



## Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών  
Τομέας Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής  
Εργαστήριο Σχεδίασης Μικροηλεκτρονικών Κυκλωμάτων

# Αναλογικά Ολοκληρωμένα Κυκλώματα Ταξινομητών Χαμηλής Κατανάλωσης Βασισμένων στο Gaussian Mixture Model με Ψηφιακές Εισόδους και Εξόδους

## Διπλωματική Εργασία

του

Στρακόση Λάζαρου

Επιβλέπων: Παύλος Π. Σωτηριάδης  
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2022



## Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών  
Τομέας Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής  
Εργαστήριο Σχεδίασης Μικροηλεκτρονικών Κυκλωμάτων

### Αναλογικά Ολοκληρωμένα Κυκλώματα Ταξινομητών Χαμηλής Κατανάλωσης Βασισμένων στο Gaussian Mixture Model με Ψηφιακές Εισόδους και Εξόδους

#### Διπλωματική Εργασία

του

Στρακόση Λάζαρου

Επιβλέπων: Παύλος Π. Σωτηριάδης  
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 5<sup>η</sup> Οκτωβρίου  
2022:

.....  
Παύλος Π. Σωτηριάδης    Νεκτάριος Κοζύρης    Ευάγγελος Χριστοφόρου  
Καθηγητής                    Καθηγητής                    Καθηγητής  
Ε.Μ.Π.                            Ε.Μ.Π.                            Ε.Μ.Π.

Αθήνα, Οκτώβριος 2022

.....  
Στρακόση Λάζαρος

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών, Ε.Μ.Π.

Copyright © Στρακόση Λάζαρος, 2022.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.



# Περίληψη

Στην παρούσα διπλωματική παρουσιάζουμε τη σχεδίαση ενός αναλογικού ταξινομητή με ψηφιακή είσοδο και έξοδο, βασισμένου στη Γκαουσιανή συνάρτηση. Συγκεκριμένα θα παρουσιαστούν ταξινομητές βασισμένοι στο Γκαουσιανό Μοντέλο Μίξης (Gaussian Mixture Model), αλλά και στο απλό Μπαγιεσιανό. Ειδικότερα παρουσιάζουμε ένα κύκλωμα που καταφέρνει να σχεδιάσει Γκαουσιανές κατανομές ανάλογα με τις τιμές τάσεων και ρευμάτων που θα του προσφερθούν. Η υλοποίηση έχει πολύ χαμηλή κατανάλωση καθώς χρησιμοποιούνται τάσεις πόλωσης από 0.5 V έως και 0.6 V και ρεύματα πόλωσης της τάξεως των  $nA$  ή και  $pA$  σε κάποιες περιπτώσεις. Η εκπαίδευση του μοντέλου γίνεται καθαρά με τη βοήθεια της γλώσσας python. Οι αρχιτεκτονικές που παρουσιάζονται επιβεβαιώνονται χρησιμοποιώντας τρία σετ δεδομένων. Η υλοποίηση και προσομοίωση του κυκλώματος έγινε με τη βοήθεια του προγράμματος σχεδίασης Cadence IC Suite σε τεχνολογία TSMC 90 nm CMOS process.

**Λέξεις Κλειδιά:** Γκαουσιανό Μοντέλο Μίξης, Μπαγιεσιανός Ταξινομητής, Αναλογικές αρχιτεκτονικές, Χαμηλής κατανάλωσης μετατροπέας ψηφιακής εισόδου σε αναλογική έξοδο, ταξινόμηση πάνω σε τσιπ, σχέδια πολύ χαμηλής κατανάλωσης, κυκλώματα Γκαουσιανής κατανομής, περιοχή υπο-κατωφλίου, πλήρως προσαρμόσιμη υλοποίηση, Winner-Take-All κυκλώματα



# Abstract

In this diploma thesis, the design of an analog classifier with digital inputs and digital outputs, based on the Gaussian function, will be presented. Specifically, the proposed classifiers' basis is the Gaussian Mixture Model, as well as the simple Bayesian Model. In particular, the implemented circuit is able to produce Gaussian distributions, depending on the values of the voltages and currents that control its parameters. The power consumption is very low, as the supply voltage is  $0.5V$  or  $0.6V$ , relying on the application, and the bias currents are from some  $pA$  to some  $nA$ . The models are trained using the program language Python. The proposed architectures are validated with the use of three real-world datasets. The implementations and simulations of these architectures is done using the design tool Cadence IC Suite in TSMC 90 nm CMOS process.

**Keywords:** Gaussian Mixture Model, Bayesian Classifier, analog hardware architecture, low power digital to analog converter, on-chip classification, Ultra-low power design, Gaussian function circuit, sub-threshold region, fully tunable implementation, Winner-Take-All circuit.





# Ευχαριστίες

Στο σημείο αυτό θα ήθελα να ευχαριστήσω την οικογένεια μου, τους φίλους μου και όλους τους ανθρώπους που στάθηκαν δίπλα μου και με στήριξαν σε αυτά τα βήματα της ζωής μου.

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου, κ. Παύλο Πέτρο Σωτηριάδη για την πολύτιμη καθοδήγηση του, τις συμβουλές του καθώς και το πραγματικό ενδιαφέρον που επέδειξε για μένα ως φοιτητή και για το αντικείμενο της Διπλωματικής μου Εργασίας. Μου δόθηκε έτσι η πολύτιμη ευκαιρία να ασχοληθώ με ένα ιδιαίτερα ενδιαφέρον ερευνητικό θέμα σε ένα πολύ υποστηρικτικό περιβάλλον.

Ακόμη, θα ήθελα να ευχαριστήσω ιδιαίτερα τον υποψήφιο διδάκτορα Βασίλειο Αλιμήση και τον διπλωματούχο Γεώργιο Γέννη για την ανεκτίμητη συνεισφορά τους. Δημιούργησαν μια πολύ δυνατή ομάδα που κατάφερε να παράξει πολλές καλές ιδέες. Η παρούσα Διπλωματική Εργασία είναι προϊόν της συνεργασίας αυτής, με πολύωρες συζητήσεις και καθημερινή συναναστροφή. Οπότε θα ήθελα να τους ευχαριστήσω θερμά για την συνεργασία μας όλο αυτό το διάστημα.

Στρακόση Λάζαρος,  
Σεπτέμβριος 2022



# Περιεχόμενα

Περίληψη	5
Abstract	7
Ευχαριστίες	9
Ευρετήριο Εικόνων	13
Κατάλογος Πινάκων	15
<b>1 Εισαγωγή</b>	<b>17</b>
1.1 Εισαγωγή στην Μηχανική Μάθηση . . . . .	17
1.2 Χρησιμότητα αναλογικής υλοποίησης εφαρμογών Μηχανικής Μάθη- σης . . . . .	18
1.3 Αναλογικές υλοποιήσεις βασισμένες στην Γκαουσιανή κατανομή	20
<b>2 Υπόβαθρο</b>	<b>23</b>
2.1 Gaussian Mixture Model . . . . .	23
2.1.1 Υπόβαθρο στη Θεωρία των Πιθανοτήτων . . . . .	23
2.1.2 Μπαγιεσιανό Μοντέλο . . . . .	24
2.1.3 K-Means . . . . .	27
2.1.4 Το Γκαουσιανό Μοντέλο Μίξης . . . . .	28
2.2 MOSFET transistor . . . . .	30
2.2.1 Βασική λειτουργία μεγάλου σήματος MOSFET . . . . .	31
2.2.2 Subthreshold region . . . . .	33
2.3 Το MOSFET ως ψηφιακό στοιχείο . . . . .	38
<b>3 Βασικά Κυκλώματα</b>	<b>41</b>
3.1 Bump Circuits . . . . .	41
3.1.1 Delbruck's bump circuit . . . . .	41

3.2	Winner Take All circuit (WTA)	44
3.3	Improved WTA	46
3.4	WTA bump circuit	48
3.5	Ελεγχόμενος καθρέφτης ρεύματος	50
3.6	Προτεινόμενη υλοποίηση	52
3.6.1	Προτεινόμενο bump	53
3.6.2	Πολυδιάστατη υλοποίηση του ταξινομητή	60
<b>4</b>	<b>Αξιοποίηση του κυκλώματος σε πραγματικές εφαρμογές</b>	<b>63</b>
4.1	Εκπαίδευση του μοντέλου	63
4.2	Υλοποίηση ταξινομητή για σκοπούς ανίχνευσης κατηγοριών θυρεοειδή	64
4.3	Υλοποίηση ταξινομητή για σκοπούς ανίχνευσης ειδικών προβλημάτων κινητήρα αυτοκινήτου (CWRU)	67
4.4	Υλοποίηση ταξινομητή για σκοπούς ανίχνευσης ομιλίας έναντι θορύβου speech over noise	68
<b>5</b>	<b>Συμπεράσματα και Μελλοντική Δουλειά</b>	<b>75</b>

# Ευρετήριο Εικόνων

1.3.1 Normal distribution . . . . .	20
2.1.1 1 dimensional comparator for Gaussian Distributions, Bayesian classifier . . . . .	26
2.2.1 Η δομή του nMOS transistor [1] . . . . .	31
2.2.2 Τα κανάλια για διάφορες τιμές των $V_{DS}$ σχετικά με το $V_{ov}$ . . . . .	34
2.2.3 Εξάρτηση ρεύματος υποδοχής $I_D$ σχετικά με το $V_{DS}$ και το $V_{ov}$ . . . . .	35
2.2.4 Χαρακτηριστικές υποκατωφλίου για MOS [2] . . . . .	36
2.2.5 Ρεύμα απαγωγού συναρτήσει των $V_{GS}$ , $V_{DS}$ . . . . .	37
3.1.1 Simple bump circuit as proposed by Delbruck [3] . . . . .	42
3.1.2 nmos current correlator, where S is the strength ratio . . . . .	43
3.2.1 nMOS WTA με 2 νευρώνες . . . . .	45
3.2.2 Ρεύματα εξόδου από τους νευρώνες του WTA σε σχέση με τα ρεύματα εισόδων τους, αλλάζοντας το ρεύμα εισόδου $I_1$ όπως φαίνεται . . . . .	47
3.3.1 Improved Winner Take All circuit [4] . . . . .	47
3.3.2 Αποτελέσματα για το ρεύμα εξόδου του improved WTA κυκλώματος σε σύγκριση και με ένα αντίστοιχα απλό WTA . . . . .	48
3.4.1 WTA bump circuit . . . . .	49
3.5.1 Ελεγχόμενος καθρέφτης ρεύματος . . . . .	51
3.5.2 Έξοδος ενός 8bit controllable current mirror που δέχεται όλους τους αριθμούς από 0 έως και 255 ( $I(1)$ ) σε αναλογία με έναν που δέχεται μόνο την τιμή 60 $I(2)$ . . . . .	53
3.6.3 Πειραματικά δεδομένα από την προτεινόμενη υλοποίηση του κυκλώματος . . . . .	61
3.6.4 Πολυδιάστατη Γκαουσιανή, παράδειγμα 5 διαστάσεων . . . . .	62
4.2.1 Αποτελέσματα από την ανάλυση Monte Carlo για την ανίχνευση θυρεοειδή . . . . .	66

4.3.1 Αποτελέσματα από την ανάλυση Monte Carlo για το CWRU dataset . . . . .	69
4.4.1 Αποτελέσματα από την ανάλυση Monte Carlo για το speech over noise dataset . . . . .	72

# Κατάλογος Πινάκων

3.1	Διαστάσεις τρανζίστορ (Σχήμα 3.4.1). . . . .	50
3.2	Διαστάσεις τρανζίστορ (Σχήμα 3.5.1). . . . .	52
3.3	Διαστάσεις τρανζίστορ (Σχήμα 3.5.1). . . . .	60
4.1	Accuracy . . . . .	66
4.2	Accuracy . . . . .	68
4.3	Accuracy . . . . .	71
4.4	Extracted features [5] . . . . .	73





# Κεφάλαιο 1

## Εισαγωγή

### 1.1 Εισαγωγή στην Μηχανική Μάθηση

Ζούμε στην εποχή του big data, όπως αναφέρεται και στο [6], στην διετία μεταξύ 2013 και 2015 ο όγκος δεδομένων ήταν μεγαλύτερος από ότι είχε δημιουργηθεί στην μέχρι τότε ανθρώπινη ιστορία. Δημιουργείται λοιπόν, η ανάγκη για την επεξεργασία αυτών των δεδομένων ώστε να αποσπαστεί κάποια σημαντική πληροφορία. Τον ρόλο αυτό έρχεται να παίξει η μηχανική μάθηση. Η μηχανική μάθηση είναι ένα αντικείμενο το οποίο έχει υποστεί τεράστια εξέλιξη τα τελευταία χρόνια, δεδομένων και των δυνατοτήτων της [7]. Πρακτικά είναι η διερεύνηση μεθόδων βασισμένων σε δεδομένα που προσδίδουν τη δυνατότητα μίμησης κατανόησης και υποστήριξης του ανθρώπου [8]. Μέσω αυτής δίνεται η απάντηση στο πρόβλημα κατασκευής υπολογιστών οι οποίοι μπορούν να βελτιώνονται αυτόματα, μαθαίνοντας από την εμπειρία τους [9]. Η χρήση της ποικίλει στην ανάπτυξη διάφορων εφαρμογών, κάποια παραδείγματα αυτών είναι η όραση υπολογιστών [10], το speech recognition [11], ο ρομποτικός έλεγχος [12], οι βιοϊατρικές εφαρμογές [13] και διάφορες άλλες [14].

Η γραμμική άλγεβρα και η θεωρία των πιθανοτήτων μεταξύ άλλων βασικών μαθηματικών κρίνονται απαραίτητα για την κατανόηση των τεχνικών μηχανικής μάθησης [7].

Μια αρκετά αναπτυσσόμενη και σημαντική κατεύθυνση του Machine Learning για την αντιμετώπιση προβλημάτων είναι η επίλυση τους βάση της Bayesian λογικής. Η προσέγγιση των προβλημάτων της μηχανικής μάθησης με την υποστήριξη της Γκαουσιανής επεξεργασίας είναι ο ακρογωνιαίος λίθος της Μπαγιεσιανής μηχανικής μάθησης.

Αρκετή σημασία φαίνεται να λαμβάνει ο τομέας των kernel machines, μιας κλάσης αλγόριθμων για pattern recognition με διασημότερο μέλος αυτών τον

Support Vector Machine (SVM) [15]. Η σημασία της μεθόδου με χρήση πυρήνων συνίσταται στις περιπτώσεις που χρησιμοποιούνται μη γραμμικές μέθοδοι αντί για γραμμικές, αξιοποιώντας έναν καθορισμένο θετικό πυρήνα και εκμεταλλευόμενος με αυτό τον τρόπο τα θετικά και των δύο. Ο πυρήνας αντιστοιχεί σε ένα βαθμωτό αποτέλεσμα σε έναν χώρο χαρακτηριστικών, όπου η μέθοδος εκτίμησης είναι γραμμική, αλλά μέσω του πυρήνα καταλήγει να μην χρειάζεται ο άμεσος υπολογισμός των ανώτερων διαστάσεων [16]. Ακόμη, αξίζει να σημειωθεί και το Deep learning, ένα υποσύνολο του Machine Learning. Πρακτικά, είναι ένα νευρωνικό δίκτυο με μεγάλο αριθμό στρωμάτων και παραμέτρων. Τα πολλαπλά στρώματα μη γραμμικών μονάδων επεξεργασίας αξιοποιούνται για την παραγωγή χαρακτηριστικών και διάφορων μεταμορφώσεων τα οποία θα παράλβουν τα χαμηλότερα στρώματα για την μάθηση απλών χαρακτηριστικών και τα υψηλότερα στρώματα για την μάθηση πολυπλοκότερων. Μπορεί, λοιπόν, με αυτό το τρόπο να επεξεργαστεί μεγάλο όγκο πληροφορίας και να παράξει τα επιθυμητά αποτελέσματα [14].

Αξίζει να σημειωθεί ότι η υλοποίηση της μηχανικής μάθησης γίνεται, ως επί το πλείστον, σε λογισμικό, ωστόσο όσο περνάει ο καιρός αυξάνονται ολοένα οι απαιτήσεις του, καθώς αυξάνεται ο όγκος πληροφοριών και συνάμα οι υπολογιστικοί πόροι και η απαραίτητη μνήμη [17, 18]. Σε πολλές περιπτώσεις οι υπολογισμοί που χρειάζονται υλοποιούνται σε κάποιο cloud το οποίο με τη σειρά του απαιτεί υψηλό ενεργειακό κόστος και περιορίζεται από προβλήματα του bandwidth. Αυτό, οδηγεί στην ανάγκη να γίνεται η ανάλυση τοπικά πάνω στον αισθητήρα [6]. Για την αντιμετώπιση λοιπόν τέτοιων προβλημάτων μπορούν να χρησιμοποιηθούν διάφορα ειδικά υπολογιστικά συστήματα όπως FPGAs (Field-Programmable Gate Arrays), ASICs (Application-Specific integrated circuit), GPUs αντί για την CPU ενός υπολογιστή που μπορεί να βρίσκεται τοπικά ή στο cloud [19, 20]. Κάθε ένα από αυτά έχει τα θετικά του αλλά αντιμετωπίζει και τα δικά του μειονεκτήματα, η επιλογή κάποιου έναντι κάποιου άλλου προσπίπτει στον ενασχολούμενο κάθε φορά, καθώς έρχεται αντιμέτωπος με το αέναες πρόβλημα του μηχανικού. "trade off". Ανάλογα την εφαρμογή και τις προτεραιότητες που θέτει ο καθένας, πρέπει να επιλέξει ανάμεσα σε απόδοση, επίδοση, κατανάλωση, υπολογιστικούς πόρους, μνήμη κοκ [20, 21].

## 1.2 Χρησιμότητα αναλογικής υλοποίησης εφαρμογών Μηχανικής Μάθησης

Η ενσωμάτωση της μηχανικής μάθησης άμεσα μετά τον αισθητήρα έχοντας raw data αντιμετωπίζει διάφορα προβλήματα αλλά προσφέρει διάφορα πλεονεκτήματα. Μεταξύ άλλων είναι και η ταχύτητα (λιγότερη καθυστέρηση) όπως

και η ασφάλεια (δεν αξιοποιείται το διαδίκτυο για μεταφορά δεδομένων) [6]. Ακόμη, πολλές υλοποιήσεις τέτοιας λογικής λειτουργούν ως επιταχυντές, κλασικό τέτοιο παράδειγμα αποτελούν τα FPGAs [19, 20, 21, 22, 23, 24] όπως θα δούμε και παρακάτω.

Τα FPGAs είναι ψηφιακά ολοκληρωμένα κυκλώματα που περιέχουν προγραμματιζόμενα μπλοκ λογικής τα οποία συνδέονται με τη σειρά τους με προγραμματιζόμενες συνδέσεις. Είναι, δηλαδή, σχεδιασμένα για τη διαμόρφωση τους από τον χρήστη ανάλογα με τις ανάγκες του. Ο προγραμματισμός του γίνεται βάση μιας γλώσσας υλικού HDL (Hardware Description Language) χαρακτηριστικό που προσφέρει συμβατότητα μεταξύ διαφορετικών τεχνολογιών [25]. Πλέον, τα FPGAs χρησιμοποιούνται αρκετά σαν hardware accelerators (επιταχυντές) για βελτίωση της διαδικασίας Machine Learning εκπαίδευσης του συστήματος όσο και στην διαδικασία εξαγωγής των αποτελεσμάτων του μοντέλου [19, 23, 24]. Ωστόσο, άμα ο σκοπός του συστήματος που πρέπει να σχεδιαστεί είναι η χαμηλότερη ενεργειακή κατανάλωση χωρίς να επηρεάζεται ιδιαίτερα η ακρίβεια, πολλοί επιλέγουν mixed-mode αρχιτεκτονικές που αξιοποιούν τόσο τα πλεονεκτήματα της αναλογικής όσο και της ψηφιακής υλοποίησης. Το αρνητικό που θα μπορούσε κάποιος να προσδώσει σε μια τέτοια αρχιτεκτονική είναι το έξτρα κόστος του hardware στην παράλληλη υλοποίηση των DAC, ADC κυκλωμάτων, ο θόρυβος από το analog mismatch και η καθυστέρηση στην επεξεργασία. Βέβαια ανάλογα την υλοποίηση αυτά μπορούν να περιοριστούν με διάφορες μεθόδους [26]. Σε περίπτωση που χρειάζεται περαιτέρω μείωση της κατανάλωσης αλλά και ίσως του χώρου on-chip γεγονός μπορεί να συμβαίνει για διάφορους σκοπούς όπως για την χρήση κάποιου ταξινομητή σε φορητή συσκευή όπου χρειάζεται χαμηλή τροφοδοσία και κατανάλωση [27]. Θα μπορούσε ακόμα να χρησιμοποιηθεί σε κάποιο σύστημα που χρησιμοποιεί το αναλογικό κύκλωμα σαν ένα wake-up circuit, ώστε να μπορεί να ρίξει σε μεγάλο βαθμό την κατανάλωση ενός power hungry ψηφιακού κυκλώματος που μπορεί να είναι αναγκαίο λόγω απαραίτητης υψηλής ακρίβειας [28].

Με τον όρο wake-up circuit εννοείται να υπάρχει ένα αναλογικό ολοκληρωμένο κύκλωμα που θα είναι μονίμως ανοιχτό χωρίς να μας επηρεάζει δραματικά στη συνολική κατανάλωση καθώς όπως αναφέρθηκε πολλάκις, έχει εξαιρετικά χαμηλή κατανάλωση σε σύγκριση με το ψηφιακό. Το κύκλωμα αυτό θα λειτουργεί συνεχώς λαμβάνοντας αποφάσεις με σκοπό να ενεργοποιήσει το ψηφιακό όταν θα κρίνει αναγκαίο να πάρει μια σημαντική απόφαση. Όπως γίνεται και στο [29], όπου ο αναλογικός ταξινομητής ανιχνεύει συνεχώς περιοχές που ο ασθενής πρόκειται να πάθει εντός καθορισμένου χρόνου κάποια κρίση και ενεργοποιεί το ψηφιακό για να κρίνει αν είναι σωστή η απόφαση του αναλογικού κυκλώματος αλλιώς να την αγνοήσει. Αντίστοιχη λογική μπορεί να ακολουθηθεί

σε οποιαδήποτε εφαρμογή κριθεί αναγκαία η εγκατάσταση κάποιου αναλογικού κυκλώματος χαμηλής κατανάλωσης ως wake-up circuit.

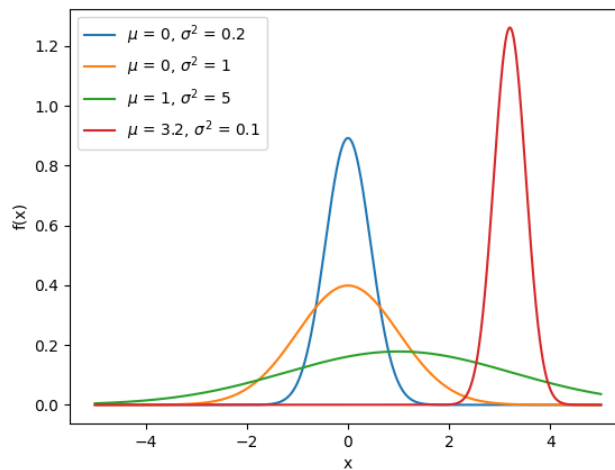
### 1.3 Αναλογικές υλοποιήσεις βασισμένες στην Γκαουσιανή κατανομή

Στις αμιγώς αναλογικές υλοποιήσεις μηχανικής μάθησης, υπάρχει μια ευρύτερη κατηγορία ταξινομητών που βασίζονται στην κανονική (Γκαουσιανή κατανομή).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (1.3.1)$$

όπου,  $\mu$  είναι η μέση τιμή της κατανομής,  $\sigma^2$  η διασπορά,  $\sigma$  η τυπική απόκλιση [30]. Αλλάζοντας το  $\mu$  μπορούμε να μεταφέρουμε το κέντρο της κατανομής μας πάνω στον  $x$  άξονα και αλλάζοντας το  $\sigma$  μπορούμε να αλλάξουμε να μεταβάλλουμε το φάρδος της καμπύλης μας όπως φαίνεται και στο σχήμα 1.3.1. Κοινώς οι Γκαουσιανές καμπύλες ονομάζονται και καμπάνες (bell curves), ονομασία που θα χρησιμοποιηθεί και μετέπειτα στη πορεία της θέσης. Περαιτέρω ανάλυση αναφορικά με την Γκαουσιανή κατανομή θα επέλθει σε επόμενο κεφαλαίο.

Σχήμα 1.3.1: Normal distribution



An example of a normal distribution with different mean values and sigmas

### 1.3. ΑΝΑΛΟΓΙΚΕΣ ΥΛΟΠΟΙΗΣΕΙΣ ΒΑΣΙΣΜΕΝΕΣ ΣΤΗΝ ΓΚΑΟΥΣΙΑΝΗ ΚΑΤΑΝΟΜΗ21

Οι προτεινόμενες υλοποιήσεις μηχανικής μάθησης σε hardware level είναι RBF NNs [31, 32, 33, 34]. Διάφορα παραδείγματα αποτελούν οι Γκαουσιανές RBF NNs (GRBF NN) [35], ο SVM [36] ή ο SVR [37].

Συγκεκριμένα παραδείγματα Γκαουσιανής κατανομής σε κυκλωματικό επίπεδο θα φανούν και παρακάτω στην διπλωματική και βασίζονται όπως θα δούμε στο bump circuit που αναφέρεται στο [3].



## Κεφάλαιο 2

# Υπόβαθρο

Στο κεφάλαιο αυτό θα εξηγηθεί με πληρότητα το απαιτούμενο υπόβαθρο για την κατανόηση της διπλωματικής. Θα εξηγηθεί αναλυτικά η λογική που βασίζονται τα building blocks της αρχιτεκτονικής που θα υλοποιηθεί στη συνέχεια. Αναφέρονται τόσο το αναγκαίο θεωρητικό μαθηματικό υπόβαθρο που αποτελεί πάτημα για την ιδέα του ταξινομητή μας όσο και τα πρακτικά μαθηματικά μοντέλα που ακολουθεί το τρανζίστορ. Στο πνεύμα αυτό θα εξηγηθεί αναλυτικά η λειτουργία του βασικού στοιχείου υλοποίησης του μεγαλύτερου αριθμού κυκλωμάτων στην αγορά, το MOSFET αλλά και η περιοχή υποκατωφλίου η οποία μας παρέχει τη δυνατότητα για υλοποιήσεις πολύ χαμηλής κατανάλωσης. Ακόμη θα γίνει μια αναφορά στην ψηφιακή λογική του τρανζίστορ. Όλα αυτά συναινοούν τελικά, ώστε να προσδοθεί το απαραίτητο υπόβαθρο για την πλήρη αντίληψη της παρούσας διπλωματικής εργασίας.

### 2.1 Gaussian Mixture Model

Παρακάτω θα αναλυθεί όλο το μαθηματικό υπόβαθρο που θα χρειαστεί για την κατανόηση του Gaussian Mixture Model και εκτός αυτού θα εξηγηθεί η λειτουργία του μοντέλου. Θα εξηγήσουμε ακόμη πως εκπαιδεύεται το μοντέλο, κάποιες διαφορές του με παρόμοια, καθώς και τα πλεονεκτήματα του έναντι αυτών.

#### 2.1.1 Υπόβαθρο στη Θεωρία των Πιθανοτήτων

Για να μπορέσουμε να κατανοήσουμε πλήρως το GMM πρέπει να μπορέσουμε να αντιληφθούμε πρώτα κάποιες βασικές αρχές της θεωρίας των πιθανοτήτων.

Η δεσμευμένη πιθανότητα του  $A$  δεδομένου του  $B$  ορίζεται ως

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad (2.1.1)$$

όπου  $P(B) > 0$  η πιθανότητα του υπό δέσμευση γεγονότος και  $P(A \cap B)$  η πιθανότητα της τομής των συνόλων  $A$  και  $B$ . Πρακτικά, το  $P(A|B)$  μας δίνει τη πιθανότητα του αποτελέσματος του ενδεχόμενου  $A$  δεδομένου ότι γνωρίζουμε το ενδεχόμενο  $B$ . Ακολουθώντας, βάση του θεωρήματος του Bayes,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)}, \quad (2.1.2)$$

όπου  $P(B)$  βάση του θεωρήματος συνολικής πιθανότητας, προϋποθέτοντας ότι τα σύνολα  $A_1, \dots, A_n$  είναι ξένα μεταξύ τους σύνολα και  $P(A_i) > 0$ , τότε

$$P(B) = P(A_1 \cap B) + \dots + P(A_n \cap B) = P(A_1)P(B|A_1) + \dots + P(A_n)P(B|A_n), \quad (2.1.3)$$

όπου  $P(A)$ ,  $P(B)$  οι πιθανότητες για τα ενδεχόμενα  $A, B$  και  $P(B|A)$  η δεσμευμένη πιθανότητα του  $B$  δεδομένου του  $A$  [38, 39]. Ειδικότερα, τα γεγονότα  $A_1, \dots, A_n$  αντιστοιχούν στις αιτίες και το  $B$  είναι το αποτέλεσμα. Οπότε όταν χρησιμοποιούμε την δεσμευμένη πιθανότητα  $A$  δεδομένου του  $B$ , επιδιώκουμε στον υπολογισμό της πιθανότητας  $P(A_i|B)$  να έχει συντελέσει στο αποτέλεσμα  $B$ , δεδομένου ότι έχουμε το αποτέλεσμα αυτό, η αιτία  $A$ . Για αυτό το λόγο, λοιπόν, αναφερόμαστε στην  $P(A_i|B)$  ως posterior probability (εκ των υστέρων πιθανότητα) του γεγονότος  $A_i$  και στην  $P(A_i)$  ως prior probability (εκ των προτέρων πιθανότητα) [40].

Έστω μια τυχαία συνεχής μεταβλητή  $X$  αν υπάρχει μια μη αρνητική συνάρτηση  $f_X(x)$  που ονομάζεται συνάρτηση πυκνότητας πιθανότητας δηλαδή probability density function (PDF) ώστε  $P(X \in B) = \int_B f_X(x)dx$  για κάθε υποσύνολο  $B \in \mathbb{R}$ . Ακόμη έχουμε ότι η πιθανότητα της τιμής  $X$  να βρίσκεται σε ένα διάστημα  $X \in (x_1, x_2)$  δίνεται από το ολοκλήρωμα  $P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f_X(x)dx$ . Τέλος, αξίζει να αναφερθεί η εξής ιδιότητα της PDF [38, 39, 40]:

$$\int_{-\infty}^{\infty} f_X(x)dx = P(-\infty < X < \infty) = 1 \quad (2.1.4)$$

### 2.1.2 Μπαγιεσιανό Μοντέλο

Σαν γενική αρχή η Γκαουσιανή κατανομή είναι μια αρκετά διαδεδομένη κατανομή για συνεχείς μεταβλητές. Όπως είδαμε και παραπάνω στην περίπτωση μιας



μεταβλητής έχουμε την μορφή

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad (2.1.5)$$

όπου  $\mu, \sigma^2$  η μέση τιμή και η διασπορά αντίστοιχα. Στην περίπτωση περισσότερων διαστάσεων, δηλαδή όταν έχουμε ένα διάνυσμα  $\mathbf{x}$  ορίζεται η παραλλαγή:

$$\mathcal{N}(X|M, \Sigma) = \frac{1}{(2\pi^{n/2})} \frac{1}{|\Sigma|^{1/2}} e^{-\frac{1}{2}(x-M)^T \Sigma^{-1}(x-M)} \quad (2.1.6)$$

με  $M$  ένα  $n$ -διάστατο διάνυσμα μέσης τιμής,  $\Sigma$  ένας  $n \times n$  πίνακας διασποράς και το  $|\Sigma|$  υποδηλώνει την ορίζουσα του  $\Sigma$  [7].

Στην περίπτωση που ο πίνακας  $\Sigma$  είναι διαγώνιος [29] έχουμε την εξής απλοποίηση:

$$\mathcal{N}(X|M, \Sigma) = \frac{1}{(2\pi\sigma_n^2)^{1/2}} e^{-\frac{1}{2} \frac{(x_n - \mu_n)^2}{\sigma_n^2}} \quad (2.1.7)$$

όπου τα  $x_n, \mu_n, \sigma_n$  είναι οι  $n$ -οστές τιμές του διανύσματος εισόδου και των πινάκων μέσης τιμής και συνδιασποράς  $M, \Sigma$  αντίστοιχα.

Αναλογιζόμενοι, λοιπόν, την PDF 2.1.7 και τους τύπους που αντλήσαμε από το θεώρημα του Bayes 2.1.2 μπορούμε να αντλήσουμε την πιθανότητα ενός  $N$ -διάστατου διανύσματος  $\mathbf{X}$  από τον εξής τύπο:

$$\mathcal{N}(X|M_i, \Sigma_i) = \prod_{n=1}^N \mathcal{N}(x_n|\mu_n, \sigma_n^2), \quad (2.1.8)$$

όπου  $N$ , ο αριθμός των διαστάσεων του διανύσματος και  $X, M_i, \Sigma_i$  είναι τα διανύσματα εισόδου, μέσης τιμής και διασποράς, ενώ τα  $x_n, \mu_n, \sigma_n^2$  είναι οι  $n$ -οστές τιμές τους αντίστοιχα, με το τελευταίο να είναι η  $(n, n)$  τιμή του πίνακα  $\Sigma_i$ .

Από αυτή την PDF 2.1.8 καταλήγουμε στον εξής τύπο για τον υπολογισμό της πιθανότητας κάποιο διάνυσμα  $X$  να ανήκει σε μια κλάση  $C_i$ :

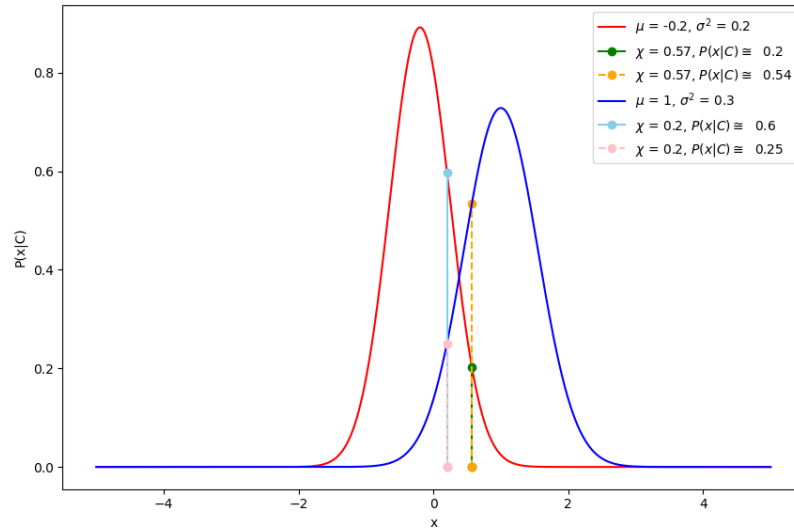
$$P(X|C_i) = \prod_{n=1}^N \frac{1}{(2\pi\sigma_n^2)^{\frac{1}{2}}} e^{-\frac{1}{2} \frac{(x_n - \mu_n)^2}{\sigma_n^2}}, \quad C_i \in [1, \dots, k], \quad (2.1.9)$$

όπου  $k$  ο αριθμός των κλάσεων.

Υπολογίζεται, λοιπόν, η δεσμευμένη πιθανότητα του διανύσματος  $X$  να βρίσκεται σε κάθε μία κλάση δεδομένης της πιθανότητας της κλάσης αυτής και

τελικά συγκρίνονται μεταξύ τους για να καταλήξει ο ταξινομητής σε ένα συμπέρασμα αναφορικά με την κλάση που ανήκει το διάνυσμα τελικά. Συγκρίνει, δηλαδή, όλες αυτές τις πιθανότητες και θεωρεί σωστή αυτή με την μεγαλύτερη τιμή.

Έστω, 2 κλάσεις  $C_1, C_2$  με 2 συναρτήσεις πυκνότητας πιθανότητας που διέπουν την κάθε μια  $\mathcal{N}(X|C_1, \Sigma_{C_1}), \mathcal{N}(X|C_2, \Sigma_{C_2})$  και μια είσοδος  $x_i$  (έστω μονοδιάστατο για επεξηγηματικού λόγους). Τότε υπολογίζουμε τις  $P(x|C_1), P(x|C_2)$  και από τις δύο η μεγαλύτερη θα μας δώσει την σωστή κλάση της εισόδου. Βλέπουμε, λοιπόν στο σχήμα 2.1.1, ότι στην περίπτωση που το διάνυ-



Σχήμα 2.1.1: 1 dimensional comparator for Gaussian Distributions, Bayesian classifier

σμα εισόδου είναι το  $x = 0.57$  η μπλε κλάση αναγνωρίζει την είσοδο σαν δικιά της αφού έχει μεγαλύτερη πιθανότητα ενώ η είσοδος  $x = 0.2$  βλέπουμε ότι επάγεται στην κόκκινη κλάση. Αντίστοιχα, θα βγάζαμε συμπεράσματα και για περισσότερες διαστάσεις. Θα έχουμε δηλαδή διάφορες γκαουσσιανές σαν αυτές που θα μας δίνουν την πιθανότητα για κάθε διάσταση και εν τέλει θα πολλαπλασιάσουμε όλες αυτές τις πιθανότητες για να έχουμε ένα αποτέλεσμα για κάθε κλάση το οποίο θα συγκριθεί με τα υπόλοιπα για να γίνει η ταξινόμηση.

### 2.1.3 K-Means

Ο k-means αλγόριθμος είναι αρκετά χρήσιμος και ένας από τους πιο διαδεδομένους αλγόριθμους για unsupervised clustering [41]. Αξιοποιείται αρκετά στους τομείς της ανάλυσης δεδομένων [42, 43], pattern recognition [7], image processing [44, 45] και αρκετά ακόμα. Ο k-means βασίζεται στον προσδιορισμό κάποιων κέντρων  $\mu_k$  των clusters, τα οποία θα σχηματίσουν έναν κύκλο εντός του οποίου πρέπει να καθοριστούν τα δεδομένα που έρχονται. Αυτό γίνεται καθορίζοντας σωστά τα διανύσματα  $\mu_k$ , έτσι ώστε να έχουν το ελάχιστο άθροισμα των τετραγώνων των αποστάσεων κάθε σημείου που έρχεται από το dataset, στο κοντινότερο  $\mu_k$ . Για κάθε σημείο  $x_n$  καθορίζεται μια δυαδική συνάρτηση  $r_{nk}$  που παίρνει τιμές 0,1,  $k = 1, \dots, K$  με  $K$  να είναι ο αριθμός των clusters. Η  $r_{nk}$  γίνεται ίση με 1 όταν το  $x_n$  ανήκει στο cluster  $k$  αλλιώς θα είναι ίση με 0. Μπορούμε με αυτό τον τρόπο να καθορίσουμε μια αντικειμενική συνάρτηση ονομαζόμενη και ως distortion measure:

$$J = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|x_n - \mu_k\|^2 \quad (2.1.10)$$

που καθορίζει το άθροισμα των τετραγώνων των αποστάσεων κάθε σημείου από κάθε σημείο του διανύσματος  $\mu_k$ . Σκοπός του αλγορίθμου είναι να μπορέσει να βρει τα κατάλληλα  $r_{nk}, \mu_k$  ώστε να ελαχιστοποιήσει την συνάρτηση 2.1.10. Αυτό καταφέρνεται μέσω μιας επαναληπτικής μεθόδου δύο βημάτων (Expectation-Maximization, EM). Ξεκινώντας ο αλγόριθμος, καθορίζει αρχικές τιμές στα  $\mu_k$ . Έπειτα, ξεκινάει η επαναληπτική διαδικασία με το πρώτο βήμα της να διατηρεί σταθερές τις  $\mu_k$  προσπαθώντας να βρει το ελάχιστο  $J$  σχετικά με το  $r_{nk}$ , ενώ στο δεύτερο κρατάει σταθερό το  $r_{nk}$  και ελαχιστοποιεί το  $J$  βάση του  $\mu_k$ . Αυτό επαναλαμβάνεται και κάποια στιγμή συγκλίνει [7, 41]. Ο k-means αντιλαμβάνομαστε ότι κάνει ένα hard-assignment των δεδομένων στα clusters, όπου κάθε ένα από αυτά θα ανήκει σε 1 cluster αποκλειστικά, σε αντίθεση με το GMM που αξιοποιεί την εκ των υστέρων πιθανότητα για τον καθορισμό των σημείων σε clusters και έτσι γίνεται ένα soft assignment [7, 46]. Καταλήγουμε, λοιπόν, ότι ο GMM είναι ανώτερος του k-means αλγόριθμου δεδομένου ότι μπορεί να φτιάξει ελλειψοειδής clusters και δεν περιορίζεται μόνο σε κυκλικά. Μπορεί, δηλαδή, να ξεπεράσει το πρόβλημα που έχει ο k-means όταν του δίνονται δεδομένα που μπορεί να μην έχουν κατανομές ίδιου μεγέθους ή σχήματος .

### 2.1.4 Το Γκαουσιανό Μοντέλο Μίξης

Γενικά ένα μοντέλο μίξης αποτελείται από 2 ή περισσότερα μοντέλα. Η μορφή της pdf ενός πεπερασμένου  $p$ -διάστατου μοντέλου μίξης είναι κατά γενική αρχή:

$$f(y) = \sum_{i=1}^n \pi_i f_i(y), \quad (2.1.11)$$

όπου τα  $\pi_i$  τα βάρη των κατανομών που πρέπει να είναι μη αρνητικά και το άθροισμα τους είναι ίσο με την μονάδα, ενώ τα  $f_i(y)$  είναι οι πυκνότητες των στοιχείων [47].

Το Γκαουσιανό μοντέλο μίξης είναι πρακτικά ένα μοντέλο μίξης που περιέχει μια ή περισσότερες Γκαουσιανές κατανομές, η κάθε μια από τις οποίες καθορίζεται από  $k \in 1, \dots, K$  όπου  $K$  είναι ο αριθμός των clusters. Κάθε κατανομή καθορίζεται από τη μέση τιμή του  $\mu_k$  η οποία ορίζει τα κέντρα των κατανομών, τη συνδιακύμανση  $\Sigma$  που καθορίζει το πλάτος της Γκαουσιανής σε μια διάσταση ή το σχήμα της ελλειψοειδούς σε μεγαλύτερες διαστάσεις και, τέλος, το  $\pi_k$  που είναι το βάρος της όπως είπαμε και παραπάνω και καθορίζει το ύψος της κάθε γκαουσιανής. Το GMM θεωρεί ότι οι κλάσεις και τα clusters διέπουν μια Γκαουσιανή κατανομή και προσπαθεί να προσδιορίσει τις βέλτιστες παραμέτρους για κάθε cluster ώστε να ταιριάζουν στα υπάρχοντα δεδομένα που δίνονται και να μπορέσει να ταξινομήσει με σωστό τρόπο καινούργια δεδομένα.

Ακολουθώντας, λοιπόν, τη σχέση 2.1.11 και θεωρώντας ένα πρόβλημα με  $N_c$  classes μπορούμε να καταλήξουμε αξιοποιώντας την Γκαουσιανή κατανομή στην εκ των υστέρων πιθανότητα μια είσοδος να προέρχεται από μια κλάση  $C_k$ :

$$p(X|C) = \sum_{k=1}^K \pi_k^c \mathcal{N}(X|M_k^c, \Sigma_k^c) \quad (2.1.12)$$

όπου το  $\pi_k^c$  είναι το σχετικό βάρος κάθε κατανομής εντός του μοντέλου για την κλάση  $C$  με  $K$  clusters, το άθροισμα των οποίων μας δίνει τη μονάδα και  $\mathcal{N}(X|M_k^c, \Sigma_k^c)$  η πολυδιάστατη Γκαουσιανή κατανομή. Τα βάρη  $\pi_k$  καθορίζονται από την πιθανότητα κάθε στοιχείου να είναι στο εκάστοτε cluster [7, 29, 48]. Σκοπός του μοντέλου, όπως και στο απλό Μπαγιεσιανό, είναι να μας δώσει τη μέγιστη τιμή της 2.1.12 για οποιοδήποτε διάνυσμα  $\mathbf{X}$  χρειαστεί να ταξινομηθεί σε κάποια από αυτά τα clusters για κάθε κλάση.

Στην περίπτωση που έχουμε ένα  $N$ -διάστατο διάνυσμα καλούμαστε να μεγιστοποιήσουμε την  $P(X|\pi, M, \Sigma)$ :

$$P(X|\pi, M, \Sigma) = \sum_{k=1}^K \pi_k \mathcal{N}(x_n|M_k, \Sigma_k), \quad (2.1.13)$$

όπου το  $\prod_{n=1}^N$  καθορίζει απλά τις διαστάσεις του διανύσματος εισόδου. Όστε να έχουμε το αποτέλεσμα μέσω της:

$$y = \operatorname{argmax}_{k \in [1, K]} \{P(C_k)P(X|C_k)\}, \quad (2.1.14)$$

Έστω μια τυχαία δυαδική μεταβλητή δείκτης  $z_{nk} \in \{0, 1\}$  η οποία είναι 1 όταν το  $x_n$  είναι στο cluster  $k$  αλλιώς είναι ίση με 0. Έχουμε δηλαδή την πιθανότητα:

$$\begin{aligned} p(z_{nk} = 1) &= \pi_k \\ \text{με } 0 &\leq \pi_k \leq 1 \\ \text{και } \sum_{k=1}^K \pi_k &= 1 \end{aligned}$$

$$p(z) = \prod_{k=1}^K \pi_k^{z_k} \quad (2.1.15)$$

Έχουμε ακόμη την συνάρτηση  $\gamma$  όπου:

$$\gamma(z_{nk}) \equiv p(z_{nk} = 1 | \mathbf{x}_n) \quad (2.1.16)$$

και μέσω των 2.1.2, 2.1.3 έχουμε:

$$\begin{aligned} \gamma(z_{nk}) &= \frac{p(x_n | z_{nk} = 1)p(z_{nk} = 1)}{p(x_n)} \\ &= \frac{p(x_n | z_{nk} = 1)p(z_{nk} = 1)}{\sum_{i=1}^K p(z_{ni} = 1)p(x_n | z_{ni} = 1)} \end{aligned} \quad (2.1.17)$$

Τέλος, υπάρχει η συνάρτηση:

$$N_k = \sum_{n=1}^N \gamma(z_{nk}) \quad (2.1.18)$$

Για να μεγιστοποιήσουμε την 2.1.13 παίρνουμε τις παραγώγους της αναφορικά με τις παραμέτρους της ( $M_k, \Sigma_k, \pi_k$ ) και εξισώνουμε με το μηδέν, από αυτές μετά από πράξεις καταλήγουμε σε τρεις τύπους 2.1.19 που τα  $M_k, \Sigma_k, \pi_k$  εξαρτώνται από το  $\gamma z_{nk}$  και όπως είδαμε το  $\gamma z_{nk}$  εξαρτάται μόνο από τα  $M_k, \Sigma_k, \pi_k$ . Έχουν δηλαδή μια σχέση αλληλεξάρτησης μεταξύ τους και έτσι οδηγείται

ο αλγόριθμος EM που αξιοποιείται για να μπορέσει το μοντέλο να έχει σωστές τιμές.

$$\begin{aligned}\mu_k^{new} &= \frac{1}{N_k} \sum_{n=1}^N \gamma(z_{nk}) x_n \\ \Sigma_k^{new} &= \frac{1}{N_k} \sum_{n=1}^N \gamma(z_{nk}) (x_n - \mu_k^{new})(x_n - \mu_k^{new})^T \\ \pi_k^{new} &= \frac{N_k}{N}\end{aligned}\tag{2.1.19}$$

Ο αλγόριθμος EM που αναφέρθηκε και στον k-means είναι μια επαναληπτική μέθοδος που καταλήγει στην εκπαίδευση των παραμέτρων του μοντέλου μας έτσι ώστε αυτό να μπορέσει να συγκλίνει σε τιμές που ταιριάζουν βέλτιστα στα δεδομένα που δίνονται από ένα train-set ώστε να μπορέσει να ταιριάζει ορθώς τα νέα δεδομένα στα clusters που αρμόζουν. Ομοίως, λοιπόν, με τον k-means προτού ξεκινήσει η επαναληπτική μέθοδος αρχικοποιούμε τις παραμέτρους που χρειάζεται να βελτιστοποιήσουμε δηλαδή τα  $\mathbf{M}$ ,  $\Sigma_k$ ,  $\pi_k$ . Εν συνεχεία, προχωρώντας στο βήμα E υπολογίζουμε τα  $\gamma$  όπως στην 2.1.17. Έχοντας τα  $\gamma$ , μπορούμε να προχωρήσουμε στο επόμενο βήμα M όπου θα επαναπροσδιορίσουμε τις παραμέτρους 2.1.19. Στο σημείο αυτό, ξεκινάει η επανάληψη των βημάτων E και M έως ότου επέλθει μια σύγκλιση, είτε αυτό είναι ένα κριτήριο λάθους που μειώνεται κάτω από ένα κατώφλι, είτε συγκεκριμένος αριθμός επαναλήψεων, είτε μέσω του ελέγχου μιας συνάρτησης ομοιότητας [7, 41, 47, 48].

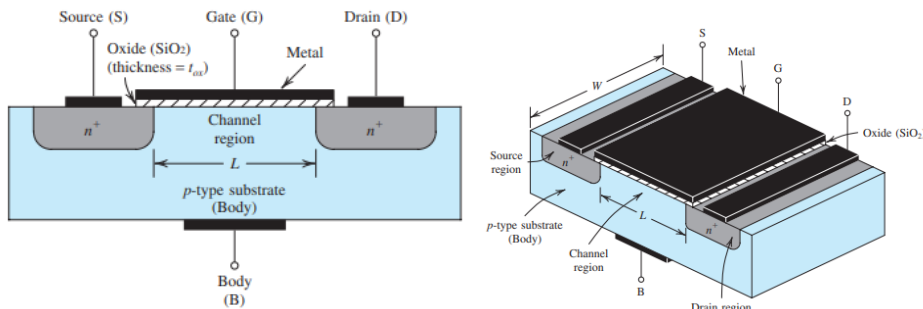
## 2.2 MOSFET transistor

Στο χώρο των ψηφιακών ολοκληρωμένων κυκλωμάτων έχουν επικρατήσει κάποια τρανζίστορ γνωστά ως MOSFET λόγω της υψηλής πυκνότητα ολοκλήρωσης που επιτρέπουν και της χαμηλής κατανάλωσης ισχύος [49]. Το MOSFET (Metal Oxide semiconductor field effect transistor) είναι ένας τύπος τρανζίστορ, όπως και το διπολικό τρανζίστορ (BJT bipolar junction transistor). Στην παρούσα διπλωματική ασχολούμαστε αποκλειστικά με το πρώτο τύπο παρά τα πλεονεκτήματα του δεύτερου στα αναλογικά κυκλώματα, καθώς ο στόχος μας είναι η υλοποίηση ενός κυκλώματος χαμηλής κατανάλωσης. Τα MOSFET είναι τα ευρύτερα χρησιμοποιούμενα ηλεκτρονικά στοιχεία στην σχεδίαση ολοκληρωμένων καθώς δύναται να προσομοιώσει αντιστάσεις και πυκνωτές, συνεπώς μπορεί να χρησιμοποιηθεί σχεδόν αποκλειστικά σαν στοιχείο. Ακόμα, συγκριτικά με τα BJT μπορούν να φτάσουν αρκετά μικρά μεγέθη, άρα έχουν

μικρότερη κατανάλωση και πιάνουν λιγότερο χώρο on-chip, ενώ συνάμα η διαδικασία κατασκευής τους είναι σχετικά απλή [1, 2]. Οπότε η σχεδίαση κυκλωμάτων με χρήση MOS μειώνει το κόστος και αυξάνει την φορητότητα για αυτό και τα αναλογικά κυκλώματα χρησιμοποιούν τεχνολογίες συμβατές με αυτά [49].

### 2.2.1 Βασική λειτουργία μεγάλου σήματος MOSFET

Για να αντιληφθούμε τη λειτουργία του τρανζίστορ πρέπει πρώτα να αναλύσουμε συνοπτικά την δομή του. Θα φέρουμε σαν παράδειγμα τα nMOS και τα ακριβώς αντίθετα θα ισχύουν για τα pMOS, δηλαδή όπου n θα έχουμε p και τούμπαλιν. Για τη δημιουργία του τρανζίστορ n καναλιού θα έχουμε, αρχικά, ένα υποστρώμα τύπου p πάνω στο οποίο νοθεύουμε δύο περιοχές τύπου n στις οποίες για να έχουν όσο το δυνατόν μικρότερη αντίσταση προσδίδουμε μεγάλη συγκέντρωση ηλεκτρονίων και συνεπώς ονομάζουμε  $n^+$ . Πάνω από αυτές θα δημιουργηθούν επαφές μετάλλου. Πάνω από την ενδιάμεση περιοχή αυτών των νοθεύσεων, δηλαδή των περιοχών για τους ακροδέκτες source, drain, βάζουμε ένα λεπτό στρώμα διοξειδίου του πυριτίου ( $SiO_2$ ) και πάνω από αυτό θα βάλουμε το μέταλλο για το ηλεκτρόδιο της πύλης (gate). Τέλος, υπάρχει και ένας ακόμα ακροδέκτης στον οποίο θα βάλουμε μια επαφή μετάλλου, στο κάτω μέρος του υποστρώματος και θα ονομαστεί σώμα (body ή bulk).



(α') Τρισδιάστατη όψη του nMOS transistor

(β') Τομή του παραπάνω μοντέλου

Σχήμα 2.2.1: Η δομή του nMOS transistor [1]

Θέτοντας θετική (αρνητική στην περίπτωση του pmos και ακολουθούμε το μοτίβο του συμπληρωματικού που αναφέρθηκε νωρίτερα) τάση στην πύλη καταφέρνουμε να απωθήσουμε τις οπές του υποστρώματος και να έλξουμε τα ηλεκτρόνια από τις νοθευμένες περιοχές, έτσι δημιουργούμε έναν πυκνωτή κάτω από το gate (ανάμεσα στα source drain) στον οποίο δημιουργείται ένα ηλεκτρικό πεδίο με εγκάρσια κατεύθυνση. Το ρεύμα που παράγεται μέσα στο κανάλι αυτό

είναι ίσο με την μεταβολή του φορτίου του καναλιού σε κάποια απόσταση  $y$  από την πηγή, το οποίο με τη σειρά του είναι ανάλογο από το πλάτος του καναλιού, μιας διαφοράς μήκους στο κανάλι και από το επαγόμενο φορτίο ηλεκτρονίων ανά μονάδα επιφάνειας του καναλιού. Αν θεωρήσουμε ότι η τάση πύλης-καναλιού ( $V_{GS}$ ) είναι μεγαλύτερη από αυτή του κατωφλίου παίρνουμε:

$$Q_1(y) = C_{ox}[V_{GS} - V(y) - V_t], \quad (2.2.1)$$

όπου  $V(y)$  είναι η τάση στο κανάλι σε απόσταση  $y$  σε σχέση με την πηγή, και η τάση πύλης-καναλιού σε αυτό το σημείο είναι  $V_{GS} - V(y)$ , ενώ  $C_{ox}$  είναι η χωρητικότητα του οξειδίου της πύλης ανά μονάδα επιφάνειας [49]. Ανάλογα με την τάση  $V_{GS}$  που δίνεται καθορίζεται και η περιοχή λειτουργίας του τρανζίστορ, αν θα είναι δηλαδή στην ενεργό περιοχή ή στην αποκοπή που σε μεγάλο βαθμό θεωρείται ότι το τρανζίστορ είναι off αλλά στην πραγματικότητα για ένα εύρος το τρανζίστορ λειτουργεί στην περιοχή της subthreshold που χωρίζεται σε ασθενή και ισχυρή και πρακτικά ορίζει την περιοχή της ασθενής αναστροφής (υπάρχει και μια ενδιάμεση περιοχή σε ισχυρή και ασθενή αναστροφή γνωστή και ως μέτρια αναστροφή). Η λειτουργία του MOSFET στην περιοχή υποκατωφλίου (**subthreshold**) θα αναλυθεί αναλυτικότερα στη συνέχεια καθώς είναι η περιοχή που λειτουργούν τα κυκλώματα στην παρούσα εργασία. Στην περίπτωση που η τάση πύλης-πηγής είναι μεγαλύτερη από την τάση κατωφλίου, ορίζονται ανάλογα με την τάση απαγωγού-πηγής δύο περιοχές γνωστές και ως περιοχή κορεσμού-στραγγαλισμού (saturation or pinched off) και ωμική-τριοδική περιοχή τις οποίες θα αναλύσουμε πολύ συνοπτικά:

1. Για  $V_{DS} < V_{GS} - V_t$  είμαστε στην τριοδική περιοχή (ή ωμική). Στην περιοχή αυτή έχουμε εμφανής εξάρτηση του ρεύματος από την τάση  $V_{DS}$  και παρατηρούμε ότι αλλάζοντας την τάση στην είσοδο (πύλη) μπορούμε να αξιοποιήσουμε το τρανζίστορ σαν αντίσταση.

$$I_D = \mu_n C_{ox} \frac{W}{L} [(V_{GS} - V_t)V_{DS} - \frac{V_{DS}^2}{2}] \quad (2.2.2)$$

όπου ,

$$V_t = V_{T0} = \gamma[\sqrt{2\phi_f + V_{SB}} - \sqrt{2\phi_f}] \quad (2.2.3)$$

Στον τύπο 2.2.3 μπορούμε να δούμε και την εξάρτηση του body στην τάση του κατωφλίου. Το  $\phi_f$  είναι ένα δυναμικό που ονομάζεται όριο Fermi και συσχετίζεται με τον ημιαγωγό, το  $\gamma$  είναι μια παράμετρος που συσχετίζεται με το οξείδιο (αντιστρόφως ανάλογη της  $C_{ox}$  που είναι η χωρητικότητα του οξειδίου της πύλης ανά μονάδα επιφάνειας),  $\mu_n$  είναι η μέση ευκίνηση των ηλεκτρονίων μέσα στο κανάλι,  $W, L$  το πλάτος



και μήκος του καναλιού ενώ  $V_{SB}$ ,  $V_{GS}$ ,  $V_{DS}$ , είναι οι τάσεις ανάμεσα σε source-body, gate-source, drain-source αντίστοιχα [49].

2. Στην περίπτωση που  $V_{DS} \geq V_{GS} - V_t$  το τρανζίστορ εισέρχεται στην περιοχή του κορεσμού (saturation region) ή αλλιώς στραγγαλισμού (pinched off). Αυτό συμβαίνει διότι το ρεύμα  $V_{DS}$  είναι μεγαλύτερο ή οριακά ίσο με την  $V_{ov}$  όπου,  $V_{ov} = V_{GS} - V_t$  και έτσι το κανάλι γίνεται pinched στην πλευρά του drain, ορίζοντας έτσι ένα καινούργιο μήκος καναλιού  $L_{eff}$  το οποίο θεωρείται προσεγγιστικά ίδιο στην βιβλιογραφία με το  $L$ . Συνεπώς ακολουθώντας την μαθηματική απόδειξη [49], καταλήγουμε:

$$I_D = \frac{\mu_n C_{ox} W}{2L} (V_{GS} - V_t)^2 \quad (2.2.4)$$

, χωρίς να συμπεριλαμβάνουμε το φαινόμενο early,

$$I_D = \frac{\mu_n C_{ox} W}{2L} (V_{GS} - V_t)^2 (1 + \lambda V_{DS}), \quad (2.2.5)$$

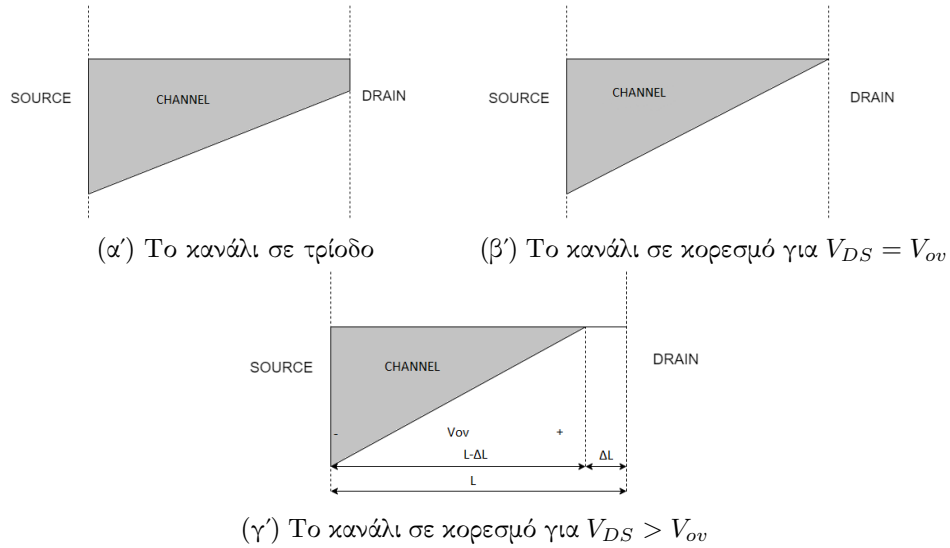
όπου  $\lambda = \frac{1}{V_A}$ ,  $V_A = L_{eff} \left( \frac{dX_d}{dV_{DS}} \right)^{-1}$ , και  $X_d$  είναι το εύρος του στρώματος απογύμνωσης. Για να καταλήξουμε στις σχέσεις 2.2.4, 2.2.5 έγιναν κάποιες παραδοχές ώστε να μπορέσουμε να έχουμε κλειστή μορφή στο ολοκλήρωμα του ρεύματος drain σχετικά με τις τάσεις και το μήκος του καναλιού. Υποτέθηκε ότι το  $V_t$  δεν εξαρτάται από το  $x$  (θέση στο κανάλι), ενώ υπάρχει εξάρτηση [49]. Αυτό δημιουργεί ασυνέχεια στις σχέσεις 2.2.2 και 2.2.5.

Στα σχήματα 2.2.3β' και 2.2.3α' βλέπουμε την εξάρτηση που έχει ανά τις περιοχές η  $I_D$  από την  $V_{DS}$ , ενώ στα 2.2.2 βλέπουμε τα κανάλια πως διαμορφώνονται ανάλογα τις περιοχές.

### 2.2.2 Subthreshold region

Το μεγαλύτερο ποσοστό των κυκλωμάτων λειτουργούν σε ισχυρή αναστροφή, δηλαδή έχουν  $V_{GS} > V_{th}$  ωστόσο ολοένα και αυξάνεται ο αριθμός των εφαρμογών που αξιοποιούν τη λειτουργία στην περιοχή υποκατωφλίου [1]. Παραπάνω αναφερθήκαμε στην ισχυρή αναστροφή και παρατήρησαμε ότι για να καταλήξουμε στις σχέσεις του ρεύματος απαγωγού ( $I_D$ ) είχαμε αναλογία με την μεταβολή του φορτίου καναλιού ( $Q$ ) και αυτό επηρεαζόταν από τις μεταβολές του  $V_{GS}$  (τύπος 2.2.1), αλλά δεν είχαμε τις ίδιες αλλαγές στο φορτίο της περιοχής απογύμνωσης (τύπος 2.2.6) :

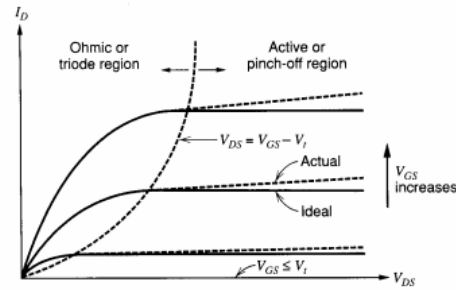
$$Q = qN_A X = \sqrt{2qN_A \epsilon \phi}, \quad (2.2.6)$$



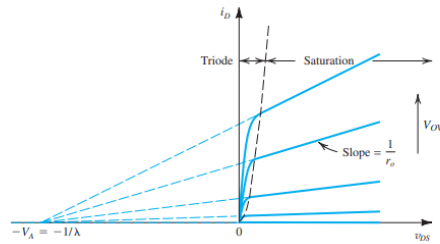
Σχήμα 2.2.2: Τα κανάλια για διάφορες τιμές των  $V_{DS}$  σχετικά με το  $V_{ov}$

όπου  $Q$  το φορτίο ανά μονάδα επιφάνειας στην περιοχή απογύμνωσης,  $\phi$  το δυναμικό στο στρώμα απογύμνωσης στη διασύνδεση του οξειδίου του πυριτίου,  $N_A$  η πυκνότητα πρόσμιξης του υποστρώματος τύπου  $p$  σε άτομα/ $cm^3$ ,  $\epsilon$  η ηλεκτρική διαπερατότητα του πυριτίου και  $q$  το φορτίο των ηλεκτρονίων. Ωστόσο, στις περιπτώσεις που η  $V_{GS} < V_{th}$  αλλά αρκετά μεγάλη ώστε να δημιουργήσει περιοχή απογύμνωσης στην επιφάνεια του πυριτίου έχουμε λειτουργία σε ασθενής αναστροφή (weak inversion). Στην περιοχή αυτή, έχουμε πολύ μικρότερο φορτίο καναλιού από ότι φορτίο περιοχής απογύμνωσης. Συνεπώς, το ρεύμα που δημιουργείται από φορείς πλειονότητας (δηλαδή το ρεύμα ολίσθησης), το οποίο ήταν υπαίτιο για το ρεύμα του απαγωγού στην ισχυρή αναστροφή, είναι αρκετά μικρότερο από το ρεύμα διάχυσης (που παράγεται λόγω των φορέων μειονότητας) και που επομένως είναι το μεγαλύτερο ποσοστό του  $I_D$  στην ασθενής αναστροφή. Ένα τρανζίστορ λοιπόν  $n$ -καναλιού θα λειτουργεί σαν ένα αντίστοιχο  $npn$  διπολικό (θυμίζουμε ότι στα διπολικά ισχύει  $I_C = I_S e^{\frac{V_{BE}}{V_T}}$ , με  $I_C$  το ρεύμα του συλλέκτη,  $I_S$  η σταθερά αναλογικότητας που ονομάζεται ρεύμα κορεσμού,  $V_{BE}$  την τάση βάσης-εκπομπού,  $V_T$  η θερμοκή τάση), με την πηγή να συμπεριφέρεται σαν εκπομπός (source as emitter), το υπόστρωμα σαν βάση (bulk as base) και τον απαγωγό σαν συλλέκτη (drain as collector) [49]. Όπως βλέπουμε και στο [2] το  $I_D$  παρουσιάζει μια εκθετική εξάρτηση από τη  $V_{GS}$ :

$$I_D = I_0 e^{\frac{V_{GS}}{nV_T}}, \quad (2.2.7)$$

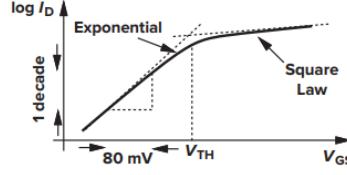


(α') Το κανάλι σε τριόδο [49]

(β') Το κανάλι σε κορεσμό για  $V_{DS} = V_{ov}$  [1]Σχήμα 2.2.3: Εξάρτηση ρεύματος υποδοχής  $I_D$  σχετικά με το  $V_{DS}$  και το  $V_{ov}$ 

με  $I_o$  ανάλογο του  $W/L$ ,  $n > 1$  ένας συντελεστής μη ιδανικότητας,  $V_T = kT/q$ . Αν αγνοήσουμε το  $n$ , η σχέση 2.2.7 είναι ίδια με την εκθετική σχέση  $I_C/V_{BE}$  ενός διπολικού τρανζίστορ. Αξίζει να σημειωθεί ότι το  $I_D$  μειώνεται με πεπερασμένο ρυθμό καθώς μειώνεται η  $V_{GS}$ , χαρακτηριστικό παράδειγμα είναι ότι σε θερμοκρασία δωματίου για τυπικές τιμές  $n$ , για να έχουμε πτώση 10db στο  $I_D$  πρέπει να πέσει 80 mV η  $V_{GS}$  όπως φαίνεται και στο σχήμα 2.2.4.

Για να διατυπωθεί μια έκφραση ακριβείς για το ρεύμα στην περιοχή υποκατωφλίου πρέπει να αναλογιστεί κάποιος το πως επηρεάζεται εκθετικά η συγκέντρωση των ηλεκτρονίων βάση της τάσης που προσδίδεται στην πύλη. Αύξηση της  $V_{GS}$  επάγεται αύξηση στο επιφανειακό δυναμικό  $\psi_s$ , που με τη σειρά του τείνει να μειώσει την αναστροφή πόλωση κατά μήκος πηγής-υποστρώματος και να αυξήσει εκθετικά την συγκέντρωση ηλεκτρονίων του υποστρώματος  $p$  στην πηγή  $n_p(0)$ . Έπειτα, κάνοντας αντικαταστάσεις και υπολογισμούς [49, 50] κα-



Σχήμα 2.2.4: Χαρακτηριστικές υποκατωφλίου για MOS [2]

ταλήγουμε στον τύπο:

$$I_D = \frac{W}{L} I_t e^{\left(\frac{V_{GS}-V_t}{nV_T}\right)} \left[1 - e^{-\left(\frac{V_{DS}}{V_T}\right)}\right] \quad \text{με} \quad (2.2.8)$$

$$I_t = qXD_n n_{p0} e^{\frac{k_2}{V_T}} \quad \text{και} \quad (2.2.9)$$

$$k_2 = k_1 + V_t/n \quad (2.2.10)$$

όπου  $k_1$  μια σταθερά για την οποία ισχύει  $\psi_s = \frac{V_{GS}}{n} + k_1$ ,  $n = (1 + C_{js}/C_{ox})$  ( $C_{js}$  χωρητικότητα περιοχής απογύμνωσης ανά μονάδα επιφάνειας),  $X$  είναι το πάχος της περιοχής που διαρρέει το ρεύμα,  $D_n$  σταθερά διάχυσης για τα ηλεκτρόνια,  $n_{p0}$  η συγκέντρωση ισορροπίας των ηλεκτρονίων στο υπόστρωμα. Αξίζει να σημειωθεί ότι, η ελάχιστη τιμή της τάσης  $V_{DS}$  που χρειάζεται το τρανζίστορ για να λειτουργήσει σαν πηγή ρεύματος δεν αλλάζει ιδιαίτερα με την αλλαγή της  $V_{ov}$  (τάσης υπεροδήγησης). Δηλαδή, αλλάζοντας την  $V_{ov}$  δεν επηρεάζει κάπως την ελάχιστη τιμή τάσης που χρειάζεται να έχει η  $V_{DS}$  ώστε το τρανζίστορ να λειτουργεί σαν πηγή ρεύματος βλέπε σχήμα 2.2.5α' [49].

Μια άλλη προσέγγιση του ρεύματος στην περιοχή υποκατωφλίου μπορούμε να βρούμε στο [51] που μας δίνει τους εξής τύπους για pMOS και nMOS:

$$I_{nmos} = I_{on} e^{k_n(V_{GS}-V_w)/V_T} (e^{(V_w-V_S)/V_T} - e^{(V_w-V_D)/V_T}) \quad (2.2.11)$$

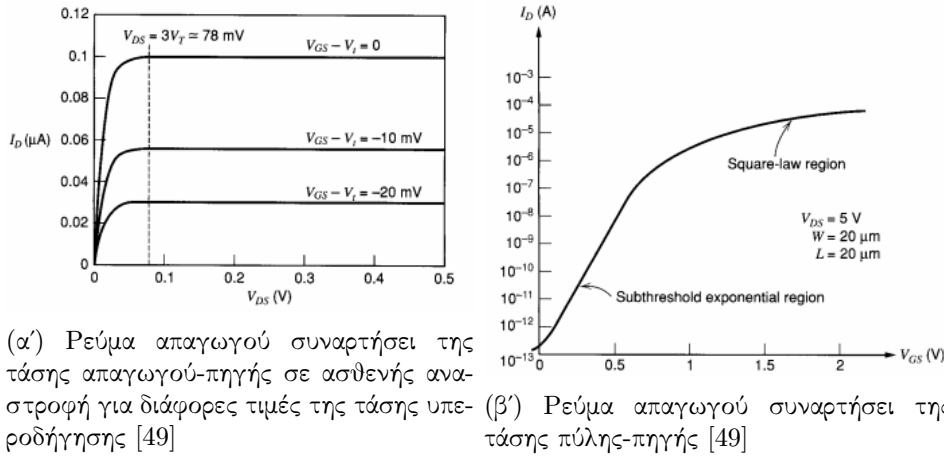
$$I_{pmos} = I_{op} e^{-k_p(V_{GS}-V_w)/V_T} (e^{-(V_w-V_S)/V_T} - e^{-(V_w-V_D)/V_T}) \quad (2.2.12)$$

όπου  $I_{op}$ ,  $I_{on}$  είναι προεκθετικοί όροι ρεύματος για τα δύο τρανζίστορ,  $V_w$  η τάση του σώματος,  $k_p$   $k_n$  τα slope factors των τρανζίστορ  $p$ -καναλιού και  $n$ -καναλιού αντίστοιχα, και γενικά έχουμε ότι το slope factor είναι:

$$k = \frac{C_{ox}}{C_{ox} + C_{js}} \quad (2.2.13)$$

είναι δηλαδή ίσο με το  $1/n$  που χρησιμοποιούσαμε στους παραπάνω τύπους.

Επεκτείνοντας τις σχέσεις 2.2.11 ,2.2.12 μπορούμε να αποφανθούμε ότι για

Σχήμα 2.2.5: Ρεύμα απαγωγού συναρτήσει των  $V_{GS}$ ,  $V_{DS}$ 

την περιοχή όπου  $V_{ds} > 4V_T$  έχουμε την περιοχή του κορεσμού στην subthreshold και επομένως στον κορεσμό, ομοίως με την ισχυρή αναστροφή, το ρεύμα του απαγωγού δεν έχει ιδιαίτερη εξάρτηση με την τάση απαγωγού-πηγής. Για αυτό μπορούμε να απλοποιήσουμε τους τύπους στον κορεσμό βγάζοντας τον παράγοντα της  $V_D$ , οπότε μετατρέπουμε την εξίσωση (για pMOS και αντίστοιχα βγαίνει και για nMOS):

$$I = I_0 e^{(-kV_G/V_T)} (e^{V_S/V_T} - e^{V_D/V_T}),$$

στην εξής:

$$I = I_0 e^{(-kV_g + V_s)/V_T}, \quad (2.2.14)$$

όπου στο  $I_0$  έχουμε συμπεριλάβει τον προεκθετικό όρο παράγοντες που μπορεί να ήταν εκτός αυτού σε προηγούμενες σχέσεις. Σε σύγκριση με τις 2.2.11, 2.2.12, στον τύπο 2.2.14 έχουμε θεωρήσει τις τάσεις πύλης, πηγής, απαγωγού σχετικά με την τάση σώματος.

Η διαγωγιμότητα στην περιοχή του υποκατωφλίου είναι η εξής:

$$g_m = \frac{I_d}{nV_T} \quad (2.2.15)$$

$$\frac{g_m}{I_D} = \frac{C_{ox}}{V_T(C_{js} + C_{ox})} \quad (2.2.16)$$

Ο τύπος 2.2.16 προκύπτει από την αντικατάσταση του  $n$  όπως είδαμε και πα-

ραπάνω. Θυμίζουμε ότι στον κορεσμό η τιμή της διαγωγιμότητας είναι:

$$g_m = \sqrt{\frac{2\mu_n C_{ox} W I_D}{L}} \quad (2.2.17)$$

$$g_m = \left(\frac{2I_D}{V_{OV}}\right) \quad (2.2.18)$$

Για την unity gain frequency  $f_T$  έχουμε ότι:

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs} + C_{gb} + C_{gd}} \quad (2.2.19)$$

Στην περιοχή του κόρου έχουμε ότι η  $C_{gs}$  είναι πολύ μεγαλύτερη των άλλων δύο. Στην περιοχή υποκατωφλίου, εφόσον το φορτίο στο στρώμα αναστροφής είναι μικρό η  $C_{gb}$  υπερικχύει των άλλων δύο. Οπότε σίγουρα έχουμε ότι το εξαρτώμενο κομμάτι της συχνότητας μοναδιαίου κέρδους από τον πυκνωτή μας δίνει μικρότερο λόγο (αφού είναι αντιστρόφως ανάλογα ποσά) για την ενεργό περιοχή. Ο λόγος  $g_m/I_d$  δεν έχει εξάρτηση από την τάση υπεροδήγησης στην περιοχή της subthreshold 2.2.16 σε αντίθεση με την ενεργό 2.2.18. Συμπεραίνουμε, λοιπόν, ότι αν αγνοήσουμε την τιμή του ρεύματος θα είχαμε υψηλότερη τιμή διαγωγιμότητας. Ωστόσο, στην περιοχή υποκατωφλίου είναι αρκετά περιορισμένο το ρεύμα του απαγωγού σε αντίθεση με την ισχυρή αναστροφή. Για αυτό και στην δεύτερη περιοχή μπορούμε να φτάσουμε μεγαλύτερες τιμές  $f_T$  για το ίδιο τρανζίστορ [2, 49, 51]. Η περιοχή του υποκατωφλίου έχει πολλά πλεονεκτήματα στη σχεδίαση ολοκληρωμένων κυκλωμάτων. Αρχικά, δίνει τη δυνατότητα για πολύ χαμηλής ισχύος εφαρμογές, της τάξης των *picoWatts* έως και *microWatts*. Έπειτα, το ρεύμα του απαγωγού έρχεται σε κορεσμό για πολύ λίγα  $kT/q$  ( $k$  η σταθερά του Boltzmann,  $T$  η θερμοκρασία), οπότε μπορεί να λειτουργήσει σε όλο το πεδίο τάσης από τη γη έως και  $V_{DD}$ . Ακόμη, η εκθετική φύση του τρανζίστορ στην περιοχή αυτή αποδεικνύεται απαραίτητη για διάφορες εφαρμογές [50]. Στον αντίποδα, μπορεί να φανεί ιδιαίτερα προβληματική σε μεγάλα κυκλώματα που η αγωγιμότητα κατωφλίου μπορεί να έχει μεγάλες απώλειες ισχύος ή απώλεια αναλογικών πληροφοριών [2]. Λογικό αν αναλογιστούμε ότι αυτά τα *pico*, *nano Ampere* αθροίζονται όταν έχουμε εκατομμύρια στοιχεία σε αρκετά υπολογίσιμη τιμή.

### 2.3 Το MOSFET ως ψηφιακό στοιχείο

Αναλύσαμε, την αναλογική λειτουργία ενός MOSFET. Είδαμε κάποιες από τις παραμέτρους μεγάλου και μικρού σήματος για να μας δώσουν τα πατήματα ώστε να μπορούμε να αντιληφθούμε την συνέχεια της παρούσας εργασίας. Στο σημείο

αυτό, κρίνεται απαραίτητο να δώσουμε μια ανάλυση του τρανζίστορ ως ψηφιακό στοιχείο. Να δούμε δηλαδή το βασικό χαρακτηριστικό που αξιοποιείται στις αντίστοιχες υλοποιήσεις, που δεν είναι άλλο παρά η ιδιότητα του να λειτουργεί σαν διακόπτης. Στην προηγούμενη ενότητα είδαμε την δομή των MOSFET σε επίπεδο φυσικής του στοιχείου. Από αυτή μπορούμε να αντιληφθούμε γιατί στην περίπτωση που έχουμε n-channel για να υπάρξει αγωγή του ρεύματος πρέπει να δοθεί μια τιμή  $V_{GS} > 0$  ώστε να μπορέσει να δημιουργηθεί ένα κανάλι αγωγής στο p-type bulk ανάμεσα στις  $n^+$  type περιοχές. Αντίστοιχα, γίνεται αντιληπτό ότι σε περίπτωση που έχουμε κανάλι τύπου p, θα υπάρχει σώμα τύπου n και ντοπαρισμένες περιοχές που θα καθορίσουν τα source drain τύπου  $p^+$ . Συνεπώς, για να μπορέσει να δημιουργηθεί ένα κανάλι τύπου p πρέπει να δοθεί εξωτερικά μια αρνητική τάση στην πύλη. Αν αναλογιστούμε, λοιπόν, τα τρανζίστορ αυτά ως ψηφιακά στοιχεία, αντλούμε το συμπέρασμα ότι λειτουργούν σχεδόν σαν ιδανικοί διακόπτες όπου στην πρώτη περίπτωση δίνουμε το λογικό "1" (που μπορεί να είναι η θετική τάση τροφοδοσίας σε αναλογική σκέψη) ώστε να κλείσει ο διακόπτης και να έχουμε αγωγή του ρεύματος ή το λογικό "0" (αρνητική τάση τροφοδοσίας ή γείωση ανάλογα το κύκλωμα) για να ανοίξει ο διακόπτης και να μην υπάρχει αγωγή ρεύματος. Για το pMOS αντιστοιχεί η συμπληρωματική λογική.

Σημαντικό είναι να γίνει μια αναφορά στο γεγονός ότι το nMOS μπορεί να περάσει από μέσα του σχεδόν ιδανικά το "0" αλλά όχι τόσο καλά το "1", για αυτό και λέμε ότι περνάει ισχυρό "0". Το pMOS αντίστοιχα περνάει ισχυρό "1" και όχι τόσο ιδανικά το "0". Για αυτό το λόγο και στις τεχνολογίες CMOS βλέπουμε ότι στο κομμάτι της θετικής τροφοδοσίας συνδέουμε τα pmos transistors και στο κομμάτι της αρνητικής (ή γείωσης ανάλογα) συνδέουμε τα nmos. Για να καταλάβουμε την αδυναμία του nmos να περάσει ισχυρό "1", αρκεί να αναλογιστούμε ένας nmos στην ενεργό περιοχή. Πρέπει να έχει  $V_{GS} > V_{th}$ , όμως  $V_{GS} = V_G - V_S = V_{DD} - V_{output} \Rightarrow V_{output} = V_{DD} - V_{th}$  οπότε βλέπουμε ότι για να έχουμε λειτουργία σαν διακόπτη του transistor δε μπορεί να περαστεί όλη η τάση που δίνεται στην είσοδο σωστά, ενώ στην περίπτωση που θέλουμε να περαστεί "0",  $V_{GS} = V_G - V_S = V_{DD} - 0 > V_{th}$  (θεωρούμε προφανές ότι έχουμε πλώσει σωστά το τρανζίστορ). Αντίστοιχα αποδεικνύεται και για τα pMOS [52].

Αυτά τα χαρακτηριστικά του MOS είναι και ο λόγος που μπορεί και αξιοποιείται η CMOS τεχνολογία ώστε να μπορεί να περάσει καλά τα υψηλά δυναμικά το pull up transistor (pmos) και τα χαμηλά το pull down (nmos). Τέλος, η λογική της τεχνολογίας αυτής είναι αρκετά απλή αφού χρησιμοποιείται η συμπληρωματικότητα του τρανζίστορ και πρακτικά γίνεται η δουλειά που χρειάζεται αξιοποιώντας το κάθε τρανζίστορ σαν έναν σχεδόν ιδανικό διακόπτη, στην περιοχή που δύναται να το κάνει [53].





## Κεφάλαιο 3

# Βασικά Κυκλώματα

Στο κεφάλαιο αυτό θα αναφερθούν κάποια βασικά κυκλώματα που αξιοποιούνται για την υλοποίηση του προτεινόμενου κυκλώματος που σχεδιάστηκε στην παρούσα διπλωματική. Θα αναλυθούν οι ιδιαιτερότητες και ο τρόπος λειτουργίας τους. Θα παρουσιαστούν πιθανές καλύτερες λύσεις των κυκλωμάτων αυτών και τροποποιήσεις που έγιναν, τόσο από τον συγγραφέα όσο και από άλλους, όπως παρουσιάστηκαν στην βιβλιογραφία, ώστε να πάρουμε τα καλύτερα δυνατά αποτελέσματα από κάθε υποκύκλωμα. Ακόμη, θα αναφέρουμε παρόμοιες υλοποιήσεις με την δική μας ώστε να δείξουμε τόσο την βάση πάνω στην οποία στηριχθήκαμε στη διπλωματική, όσο και να δώσουμε κάποια σύγκριση με αντίστοιχες υλοποιήσεις.

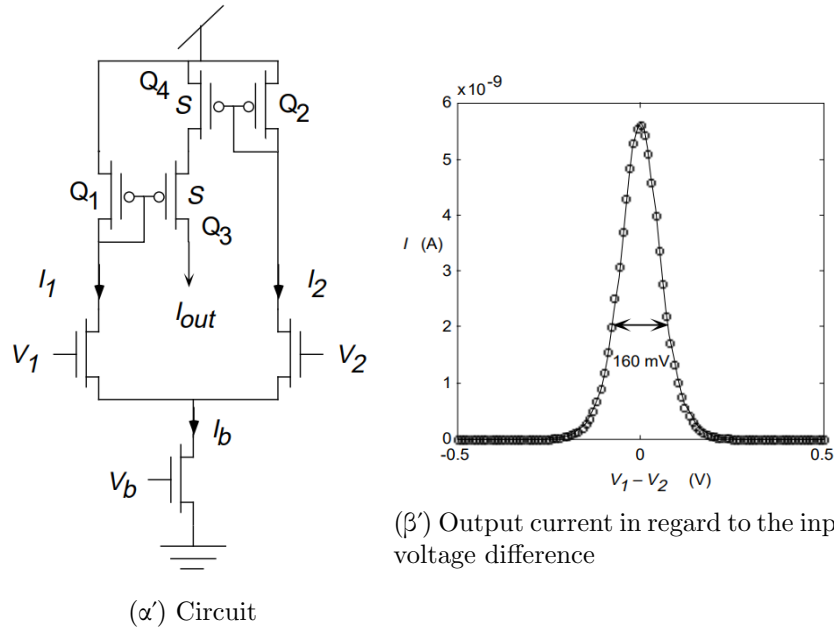
### 3.1 Bump Circuits

Bump circuits ονομάζονται τα κυκλώματα που καταφέρνουν να έχουν στην έξοδο τους μια Γκαουσιανή κατανομή (bell-shaped output). Για την ακρίβεια όπως αναφέρεται και στο [3], η ανάγκη για τον υπολογισμό ομοιότητας ή απόστασης δύο εισόδων, οδήγησε στη δημιουργία μιας κλάσης κυκλωμάτων που λέγονται bump circuits. Αυτή η ανάγκη δημιουργήθηκε λόγω των RBF (radial-basis-functions) που είναι μια κλάση δικτύων, που αξιοποιείται στη θεωρία των νευρωνικών δικτύων, η οποία βασίζεται σε μονάδες που υπολογίζουν την ομοιότητα μεταξύ διανυσμάτων εισόδου και διανυσμάτων που έχουν αποθηκευμένα [7, 54].

#### 3.1.1 Delbruck's bump circuit

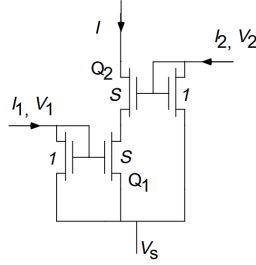
Το βασικό κύκλωμα πάνω στο οποίο πατήσαμε για την υλοποίηση που θα επέλθει στη συνέχεια αλλά και το βασικό κύκλωμα στο οποίο πατήσαν οι επόμενες

ενότητες που αναφέρονται σε αντίστοιχα bump-circuits είναι το simple bump circuit που προτάθηκε από τον T. Delbruck [3]. Το κύκλωμα αυτό μπορούμε να το δούμε στο σχήμα 3.1.1α' ενώ το ρεύμα εξόδου του προκύπτει βάση της διαφοράς στην τάση εισόδου του όπως φαίνεται στο σχήμα 3.1.1β'.



Σχήμα 3.1.1: Simple bump circuit as proposed by Delbruck [3]

Βλέποντας το σχήμα 3.1.1α' μπορούμε και διακρίνουμε τις επιμέρους τοπολογίες του. Στο κάτω μέρος βλέπουμε ότι υπάρχει ένα τρανζίστορ που λειτουργεί σαν πηγή ρεύματος και παράγει το ρεύμα  $I_b$  που καθορίζεται βάση μιας τάσης πόλωσης  $V_b$ . Έπειτα βλέπουμε έναν απλό διαφορικό ενισχυτή που δέχεται εισόδους  $V_1$   $V_2$  και βάση της διαφοράς τους μας δίνει κάποια σιγμοειδή ρεύματα, δηλαδή έχουμε συγκρίσιμα ρεύματα μόνο όταν οι τιμές των τάσεων αυτών είναι κοντινές, οπότε η διαφορά τους  $\Delta V$  τείνει στο μηδέν. Όταν, δηλαδή, έχουμε  $\Delta V$  μεγαλύτερο από λίγα  $kT/q$  το ένα εκ των δύο ρευμάτων (αυτό που έχει την μεγαλύτερη τάση), βγαίνει στην έξοδο, τραβάει δηλαδή όλο το ρεύμα που δίνεται από την πηγή και το άλλο "πόδι" του διαφορικού ενισχυτή βγαίνει εκτός λειτουργίας [1, 49]. Ανεβαίνοντας μετά το διαφορικό ζεύγος, έχουμε τον συσχετιστή ρεύματος δηλαδή τον current correlator που αποτελείται από τα τρανζίστορ  $Q_1$   $Q_2$   $Q_3$   $Q_4$ , ο οποίος παίρνει δύο ρεύματα και τα συσχετίζει μεταξύ τους. Υποθέτουμε, ότι τα τρανζίστορ  $Q_3$   $Q_4$  είναι  $S$  φορές ισχυρότερα από τα  $Q_1$   $Q_2$ , για να το αντιληφθούμε αυτό πρέπει να ριζούμε μία ματιά σε



Σχήμα 3.1.2: nmos current correlator, where  $S$  is the strength ratio

μεγαλύτερο βάθος στον current correlator όπως παρουσιάζεται στο [3].

Θα αναλύσουμε τον nMOS current correlator που είναι ο συμπληρωματικός του ανώτερου και φαίνεται στο σχήμα 3.1.2. Στην περιοχή της subthreshold τα τρανζίστορ που είναι συνδεδεμένα σε σειρά λειτουργούν σαν συνδυασμός αναλογικών πυλών AND. Σε περίπτωση που κάποια τάση πύλης είναι χαμηλή τότε το ρεύμα εξόδου  $I$  κλείνει, ενώ όταν και οι δύο τάσεις στις πύλες είναι υψηλές έχουμε αξιόλογο ρεύμα εξόδου. Στην ενδιάμεση κατάσταση, το κύκλωμα υπολογίζει έναν τύπο πολλαπλασιασμού των ρευμάτων εισόδου  $I_1 I_2$ . Σαν strength του τρανζίστορ ορίζεται ο λόγος  $W/L$  και συγκεκριμένα για τον current correlator έχουμε ότι:

$$S = \frac{(W/L)_{middle}}{(W/L)_{outer}} \quad (3.1.1)$$

Η παράμετρος αυτή είναι πολύ σημαντική στη λειτουργία του κυκλώματος καθώς στην περίπτωση που έχουμε κορεσμό μπορούμε να καταλήξουμε αξιοποιώντας και τις σχέσεις 2.2.8 στην εξής σχέση για το ρεύμα εξόδου:

$$\begin{aligned} I &= S e^{-V_s} \frac{e^{V_1} e^{V_2}}{e^{V_1} + e^{V_2}} \\ &= S \frac{I_1 I_2}{I_1 + I_2} \end{aligned} \quad (3.1.2)$$

Όσο το κύκλωμα λειτουργεί στην περιοχή υποκατωφλίου υπάρχει συμμετρία αναφορικά με τα ρεύματα εισόδου. Το προϊόν που παράγεται από τον συσχετιστή ρεύματος είναι μια self-normalized συσχέτιση, όπου βλέπουμε ότι το ρεύμα εξόδου είναι ανάλογο γινόμενο των ρευμάτων εισόδου διαιρούμενο με το άθροισμα τους [3].

Επιστρέφοντας, λοιπόν, στο simple bump, ο υπολογισμός του κάθε διαφο-

ρικού ρεύματος βγαίνει από την σχέση Fermi σχετικά με το  $\Delta V$ :

$$I_1 = \frac{I_b}{1 + e^{-k\Delta V}} \quad (3.1.3)$$

και μέσω της απλής σχέσης 3.1.2, παίρνουμε:

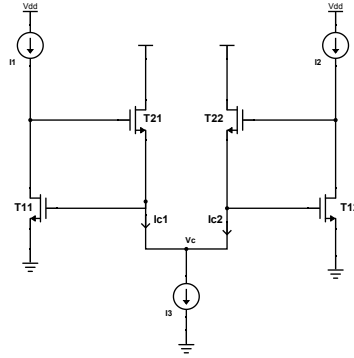
$$\begin{aligned} I_{out} &= I_b \frac{S}{4} \operatorname{sech}^2\left(\frac{k\Delta V}{2}\right) \\ &= \frac{I_b}{\frac{4}{S} \cosh^2 \frac{k\Delta V}{2}} \end{aligned} \quad (3.1.4)$$

Από την σχέση αυτή βλέπουμε ότι η έξοδος θα είναι μια "καμπάνα" η οποία έχει κέντρο  $\Delta V = 0$  και μέγιστο ύψος  $SI_b/4$ . Η σχέση αυτή μας προσδίδει ότι φαίνεται η ομοιότητα των εισόδων  $V_1$ ,  $V_2$  και συνεπώς αλλάζοντας εκείνες τις τιμές μπορούμε να μεταβάλλουμε και το κέντρο της καμπάνας. Αυξάνοντας το  $I_b$  αυξάνεται το ύψος του bump. Στο σχήμα 3.1.1β', βλέπουμε την έξοδο του simple bump συναρτήσει της διαφοράς  $\Delta V$  όπως παρουσιάστηκε από τον Delbruck [3].

### 3.2 Winner Take All circuit (WTA)

Στις προηγούμενες ενότητες, πήραμε μια ιδέα αναφορικά με το πως σχεδιάζεται ένα αναλογικό κύκλωμα bump στην περιοχή υποκατωφλίου. Αν, θυμηθούμε τη θεωρία που αναφέραμε για τους Γκαουσιανούς βασισμένους ταξινομητές πέρα από το κομμάτι του bump, δηλαδή της Γκαουσιανή, χρειαζόμαστε και έναν συγκριτή που θα παίρνει την πιθανότητα που υπολογίστηκε, στην περίπτωση μας θα είναι ένα ρεύμα, και να βγάζει στην έξοδο του το μεγαλύτερο. Χρειαζόμαστε, δηλαδή μια συνάρτηση  $\operatorname{argmax}$ . Το ρόλο αυτό έρχεται να παίξει ένα κύκλωμα που προτάθηκε από τον Lazzaro και ονομάζεται Winner Take All (WTA) [55], αυτό φαίνεται στο σχήμα 3.2.1. Το σχήμα 3.2.1 είναι συγκεκριμένα για 2 νευρώνες αλλά ακριβώς την ίδια αναλογία ακολουθούμε και για περισσότερους, δηλαδή για περισσότερες εισόδους εξόδους. Ακόμα, το κύκλωμα αυτό λειτουργεί και για pmos και για nmos transistors, για υλοποίηση με pmos έχουμε ακριβώς την συμπληρωματική λογική.

Συνοπτικά, το κύκλωμα δέχεται δύο ρεύματα  $I_1$  και  $I_2$  τα οποία συγκρίνει μεταξύ τους και μας δίνει στην έξοδο του νευρώνα του οποίου το ρεύμα είναι μεγαλύτερο, όλο το ρεύμα πόλωσης  $I_c$  που δέχεται. Αντίστοιχη λογική έχει και σε περισσότερες εισόδους, δηλαδή συγκρίνει όλα τα ρεύματα  $I_i$  και δίνει το ρεύμα πόλωσης εξ'ολοκλήρου στον νευρώνα που έχει μεγαλύτερο ρεύμα  $I_i$ . Στο σχήμα 3.2.1, βλέπουμε δύο νευρώνες που αποτελούνται από 2 τρανζίστορ



Σχήμα 3.2.1: nMOS WTA με 2 νευρώνες

ο καθένας  $T_{11}$ ,  $T_{21}$  και  $T_{12}$ ,  $T_{22}$ .

Αναλυτικότερα, το κύκλωμα αυτό έχει τρεις περιπτώσεις λειτουργίας. Όταν τα ρεύματα εισόδου είναι μεταξύ τους ίσα και όταν κάποιο εκ των δύο είναι μεγαλύτερο. Θα εξηγήσουμε την περίπτωση  $I_1 = I_2$  και την περίπτωση που  $I_1 > I_2$ , για την περίπτωση που  $I_1 < I_2$  θα έχουμε την συμπληρωματική διαδικασία με την δεύτερη περίπτωση. Αξίζει να σημειωθεί ότι η αναλογία σε όλα τα τρανζίστορ είναι 1:1, σε διαφορετική περίπτωση ο νευρώνας που έχει μεγαλύτερη δύναμη, για παράδειγμα  $m:1$ , θα οδηγήσει το κύκλωμα σε μια λειτουργία που θα αναλογούσε με αυτή που θα είχαμε αν αντί για αύξηση κατά  $m$  του μεγέθους των τρανζίστορ, είχαμε μείωση στο ρεύμα, της τάξης  $I_{in}/m$ , οπότε για να πάρει όλο το ρεύμα πόλωσης ο νευρώνας αυτός θέλει  $m$  φορές μεγαλύτερο ρεύμα από ότι ήθελε στην περίπτωση που η αναλογία ήταν 1:1. Αυτό συμβαίνει γιατί η τάση πύλης του μεγαλύτερου τρανζίστορ θα είναι μικρότερη (μεγαλύτερο τρανζίστορ αλλά το ρεύμα απαγωγού ίδιο με πριν), οπότε θα χρειάζεται να εξισορροπήσει την τάση του με την προηγούμενη που γίνεται για ρεύμα  $m$  φορές μεγαλύτερο από την προηγούμενη κατάσταση του.

Θεωρούμε ότι  $I_1 = I_2$ , τότε  $V_{g11} = V_{g12}$  αφού είναι βραχυκυκλωμένα μεταξύ τους και  $V_s$  είναι ίδια μεταξύ τους, αφού είναι συνδεδεμένα (στο σχήμα 3.2.1) στην γη, ή γενικότερα στο χαμηλό δυναμικό που θα του παρέχεται. Επομένως, η τάση του απαγωγού θα είναι ίδια στα δύο αυτά τρανζίστορ. Αυτό συνεπάγεται ότι  $V_1 = V_2$  που είναι και η τάση στις πύλες των τρανζίστορ  $T_{21}$ ,  $T_{22}$ . Τα τρανζίστορ αυτά, συνεπώς, έχουν ίδια τάση στις πύλες, πηγές και απαγωγούς τους και έτσι δίνουν το ίδιο ρεύμα εξόδου που θα είναι  $I_{out1} = I_{out2} = I_c/2$ .

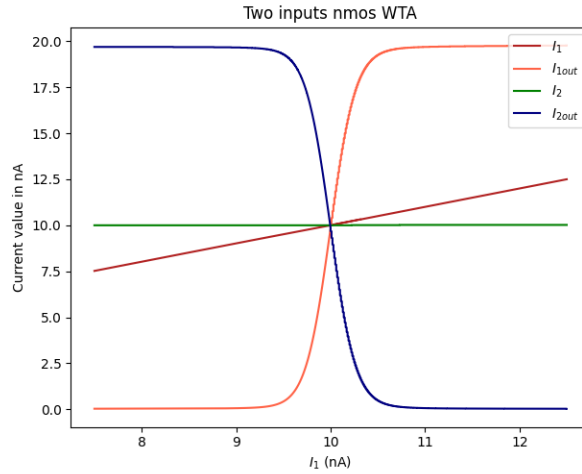
Στην περίπτωση που έχουμε  $I_1 = I_m + \delta_i$ ,  $I_2 = I_m$  τότε το τρανζίστορ  $T_{11}$  καλείτε να οδηγήσει μεγαλύτερο ρεύμα από ότι στην προηγούμενη περίπτωση

οπότε αυξάνεται λίγο η τάση της πύλης του που συνεπάγεται αύξηση της τάσης της πύλης του τρανζίστορ  $T_{12}$  το οποίο όμως δέχεται εξωτερικά ρεύμα  $I_2$  στον απαγωγό του, οπότε δεν μπορεί να αλλάξει την τιμή του ρεύματος ανάλογα με αυτή που του υποδεικνύεται από την τάση αυτή. Για αυτό το λόγο, ώστε να μπορέσει να επέλθει ισορροπία ρίχνει λίγο την τάση στον απαγωγό του  $V_2$ . Για μικρές μεταβολές  $\delta_i$  (η τιμή των μεταβολών  $\delta_i$  εξαρτάται άμεσα από τα ρεύματα που δίνονται στους νευρώνες του WTA) η τάση στον απαγωγό μειώνεται μέσω του φαινομένου *early*, ωστόσο φτάνοντας μεγαλύτερες τιμές, το τρανζίστορ  $T_{12}$  βγαίνει από τον κορεσμό και επομένως η  $V_2$  γίνεται ίση με το μηδέν, οπότε το  $T_{22}$  σβήνει σαν τρανζίστορ και δεν έχουμε αγωγή ρεύματος από αυτόν τον νευρώνα, για αυτό και όλο το ρεύμα πόλωσης  $I_c$  περνάει από το τρανζίστορ  $T_{21}$ .

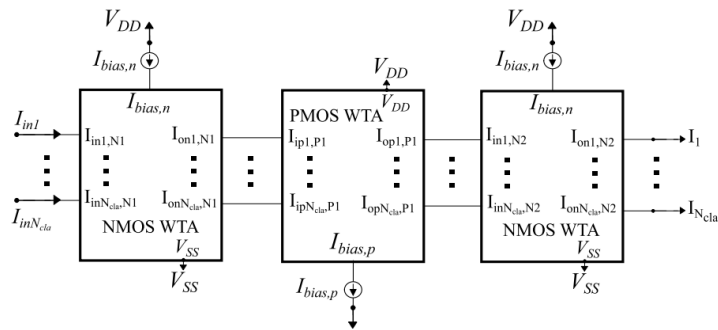
Συμπεραίνουμε λοιπόν ότι, όπως φαίνεται και στο σχήμα 3.2.2 στην περίπτωση που έχουμε ένα ρεύμα κάποιου νευρώνα μεγαλύτερο από τα υπόλοιπα του wta, το οποίο είναι αρκετά μεγάλο ώστε να μην μπορεί να αντισταθμίζεται η αλλαγή που επιφέρει τότε στην έξοδο του κυκλώματος παίρνουμε το ρεύμα πόλωσης από τον νευρώνα του ρεύματος αυτού. Σε περίπτωση που είναι ίσα τα ρεύματα εισόδου, θα έχουμε ίση κατανομή του ρεύματος πόλωσης του κυκλώματος στους νευρώνες αυτούς, ενώ στην κατάσταση που κάποιο ρεύμα είναι μεγαλύτερο από τα άλλα, αλλά όχι αρκετά ώστε να βγάλει τους υπόλοιπους νευρώνες σε αποκοπή, έχουμε μια ενδιάμεση κατάσταση όπου η μετάβαση από κάποια έξοδο σε κάποια άλλη γίνεται με έναν γραμμικό τρόπο. Αυτή η ενδιάμεση κατάσταση προκαλεί σε κάποιες περιπτώσεις *multiple winners* που μπορεί να μας προκαλέσει διάφορα σφάλματα, για αυτό έχει προταθεί μια βελτίωση του WTA η οποία θα φανεί στην επόμενη ενότητα [4]. Ωστόσο, η περιοχή που δεν έχουμε απόλυτη απόφαση για κάποιον νικητή μπορεί να αξιοποιηθεί σε κάποιο διαφορετικό κύκλωμα για συγκεκριμένους σκοπούς όπως θα δούμε και παρακάτω στην προτεινόμενη υλοποίηση.

### 3.3 Improved WTA

Όπως βλέπουμε και στο [4], για να μπορέσει το κύκλωμα να βγάλει στην έξοδο την σωστή απόφαση χωρίς να έχει πολλαπλούς νικητές όπως γίνεται στην περίπτωση που τα ρεύματα είναι συγκρίσιμα στην γραμμική περιοχή λειτουργίας του, όπως είδαμε και παραπάνω, αξιοποιεί τρία κυκλώματα WTA. Παίρνει δηλαδή τα ρεύματα εισόδου και τα προσδίδει σε ένα κύκλωμα WTA, *pmos* ή *nmos* ανάλογα με την έξοδο από την οποία προήλθαν τα ρεύματα. Έπειτα, τα ρεύματα εξόδου τα περνάει από ένα συμπληρωματικό WTA κύκλωμα και τα ρεύματα αυτουνού από ένα συμπληρωματικό ως προς αυτό. Έχει δηλαδή μια διάταξη που δίνει αρχικά ένα ρεύμα έστω σε *nmos* WTA, αυτό με τη σειρά του

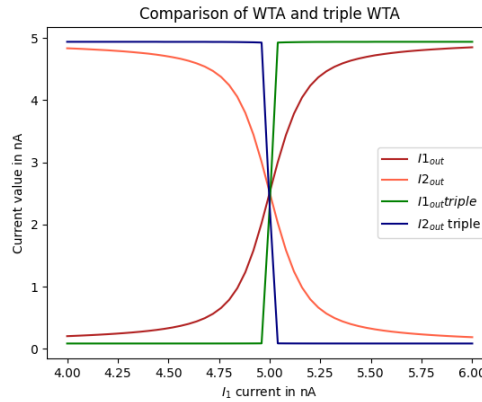


Σχήμα 3.2.2: Ρεύματα εξόδου από τους νευρώνες του WTA σε σχέση με τα ρεύματα εισόδων τους, αλλάζοντας το ρεύμα εισόδου  $I_1$  όπως φαίνεται



Σχήμα 3.3.1: Improved Winner Take All circuit [4]

τα παράγωγα ρεύματα του τα δίνει σαν εισόδους σε ένα pmos WTA το οποίο δίνει τα ρεύματα των εξόδων του σε ένα τρίτο και τελευταίο nmos WTA το οποίο θα δώσει τα τελικά ρεύματα εξόδου με μια αρκετά καλύτερη ακρίβεια από ότι είχαμε πρωτύτερα στο απλό WTA. Μπορούμε να δούμε και μια σύγκριση που κάναμε προσομοιώνοντας τις δύο διατάξεις αυτές στο σχήμα 3.3.2

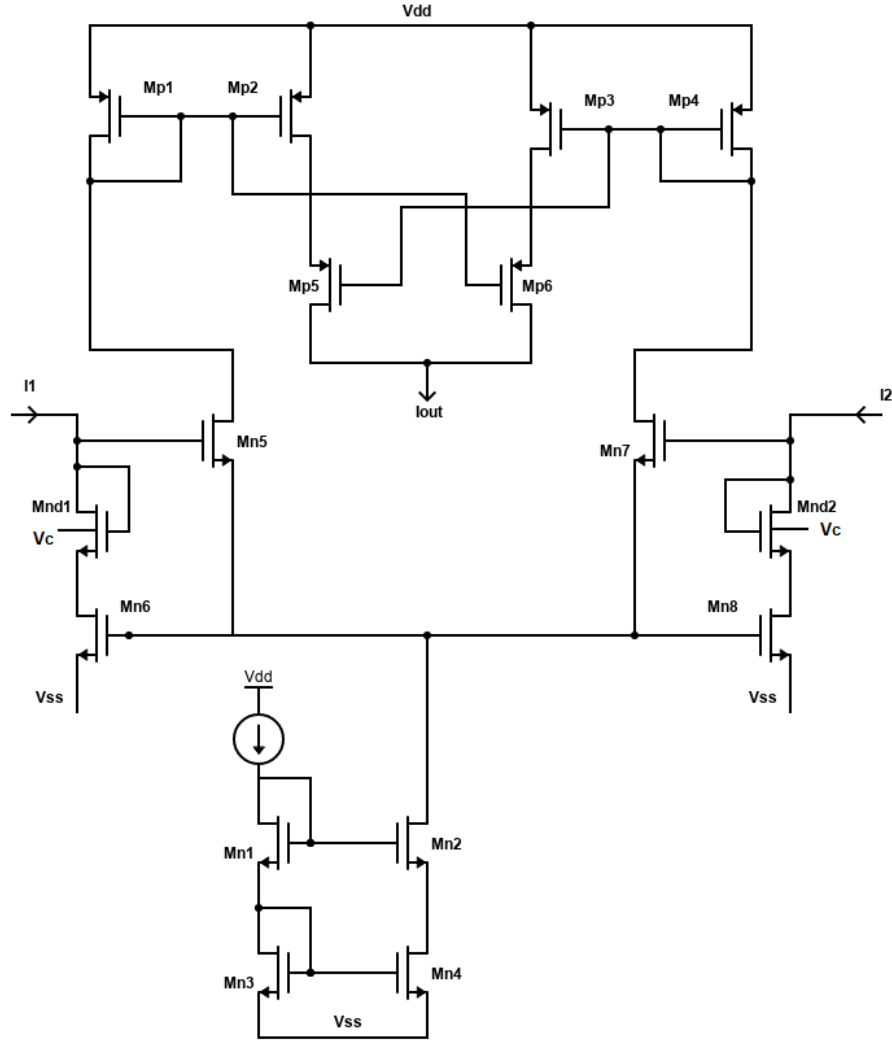


Σχήμα 3.3.2: Αποτελέσματα για το ρεύμα εξόδου του improved WTA κυκλώματος σε σύγκριση και με ένα αντίστοιχα απλό WTA

### 3.4 WTA bump circuit

Μια παραλλαγή του bump circuit αξιοποιώντας το simple WTA φαίνεται στο σχήμα 3.4.1 και αποτελεί πρακτικά ένα απλό bump circuit όπως είδαμε παραπάνω με την διαφορά ότι αυτό το κύκλωμα έχει καταφέρει να δέχεται ρεύματα αντί για τάσεις στην είσοδο του και βάση αυτών δημιουργεί την Γκαουσιανή με τα ανάλογα κέντρα της. Στο κάτω μέρος του κυκλώματος βλέπουμε έναν κασχοδικό καθρέφτη ρεύματος που μας δίνει καλύτερο λόγο καθρεπτισμού τόσο σε μικρά όσο και σε μεγάλα ρεύματα, αφού ο κασχοδικός καθρέφτης μας προσομοιάζει καλύτερα μια ιδανική πηγή. Έχοντας υψηλή αντίσταση εξόδου δεν επηρεάζεται πολύ από την τάση εξόδου που έρχεται από το υπόλοιπο κύκλωμα. Ανεβαίνοντας στο κύκλωμα, βλέπουμε ότι πλέον δεν υπάρχει το διαφορικό ζεύγος που είχε το simple bump, αλλά έχουμε τους νευρώνες που είδαμε στο WTA, ενδιάμεσα τους έχουμε μια δίοδο που έχει συνδεδεμένο το σώμα της σε μία εξωτερική τάση  $V_c$  και μέσω αυτής μπορούμε να ελέγξουμε την διασπορά του κυκλώματος, το πάχος δηλαδή της καμπάνας. Πέρα από αυτό, η δίοδος στο συγκεκριμένο σημείο μας δίνει πολύ πιο απότομες Γκαουσιανές, αυτή η λειτουργία σε συνεργασία με το γεγονός ότι ελέγχεται από τα ρεύματα το συγκεκριμένο κύκλωμα μας επιτρέπει να έχουμε πιο πυκνές Γκαουσιανές και δεν περιοριζόμαστε στο εύρος όπως γινόταν με το simple bump. Τέλος, αφού περάσουμε το στάδιο των νευρώνων το ρεύμα εξόδου που θα έρθει από αυτά, μπαίνει σε έναν συμμετρικό current correlator. Το κομμάτι αυτό του κυκλώματος, έγινε συμμετρικό για σκοπούς διατήρησης της συμμετρίας της Γκαουσιανής καμπύλης και πρακτικά, η συμμετρία για τις εισόδους γύρω από τη μέση τιμή



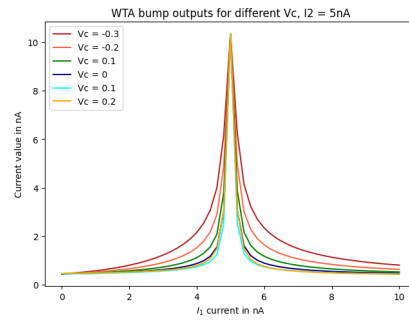
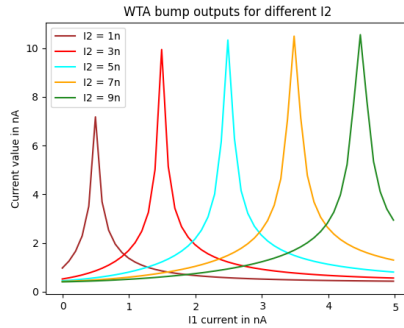


Σχήμα 3.4.1: WTA bump circuit

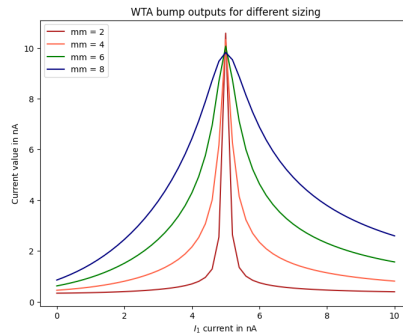
μπορεί και διατηρείται ακόμη και για πολύ μικρά ρεύματα. Με αυτές τις αλλαγές του simple bump το κύκλωμα αυτό καταφέρνει και λειτουργεί τόσο για χαμηλές τιμές ρεύματος πόλωσης, εισόδου όσο και για χαμηλές τιμές των ρευμάτων στην είσοδο του συσχετιστή ρεύματος.

Πίνακας 3.1: Διαστάσεις τρανζίστορ (Σχήμα 3.4.1).

nMOS transistors	W/L ( $\mu m/\mu m$ )	pMOS transistors	W/L ( $\mu m/\mu m$ )
$M_{n1}-M_{n3}$	0.4/1.6	$M_{p1}, M_{p6}$	0.4/1.6
$M_{n4}$	0, 8/1.6	-	-
$M_{n5}, M_{n7}$	1.2/1.6	-	-
$M_{n6}, M_{n8}$	1.2/3.2	-	-
$M_{nd1}, M_{nd2}$	0.4/0.2	-	-



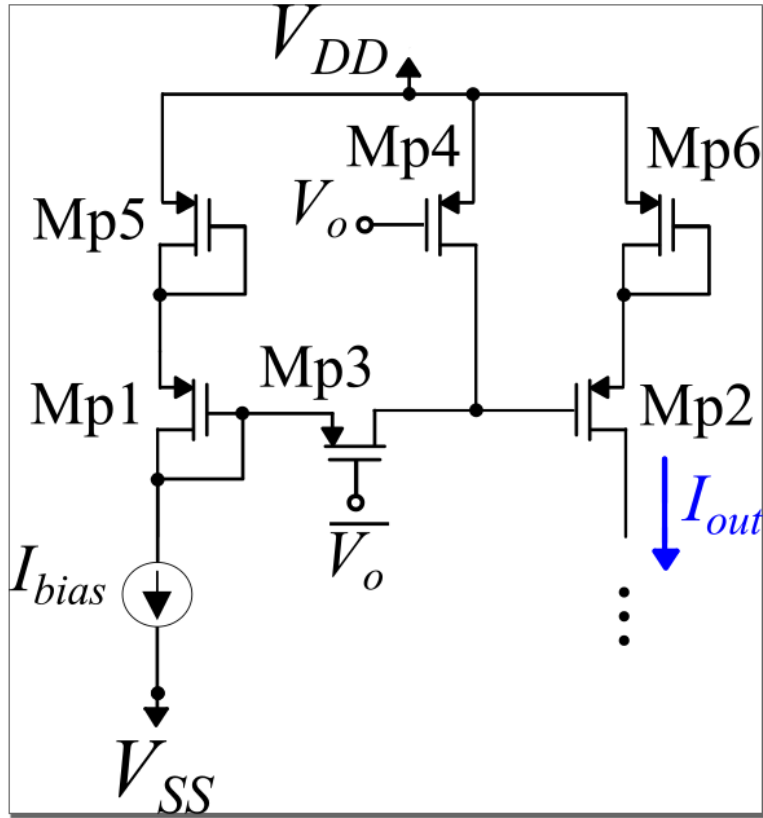
( $\alpha$ ) WTA bump circuit με διαφορετικές ( $\beta$ ) WTA bump circuit με διαφορετικές τιμές διασποράς  $V_c$   
 μέσες τιμές  $I_2$



( $\gamma$ ) WTA bump circuit με διαφορετικές τιμές μεγέθους τρανζίστορ νευρώνων mm

### 3.5 Ελεγχόμενος καθρέφτης ρεύματος

Στο σημείο αυτό θα αναλυθεί ένα κύκλωμα καθρέφτη ρεύματος το οποίο μπορεί να ελέγχεται εξωτερικά από κάποιες τάσεις που δίνει ο χρήστης. Αυτές,



Σχήμα 3.5.1: Ελεγχόμενος καθρέφτης ρεύματος

επιτρέπουν στο κύκλωμα να καθρεφτίσει ή όχι το ρεύμα της εισόδου του στην έξοδο. Το κύκλωμα αυτό είναι πρακτικά ένας απλός καθρέφτης ρεύματος με δύο διόδους πάνω από τα δύο τρανζίστορ του βασικού κυκλώματος. Ο σκοπός αυτών των διόδων είναι να ρίξουν το offset που είχαμε στο κύκλωμα καθώς και να ρίξουν το συνολικό εύρος ρεύματος εξόδου που μας έδινε από on σε off, δηλαδή να μειώσουν τη διαφορά του ρεύματος που παίρναμε στην έξοδο όταν στην είσοδο είχαμε τον καθρέπτη να καθρεφτίζει το ρεύμα του ή όχι  $I_{on} - I_{off}$ . Ενδιάμεσα από το κύκλωμα του απλού καθρέφτη βλέπουμε ότι έχουν μπει 2 τρανζίστορ που λαμβάνουν τις εισόδους σε συμπληρωματική μορφή. Το κάτω τρανζίστορ  $M_{p3}$  δέχεται στην πύλη του μία τάση  $V_0'$  η οποία είναι η τάση του ελέγχου, αφού περαστεί από έναν κλασικό CMOS αντιστροφέα. Αντίστοιχα, το πάνω τρανζίστορ  $M_{p4}$  δέχεται στην πύλη του την τάση ελέγχου απευθείας. Πρακτικά, αν δώσουμε εξωτερικά μια τάση  $V_0 = "0" = V_{ss}$  θα βγει σε αποκοπή

Πίνακας 3.2: Διαστάσεις τρανζίστορ (Σχήμα 3.5.1).

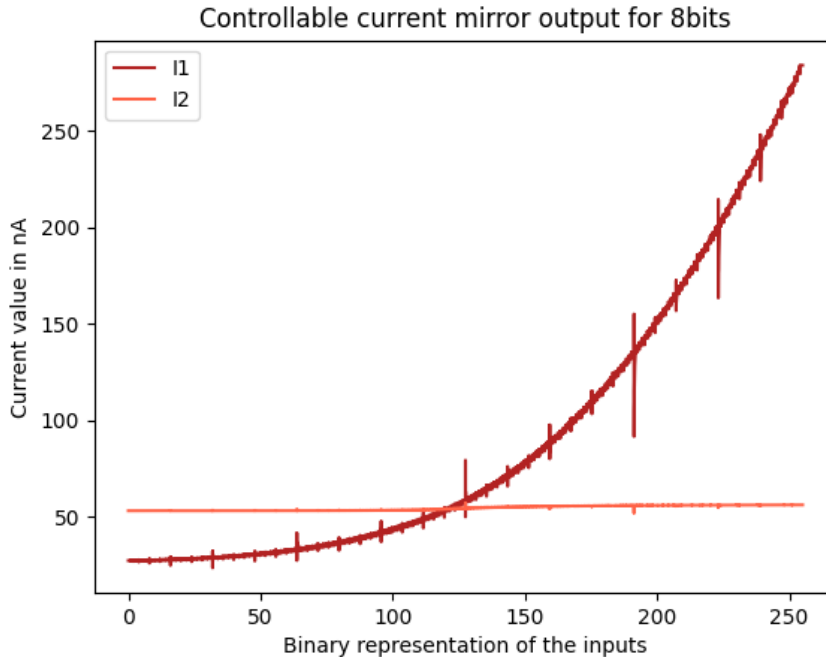
pMOS tran- sistors	W/L ( $\mu m/\mu m$ )
$M_{p1}-M_{p6}$	0.4/0.5

το κάτω τρανζίστορ (απενεργοποιηθεί ο MOSFET διακόπτης) οπότε δεν θα έχουμε καθρεπτισμό του ρεύματος στην έξοδο και ακόμα θα ενεργοποιηθεί ο πάνω διακόπτης MOSFET ο οποίος με αυτό τον τρόπο φέρνει την τάση  $V_{dd}$  (υψηλή τάση του κυκλώματος) στην πύλη του τρανζίστορ στην έξοδο, το οποίο με αυτό το τρόπο θα βγει σε αποκοπή και θα έχουμε στην έξοδο ένα σχετικά μικρό ρεύμα leakage. Σε αντίθετη περίπτωση αν δοθεί εξωτερικά  $V_0 = "1" = V_{dd}$  στο κύκλωμα του ελεγχόμενου καθρέφτη, το κάτω τρανζίστορ θα λειτουργήσει σαν κλειστός διακόπτης και το πάνω σαν ανοιχτός, οπότε στην πύλη του τρανζίστορ εξόδου θα έχουμε την τάση της πύλης του τρανζίστορ εισόδου και συνεπώς θα έχουμε καθρεπτισμό του ρεύματος. Αξίζει να σημειωθεί ότι οι είσοδοι στο κομμάτι αυτό του κυκλώματος θεωρούνται καθαρά ψηφιακές έχουμε δηλαδή σαν λογικό "0" το high voltage του κυκλώματος, δηλαδή την θετική τροφοδοσία και σαν "1" την αρνητική τροφοδοσία low voltage. Αυτή η "ψηφιακή" λογική του κυκλώματος θα μας φανεί αρκετά χρήσιμη στην υλοποίηση μας.

Αξίζει να σημειωθεί ότι το συγκεκριμένο κύκλωμα από μόνο του έχει κατανάλωση  $0.826\mu W$  στην περίπτωση που έχουμε 8 bits είσοδο για τάση τροφοδοσίας  $0.5V$ , ενώ για  $0.6V$  έχει κατανάλωση  $1.24\mu W$ . Αξίζει να αναφερθεί ότι το κύκλωμα αυτό υπερτερεί ή έστω συναγωνίζεται της βιβλιογραφίας όπως είδαμε και στα [56, 57, 58, 59] τα οποία μας δίνουν αντίστοιχα  $2mW$ ,  $850\mu W$ ,  $561\mu W$  και το πιο low power που βρήκαμε, το οποίο λειτουργούσε με τάση τροφοδοσίας  $0.5V$  είχε κατανάλωση του DAC  $0.9\mu W$ .

### 3.6 Προτεινόμενη υλοποίηση

Στο κεφάλαιο αυτό θα παρουσιάσουμε την προτεινόμενη υλοποίηση που έγινε για τους σκοπούς της διπλωματικής. Θα παρουσιαστεί, δηλαδή, ο αναλογικός ταξινομητής που δέχεται ψηφιακή είσοδο και βγάζει ψηφιακή έξοδο. Θα αναφερθούμε ακόμη και στα πλεονεκτήματα του κυκλώματος αυτού έναντι σχετικών κυκλωμάτων και κάποιοι πιθανοί περιορισμοί του. Θα εξηγηθεί σε βάθος η λειτουργία του και ενδεχόμενες απορίες που μπορεί να προκύψουν για τη συγκεκριμένη υλοποίηση.

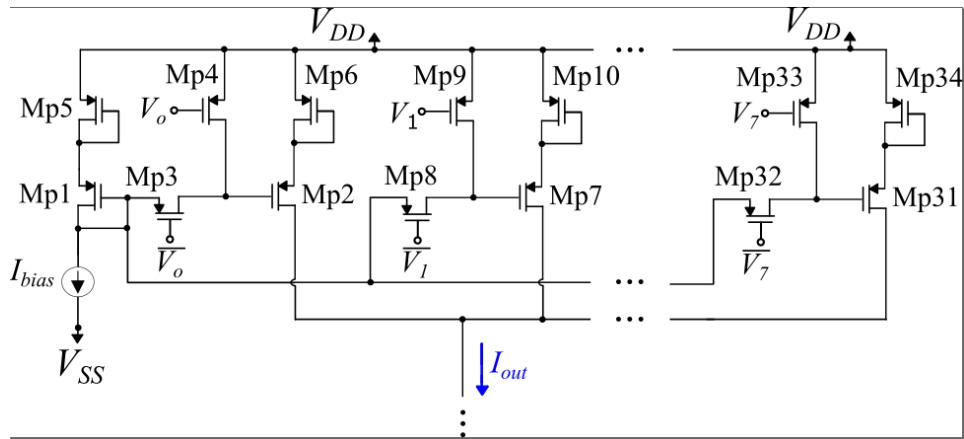


Σχήμα 3.5.2: Έξοδος ενός 8bit controllable current mirror που δέχεται όλους τους αριθμούς από 0 έως και 255 ( $I_1$ ) σε αναλογία με έναν που δέχεται μόνο την τιμή 60  $I_2$ )

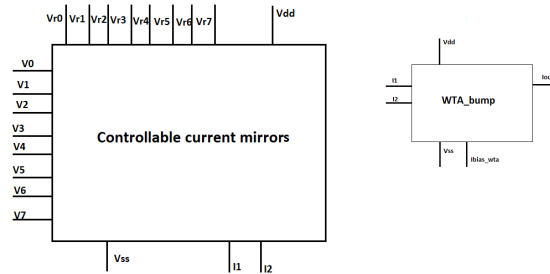
### 3.6.1 Προτεινόμενο bump

Σε αυτό το σημείο θα αναφερθούμε στην υλοποίηση που επιδιώξαμε να κάνουμε. Ο σκοπός του κυκλώματος μας είναι να μπορεί να παίρνει απευθείας τα ψηφιακά χαρακτηριστικά που θα γίνουν extracted από ένα dataset, για να κάνει κάποια ταξινόμηση, χωρίς αυτά να πρέπει να περάσουν πρώτα από έναν Digital to Analog Converter (DAC). Με αυτό το τρόπο η κατανάλωση του κυκλώματος μπορεί να μειωθεί ραγδαία καθώς εξαλείφουμε τελείως το power hungry ψηφιακό κύκλωμα που θα μας έκανε την μετατροπή των εισόδων σε αναλογικές τάσεις. Το κύκλωμα που παρουσιάζεται έχει την καινοτομία, λοιπόν, να χρησιμοποιεί ψηφιακές εισόδους χωρίς την αξιοποίηση κάποιου καθαρά ψηφιακού κυκλώματος, να αξιοποιεί την χαμηλή κατανάλωση που μπορεί να έχει μόνο ένα αναλογικό κύκλωμα και να δίνει στην έξοδο ξανά ένα ψηφιακό σήμα (ρεύμα).

Πρακτικά το κύκλωμα που υλοποιεί το bump αποτελείται από κάποιους ελεγχόμενους διακόπτες των οποίων οι έξοδοι εισέρχονται σε ένα WTA bump



(α') 8 bits ελεγχόμενος καθρέφτης

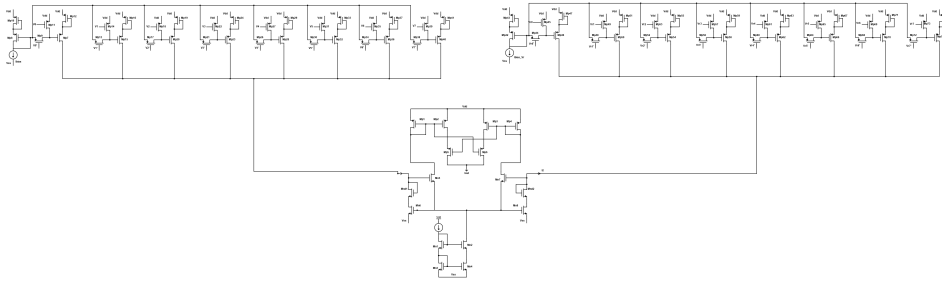


(β') Σύμβολο του 8 bits ελεγχόμενου καθρέφτη

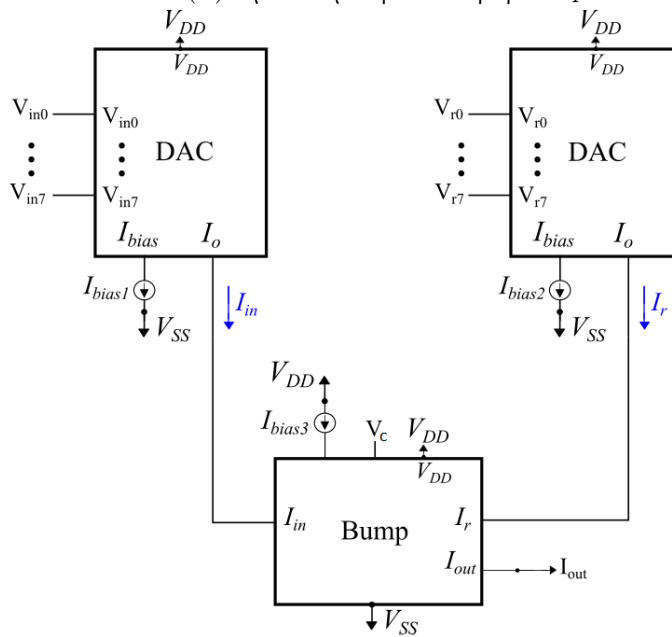
(γ') Σύμβολο του WTA-bump

το οποίο θα μας δώσει το αποτέλεσμα. Η έξοδος αυτή μετέπειτα θα μπει σε ένα WTA κύκλωμα για να συγκριθεί με τις αντίστοιχες εξόδους των υπόλοιπων κυκλωμάτων ώστε να εξαχθεί ένα αποτέλεσμα.

Έπειτα από πειραματισμό για τον κατάλληλο αριθμό των διακοπών που θα έχουμε στην είσοδο καταλήξαμε ότι η χρυσή τομή στο κύκλωμα μας είναι για 8 bits. Πρακτικά, δηλαδή όπως φαίνεται και στο σχήμα 3.6.2α' παίρνουμε κάθε bit εισόδου και το δίνουμε αντίστοιχα σε έναν ελεγχόμενο καθρέφτη ρεύματος (και στις 2 εισόδους του συμπληρωματικά όπως είδαμε παραπάνω). Αυτοί οι διακόπτες έχουν μια αναλογία στα μεγέθη τους έτσι ώστε να μπορούν να μας δώσουν στην έξοδο ανάλογα με την είσοδο τους ένα ρεύμα που θα μπορέσει να μας δημιουργήσει μια κάπως κβαντισμένη ράμπα. Δηλαδή το lsb θα ελέγξει τον ελεγχόμενο καθρέφτη (σε υψηλό επίπεδο ως σχεφτόμαστε έναν διακόπτη) του κυκλώματος που έχει τις μικρότερες διαστάσεις ώστε να παράξει και το μικρότε-



(α') Προτεινόμενη υλοποίηση bump



(β') Σύμβολο του προτεινόμενου bump

ρο δυνατό ρεύμα, όντας το λιγότερο σημαντικό μπιτ. Το αμέσως επόμενο μπιτ θα εισέλθει στον επόμενο ελεγχόμενο καθρέφτη ο οποίος θα έχει διπλασίου μεγέθους τρανζίστορ ώστε να παράξει διπλάσιο ρεύμα, δηλαδή να μπορέσει να έχει στην έξοδο του ένα ρεύμα που όταν θα έρθει λογικό "1" σαν δεύτερο lsb αλλά λογικό "0" σαν lsb να μπορεί να μας δώσει ένα ρεύμα στην έξοδο που θα είναι μεγαλύτερο του πρώτου. Αντίστοιχα υλοποιήθηκαν και οι υπόλοιποι διακόπτες για τα 8 bits. Οι αναλογίες των διακοπών θα φανούν στον πίνακα 3.3. Στο σημείο αυτό, ωστόσο, θα αναρωτηθεί κάποιος γιατί βάλουμε την αναλογία να είναι διπλάσια σε κάθε διακόπτη από τον προηγούμενο και δεν δώσαμε μια μικρότερη αναλογία για παράδειγμα  $1 : \frac{3}{2}$ . Η απάντηση στο ερώτημα αυτό

απαντάτε στο επόμενο κομμάτι του κυκλώματος που θα δεχτεί το ρεύμα εξόδου που έρχεται από τους διακόπτες και όπως βλέπουμε αθροίζεται σε ένα καλώδιο, το κύκλωμα αυτό είναι το WTA bump. Το συγκεκριμένο σημείο θα το αναλύσουμε παρακάτω στην ενότητα αυτή, ωστόσο, πολύ συνοπτικά, η σύγκριση των ρευμάτων που θα δούμε ότι κάνει το WTA bump και που αξιοποιούμε την γραμμική περιοχή που δεν έχουμε ξεκάθαρο αποτέλεσμα, μας δίνει πιο παχιές καμπύλες για μικρότερη αναλογία από 1:2 και ακόμα σε κάποιες περιπτώσεις που τα ρεύματα που θα έρθουν από την "σκάλα" (την ονομάζουμε σε εισαγωγικά σκάλα γιατί είναι μια κβαντισμένη ράμπα λόγω του μικρού μεγέθους που έχουμε σαν είσοδο, θυμίζουμε μόνο 8 μπιτς), είναι αρκετά κοντινές τιμές όπως γίνεται για αναλογίες κάτω από 1:2, έχουμε στην έξοδο μας ένα bump με μεγάλη αστάθεια. Βγάζει, δηλαδή λάθος πιθανότητες αφού μπορεί να έχει κάποια αξιολογικά εύρη που θα κάνει μικρά bumps.

Η έξοδος αυτού του κυκλώματος είδαμε ότι μπαίνει σε ένα WTA bump, σαν ρεύμα εισόδου  $I_1$ , ωστόσο στο WTA bump, θέλουμε και μια ακόμα είσοδο ρεύματος στον δεύτερο νευρώνα για να μπορέσει να κάνει τη σύγκριση. Το ρεύμα αυτό προέρχεται από ένα άλλο κύκλωμα συμμετρικό με εκείνο των ελεγχόμενων διακοπών που είχαμε για την είσοδο και σε αυτό εισέρχεται η μέση τιμή του κυκλώματος. Αυτή η μέση τιμή δίνεται εξωτερικά σε τάσεις high, low όπως ακριβώς και στις εισόδους. Πρακτικά, μέσω αυτής μπορούμε να ελέγξουμε τη μέση τιμή της Γκαουσιανής μας, αφού αυτές οι μέσες τιμές θα καθορίσουν την τιμή του ρεύματος που θα δοθεί σαν  $I_2$  στο WTA bump, και υπενθυμίζουμε ότι πρακτικά το  $I_1$  παίζει το ρόλο της εισόδου ενώ το  $I_2$  θα είναι το ρεύμα που παίζει το ρόλο της μέσης τιμής. Αυτό όπως είχαμε αναφέρει και πρωτύτερα γίνεται λόγω της λειτουργίας των νευρώνων που θα βγάζουν το ρεύμα πόλωσης στην έξοδο του νευρώνα με το μεγαλύτερο ρεύμα, εκτός από μια γραμμική περιοχή λειτουργίας που έχουμε τιμές ρεύματος και από τους δύο νευρώνες. Πρακτικά, αυτό είναι το χαρακτηριστικό που εκμεταλλευόμαστε, καθώς τα ρεύματα της γραμμικής περιοχής είναι και αυτά που θα εκμεταλλευτούμε στο στάδιο του current correlator, αφού δέχεται δύο ρεύματα που το ένα φθίνει ενώ παράλληλα το άλλο αυξάνεται, έχουν ίδιο ρυθμό ανάπτυξης και φτάνουν από ίδιες τιμές minimum σε maximum. Έτσι, ο τύπος 3.1.2 του current correlator μας δίνει σαν αποτέλεσμα μια Γκαουσιανή κατανομή.

Ενδιαφέρον έχει να αναλογιστούμε πως μπορούμε να θέσουμε τις παραμέτρους της Γκαουσιανής σε μια τέτοια εφαρμογή. Όπως αναφέραμε νωρίτερα, η μέση τιμή του κυκλώματος καθορίζεται από το ρεύμα που έρχεται στον δεύτερο νευρώνα του κυκλώματος (και προέρχεται από ίδιους διακόπτες με αυτούς που επεξεργάζονται την είσοδο), επομένως αφού έχουμε και συμμετρικό κύκλωμα η στάθμη του ρεύματος που θα δοθεί στην έξοδο θα είναι ίδια και στις δύο πλευρές. Έχοντας ακόμη μια ακρίβεια της τάξης των 8 bits, και τις αναλογίες



του κυκλώματος με τον τρόπο που είναι παίρνουμε κοντινές τιμές ρεύματος από την είσοδο και την μέση τιμή οπότε καταφέρνει να δημιουργηθεί μια σχετικά αξιόλογη γραμμική περιοχή από το WTA-bump και έπειτα ο current correlator μπορεί να βγάλει μια αξιοσημείωτη Γκαουσιανή στην έξοδο. Πρακτικά ανάλογα με την θέση του ρεύματος που καθορίζει την μέση τιμή, αυτή θα έχει διαφορετική αναλογία με την είσοδο (που στην απλή περίπτωση που θέλουμε να δούμε μια μονοδιάστατη Γκαουσιανή έχουμε απλά μια σκάλα, για μεγαλύτερη ακρίβεια όπως αυτή των 8 bits έχουμε μια χβαντισμένη καμπύλη που είναι εκθετικής φύσεως, όπως μπορεί να γίνει εύκολα αντιληπτό από τις διαστάσεις). Αυτή η διαφορά αναλογίας (διαφορετική τιμή του ρεύματος μέσης τιμής σε σχέση με την είσοδο), καταφέρνει να μετατοπίσει την περιοχή που θα έχουμε συγκρίσιμα ρεύματα. Παραδείγματος χάρη, αν δώσουμε ένα ρεύμα πόλωσης  $1nA$  στους διακόπτες θα έχουμε για μια συγκεκριμένη μέση τιμή τιμή εξόδου  $10nA$  για μια άλλη θα έχουμε  $50nA$ , οπότε το WTA bump θα δημιουργήσει γραμμική περιοχή δεξιάτερα από ότι πριν αφού μετατοπίστηκε η σύγκριση από την είσοδο με την τιμή των  $10nA$  σε αυτήν της εισόδου με τα  $50nA$  που θα παίρνουμε για μεγαλύτερη είσοδο από πριν.

Η παράμετρος του βάρους σε κάθε Γκαουσιανή μπορεί να δοθεί μέσω του ρεύματος πόλωσης του κασκοδικού καθρέφτη ρεύματος που πολώνει το WTA-bump, καθώς αυτό είναι που θα βγει στην έξοδο από τους νευρώνες.

Τέλος, σημαντική είναι η μέθοδος που αξιοποιήσαμε για να μπορούμε να έχουμε διαφορετικές διασπορές στην έξοδο μας και να μπορούμε να έχουμε από πολύ λεπτές γκαουσιανές έως και αρκετά φαρδιές. Η πρώτη μας σκέψη ήταν η αξιοποίηση της διόδου που έχουμε βάλει η οποία καταφέρνει και μας ρίχνει τα ρεύματα εισόδου που έχουν ένα DC offset, όπως και είναι λογικό καθώς το leakage current του ελεγχόμενου καθρέφτη που έχει μεγάλες διαστάσεις όπως αυτό του 8<sup>ου</sup> bit, είναι αρκετά μεγάλο ειδικά αν συνυπολογίσουμε ότι ανθροίζονται όλα τα leakage currents, μπορούμε να πάρουμε πολύ μεγάλο ρεύμα leakage. Οπότε, το κύκλωμα μας δουλεύει με ένα dc offset το οποίο μπορεί και αντισταθμίζει με καλό τρόπο η διόδος που μπαίνει στο σημείο εκείνο. Ωστόσο, αυτές οι διόδοι, σε περίπτωση που τους προσφερθεί κάποια τάση στο bulk μπορούν και μετατοπίζουν με την αξιοποίηση του φαινομένου σώματος, την διασπορά του κυκλώματος. Βέβαια οι διαφορές που βλέπαμε παρότι αξιοσημείωτες, σε πολλές περιπτώσεις δε μας δίνανε την σωστή τιμή της γκαουσιανής που μπορεί να χρειαζόμασταν και έτσι έβγαζε το κύκλωμα αρκετά σφάλματα όταν δοκιμάζαμε την πολυδιάστατη Γκαουσιανή όπως θα δούμε και στο επόμενο κεφάλαιο. Για αυτό το λόγο, παρατηρήσαμε ότι η διασπορά του κυκλώματος μπορεί να αλλάξει και μέσω άλλης μεθόδου, αξιοποιώντας δηλαδή το ρεύμα εισόδου που δίνουμε στους ελεγχόμενους καθρέφτες τόσο για την είσοδο όσο και για την διασπορά (δίνουμε ίδια τιμή ρεύματος και στις δύο), μας αλλάζει η κλίση που έχει η συ-

νάρτηση της εισόδου που παράγεται από τους διακόπτες. Αυξάνοντας δηλαδή το ρεύμα εισόδου παρατηρούμε ότι κάνουμε πιο απότομη την κλίση της εκθετικής σχέσης οπότε η γραμμική περιοχή μειώνεται όσο ανεβαίνουμε σε στάθμες ρεύματος. Παρακάτω στο σχήμα 3.6.3 βλέπουμε την απόκριση του κυκλώματος για διαφορετικές τιμές ρεύματος πόλωσης διακοπτών και για διαφορετικές τάσεις  $V_c$ .

Ας εξηγήσουμε αυτή την περίπτωση με ένα παράδειγμα. Υπενθυμίζω ότι έχουμε 8 bits, συνεπώς έχουμε  $2^8 - 1 = 255$  ενδεχόμενους δεκαδικούς που μπορούν να μπου σαν είσοδοι. Όπως είπαμε και πρωτότερα, το κύκλωμα των διακοπτών (καθρεφτών) μας έχει αναπόφευκτα κάποια τιμή ρεύματος και στην περίπτωση που είναι off κάποιος διακόπτης (πιο συγκεκριμένα μέσω πειραματικών μετρήσεων είδαμε ότι το κύκλωμα έχει ρεύμα της τάξης των  $114\mu A$  σαν leakage για το lsb, αυτό συνεπάγεται ότι για τους μεγαλύτερους διακόπτες αυτό πολλαπλασιάζεται και φτάνει για το msb να είναι στην τάξη των  $14.6nA$  που είναι τιμή ρεύματος συγκρίσιμη με αυτή που θα έδιναν σε ενεργή περιοχή οι διακόπτες που οδηγούν τα bits από 0 έως και 4 μαζί). Αναλογιζόμενοι τις διαστάσεις που έχει το MSB σε σχέση με το LSB και τα ενδιάμεσα στάδια (η αναλογία είναι 128:1) αντιλαμβανόμαστε ότι η περίπτωση που ο διακόπτης MSB και οι αντίστοιχοι μεγαλύτεροι είναι κλειστοί, παρέχουν στην έξοδο του κυκλώματος μεγαλύτερο ρεύμα από ότι οι αντίστοιχοι διακόπτες στις θέσεις lsb όπως είδαμε και από τις μετρήσεις μας. Επομένως, γίνεται εύκολα αντιληπτό ότι για τις τιμές αυτές των τάσεων που θα έχουν κλειστούς τους μεγάλους διακόπτες αλλά ανοιχτούς τους μικρούς, όταν δοθεί κάποιο μικρό ρεύμα πόλωσης τότε στις περιοχές λειτουργίας που έχουν σημαντική αξία οι μικροί διακόπτες δεν θα έχουμε μεγάλη μεταβολή στο ρεύμα εξόδου. Αυτό συνεπάγεται μεγαλύτερες περιοχές όπου τα ρεύματα που έρχονται να συγκριθούν από την μέση τιμή και τις εισόδους μας, θα έχουν παραπλήσιες τιμές και για αυτό μας δίνουν στην έξοδο αρκετά πιο παχιές Γκαουσιανές κατανομές (έχουν μεγαλύτερο  $\sigma$ ). Σε αντίθετη περίπτωση όσο αυξάνουμε την τιμή του ρεύματος που πολώνονται οι διακόπτες τότε οι καταστάσεις που οι διακόπτες lsb είναι ενεργοί δίνουν αξιόλογο ρεύμα στην έξοδο. Για αυτό μπορούν να δοθούν από τους καθρέφτες που διαχειρίζονται την είσοδο μεγαλύτερα ρεύματα με κυριότερο χαρακτηριστικό μεγαλύτερη διαφορά από κατάσταση σε κατάσταση. Οπότε όσο δεν αυξάνεται σε βαθμό που να αλλάξει περιοχή λειτουργίας το κύκλωμα μας έχουμε καταφέρει να μειώσουμε την γραμμική περιοχή που θα βγάλουν οι νευρώνες του WTA-bump, και επομένως θα έχουμε πιο στενή γκαουσιανή (μικρότερο  $\sigma$ ). Προχωρώντας σε μεγαλύτερες τιμές για την μέση τιμή, και ξεπερνώντας ειδικά το σημείο που οι διακόπτες δεν συνεισφέρουν τόσο μεγάλες διαφορές είτε είναι on είτε είναι off. Σε περιπτώσεις με μικρά ρεύματα πόλωσης έχουμε αντίστοιχο θέμα, αλλά σε αρκετά μικρότερο βαθμό αφού η περιοχή που η σύγκριση των νευρώνων

δίνει γραμμική περιοχή είναι σαφώς μικρότερη αλλά υπάρχει. Όσο αυξάνουμε τα ρεύματα έως και ένα κατώφλι έχουμε μείωση αυτής (μείωση  $\sigma$  Γκαουσιανής). Ωστόσο, όταν ξεπεράσουμε ένα κατώφλι ρεύματος παρότι η περιοχή αυτή ολοένα και μειώνεται, οπότε θα περίμενε κανείς να έχουμε ακόμα μικρότερες διασπορές, έχουμε αύξηση του  $\sigma$ , η οποία όμως δεν είναι συμμετρική. Οπότε καταλαβαίνουμε ότι το ρεύμα που μπορούσε να διαχειριστεί το WTA-bump έχει ένα κατώφλι πριν επηρεάσει το κύκλωμα σε βαθμό που δε μπορεί να λειτουργήσει με τον ιδανικό τρόπο που εξηγήσαμε. Πρακτικά, αν αναλογιστούμε τους τύπους που είδαμε για την περιοχή υποκατωφλίου μπορούμε να καταλάβουμε την εκθετική σχέση που έχει το ρεύμα απαγωγού με τις τάσεις του τρανζίστορ ανάλογα και την περιοχή (τρίοδο ή κορεσμό για να συνυπολογίσουμε ή όχι την τάση απαγωγού). Επομένως είναι σαφές ότι οι τάσεις που επηρεάζουν την λειτουργία ενός WTA όπως είδαμε αλλάζουν λογαριθμικά σε συνάρτηση με το ρεύμα που τους δίνεται στους απαγωγούς. Για αυτόν ακριβώς το λόγο οι τιμές των ρευμάτων που δίνονται στους νευρώνες θα έχουν διαφορετικές τιμές  $\delta_i$  που επηρεάζουν τη γραμμική περιοχή ανάλογα με το ρεύμα που τους δίνεται στην είσοδο. Για αυτό και όταν δίνουμε μεγάλο ρεύμα πόλωσης στους ελεγχόμενους καθρέφτες, οπότε παίρνουμε μεγάλα ρεύματα σύγκρισης στο WTA-bump, δημιουργείται ένα σημείο πάνω από το οποίο η αύξηση του ρεύματος μας δίνει τις αντίθετες επιπτώσεις από αυτές που επιδιώκουμε. Λειτουργεί δηλαδή με μια αντίστοιχη αναλογία που λειτουργούν τα μικρά ρεύματα πόλωσης στις μικρές μέσες τιμές. Αντιλαμβανόμαστε, λοιπόν, ότι όσο αυξάνεται η μέση τιμή τόσο λιγότερο ρεύμα χρειάζεται να δοθεί στο κύκλωμα, καθώς και με μικρά ρεύματα πόλωσης το ρεύμα στις περιοχές με μεγάλη μέση τιμή θα έχει μεγάλο ρεύμα εξόδου στους νευρώνες λόγω της αναλογίας στα μεγέθη που συζητήθηκε πιο πάνω. Επίσης, το αρνητικό με τις περιοχές αυτές είναι ότι επειδή ούτως ή άλλως έχουν υψηλή τιμή ρεύματος, η μικρότερη δυνατή γκαουσιανή μπορεί να μην καλύπτει πλήρως τις ανάγκες του ταξινομητή. Εν μέρη, βέβαια κάτι τέτοιο μπορεί να λυθεί εύκολα μειώνοντας το εύρος που δίνονται οι μέσες τιμές από  $[0,255]$  σε μια τάξη που θα είναι πιο ευνοϊκή ανάλογα και την εφαρμογή. Αυτά, βέβαια, θα συζητηθούν και στο επόμενο κεφάλαιο.

Κάποια spikes που μπορεί να διακρίνονται στην Γκαουσιανή της εξόδου, αξίζει να αναφερθεί ότι είναι φραγμένα από την πόλωση του WTA bump, οπότε δεν μπορούν να μας δημιουργήσουν κάποιο πρόβλημα στο ολοκληρωμένο κύκλωμα, ακόμη στην ταξινόμηση δε μας επηρεάζει καθώς είναι αρκετά μεταβατικό φαινόμενο και προέρχεται από την αδυναμία των πυλών να έχουν ακαριαίους χρόνους rise, fall. Οπότε για αυτό το λόγο και εμείς επιλέξαμε να μην παρουσιάσουμε αποτελέσματα από κάποια ιδανική πηγή εισόδου αλλά από μια ρεαλιστική υλοποίηση που μας δίνει αυτά τα μεταβατικά φαινόμενα. Αυτό συμβαίνει ακόμη, εξαιτίας της σύγκρισης που μπορεί να δίνει το WTA bump αφού

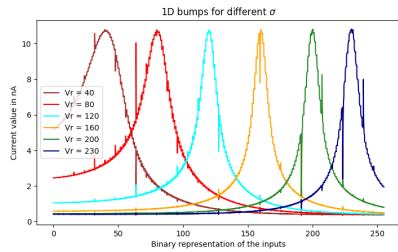
περάσει το επίπεδο της μέσης τιμής η είσοδος. Λογικά αν το σκεφτεί κάποιος μπορεί να το αντιληφθεί εύκολα, καθώς όταν οι είσοδοι γίνονται από high-low, η τιμή του ρεύματος που βγαίνει από την είσοδο μπορεί να πέσει σε συγκρίσιμες τιμές με αυτήν της μέσης τιμής ή ακόμα και μικρότερες ειδικά για τις περιοχές που είμαστε άμεσα κοντά στην μέση τιμή, οπότε να βγάζει έτσι κάποιο spike ρεύματος στην έξοδο.

Πίνακας 3.3: Διαστάσεις τρανζίστορ (Σχήμα 3.5.1).

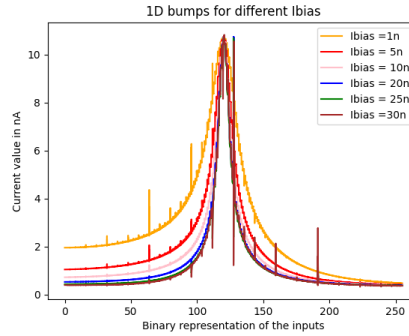
Ελεγχόμενοι διακόπτες είσοδου	W/L ( $\mu\text{m}/\mu\text{m}$ )	Ελεγχόμενοι διακόπτες μέσης τιμής	W/L ( $\mu\text{m}/\mu\text{m}$ )
$M_{p8}-M_{p12}$	0.4/0.5	$M_{p42}-M_{p47}$	0.4/0.5
$M_{p13}-M_{p16}$	0.8/0.5	$M_{p48}-M_{p51}$	0.8/0.5
$M_{p17}-M_{p20}$	1.6/0.5	$M_{p52}-M_{p55}$	1.6/0.5
$M_{p21}-M_{p24}$	3.2/0.5	$M_{p56}-M_{p59}$	3.2/0.5
$M_{p26}-M_{p29}$	6.4/0.5	$M_{p60}-M_{p63}$	6.4/0.5
$M_{p30}-M_{p33}$	12.8/0.5	$M_{p64}-M_{p67}$	12.8/0.5
$M_{p34}-M_{p37}$	25.6/0.5	$M_{p68}-M_{p71}$	25.6/0.5
$M_{p38}-M_{p41}$	51.2/0.5	$M_{p72}-M_{p75}$	51.2/0.5

### 3.6.2 Πολυδιάστατη υλοποίηση του ταξινομητή

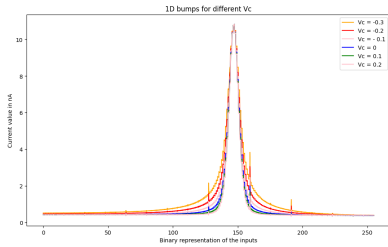
Η πολυδιάστατη μορφή του ταξινομητή μας είναι αρκετά απλή. Το μόνο που χρειάζεται να γίνει ώστε να πάρουμε το γινόμενο που θα θέλαμε μαθηματικά είναι να ορίσουμε την έξοδο του ενός bump σαν είσοδο στο επόμενο. Η έξοδος είναι προφανώς μία για την κάθε διάσταση αλλά οι είσοδοι είναι τρεις (αναφορικά με τα ρεύματα μόνο), σε κάθε κύκλωμα. Για να λειτουργήσει η λογική όμως του πολλαπλασιασμού όμως όπως θα θέλαμε, να πολλαπλασιαστούν δηλαδή οι πιθανότητες που βγάζει η κάθε διάσταση μεταξύ τους και να έχουμε μια συνολική πιθανότητα για κάθε είσοδο, θα θέλαμε να μπαίνει η έξοδος του κάθε bump σαν είσοδος στο ρεύμα που πολώνει το WTA-bump καθώς αυτή είναι η πιθανότητα που επιζητούμε. Οπότε, ανάλογα με τις διαστάσεις που θέλουμε μπορούμε να ορίσουμε με αυτό τον τρόπο μια πολυδιάστατη γκαουσιανή όπου κάθε bump circuit θα δέχεται τις εισόδους του κυκλώματος και την αντίστοιχη μέση τιμή της και διασπορά της κάθε περιοχής χαρακτηριστικού και θα βγάζει έτσι μια πιθανότητα συνυπολογίζοντας την αμέσως προηγούμενη διάσταση. Αυτό μπορούμε να το δούμε και με ένα διάγραμμα όπως 3.6.4.



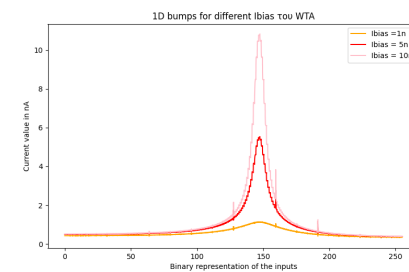
(α') Αποτελέσματα του μονοδιάστατου προτεινόμενου κυκλώματος για διαφορετικές τιμές  $V_r$ , δηλαδή μέσης τιμής, όπως αυτές φαίνονται σε δεκαδική μορφή, με ίδιο ρεύμα bias για όλα τα bumps



(β') Ρεύμα εξόδου προτεινόμενου κυκλώματος με διαφορετικές τιμές  $I_{bias}$  που επηρεάζει άμεσα το  $\sigma$

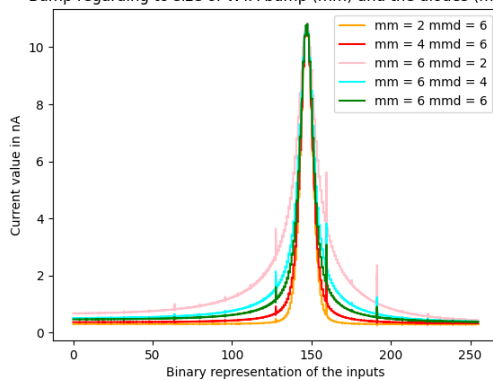


(γ') Αποτελέσματα του μονοδιάστατου προτεινόμενου κυκλώματος για διαφορετικές τιμές  $V_c$ , με  $V_r = 140$  και  $I_{bias} = 10nA$



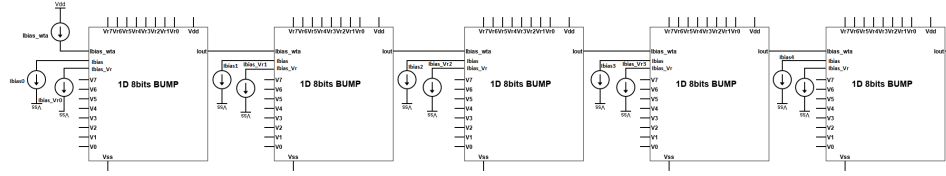
(δ') Αποτελέσματα του μονοδιάστατου προτεινόμενου κυκλώματος για διαφορετικές τιμές ρεύματος πόλωσης του wta bump

Bump regarding to size of WTA-bump (mm) and the diodes (mmd)



(ε') Αποτελέσματα του μονοδιάστατου προτεινόμενου κυκλώματος για διαφορετικές τιμές του μεγέθους των τρανζίστορ των νευρώνων του WTA-bump και των διόδων του. Ως mm και mmd έχουμε τον τύπο  $W = 200n \times mm$  και  $W = 200n \times mmd$

Σχήμα 3.6.3: Πειραματικά δεδομένα από την προτεινόμενη υλοποίηση του κυκλώματος



Σχήμα 3.6.4: Πολυδιάστατη Γκαουσιανή, παράδειγμα 5 διαστάσεων

## Κεφάλαιο 4

# Αξιοποίηση του κυκλώματος σε πραγματικές εφαρμογές

Στο κεφάλαιο αυτό θα συζητηθεί η λειτουργία του κυκλώματος που παρουσιάστηκε παραπάνω σε πραγματικές εφαρμογές. Θα αναλυθεί δηλαδή η λειτουργία του, σε πραγματικά dataset. Οι υλοποιήσεις αυτές μπορεί να διαφέρουν μεταξύ τους σχετικά με την λογική που θα ακολουθήσουμε. Θα φτιάξουμε δηλαδή κάποιους ταξινομητές οι οποίοι ανάλογα το dataset στο οποίο θα εκπαιδευτούν, θα ακολουθούν είτε την λογική του Μπαγιεσιανού μοντέλου είτε τη λογική του Gaussian Mixture Model ανάλογα με τις απαιτήσεις που υπάρχουν. Όταν, έχουμε πιο απλά dataset θα χρησιμοποιούμε την λογική του Bayes οπότε θα έχουμε αρκετά μικρότερη κατανάλωση αφού εξοικονομούμε τουλάχιστον τους μισούς πόρους από αυτούς που θα χρησιμοποιούσαμε. Θα δείξουμε ακόμα τα αποτελέσματα που βγάζει το κύκλωμα καθώς και την συνολική κατανάλωση που χρειάζεται.

### 4.1 Εκπαίδευση του μοντέλου

Για την εκπαίδευση του μοντέλου αξιοποιήσαμε διάφορες βιβλιοθήκες σε γλώσσα python. Αξιοποιήθηκαν προϋπάρχοντες κώδικες που είχαν γραφτεί για την επεξεργασία των δεδομένων που περάστηκαν στο πρόγραμμα προσομοίωσης Cadence IC Suite για κυκλώματα που σχεδιάστηκαν σε τεχνολογία TSMC 90 nm CMOS process, το οποίο και αξιοποιείται καθ'όλη τη διάρκεια της παρούσας διπλωματικής για την υλοποίηση του κυκλώματος. Ακόμη γράφτηκαν διάφοροι κώδικες για να μπορέσει να εκπαιδευτεί το μοντέλο αναφορικά με τις παραμέτρους αλλά και τις κατάλληλες υπερπαραμέτρους. Αφού δημιουργήσαμε τα κατάλληλα αρχεία που περάσαμε στο πρόγραμμα σχεδίασης και τρέξουμε τις

προσομοιώσεις που χρειάστηκε εξάγαμε τα δεδομένα που χρειαζόμασταν για τις αντίστοιχες επεξεργασίες που τους έγιναν (γραφικές, σύγκριση κώδικα και υλοποίησης). Το μοντέλο που χρησιμοποιείται είναι το Γκαουσιανό μοντέλο μίξης και μέσω αυτού μπορούμε να βρούμε τις μέσες τιμές και τις διασπορές που θα αξιοποιήσουμε για να τεστάρουμε τόσο τον ταξινομητή μας σε επίπεδο python όσο και στο chip που σχεδιάσαμε. Ωστόσο, όπως είναι λογικό, για το IC θα χρησιμοποιηθούν εισόδοι οι οποίες θα περάσουν από μια μικρή προεργασία, θα μετατραπούν, δηλαδή, από ακέραιους ή πραγματικούς αριθμούς σε δυαδικούς. Προτιμήθηκε η αξιοποίηση ενός ταξινομητή με 8 bits εισόδους, ωστόσο είναι σαφές ότι με ανάλογη εκπαίδευση θα μπορούσαμε να αξιοποιήσουμε μεγαλύτερους αλλά και μικρότερους ταξινομητές (αναφορικά με το πλήθος εισόδων). Καταλήξαμε στα 8 bits εξαιτίας της μορφής της Γκαουσιανής καμπύλης η οποία παραγόταν από αυτά, αλλά και συναρτήσει της χαμηλότερης κατανάλωσης που μας έδινε συγκριτικά με περισσότερες εισόδους. Οι εισόδοι και οι μέσες τιμές δίνονται στο μοντέλο μας με την μορφή τάσης bitwise ενώ η διασπορά που αποφασίσαμε να ελέγχεται μέσω διαφορετικών τιμών εισόδου ρεύματος  $I_{bias}$  σε συνδυασμό με την αλλαγή του  $V_c$ , είχε μια δυσκολία στην μεταφορά των τιμών στο Cadence. Ωστόσο, καταφέραμε να αντιστοιχίσουμε τις τιμές που χρειαζόμαστε για μια αντίστοιχη διασπορά ώστε να περάσουμε το αντίστοιχο αρχείο που περιείχε τις τιμές ρεύματος που δίνονται σε κάθε bump.

## 4.2 Υλοποίηση ταξινομητή για σκοπούς ανίχνευσης κατηγοριών θυρεοειδή

Ο θυρεοειδής είναι ένα μικρός αδένας σε σχήμα πεταλούδας που βρίσκεται μπροστά από την τραχεία. Η λειτουργικότητα του συμβαδίζει με την αρμονική λειτουργία του μεταβολισμού, εκκρίνει δηλαδή ορμόνες οι οποίες είναι υπαίτιες για τον έλεγχο του. Η επιτήρηση της εύρυθμης λειτουργίας του αναφορικά με την ποσότητα ορμονών που παράγει γίνεται από την υπόφυση που βρίσκεται στη βάση του κρανίου κάτω από τον εγκέφαλο και ελέγχει τις ορμόνες μέσω της ροής του αίματος. Η παραγωγή περισσότερης ή λιγότερης ποσότητας από αυτή την ορμόνη μπορεί να προκαλέσει έντονες διαταραχές στον άνθρωπο και πιθανότατα σοβαρά προβλήματα που μπορεί να χρειαστούν αγωγή. Ένας υπερδραστικός θυρεοειδής μπορεί να επηρεάσει οποιονδήποτε αλλά κυρίως γυναίκες παρά άντρες, τυπικά παρατηρείται να υποφέρουν από αυτόν άτομα ηλικιακής ομάδας μεταξύ 20 και 40 χρονών. Γενικότερα υπάρχουν πολλές ασθένειες σχετικές με τον θυρεοειδή αλλά κυρίως εμφανίζονται δύο τύποι που μπορούμε να τους χωρίσουμε ανάλογα με την παραγωγή των θυρεοειδικών ορμονών. Η παραπάνω παραγωγή ορμονών οδηγεί στον υπερθυρεοειδισμό ενώ η λιγότερη



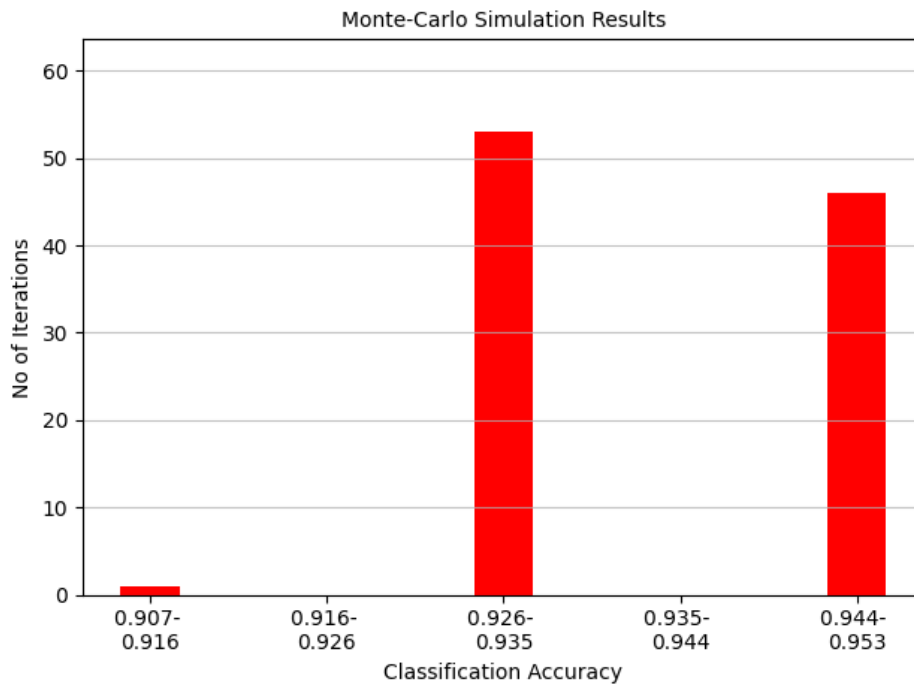
#### 4.2. ΥΛΟΠΟΙΗΣΗ ΤΑΞΙΝΟΜΗΤΗ ΓΙΑ ΣΚΟΠΟΥΣ ΑΝΙΧΝΕΥΣΗΣ ΚΑΤΗΓΟΡΙΩΝ ΘΥΡΕΟΕΙΔΗ65

στον υποθυρεοειδισμό. Τα συμπτώματα της πρώτης μορφής είναι κυρίως άγχος, ευερεθιστότητα ή διάθεση, υπερκινητικότητα, εφίδρωση ή ευαισθησία σε υψηλές θερμοκρασίες, τρέμουλο χεριών, απώλεια μαλλιών ή ελαφριά εμμηνόρροια. Στην άλλη πλευρά, για την δεύτερη αντιμετωπίζουμε κούραση, προβλήματα ύπνου, κατάθλιψη, ευαισθησία σε χαμηλές θερμοκρασίες, ξηροδερμία και τριχόπτωση, βραχνάδα, δυσκοιλιότητα, αύξηση βάρους, υψηλή χοληστερόλη, βραδυκαρδία, μυαλγίες, δυσκολία συγκέντρωσης και βαριές περιόδους. Ανάλογα την μορφή της νόσου η αγωγή μπορεί να περιλαμβάνει φάρμακα, ραδιενεργό ιώδιο ή ακόμη και χειρουργική επέμβαση [60, 61].

Για τον έλεγχο του ταξινομητή μας αναφορικά με την ακρίβεια του, αρχικά δοκιμάσαμε ένα σχετικά απλό dataset σχετικό με την διάγνωση θυρεοειδή. Αυτό περιείχε 5 δεδομένα τα οποία εξήγαγαν αποτελέσματα σχετικά με το αν ο ασθενής από τον οποίο έχουν παρθεί οι μετρήσεις αυτές δεν έχει πρόβλημα, αν έχει υπερθυρεοειδισμό ή υποθυρεοειδισμό. Οι μετρήσεις οι οποίες αποτελούν τα χαρακτηριστικά που θα εκπαιδεύσουν αλλά και τεστάρουν το μοντέλο μας πάρθηκαν από το UCI [62]. Το συγκεκριμένο σετ δεδομένων ήταν σχετικά απλοϊκό, όπως και φαίνεται αν αναλογιστούμε το γεγονός ότι έχουμε 5 χαρακτηριστικά ανά είσοδο. Ακόμη ένας παράγοντας διευκόλυνσης ήταν ότι δεν χρειάστηκε κάποιο feature extraction πάνω στο dataset αφού το ίδιο αποτελούσε τα χαρακτηριστικά που χρειάζονταν. Για το συγκεκριμένο μοντέλο είδαμε ότι είχαμε αρκετά ικανοποιητικά αποτελέσματα αξιοποιώντας τη Μπαγιεσιανή λογική, έχοντας δηλαδή μόνο ένα cluster ανά κλάση. Έτσι, είχαμε αρκετά καλή κατανάλωση και συγκρίσιμη ακρίβεια με το αντίστοιχο μοντέλο που τεστάρουμε σε python. Για να αξιοποιήσουμε τα δεδομένα με τον ταξινομητή μας, ο οποίος είναι ψηφιακής εισόδου και όχι αναλογικής ενώ το dataset έχει τιμές σε δεκαδικούς αριθμούς αξιοποιήσαμε μια μετατροπή με τη βοήθεια της python από δεκαδικό σύστημα αρίθμησης σε δυαδικό. Φτιάχτηκε, με αυτόν τον τρόπο ένα csv το οποίο μπορούσαμε να εισάγουμε στο Cadence.

Για την εξακρίβωση της ορθότητας των αποτελεσμάτων που εξάγαμε τρέξαμε 20 επαναλήψεις τόσο το μοντέλο του software όσο και την προσομοίωση σε hardware. Πήραμε, δηλαδή, το dataset και το κόψαμε σε 20 διαφορετικά train και test sets, τα οποία αξιοποιήσαμε για την εκπαίδευση και επαλήθευση των μοντέλων μας (hardware and software). Τα αποτελέσματα που παράχθηκαν από τις προσομοιώσεις αυτές φαίνονται στον πίνακα 4.1. Παρατηρούμε ότι, η βέλτιστη, η χειριστη και η μέση τιμή, των περιπτώσεων που εξετάστηκαν, του software based μοντέλου είναι ελάχιστα καλύτερες από του hardware based. Ωστόσο, αυτό είναι αναμενόμενο και αρκετά ενδιαφέρον αν αναλογιστεί κανείς και την διαφορά σε κατανάλωση που έχει ένας υπολογιστής με το κύκλωμα που παρουσιάζουμε. Η κατανάλωση του ταξινομητή είναι για τάση τροφοδοσίας 0.6V, 18.28μW.

Ακόμη, εκτός από την προσομοίωση στο Cadence IC Suite, για να μπορέσουμε να επαληθεύσουμε και την ευστάθεια του ταξινομητή μας, τρέξαμε μια Monte Carlo analysis πάνω σε ένα από τα 20 iterations, οπότε τα αποτελέσματα που θα παρουσιάσουμε παρακάτω έχουν εξαχθεί από την ανάλυση Monte Carlo, σχήμα 4.2.1. Η χαμηλότερη τιμή που μας δίνει η ανάλυση βλέπουμε ότι είναι 90.698%, η υψηλότερη 95.349%, η μέση τιμή είναι στο 94.07%, ενώ η τυπική απόκλιση είναι της τάξης του 1.203%.



Σχήμα 4.2.1: Αποτελέσματα από την ανάλυση Monte Carlo για την ανίχνευση θυρεοειδή

Πίνακας 4.1: Accuracy

Method	Best	Worst	Mean Value	Standard Deviation
Software	0.9535	0.924	0.9436	0.0102
Hardware	0.9435	0.9069	0.932	0.0137

### 4.3 Υλοποίηση ταξινομητή για σκοπούς ανίχνευσης ειδικών προβλημάτων κινητήρα αυτοκινήτου (CWRU)

Στο σημείο αυτό θέλαμε να δοκιμάσουμε ένα σύνολο δεδομένων με αρκετά χαρακτηριστικά και ενδιαφέρον για να δούμε αν μπορεί να ανταποκριθεί ο ταξινομητής μας. Για αυτό αποφασίσαμε να δοκιμάσουμε το σύνολο δεδομένων του Case Western Reserve University (CWRU) [63], το οποίο έχει δεδομένα από ένα επιταχυνσιόμετρο τοποθετημένο στο κινητήρα. Τα δεδομένα αυτά μας δίνουν τη δυνατότητα να μπορέσουμε να αντιληφθούμε εάν ο κινητήρας λειτουργεί αβλαβής ή έχει 3 συγκεκριμένες περιπτώσεις βλάβης, δηλαδή αν έχει βλάβη σε τρία διαφορετικά σημεία. Έτσι, μπορούμε και κατηγοριοποιούμε τον ταξινομητή μας σε 4 διαφορετικές κλάσεις, η κάθε μία από τις οποίες μας δηλώνει κατά πόσο ο κινητήρας είναι λειτουργικός ή έχει κάποια από τις 3 περιπτώσεις βλάβης που αντιλαμβάνεται το επιταχυνσιόμετρο. Αυτή η εφαρμογή έχει ιδιαίτερο ενδιαφέρον αν αναλογιστούμε το εύρος της χρήσης της στην αυτοκινητοβιομηχανία. Κάθε κλάση περιέχει δεδομένα από καταχωρήσεις δεδομένων 20 δευτερολέπτων, οι οποίες χωρίζονται σε 20 ίσα τμήματα δημιουργώντας δεδομένα 1 δευτερολέπτου που θα επιβληθούν σε περαιτέρω επεξεργασία για την παραγωγή των 9 χαρακτηριστικών που βλέπουμε στον πίνακα 4.4. Έτσι, καταλήγουμε σε ένα 9-διάστατο ταξινομητή που δέχεται ψηφιακή είσοδο και βγάζει ψηφιακή έξοδο. Τα δεδομένα που θα εξαχθούν όπως αναφέραμε είναι προφανώς μη ακέραιας φύσης και συνεπώς υπάγονται σε μια κανονικοποίηση ώστε να έρθουν στην μορφή των 8 bits που θα χρειαστούμε για τον ταξινομητή μας, αφού έρθουν στο κατάλληλο εύρος, λοιπόν, θα μετασχηματιστούν σε δυαδικές εισόδους όπως αναφέραμε και πρωτύτερα. Το συγκεκριμένο dataset, έχει αρκετά μεγάλο και σύνθετο σύνολο δεδομένων, καθώς οι τιμές σε αρκετά χαρακτηριστικά δεν ήταν ομοιόμορφα ή κάπως ομοιόμορφα κανονικοποιημένες, ακόμα υπήρξαν αρκετοί outliers και δεν μπορούσαμε να έχουμε την βέλτιστη δυνατή κανονικοποίηση. Αφού καταφέραμε να κάνουμε μια κατάλληλη κανονικοποίηση χρειάστηκε να δούμε την καλύτερη μορφή που θα γινόταν το clustering. Παρατηρήσαμε ότι με λιγότερο από τέσσερα clusters δεν είχαμε αρκετά ικανοποιητικά αποτελέσματα και για αυτό το λόγο σπάσαμε τα δεδομένα μας ώστε κάθε κλάση να έχει τέσσερα clusters. Επομένως, αντιλαμβανόμαστε ότι για την συγκεκριμένη εφαρμογή αξιοποιήσαμε το Γκαουσιανό μοντέλο μίξης και όχι ένα Μπαγιεσιανό όπως πρωτύτερα.

Για την εξακρίβωση της ορθότητας των αποτελεσμάτων που εξάγαμε τρέξαμε, όπως και πριν 20 επαναλήψεις τόσο το μοντέλο του software όσο και την προσομοίωση σε hardware. Τα αποτελέσματα που παράχθηκαν από τις προσομοιώσεις αυτές φαίνονται στον πίνακα 4.2. Παρατηρούμε ότι, όπως και πριν, η

βέλτιστη, η χειρίστη και η μέση τιμή, των περιπτώσεων που εξετάστηκαν, του software based μοντέλου είναι ελάχιστα καλύτερες από του hardware based. Η διαφορά με την προηγούμενη εφαρμογή, ωστόσο, είναι ότι στη συγκεκριμένη περίπτωση η μέση τιμή του analog ταξινομητή είναι κοντά στην χειρίστη του software. Η διαφορά αυτή, συνολικά, δεν υποβαθμίζει την συγκεκριμένη υλοποίηση καθώς σε κάθε περίπτωση οι αποκλίσεις είναι μικρές και η ελάχιστη μικρότερη ακρίβεια του αναλογικού ταξινομητή προέρχεται από, πιθανόν, μη βελτιστοποιημένο tuning και ιδιαίτερος στον έλεγχο της διασποράς μέσω των ρευμάτων που τονίσαμε και σε προηγούμενο κεφάλαιο. Αυτό επηρεάζει και την ανάλυση Monte Carlo όπως θα δούμε παρακάτω.

Τρέχοντας την ανάλυση Monte Carlo βλέπουμε ότι έχουμε μια μέση ακρίβεια της τάξης του 90.029 % ενώ παρατηρούμε ότι ο ταξινομητής μας βγάζει σε κάποιες περιπτώσεις πολύ μικρή ακρίβεια η μικρότερη των οποίων είναι 75 % και αρκετά αποτελέσματα στο φάσμα του 91.6-95.7 %. Ακόμα, αξίζει να αναφερθεί ότι λόγω του μεγάλου χρόνου προσομοίωσης δεν υπήρχε η δυνατότητα για Monte Carlo 100 σημείων όπως έγινε στις υπόλοιπες εφαρμογές αλλά για να υπάρχει πληρότητα αναφορικά με την λειτουργικότητα του ταξινομητή, τρέξαμε μια προσομοίωση Monte Carlo 30 σημείων. Αναφορικά με την κατανάλωση του ο ταξινομητής χρειάζεται 117.2  $\mu W$  για τους ανωτέρω υπολογισμούς.

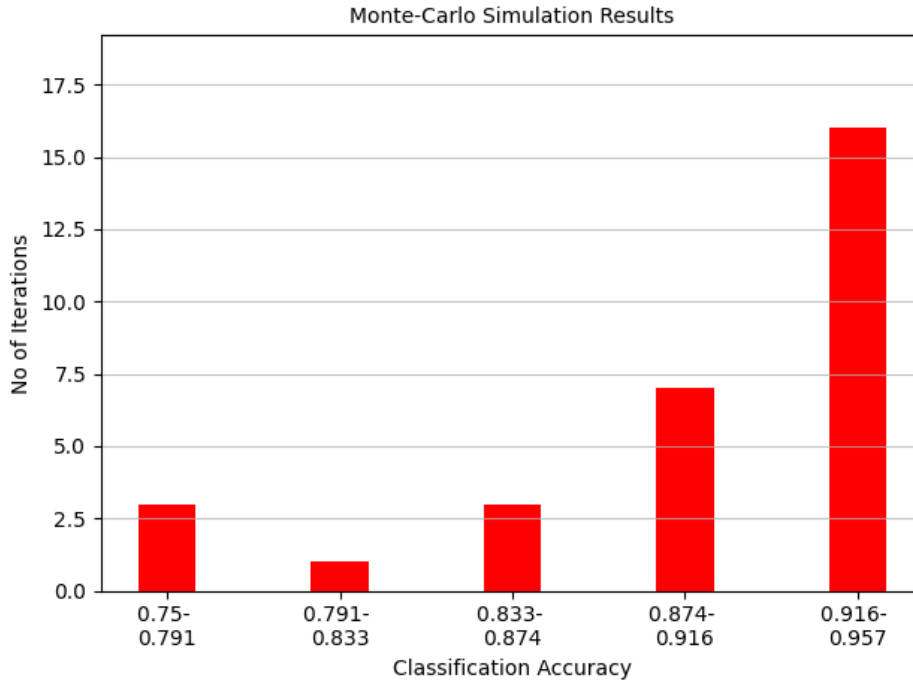
Πίνακας 4.2: Accuracy

Method	Best	Worst	Mean Value	Standard Deviation
Software	0.9858	0.924	0.9588	0.0112
Hardware	0.9535	0.8925	0.9277	0.0137

#### 4.4 Υλοποίηση ταξινομητή για σκοπούς ανίχνευσης ομιλίας έναντι θορύβου speech over noise

Το τελευταίο dataset που χρησιμοποιήσαμε ήταν για σκοπούς αναγνώρισης φωνής έναντι θορύβου. Δώσαμε, όπως θα δούμε, ένα dataset που περιείχε φωνή, δηλαδή, μια σειρά βίντεο που λέγανε κάποιες φράσεις. Αυτό είναι το TIMIT dataset [64] το οποίο μας παρείχε τα δεδομένα της φωνής που αξιοποιήσαμε και πρακτικά είναι διάφορες φράσεις που λένε διάφοροι άνθρωποι με διαφορετικές προφορές. Για την ακρίβεια είχε 8 διαφορετικές διαλέκτους και συνολικά 438 άντρες ομιλητές και 192 γυναίκες. Εμείς, για τους σκοπούς της διπλωματικής δεν αξιοποιήσαμε σαφώς όλους τους ομιλητές γιατί μας έδιναν πάρα πολλά δεδο-

#### 4.4. ΥΛΟΠΟΙΗΣΗ ΤΑΞΙΝΟΜΗΤΗ ΓΙΑ ΣΚΟΠΟΥΣ ΑΝΙΧΝΕΥΣΗΣ ΟΜΙΛΙΑΣ ΕΝΑΝΤΙ ΘΟΡΥΒΟΥ SPEE



Σχήμα 4.3.1: Αποτελέσματα από την ανάλυση Monte Carlo για το CWRU dataset

μένα για τα οποία δεν μπορούσαμε να βγάλουμε συμπεράσματα σε πραγματικό χρόνο στο πρόγραμμα προσομοίωσης, για αυτούς τους σκοπούς θα δίνουμε έναν καθαρά ψηφιακό τρόπο σχεδίασης όπως και είναι λογικό. Ακόμη, για να έχουμε αληθοφανείς αποτελέσματα χρησιμοποιήσαμε και ένα dataset θορύβου το MUSAN [65]. Από αυτά τα δύο datasets πήραμε 10 τυχαία δείγματα θορύβου στα οποία κάναμε feature extraction και 10 δείγματα βίντεο που κάναμε και από αυτά την ίδια διαδικασία. Αξίζει να σημειωθεί ότι τα βίντεο που είχαν τα δείγματα φωνής, τα εμπλουτίσαμε με θόρυβο ο οποίος ήταν της τάξης του 20%, αφού επήλθε κανονικοποίηση και στο dataset του ήχου και σε αυτό του θορύβου. Έτσι, τα δεδομένα μας μπορούμε να δηλώσουμε με βεβαιότητα ότι ανταποκρίνονται σε ένα σύνολο πραγματικών δεδομένων που θα μπορούσαν να έρχονται άμεσα από κάποιο μικρόφωνο. Τα χαρακτηριστικά που χρησιμοποιήσαμε για την εξαγωγή των δεδομένων ήταν κάποια από αυτά που χρησιμοποιήσαμε στο CWRU dataset όπως φαίνονται στον πίνακα 4.4. Χρησιμοποιήθηκαν, δηλαδή, μόνο τα SRA, KV, SV, PPV, CF, IF, MF χαρακτηριστικά του dataset αφού μέσω

#### 70 ΚΕΦΑΛΑΙΟ 4. ΑΞΙΟΠΟΙΗΣΗ ΤΟΥ ΚΥΚΛΩΜΑΤΟΣ ΣΕ ΠΡΑΓΜΑΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ

της χρήσης τους μπορούσαμε να εξαγάγουμε αποτελέσματα χωρίς να έχουμε κάποιο βαρύ τίμημα στην ακρίβεια αλλά μειώσαμε την κατανάλωση αφαιρώντας χαρακτηριστικά που θα δίνονταν ως παραπάνω διαστάσεις στο bump

Για την εκπαίδευση του μοντέλου μας, χρησιμοποιήθηκε μια μέθοδος που μπορεί να θεωρηθεί ότι αξιολογείται για παράδειγμα, σε κάποιο μικρόφωνο που χρησιμοποιείται από διάφορους ομιλητές. Κάναμε, δηλαδή, ένα feature extraction από διάφορα βίντεο που πήραμε δειγματικά για 10 τυχαία δείγματα φωνής από το dataset και αντίστοιχα τυχαία δέκα δείγματα θορύβου, τα οποία στη συνέχεια υποβάλαμε σε εκπαίδευση και εξακρίβωση. Με αυτό τον τρόπο μπόρεσαμε να αποδείξουμε την καλή λειτουργία του κύκλωματος για διαφορετικά άτομα με τα χαρακτηριστικά που είχε ο καθένας τους. Έτσι φαίνεται εκ του αποτελέσματος ότι το κύκλωμα λειτουργεί σε κάθε περίπτωση και δεν μας δίνει ορθά αποτελέσματα μόνο στην περίπτωση που εκπαιδεύεται για ένα άτομο συγκεκριμένα.

Για την εξαγωγή των χαρακτηριστικών του συγκεκριμένου dataset, κόψαμε κάθε βίντεο σε παράθυρα από τα οποία βρήκαμε κορυφές και κοιλίες που είχαν τα σήματα φωνής. Εν συνεχεία, πήραμε μέγιστα, ελάχιστα και μέσες τιμές από κάθε παράθυρο ώστε να κόψουμε τις παύσεις που είχαν μέσα τα βίντεο και να ξεχωρίσουμε με αυτό τον τρόπο τον ήχο. Αφού, είχαμε τα κομμάτια που περιείχαν τον ήχο αποκλειστικά, καταφέραμε να εξάγουμε τα χαρακτηριστικά που αναφέρθηκαν ανωτέρω. Αντίστοιχη διαδικασία ακολουθήσαμε για να βγάλουμε συμπεράσματα χαρακτηριστικών για τον θόρυβο, κόψαμε δηλαδή τα βίντεο ώστε να ξεχωρίσουμε θόρυβο από παύσεις και μετά κάναμε το ίδιο feature extraction που αξιοποιήσαμε για την φωνή. Έτσι, δημιουργήσαμε τα dataset φωνής και θορύβου για να μπορέσουμε να εκπαιδεύσουμε αλλά και να εξακριβώσουμε το μοντέλο μας.

Στην συγκεκριμένη υλοποίηση αξιοποιήσαμε ένα Μπαγιασιανό μοντέλο ταξινομητή. Δεν επεκταθήκαμε σε GMM καθώς παρατηρήσαμε ότι δεν υπήρχε μεγάλη βελτίωση αναφορικά με την αύξηση της ακρίβειας αναλογικά με την αύξηση των clusters. Συνεπώς, δεν άξιζε το trade off της κατανάλωσης όσο και του χώρου on-chip για να δικαιολογήσουμε τέτοια ακρίβεια στον ταξινομητή. Ακόμη, η υλοποίηση μας είναι τέτοια που το κύκλωμα καταναλώνει πολύ λιγότερη ενέργεια από ότι θα έκανε κάποιο ανάλογο του και έτσι μπορούμε να ρίξουμε το κόστος του ταξινομητή ώστε να επιτύχουμε μια αρκετά καλή ακρίβεια ενώ συνάμα έχουμε πολύ καλή κατανάλωση. Ακόμη, στο συγκεκριμένο μοντέλο καταφέραμε να ρίξουμε την τάση τροφοδοσίας στα 0.5V αντί για 0.6V που χρησιμοποιούταν στην βιβλιογραφία και στις προηγούμενες εφαρμογές που είδαμε. Με αυτό τον τρόπο είχαμε συνολική κατανάλωση του ταξινομητή μας ίση με  $15.53\mu W$ , ενώ όταν τροφοδοτούσαμε το κύκλωμα μας με 0.6V είχαμε κατανάλωση ίση με  $26.07\mu W$ . Ρίχνοντας την κατανάλωση ρίξαμε και σε

#### 4.4. ΥΛΟΠΟΙΗΣΗ ΤΑΞΙΝΟΜΗΤΗ ΓΙΑ ΣΚΟΠΟΥΣ ΑΝΙΧΝΕΥΣΗΣ ΟΜΙΛΙΑΣ ΕΝΑΝΤΙ ΘΟΡΥΒΟΥ SPEECH

ένα μικρό ποσοστό την ακρίβεια ωστόσο αν αναλογιστούμε ότι η κατανάλωση μειώθηκε σχεδόν στην μισή θεωρήσαμε ότι άξιζε το trade off. Σημειώνουμε ότι το κυκλώμα των ελεγχόμενων καθρεπτών ρεύματος (διακοπών), μας δίνει κατανάλωση της τάξης των  $1.4\mu W$ , ενώ με την μείωση της τροφοδοσίας τον ρίξαμε στα  $0.93\mu W$  καταφέροντας έτσι να έχουμε έναν από τους πιο χαμηλής κατανάλωσης Digital to Analog Converters στην βιβλιογραφία, όπως προαναφέραμε.

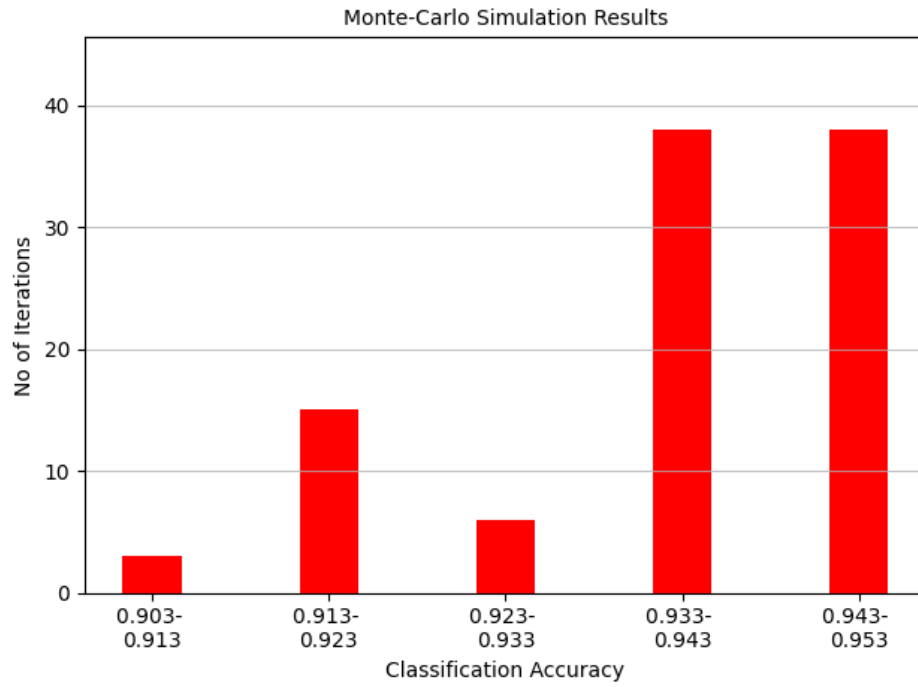
Ακολουθώντας το μοτίβο που έχουμε θέσει καθ' όλη τη διάρκεια της διπλωματικής, αποδεικνύουμε την λειτουργικότητα του αναλογικού μας ταξινομητή μέσω προσομοίωσης στο Cadence IC Suite και ακόμη Monte Carlo analysis για να έχουμε μια ακριβή εικόνα του ταξινομητή. Για την ακρίβεια μέσω της εκπαίδευσης του μοντέλου μας σε 20 επαναλήψεις βλέπουμε ότι έχουμε αρκετά κοντινές τιμές με το μοντέλο του software, τόσο στην βέλτιστη, στην ελάχιστη όσο και στην μέση τιμή των αποτελεσμάτων. Τα αποτελέσματα αυτά φαίνονται και στον πίνακα 4.3.

Ακόμη, μέσω της ανάλυσης Monte Carlo βγάλαμε σαν συμπεράσματα ότι η μέγιστη ακρίβεια που επιτυγχάνεται είναι 95.333%, η μικρότερη είναι 90.333% και η μέση τιμή βγαίνει 93.63% με τυπική απόκλιση 1.15%. Μειώνοντας την τάση τροφοδοσίας του αναλογικού ταξινομητή μας πρακτικά καταφέραμε να διατηρήσουμε τον ίδιο αριθμό σωστών προβλέψεων αναφορικά με το θόρυβο αλλά αυξήσαμε λίγο τον αριθμό των λάθους προβλέψεων στην φωνή. Παρατηρούμε, λοιπόν, και στη γραφική 4.4.1 ότι ο αναλογικός ταξινομητής λειτουργεί εξίσου καλά με τον ταξινομητή που είχαμε σε software, τόσο στην απλή προσομοίωση όσο και στην βαθύτερη που ελέγχουμε τα PVT variations και επαληθεύουμε έτσι την ευστάθεια του κυκλώματος.

Πίνακας 4.3: Accuracy

Method	Best	Worst	Mean Value	Standard Deviation
Software	0.98	0.9267	0.9588	0.0128
Hardware	0.96	0.9133	0.9397	0.0148

72ΚΕΦΑΛΑΙΟ 4. ΑΞΙΟΠΟΙΗΣΗ ΤΟΥ ΚΥΚΛΩΜΑΤΟΣ ΣΕ ΠΡΑΓΜΑΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ



Σχήμα 4.4.1: Αποτελέσματα από την ανάλυση Monte Carlo για το speech over noise dataset



4.4. ΥΛΟΠΟΙΗΣΗ ΤΑΞΙΝΟΜΗΤΗ ΓΙΑ ΣΚΟΠΟΥΣ ΑΝΙΧΝΕΥΣΗΣ ΟΜΙΛΙΑΣ ΕΝΑΝΤΙ ΘΟΡΥΒΟΥ SPEECH

Πίνακας 4.4: Extracted features [5]

Statistic	Equation	Statistic	Equation
Root mean square	$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$	Crest factor	$CF = \frac{\max(x_i)}{RMS}$
Square root of amplitude	$SRA = \left( \frac{1}{N} \sum_{i=1}^N \sqrt{ x_i } \right)^2$	Impulse factor	$IF = \frac{N \cdot \max(x_i)}{\sum_{i=1}^N  x_i }$
Kurtosis value	$KV = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \mu_x}{\sigma_x} \right)^4$	Margin factor	$MF = \frac{\max(x_i)}{SRA}$
Skewness value	$SV = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \mu_x}{\sigma_x} \right)^3$	Peak-to-peak value	$PPV = \max(x_i) - \min(x_i)$
Shape factor	$SF = \frac{\max(x_i)}{SV}$	Kurtosis factor	$KF = \frac{KV}{RMS^4}$

Το RMS δεν το αξιοποιούμε απλά το δείχνουμε στον πίνακα για πληρότητα καθώς αξιοποιείται από άλλα χαρακτηριστικά



## Κεφάλαιο 5

# Συμπεράσματα και Μελλοντική Δουλειά

Στην παρούσα διπλωματική εργασία παρουσιάστηκε μια πλήρως αναλογική υλοποίηση ενός ταξινομητή βασισμένου στο GMM. Ο ταξινομητής αυτός, είχε ενσωματωμένο έναν DAC και μπορούσε να δέχεται μια ψηφιακή είσοδο ώστε να πραγματοποιήσει μια αναλογική επεξεργασία και να εξάγει ένα ψηφιακό αποτέλεσμα αφού επέλεγε νικητή κάθε φορά μέσω ενός WTA κυκλώματος. Η μέθοδος αυτή καταφέρνει να μας δώσει ένα αρκετά χαμηλής ισχύος κύκλωμα. Αναλύθηκαν διάφορες αρχιτεκτονικές και παρουσιάστηκαν παραλλαγές τους που και αξιοποιήθηκαν για την προτεινόμενη αρχιτεκτονική. Συγκεκριμένα, εστιάσαμε στα εξής κυκλώματα: το WTA, το bump του Delbruck το οποίο και δείξαμε μετέπειτα σε μια μορφή που περιπλέκεται με το WTA και δημιουργεί το WTA-bump, τον ελεγχόμενο καθρέφτη ρεύματος που αξιοποιήσαμε σαν DAC και τέλος την προτεινόμενη υλοποίηση. Η τελευταία μπορεί να καταφέρει αυτόνομη ταξινόμηση με εντελώς αναλογικό τρόπο, εφόσον προηγουμένως έχει γίνει η εκπαίδευση του μοντέλου σε ένα εξωτερικό σύστημα. Τα πειραματικά αποτελέσματα αποδεικνύουν τη σωστή λειτουργία του σχεδιασμένου ταξινομητή και ως προς την ακρίβεια της ταξινόμησης (συγκριτικά με μια παραδοσιακή υλοποίηση λογισμικού) και ως προς την ευαισθησία του κυκλώματος.

Μια μελλοντική δουλειά που θα μπορούσε να συσχετιστεί με την συγκεκριμένη διπλωματική είναι η κατασκευή ενός πραγματικού chip. Για να γίνει αυτό χρειάζεται ακόμα να σχεδιαστούν αναλογικά και χαμηλής ισχύος κυκλώματα για τη διασύνδεση μνήμης με την προτεινόμενη αρχιτεκτονική, με στόχο την αποθήκευση τιμών παραμέτρων. Ακόμη, θα μπορούσε να γίνει ένα καλύτερο tuning στον ταξινομητή ώστε να μπορέσει να δώσει μεγαλύτερης ακρίβειας αποτελέσματα. Εναλλακτικά, εκτός από την υλοποίηση του GMM, τα βασικά

δομικά στοιχεία αυτής της αρχιτεκτονικής θα μπορούσαν να τροποποιηθούν ώστε να χρησιμοποιηθούν σε άλλες υλοποιήσεις υλικού μηχανικής εκμάθησης, όπως SVM, K-means και άλλα.

# Bibliography

- [1] A. S. Sedra, K. C. Smith, T. C. Carusone, and V. Gaudet, *Microelectronic circuits*, vol. 4. Oxford university press New York, 2004.
- [2] B. Razavi, *Design of analog CMOS integrated circuits / Behzad Razavi*. New York, NY: McGraw-Hill Education, second edition. ed., 2017 - 2017.
- [3] T. Delbruck and C. Mead, “Bump circuits,” in *Proceedings of International Joint Conference on Neural Networks*, vol. 1, pp. 475–479, 1993.
- [4] V. Alimisis, G. Gennis, K. Touloupas, C. Dimas, M. Gourdouparis, and P. P. Sotiriadis, “Gaussian mixture model classifier analog integrated low-power implementation with applications in fault management detection,” *Microelectronics Journal*, vol. 126, p. 105510, 2022.
- [5] B. Panić, J. Klemenc, and M. Nagode, “Gaussian mixture model based classification revisited: Application to the bearing fault classification,” *Strojnikski Vestnik/Journal of Mechanical Engineering*, vol. 66, no. 4, 2020.
- [6] V. Sze, Y.-H. Chen, J. Emer, A. Suleiman, and Z. Zhang, “Hardware for machine learning: Challenges and opportunities,” in *2017 IEEE Custom Integrated Circuits Conference (CICC)*, pp. 1–8, IEEE, 2017.
- [7] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*, vol. 4. Springer, 2006.
- [8] D. Barber, *Bayesian reasoning and machine learning*. Cambridge University Press, 2012.
- [9] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science*, vol. 349, no. 6245, pp. 255–260, 2015.

- [10] N. Sebe, I. Cohen, A. Garg, and T. S. Huang, *Machine learning in computer vision*, vol. 29. Springer Science & Business Media, 2005.
- [11] J. Padmanabhan and M. J. Johnson Premkumar, “Machine learning in automatic speech recognition: A survey,” *IETE Technical Review*, vol. 32, no. 4, pp. 240–251, 2015.
- [12] S. Wang, W. Chaovaitwongse, and R. Babuska, “Machine learning algorithms in bipedal robot control,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 5, pp. 728–743, 2012.
- [13] C. Park, C. C. Took, and J.-K. Seong, “Machine learning in biomedical engineering,” *Biomedical Engineering Letters*, vol. 8, no. 1, pp. 1–3, 2018.
- [14] P. P. Shinde and S. Shah, “A review of machine learning and deep learning applications,” in *2018 Fourth international conference on computing communication control and automation (ICCUBEA)*, pp. 1–6, IEEE, 2018.
- [15] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, vol. 2. MIT press Cambridge, MA, 2006.
- [16] T. Hofmann, B. Schölkopf, and A. J. Smola, “Kernel methods in machine learning,” *The annals of statistics*, vol. 36, no. 3, pp. 1171–1220, 2008.
- [17] E. Strubell, A. Ganesh, and A. McCallum, “Energy and policy considerations for deep learning in nlp,” *arXiv preprint arXiv:1906.02243*, 2019.
- [18] P. Chi, S. Li, C. Xu, T. Zhang, J. Zhao, Y. Liu, Y. Wang, and Y. Xie, “Prime: A novel processing-in-memory architecture for neural network computation in reram-based main memory,” *ACM SIGARCH Computer Architecture News*, vol. 44, no. 3, pp. 27–39, 2016.
- [19] K. Kara, D. Alistarh, G. Alonso, O. Mutlu, and C. Zhang, “Fpga-accelerated dense linear machine learning: A precision-convergence trade-off,” in *2017 IEEE 25th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, pp. 160–167, IEEE, 2017.

- [20] E. Nurvitadhi, J. Sim, D. Sheffield, A. Mishra, S. Krishnan, and D. Marr, “Accelerating recurrent neural networks in analytics servers: Comparison of fpga, cpu, gpu, and asic,” in *2016 26th International Conference on Field Programmable Logic and Applications (FPL)*, pp. 1–4, IEEE, 2016.
- [21] E. Nurvitadhi, D. Sheffield, J. Sim, A. Mishra, G. Venkatesh, and D. Marr, “Accelerating binarized neural networks: Comparison of fpga, cpu, gpu, and asic,” in *2016 International Conference on Field-Programmable Technology (FPT)*, pp. 77–84, IEEE, 2016.
- [22] T. Chen, Z. Du, N. Sun, J. Wang, C. Wu, Y. Chen, and O. Temam, “Diannao: A small-footprint high-throughput accelerator for ubiquitous machine-learning,” *ACM SIGARCH Computer Architecture News*, vol. 42, no. 1, pp. 269–284, 2014.
- [23] M. B. Rabieah and C.-S. Bouganis, “Fpgasvm: A framework for accelerating kernelized support vector machine,” in *Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications*, pp. 68–84, PMLR, 2016.
- [24] C. Farabet, Y. LeCun, K. Kavukcuoglu, E. Culurciello, B. Martini, P. Akselrod, and S. Talay, “Large-scale fpga-based convolutional networks,” *Scaling up machine learning: parallel and distributed approaches*, vol. 13, no. 3, pp. 399–419, 2011.
- [25] C. Ünsalan and B. Tar, *Digital system design with FPGA: Implementation using Verilog and VHDL*. McGraw-Hill Education, 2017.
- [26] J. Oh, S. Lee, and H.-J. Yoo, “1.2-mw online learning mixed-mode intelligent inference engine for low-power real-time object recognition processor,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 5, pp. 921–933, 2012.
- [27] V. Alimisis, G. Gennis, C. Dimas, M. Gourdouparis, and P. P. Sotiriadis, “An ultra low power analog integrated radial basis function classifier for smart iot systems,” *Analog Integrated Circuits and Signal Processing*, pp. 1–12, 2022.
- [28] V. Alimisis, M. Gourdouparis, G. Gennis, C. Dimas, and P. P. Sotiriadis, “Analog gaussian function circuit: Architectures, operating principles and applications,” *Electronics*, vol. 10, no. 20, p. 2530, 2021.

- [29] V. Alimisis, G. Gennis, K. Touloupas, C. Dimas, N. Uzunoglu, and P. P. Sotiriadis, “Nanopower integrated gaussian mixture model classifier for epileptic seizure prediction,” *Bioengineering*, vol. 9, no. 4, p. 160, 2022.
- [30] M. Ahsanullah, B. Kibria, and M. Shakil, “Normal distribution,” in *Normal and Student st Distributions and Their Applications*, pp. 7–50, Springer, 2014.
- [31] A. R. Mohamed, L. Qi, Y. Li, and G. Wang, “A generic nano-watt power fully tunable 1-d gaussian kernel circuit for artificial neural network,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 9, pp. 1529–1533, 2020.
- [32] J. Anderson, J. Platt, and D. B. Kirk, “An analog vlsi chip for radial basis functions,” *Advances in Neural Information Processing Systems*, vol. 5, 1992.
- [33] S. S. Watkins, P. M. Chau, and R. Tawel, “A radial basis function neurocomputer implemented with analog vlsi circuits,” in *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks*, vol. 2, pp. 607–612, IEEE, 1992.
- [34] M. Verleysen, P. Thissen, J.-L. Voz, and J. Madrenas, “An analog processor architecture for a neural network classifier,” *IEEE Micro*, vol. 14, no. 3, pp. 16–28, 1994.
- [35] S. Guo, L. Peters, and H. Surmann, “Design and application of an analog fuzzy logic controller,” *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 4, pp. 429–438, 1996.
- [36] S.-Y. Peng, B. A. Minch, and P. Hasler, “Analog vlsi implementation of support vector machine learning and classification,” in *2008 IEEE International Symposium on Circuits and Systems*, pp. 860–863, IEEE, 2008.
- [37] R. Zhang, N. Uetake, T. Nakada, and Y. Nakashima, “Design of programmable analog calculation unit by implementing support vector regression for approximate computing,” *IEEE Micro*, vol. 38, no. 6, pp. 73–82, 2018.
- [38] H. Bauer, *Probability theory*, vol. 23. Walter de Gruyter, 2011.
- [39] A. Spanos, *Probability theory and statistical inference: Empirical modeling with observational data*. Cambridge University Press, 2019.



- [40] G. Schay, *Introduction to probability with statistical applications*. Birkhäuser, 2016.
- [41] V.-E. Neagoe and V. Chirila-Berbentea, “Improved gaussian mixture model with expectation-maximization for clustering of remote sensing imagery,” in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3063–3065, IEEE, 2016.
- [42] H.-H. Bock, “Clustering methods: a history of k-means algorithms,” *Selected contributions in data analysis and classification*, pp. 161–172, 2007.
- [43] Y. Lu, S. Lu, F. Fotouhi, Y. Deng, and S. J. Brown, “Incremental genetic k-means algorithm and its application in gene expression data analysis,” *BMC bioinformatics*, vol. 5, no. 1, pp. 1–10, 2004.
- [44] R. Anand, S. Veni, and J. Aravinth, “An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method,” in *2016 international conference on recent trends in information technology (ICRTIT)*, pp. 1–6, IEEE, 2016.
- [45] Z. Lv, Y. Hu, H. Zhong, J. Wu, B. Li, and H. Zhao, “Parallel k-means clustering of remote sensing images based on mapreduce,” in *International Conference on Web Information Systems and Mining*, pp. 162–170, Springer, 2010.
- [46] J. Wu, “Cluster analysis and k-means clustering: an introduction,” in *Advances in K-means Clustering*, pp. 1–16, Springer, 2012.
- [47] G. J. McLachlan, S. X. Lee, and S. I. Rathnayake, “Finite mixture models,” *Annual Review of Statistics and Its Application*, vol. 6, no. 1, pp. 355–378, 2019.
- [48] N. Bouguila and W. Fan, *Mixture models and applications*. Springer, 2020.
- [49] P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and design of analog integrated circuits*. John Wiley & Sons, 2009.
- [50] C. Mead, “Analog vlsi and neural systems,” *NASA STI/Recon Technical Report A*, vol. 90, p. 16574, 1989.
- [51] S.-C. Liu, J. Kramer, G. Indiveri, T. Delbrück, and R. Douglas, *Analog VLSI: circuits and principles*. MIT press, 2002.

- [52] N. H. Weste and K. Eshraghian, *Principles of CMOS VLSI design: a systems perspective*. Addison-Wesley Longman Publishing Co., Inc., 1985.
- [53] J. P. Uyemura, *Circuit design for CMOS VLSI*. Springer Science & Business Media, 2012.
- [54] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
- [55] J. Lazzaro, S. Ryckebusch, M. Mahowald, and C. A. Mead, “Winner-take-all networks of  $o(n)$  complexity,” in *Advances in Neural Information Processing Systems* (D. Touretzky, ed.), vol. 1, Morgan-Kaufmann, 1988.
- [56] R. Sehgal and S. Rajput, “A low voltage 8-bit digital-to-analog converter using floating gate mosfets,” *Analog Integrated Circuits and Signal Processing*, vol. 56, no. 3, pp. 199–204, 2008.
- [57] L. Wong, C. Kwok, and G. Rigby, “1-v cmos d/a converter with multi-input floating-gate mosfet,” *Solid-State Circuits, IEEE Journal of*, vol. 34, pp. 1386 – 1390, 11 1999.
- [58] H.-C. Wang, H. Kao, and T. Lee, “An 8-bit 2-v 2-mw 0.25-mm<sup>2</sup> cmos dac,” pp. 102–105, 01 2004.
- [59] K.-T. Lin, Y.-W. Cheng, and K.-T. Tang, “A 0.5 v 1.28-ms/s 4.68-fj/conversion-step sar adc with energy-efficient dac and trilevel switching scheme,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 24, no. 4, pp. 1441–1449, 2016.
- [60] M. P. Vanderpump, “The epidemiology of thyroid disease.,” *British medical bulletin*, vol. 99, no. 1, 2011.
- [61] A. Jabbar, A. Pingitore, S. H. Pearce, A. Zaman, G. Iervasi, and S. Razvi, “Thyroid hormones and cardiovascular disease,” *Nature Reviews Cardiology*, vol. 14, no. 1, pp. 39–55, 2017.
- [62] D. Dua and C. Graff, “UCI machine learning repository,” 2017.
- [63] K. A. Loparo, “Bearing data center,” *Case Western Reserve University*, vol. 338, 2013.

- [64] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, and D. S. Pallett, "Darpa timit acoustic-phonetic continuous speech corpus cd-rom. nist speech disc 1-1.1," *NASA STI/Recon technical report n*, vol. 93, p. 27403, 1993.
- [65] D. Snyder, G. Chen, and D. Povey, "Musan: A music, speech, and noise corpus," *arXiv preprint arXiv:1510.08484*, 2015.