρώτου θυγατρικού κυττάρου από το περιεχόμενο του μητρικού κυττάρου, λόγω της διατήρησης της μάζας.

2.5. «Αποκατάσταση» του κυτταρικού πληθυσμού

Κάθε κυτταρική διαίρεση οδηγεί στον αφανισμό ενός μητρικού κυττάρου και στη γένεση δύο νέων, θυγατρικών κυττάρων. Μετά τον καθορισμό του περιεχομένου καθενός από τα δύο θυγατρικά κύτταρα, το ενδοκυτταρικό περιεχόμενο του μητρικού κυττάρου αντικαθίσταται από αυτό του πρώτου θυγατρικού κυττάρου.

Τη διαδικασία αυτή ακολουθεί έλεγχος του μεγέθους του πληθυσμού. Αν η τιμή του είναι μικρότερη από μία προκαθορισμένη μέγιστη τιμή, τότε το μέγεθος του πληθυσμού και η κυτταρική πυκνότητα αυξάνονται κατά μία μονάδα, ώστε να

αντικατοπτριστεί το πρόσφατο συμβάν διαίρεσης. Το περιεχόμενο του νέου κυττάρου του πληθυσμού τίθεται ίσο με το περιεχόμενο του δεύτερου θυγατρικού κυττάρου.

Διαφορετικά, αν ο κυτταρικός πληθυσμός είναι ίσος με τη μέγιστη προκαθορισμένη τιμή, το μέγεθός του παραμένει αμετάβλητο κατά την επόμενη κυτταρική διαίρεση, ενώ η κυτταρική πυκνότητα αυξάνεται κατά $m_o(\tau)/N_s$ μονάδες, όπου $m_o(\tau)$ ο αριθμός των κυττάρων πριν την κυτταρική διαίρεση. Από τη στιγμή που γεννάται ένα νέο κύτταρο κατά τη διαίρεση και ο πληθυσμός των κυττάρων δεν μπορεί να υπερβεί μία μέγιστη τιμή, ο πληθυσμός πρέπει να «αποκατασταθεί» προκειμένου να συμπεριληφθεί σε αυτόν και το δεύτερο νεογέννητο κύτταρο. Για το λόγο αυτό, γεννάται ένας τρίτος τυχαίος αριθμός από την ομοιόμορφη κατανομή στο διάστημα $(0,1)$ ο οποίος επιλέγει τυχαία ένα κύτταρο από τον πληθυσμό. Το κύτταρο που επιλέγεται απομακρύνεται από τον πληθυσμό και αντικαθίσταται από το δεύτερο θυγατρικό κύτταρο. Όλα τα κύτταρα το πληθυσμού έχουν ίσες πιθανότητες να επιλεγούν προς αντικατάσταση.

Με αυτόν τον τρόπο, ο αλγόριθμος συμπεριφέρεται ως υβρίδιο μεταξύ σταθερού όγκου και σταθερού αριθμού kinetic Monte Carlo. Όσο το μέγεθος του πληθυσμού είναι μικρότερο από το προκαθορισμένο, ο αλγόριθμος δρα σαν σταθερού όγκου Monte Carlo αυξάνοντας το μέγεθος του πληθυσμού μετά από κάθε κυτταρική διαίρεση και ακολουθώντας τη δυναμική της κυτταρικής πυκνότητας. Όταν, όμως, ο πληθυσμός φτάσει τη μέγιστη τιμή του, διατηρείται σταθερός για το υπόλοιπο της προσομοίωσης όπως γίνεται σε προσομοιώσεις με αλγόριθμο Monte Carlo σταθερού αριθμού.

Τα βήματα που περιγράφηκαν παραπάνω επαναλαμβάνονται μέχρι το σύστημα να φτάσει στον επιθυμητό τελικό χρόνο. Μετά το πέρας της προσομοίωσης υπολογίζεται η συνάρτηση πυκνότητας αριθμών και αποθηκεύεται σε αρχεία.

Ο αλγόριθμος που περιγράφηκε υπερτερεί αλγορίθμων Monte Carlo σταθερού αριθμού ή Monte Carlo σταθερού όγκου ως προς τον υπολογιστικό χρόνο που απαιτείται για να εκτελεστεί η προσομοίωση [16].

Εντούτοις, η χρήση στοχαστικών αλγορίθμων είναι μια χρονοβόρος διαδικασία. Ενδεικτικά, αναφέρεται ότι για να φτάσουμε σε μόνιμη κατάσταση προσομοιώνοντας τη δυναμική ενός πληθυσμού $N_s = 10000$ με τιμή παραμέτρου εξωκυτταρικής συγκέντρωσης $\rho = 0.06$ απαιτείται χρόνος περίπου 5 λεπτών για κάθε αντίγραφο. Για να έχουμε μια αρκούντως αντιπροσωπευτική συμπεριφορά του πληθυσμού των

κυττάρων χρησιμοποιούνται 100 αντίγραφα, οπότε ο συνολικός χρόνος που απαιτείται για να φτάσει το σύστημα σε μόνιμη κατάσταση ανέρχεται περίπου στα 500 λεπτά. Επιπροσθέτως, ενδιαφερόμαστε όχι για τη συμπεριφορά κάθε κυττάρου ξεχωριστά, αλλά για τη συμπεριφορά του πληθυσμού ως συνόλου.

Για τους παραπάνω λόγους, καθώς επίσης και για τη μελέτη της συμπεριφοράς του πληθυσμού για διάφορες τιμές παραμέτρων, καταφεύγουμε σε μεθόδους αδρομερούς ανάλυσης.

Κεφάλαιο 3: Αδρομερείς υπολογισμοί δυναμικών συστημάτων

3.1. Εισαγωγή

Ένας από τους βασικούς στόχους της παρούσας εργασίας είναι η ανάπτυξη ενός υπολογιστικού πλαισίου, το οποίο θα επιτρέπει την πραγματοποίηση μεθόδων συστημικής ανάλυσης χρησιμοποιώντας μοντέλα μικροκλίμακας. Το υπολογιστικό αυτό πλαίσιο επιτρέπει την αποτελεσματική μελέτη της μακροσκοπικής συμπεριφοράς συστημάτων, η περιγραφή των οποίων είναι διαθέσιμη μόνο σε επίπεδο μικρο-κλίμακας, χωρίς την ανάγκη εξαγωγής μοντέλου (διαφορικών εξισώσεων) που θα περιγράφει τη δυναμική συμπεριφορά σε μακροσκοπική κλίμακα. Μια τέτοια τεχνική είναι η «ελεύθερη εξισώσεων» μεθοδολογία (equation-free) στην οποία εξάγεται μακροσκοπική πληροφορία από κατάλληλα διαμορφωμένα υπολογιστικά πειράματα, που διεξάγονται σε επίπεδο μικρο-κλίμακας.

Η βασική ιδέα της υπολογιστικής μεθοδολογίας είναι η παράκαμψη εξαγωγής του μακροσκοπικού μοντέλου και η επιτυχία της διασφαλίζεται μόνο εφόσον μπορεί να υπάρξει μια μακροσκοπική περιγραφή, ακόμα και αν αυτή δεν είναι διαθέσιμη σε κλειστή μορφή [11,12,22,27]. Σε περιπτώσεις κατά τις οποίες μακροσκοπικά μοντέλα είναι διαθέσιμα με τη μορφή διαφορικών εξισώσεων, η υπολογιστική ανάλυση πρέπει να γίνεται απευθείας σε αυτά. Η προτεινόμενη τεχνική εφαρμόζεται σε περιπτώσεις που εξετάζονται πολύπλοκα συστήματα, η δυναμική περιγραφή των οποίων παρέχεται σε επίπεδο μικρο-κλίμακας όπου η εξαγωγή πληροφορίας σε επίπεδο μακρο-κλίμακας απαιτεί μεγάλο υπολογιστικό κόστος. Η μεθοδολογία που αναπτύσσεται αποτελεί μια συστηματική διαδικασία για την άντληση πληροφορίας από μικροσκοπικούς προσομοιωτές άμεσα και αποτελεσματικά σε μακροσκοπικό επίπεδο και παράγει την ίδια πληροφορία που θα αντλούσαμε αν το μακροσκοπικό μοντέλο υπήρχε και ήταν διαθέσιμο σε κλειστή μορφή.

Η «ελεύθερη εξισώσεων» μεθοδολογία είναι ικανή να επιταχύνει σε πολύ μεγάλο βαθμό τους υπολογισμούς, καθώς δε βασίζεται στη λεπτομερή περιγραφή της μικρο-κατάστασης (π.χ. πληθυσμός 10000 κυττάρων) αλλά σε έναν μικρό αριθμό στατιστικών μεγεθών που περιγράφουν τη μικρο-κατάσταση (ροπές χαμηλότερης τάξης). Ο αριθμός των ροπών που επιλέγεται είναι πολύ μικρός σε σχέση για παράδειγμα με τον αριθμό των κυττάρων και ταυτόχρονα παρέχει την απαραίτητη

ακρίβεια στους υπολογισμούς των κατανομών, οι οποίες είναι συνεπείς με τις κατανομές που προκύπτουν από τα στοχαστικά μοντέλα.

3.2. Αδρομερής χρονικός ολοκληρωτής

Ένα από τα μεγάλα προβλήματα στις σύγχρονες υπολογιστικές τεχνικές είναι η γεφύρωση των πολλαπλών χωρικών και χρονικών κλιμάκων που συναντώνται σε μεγάλα και περίπλοκα συστήματα. Πολύ συχνά η καλύτερη δυνατή περιγραφή τους είναι διαθέσιμη σε επίπεδο μικρο-κλίμακας και για τη δυναμική προσομοίωσή τους εφαρμόζονται τεχνικές Monte Carlo, Μοριακής Δυναμικής, Brownian Δυναμικής, κ.ά. Η «ελεύθερη εξίσωση» μεθοδολογία αποτελεί μια ελπιδοφόρα προσέγγιση στη μελέτη της μακροσκοπικής περιγραφής ενός περίπλοκου συστήματος, η δυναμική συμπεριφορά του οποίου περιγράφεται σε μικροσκοπικό επίπεδο [11].

Το βασικό εργαλείο που επιτρέπει την εκτέλεση αριθμητικών εφαρμογών σε επίπεδο μακρο-κλίμακας, χρησιμοποιώντας μικροσκοπικούς προσομοιωτές είναι ο αδρομερής χρονικός ολοκληρωτής, T . Ο αδρομερής χρονικός ολοκληρωτής παρέχει μια προσέγγιση του χρόνου τ που εμπλέκεται στην εξίσωση που διέπει το μακροσκοπικό μοντέλο, αλλά δεν είναι διαθέσιμη σε κλειστή μορφή. Ουσιαστικά, δίνει τη μακροσκοπική εικόνα του εξεταζόμενου συστήματος ανά τακτά χρονικά διαστήματα. Τα διαστήματα αυτά είναι μικρά αν τα παρατηρούμε από τη μακροσκοπική κλίμακα, είναι όμως μεγάλα αν η οπτική μας είναι από τη μικροσκοπική κλίμακα. Ο χρόνος $\tau > 0$ αποτελεί το χρονικό ορίζοντα του αδρομερούς χρονικού ολοκληρωτή και είναι τόσο μικρός ώστε να κάνει τον υπολογισμό του πραγματικού προβλήματος εύκολο και τόσο μεγάλος ώστε να επιτρέπει αρκετές αλλαγές της δυναμικής περιγραφής σε μακροσκοπικό επίπεδο για να μπορούμε να τις εκτιμήσουμε τοπικά [11].

Ο αδρομερής χρονικός ολοκληρωτής στηρίζεται στο λεπτομερή χρονικό ολοκληρωτή, τη λεπτομερή μικροσκοπική εξέλιξη του συστήματος, καθώς και σε έναν κατάλληλο, μη-μοναδικό τελεστή άρσης μ (lifting operator), ο οποίος απεικονίζει τη μακροσκοπική περιγραφή U σε μια συνεπή μικροσκοπική περιγραφή u . Αν M είναι ένας συντελεστής περιορισμού (restriction operator), που καθορίζει τη επιλογή των αδρομερών μεταβλητών και αντιστοιχίζει τη μικροσκοπική περιγραφή στη μακροσκοπική, τότε πρέπει να ισχύει $M\mu = I$. Δηλαδή, κάνοντας άρση (lifting) από τη μακρο-κλίμακα στη μικρο-κλίμακα και στη συνέχεια την αντίστροφη

διαδικασία (restriction), θα πρέπει να μην παρατηρούνται επιδράσεις στο σύστημα. Για πολλές μικροσκοπικές περιγραφές η αδρομερής στατιστική απεικόνιση στηρίζεται στις πρώτες/χαμηλότερης τάξης ροπές των μικροσκοπικών κατανομών. Για παράδειγμα σε μια προσομοίωση της δυναμικής των ρευστών σε επίπεδο σωματιδίων, η πυκνότητα είναι ίση με τη μηδενική ροπή της κατανομής των ταχυτήτων των σωματιδίων [11].

Τα βασικά βήματα της προτεινόμενης μεθοδολογίας είναι συνοπτικά τα εξής [11,27]:

- i.** Επιλέγονται τα κατάλληλα στατιστικά μεγέθη των μικροσκοπικών κατανομών, τα οποία περιγράφουν τη συμπεριφορά σε μακροσκοπικό επίπεδο. Στο παραπάνω παράδειγμα προσομοίωσης της δυναμικής των ρευστών τα μεγέθη αυτά είναι η πίεση, η πυκνότητα και η ταχύτητα. Η μακροσκοπική απεικόνιση συμβολίζεται με το διάνυσμα U . Στο σύστημα που εξετάζεται στην εργασία επιλέγονται οι ροπές χαμηλής τάξης.
- ii.** Επιλέγεται ένας κατάλληλος τελεστής άρσης μ (lifting operator) από τα μακροσκοπικά μεγέθη σε μία ή περισσότερες αντίστοιχες μικροσκοπικές κατανομές u , μέσω της σχέσης $u = \mu U$.
- iii.** Πραγματοποιείται η χρονική ολοκλήρωση σε επίπεδο μικροσκοπικού προσομοιωτή με αρχικές συνθήκες ένα σύνολο μικροσκοπικών κατανομών, οι οποίες είναι συμβατές με τα αντίστοιχα μακροσκοπικά μεγέθη. Η χρονική ολοκλήρωση γίνεται για ένα σχετικά σύντομο χρονικό ορίζοντα, που επιλέγεται κατάλληλα από το χρήστη έτσι ώστε να πληροί τις προϋποθέσεις που αναφέρθηκαν παραπάνω.
- iv.** Μετά τη χρονική ολοκλήρωση, υπολογίζονται τα αδρομερή μεγέθη εφαρμόζοντας έναν τελεστή περιορισμού M (restriction operator) στις νέες μικροσκοπικές κατανομές. Για τον τελεστή περιορισμού ισχύει $U = Mu$.

Τα βήματα (i)-(iv) αποτελούν μια απεικόνιση Φ_t , όπου t ο χρονικός ορίζοντας του προσομοιωτή της μικρο-κλίμακας, σύμφωνα με την οποία:

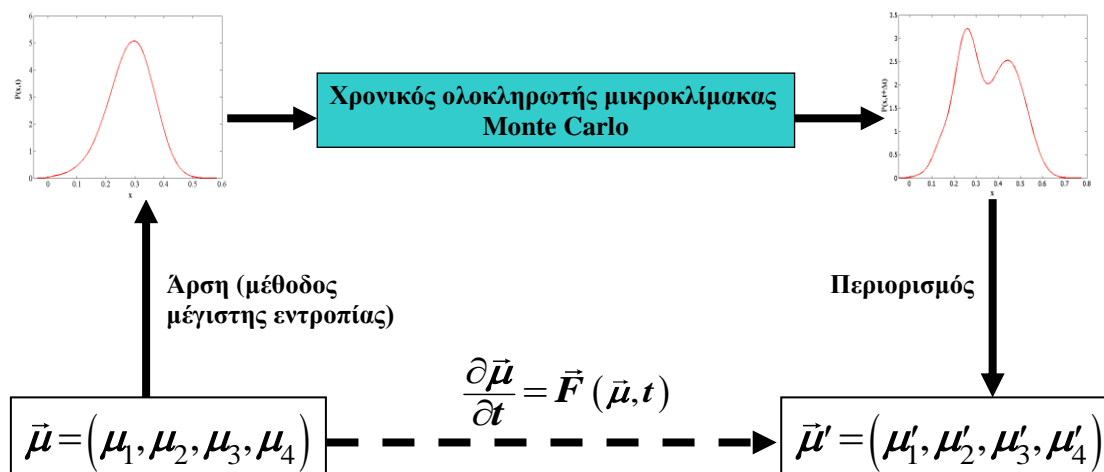
$$u(k+1) = \Phi_t(u(k)), \quad (3.1)$$

όπου $u(k)$ η λύση που αντιστοιχεί στο k βήμα του αδρομερούς χρονικού ολοκληρωτή (3.1).

Ο αδρομερής χρονικός ολοκληρωτής παρέχει ένα πλαίσιο που επιτρέπει τον υπολογισμό αριθμητικών μεγεθών, όπως είναι το δεξί μέλος εξισώσεων, το γινόμενο του Ιακωβιανού πίνακα με διανύσματα, αδρομερών παραγώγων ως προς παραμέτρους του συστήματος, κ.ά. Οι ποσότητες αυτές χρησιμοποιούνται για την εύρεση μόνιμων καταστάσεων με την εφαρμογή επαναληπτικών μεθόδων (π.χ. Newton-Raphson) σε συνδυασμό με γραμμικούς επιλύτες τύπου Krylov [9], για την ανάλυση ευστάθειας, την ανάλυση διακλάδωσης και το σχεδιασμό ρυθμιστών.

Τα βασικά τμήματα του αδρομερούς χρονικού ολοκληρωτή που χρησιμοποιείται για τον υπολογισμό μόνιμων καταστάσεων περιγράφονται συνοπτικά στα παρακάτω τρία βήματα και παρουσιάζονται σχηματικά στο διάγραμμα 3.1.:

- i. Η άρση (lifting) από τα μακροσκοπικά μεγέθη στις αντίστοιχες μικροσκοπικές κατανομές.
- ii. Η ολοκλήρωση που γίνεται με τη χρήση του αλγορίθμου Monte Carlo, με αρχική συνθήκη την κατανομή του βήματος (i).
- iii. Ο περιορισμός (restrict) για τον εκ νέου υπολογισμό των ροπών που χαρακτηρίζουν την κατανομή που υπολογίστηκε.



Σχήμα 3.1.: Αδρομερής χρονικός ολοκληρωτής για τη χρονική εξέλιξη του κυτταρικού πληθυσμού.

3.3. Άρση με τη μέθοδο της μέγιστης εντροπίας (lifting with the maximum entropy method)

Στην παρούσα εργασία η διαδικασία της άρσης πραγματοποιείται με τη βοήθεια της μεθόδου της μέγιστης εντροπίας [18]. Η μέθοδος αυτή αποτελεί μια εναλλακτική προσέγγιση δημιουργίας διανυσμάτων μικροκαταστάσεων συνεπών με την αδρομερή περιγραφή, η οποία παρέχεται από τις ροπές μικρότερης τάξης μιας μικρο-κατανομής (κυτταρικός πληθυσμός). Ουσιαστικά, πρόκειται για ένα πρόβλημα ροπών, όπου αναζητούνται κατανομές με δεδομένες ροπές χαμηλότερης τάξης και η μέθοδος της μέγιστης εντροπίας δίνει την κατανομή εκείνη που έχει τη μέγιστη εντροπία.

Οι ροπές χαμηλότερης τάξης προκύπτουν από τις κατανομές των στοχαστικών προσομοιώσεων εφαρμόζοντας τη διαδικασία του περιορισμού (restrict). Η σχέση υπολογισμού των ροπών είναι η ακόλουθη:

$$\mu_i = \frac{1}{N_s} \sum_{j=1}^{N_s} X_j^{i-1} \quad i=1, \dots, N. \quad (3.2)$$

Η μεγιστοποίηση της εντροπίας για μια μονοδιάστατη κατανομή μπορεί να τεθεί με τέτοιο τρόπο ώστε να ικανοποιεί τη συνάρτηση Lagrange, υπό τον περιορισμό ότι οι ροπές που χρησιμοποιούνται είναι αυτές που έχουν προκύψει από την αδρομερή περιγραφή της μικροκατάστασης:

$$L(n^{lift}, \underline{\lambda}) = -\int n^{lift}(x) \ln n^{lift}(x) dx + \sum_{i=0}^N \lambda_i \left(\int x^i n^{lift}(x) dx - \mu_i \right). \quad (3.3)$$

Στην εξίσωση (3.3) $n^{lift}(x)$ είναι η προσέγγιση της συνάρτησης πυκνότητας πιθανότητας των κυττάρων ως συνάρτηση του ενδοκυτταρικού περιεχομένου, x . μ_i είναι η ροπή i -ιοστής τάξης, υπολογισμένη από την κατανομή των κυττάρων όπως λαμβάνεται από το μικροσκοπικό προσομοιωτή και $\underline{\lambda}$ το διάνυσμα των πολλαπλασιαστών Lagrange. Σύμφωνα με τη μέθοδο της μέγιστης εντροπίας η συνάρτηση πυκνότητας αριθμών δίνεται από τη σχέση:

$$n^{lift}(x) = \exp\left(\sum_{i=0}^N \lambda_i x^i - 1\right). \quad (3.4)$$

Με δεδομένο ένα πεπερασμένο αριθμό ροπών που προκύπτει από την κατανομή των κυττάρων, η άρση μπορεί να γίνει ακολουθώντας τα εξής βήματα:

i. Αρχικά, υπολογίζονται οι πολλαπλασιαστές Lagrange που βελτιστοποιούν την αντίστοιχη συνάρτηση (εξ. 3.3). Ο υπολογισμός γίνεται με τη χρήση σχημάτων ολοκλήρωσης (π.χ Gauss-Legendre). Με αυτόν τον τρόπο, μια προσέγγιση της εντροπίας Shannon είναι:

$$S(x) = -\int n^{lift}(x) \ln n^{lift}(x) dx \approx -\sum_{i=1}^n w_i n^{lift}(x_i) \ln n^{lift}(x_i), \quad (3.5)$$

όπου w_i και x_i είναι τα βάρη και οι συντεταγμένες (weights and abscissas) αντίστοιχα, του εφαρμοζόμενου σχήματος ολοκλήρωσης. Οι πολλαπλασιαστές Lagrange υπολογίζονται από τη λύση της εξίσωσης:

$$R_k \equiv \sum_{i=1}^n w_i x_i^k \exp\left(\sum_{j=0}^N (\lambda_j x_i^j - 1)\right) - \mu_k = 0, \quad k = 0, 1, 2, \dots, N. \quad (3.6)$$

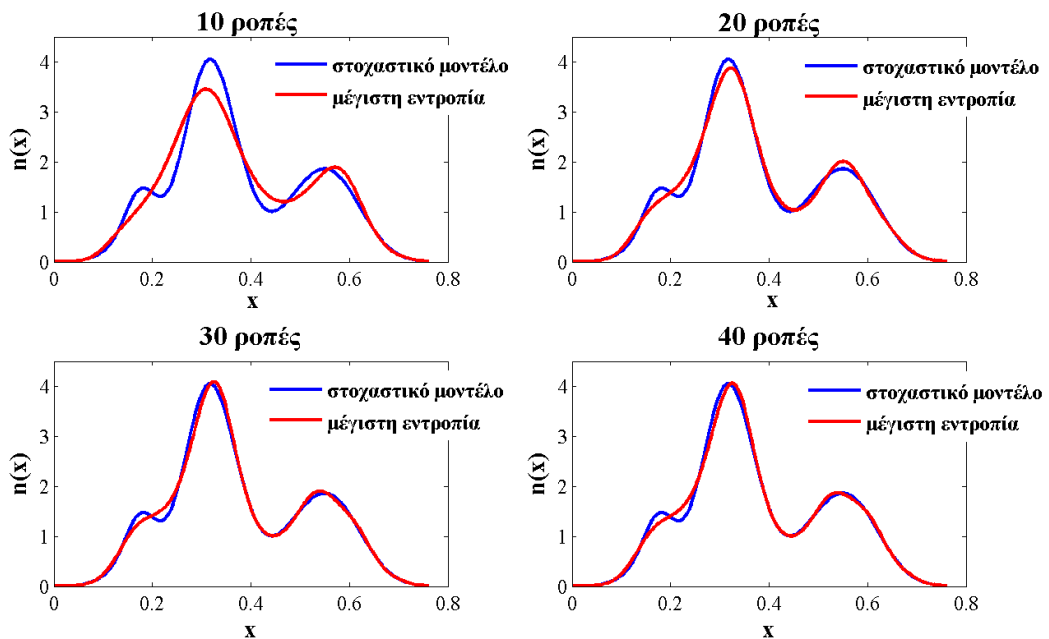
Το σύστημα (3.6) είναι ένα μη γραμμικό σύστημα εξισώσεων και μπορεί να επιλυθεί με επαναληπτικές μεθόδους (Newton-Raphson). Σε κάθε επανάληψη υπολογίζεται μια ανανέωση $\delta \underline{\lambda}$ του διανύσματος των πολλαπλασιαστών Lagrange, μέσω της επίλυσης του γραμμικού συστήματος:

$$\underline{\underline{H}} \delta \underline{\lambda} = -R, \quad (3.7)$$

όπου $\underline{\underline{H}}$ είναι το Εσσιανό μητρώο (Hessian matrix) που δίνεται από :

$$\underline{\underline{H}} = \frac{\partial R_k}{\partial \lambda_j} \sum_{i=1}^n w_i x_i^j x_i^k \exp\left(\sum_{j=0}^N (\lambda_j x_i^j - 1)\right), \quad j, k = 0, 1, 2, \dots, N. \quad (3.8)$$

Οι ζητούμενοι πολλαπλασιαστές Lagrange ανανεώνονται σε κάθε επανάληψη μέσω της σχέσης $\underline{\lambda}^{(m+1)} = \underline{\lambda}^{(m)} + \delta\underline{\lambda}$, όπου το m συμβολίζει τα επαναληπτικό βήμα. Μόλις επιτευχθεί η σύγκλιση, οι προκύπτοντες πολλαπλασιαστές Lagrange χρησιμοποιούνται στην εξίσωση (3.4) για την ανακατασκευή της συνάρτησης πυκνότητας πιθανότητας $n^{lift}(x)$. Σε αυτό το σημείο πρέπει να αναφερθεί ότι η προσέγγιση και η ανακατασκευή της κατανομής γίνεται πολύ ευαίσθητη ανάλογα με την επιλογή της αρχικής τιμής των πολλαπλασιαστών Lagrange, καθώς και με τον αριθμό των ροπών που χρησιμοποιούνται. Όσο αυξάνεται ο αριθμός των ροπών, δυσκολεύει η σύγκλιση της μεθόδου και μπορεί να οδηγηθούμε ακόμα και σε απόκλιση, καθώς το Εσσιανό μητρώο τείνει να γίνει ιδιάζον [26]. Πέρα όμως από αυτή τη δυσκολία, με την αύξηση του αριθμού των ροπών, αυξάνεται και η ακρίβεια προσέγγισης της κατανομής, μέχρι ενός σημείου, όπου περαιτέρω αύξηση των ροπών δεν έχει καμία επίδραση στην ακρίβεια της λύσης (σχήμα 3.2.).



Σχήμα 3.2.: Ανακατασκευή κατανομών μικροκαταστάσεων με τη μέθοδο της μέγιστης εντροπίας, χρησιμοποιώντας 10, 20, 30 και 40 ροπές. Οι μπλε γραμμές παριστάνουν τις κατανομές που έχουν προκύψει από το στοχαστικό μοντέλο, ενώ οι κόκκινες αυτές από την άρση με τη μέγιστη εντροπία.

ii. Υπολογίζεται η αντίστοιχη συνάρτηση αθροιστικής κατανομής, $F(x_j)$, σύμφωνα με τον τύπο:

$$F(x_j) = \frac{\sum_{i=1}^j w_i n^{lift}(x_i)}{\sum_{i=1}^n w_i n^{lift}(x_i)}. \quad (3.9)$$

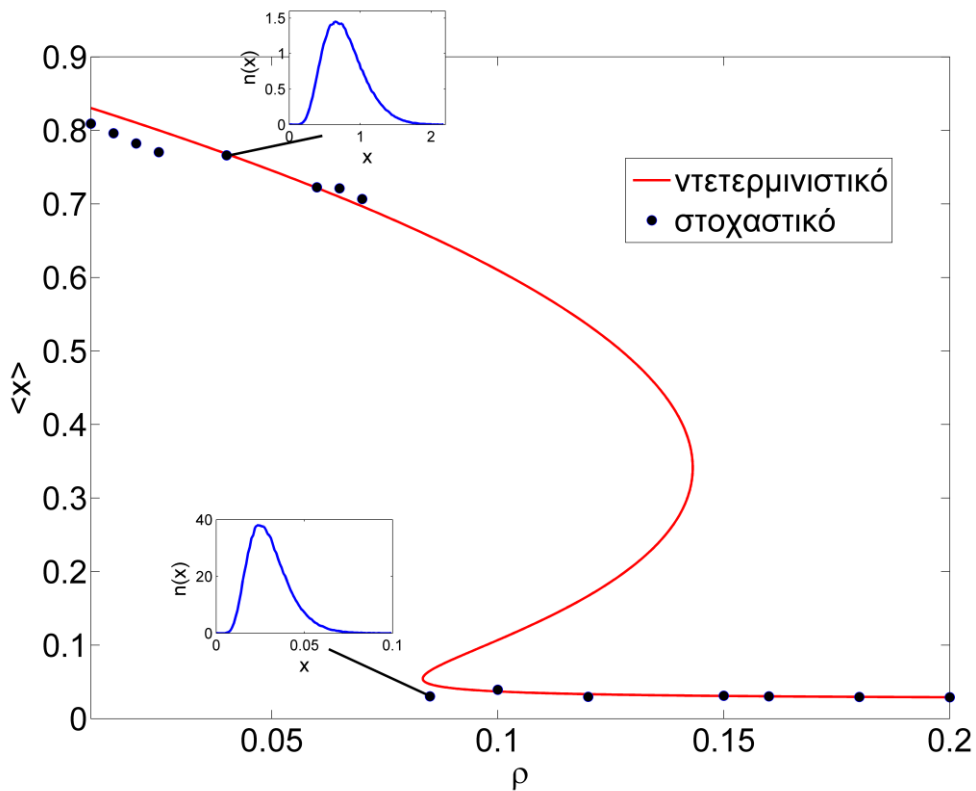
και δημιουργείται μια συνεπής μικροσκοπική κατανομή. Για κάθε ένα από τα σωματίδια, N_s , της κατανομής γεννάται ένας τυχαίος αριθμός από την ομοιόμορφη κατανομή (0,1). Στη συνέχεια, στο σωματίδιο k ($1 \leq k \leq N_s$) αποδίδεται τιμή ενδοκυτταρικού περιεχομένου x_j , αν ισχύει $F(x_{j-1}) < p < F(x_j)$ (ορίζεται $F(x_0) = 0$). Με αυτόν τον τρόπο, παράγεται μια κατανομή που είναι συνεπής με την αδρομερή περιγραφή (πρώτες ροπές) της κατανομής των κυττάρων, η οποία προέκυψε από τη στοχαστική προσομοίωση, έχοντας περιορίσει σημαντικά το υπολογιστικό κόστος σε σχέση με τις στοχαστικές προσομοιώσεις.

Λόγω του μικρότερου υπολογιστικού χρόνου που απαιτούν οι αδρομερείς προσομοιώσεις, οι υπολογισμοί μόνιμων καταστάσεων γίνονται με τη χρήση της «ελεύθερης εξισώσεων» μεθοδολογίας. Κατασκευάζεται ένας αλγόριθμος βασισμένος στη μέθοδο της μέγιστης εντροπίας, ο οποίος είναι ικανός να υπολογίσει καταστάσεις σε συνθήκες μόνιμης κατάστασης για διάφορες τιμές παραμέτρων. Οι υπολογισμοί αδρομερών μόνιμων καταστάσεων αναπτύσσονται αναλυτικά στην παράγραφο που ακολουθεί.

3.4. Υπολογισμοί αδρομερών μόνιμων καταστάσεων

Σε περιπτώσεις που η μακροσκοπική δυναμική του συστήματος προσεγγίζει μια λύση μόνιμης κατάστασης, ο αδρομερής χρονικός ολοκληρωτής μπορεί να χρησιμοποιηθεί για τον υπολογισμό μόνιμων καταστάσεων, για την παραμετρική ανάλυση ή για την ανάλυση διακλάδωσης [7,14] (σχήμα 3.3.) σε μόνιμη κατάσταση χρησιμοποιώντας μεθόδους μήκους τόξου (arc-length continuation). Στο σχήμα 3.3. παρουσιάζεται η ασυμπτωτική συμπεριφορά του μέσου κυτταρικού πληθυσμού ($N_s = 10000$ κύτταρα) σε σχέση με την παράμετρο εξωκυτταρικής συγκέντρωσης ρ , που έχει προκύψει από το ντετερμινιστικό μοντέλο (κόκκινη γραμμή) και από τη

χρήση του αδρομερούς χρονικού ολοκληρωτή σε συνδυασμό με τη Newton-Raphson (μαύροι κύκλοι). Παρατηρούμε ότι με τη χρήση του αδρομερούς χρονικού ολοκληρωτή, η ακρίβεια υπολογισμού της μέσης τιμής του πληθυσμού σε μόνιμη κατάσταση είναι μεγάλη και οι υπολογισμοί με αυτόν τον τρόπο προσεγγίζουν ικανοποιητικά τη δυναμική συμπεριφορά του πληθυσμού των κυττάρων, σε σημαντικά μικρότερο χρόνο σε σχέση με την ευθεία εφαρμογή του στοχαστικού αλγορίθμου.



Σχήμα 3.3.: Ασυμπτωτική συμπεριφορά του μέσου κυτταρικού πληθυσμου συναρτήσει της αδιάστατης παραμέτρου ρ όπως έχει προβλεφθεί από το ντετερμινιστικό μοντέλο (κόκκινη συνεχής γραμμή) και από το στοχαστικό μοντέλο (μαύροι κύκλοι).

Στη μόνιμη κατάσταση από την εξίσωση (3.1) θα ισχύει:

$$u(k+1) = \Phi_t(u(k)) \rightarrow u' = \Phi_t(u') \rightarrow$$

$$\rightarrow R_{coarse} = u' - \Phi_t(u') = 0, \quad (3.10)$$

όπου u' είναι η μακροσκοπική λύση της μόνιμης κατάστασης. Σε κάθε επανάληψη της επαναληπτικής μεθόδου Newton-Raphson που χρησιμοποιείται επιλύεται το γραμμικό σύστημα:

$$\frac{\partial R_{coarse}}{\partial u'} \delta u = -R_{coarse} \rightarrow \left(I - \frac{\partial \Phi_t(u')}{\partial u'} \right) \delta u = \Phi_t(u') - u'. \quad (3.11)$$

Όπως γίνεται κατανοητό από τη μορφή της εξίσωσης (3.4), για την επίλυση του συστήματος είναι απαραίτητος ο υπολογισμός της παραγώγου $\frac{\partial \Phi_t(u')}{\partial u'}$. Όμως, η Φ_t δεν περιγράφεται αναλυτικά και έτσι δεν είναι γνωστή η ακριβής μορφή της. Για να ξεπεραστεί αυτή η δυσκολία, καταφεύγουμε σε προσεγγιστικές μεθόδους υπολογισμού της παραπάνω παραγώγου. Πιο συγκεκριμένα χρησιμοποιούνται σχήματα πεπερασμένων διαφορών ως εξής:

$$\frac{\partial \Phi_t(u')}{\partial u'} = \frac{\Phi_t(u' + \varepsilon) - \Phi_t(u')}{\varepsilon}, \quad (3.12)$$

όπου ε είναι κατάλληλα επιλεγμένη μικρή ποσότητα (παράγοντας διαταραχής-perturbation factor). Για την επιλογή του ε υπολογίζεται η τυπική απόκλιση των ροπών, σε μόνιμη κατάσταση, μεταξύ των αντιγράφων που χρησιμοποιούνται στον αλγόριθμο Monte Carlo και στον παράγοντα διαταραχής δίνεται τιμή λίγο μεγαλύτερη από την υπολογισμένη τυπική απόκλιση.

Το διάγραμμα ροής της μεθόδου Newton-Raphson που χρησιμοποιείται για τον υπολογισμό λύσεων μόνιμης κατάστασης παρουσιάζεται στο σχήμα 3.4. Αρχικά, γίνεται η άρση με τη μέθοδο της μέγιστης εντροπίας, βάσει των ροπών N της κατανομής και στη συνέχεια καλείται ο αλγόριθμος Monte Carlo, ο οποίος διαβάζει την κατανομή που έχει προκύψει από την άρση σαν αρχική εκτίμηση. Ο αλγόριθμος εκτελείται για ένα μικρό χρονικό ορίζοντα και μετά το πέρας του εφαρμόζεται η διαδικασία του περιορισμού (restrict) για τον εκ νέου υπολογισμό των ροπών της καινούριας κατανομής. Υπολογίζεται έτσι το δεξί μέλος της εξίσωσης (3.11) ως:

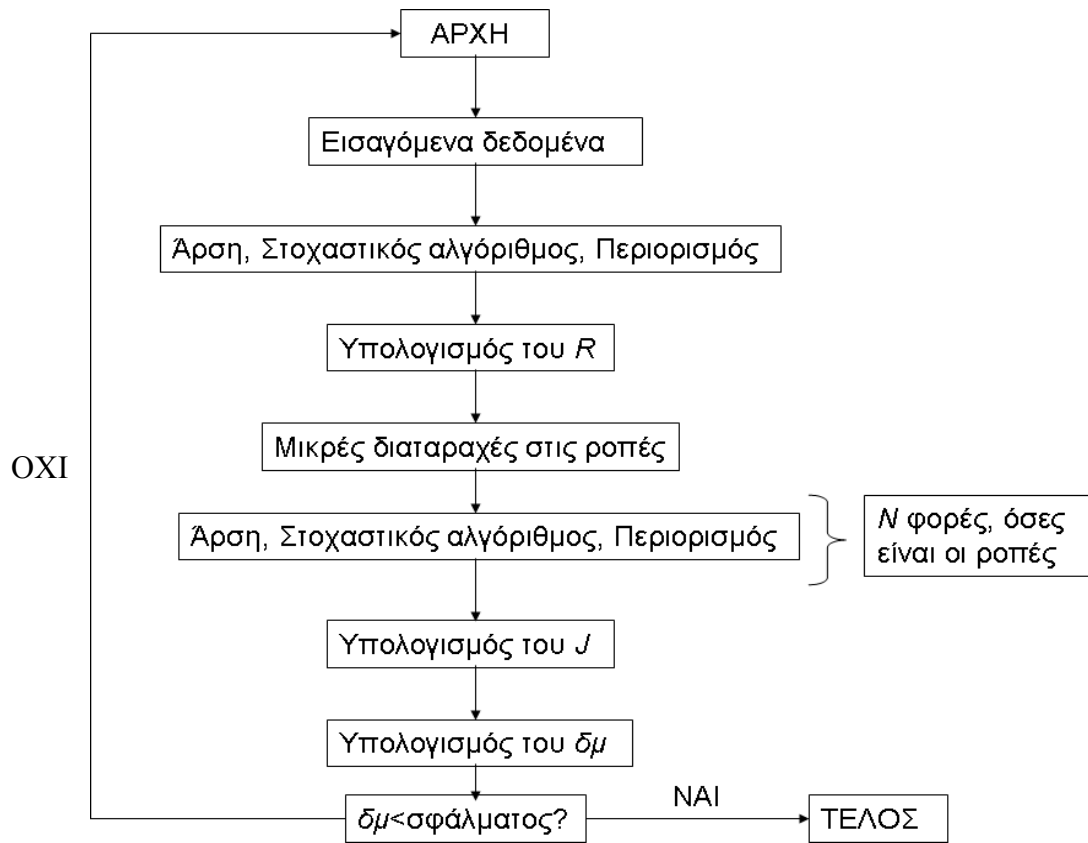
$$R(i) = \underline{\mu}'(i) - \underline{\mu}(i), \quad i = 0, 1, \dots, N. \quad (3.13)$$

Στη συνέχεια, εφαρμόζεται η ίδια διαδικασία αφού γίνει διαταραχή στην πρώτη ροπή. Η διαδικασία αυτή επαναλαμβάνεται τόσες φορές όσος είναι ο αριθμός των ροπών, ώστε να εφαρμοστούν μικρές διαταραχές σε όλες τις ροπές. Αποτέλεσμα αυτής της διαδικασίας είναι ο υπολογισμός ενός Ιακωβιανού μητρώου $\underline{\underline{J}}$ βάσει της σχέσης (3.12):

$$\begin{aligned}
 J(i, j) &= \frac{\partial R(i)}{\partial \mu(j)} = \frac{R_i(\mu'(j) + \varepsilon) - R_i(\mu'(j))}{\varepsilon} = \\
 &= \frac{(\Phi_t(\mu'_{pert}) - \mu'_{pert}) - (\Phi_t(\mu') - \mu')}{\varepsilon}.
 \end{aligned} \tag{3.14}$$

Σε κάθε επανάληψη της μεθόδου λύνεται τελικά το γραμμικό σύστημα $\underline{\underline{J}}\underline{\underline{\delta\mu}} = -\underline{\underline{R}}$. Η λύση $\underline{\underline{\delta\mu}}$ του συστήματος χρησιμοποιείται για την ανανέωση της λύσης μέσω της σχέσης $\underline{\underline{\mu}}^{(n+1)} = \underline{\underline{\mu}}^{(n)} + \underline{\underline{\delta\mu}}$, όπου με n συμβολίζεται το επαναληπτικό βήμα.

Επιπλέον της μεθόδου Newton-Raphson είναι δυνατόν να εφαρμοστεί και η μέθοδος παραμετρικού βηματισμού μήκους-τόξου ώστε να επιτυγχάνεται σύγκλιση και σε μη ευσταθείς καταστάσεις [8]. Η εύρεση των παραγώγων και σε αυτή τη μέθοδο γίνεται προσεγγιστικά, με τον ίδιο τρόπο που γίνεται και στη Newton-Raphson.



Σχήμα 3.4.: Διάγραμμα ροής της επαναληπτικής διαδικασίας Newton-Raphson που χρησιμοποιήθηκε.

Κεφάλαιο 4: Αποτελέσματα υπολογιστικής ανάλυσης

4.1. Εισαγωγή

Στο κεφάλαιο αυτό παρατίθενται τα αποτελέσματα της υπολογιστικής ανάλυσης των ΙΚΠ. Οι κώδικες που δημιουργήθηκαν για την πραγματοποίηση της υπολογιστικής μελέτης είναι γραμμένοι σε γλώσσα προγραμματισμού C++ και χρησιμοποιείται το εμπορικό λογισμικό MATLAB. Οι προσομοιώσεις έγιναν τόσο σε περιβάλλον windows όσο και σε linux.

Αρχικά, παρουσιάζονται τα αποτελέσματα που προέκυψαν με τη χρήση του στοχαστικού αλγορίθμου Monte Carlo και παρατηρείται το πότε το σύστημα οδηγείται σε μόνιμη κατάσταση. Τα αποτελέσματα αυτά είναι η βάση για την επιλογή αδρομερών μεγεθών (ροπές). Οι ροπές χρησιμοποιούνται στη συνέχεια για προσομοιώσεις με τη μέθοδο της μέγιστης εντροπίας και τα αποτελέσματα συγκρίνονται με αυτά που έχουν προκύψει από την επίλυση του στοχαστικού μοντέλου.

Στις προσομοιώσεις με τη χρήση του αλγορίθμου Monte Carlo χρησιμοποιήθηκαν 100 αντίγραφα για τη μείωση του στατιστικού σφάλματος. Συνεπώς, οι κατανομές που παρουσιάζονται είναι ο μέσος όρος μεταξύ των 100 αντιγράφων. Κάτι ανάλογο εφαρμόστηκε και στη διαδικασία της άρσης με τη μέθοδο της μέγιστης εντροπίας.

Ο υπολογιστικός χρόνος που απαιτείται για τις στοχαστικές προσομοιώσεις ποικίλλει ανάλογα με το χρονικό ορίζοντα για τον οποίο ο αλγόριθμος εκτελεί την προσομοίωση. Για χρονικό ορίζοντα (αδιάστατος χρόνος) από 0 έως 0.1 απαιτείται υπολογιστικός χρόνος ίσος με 7 δευτερόλεπτα για κάθε αντίγραφο, ενώ αντίθετα για χρονικό ορίζοντα από 9 έως 10 χρειάζονται περίπου 12 λεπτά για κάθε αντίγραφο. Αντίθετα, το συνολικό υπολογιστικό κόστος της άρσης με τη μέθοδο της μέγιστης εντροπίας περιορίζεται στα 1-2 λεπτά.

4.2. Αποτελέσματα του στοχαστικού μοντέλου

Όπως αναφέρθηκε και στα κεφάλαια 1 και 2, ο στοχαστικός προσομοιωτής στηρίζεται στη γενετική αρχιτεκτονική με θετική ανάδραση. Το δίκτυο αντιδράσεων της αρχιτεκτονικής αυτής εκφράζεται με τη σχέση:

$$R(x) = \frac{\pi\rho + x^2}{\rho + x^2} - \delta x. \quad (1.11)$$

και ο ρυθμός διαίρεσης των κυττάρων προς τη δημιουργία θυγατρικών με διαφορετικό ενδοκυτταρικό περιεχόμενο με τη σχέση:

$$\Gamma(x) = \left(\frac{x}{\langle x \rangle} \right)^L. \quad (1.15)$$

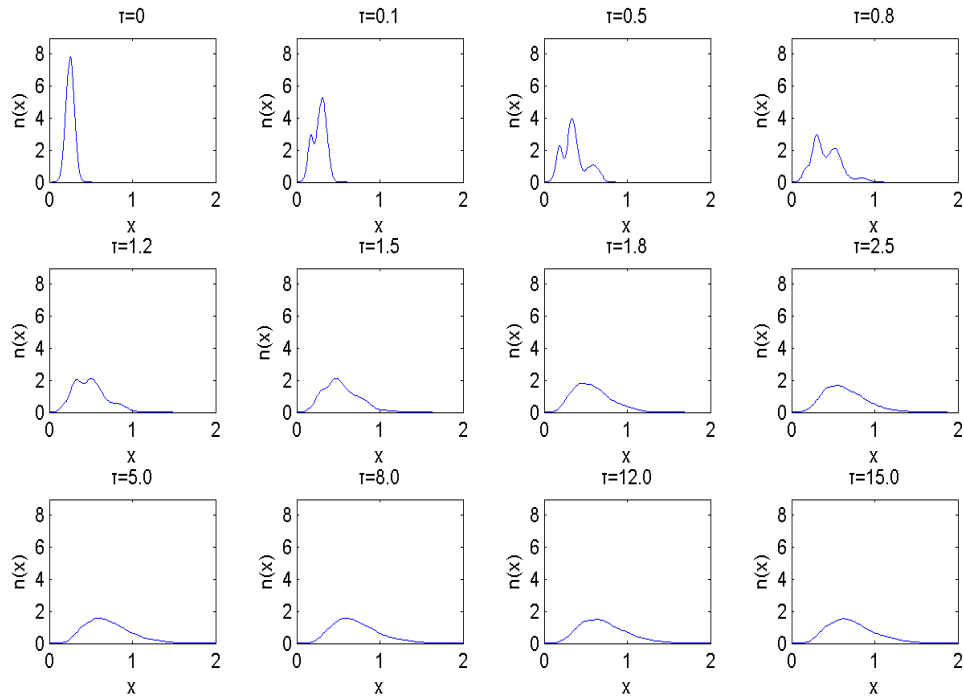
ενώ το ενδοκυτταρικό περιεχόμενο των θυγατρικών κυττάρων καθορίζεται από τη συνάρτηση πυκνότητας πιθανότητας

$$P(x, x') = \frac{1}{2f} \delta(fx' - x) + \frac{1}{2(1-f)} \delta((1-f)x' - x). \quad (1.16)$$

Για τις στοχαστικές προσομοιώσεις που πραγματοποιήθηκαν χρησιμοποιούνται οι εξής τιμές παραμέτρων: $\pi = 0.03$, $\delta = 0.05$, $L = 2$ και $f = 0.5$, εξετάζουμε, δηλαδή, την περίπτωση κατά την οποία η συνάρτηση πυκνότητας πιθανότητας είναι συμμετρική. Για την παράμετρο ρ χρησιμοποιήθηκαν τιμές $\rho = 0.06$, $\rho = 0.04$ και $\rho = 0.15$. Ο λόγος που εξετάζονται οι τρεις αυτές τιμές παραμέτρου εξηγείται παρακάτω, στην παράγραφο 4.3. Ο πληθυσμός κυττάρων που χρησιμοποιείται σε όλες αυτές τις περιπτώσεις είναι $N_s = 10000$.

Σε κάθε προσομοίωση χρησιμοποιείται ως αρχική συνθήκη μια μετατοπισμένη κατανομή Gauss με μέση τιμή $\mu = 0.25$ και τυπική απόκλιση $\sigma = 0.05$.

Στο σχήμα 4.1., παρουσιάζονται τα αποτελέσματα από το στοχαστικό μοντέλο, για τις τιμές παραμέτρων που αναφέρθηκαν και για $\rho = 0.06$. Κάθε διάγραμμα απεικονίζει τη μεταβολή της κατανομής των κυττάρων $n(x)$ συναρτήσει του ενδοκυτταρικού περιεχομένου σε διάφορες χρονικές στιγμές και εξετάζεται πώς μεταβάλλεται το ενδοκυτταρικό περιεχόμενο στην εξέλιξη του χρόνου.



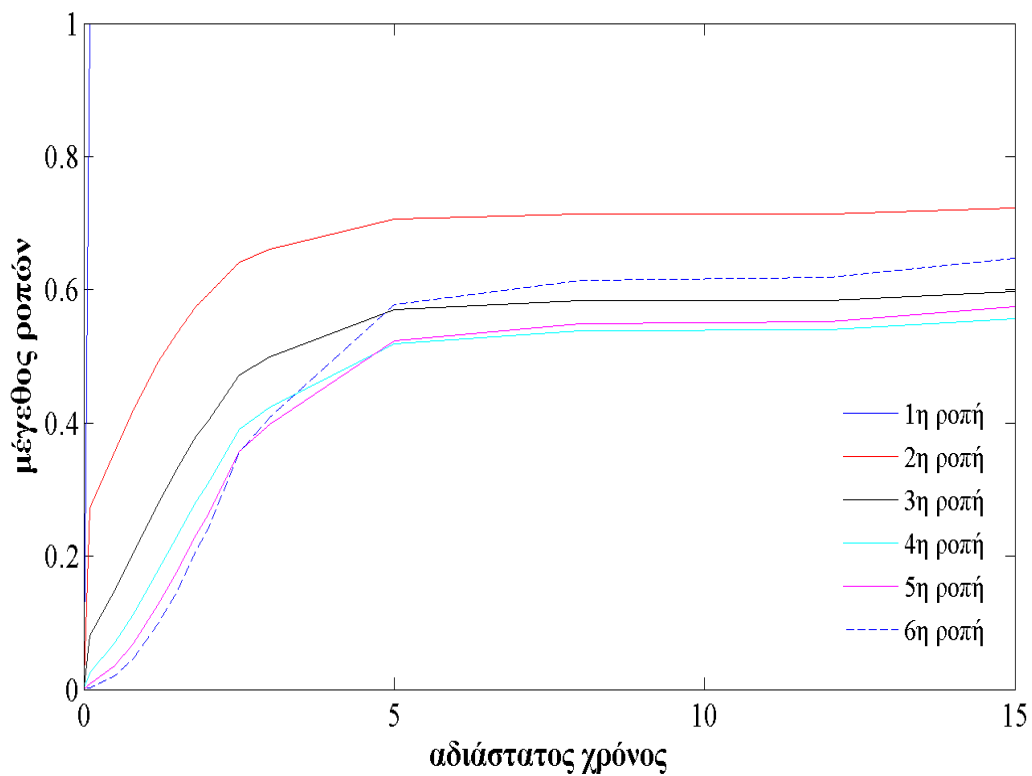
Σχήμα 4.1.: Στιγμιότυπα της στοχαστικής προσομοίωσης για ένα πληθυσμό κυττάρων που χρησιμοποιεί την γενετική αρχιτεκτονική με θετική ανάδραση. Τιμές των παραμέτρων που χρησιμοποιήθηκαν: $\pi = 0.03$, $\delta = 0.05$, $\rho = 0.06$, $L = 2$ και $f = 0.5$. Το μέγεθος του πληθυσμού είναι σταθερό και ίσο με $N_s = 10000$.

Στο παραπάνω σχήμα, παρατηρούμε ότι οι μεταβολές στην κατανομή των κυττάρων είναι απότομες στα πρώτα χρονικά βήματα εξέλιξης της προσομοίωσης παρουσιάζοντας δύο ή και τρεις κορυφές. Όσο, όμως, εξελίσσεται η προσομοίωση, οι κατανομές γίνονται πιο ομαλές χωρίς απότομες αλλαγές και από ένα χρόνο και έπειτα δεν εμφανίζουν μεταβολές. Αυτός ο χρόνος είναι ίσος με $\tau=2.5$ και από τότε και μετά θεωρούμε ότι βρισκόμαστε σε μόνιμη κατάσταση (steady state).

Από τις κατανομές του σχήματος 4.1. υπολογίζονται τα αδρομερή μεγέθη τους, οι ροπές. Υπολογίζονται 6 ροπές μέσω της σχέσης:

$$\mu_i = \frac{1}{N_s} \sum_{j=1}^{N_s} X_j^{i-1}, \quad i = 1, \dots, 6. \quad (3.2)$$

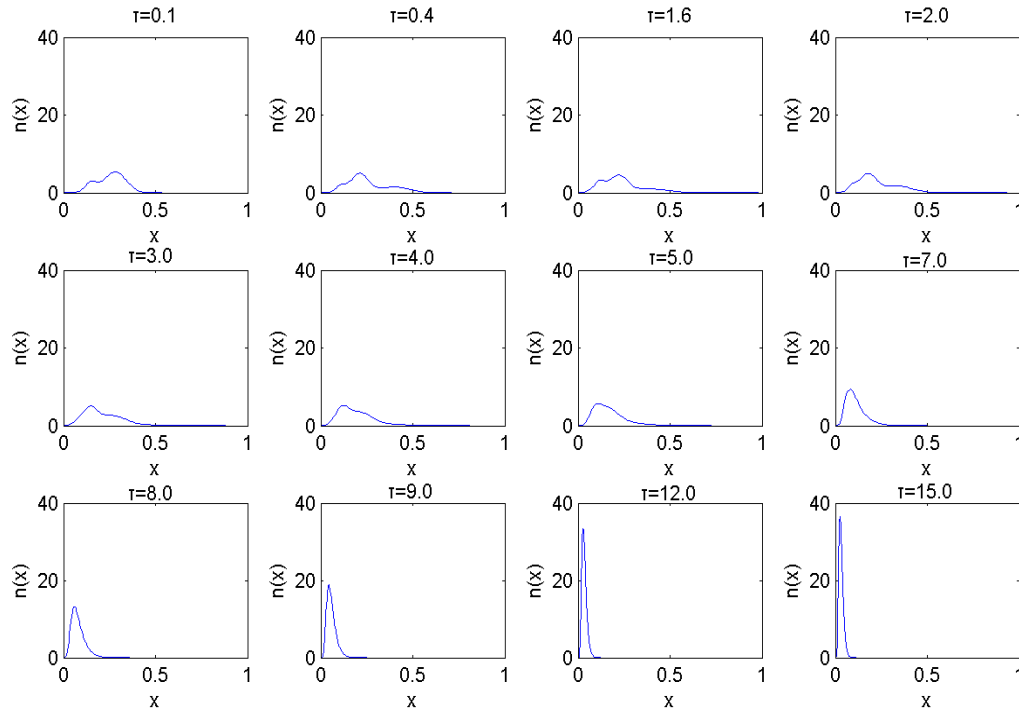
και η χρονική εξέλιξή τους φαίνεται στο σχήμα 4.2.



Σχήμα 4.2.: Χρονική εξέλιξη των 6 πρώτων ροπών που προέκυψαν από τις κατανομές των στοχαστικών προσομοιώσεων για τιμές παραμέτρων $\pi = 0.03$, $\delta = 0.05$, $\rho = 0.06$, $L = 2$ και $f = 0.5$ και πληθυσμό κυττάρων $N_s = 10000$.

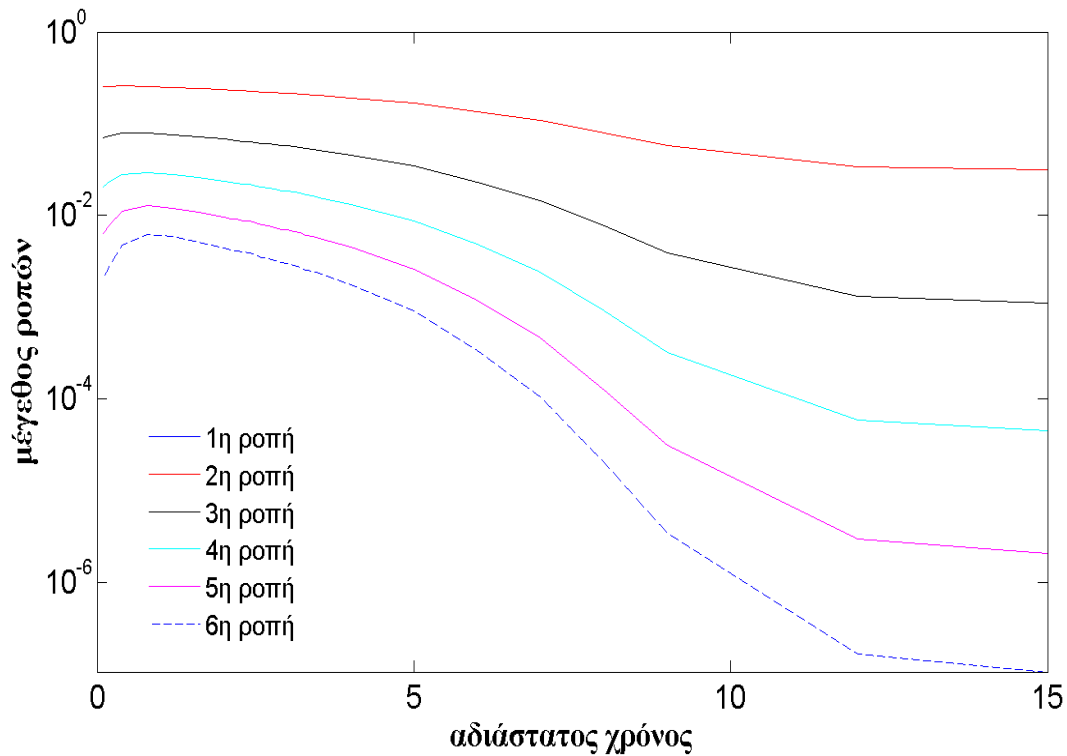
Από το σχήμα 4.2. παρατηρούμε ότι οι ροπές αυξάνονται όσο αυξάνεται ο χρόνος μέχρι ένα σημείο από εκεί και ύστερα παραμένουν σταθερές.

Στο σχήμα 4.3. παρουσιάζεται το ίδιο διάγραμμα αλλά για τιμή παραμέτρου $\rho = 0.15$.



Σχήμα 4.3.: Στιγμιότυπα της στοχαστικής προσομοίωσης για ένα πληθυσμό κυττάρων που χρησιμοποιεί την γενετική αρχιτεκτονική με θετική ανάδραση. Τιμές των παραμέτρων που χρησιμοποιήθηκαν: $\pi = 0.03$, $\delta = 0.05$, $\rho = 0.15$, $L = 2$ και $f = 0.5$. Το μέγεθος του πληθυσμού είναι σταθερό και ίσο με $N_s = 10000$.

Από την εικόνα 4.3. γίνεται κατανοητό πόσο επηρεάζει το σύστημα η παράμετρος ρ . Το ενδοκυτταρικό περιεχόμενο παίρνει μικρότερες τιμές ενώ οι κατανομές των κυττάρων μεγαλύτερες. Στα αρχικά χρονικά, βήματα παρατηρούνται και εδώ απότομες αλλαγές στην κατανομή, οι οποίες φθίνουν όσο εξελίσσεται η προσομοίωση. Το σύστημα φαίνεται ότι περνάει σε μόνιμη κατάσταση σε αδιάστατο χρόνο $\tau=12$. Ο χρόνος που απαιτείται για να φτάσει το σύστημα σε μόνιμη κατάσταση για αυτή την τιμή παραμέτρου είναι σημαντικά μεγαλύτερος σε σχέση με την προηγούμενη περίπτωση. Αυτό συμβαίνει διότι η αρχική κατανομή που δίνουμε έχει μέση τιμή και τυπική απόκλιση, οι οποίες είναι αρκετά διαφορετικές από αυτές του πληθυσμού στη μόνιμη κατάσταση, με αποτέλεσμα να απαιτείται παραπάνω χρόνος για την επίτευξή της. Η στιγμή που το σύστημα περνάει στη μόνιμη κατάσταση φαίνεται και στο ημιλογαριθμικό διάγραμμα 4.4., όπου απεικονίζεται η χρονική εξέλιξη των 6 πρώτων ροπών που υπολογίστηκαν από τις κατανομές του στοχαστικού μοντέλου για τιμή παραμέτρου εξωκυτταρικής συγκέντρωσης $\rho = 0.15$.

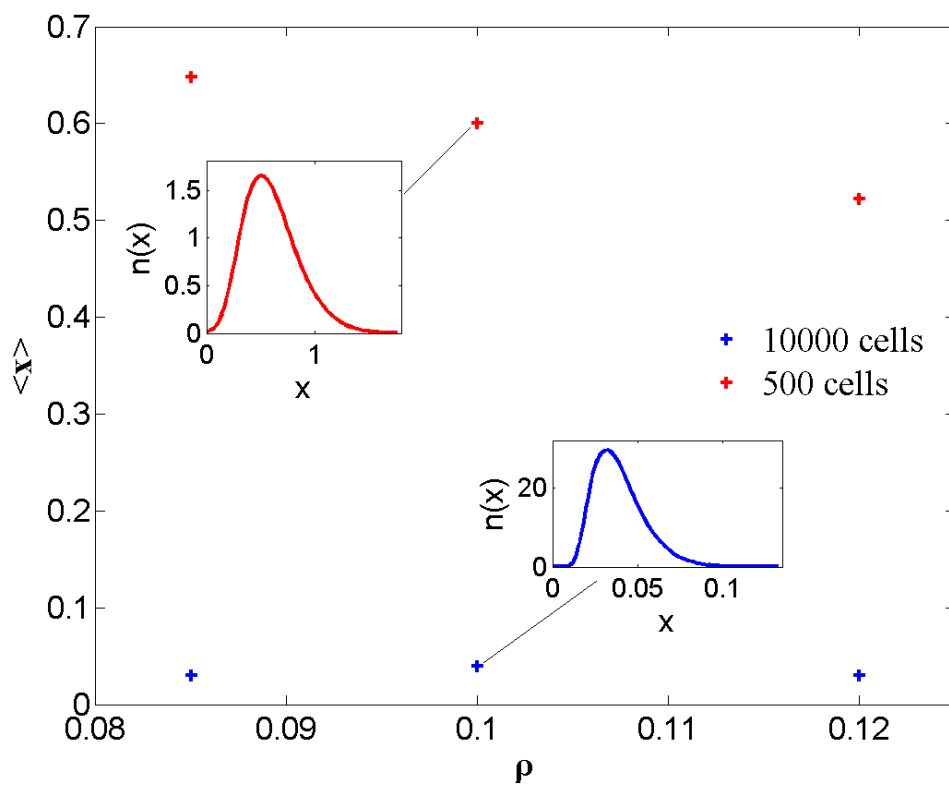


Σχήμα 4.4.: Ημιλογαριθμικό διάγραμμα για τη χρονική εξέλιξη των 6 πρώτων ροπών που προέκυψαν από τις κατανομές των στοχαστικών προσομοιώσεων για τιμές παραμέτρων $\pi = 0.03$, $\delta = 0.05$, $\rho = 0.15$, $L = 2$ και $f = 0.5$ και πληθυσμό κυττάρων $N_s = 10000$.

Στο σχήμα 4.4. παρατηρείται ότι το σύστημα περνάει στη μόνιμη κατάσταση σε αδιάστατο χρόνο $\tau=12$. Από το σχήμα αυτό φαίνεται ότι οι 6 πρώτες ροπές είναι πολύ μικρότερες σε σχέση με αυτές που υπολογίστηκαν από τη μόνιμη κατάσταση για τιμή παραμέτρου $\rho = 0.06$. Η μέση τιμή (δεύτερη ροπή) του πληθυσμού στη μόνιμη κατάσταση είναι $\mu = 0.0311$ και είναι πολύ μικρότερη από τη μέση τιμή $\mu = 0.25$ που έχει η μετατοπισμένη κατανομή Gauss που αποτελεί την αρχική συνθήκη. Από το γεγονός αυτό εξηγείται και το γιατί χρειάζεται παραπάνω χρόνος για να φτάσουμε σε μόνιμη κατάσταση με αυτή την τιμή παραμέτρου.

Στο σχήμα 4.5. παρουσιάζονται τα αποτελέσματα του στοχαστικού μοντέλου για τη μέση τιμή δύο διαφορετικών πληθυσμών κυττάρων, 500 και 10000 κύτταρα, για τιμές παραμέτρων $\pi = 0.03$, $\delta = 0.05$, $L = 2$ και $f = 0.5$ και για $\rho = 0.085, 0.10, 0.12$, με την ίδια αρχική μετατοπισμένη κατανομή Gauss και με τη χρήση 1000 και 100 αντιγράφων, αντίστοιχα. Λαμβάνοντας υπόψη το ντετερμινιστικό διάγραμμα διακλάδωσης του σχήματος 3.3. η μέση τιμή των πληθυσμών για αυτές τις παραμέτρους ρ που χρησιμοποιήθηκαν θα έπρεπε να είναι

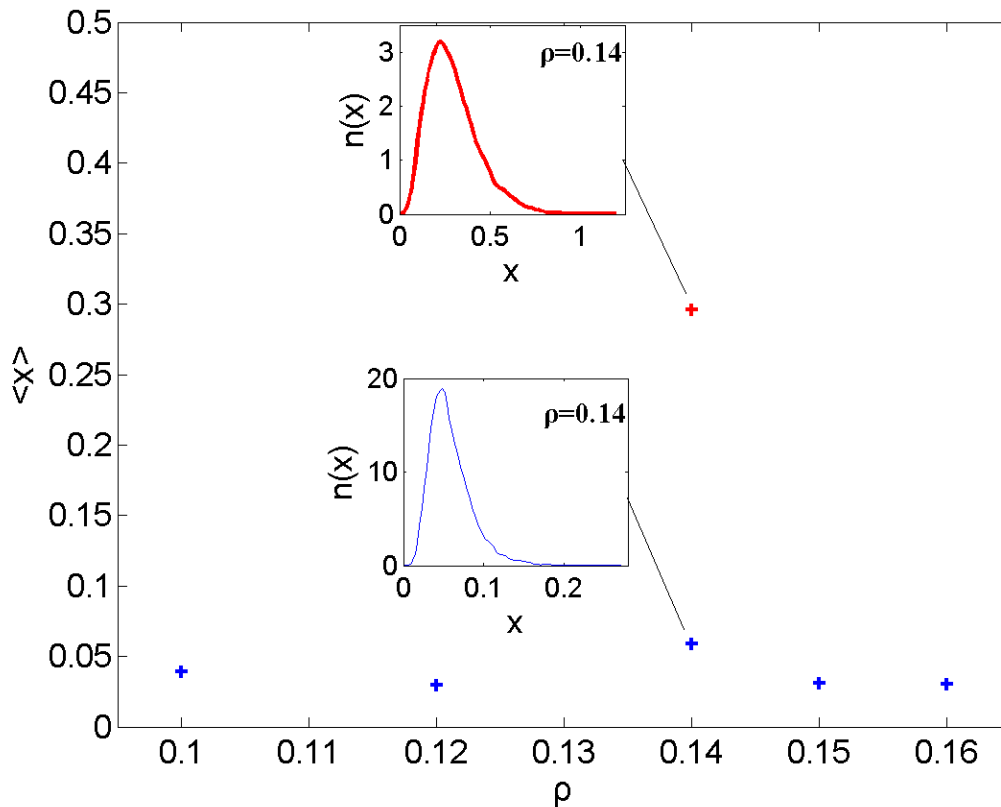
στον κάτω κλάδο ευσταθών καταστάσεων. Παρατηρούμε ότι αυτό συμβαίνει για τον πληθυσμό των 10000 κυττάρων όχι όμως και για τον αντίστοιχο των 500 κυττάρων. Για την επιβεβαίωση των αποτελεσμάτων αυτών πραγματοποιήθηκαν τρεις ανεξάρτητες προσομοιώσεις για τον κάθε πληθυσμό, οι οποίες έδωσαν τα ίδια αποτελέσματα. Το γεγονός αυτό καταδεικνύει την επίδραση του αριθμού των κυττάρων στη συμπεριφορά του κυτταρικού πληθυσμού, αλλά και ότι για αυτές τις τιμές παραμέτρου εξωκυτταρικής συγκέντρωσης IPTG, το σύστημα μπορεί να παρουσιάζει πολλαπλότητα λύσεων (διπλο-ευστάθεια).



Σχήμα 4.5.: Η μέση τιμή του κυτταρικού πληθυσμού για δύο πληθυσμούς κυττάρων, με κόκκινους σταυρούς για 500 κύτταρα και με μπλε σταυρούς για 10000 κύτταρα. Σαν αρχική συνθήκη χρησιμοποιήθηκε μια μετατοπισμένη κατανομή Gauss με μέση τιμή $\mu = 0.5$ και τυπική απόκλιση $\sigma = 0.25$. Οι τιμές των παραμέτρων είναι $\pi = 0.03$, $\delta = 0.05$, $L = 2$ και $f = 0.5$ και $\rho = 0.085, 0.10, 0.12$. Χρησιμοποιήθηκαν 1000 και 100 αντίγραφα για κάθε πληθυσμό, αντίστοιχα.

Η διπλο-ευστάθεια φαίνεται καλύτερα στο σχήμα 4.6. όπου σε μόνιμη κατάσταση για την ίδια τιμή παραμέτρου ρ ($\rho = 0.14$) και τον ίδιο πληθυσμό κυττάρων

($N_s = 10000$), η μέση τιμή του πληθυσμού μπορεί να βρίσκεται είτε στον άνω είτε στον κάτω κλάδο ευσταθών καταστάσεων.



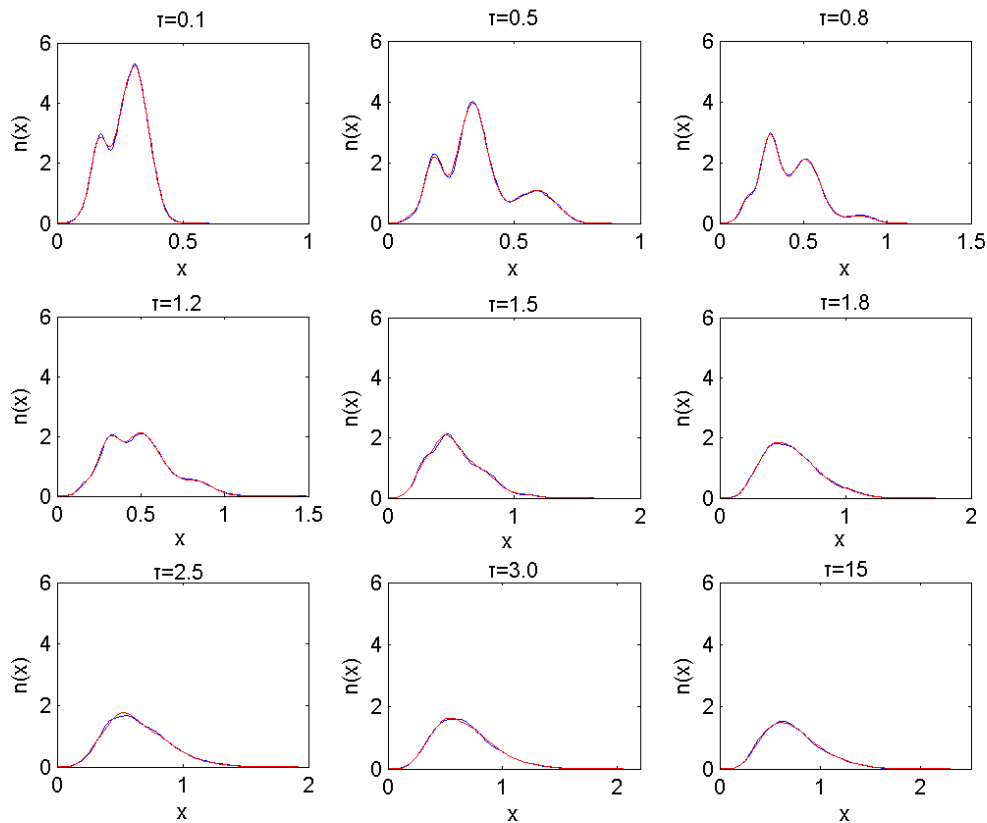
Σχήμα 4.6.: Η μέση τιμή του κυτταρικού πληθυσμού συναρτήσει της παραμέτρου ρ . Για τιμή παραμέτρου $\rho = 0.14$ υπάρχει πολλαπλότητα λύσεων (διπλο-ευστάθεια). Ο κόκκινος σταυρός αποτελεί φαινότυπο του άνω ευσταθή κλάδου καταστάσεων, ενώ ο αντίστοιχος μπλε του κάτω ευσταθή κλάδου καταστάσεων.

Η παρατηρούμενη διπλο-ευστάθεια για αυτή την τιμή παραμέτρου οφείλεται πιθανώς στο ότι το σύστημα βρίσκεται κοντά σε σημείο στροφής. Για να υπάρχει μια πιο σαφής εικόνα για το τι συμβαίνει σε αυτό το σημείο, χρειάζεται, και θα γίνει μελλοντικά, η εφαρμογή της μεθόδου μήκους-τόξου.

4.3. Αποτελέσματα της άρσης με τη μέθοδο της μέγιστης εντροπίας

Οι ροπές που υπολογίζονται από τις κατανομές μικροκαταστάσεων, οι οποίες έχουν προκύψει από το στοχαστικό μοντέλο, τροφοδοτούνται στη μέθοδο της μέγιστης εντροπίας, ώστε να πραγματοποιηθεί η διαδικασία της άρσης και να γίνει

ανακατασκευή των κατανομών. Δουλεύοντας, αρχικά, με τις 6 ροπές που υπολογίζονται από τις κατανομές του σχήματος 4.1. και εφαρμόζοντας τη μέθοδο της μέγιστης εντροπίας κατασκευάζονται οι κατανομές του σχήματος 4.7. που συγκρίνονται με αυτές του σχήματος 4.1.



Σχήμα 4.7.: Ανακατασκευή των μικροκατανομών με τη μέθοδο της μέγιστης εντροπίας (κόκκινες γραμμές) και σύγκρισή τους με τις κατανομές που προέκυψαν από την εφαρμογή του στοχαστικού αλγορίθμου (μπλε γραμμές).

Όπως φαίνεται και από το σχήμα 4.7., με τη μέθοδο της μέγιστης εντροπίας ανακατασκευάζονται δομές μικροκαταστάσεων, οι οποίες είναι αρκετά συνεπείς με αποτελέσματα που προκύπτουν από ευθεία εφαρμογή του αλγορίθμου Monte Carlo που χρησιμοποιήθηκε σε πρώτη φάση. Εκτός από τη σχηματική σύγκριση των δύο μεθόδων με την οποία μπορούμε να κατανοήσουμε το κατά πόσο η μέθοδο άρσης που χρησιμοποιούμε είναι ακριβής ή όχι, ένας άλλος τρόπος προς αυτήν την κατεύθυνση είναι να συγκρίνουμε τα αδρομερή μεγέθη τα οποία έχουν προκύψει με κάθε μία από τις μεθόδους. Στον πίνακα 4.1., αναγράφονται οι 10 πρώτες ροπές που έχουν υπολογιστεί από τις μικροκατανομές των δύο μεθόδων στη μόνιμη κατάσταση σε

χρόνο $\tau=12$. Οι τιμές αυτές είναι πολύ κοντά μεταξύ τους και το σχετικό σφάλμα ικανοποιητικά μικρό, επομένως η άρση με τη μέθοδο της μέγιστης εντροπίας είναι ικανή να παράξει μικροκατανομές συνεπείς με τις μικροκαταστάσεις που προέρχονται από τον αλγόριθμο Monte Carlo και χαρακτηρίζονται από τα ίδια αδρομερή μεγέθη με αυτές.

Πίνακας 4.1.: Οι 10 πρώτες ροπές που έχουν υπολογιστεί από το στοχαστικό μοντέλο και από την άρση με τη μέγιστη εντροπία

Ροπές	Στοχαστικό μοντέλο	Άρση με τη μέθοδο της μέγιστης εντροπίας	% σφάλμα
1 ^η	1.00	1.00	0
2 ^η	0.720853	0.721117	3.67×10^{-4}
3 ^η	0.595281	0.595565	4.75×10^{-4}
4 ^η	0.552973	0.553147	3.14×10^{-4}
5 ^η	0.568533	0.568415	-2.09×10^{-4}
6 ^η	0.637805	0.637005	-1.25×10^{-3}
7 ^η	0.771155	0.768826	-3.02×10^{-3}
8 ^η	0.994388	0.988707	-5.71×10^{-3}
9 ^η	1.35546	1.342551	-9.52×10^{-3}
10 ^η	1.938712	1.910409	-1.46×10^{-2}

4.4. Επιλογή του κατάλληλου αριθμού ροπών

Στη συνέχεια, παρουσιάζεται ο τρόπος με τον οποίο γίνεται η επιλογή του πεπερασμένου αριθμού των ροπών που απαιτούνται για την πραγματοποίηση των προσομοιώσεων με την άρση με τη μέθοδο της μέγιστης εντροπίας.

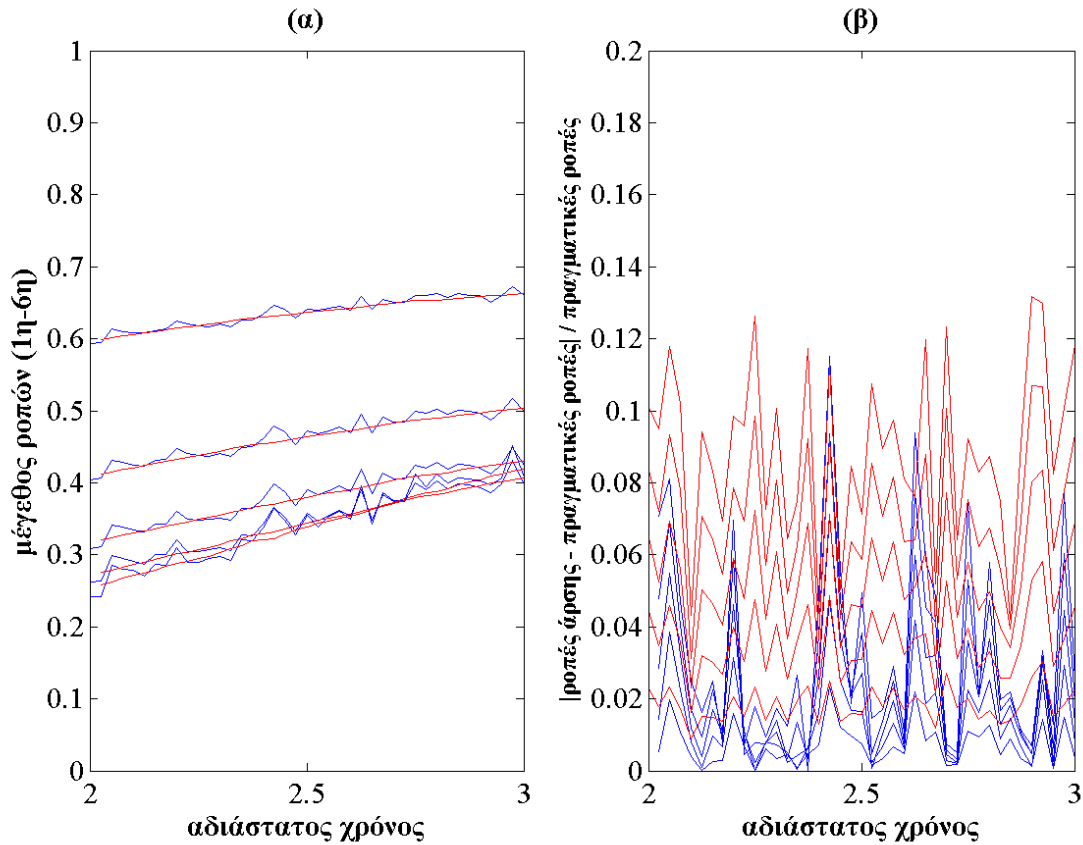
Αν ο αριθμός των κυρίαρχων ροπών έχει επιλεγεί σωστά, η επίδραση της αρχικοποίησης ροπών υψηλότερης τάξης από αυτών που έχουν επιλεγεί πρέπει να είναι αμελητέα και να αποσβέννεται όσο εξελίσσεται η προσομοίωση. Δηλαδή οι ροπές υψηλότερης τάξης πρέπει σε σύντομο χρονικό διάστημα να υποτάσσονται στις κυρίαρχες. Αυτό είναι ένα φαινόμενο που σχετίζεται με το διαχωρισμό των χρονικών κλιμάκων και αναφέρεται ως επούλωση («healing») και ο χρόνος που απαιτείται για να επιτευχθεί ως χρόνος επούλωσης («healing time»). Το ζητούμενο, λοιπόν, είναι η εύρεση του χρόνου αυτού ανάλογα με το πόσες ροπές χρησιμοποιούμε, με σκοπό να

επιλέξουμε τελικά τον αριθμό ροπών που απαιτείται για να είναι αυτός ο χρόνος ικανοποιητικά μικρός και η ανακατασκευή των κατανομών ακριβής [26].

Αρχικά, επιλέγεται ο αριθμός των κυρίαρχων-χαμηλότερης τάξης ροπών να είναι 6 και με βάση αυτές κατασκευάζονται κατανομές μικροκαταστάσεων εφαρμόζοντας την άρση. Από κάθε άρση υπολογίζονται οι ροπές υψηλότερης τάξης και εξετάζουμε τότε οι κυρίαρχες ροπές, οι οποίες εκτιμώνται στο τέλος του Monte Carlo που έχει εκτελεστεί για μικρό χρονικό ορίζοντα, σταματούν πρακτικά να δέχονται επίδραση από την άρση με χρήση και των ροπών υψηλότερης τάξης.

Συνεχίζοντας το τεστ, από τη μία «παγώνουμε» το стоχαστικό αλγόριθμο στο χρονικό διάστημα (2,3) (π.χ. σε αδιάστατο χρόνο $\tau=2$, $\tau=2.1$, κτλ.) και υπολογίζουμε τις τιμές των κυρίαρχων ροπών. Από την άλλη, εφαρμόζουμε την άρση για τις χρονικές αυτές στιγμές με στόχο να υπολογίσουμε τις ροπές υψηλότερης τάξης [26].

Στο σχήμα 4.8. (α) παρουσιάζεται η χρονική εξέλιξη των 6 πρώτων ροπών οι οποίες υπολογίστηκαν από τις κατανομές των δύο διαφορετικών μεθόδων στο χρονικό διάστημα 2-3. Στο 4.8. (β) απεικονίζεται το σχετικό σφάλμα μεταξύ των ροπών αυτών.



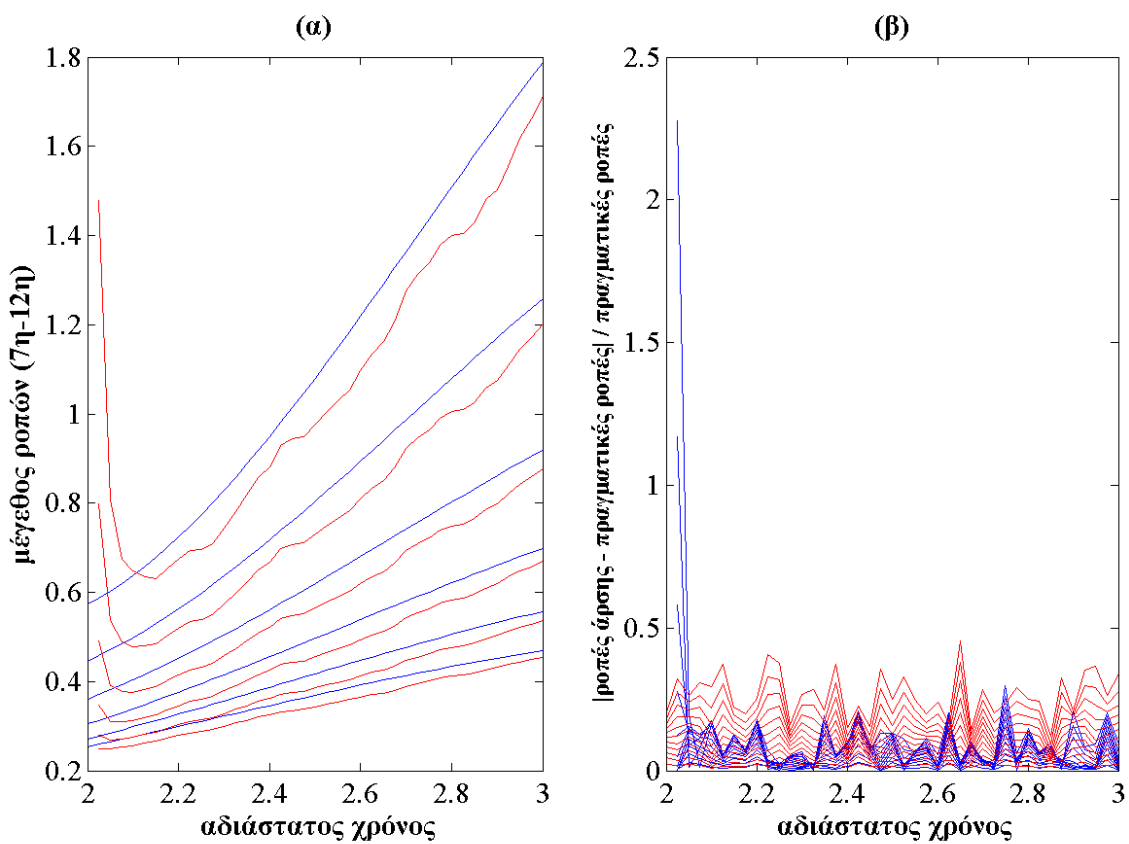
Σχήμα 4.8.: (α) Εξέλιξη των 6 πρώτων ροπών που προέκυψαν από την άμεση εφαρμογή του αλγορίθμου Monte Carlo (πραγματικές ροπές-μπλε γραμμές) και των ροπών που προέκυψαν από την άρση χρησιμοποιώντας τη μέθοδο της μέγιστης εντροπίας (κόκκινες γραμμές). (β) Σχετικό σφάλμα μεταξύ των 6 πραγματικών ροπών και των 6 ροπών που υπολογίστηκαν από την άρση (μπλε γραμμές).

Με κόκκινες γραμμές παριστάνεται η ποσότητα $2\sigma_M / M^{\text{πραγματικές}}$, όπου σ_M είναι η τυπική απόκλιση του στοχαστικού θορύβου των ροπών που εκτιμήθηκαν με τη χρήση 100 αντιγράφων.

Παρατηρούμε στο σχήμα 4.8.(α) ότι οι πραγματικές ροπές είναι πολύ κοντά με τις ροπές άρσης και παρουσιάζουν παρόμοια χρονική εξέλιξη. Η ακρίβεια των ροπών άρσης επιβεβαιώνεται και από το σχήμα 4.8.(β), όπου το σχετικό σφάλμα μεταξύ των τιμών των ροπών των δύο μεθόδων δεν ξεπερνάει το 0.12% και είναι της τάξης της ποσότητας $2\sigma_M / M^{\text{πραγματικές}}$, όπου σ_M είναι η τυπική απόκλιση των ροπών που εκτιμήθηκαν με βάση τα 100 αντίγραφα του στοχαστικού αλγορίθμου. Δηλαδή, το σχετικό σφάλμα είναι μικρότερο από το στατιστικό-στοχαστικό σφάλμα, με αποτέλεσμα οι 6 ροπές να είναι τα κυρίαρχα αδρομερή μεγέθη του συστήματος.

Όσον αφορά τις ροπές άρσης υψηλότερης τάξης, $7^{\text{η}}$ - $12^{\text{η}}$, παρατηρούμε ότι στο σχήμα 4.9. (α) οι τιμές αυτών των ροπών προσεγγίζουν πολύ γρήγορα τις «πραγματικές» τιμές, περίπου σε έναν αδιάστατο χρόνο $t = 0.075$, ενώ στο σχήμα

4.9. (β) που είναι αντίστοιχο με το 4.8. (β), φαίνεται ότι το σχετικό σφάλμα γίνεται γρήγορα μικρό σε σχέση με τις στοχαστικές διακυμάνσεις. Αυτό επιβεβαιώνει την υπόθεση ότι η ανακατασκευή κατανομών μικροκαταστάσεων με την άρση με τη μέθοδο της μέγιστης εντροπίας και τη χρήση 12 ροπών, μπορεί να αναπαραχθεί πολύ καλά με την άρση με τη μέθοδο της μέγιστης εντροπίας και τη χρήση 6 μόνο ροπών. Έτσι, η άρση με τη μέθοδο της μέγιστης εντροπίας που βασίζεται σε 6 μόνο ροπές φαίνεται επαρκής για την εφαρμογή της «ελεύθερης εξισώσεων» μεθοδολογίας για την επιτάχυνση των υπολογισμών.



Σχήμα 4.9.: (α) Εξέλιξη των ροπών υψηλότερης τάξης, 7^{11} - 12^{11} , που έχουν προκύψει από την εφαρμογή του αλγορίθμου Monte Carlo (μπλε γραμμές) και από την άρση με τη μέθοδο της μέγιστης εντροπίας (κόκκινες γραμμές). (β) Σχετικό σφάλμα μεταξύ των 12 πραγματικών ροπών και των 12 ροπών που υπολογίστηκαν από την άρση (μπλε γραμμές). Με κόκκινες γραμμές παριστάνεται η ποσότητα $2\sigma_M / M^{\text{πραγματικές}}$, όπου σ_M είναι η τυπική απόκλιση του στοχαστικού θορύβου των ροπών που εκτιμήθηκαν με τη χρήση 100 αντιγράφων.

Από τα διαγράμματα του σχήματος 4.9. μπορούμε να βρούμε το χρόνο επούλωσης, ο οποίος είναι ίσος με $t=0.075$, κάτι που σημαίνει ότι οι επιδράσεις των ροπών υψηλότερης τάξης αποσβέννυνται σε αδιάστατο χρόνο ίσο με 0.075.

4.5. Παραμετρική ανάλυση

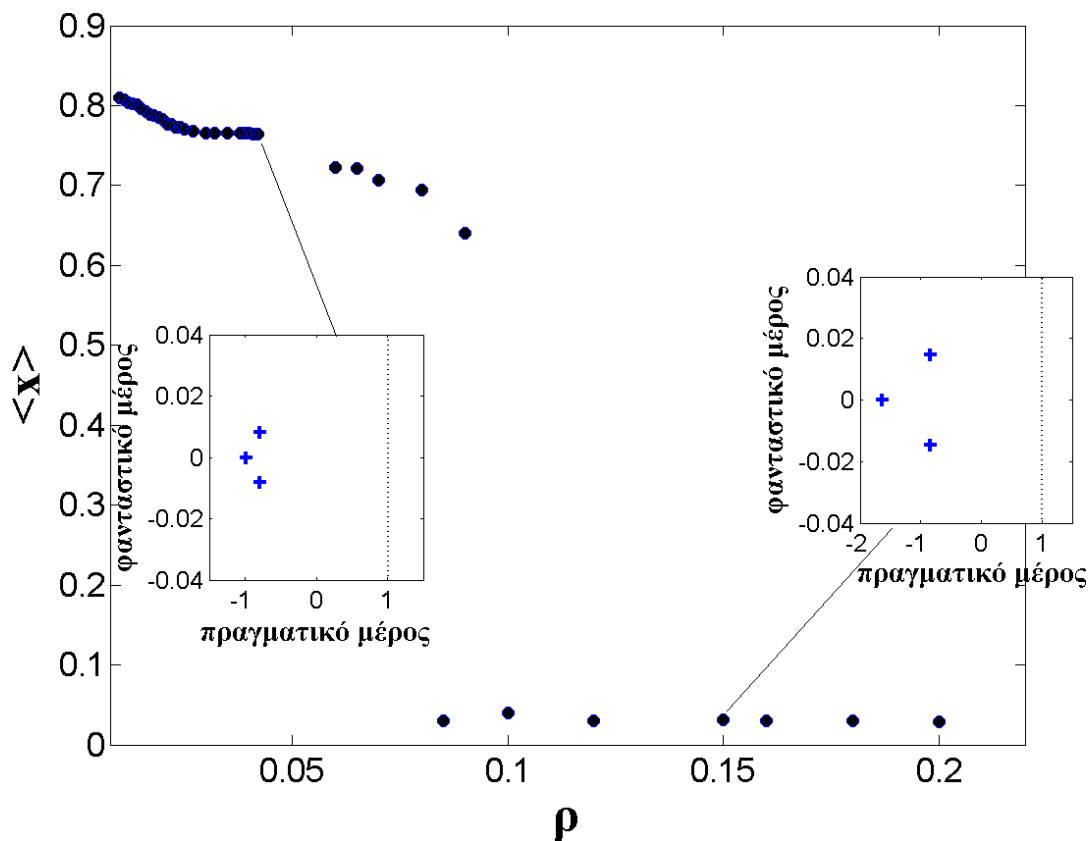
Η παραμετρική ανάλυση σε ισοζύγια κυτταρικών πληθυσμών που φέρουν το ίδιο ρυθμιστικό γενετικό πραγματοποιείται για να διερευνηθεί η επίδραση της ετερογένειας στο φαινότυπο του κυτταρικού πληθυσμού και να εξεταστεί αν το σύστημα εμφανίζει περιοχές αστάθειας ή περιοχές διπλο-ευστάθειας για κάποιες τιμές παραμέτρων.

Όπως αναφέρθηκε και στο κεφάλαιο 3, για την παραμετρική ανάλυση εφαρμόζεται η επαναληπτική μέθοδος Newton-Raphson που στηρίζεται στα αδρομερή μεγέθη, συνεπώς δεν υπάρχουν εξισώσεις σε κλειστή μορφή και αναγκαζόμαστε να καταφύγουμε σε προσεγγιστικές μεθόδους υπολογισμού του Ιακωβιανού πίνακα (πχ. σχήμα πεπερασμένων διαφορών, εξ. 3.13), αφού έχουμε δώσει μικρές διαταραχές στις ροπές. Το διάγραμμα ροής που ακολουθείται για την εκτέλεση της επαναληπτικής μεθόδου είναι αυτό που παρουσιάζεται στο σχήμα 3.4.

Ο χρονικός ορίζοντας για τον οποίο απαιτείται να εκτελέσει ο αλγόριθμος Monte Carlo την προσομοίωση, όταν καλείται μέσα στην επαναληπτική διαδικασία επιλέγεται με βάση το χρόνο επούλωσης, ο οποίος έχει υπολογιστεί νωρίτερα, κατά την εύρεση της επίδρασης ροπών υψηλότερης τάξης στις κυρίαρχες ροπές. Αν ο χρόνος επούλωσης είναι αρκούντως μικρός ώστε σφάλματα από τον υπολογισμό ροπών υψηλότερης τάξης να «θεραπεύονται» γρήγορα, και αρκούντως μεγάλος ώστε τα σφάλματα αυτά να προλαβαίνουν να «ξεχαστούν», τότε ο χρονικός ορίζοντας της στοχαστικής προσομοίωσης πρέπει να είναι $\Delta t > t^{healing}$ [14,25].

Το σχήμα 4.10. απεικονίζει τα αποτελέσματα που προέκυψαν με την περιγραφείσα μέθοδο. Σε αυτό εξετάζεται η ασυμπτωτική συμπεριφορά του μέσου κυτταρικού πληθυσμού συναρτήσει της παραμέτρου ρ , η οποία όπως αναφέρθηκε εκφράζει την εξωκυτταρική συγκέντρωση του παράγοντα IPTG. Όπως φαίνεται, κινούμαστε στον πάνω και στον κάτω κλάδο ευστάθειας. Όπως φαίνεται από το σχήμα, για κάποιες τιμές παραμέτρων το σύστημα εμφανίζει διπλο-ευστάθεια, κάτι που παρατηρήθηκε και παραπάνω. Ο υπολογισμός φαινοτύπων που ανήκουν στον κλάδο ασταθών καταστάσεων (σχήμα 3.3.-γραμμή ντετερμινιστικού μοντέλου) δεν

έχει μέχρι στιγμής επιτευχθεί. Για να κινηθούμε σε αυτόν τον κλάδο η μέθοδος Newton-Raphson δεν αρκεί και πρέπει να εφαρμοστεί η μέθοδος μήκους τόξου (arc-length continuation). Οι υπολογισμοί με τιμή παραμέτρου $\rho=0.15$, που δίνει φαινοτύπους στον κάτω ευσταθή κλάδο, έγιναν ώστε να φτάσει το σύστημα σε μόνιμη κατάσταση και να υπολογιστούν επαναληπτικά φαινότυποι και σε αυτόν τον κλάδο. Στο σχήμα 4.10. φαίνονται επίσης και οι ιδιοτιμές του Ιακωβιανού πίνακα της επαναληπτικής διαδικασίας, για δύο καταστάσεις που ανήκουν στον άνω και στον κάτω ευσταθή κλάδο, αντίστοιχα. Σε αυτά τα ένθετα σχήματα παρατηρούμε ότι το πραγματικό μέρος των εν λόγω ιδιοτιμών είναι μικρότερο της μονάδας [9], δηλαδή επιβεβαιώνεται ότι σε αυτές τις τιμές παραμέτρων οι προκύπτουσες καταστάσεις είναι ευσταθείς. Στα σχήματα αυτά παρουσιάζονται οι ιδιοτιμές των 3 μεγαλύτερων ροπών.



Σχήμα 4.10.: Μεταβολή του μέσου κυτταρικού πληθυσμού συναρτήσει της αδιάστατης παραμέτρου ρ , για τον άνω κλάδο ευστάθειας. Οι υπολογισμοί πραγματοποιήθηκαν με τη μέθοδο Newton-Raphson, κατάλληλα διαμορφωμένη ώστε να περικλείει την αδρομερή περιγραφή του συστήματος.

Κεφάλαιο 5: Συμπεράσματα

Ο βασικός στόχος της παρούσας εργασίας είναι η ανάπτυξη μιας υπολογιστικής μεθοδολογίας που επιτρέπει την αποτελεσματική μελέτη της μακροσκοπικής συμπεριφοράς συστημάτων, η περιγραφή των οποίων είναι διαθέσιμη σε επίπεδο μικρο-κλίμακας, χωρίς την ανάγκη εξαγωγής ενός μοντέλου διαφορικών εξισώσεων που να περιγράφει τη δυναμική συμπεριφορά σε επίπεδο μακρο-κλίμακας.

Σε αυτά τα πλαίσια λοιπόν εξετάστηκε η δυναμική συμπεριφορά κυτταρικών πληθυσμών που υπακούουν σε μαθηματικά μοντέλα ΙΚΠ, όπως αυτά περιγράφηκαν στο 1^ο κεφάλαιο.

Αρχικά έγιναν υπολογισμοί με την εφαρμογή στοχαστικού αλγορίθμου kinetic Monte Carlo και παρατηρήθηκε η μεταβολή του ενδοκυτταρικού περιεχομένου σε σχέση με το χρόνο. Στα πρώτα χρονικά στάδια της στοχαστικής προσομοίωσης παρατηρούμε απότομες αλλαγές στη μεταβολή του ενδοκυτταρικού περιεχομένου, οι οποίες όσο εξελίσσεται η προσομοίωση χάνονται και το σύστημα οδηγείται σε μόνιμη κατάσταση. Για τη λήψη όσο το δυνατόν πιο αντιπροσωπευτικών αποτελεσμάτων (μείωση στοχαστικού σφάλματος) στις προσομοιώσεις κάθε χρονικής στιγμής γίνεται χρήση 100 αντιγράφων για πληθυσμό $N_s = 10000$ κυττάρων και 1000 αντιγράφων για πληθυσμούς $N_s = 500, 1000$ κυττάρων και η τελική κατανομή υπολογίζεται από το μέσο όρο αυτών.

Οι προσομοιώσεις με αλγόριθμο Monte Carlo δίνουν ακριβή αποτελέσματα, όμως απαιτούν μεγάλο υπολογιστικό χρόνο για την εκτέλεσή τους. Για το λόγο αυτό πραγματοποιήθηκαν επιπλέον υπολογισμοί που βασίζονται στα αδρομερή μεγέθη τα οποία περιγράφουν τις κατανομές που προέκυψαν από το στοχαστικό μοντέλο. Η αδρομερής μέθοδος που χρησιμοποιήθηκε είναι η άρση με τη μέθοδο της μέγιστης εντροπίας και εντάσσεται στην «ελεύθερη εξισώσεων» μεθοδολογία. Για τις προσομοιώσεις με αυτή τη μέθοδο υπολογίστηκαν οι ροπές των κατανομών του στοχαστικού μοντέλου, βάση των οποίων η μέθοδος της μέγιστης εντροπίας κάνει ανακατασκευή των μικροκαταστάσεων.

Οι νέες κατανομές οι οποίες προέκυψαν από αυτή τη μέθοδο φαίνεται ότι περιγράφουν αρκετά καλά τη δυναμική του κυτταρικού πληθυσμού αφού έχουν πολύ μικρές διαφορές με τις κατανομές του στοχαστικού μοντέλου, ενώ και οι ροπές που υπολογίστηκαν από αυτές διαφέρουν ελάχιστα από τις πραγματικές ροπές (σφάλματα της τάξης του 0.12%).

Για την εύρεση του πεπερασμένου αριθμού ροπών που πρέπει να χρησιμοποιηθεί ώστε οι νέες κατανομές να είναι συνεπείς με αυτές του αλγορίθμου Monte Carlo, πραγματοποιήθηκαν υπολογισμοί που περιλάμβαναν τον υπολογισμό ροπών υψηλότερης τάξης ώστε να ευρεθεί σε πόσο χρόνο αυτές προσεγγίζουν τα επίπεδα των κυρίαρχων ροπών, να ευρεθεί δηλαδή ο λεγόμενος χρόνος επούλωσης (healing time). Αυτός ο χρόνος αν είναι πολύ μικρός δείχνει ότι ο αριθμός ροπών που χρησιμοποιούμε μπορεί να περιγράψει το σύστημα επαρκώς με τις αντίστοιχες διαφορικές εξισώσεις (που δεν τις έχουμε σε κλειστή μορφή), ενώ αν είναι μεγάλος σημαίνει ότι ενδεχομένως πρέπει να αυξηθούν οι χαμηλότερης τάξης ροπές που χρησιμοποιούμε για την προσομοίωση της δυναμικής συμπεριφοράς του συστήματος, ώστε να έχουμε ακριβέστερα αποτελέσματα. Με τον τρόπο αυτό, υπολογίστηκε χρόνος επούλωσης, $t^{healing} = 0.075 \text{ sec}$, που κρίνεται αρκούντως μικρός και ο αριθμός των 6 ροπών που χρησιμοποιούμε κρίνεται ικανοποιητικός.

Τέλος, πραγματοποιήθηκε παραμετρική ανάλυση, με στόχο να διερευνηθεί η επίδραση της ετερογένειας στο φαινότυπο του κυτταρικού πληθυσμού. Σε παλαιότερες εργασίες [8,17,27] έχει διαπιστωθεί ότι υπάρχει περιοχή τιμών εξωκυτταρικής συγκέντρωσης του ενεργοποιητή IPTG του δικτύου, για τις οποίες παρατηρείται διπλο-ευστάθεια, κάτι που παρατηρήθηκε και στην παρούσα διατριβή (σχήματα 4.5., 4.5., 4.10.). Η περιοχή αυτή διπλο-ευστάθειας παρατηρείται και στο επίπεδο του ενός κυττάρου, το εύρος της όμως περιορίζεται σε επίπεδο κυτταρικού πληθυσμού λόγω της ετερογένειας. Στην παρούσα εργασία, ξεκινώντας αρχικά με τιμή παραμέτρου $\rho = 0.06$ για τη συγκέντρωση του IPTG, πραγματοποιήθηκε ο υπολογισμός φαινοτύπων που ανήκουν στον άνω κλάδο ευσταθών καταστάσεων. Σε επόμενο βήμα, έγιναν υπολογισμοί μόνιμης κατάστασης με τιμή παραμέτρου $\rho = 0.15$, ώστε να υπολογιστούν αντίστοιχες καταστάσεις στον κάτω κλάδο ευστάθειας. Το γεγονός ότι βρισκόμαστε σε ευσταθείς κλάδους επιβεβαιώνεται και από τις ιδιοτιμές του Ιακωβιανού πίνακα, οι οποίες είναι μικρότερες της μονάδας. Μελλοντικά είναι προγραμματισμένη και η εφαρμογή της μεθόδου μήκους-τόξου για τον υπολογισμό ασταθών καταστάσεων και την εύρεση σημείων στροφής, με τον ίδιο τρόπο που εφαρμόστηκε η Newton-Raphson.

Η χρησιμοποιούμενη επαναληπτική μέθοδος καλεί την άρση με τη μέγιστη εντροπία, όπως έχει περιγραφεί στο σχήμα 3.4. Επίσης καθότι δεν υπάρχουν εξισώσεις σε αναλυτική μορφή, ο υπολογισμός του Ιακωβιανού πίνακα γίνεται

προσεγγιστικά χρησιμοποιώντας σχήματα πεπερασμένων διαφορών, για τα οποία δίνονται μικρές διαταραχές στις ροπές. Ωστόσο, η μέθοδος της μέγιστης εντροπίας δυσκολεύεται αρκετά να συγκλίνει όταν χρησιμοποιεί τις διαταραγμένες ροπές. Η δυσκολία σύγκλισης της άρσης με τη μέγιστη εντροπία έγκειται στην ευαισθησία υπολογισμού των πολλαπλασιαστών Lagrange και παρόλο που εφαρμόστηκαν μέθοδοι με σκοπό την καλύτερη αρχική εκτίμησή τους [5,22], δεν μπόρεσε να γίνει σταθεροποίηση της μεθόδου.

Η μέθοδος της μέγιστης εντροπίας μας απαλλάσσει από το μεγάλο υπολογιστικό κόστος που απαιτούν οι στοχαστικές προσομοιώσεις, όπως ο αλγόριθμος Monte Carlo που χρησιμοποιήθηκε και αυτό γιατί αντί να λειτουργεί με ολόκληρο τον πληθυσμό κυττάρων (500,1000,10000 κύτταρα), είναι σε θέση να κατασκευάζει ακριβείς κατανομές μικροκαταστάσεων χρησιμοποιώντας λίγους μόνο αριθμούς (4,6,8,10,20,κτλ), δηλαδή τις κυρίαρχες πρώτες ροπές που χαρακτηρίζουν την κατανομή.

Βιβλιογραφία

1. Beckwith, J.R. and D.Zipser, *The Lactose Operon*. Cold Spring Harbor, ed. C.S.H. Laboratory, 1970, NY.
2. Blake, W.J., M. Kaern, C.R. Cantor and J.J. Collins, *Noise in eukaryotic gene expression*. Nature, 2003. **422**: p. 633-637.
3. Eakman, J.M., A.G. Fredrickson, H.M. Tsuchyia, *Statistics and dynamics of microbial cell populations*. Chemical Engineering Progress, 1966. **62**: p. 37-49.
4. Elowitz, M.B., A.J. Levine, E.D. Siggia, P.S. Swain, *Stochastic gene expression in a single cell*. Science, 2002. **297**: p. 1183-1186.
5. Erdogmus, D., K.E. Hild, Y.N. Rao and J.C. Principe, *Minimax Mutual Information Approach for Independent Component Analysis*. Neural Computation, 2011. **16**: p. 1235-1252.
6. Fredrickson, A.G., D. Ramkrishna, H.M. Tsuchyia, *Statistics and dynamics of prokaryotic cell populations*. Mathematical Biosciences, 1967. **1**: p. 327-374.
7. Gear, C.W., I.G. Kevrekidis, C. Theodoropoulos, *'Coarse' integration/bifurcation analysis via microscopic simulators: micro-Galerkin methods*. Computers and Chemical Engineering, 2002. **26**: p. 941-963.
8. Kavousanakis, M.E., N.V. Mantzaris, A.G. Boudouvis, *A novel free boundary algorithm for the solution of cell population balance models*. Chemical Engineering Science, 2009. **64**: p. 4247-4261.
9. Kelley, C.T., I.G. Kevrekidis and L. Qiao, *Newton-Krylov solvers for timesteppers*. SIAM Journal of Scientific Computing, 2004. **19**: p. 1188-1209.

10. Kepler, T.B. and T.C. Elston, *Stochasticity in transcriptional regulation: origins, consequences and mathematical representations*. Biophysical Journal, 2001. **81**: p. 3116-3136.
11. Kevrekidis, I.G., C. Gear, J. Hyman, P. Kevrekidis, O. Runb ö rg and K. Theodoropoulos, *Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis*. Communications in Mathematical Science, 2003. **1**: p. 715-762.
12. Kevrekidis, I.G., C.W. Gear and G. Hummer, *Equation-free: the computer-assisted analysis of complex, multiscale systems*. AIChE Journal, 2004. **50**(7): p. 1346-1354.
13. Lin, Y., K. Lee, T. Masoukas, *Solution of the population balance equation using constant-number Monte Carlo*. Chemical Engineering Science, 2002. **57**: p. 2241-2252.
14. Makeev, A.G., D. Maroudas and I.G. Kevrekidis, *Coarse Stability and Bifurcation Analysis Using Stochastic Simulators: Kinetic Monte Carlo Examples*. Journal of Computational Physics, 2002. **117**: p. 8229-8240.
15. Mantzaris, N.V., *A cell population balance model describing positive feedback loop expression dynamics*. Computers and Chemical Engineering, 2005. **29**: p. 897-909.
16. Mantzaris, N.V., *Stochastic and deterministic simulations of heterogeneous cell population dynamics*. Journal of Theoretical Biology, 2006. **241**: p. 690-706.
17. Mantzaris, N.V., *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, 2007. **92**: p. 4271-4288.

18. Mead, L.R. and N. Papanikolaou, *Maximum entropy in the problem of moments*. Journal of Mathematical Physics, 1984. **25**: p. 2404-2417.
19. Miller, J.H. and W.S. Reznikoff, *The Operon*. Cold Spring Harbor, ed. C.S.H. Laboratory, 1978, NY.
20. Ramkrishna, D., *Population Balances: Theory and applications to particulate systems in engineering*. Academic Press, 2000. San Diego, CA.
21. Saad, T., *The Maximum Entropy Method for Reconstructing Density Distributions*. Post Doctoral Fellow, Institute for Clean Energy and Security, University of Utah, 2011.
22. Siettos, C.I., A. Armaou, A.G. Makeev and I.G. Kevrekidis, *Microscopic/stochastic timesteppers and coarse control: a kinetic Monte Carlo example*. AIChE Journal, 2003. **49**: p. 1922-1926.
23. Smith, M. and T. Matsoukas, *Constant-number Monte Carlo simulation of population balances*. Chemical Engineering Science, 1998. **53**: p. 1777-1786.
24. Tsuchiya, H.M., A.G. Fredrickson and R. Aris, *Dynamics of microbial cell populations*. Advances in Chemical Engineering, 1966. **6**: p. 125-206.
25. Van Leemput, P., K.W.A. Lust, I.G. Kevrekidis, *Coarse-grained numerical bifurcation analysis of lattice Boltzmann models*. Physica D, 2005. **210**: p. 58-76.
26. Zou, Y., M.E. Kavousanakis, I.G. Kevrekidis, R.O. Fox, *Coarse-grained computations for particle coagulation and sintering processes by linking Quadrature Method of Moments with Monte-Carlo*. Journal of Computational Physics, 2010. **229**: p. 5299-5314.
27. Μ.Ε. Καβουσανάκης, *Αδρομερής Προσομοίωση Περίπλοκων Φυσικοχημικών και Βιολογικών Συστημάτων με Μεθόδους Ανάλυσης Πολλαπλών Κλιμάκων, Διακλάδωσης και Ομοιοτυπίας*. Διδακτορική Διατριβή, ΕΜΠ, Αθήνα 2007.