

Memristors

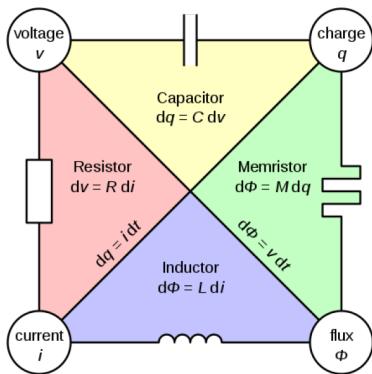
1^η Αβράμη Δήμητρα
Ηλεκτρολόγων Μηχανικών
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Θεσσαλονίκη, Ελλάδα
dimitavram@ece.auth.gr

2^{ος} Στεφανίδης Ιωάννης
Ηλεκτρολόγων Μηχανικών
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Θεσσαλονίκη, Ελλάδα
istefanidi@ece.auth.gr

Περίληψη--Σε αυτό το σύγγραμα παρουσιάζονται η ιστορία, η πρώτη υλοποίηση από την HP κι κάποιες πειραματικές εφαρμογές του memristor.

I. Εισαγωγή

Όποιος έχει παρακολουθήσει ποτέ εργαστηριακό μάθημα ηλεκτρονικής θα είναι εξοικειωμένος με τα θεμελιώδη παθητικά στοιχεία του κυκλώματος: την αντίσταση, τον πυκνωτή και το πηνίο. Ωστόσο, το 1971 ο Leon Chua συμπέρανε από επιχειρήματα συμμετρίας ότι θα έπρεπε να υπάρχει και ένα τέταρτο θεμελιώδες στοιχείο, το οποίο ονόμασε memristor (συντομογραφία των λέξεων memory resistor). Πιο συγκεκριμένα, ο Chua σημείωσε ότι υπάρχουν έξι διαφορετικές μαθηματικές σχέσεις που συνδέουν τα ζεύγη των τεσσάρων θεμελιώδων μεταβλητών του κυκλώματος: ηλεκτρικό ρεύμα i , τάση v , φορτίο q και μαγνητική ροή ϕ . Μία από αυτές τις σχέσεις προσδιορίζεται από τους ορισμούς δύο μεταβλητών και μία άλλη προσδιορίζεται από το νόμο επαγωγής του Faraday. Έτσι, θα πρέπει να υπάρχουν τέσσερα βασικά στοιχεία του κυκλώματος που περιγράφονται από τις υπόλοιπες σχέσεις μεταξύ των μεταβλητών (Σχ. 1). Το στοιχείο που "λείπει" - το memristor, με memristance M - παρέχει μια σχέση μεταξύ φορτίου και ροής, $d\phi = M dq$.



Σχήμα 1: Τα τέσσερα θεμελιώδη ηλεκτρικά στοιχεία δύο ακροδεκτών: αντίσταση, πυκνωτής, πηνίο και memristor.

Το memristor λοιπόν θεωρείτε το 4^ο θεμελιώδες ηλεκτρικό στοιχείο [1] μαζί με την αντίσταση, τον πυκνωτή κι το πηνίο και χρησιμοποιείται κυρίως για τον έλεγχο της ροής του ηλεκτρικού ρεύματος. Το σημαντικότερο πλεονεκτήματα

ενός memristor είναι ότι μπορεί να θυμάται την προηγούμενη κατάσταση του χωρίς παροχή ενέργειας. Η μνημονική αντίσταση είναι το είδος της ηλεκτρικής αντίστασης που διαφέρει με βάση το ηλεκτρικό φορτίο που διέρρευσε στο κύκλωμα προηγουμένως. Επιπλέον, κάθε memristor ως συσκευή χαρακτηρίζεται με τη memristance συνάρτηση εξηγώντας την εξάρτηση του φορτίου με τον ρυθμό ροής. Ένα memristor είναι ένα στοιχείο με δύο ακροδέκτες του οποίου η αντίσταση επηρεάζεται από το πλάτος, την κατεύθυνση και τη διάρκεια της εφαρμοζόμενης τάσης. Ένα memristor μπορεί να παράγει δυναμική-αρνητική αντίσταση διατηρώντας την πιο πρόσφατη κατάστασή του όταν η τάση διακοπεί μέχρι την επόμενη φορά που η τάση εφαρμοστεί ξανά. Το memristor για την ανάλυση κυκλωμάτων μπορεί να οριστεί μαθηματικά ως εξής:

$$v = R(w)i \quad (1)$$

$$\frac{dw}{dt} = f(w, i) \quad (2)$$

όπου

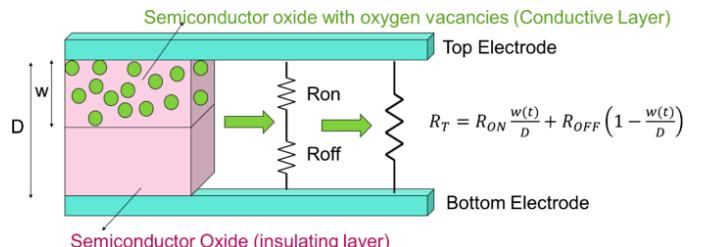
w η μεταβλητή κατάστασης της συσκευής,

R η γενικευμένη αντίσταση.

R και f είναι οι ρητές συναρτήσεις του χρόνου.

Με δύο μεταβλητές αντιστάσεις που συνδέονται σε σειρά (όπως στο Σχ. 2) μπορούμε να προσομοιώσουμε την αντίσταση ενός memristor. Η συνολική τάση στο memristor μπορεί να υπολογιστεί ως εξής:

$$V(t) = (R_{on} \frac{w(t)}{D} + R_{off}(1 - \frac{w(t)}{D}))I(t) \quad (3)$$



Σχήμα 2: Memristor διοξειδίου/οξειδίου του μετάλλου

Κάποιες συμβάσεις που κάνουμε είναι ότι θεωρούμε ότι έχουμε ωμική αγωγμότητα, γραφική ολίσθηση (drift) ιόντων σε ομογενές πεδίο κι ότι η μέση κινητικότητα των ιόντων είναι μ_ν .

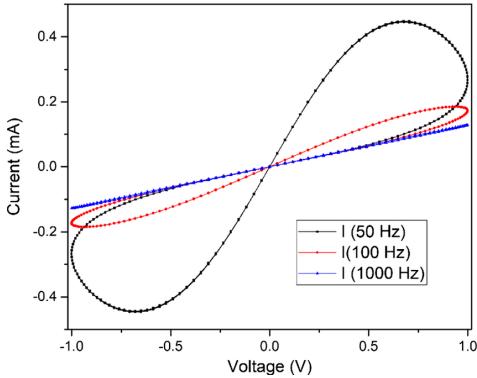
$$\frac{dw(t)}{dt} = \mu_\nu \frac{R_{on}}{D} i(t) \quad (4)$$

To $w(t)$ μπορεί να εκφραστεί ως:

$$w(t) = \mu_\nu \frac{R_{on}}{D} q(t) \quad (5)$$

Συνδυάζοντας τις σχέσεις (4) και (5), μπορούμε να εκφράσουμε της μνημονική αντίσταση της συσκευής ως:

$$m(q) = R_{off}(1 - \frac{\mu_\nu R_{on}}{D^2} q(t)) \quad (6)$$



Σχήμα 3: Χαρακτηριστική memristor.

Ένα ακόμα χαρακτηριστικό του memristor ως ηλεκτρικό στοιχείο είναι ότι η χαρακτηριστική του είναι παραμετρική σχέση που περιγράφεται από 2 εξισώσεις κι σχηματίζει ένα "pinched loop" (Σχ 3). Αυτό που παρατηρήθηκε είναι όσο αυξάνεται η συχνότητα, η κλίση μειώνεται και ο βρόγχος γίνεται όλο κι πιο στενός έως ότου γίνει σχεδόν μια γραμμή, δηλαδή το memristor προσομοιώνει μια αντίσταση.

II. Ιστορία των MEMRISTOR

Αυτός που διατύπωσε πρώτος τις σχέσεις που περιγράφουν ένα memristor ήταν ο Leon Chua το 1971 στο σύγγραμμά του "Memristor-The Missing Circuit Element" [1]. Το πρώτο όμως memristor κατασκευάστηκε 37 χρόνια μετά από την HP [2] το 2008 με ταχύτητα αλλαγής κατάστασης 1ns ($\sim 1\text{GHz}$) κι διαστάσεις 3nm \times 3nm. Ένας από τους λόγους αυτού του μεγάλου χρονικού κενού είναι ότι το memristance γίνεται σημαντικό στην κλίμακα των νανομέτρων λόγω του D^2 στην σχέση (6).

Η πρώτη υλοποίηση λοιπόν από την HP είναι όπως αυτή που φαίνεται στο Σχ. 2 με χρήση TiO_2 ως διοξείδιο του μετάλλου. Όπου μειώνοντας την ποσότητα του TiO_2 μειώνεται η τιμή της αντίστασης του memristor.

III. Τύποι Memristor

Τα memristor χωρίζονται σε αρκετές κατηγορίες βάση των υλικών από τα οποία κατασκευάζονται καθώς επίσης και τον μηχανισμό μεταβολής της αντίστασης με τον οποίο λειτουργούν [3]. Ανάμεσα στα είδη των Memristor εκείνο του δι-οξειδίου μετάλλου αποτελεί ένα από τα πιο συνήθη και είναι εκείνο που συναντήσαμε σε όλες τις εφαρμογές που μελετήσαμε. Τα φαίνομενα μεταβολής της αντίστασης παρατηρούνται σε κατηγορίες δυαδικών οξειδίων όπως TiO_2 , CuO , NiO , CoO , Fe_2O_3 , MoO_3 , VO_2 κ.λπ. Αυτά τα φαίνομενα βασίζονται στη μετάβαση μεταξύ της κατάστασης υψηλής αντίστασης (HRS) και της κατάστασης χαμηλής αντίστασης (LRS), όπου η LRS αντιστοιχεί στο λογικό "1" και η HRS αντιστοιχεί στο λογικό "0". Η εναλλαγή της αντίστασης γίνεται με τρεις τρόπους

- 1) Διπολική εναλλαγή αντίστασης - Απαιτούνται τόσο θετικές όσο και αρνητικές πολικότητες τάσης για τη μεταβολή της διάταξης από LRS σε HRS και αντίστροφα.
- 2) Μονοπολική μεταγωγή αντίστασης: Βασίζεται στο θερμικό φαινόμενο. Η τάση εισόδου οδηγήθηκε σε μία πολικότητα, είτε σε θετικούς είτε σε αρνητικούς κλάδους τάσης. Η μεταγωγή βασίζεται στις τάσεις μετάβασης SET (μετάβαση από OFF σε ON) και RESET (ON σε OFF).
- 3) Μη αναστρέψιμη μεταγωγή: Αυτός ο τύπος επιτρέπει αυστηρά τη μεταγωγή σε διαφορετική κατάσταση μη αναστρέψιμα σε μονόδρομη κατεύθυνση. Π.χ.: Μη πτητική μνήμη και μνήμη WORM (Write Once Read Many times). Αυτός ο τύπος παρατηρείται σε οργανικές διατάξεις.

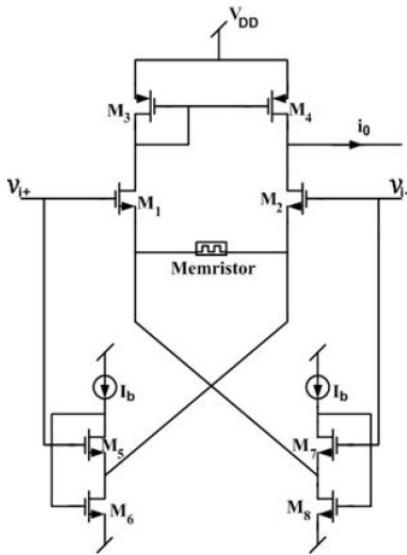
IV. Εφαρμογές

1. OTA

Τα Memristor χωρίζονται επίσης ως προς τον τρόπο με τον οποίο χρησιμοποιούνται. Ανάλογα με την εφαρμογή χρησιμοποιούνται είτε ως μεμονωμένα στοιχεία είτε σε διάταξη Crossbar.

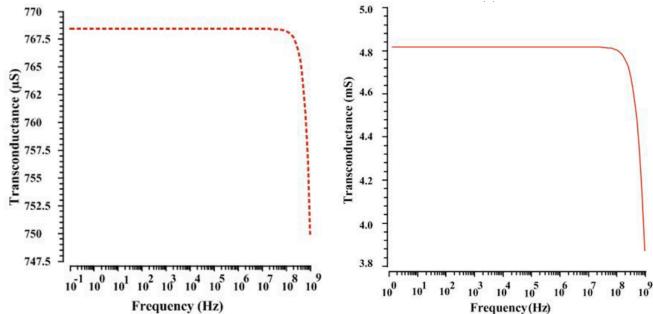
Παράδειγμα μεμονωμένης χρήσης τους αποτελεί η ενσωμάτωση τους σε ένα κύκλωμα ενισχυτή CMOS για βελτίωση της γραφικότητας που μελετήθηκε. Στον εν λόγω παράδειγμα έχουμε έναν Τελεστικό Ενισχυτή Διαγωγμότητας (Operational Transconductance Amplifier) [4]

Όπου χρησιμοποιούμε την Τεχνική Εκφυλισμού Πηγής (Source Degeneration Technique). Η συγκεκριμένη τεχνική βασίζεται στο γεγονός ότι το κέρδος ενός (μεταγωγέα) μίας εισόδου δίνεται από τον τύπο $-gmRD$. Με την κατάλληλη προσθήκη ενός RS, η διαγωγμότητα του κυκλώματος γίνεται $\frac{gm}{1+gmRS}$ συνεπώς το $Gain \approx -\frac{RD}{RS}$ όπου $gmRS \gg 1$, που είναι μια γραφική προσέγγιση. Το memristor σε αυτή την περίπτωση αντικαθιστά την αντίσταση RS. Ως εκ τούτου, η γραφικότητα βελτιώνεται. Παράλληλα χρησιμοποιείται η τεχνική προσαρμοστικής πόλωσης (Adaptive biasing technique) η οποία παρέχει ένα πολύ χαμηλό ρεύμα ηρεμίας, το οποίο μειώνει τη διαρροή στατικής ισχύος. Οι τεχνικές



Σχήμα 4: Hybrid OTA

που χρησιμοποιούνται οδηγούν σε αύξηση της γραμμικότητας, της συνολικής διαγωγμότητας του OTA, της συχνότητας μοναδιαίου κέρδους, του κέρδους εύρους ζώνης και της μείωσης κατανάλωσης ενέργειας. Αυτό φαίνεται από τα αποτελέσματα του πειράματος: στο Σχήμα 3 παρουσιάζεται το αποτέλεσμα προσομοίωσης της ασ πάρκρισης του προτεινόμενου κυκλώματος. Η προσομοιωμένη καμπύλη δείχνει την τιμή της διαγωγμότητας ως 4,82mS μέχρι την περιοχή συχνοτήτων GHz. Το Σχ. 5α' δείχνει τη συνολική

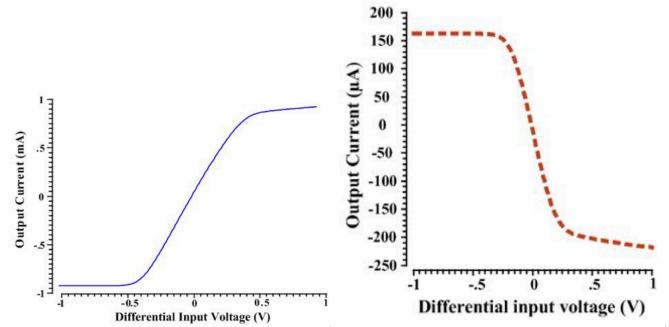


(α') Η συνολική καμπύλη διαγω- (β') Η συνολική καμπύλη διαγμότητας της τεχνικής εκφυλ- γωγμότητας του προτεινόμενου σμού πηγής με βάση το MOS κυκλώματος.

Σχήμα 5: Αποτέλεσμα προσομοίωσης της διαγωγμότητας σε σχέση με την απόκριση συχνότητας.

διαγωγμότητα του OTA με ένα NMOS που λειτουργεί ως αντίσταση για την τεχνική εκφυλισμού πηγής με τιμή διαγωγμότητας 768μS. Η τιμή της διαγωγμότητας παρατηρείται να αυξάνεται από 768μS σε 4,82mS μέσω της αντικατάστασης του OTA που βασίζεται σε MOS με το memristor, όπως φαίνεται στο Σχ. 5β'. Δεν υπάρχει απόκλιση στην τιμή της διαγωγμότητας μέχρι την περιοχή συχνοτήτων GHz. Το προτεινόμενο κύκλωμα τροφοδοτεί ρεύμα υψηλής τιμής στην έξοδο ακόμη και όταν έχει υψηλότερο φορτίο. Το κύκλωμα

έχει κέρδος DC 27,1 dB και συχνότητα μοναδιαίου κέρδους 9,7 MHz, παρότι μπορεί να δίνει μικρότερο κέρδος DC λόγω μίας μόνο βαθμίδας, μπορεί να βελτιωθεί περαιτέρω με τη χρήση πολλαπλών βαθμίδων. Επίσης το περιθώριο φάσης είναι περίπου 86° για χωρητικότητα φορτίου 80pF, γεγονός που σηματοδοτεί τη σταθερότητα του συστήματος. Τέλος Σχ. 6α' η καμπύλη της συνάρτησης μεταφοράς της προσομοίωσης δείχνει ότι το gm είναι σταθερό κατά τη διάρκεια της εναλλαγής τάσης εισόδου ±400 mV. Στο Σχ. 6β' παρουσιάζονται οι χαρακτηριστικές μεταφοράς του συμβατικού OTA με σταθερό gm σε εύρος ±150 mV. Η τιμή γραμμικότητας αυξάνεται σε μια ευρύτερη τάση από ±150 mV έως ±400 mV μέσω της αντικατάστασης του κυκλώματος MOS με το memristor (Πίνακας I).



Σχήμα 6: Προσομοιωμένο αποτέλεσμα της διαφορικής μεταβολής τάσης ρεύματος εξόδου.

Πίνακας I: Performance Summary of this work

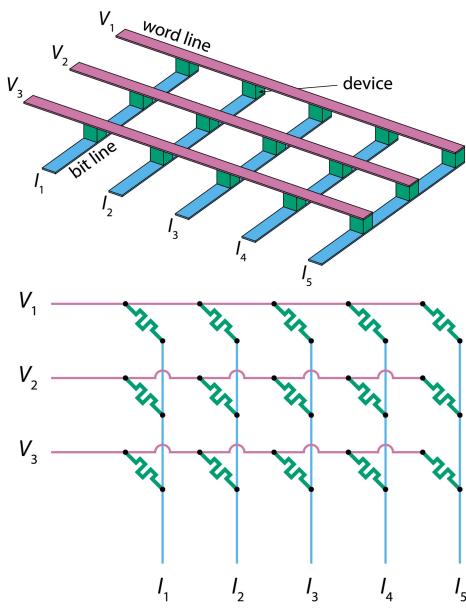
Parameter	This Work
Technology	0.18 μm CMOS
Linearity	0.4 V to +0.4 V
Gain (dB)	27.1
UGF (MHz)	9.7
PM (deg)	86
CMRR (dB)	51.7

Αυτή η εργασία μπορεί να επεκταθεί στην υλοποίηση δομών φίλτρων υψηλότερης τάξης με την κλιμάκωση της δεδομένης δομής. Το προτεινόμενο κύκλωμα μπορεί να βρει εφαρμογή στη σχεδίαση μπροστινής πλευράς ασύρματου δέκτη. Πιο προηγμένες τεχνολογίες μπορούν να εφαρμοστούν στη χρήση του προτεινόμενου υβριδικού κυκλώματος OTA σε λειτουργίες υψηλής συχνότητας στο μέλλον.

2. Crossbar

Ο άλλος τρόπος χρήσης των memristor είναι σε διάταξη Crossbar. Η διάταξη Crossbar σχηματίζεται από οριζόντια και κάθετα ηλεκτρικά-απομονωμένα μεταξύ τους ηλεκτρόδια. Τα οριζόντια αποτελούν έστω το άνω ηλεκτρόδιο τα κάθετα το κάτω, όπως φαίνεται στο Σχήμα 7 με τον μονωτή (π.χ. κάποιο οξείδιο) να βρίσκεται ανάμεσα σε αυτά

τα ηλεκτρόδια σχηματίζοντας τα memristor (δηλαδή την μεταβλητή αντίσταση) στις τομές τους.



Σχήμα 7: Διάταξη memristor σε crossbar.

Η λειτουργία τους ελέγχεται από τις τάσεις εισόδου V_i , I_j για ευκολή ευρεση του Memristance. Ιδιαίτερο χαρακτηριστικό τους αποτελεί η εκτέλεση πράξεων όπως το εσωτερικό γινόμενο με απλή ρύθμιση της προϋπάρχουσας αντίστασης και του σήματος που έρχεται από την είσοδο V δίνοντας το εσωτερικό γινόμενο πινάκων στην έξοδο I . Το Memristor Crossbar λειτουργεί όπως το σύνολο των συνάψεων μεταξύ των νευρώνων στα στρώματα ενός νευρωνικού δικτύου. Οι συνάψεις που χρησιμοποιούν την τεχνολογία των Memristor μπορούν επίσης να έχουν δυαδικές ή μη δυαδικές καταστάσεις ανάλογα με τις εφαρμογές και τους πρακτικούς περιορισμούς του προγραμματισμού των εκάστοτε διατάξεων.

Επίσης είναι μια διάταξη που προβλέπεται να χρησιμοποιηθεί για την κατασκευή μνήμης αλλά και για την αντικατάσταση ξεχωριστών μονάδων λογισμικού και κυκλωμάτων όπως FPGA με αναλογικά on-chip κυκλώματα αποτελούμενα από memristor. Το μικρό μεγεθός τους οδηγεί στην ελαχιστοποίηση της απαιτούμενης επιφάνειας, οι μικρές απαιτήσεις ενέργειας και η μη πτητική τους μνήμη τα καθιστούν ιδιαίτερα χρήσιμα στο σχεδιασμό on-chip τεχνολογιών απλοποιώντας και επιταχύνοντας αυτές τις συσκευές.

Παράλληλα λόγο της έμφυτης ικανότητας των memristor να διατηρούν στάθμες (ακόμη και όταν αποσυνδέονται από το ρεύμα) μπορούν να αποθηκεύσουν σε ένα μόνο memristor περισσότερα του ενός bit. Αυτή η ιδιότητα τα κάνει πολύ χρήσιμα για την δημιουργία μνήμης από Memristor Crossbar. Ωστόσο, παρότι μια ιδανική μη πτητική συσκευή memristor μπορεί να αποθηκεύσει και να προγραμματιστεί σε οποιαδήποτε συγκεκριμένη τιμή μεταξύ RON και ROFF, οι πραγματικές συσκευές memristor μπορεί να έχουν προ-

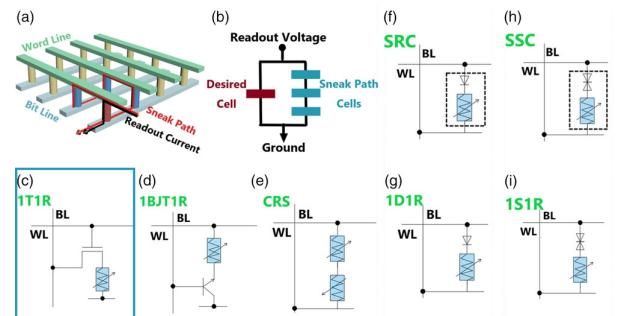
βλήματα με τη μετάβαση στις ενδιάμεσες τιμές αντίστασης και τη συνέπεια με την οποία παίρνουν τις ενδιάμεσες αυτές τιμές.

Ακόμη ένα σημαντικό στοιχείο αυτής της διάταξης είναι ότι η φυσική λειτουργία ενός Memristor Crossbar Array μοιάζει με το σύνολο των συνάψεων μεταξύ των νευρώνων στα στρώματα ενός νευρωνικού δικτύου. Οι συνάψεις που χρησιμοποιούν την τεχνολογία των Memristor μπορούν επίσης να έχουν δυαδικές ή μη δυαδικές καταστάσεις ανάλογα με τις εφαρμογές και τους πρακτικούς περιορισμούς του προγραμματισμού των εκάστοτε διατάξεων. Αυτή η εκ φύσεως συμπεριφορά τα καθιστά κατάλληλα για ανάπτυξη hardware για εφαρμογές νευρωνικών δικτύων.

Τέλος ένα πιο τεχνικό κυρίως χαρακτηριστικό είναι ότι τα memristor είναι στοιχεία δύο ακροδεκτών το οποίο κάνει την δημιουργία του Crossbar μια αρκετά απλή διάταξη από κατασκευαστική σκοπιά. Οι απλές δομές δύο ακροδεκτών μετάλλου-μονωτή-μετάλλου (MIM) των μεμριστών τα καθιστούν ικανά να ενσωματωθούν σε πυκνές συστοιχίες Crossbar array.

3. Leakage Reduction

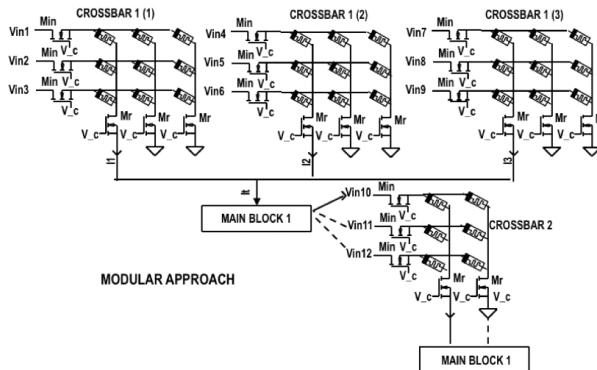
Ένα από τα πιο συχνά προβλήματα των crossbar arrays μεγάλης κλίμακας είναι τα ρεύματα διαρροής. Για την επίλυση αυτού του προβλήματος εχουν ερευνηθεί διάφορες λύσεις. Η πρώτη έχει να κάνει με συνδυαστική αρχιτεκτονική, δηλαδή η κατασκευή όπου σε ένα Crossbar όπου υπάρχει ένα memristor υπάρχει ένα ακόμη ηλεκτρικό στοιχείο (π.χ. δίοδος, τρανζίστορ) συνδεδεμένο σε σειρά όπως φαίνεται στο Σχήμα 8. Η πιο ευρέως χρησιμοποιούμενη αρχιτεκτονική για αυτή την μέθοδο είναι εκείνη με 1 τρανζίστορ 1 memristor, 1T 1R. Ωστόσο αυτή η μέθοδος αυξάνει τόσο την απαραίτητη επιφάνεια όσο και την κατανάλωση ενέργειας, η ελαχιστοποίηση των οποίων ήταν ένας βασικός λόγος της χρήσης Memristor Crossbar Arrays έναντι των υπαρχόντων τεχνολογιών.



Σχήμα 8: Συνδιαστικές λύσεις για τη μείωση του ρεύματος διαρροής.

Η δεύτερη λύση είναι εκείνη της κατανεμημένης προσέγγισης. Αυτή η προσέγγιση χρησιμοποιείται σε ένα από τα παραδείγματα χρήσεις που θα ακολουθήσουν. Ουσιαστικά σε αυτή την μέθοδο ένα μεγάλο crossbar χωρίζεται σε μικρότερα crossbars, όπως φαίνεται στο Σχήμα 9, και το

ρεύμα από όλα τα επιμέρους crossbars αθροίζεται για να επεξεργαστεί μέσω της συνάρτησης ενεργοποίησης στο βασικό μπλοκ που εκτελεί τον αλγόριθμο της εφαρμογής (MB1 σε αυτήν την περίπτωση). Όπως φαίνεται στα αποτελέσματα προσομοίωσης, αυτή η προσέγγιση επιτρέπει την επίτευξη παρόμοιας ακρίβειας απόδοσης, όπως η προσέγγιση single crossbar. Επιπλέον, εάν το δίκτυο κλιμακωθεί, η διαδοχική επεξεργασία μπορεί να εισάγει τον περιορισμό στο σύστημα με τη μορφή μειωμένης ταχύτητας επεξεργασίας. Σε αυτή την περίπτωση, μπορεί να εισαχθεί η παράλληλη επεξεργασία, η οποία περιλαμβάνει τον ταυτόχρονο υπολογισμό και την ταυτόχρονη εκτέλεση των υπολογισμών εξόδου. Η κατανεμημένη προσέγγιση μπορεί επίσης να είναι χρήσιμη, ωστόσο, κάθε επιμέρους crossbar θα πρέπει να διαθέτει αντίστοιχα μπλοκ επεξεργασίας για τον αλγόριθμο της εφαρμογής. Αυτό εισάγει πρόσθετη πολυπλοκότητα για το σύστημα και αυξάνει την επιφάνεια και την ισχύ του chip, αλλά μειώνει το χρόνο επεξεργασίας. Η κατανεμημένη προσέγγιση μπορεί επίσης να επιτρέψει την αφάρεση των αναλογικών μονάδων αποθήκευσης. Καθώς θα μειωθεί το μέγεθος της εγκάρσιας ράβδου, αντί της αναλογικής μονάδας αποθήκευσης μπορεί να χρησιμοποιηθεί ένα σήμα με χρονική καθυστέρηση που παράγεται από ένα κύκλωμα καθυστέρησης σήματος.



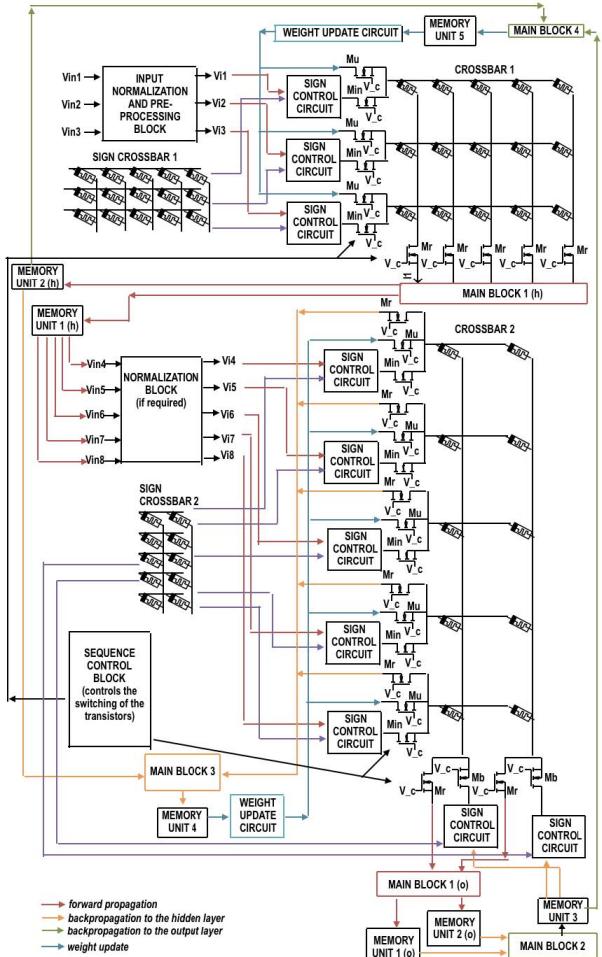
Σχήμα 9: Κατανεμημένη προσέγγιση για τη μείωση του ρεύματος διαρροής και της πολυπλοκότητας στον προγραμματισμό της συστοιχίας 1M.

4. Memristive Neural Network

Μια ακόμη πολύ σημαντική εφαρμογή των memristor είναι η χρήση τους σε κυκλώματα νευρωνικών δικτύων.

Υπάρχουν αρκετές μελέτες που προτείνουν την υλοποίηση μνημονικού νευρωνικού δικτύου με backpropagation σε ψηφιακό και μικτό σήμα. Ωστόσο, τα αναλογικά κυκλώματα εκμάθησης που βασίζονται στον συμβατικό αλγόριθμο εκμάθησης backpropagation σε memristive crossbars δεν έχουν υλοποιηθεί πλήρως. Η υλοποίηση ενός τέτοιου αλγορίθμου εκμάθησης δίνει την ευκαιρία να δημιουργηθεί μια αναλογική αρχιτεκτονική εκμάθησης βασισμένη σε hardware. Αυτό θα μετέφερε τους αλγορίθμους εκμάθησης από τις ξεχωριστές μονάδες λογισμικού και FPGA σε αναλογικά κυκλώματα εκμάθησης on-chip, τα οποία μπορούν να απλοποιήσουν και να επιταχύνουν τη διαδικασία εκμάθησης.

Οι αλγεβρικές και ολοκληρωδιαφορικές πράξεις του αλγορίθμου εκμάθησης backpropagation, οι οποίες είναι δύσκολο να υλοποιηθούν με ακρίβεια σε ένα ψηφιακό σύστημα, είναι εγγενώς διαθέσιμες σε ένα αναλογικό υπολογιστικό σύστημα. Παρακάτω προτείνετε ένα αναλογικό backpropagation κύκλωμα εκμάθησης [5] (Σχ. 10) που επιτρέπει μια φυσική υλοποίηση αναλογικής αρχιτεκτονικής νευρωνικού δικτύου on-chip, η οποία είναι επωφελής από την άποψη της ταχύτητας επεξεργασίας, της μείωσης της συνολικής ισχύος και της μικρότερης πολυπλοκότητας, σε σύγκριση με τα αντίστοιχα ψηφιακά κυκλώματα.

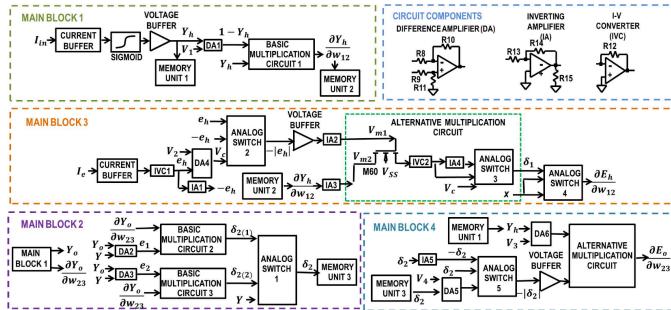


Σχήμα 10: Αρχιτεκτονική αναλογικού backpropagation κυκλώματος εκμάθησης με memristive crossbars.

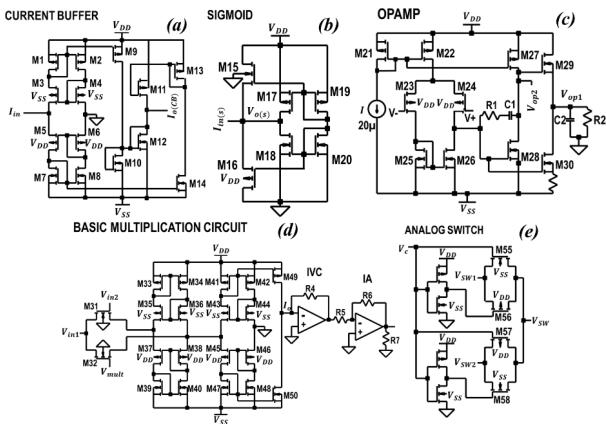
Στο κύκλωμα του σχήματος 10 τα memristive crossbar arrays μιμούνται το σύνολο των συνάψεων μεταξύ των νευρώνων στα στρώματα του νευρωνικού δικτύου. Οι συνάψεις μπορούν να είναι δυαδικές ή μη δυαδικές ανάλογα με τις εφαρμογές και τους πρακτικούς περιορισμούς του προγραμματισμού των καταστάσεων της συσκευής memristor. Ενώ μια ιδανική μη πτητική συσκευή memristor μπορεί να αποθηκεύσει και να προγραμματιστεί σε οποιαδήποτε συγκεκριμένη τιμή μεταξύ R_{ON} και R_{OFF} , οι πραγματικές συσκευές memristor μπορεί να έχουν προβλήματα με τη

μετάβαση στις ενδιάμεσες τιμές αντίστασης. Είναι ευκολότερο και απλούστερο να μεταβεί το memristor είτε στην κατάσταση R_{ON} είτε στην κατάσταση R_{OFF} .

Το παράδειγμα που παρουσιάζεται στο Σχήμα 10 δείχνει ένα βασικό ANN (Artificial Neural Networks) τριών επιπέδων με την προτεινόμενη αρχιτεκτονική backpropagation, το κύκλωμα ελέγχου και τα κυκλώματα ενημέρωσης βάρους. Το νευρωνικό δίκτυο έχει τρεις νευρώνες εισόδου, δύο νευρώνες εξόδου και πέντε νευρώνες σε ένα κρυφό στρώμα. Η λειτουργία του crossbar και η εναλλαγή μεταξύ των λειτουργιών forward propagation, backpropagation και weight update ελέγχεται από τα τρανζίστορ M_{in} , M_u και M_r , τα οποία με τη σειρά τους ελέγχονται από το μπλοκ ελέγχου ακολουθίας. Το CROSSBAR 1 αντιστοιχεί στο σύνολο των συναπτικών συνδέσεων μεταξύ του στρώματος εισόδου και του κρυφού στρώματος και το CROSSBAR 2 αντιπροσωπεύει τις συνάψεις που συνδέουν το κρυφό στρώμα με το στρώμα εξόδου. Τα τρία σήματα εισόδου απεικονίζονται ως V_{in1} , V_{in2} και V_{in3} και τα αντίστοιχα κανονικοποιημένα σήματα εισόδου απεικονίζονται ως V_{i1} , V_{i2} και V_{i3} , αντίστοιχα. Το εύρος των σημάτων εξόδου από το κύκλωμα κανονικοποίησης εξαρτάται από την εφαρμογή, τους περιορισμούς των memristors και τη γραμμικότητα των τρανζίστορ.

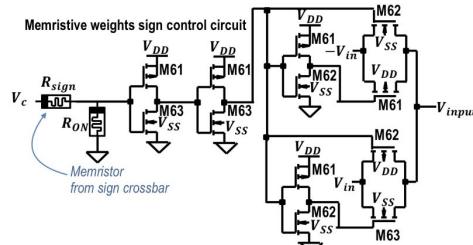


Σχήμα 11: Η αρχιτεκτονική σε επίπεδο κυκλώματος της προτεινόμενης υλοποίησης backpropagation.



Σχήμα 12: Στοιχεία κυκλώματος της προτεινόμενης εφαρμογής backpropagation.

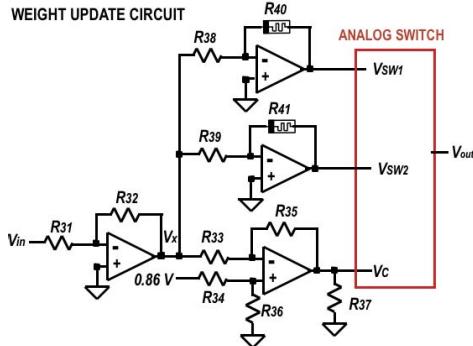
Η υλοποίηση σε επίπεδο κυκλώματος όλων των μπλοκ backpropagation απεικονίζεται στο Σχ. 11, ενώ τα στοιχεία που χρησιμοποιούνται στα κυκλώματα φαίνονται στο Σχ. 12. Το MB1 εκτελεί forward propagation για τη συμβατική αρχιτεκτονική backpropagation. Το MB2 εκτελεί τη διαδικασία backpropagation μέσω του στρώματος εξόδου, η οποία ολοκληρώνεται από το MB4. Το MB3 εκτελεί την backpropagation μέσω του κρυμμένου στρώματος. Οι υλοποίησεις σε επίπεδο κυκλώματος των στοιχείων από τα κύρια μπλοκ backpropagation παρουσιάζονται στο Σχήμα 12. Το Σχ. 12(a) απεικονίζει την υλοποίηση του ρυθμιστή ρεύματος. Στο Σχ. 12(b) παρουσιάζεται η υλοποίηση της σιγμοειδούς συνάρτησης. Στο Σχ. 12(c) παρουσιάζεται το κύκλωμα OpAmp. Στο Σχ. 12(d) απεικονίζεται το κύκλωμα πολλαπλασιασμού με βάση τη διαφορά ρεύματος στα τρανζίστορ M_{31} και M_{32} . Στο Σχ. 12(e) παρουσιάζεται η υλοποίηση του κυκλώματος αναλογικού διακόπτη.



Σχήμα 13: Κύκλωμα ελέγχου προσήμου memristive βάρους

Καθώς τα βάρη του νευρωνικού δικτύου μπορούν να είναι τόσο θετικά όσο και αρνητικά και το αρνητικό βάρος δεν μπορεί πρακτικά να υλοποιηθεί από το memristor, απαιτείται η υλοποίηση του επιπλέον κυκλώματος ελέγχου προσήμου του βάρους. Για κάθε αρνητικό βάρος, το πρόσημο της τάσης εισόδου αλλάζει. Η λύση είναι να αποθηκεύεται το πρόσημο για κάθε βάρος στα επιπλέον memristive crossbar στοιχεία. Προτείνεται λοιπόν το κύκλωμα ελέγχου του πρόσημου των memristive βαρών που παρουσιάζεται στο Σχήμα 13. Το πρόσημο κάθε memristor στο crossbar αποθηκεύεται στο memristive crossbar ή σε ξεχωριστά memristors ως R_{ON} ή R_{OFF} . Για την επίτευξη του καλύτερου δυνατού trade-off μεταξύ της απαιτούμενης επιφάνειας, της ισχύος και του χρόνου επεξεργασίας στο κύκλωμα, επιλέγετε υλοποίηση η οποία επιτρέπει την ανάγνωση του προσήμου όλων των memristors με μια μόνο στήλη. Η υλοποίηση αυτή απαιτεί τόσα κυκλώματα ανάγνωσης προσήμου όσα οι γραμμές στο crossbar.

Τέλος η υλοποίηση του κυκλώματος ενημέρωσης memristive βάρους απεικονίζεται στο Σχήμα 14. Το κύκλωμα ενημέρωσης βάρους προσδιορίζει το πλάτος του παλμού που απαιτείται για τον προγραμματισμό του memristor σε μια συστοιχία crossbar, ανάλογα με την υπολογισμένη τιμή ενημέρωσης βάρους από τα MB2 και MB4. Το κύκλωμα είναι προσαρμόσιμο για διαφορετικές τιμές εκμάθησης λόγω της εφαρμογής των memristive διατάξεων R_{40} και R_{41} στους ενισχυτές. Όλες οι αντιστάσεις στο κύκλωμα ενημέρωσης βάρους είναι 1k και τα memristors προγραμματίζονται

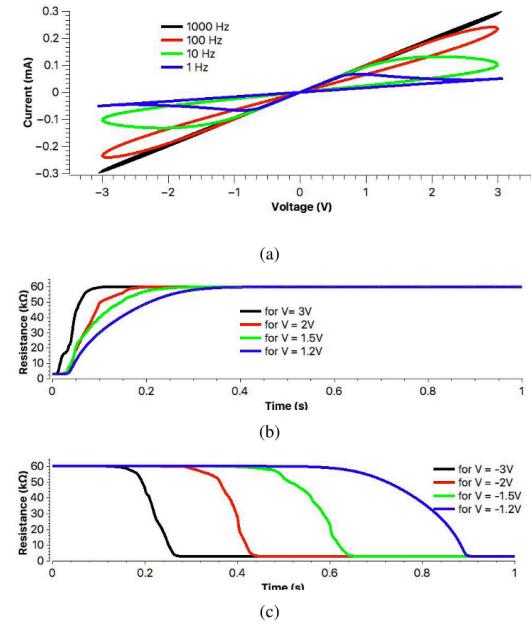


Σχήμα 14: Κύκλωμα ενημέρωσης memristive βάρους

λαμβάνοντας υπόψη τον απαιτούμενο ρυθμό εκμάθησης. Καθώς το αρνητικό (για να μεταβούμε από R_{OFF} σε R_{ON}) και το θετικό (για να μεταβούμε από R_{ON} σε R_{OFF}) πλάτος προγραμματισμού δεν έχουν το ίδιο πλάτος για τα memristors, ο αναλογικός διακόπτης επιλέγει το πλάτος του σήματος με βάση το πρόσθιμο της τάσης εισόδου από τα MB1 και MB4. Η υλοποίηση του αναλογικού διακόπτη φαίνεται στο Σχήμα 12(ε). Το μετατοπισμένο σήμα εισόδου εφαρμόζεται στον έλεγχο του αναλογικού διακόπτη V_c που καθορίζει, ποια είσοδος στον διακόπτη πρέπει να επιλέγει V_{SW1} ή V_{SW2} . Η είσοδος στο V_{SW1} αντιστοιχεί στη θετική τάση εισόδου, ενώ το V_{SW2} αντιστοιχεί στην αρνητική τάση εισόδου.

Το μοντέλο του memristor που χρησιμοποιήθηκε στις προσομοιώσεις είναι το τροποποιημένο μοντέλο S της Biolek [6] για memristor HP TiO₂ με τάση κατωφλίου V_{th} 1V [2]. Αυτό το μοντέλο memristor έχει αναπτυχθεί για προσομοιώσεις μεγάλης κλίμακας για την απλοποίηση του υπολογισμού και της επεξεργασίας [6]. Τα χαρακτηριστικά του memristor και ο χρόνος μεταγωγής για $R_{ON} = 3k\Omega$ και $R_{OFF} = 62k\Omega$ παρουσιάζονται στο Σχήμα 15. Το Σχ 15 (b) και το Σχ 15 (γ) απεικονίζουν τη διαδικασία ενημέρωσης του memristor εφαρμόζοντας παλμό 1s με διαφορετικά πλάτη. Ο χρόνος μεταγωγής είναι μεγάλος και η ταχύτητα της διαδικασίας εκμάθησης με τα μνημονικά στοιχεία είναι αργή. Ωστόσο, η διαδικασία εκμάθησης και εκπαίδευσης είναι μια διαδικασία που πραγματοποιείται μία φορά στο νευρωνικό δίκτυο. Μετά την εκπαίδευση κατά το στάδιο της δοκιμής, ο χρόνος ανάγνωσης είναι μικρός και η επεξεργασία των δεδομένων είναι γρήγορη.

Ο Πίνακας II παρουσιάζει την επιφάνειας του chip και της μέγιστης διάχυσης ισχύος για τα ξεχωριστά στοιχεία για την υλοποίηση του αναλογικού κυκλώματος εκμάθησης και πρόσθετων στοιχείων και συναρτήσεων ενεργοποίησης. Επίσης, παρουσιάζεται η επιφάνειας και η διάχυση ισχύος για ένα μικρό crossbar. Ο Πίνακας III παρουσιάζει την επιφάνεια στο chip και τη μέγιστη διάχυση ισχύος για ξεχωριστά εξαρτήματα για τα κύρια τμήματα της προτεινόμενης υλοποίησης του αναλογικού κυκλώματος εκμάθησης backpropagation.



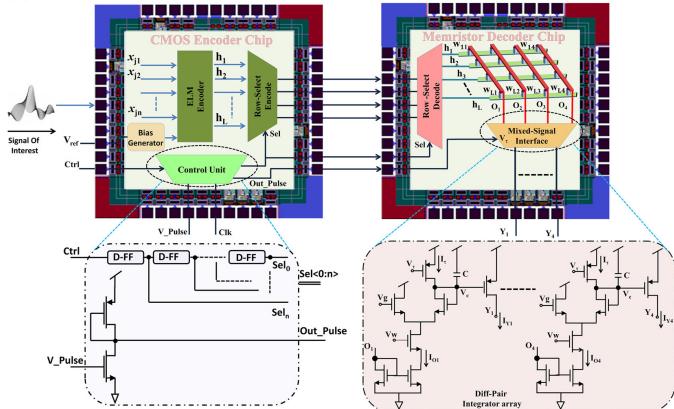
Σχήμα 15: Χαρακτηριστικά memristor: (a) hysteresis για διαφορετικές συχνότητες με $R_{initial} = 10k\Omega$, (b) χρόνος μεταγωγής από $R_{ON} = 3k\Omega$ σε $R_{OFF} = 62k\Omega$ για διάφορα πλάτη τάσεων, (c) χρόνος μεταγωγής από $R_{OFF} = 62k\Omega$ σε $R_{ON} = 3k\Omega$ για διάφορα πλάτη τάσεων.

Πίνακας II: Power consumption and on-chip area calculation for the separate circuit components

Circuit component	Power consumption	On-chip area
Crossbar (4 input neurons and output neurons)	10 μW	1.36 μm^2
Crossbar with control switches	1200 μW	115.3 μm^2
Weight sign control circuit	195.1 μW	16.64 μm^2
Sigmoid	11.4 μW	184.00 μm^2
Current buffer	149.0 μW	280.00 μm^2
OpAmp (maximum)	39.8 mW	2801.76 mm^2
Analog switch current	162.3 mW	1.55 μm^2
Approximate driven sigmoid/tangent	52.9 mW	2118.00 μm^2
Approximate voltage driven sigmoid/tangent	41.2pW	0.40 μm^2
Linear activation units with diode	963.7 μW	244 μm^2
Linear activation unit with switch	23.214 mW	951.06 μm^2
Weight update circuit	14.34 mW	1269.63 μm^2

Πίνακας III: Area and power calculations for the main blocks

Configuration	Area (μm^2)	Maximum Power (mW)
MB1 (hidden layer)	4885.86	3.70
MB2 + MB1 (output layer)	8264.88	10.64
MB3	15238.69	61.78
MB4	9734.33	39.53
Total	38123.76	115.65



Σχήμα 16: CMOS encoder chip and Memristor decoder chip

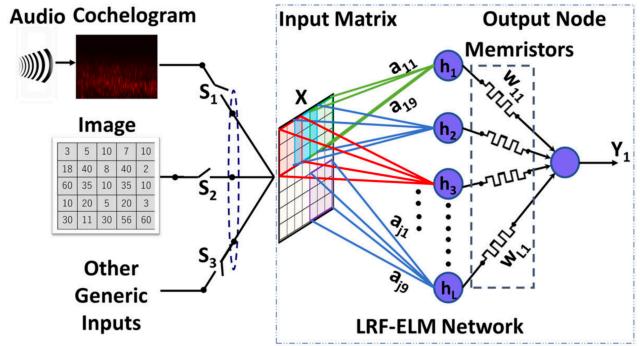
5. Encoder/Decoder

Παρακάτω θα περιγραφεί η λειτουργία του Memristor Crossbar Array σε μία εφαρμογή υβριδικής αρχιτεκτονικής βασισμένης σε δισδιάστατη συστοιχία memristor crossbar και ολοκληρωμένο κύκλωμα CMOS για edge computing [7].

Η προτεινόμενη αρχιτεκτονική αποτελείται από ένα τσιπ κωδικοποιητή CMOS, ακολουθόμενο από το τσιπ αποκωδικοποιητή Memristor, όπως φαίνεται στο Σχήμα 16. Οι δύο αυτές μονάδες λειτουργούν παράλληλα, αλλά απομονωμένες μεταξύ τους. Το τσιπ κωδικοποιητή CMOS αποτελείται από: έναν κωδικοποιητή ELM, μια μονάδα κωδικοποίησης επιλογής σειράς, μια μονάδα γεννήτριας πόλωσης και μια μονάδα ελέγχου. Ο πυρήνας της προτεινόμενης αρχιτεκτονικής χρησιμοποιεί τον αλγόριθμο LRF-ELM ο οποίος είναι μια παραλλαγή του αλγορίθμου ELM όπου τα βάρη μεταξύ εισόδου και κρυφού στρώματος είναι τοπικά και τυχαία. Τα μόνα εκπαιδεύσιμα βάρη είναι τα βάρη εξόδου, τα οποία μπορούν να μάθουν με τη μέθοδο των ελαχίστων τετραγώνων για μια δεδομένη εργασία ταξινόμησης ή παλινδρόμησης. Η έξοδος αυτού του κωδικοποιητή ELM (μια γραμμή του πίνακα H) (Σχ. 16) περνά ως είσοδος στη μονάδα κωδικοποίησης επιλογής σειράς. Ανάλογα με το σήμα επιλογής ($Sel < 0 : n >$), η έξοδος του κρυφού κόμβου περνάει στο chip Memristor Decoder.

Το τσιπ Memristor Decoder με τη σειρά του αποτελείται από: έναν αποκωδικοποιητή επιλογής σειράς, μια συστοιχία memristor crossbar και μια μονάδα διεπαφής μικτού σήματος. Ο πίνακας βάρους W υλοποιείται με τη χρήση μιας συστοιχίας memristor crossbars από υλικά 2D, όπως φαίνεται στο τσιπ Memristor Decoder του Σχήματος 16. Κάθε κόμβος εξόδου του δικτύου LRF-ELM του Σχ. 17 είναι ένας διαφορικός ολοκληρωτής ζεύγους (DPI), του οποίου η συστοιχία υλοποιείται στη μονάδα διεπαφής μικτού σήματος του chip Memristor Decoder. Κατά τη διάρκεια της εκπαίδευσης και της εξαγωγής συμπερασμάτων καταγράφηκε η μη γραμμικότητα του κυκλώματος CMOS και τα βάρη αποθηκεύτηκαν στα memristor ως κβαντισμένες καταστάσεις αγωγιμότητας.

Τα εκπαιδευμένα πλέον βάρη μεταξύ του κρυμμένου κόμ-



Σχήμα 17: Υβριδική αρχιτεκτονική ταξινόμησης LRF-ELM η οποία μπορεί να χρησιμοποιηθεί για δεδομένα υψηλών διαστάσεων εισόδου, όπως εικόνα και φασματογράφημα ήχου.

βου και του κόμβου εξόδου προγραμματίστηκαν και βρίσκονται στο memristor crossbar, όπου κάθε συσκευή memristor μπορεί να έχει πολλαπλές κβαντισμένες καταστάσεις αγωγιμότητας. Το memristor crossbar εκτελεί τη λειτουργία MAC μεταξύ της τάσης εξόδου των κρυφών κόμβων και της αγωγιμότητας των memristor. Επίσης, αποκωδικοποιεί το χαρακτηριστικό εισόδου από το κρυφό στρώμα σε έναν χώρο υψηλότερων διαστάσεων.

Οι τάσεις των κρυφών κόμβων h_1-h_n του δικτύου LRF-ELM περνούν μέσω της μονάδας κωδικοποίησης επιλογής σειράς στη μονάδα αποκωδικοποίησης επιλογής σειράς σε χρονικά πολλαπλάσια βάση. Ο χρονισμός επιλογής σειράς ελέγχεται με ακρίβεια χρησιμοποιώντας το μπλοκ ελέγχου χρονισμού του τσιπ CMOS. Αυτές οι τάσεις διαβιβάζονται στη συνέχεια ως είσοδος στο memristor Crossbar, όπου κάθε διάταξη του είναι προγραμματίζενη σε μία από τις 26 πιθανές καταστάσεις που αντιστοιχούν σε 26 προ-εκπαιδευμένες καταστάσεις αγωγιμότητας. Η έξοδος κάθε στήλης του crossbar αποθηκεύεται ως φορτίο σε έναν πυκνωτή ενός κυκλώματος ολοκληρωτή DPI (Σχ. 16). Σε κάθε χρονικό βήμα, ο πυκνωτής (C στο Σχ. 17) στη συστοιχία DPI συνεχίζει να ενσωματώνει την έξοδο από τα Crossbar έως ότου ανιχνευθούν όλοι οι κρυφοί κόμβοι για συγκεκριμένες εισόδους δεδομένων. Τα τελικά αποτελέσματα λαμβάνονται στη συνέχεια με τη σύγκριση των τάσεων στους κόμβους εξόδου DPI.

Τώρα τα βασικά πλεονεκτήματα του προτεινόμενου σχεδιασμού είναι: (1) Το πλαίσιο είναι ανεκτικό στη μεταβλητότητα της συσκευής και χρησιμοποιεί την εγγενή αναντιστοιχία της συσκευής MOS ως τυχαία βάρη προς όφελός του, άρα ξεφεύγει από κάθε απαίτηση αποθήκευσης μνήμης για τα βάρη του πρώτου στρώματος. (2) Η αρχιτεκτονική είναι ανθεκτική στις μεταβολές της διεργασίας και επομένως μπορεί να σχεδιαστεί σε πολύ χαμηλότερο τεχνολογικό κόμβο CMOS, οδηγώντας σε περαιτέρω βελτίωση των παραμέτρων απόδοσης του συστήματος. (3) Η συνολική ενεργειακή κατανάλωση του προτεινόμενου συστήματος είναι

πολύ χαμηλή- συνεπώς, ένας καλός υποψήφιος για εφαρμογές υπολογισμού ακμών. (4) Στο προτεινόμενο υβριδικό σύστημα, το CMOS Encoder Chip και το Memristor Decoder Chip είναι δύο σημαντικά μέρη του πλαισίου ELM και κάθε μέρος μπορεί να βελτιστοποιηθεί ξεχωριστά. Αυτή η διάταξη επιτρέπει τη διερεύνηση άλλων τύπων αναδυόμενων μνημών απλώς με την αντικατάσταση του chip αποκωδικοποιητή, χωρίς να επηρεάζεται το συνολικό σύστημα. (5) Η συστοιχία memristor εκτελεί τη λειτουργία πολλαπλασιασμού και συσσώρευσης (MAC) μεταξύ του χάρτη χαρακτηριστικών εισόδου υψηλότερων διαστάσεων και των αποθηκευμένων κβαντισμένων βαρών σε χρονική βάση. (6) Το πλαίσιο LRF-ELM για την ταξινόμηση συνόλων δεδομένων σε πραγματικό χρόνο χρησιμοποιεί κύκλωμα CMOS 9D Gaussian που λειτουργεί σε καθεστώς υποκατωφλίου. Το κύκλωμα, ως εκ τούτου, προσφέρει την επιθυμητή μη γραμμικότητα με ελάχιστο αριθμό τρανζίστορ, και συνεπώς προσφέρει μείωση της επιφάνειας της ισχύος. (7) Ο σχεδιασμός προσφέρει ένα υποκατάστατο για την ψηφιακή μνήμη με τη χρήση memristor clobars πολλαπλών καταστάσεων, και ως εκ τούτου θα μπορούσε να εξοικονομήσει ενέργεια και έκταση.

Αναφορές

- [1] L. Chua, “Memristor-the missing circuit element,” *IEEE Transactions on Circuit Theory*, vol. 18, no. 5, pp. 507–519, 1971, ISSN: 0018-9324. doi: [10.1109/TCT.1971.1083337](https://doi.org/10.1109/TCT.1971.1083337). [Online]. Available: [http://ieeexplore.ieee.org/document/1083337/](https://ieeexplore.ieee.org/document/1083337/) (visited on 11/01/2022).
- [2] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, “The missing memristor found,” *Nature*, vol. 453, no. 7191, pp. 80–83, May 1, 2008, ISSN: 0028-0836, 1476-4687. doi: [10.1038/nature06932](https://doi.org/10.1038/nature06932). [Online]. Available: [http://www.nature.com/articles/nature06932](https://www.nature.com/articles/nature06932) (visited on 11/01/2022).
- [3] N. Nithya and K. Paramasivam, “A comprehensive study on the characteristics, complex materials and applications of memristor,” in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India: IEEE, Mar. 2020, pp. 171–176, ISBN: 978-1-72815-196-0 978-1-72815-197-7. doi: [10.1109/ICACCS48705.2020.9074392](https://doi.org/10.1109/ICACCS48705.2020.9074392). [Online]. Available: <https://ieeexplore.ieee.org/document/9074392/> (visited on 11/08/2022).
- [4] S. Sahoo, G. Ramana Murthy, S. Ramesh, and G. Anitha, “Hybrid CMOS-memristor based operational transconductance amplifier for high frequency applications,” *Sustainable Energy Technologies and Assessments*, vol. 53, p. 102506, Oct. 2022, ISSN: 22131388. doi: [10.1016/j.seta.2022.102506](https://doi.org/10.1016/j.seta.2022.102506). [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2213138822005562> (visited on 11/02/2022).
- [5] O. Krestinskaya, K. N. Salama, and A. P. James, “Learning in memristive neural network architectures using analog backpropagation circuits,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 2, pp. 719–732, Feb. 2019, ISSN: 1549-8328, 1558-0806. doi: [10.1109/TCSI.2018.2866510](https://doi.org/10.1109/TCSI.2018.2866510). [Online]. Available: <https://ieeexplore.ieee.org/document/8468181/> (visited on 11/08/2022).
- [6] D. Bolek, Z. Kolka, V. Biolkovalova και Z. Bolek, ``Memristor models for spice simulation of extremely large memristive networks”, στο *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, IEEE, 2016, σσ. 389–392.
- [7] P. Kumar, K. Zhu, X. Gao, S.-D. Wang, M. Lanza, and C. S. Thakur, “Hybrid architecture based on two-dimensional memristor crossbar array and CMOS integrated circuit for edge computing,” *npj 2D Materials and Applications*, vol. 6, no. 1, p. 8, Jan. 21, 2022, ISSN: 2397-7132. doi: [10.1038/s41699-021-00284-3](https://doi.org/10.1038/s41699-021-00284-3). [Online]. Available: [http://www.nature.com/articles/s41699-021-00284-3](https://www.nature.com/articles/s41699-021-00284-3) (visited on 11/13/2022).