



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ
ΑΤΤΙΚΗΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΥΠΟΛΟΓΙΣΤΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΣΥΓΧΡΟΝΕΣ ΚΑΙ ΑΠΟΔΟΤΙΚΕΣ
ΜΕΘΟΔΟΙ ΨΗΦΙΑΚΗΣ ΕΠΕΞΕΡΓΑΣΙΑΣ
ΥΨΗΡΡΥΘΜΩΝ ΣΥΣΤΗΜΑΤΩΝ
ΟΠΤΙΚΩΝ ΕΠΙΚΟΙΝΩΝΙΩΝ

Σταύρος Λ. Δεληγιαννίδης

ΑΘΗΝΑ

Σεπτέμβριος 2023

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΣΥΓΧΡΟΝΕΣ ΚΑΙ ΑΠΟΔΟΤΙΚΕΣ ΜΕΘΟΔΟΙ ΨΗΦΙΑΚΗΣ ΕΠΕΞΕΡΓΑΣΙΑΣ ΥΨΗΡΡΥΘΜΩΝ ΣΥΣΤΗΜΑΤΩΝ ΟΠΤΙΚΩΝ ΕΠΙΚΟΙΝΩΝΙΩΝ

Σταύρος Λ. Δεληγιαννίδης

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:

Αντώνιος Μπόγρης, **Καθηγητής ΠΑΔΑ**

ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΕΠΙΤΡΟΠΗ

Αντώνιος Μπόγρης, **Καθηγητής ΠΑΔΑ**

Παναγιώτης Καρκαζής, **Αναπληρωτής Καθηγητής ΠΑΔΑ**

Σωτήριος Καραμπέτσος, **Αναπληρωτής Καθηγητής ΠΑΔΑ**

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Αντώνιος Μπόγρης, Καθηγητής ΠΑΔΑ

Παναγιώτης Καρκαζής, Αναπληρωτής Καθηγητής, ΠΑΔΑ

Σωτήριος Καραμπέτσος, Αναπληρωτής Καθηγητής ΠΑΔΑ

Θωμάς Καμαλάκης, Καθηγητής, Χαροκόπειο Πανεπιστήμιο

Χάρης Μεσαριτάκης, Αναπληρωτής Καθηγητής, Πανεπιστήμιο Αιγαίου

Έκτορας Νισταζάκης, Καθηγητής, ΕΚΠΑ

Περικλής Πετρόπουλος, Καθηγητής, University of Southampton

Ημερομηνία υποστήριξης : 8/9/2023

Μέλη Τριμελούς Συμβουλευτικής Επιτροπής συμπεριλαμβανομένου και του Επιβλέποντα

A/A	ΟΝΟΜΑ ΕΠΩΝΥΜΟ	ΒΑΘΜΙΔΑ/ΙΔΙΟΤΗΤΑ/ΤΜΗΜΑ	ΨΗΦΙΑΚΗ ΥΠΟΓΡΑΦΗ
1	Αντώνιος Μπόγρης	Καθηγητής / Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών	
2	Παναγιώτης Καρκαζής	Αναπληρωτής Καθηγητής / Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών	
3	Σωτήριος Καραμπέτσος	Αναπληρωτής Καθηγητής / Τμήμα Ηλεκτρολόγων & Ηλεκτρονικών Μηχανικών	

Μέλη Επταμελούς Εξεταστικής Επιτροπής

A/A	ΟΝΟΜΑ ΕΠΩΝΥΜΟ	ΒΑΘΜΙΔΑ/ΙΔΙΟΤΗΤΑ/ΤΜΗΜΑ/ΣΧΟΛΗ/ΠΑΝΕΠΙΣΤΗΜΙΟ
1	Αντώνιος Μπόγρης	Καθηγητής / Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών / Σχολή Μηχανικών / Πανεπιστήμιο Δυτικής Αττικής
2	Παναγιώτης Καρκαζής	Αναπληρωτής Καθηγητής / Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών / Σχολή Μηχανικών / Πανεπιστήμιο Δυτικής Αττικής
3	Σωτήριος Καραμπέτσος	Αναπληρωτής Καθηγητής / Τμήμα Ηλεκτρολόγων & Ηλεκτρονικών Μηχανικών / Σχολή Μηχανικών / Πανεπιστήμιο Δυτικής Αττικής
4	Θωμάς Καμαλάκης	Καθηγητής / Τμήμα Πληροφορικής και Τηλεματικής / Σχολή Ψηφιακής Τεχνολογίας / Χαροκόπειο Πανεπιστήμιο
5	Χάρης Μεσαριτάκης	Αναπληρωτής Καθηγητής / Τμήμα Μηχανικών Πληροφορικών και Επικοινωνιακών Συστημάτων / Πολυτεχνική Σχολή / Πανεπιστήμιο Αιγαίου
6	Έκτορας – Εμμανουήλ Νισταζάκης	Καθηγητής / Τμήμα Φυσικής / Σχολή Θετικών Επιστημών / Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών
7	Περικλής Πετρόπουλος	Professor of Optical Communications, Deputy Director of the Optoelectronics Research Centre (ORC) / University of Southampton

Η έγκριση της διδακτορικής διατριβής από το Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών του Πανεπιστημίου Δυτικής Αττικής δεν υποδηλοί αποδοχή των γνώμων του συγγραφέα
(Ν. 5343/32, Άρθρο 202).

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ

Ο κάτωθι υπογεγραμμένος Δεληγιαννίδης Σταύρος του Λαζάρου, υποψήφιος/α διδάκτορας του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών της Σχολής Μηχανικών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι:

«Είμαι συγγραφέας και δικαιούχος των πνευματικών δικαιωμάτων επί της διατριβής και δεν προσβάλλω τα πνευματικά δικαιώματα τρίτων. Για τη συγγραφή της διδακτορικής μου διατριβής δεν χρησιμοποίησα ολόκληρο ή μέρος έργου άλλου δημιουργού ή τις ιδέες και αντιλήψεις άλλου δημιουργού χωρίς να γίνεται αναφορά στην πηγή προέλευσης (βιβλίο, άρθρο από εφημερίδα ή περιοδικό, ιστοσελίδα κ.λπ.). Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του διδακτορικού διπλώματος μου».

Ο Δηλών

Δεληγιαννίδης Σταύρος



Περίληψη

Τα σύγχρονα συστήματα οπτικών επικοινωνιών παρουσιάζουν ραγδαία ανάπτυξη εξαιτίας της εμφάνισης αρχιτεκτονικών δικτύωσης υπολογιστικού νέφους/άκρου που σχετίζονται με σύγχρονες εφαρμογές και υπηρεσίες και καθορίζουν πώς θα εξελιχθεί το διαδίκτυο τις επόμενες δεκαετίες. Καθώς οι ρυθμοί μετάδοσης μεγαλώνουν διαρκώς πλησιάζοντας τα φυσικά όρια της μονότροπης οπτικής ίνας, ο στόχος της επίτευξης εξαιρετικά υψηλής χωρητικότητας ανά οπτική ίνα είναι επιτακτική ανάγκη και χωρίς ουσιαστική αύξηση της κατανάλωσης ενέργειας, αποτελεί σημαντική πρόκληση.

Στα συστήματα μετάδοσης μεγάλων αποστάσεων, όπου επικρατούν οι ψηφιακές σύμφωνες λύσεις, η βελτίωση της φασματικής απόδοσης παρεμποδίζεται κυρίως από τα μη γραμμικά φαινόμενα που αποδίδονται στο φαινόμενο Kerr. Στα συστήματα διαμόρφωσης έντασης/άμεσης ανίχνευσης (Intensity Modulation/Direct Detection, IM/DD) που προτιμώνται στις μικρές αποστάσεις η κύρια υποβάθμιση προέρχεται από το φαινόμενο εξασθένισης ισχύος, που προκαλείται από τη χρωματική διασπορά, από τους περιορισμούς του εύρους ζώνης και από τις μη-γραμμικότητες των πομποδεκτών.

Η εξέλιξη των αλγορίθμων μηχανικής μάθησης και των τεχνικών βαθιάς μάθησης δίνει τη δυνατότητα εφαρμογής τους σε ψηφιακούς επεξεργαστές τελευταίας τεχνολογίας, προκειμένου να επιλύσουν αποτελεσματικά, σύνθετα μη γραμμικά προβλήματα. Συνεπώς η υιοθέτησή τους για την αντιμετώπιση των μη-γραμμικών φαινομένων της οπτικής διάδοσης κερδίζει σταδιακά έδαφος.

Τα κύρια αποτελέσματα της διατριβής ήταν η αντιμετώπιση των μη γραμμικοτήτων της οπτικής ίνας μέσω της εφαρμογής επαναλαμβανόμενων νευρωνικών δικτύων (recurrent neural networks, RNN) και η προσομοίωση μη-γραμμικών φωτονικών συστημάτων με τη χρήση αλγορίθμων μηχανικής μάθησης. Ειδικότερα, η χρήση αμφίδρομων RNN επιφέρει τη σημαντική βελτίωση της απόδοσης των υψίρρυθμων συστημάτων οπτικών επικοινωνιών. Η διατριβή βασίστηκε σε αναλύσεις, προσομοιώσεις και πειράματα που πραγματοποιήθηκαν σε διάφορα σενάρια εφαρμογής, τόσο σε δίκτυα μεγάλων αποστάσεων (>1000 km) που χρησιμοποιούν σύμφωνο δέκτη, όσο και σε οπτικά δίκτυα μικρών αποστάσεων (< 120 km) που βασίζονται σε IM/DD συστήματα.

Οι σημαντικότερες θεματικές που αναλύθηκαν κατά τη διάρκεια εκπόνησης της συγκεκριμένης διατριβής είναι οι ακόλουθες:

- Παρουσιάστηκε για πρώτη φορά η χρήση νευρωνικού δικτύου μακριάς βραχυπρόθεσμης μνήμης (Long Short Term Memory Neural Network, LSTM) για την αντιστάθμιση της μη γραμμικότητας των οπτικών ινών σε ψηφιακά σύμφωνα συστήματα. Πραγματοποιήθηκαν αριθμητικές προσομοιώσεις διάδοσης στη ζώνη-C και τη ζώνη-O, διαμόρφωσης 16-QAM με πολυπλεξία πόλωσης, ενός ή πολλών καναλιών. Αναλύθηκε λεπτομερώς η επίδραση του αριθμού των κρυφών μονάδων και του μήκους της λέξης των συμβόλων. Μελετήθηκε εκτενώς η απόδοση και τα όρια λειτουργίας ενώ αναδείχθηκε η ανθεκτικότητα του προτεινόμενου ισοσταθμιστή σε διάφορα σενάρια μεταβολής των συνθηκών διάδοσης. Οι προσομοιώσεις απέδειξαν πως τα νευρωνικά δίκτυα τύπου LSTM μπορούν να είναι πολύ αποτελεσματικά ως ψηφιακοί επεξεργαστές σήματος στον δέκτη και την ίδια στιγμή είναι λιγότερο πολύπλοκα συγκρινόμενα με τους ισοσταθμιστές τύπου Digital Back Propagation (DBP), ειδικά στο σενάριο διάδοσης πολλαπλών καναλιών σε μεγάλες αποστάσεις (> 1000 km.).
- Μελετήθηκε η πολυπλοκότητα και η απόδοση μονάδων επεξεργασίας για την άρση των μη γραμμικών φαινομένων. Συγκρίθηκαν οι επικρατέστεροι τύποι RNN, bi-LSTM, bi-GRU και bi-Vanilla-RNN ως προς την πολυπλοκότητα και τις επιδόσεις τους και αποδείχθηκε πως όλοι οι τύποι διαθέτουν τη δυνατότητα να επιτύχουν μη γραμμική εξισορρόπηση ιδιαίτερα σε συστήματα μεγάλης διασποράς σε ψηφιακά σύμφωνα συστήματα, που μεταφέρουν σήματα διαμόρφωσης 16 και 32 QAM, με πολυπλεξία πόλωσης, ενός ή πολλών καναλιών. Οι προσομοιώσεις έδειξαν ότι τα τρία μοντέλα παρέχουν παρόμοια απόδοση αντιστάθμισης, επομένως, σε πραγματικά συστήματα, πρέπει να προτιμηθεί το απλούστερο σχήμα που βασίζεται στο μοντέλο Vanilla-RNN. Προτάθηκε η προσέγγιση πολλαπλών εισόδων πολλαπλών εξόδων (many to many) ως μια ιδιαίτερα αποδοτική μέθοδος ελαχιστοποίησης της πολυπλοκότητας. Πραγματοποιήθηκε σύγκριση με τους μη γραμμικούς ισοσταθμιστές τύπου Volterra και επιδείχθηκε η υπεροχή του Vanilla-RNN τόσο από την άποψη των επιδόσεων όσο και της πολυπλοκότητας.

- Με βάση ένα πειραματικό σύστημα διάδοσης με διαίρεση μήκους κύματος (CWDM) στη ζώνη-O, διαμόρφωσης έντασης και άμεσης ανίχνευσης (IM/DD) και ρυθμού $4 \times 50\text{-Gb/s}$, αποδείχθηκε η σημαντικά καλύτερη επίδοση ρυθμού σφάλματος (bit error rate, BER) του σχήματος Vanilla-RNN σε σχέση με τον συμβατικό ισοσταθμιστή ανάδρασης απόφασης (DFE) για διαμορφώσεις OOK και PAM4. Αναδείχθηκε η ικανότητα του ισοσταθμιστή Vanilla-RNN να αντισταθμίζει τόσο τις γραμμικές όσο και τις μη γραμμικές αλλοιώσεις που προκαλούνται από τον πομποδέκτη και τη μονότροπη ίνα. Ως απότοκο της δράσης του Vanilla-RNN ισοσταθμιστή, επετεύχθη οπτική διάδοση έως και 100 km και 75 km για τις OOK και PAM4 διαμορφώσεις αντίστοιχα. Επιπλέον, μέσω της σύγκρισης με άλλα σχήματα εξισορρόπησης, συμπεριλαμβανομένου του γραμμικού ισοσταθμιστή, του ισοσταθμιστή 3ης τάξης Volterra και του Volterra+DFE, αποδείχθηκε ότι ο ισοσταθμιστής Vanilla-RNN επιτυγχάνει την καλύτερη επίδοση BER παρουσιάζοντας επίσης χαμηλότερη πολυπλοκότητα υλοποίησης σε σύγκριση με σχήματα που βασίζονται στο Volterra.
- Προτάθηκε και εφαρμόστηκε η πολυκαναλική ισοστάθμιση WDM καναλιών προκειμένου να εξισορροπηθούν τα μη γραμμικά φαινόμενα μεταξύ των καναλιών, τροποποιώντας κατάλληλα τους επεξεργαστές bi-RNN των προηγούμενων εργασιών. Ο ισοσταθμιστής χρησιμοποιεί την αρχιτεκτονική επεξεργασίας πολλαπλών εισόδων πολλαπλών εξόδων (multi-input multi-output, MIMO) για την αποκωδικοποίηση των WDM καναλιών. Με αυτόν τον τρόπο, εκμεταλλεύεται τη χρήσιμη πληροφορία από τα γειτονικά κανάλια για την καλύτερη αναγνώριση και άρση των διακαναλικών αλληλεπιδράσεων. Αποδείχθηκε μέσω εκτεταμένων αριθμητικών προσομοιώσεων και πειραματικών αποτελεσμάτων ότι η προτεινόμενη προσέγγιση ξεπερνά τόσο την τυπική εξισορρόπηση πολλαπλών καναλιών με τη μορφή προσαρμοστικών ισοσταθμιστών, όσο και τους ισοσταθμιστές τύπου DBP αλλά και το μονοκαναλικό bi-RNN.
- Χρησιμοποιήθηκαν οι προηγούμενες δομές RNN για την εξομοίωση της δυναμικής συμπεριφοράς ενός πραγματικού οπτικού συστήματος διάδοσης, βασιζόμενου σε laser επιφανειακής εκπομπής, VCSEL ρυθμού 106.25 Gb/s και PAM4 διαμόρφωσης. Αποδείχθηκε πως τα RNN δίκτυα είναι ικανά να αναπαράγουν τα

δυναμικά χαρακτηριστικά του οπτικού συστήματος με ακρίβεια πρόβλεψης που πλησιάζει το 100%. Μελετήθηκε τέλος η ικανότητα εξομοίωσης περιοχών λειτουργίας του οπτικού συστήματος στις οποίες το διαθέσιμο νευρωνικό δίκτυο δεν έχει εκπαιδευτεί.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Οπτικές Επικοινωνίες, Ψηφιακή Επεξεργασία Σήματος και Μηχανική Μάθηση

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Ψηφιακά σύμφωνα συστήματα, διαμόρφωσης έντασης και άμεση ανίχνευση, μη-γραμμική επεξεργασία σήματος, διασπορά οπτικών ινών, επαναλαμβανόμενα νευρωνικά δίκτυα, βαθιά μάθηση

Abstract

Modern optical communication systems exhibit rapid growth due to the advent of edge-cloud networking architectures that fuel the future of 5G/B5G define how internet will evolve in the next decades. As transmission rates continue to grow approaching the physical limits of single-mode fiber, the goal of high capacity without substantially increasing power consumption is a significant challenge.

In long haul systems, where digital coherent solutions prevail, the improvement of spectral efficiency is mainly hindered by the nonlinear effects attributed to the Kerr effect. In intensity modulation/direct detection (IM/DD) systems the main degradation comes from power fading, caused by chromatic dispersion, bandwidth limitations and transceiver nonlinearities.

The evolution of machine learning algorithms and deep learning techniques makes it possible to apply them to state-of-the-art digital processors in order to efficiently solve complex nonlinear problems. Hence, machine learning becomes a significant candidate for deal with nonlinear transmission impairments in optical fibers.

The main results of this thesis were the treatment of optical fiber nonlinearities through the application of recurrent neural networks (RNN) and the emulation of photonic systems using machine learning algorithms.

The use of bidirectional RNNs brings about the significant improvement of the performance of high-speed optical communication systems. The thesis was based on analyses, simulations and experiments carried out in various application scenarios, both in long haul networks (>1000 km) using a coherent receiver and in short reach optical networks (< 120 km) based on IM/DD systems.

The most important topics analyzed during the preparation of this thesis are the following:

- The use of a Long Short Term Memory Neural Network (LSTM) to compensate for optical fiber nonlinearity in digital coherent systems was presented for the first time. Numerical simulations of C-band and O-band transmission systems for single

channel and multi-channel 16-QAM modulation format with polarization multiplexing were performed. The effect of the number of hidden units and the length of the word of symbols were analyzed in detail. The performance and operating limits were extensively studied while the robustness of the proposed equalizer in various scenarios of channel variability was highlighted. The simulations proved that LSTM-type neural networks can be very efficient as digital processors, providing superior performance and at the same time are less complex compared to Digital Back Propagation (DBP) equalizers, especially in the long-distance multi-channel scenario (> 1000 km).

- The complexity and performance of processing units for the removal of non-linear effects were studied. The dominant types of RNN, bi-LSTM, bi-GRU and bi-Vanilla-RNN were compared in terms of complexity and performance, and it was shown that all have the potential to achieve nonlinear compensation especially in digital coherent systems with no in-line dispersion compensation, which carry 16 and 32 QAM modulation signals, with polarization multiplexing, single or multi-channel. Simulations show that the three models provide similar compensation performance, so in real systems, the simpler scheme based on the Vanilla-RNN model should be preferred. The many-to-many approach was proposed as a particularly efficient method of minimizing complexity. A comparison with non-linear Volterra-type equalizers was performed and the superiority of Vanilla-RNN was demonstrated in terms of both performance and complexity.
- Based on an experimental 4×50 -Gb/s IM/DD O-band CWDM system, the significantly better BER performance was demonstrated of the Vanilla-RNN scheme versus the conventional decision feedback equalizer (DFE) for OOK and PAM4 configurations. The ability of the Vanilla-RNN equalizer to compensate for both linear and non-linear impairments induced by the transmitter and single-mode fiber was demonstrated. As a result, optical transmission of up to 100 km and 75 km for the OOK and PAM4 modulation formats respectively can be realized. Furthermore, through comparison with other equalization schemes including the linear equalizer, the 3rd-order Volterra equalizer and the Volterra+DFE, it was shown that the Vanilla-RNN equalizer achieves the best performance.

-
- Multi-channel equalization of WDM channels was proposed and implemented in order to compensate for the non-linear effects between channels by suitably modifying the bi-RNN processors of the previous works. The equalizer used the multi-input multi-output (MIMO) processing architecture to decode the WDM channels. In this way, it exploited useful information from neighboring channels to better identify and remove inter-channel crosstalk. It was demonstrated through extensive numerical simulations and experimental results that the proposed approach outperformed both standard multi-channel equalization in the form of adaptive equalizers, DBP-type equalizers, and single-channel bi-RNN.
 - The previous RNN structures were used to emulate the dynamic behavior of a real-life 106.25 Gb/s PAM4 optical transmission system based on a vertical cavity surface-emitting laser (VCSEL). It was shown that RNN networks are capable of reproducing the dynamic characteristics of the optical system with a prediction accuracy approaching 100%. Finally, the ability to emulate functional areas of the system in which the neural network has not been trained was studied.

SUBJECT AREA: Optical communications, Digital Signal Processing and Machine Learning

KEYWORDS: Digital coherent systems, Intensity-modulation and Direct-detection, nonlinear signal processing, optical fibre dispersion, recurrent neural networks, deep learning

Ευχαριστίες

Αρχικά θα ήθελα να εκφράσω την απέραντη εκτίμηση και ευγνωμοσύνη μου στον επιβλέποντα καθηγητή μου και φίλο κ. Αντώνη Μπόγρη για την άψογη καθοδήγηση του σε όλη την πορεία των διδακτορικών μου χρόνων. Ο κ. Αντώνης Μπόγρης, Καθηγητής του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών του Πανεπιστημίου Δυτικής Αττικής, είναι ο άνθρωπος που έχει αφήσει το μεγαλύτερο αποτύπωμα στην ερευνητική μου διαδρομή, καθώς η συνεργασία μας ξεκίνησε πολλά χρόνια πριν από την εκπόνηση αυτής της διατριβής. Αποτελεί υπόδειγμα επιστήμονα, καθηγητή και δοτικού ανθρώπου, όλα σε υπερθετικό βαθμό. Τον ευχαριστώ από τα βάθη της καρδιάς μου που ήταν πάντοτε διαθέσιμος να βοηθήσει αποτελεσματικά, κάνοντας με τη γνώση και τη διδακτική του ικανότητα ακόμα και τα πλέον δυσεπίλυτα προβλήματα να φαντάζονται εύκολα.

Ένα μεγάλο ευχαριστώ στον Αναπληρωτή Καθηγητή του Τμήματος Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων, κ. Χάρη Μεσαριτάκη, άλλοτε συμφοιτητή μου στο μεταπτυχιακό και από τότε φίλο, σήμερα επικεφαλής, μαζί με τον κ. Αντώνη Μπόγρη, της ερευνητικής μας ομάδας RNCP. Τον ευχαριστώ τόσο για την υποστήριξη του όσο και για την αξιοζήλευτη ικανότητά του να μεταδίδει την επιστημονική του γνώση και να εμπνέει.

Ευχαριστώ ιδιαίτερα τους συναδέλφους και φίλους Κώστα Σώζο, Γιάννη Τσιλικά, Δημήτρη Δερμάνη, Γιώργο Σαράντογλου, Μενέλαο Σκοντράνη και Άρη Τσιριγώτη για το εξαιρετικό κλίμα συνεργασίας και τις ατελείωτες εποικοδομητικές συζητήσεις που είχαμε σε ερευνητικά (και μη) θέματα.

Ευχαριστώ πολύ τους ερευνητές Kyle R. H. Bottrill, Yang Hong και βέβαια τον Καθηγητή κ. Περικλή Πετρόπουλο του Πανεπιστημίου Southampton για την άψογη συνεργασία στα απαιτητικά πειράματα της διατριβής, όπως επίσης τους ερευνητές της NVIDIA Δρ. Νίκο Αργύρη, Δρ. Στέφανο Δρη, Δρ. Δημήτρη Καλαβρουζιώτη και Δρ. Παρασκευά Μπακόπουλο.

Τέλος, θα ήθελα να πω ένα μεγάλο ευχαριστώ στους δικούς μου ανθρώπους και ιδιαίτερα στους γονείς μου Λάζαρο και Ευφροσύνη και την αδελφή μου Μαρία, για την ηθική συμπαράσταση, την ενθάρρυνση και τη βοήθεια που ακούραστα παρείχαν.

Κλείνοντας το μεγαλύτερο ευχαριστώ το οφείλω στη σύντροφό μου Κατερίνα, για την τεράστια κατανόηση, την αστείρευτη υπομονή και υποστήριξή της όλα αυτά τα χρόνια και βέβαια στο παιδί μας, τον Άρη, που γέμισε με χαρά και χαμόγελα τη ζωή μας. Τους υπόσχομαι πως τώρα που κλείνει αυτός ο κύκλος, θα προσπαθήσω να αναπληρώσω τα ... χαμένα ταξίδια και παιχνίδια!

Στον Άρη και στην Κατερίνα αφιερώνω τη διδακτορική μου διατριβή.

Σταύρος Δεληγιαννίδης

Δημοσιεύσεις

Κατά τη διάρκεια εκπόνησης της συγκεκριμένης διδακτορικής διατριβής, έχουν δημοσιευτεί δέκα (10) δημοσιεύσεις σε επιστημονικά περιοδικά με τη διαδικασία του peer review και εννιά (9) δημοσιεύσεις σε επιστημονικά συνέδρια. Παρακάτω παρατίθενται αναλυτικά οι σχετικές δημοσιεύσεις, ενώ σε παρένθεση αναγράφεται ο αριθμός των εργασιών που την αναφέρουν. Συνολικά, μέχρι την ημερομηνία ολοκλήρωσης της συγγραφής (30/8/2023) ο αριθμός των παραθέσεων - αναφορών στις παρακάτω εργασίες είναι 180.

Δημοσιεύσεις σε επιστημονικά περιοδικά που άπτονται της διατριβής

1. Deligiannidis, S., Bogris, A., Mesaritakis, C., & Kopsinis, Y. (2020). Compensation of fiber nonlinearities in digital coherent systems leveraging long short-term memory neural networks. *Journal of Lightwave Technology*, 38(21), 5991-5999. (Αναφορές : 79) [1]
2. Deligiannidis, S., Mesaritakis, C., & Bogris, A. (2021). "Performance and Complexity Analysis of Bi-Directional Recurrent Neural Network Models Versus Volterra Nonlinear Equalizers in Digital Coherent Systems" *Journal of Lightwave Technology*, 39(18), 5791-5798 (Αναφορές : 37) [2]
3. Hong, Y., Deligiannidis, S., Taengnoi, N., Bottrill, K. R., Thipparapu, N. K., Wang, Y., Sahu, J. K., Richardson D. J., Mesaritakis, C., Bogris, A. & Petropoulos, P. ML-assisted Equalization for 50-Gb/s/λ O-band CWDM Transmission over 100-km SMF, *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 4, 2022, doi:10.1109/JSTQE.2022.3155990, (Αναφορές: 8) [3]
4. K. Sozos, S. Deligiannidis, G. Sarantoglou, and C. Mesaritakis and A. Bogris, "Recurrent Neural Networks and Recurrent Optical Slicers as Equalizers in High Baud Rate Optical Transmission Systems", *Journal of Lightwave Technology (Invited paper)*, (early access, doi:10.1109/JLT.2023.3282999), (Αναφορές: 1) [4]
5. Deligiannidis, S., Bottrill, K. R., Sozos, K., Mesaritakis, C., Petropoulos, P. and Bogris, A., "Multichannel Nonlinear Equalization in Coherent WDM Systems based on Bi-directional

Recurrent Neural Networks”, *Journal of Lightwave Technology*, May 2023, (υποβλήθηκε - υπό εξέταση) [5]

Δημοσιεύσεις σε επιστημονικά συνέδρια που άπτονται της διατριβής

1. Deligiannidis, S., Mesaritakis, C., & Bogris, A. (2020, December). Performance and complexity evaluation of recurrent neural network models for fibre nonlinear equalization in digital coherent systems. In *2020 European Conference on Optical Communications (ECOC)* (pp. 1-4). IEEE. (Αναφορές: 7) [6]
2. Hong, Y., Deligiannidis, S., Taengnoi, N., Bottrill, K. R., Thipparapu, N. K., Wang, Y., Sahu, J. K., Richardson D. J., Mesaritakis, C., Bogris, A. & Petropoulos, P. (2021, May). Performance-enhanced Amplified O-band WDM Transmission using Machine Learning based Equalization. In *2021 Conference on Lasers and Electro-Optics (CLEO)* (pp. 1-2). IEEE (Αναφορές: 2) [7]
3. A. Bogris, K. Sozos, S. Deligiannidis, G. Sarantoglou, and C. Mesaritakis, “Machine Learning and Neuromorphic Computing Approaches for the mitigation of transmission impairments in high baud rate transmission systems,” in *2022 European Conference on Optical Communication (ECOC)*, Sep. 2022 (Αναφορές: 2) [8]
4. S. Deligiannidis, N. Argyris, S. Dris, D. Kalavrouziotis, P. Bakopoulos, C. Mesaritakis and A. Bogris, “Deep - Learning – based VCSEL transmitter emulator,” in *CLEO-Europe, 2023* [9]

Άλλες Δημοσιεύσεις σε επιστημονικά περιοδικά

1. K. Sozos, A. Bogris, P. Bienstman, G. Sarantoglou, S. Deligiannidis, and C. Mesaritakis, “High-speed photonic neuromorphic computing using recurrent optical spectrum slicing neural networks,” *Communications Engineering*, vol. 1, no. 1, p. 24, Oct. 2022, doi: 10.1038/s44172-022-00024-5. (Αναφορές: 8) [10]
2. Bogris, A., Mesaritakis, C., Deligiannidis, S., & Li, P. (2020). Fabry-Perot lasers as enablers for parallel reservoir computing. *IEEE Journal of Selected Topics in Quantum Electronics*, 27(2), 1-7. (Αναφορές: 27) [11]
3. Skontranis, M., Sarantoglou, G., Deligiannidis, S., Bogris, A., & Mesaritakis, C. (2021). Time-Multiplexed Spiking Convolutional Neural Network Based on VCSELs for Unsupervised Image Classification. *Applied Sciences*, 11(4), 1383 (Αναφορές: 6) [12]

-
4. K. Sozos, S. Deligiannidis, C. Mesaritakis and A. Bogris (2023), Self-Coherent Receiver Based on a Recurrent Optical Spectrum Slicing Neuromorphic Accelerator, *accepted for publication in Journal of Lightwave Technology*, vol.: 41, 1-9, 2023, 10.1109/JLT.2023.3235278 (Αναφορές: 1) [13]
 5. A. Tsirigotis, G. Sarantoglou, M. Skontranis, S. Deligiannidis, K. Sozos, G. Tsilikas, D. Dermanis, A. Bogris & C. Mesaritakis, C. “Unconventional Integrated Photonic Accelerators for High Throughput Convolutional Neural Networks,” *Intell. Comput.*, pp. 1–33, Jun. 2023 [14]

Άλλες Δημοσιεύσεις σε επιστημονικά συνέδρια

1. Skontranis, M., Sarantoglou, G., Deligiannidis, S., Bogris, A., & Mesaritakis, C. (2020, December). Unsupervised Image Classification Through Time-Multiplexed Photonic Multi-Layer Spiking Convolutional Neural Network. In *2020 European Conference on Optical Communications (ECOC)* (pp. 1-4). IEEE. [15]
2. K. Sozos, S. Deligiannidis, C. Mesaritakis, and A. Bogris, “Unconventional Computing based on Four Wave Mixing in Highly Nonlinear Media,” *CLEO-Europe, 2023* [16]
3. I. Tsilikas, S. Deligiannidis, A. Tsirigotis, G. N. Tsigaridas, and A. Bogris, “Neuromorphic Camera Assisted High-Flow Imaging Cytometry for Particle Classification,” in *CLEO-Europe, 2023*. [17]
4. I. Tsilikas, A. Tsirigotis, S. Deligiannidis, G. N. Tsigaridas, and A. Bogris, “Time-Stretched Imaging Flow Cytometry and Photonic Neuromorphic Processing for Particle Classification,” in *CLEO-Europe, 2023* [18]
5. A. Bogris, K. Sozos, G. Sarantoglou, S. Deligiannidis and C. Mesaritakis, “Neuromorphic computing by means of recurrent spectrum slicing for next generation high baud rate transmission systems” 2023 IEEE Photonics Society Summer Topicals Meeting Series (SUM), Sicily, Italy, 2023 [19]

Η εκπόνηση της διδακτορικής διατριβής υποστηρίχθηκε οικονομικά από το Ελληνικό Ίδρυμα Έρευνας & Καινοτομίας (ΕΛ.Ι.Δ.Ε.Κ), Αρ. Πρ. 2901: Next generation Optical communication systems in the O-band: A Key enabler for capacity enhancement in existing fibre links, (NOOK) Ε.Υ. Μπόγγρης Αντώνιος, Φορέας ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ.

Περιεχόμενα

Περίληψη	ii
Abstract	viii
Ευχαριστίες	xi
Δημοσιεύσεις	xiii
1 ^ο ΚΕΦΑΛΑΙΟ	1
Εισαγωγή – Επισκόπηση	1
1.1 Εισαγωγή	1
1.2 Συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων	4
1.3 Συστήματα οπτικών επικοινωνιών μικρών αποστάσεων	6
1.4 Μηχανική μάθηση στα συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων	8
1.5 Μηχανική μάθηση στα συστήματα οπτικών επικοινωνιών μικρών αποστάσεων	9
1.6 Συμβολή της παρούσας διατριβής	10
1.7 Δομή της διατριβής	11
2 ^ο ΚΕΦΑΛΑΙΟ	14
Συστήματα οπτικών επικοινωνιών	14
2.1 Εισαγωγή	14
2.2 Βασικά φαινόμενα διάδοσης του φωτός σε οπτική ίνα	15
2.2.1 Απώλειες στην ίνα	17
2.2.2 Διασπορά	18
2.2.2.1 Χρωματική Διασπορά	19
2.2.3 Φαινόμενα Πόλωσης	22
2.2.4 Αντιστάθμιση Γραμμικών Φαινομένων	24
2.3 Μη γραμμικά φαινόμενα	24
2.3.1 Τεχνικές αντιμετώπισης των μη γραμμικών φαινομένων	30
2.3.1.2 Ψηφιακή αντίστροφη διάδοση (Digital back-propagation - DBP)	30
2.3.1.2 Μη γραμμικός ισοσταθμιστής βασισμένος στη σειρά Volterra	33
2.3.1.3 Οπτική συζυγία φάσης (Optical Phase Conjugation -OPC)	34
2.3.1.4 Μη γραμμική αντιστάθμιση βασισμένη σε Διαταραχή (Perturbation-based Non Linear Compensation, PB-NLC)	36
2.3.1.5 Μη γραμμικός μετασχηματισμός Fourier	37
2.4 Συστήματα Οπτικών Επικοινωνιών Διαμόρφωσης Έντασης – Άμεσης Φώρασης	40
2.4.1 Επιβαρύνσεις στα IM/DD συστήματα	43

2.5 Σύμφωνα και μη Σύμφωνα Συστήματα Οπτικών Επικοινωνιών	48
2.5.1 Δέκτες	48
2.5.1.1 Ετερόδυνη ανίχνευση.....	52
2.5.1.2 Ομόδυνη ανίχνευση	53
2.5.1.3 Ομόδυνη ανίχνευση – 90° οπτικό υβρίδιο	54
2.5.2 Σχήματα διαμόρφωσης.....	57
2.5.3 Φασματική απόδοση	57
2.6 Ψηφιακή Επεξεργασία Σήματος σε Σύμφωνα Συστήματα	60
2.6.1 ADC	60
2.6.2 Προσαρμοστικός Ισοσταθμιστής (Adaptive Equalizer).....	61
2.6.3 Εξισορρόπηση στο πεδίο των συχνοτήτων (Frequency-Domain Equalization – FDE).....	66
2.6.3 Εκτίμηση Φάσης Φέρουσας (Carrier-Phase Estimation – CPE).....	67
2.6.4 Ανάκτηση συχνότητας ρολογιού	68
3 ^ο ΚΕΦΑΛΑΙΟ	69
3.1 Εισαγωγή	69
3.2 Είδη Μηχανικής Μάθησης.....	70
3.3 Βαθιά Μάθηση – Deep Learning	74
3.4 Προσωτροφοδοτούμενα Νευρωνικά Δίκτυα– Βασικές τεχνικές εκπαίδευσης.....	75
3.5 Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks – CNN)	83
3.6 Επαναλαμβανόμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks – RNN).....	85
3.6.1 Vanilla – RNN (VRNN)	88
3.6.2 Long-Short Term Memory Networks (LSTM)	89
3.6.3 Gated Recurrent Unit Networks (GRU).....	92
3.6.4 Αρχιτεκτονικές βασισμένες σε RNN	95
3.6.5 Τοπολογίες εκπαίδευσης μοντέλων βασισμένων σε RNN.....	100
4 ^ο ΚΕΦΑΛΑΙΟ	104
Ισοστάθμιση Βασισμένη σε Επαναλαμβανόμενα Νευρωνικά Δίκτυα στα Συστήματα Οπτικών Επικοινωνιών Μεγάλων Αποστάσεων.....	104
4.1 Εισαγωγή	104
4.2 Αντιστάθμιση μη γραμμικότητας οπτικών ινών σε ψηφιακά σύμφωνα συστήματα που αξιοποιούν νευρωνικά δίκτυα (LSTM).....	106
4.2.1 Ανάλυση Μοντέλου Οπτικής Διάδοσης	108
4.2.1 Βασικά Χαρακτηριστικά του Μοντέλου Bi- LSTM	110
4.2.3 Παρουσίαση Αποτελεσμάτων για διάδοση σε ένα ή πολλά κανάλια	115
4.2.3 Ανθεκτικότητα της εκπαιδευμένου μοντέλου στις διακυμάνσεις του καναλιού	122

4.2.4 Ανάλυση πολυπλοκότητας bi-LSTM.....	125
4.2.5 Συμπέρασμα	129
4.3 Σύγκριση πολυπλοκότητας και απόδοσης διαφόρων επαναλαμβανόμενων νευρωνικών δικτύων για την εξισορρόπηση των μη γραμμικών φαινομένων της ίνας σε ψηφιακά σύμφωνα συστήματα	130
4.3.1 Εισαγωγή.....	130
4.3.2 Μοντελοποίηση συστήματος	131
4.3.2 Σύγκριση απόδοσης μοντέλων bi-RNN	138
4.3.3 Σύγκριση ισοσταθμιστών βασισμένων σε bi-RNN και Volterra	143
4.3.4 Ανάλυση υπολογιστικής πολυπλοκότητας.....	147
4.3.5 Συμπέρασμα	151
4.4 Πολυκαναλική ανίχνευση βασισμένη σε αμφίδρομα επαναλαμβανόμενα νευρωνικά δίκτυα για την εξομάλυνση διακαναλικών μη γραμμικών φαινομένων σε ψηφιακά σύμφωνα συστήματα	153
4.4.1 Εισαγωγή.....	153
4.4.2 Προτεινόμενο σύστημα πολυκαναλικής ισοστάθμισης	154
4.4.3 Παρουσίαση αποτελεσμάτων εξομίωσης της πολυκαναλικής επεξεργασίας	162
4.4.3.2 Εξάρτηση του bi-RNN από τον αριθμό των κρυφών μονάδων και των εποχών εκπαίδευση	168
4.4.4 Πειραματική επιβεβαίωση αποτελεσμάτων πολυκαναλικής επεξεργασίας βασισμένης σε RNN	169
4.4.4.1 Πειραματική διάταξη	169
4.4.4.2 Ψηφιακή Επεξεργασία Σήματος	171
4.4.4.3 Πολυκαναλική και μονακαναλική bi-RNN ισοστάθμιση έναντι DBP	174
4.4.5 Ανάλυση Υπολογιστικής Πολυπλοκότητας.....	178
4.5 Συμπεράσματα	186
5° ΚΕΦΑΛΑΙΟ	187
Ισοστάθμιση Βασισμένη σε Επαναλαμβανόμενα Νευρωνικά Δίκτυα στα Συστήματα Οπτικών Επικοινωνιών Μικρών Αποστάσεων	187
5.1 Εισαγωγή	187
5.2 Ισοστάθμιση υποβοηθούμενη από bi-RNN για CWDM 100km διάδοση 50-Gb/s/λ στη ζώνη O..	188
5.2.1 Εισαγωγή.....	188
5.2.2 Αρχή λειτουργίας του προτεινόμενου συστήματος εξισορρόπησης.....	190
5.2.3 Πειραματική διάταξη	193
5.2.4 Πειραματικά αποτελέσματα για την CWDM διάδοση στη ζώνη O	199
5.2.4.1 Διάδοση B2B	199
5.2.4.2 Διάδοση 100 km για διαμόρφωση OOK και 75 km για διαμόρφωση PAM4.....	200
5.2.5 Σύγκριση επιδόσεων μεταξύ διαφόρων σχημάτων ισοστάθμισης.....	204

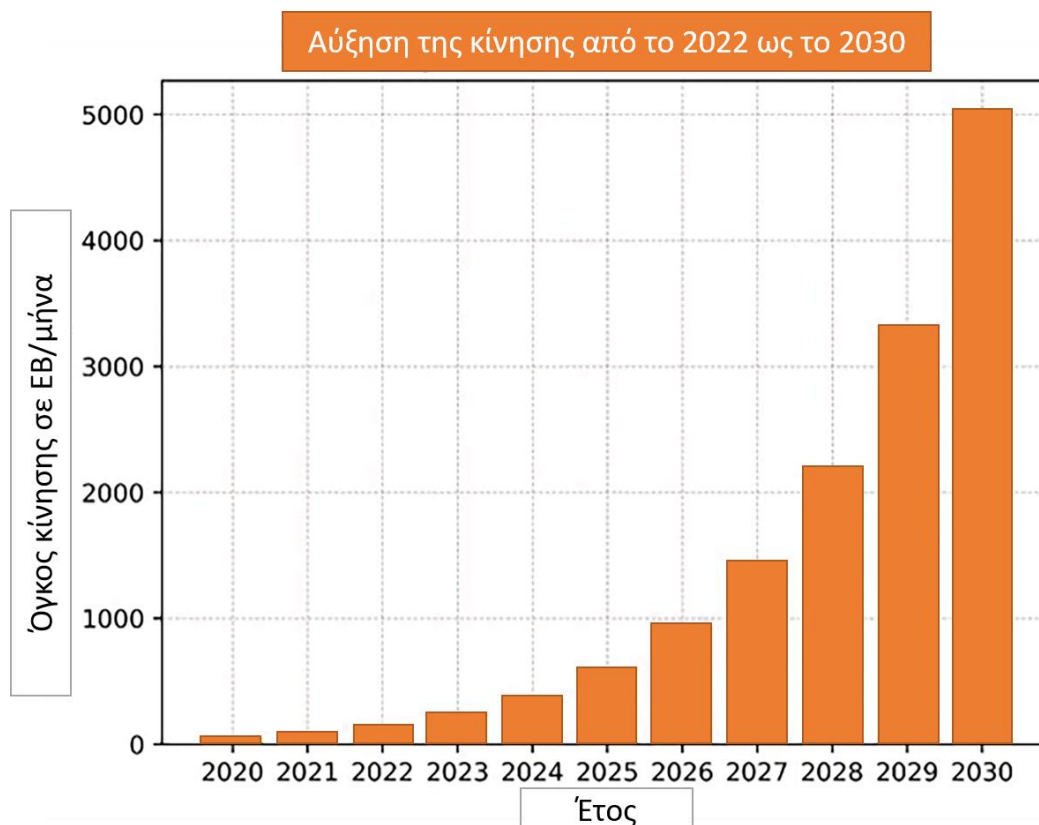
5.2.5.1 Σύγκριση επιδόσεων BER.....	205
5.3 Συμπεράσματα	212
6° ΚΕΦΑΛΑΙΟ	213
Προσομοίωση Οπτικών Συστημάτων με τη χρήση Επαναλαμβανόμενων Νευρωνικών Δικτύων	213
6.1 Εισαγωγή	213
6.2 Πειραματική Διάταξη – Αρχιτεκτονική RNN.....	215
6.3 Προοδευτική Εκπαίδευση	218
6.4 Συμπεράσματα	224
7° ΚΕΦΑΛΑΙΟ	225
Σύνοψη.....	225
7.1 Συνεισφορά της Διατριβής στις εφαρμογές των οπτικών επικοινωνιών.....	225
7.2 Συμπεράσματα και μελλοντικά βήματα.....	229
ΠΑΡΑΡΤΗΜΑ.....	231
Παράλληλη Επεξεργασία.....	231
Π.1 Εισαγωγή στην Παράλληλη Επεξεργασία – GPGPU	231
Π.2 Χρήση Προγραμματισμού GPU στην προσομοίωση της Οπτικής Διάδοσης.....	234
Π.3 Εκπαίδευση σύνθετων νευρωνικών δικτύων με τη χρήση GPU	236
Αναφορές	239
Ακρωνύμια.....	260

1° ΚΕΦΑΛΑΙΟ

Εισαγωγή – Επισκόπηση

1.1 Εισαγωγή

Τα τηλεπικοινωνιακά δίκτυα αποτελούν θεμέλιο για τη σύγχρονη κοινωνία, επιτρέποντας την ανταλλαγή πληροφοριών και την επικοινωνία μεταξύ ανθρώπων και συσκευών σε παγκόσμιο επίπεδο. Με την εξέλιξη της τεχνολογίας και την αυξημένη ζήτηση για υψηλής

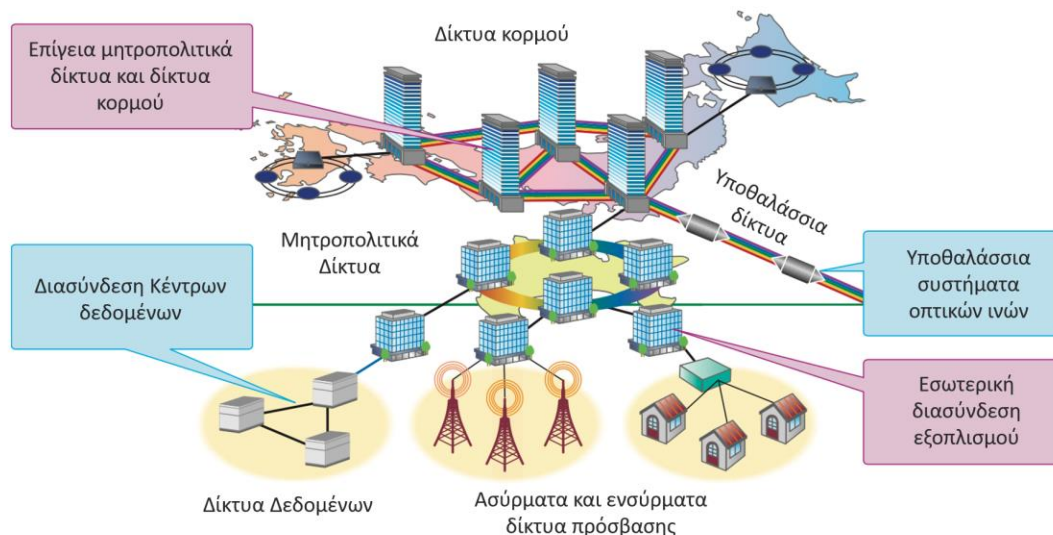


Σχήμα 1.1: Κίνηση δεδομένων από το 2020 ως το 2030 [20]

ποιότητας επικοινωνία, έχουν αναπτυχθεί σύγχρονα τηλεπικοινωνιακά δίκτυα που επιτρέπουν την αποτελεσματική μεταφορά δεδομένων.

Ο όγκος της κίνησης των δεδομένων σε παγκόσμια κλίμακα παρουσιάζει εκθετική αύξηση τα τελευταία χρόνια και η τάση αυτή φαίνεται πως θα διατηρηθεί τα αμέσως επόμενα, όπως φαίνεται στο σχήμα 1.1 [20], με γεγονότα όπως η πρόσφατη πανδημία του COVID-19 να την επιτείνουν.

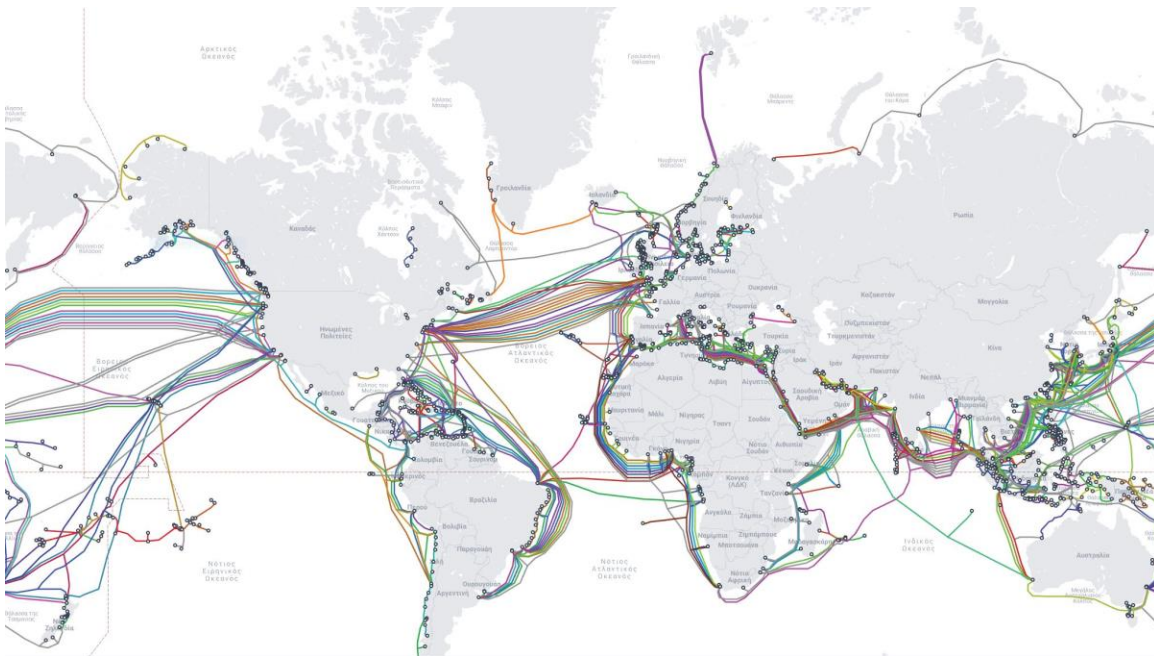
Ένα από τα σημαντικότερα επιτεύγματα στον τομέα των τηλεπικοινωνιών είναι τα δίκτυα οπτικών επικοινωνιών. Οι οπτικές ίνες αποτελούν τη βασική τεχνολογία που χρησιμοποιείται σε αυτά τα δίκτυα, επιτρέποντας τη μεταφορά μεγάλου όγκου πληροφοριών με υψηλή ταχύτητα. Η ευρεία χρήση των δικτύων οπτικών επικοινωνιών



Σχήμα 1.2: Πεδία εφαρμογής των συστημάτων οπτικών επικοινωνιών [21]

έχει ανοίξει τον δρόμο για την ανάπτυξη πολλαπλών πολυμεσικών εφαρμογών, όπως είναι οι τηλεδιασκέψεις, οι τηλεματικές εφαρμογές, η τηλεργασία, η ψυχαγωγία, οι εφαρμογές εικονικής και επαυξημένης πραγματικότητα και άλλες. Η εμβληματική ευρυζωνικότητα (> 100 THz ανά οπτική ίνα) που παρέχουν οι οπτικές ίνες, σε συνδυασμό με την ανθεκτικότητά τους στις ηλεκτρομαγνητικές παρεμβολές και άλλες ανεπιθύμητες επιδράσεις, τις καθιστούν ιδανικές για τη μεταφορά μεγάλου όγκου δεδομένων.

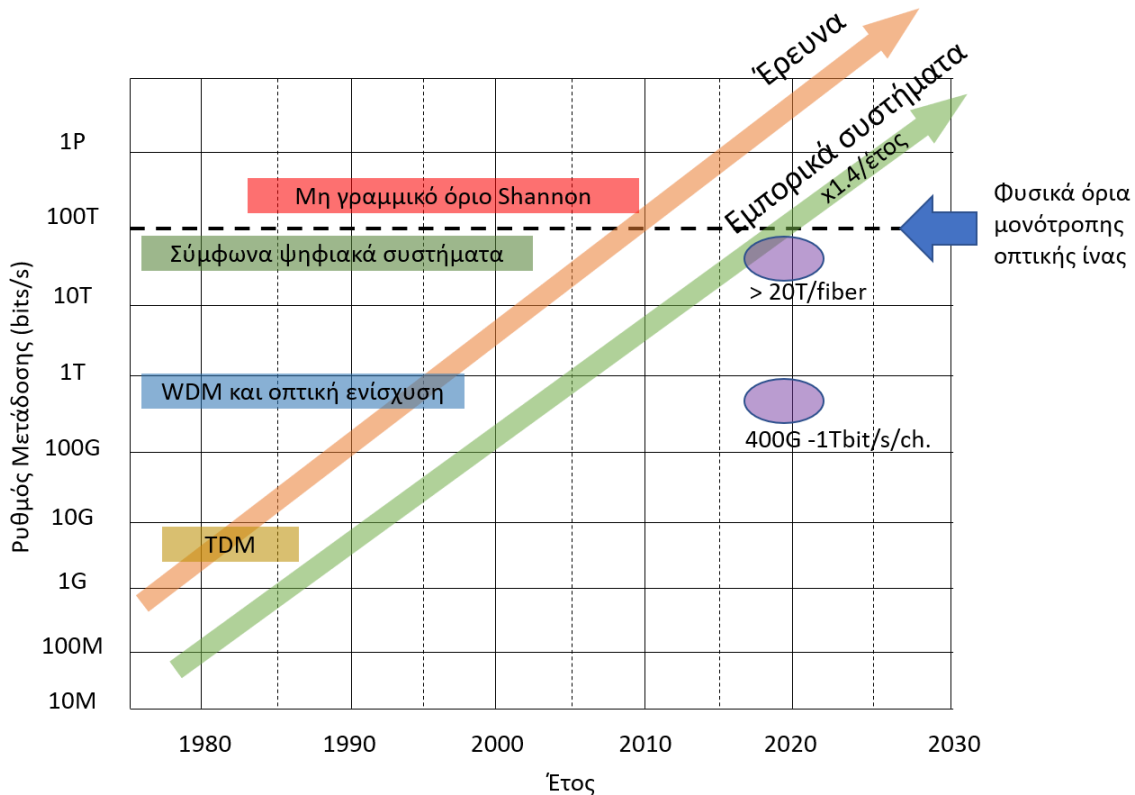
Οι οπτικές επικοινωνίες καλύπτουν ένα μεγάλο φάσμα εφαρμογών που εκτείνεται από την εσωτερική δικτύωση εγκαταστάσεων, τα δίκτυα διασύνδεσης κέντρων δεδομένων (Data Center Interconnects, DCI), ως τα μητροπολιτικά δίκτυα και τα υποθαλάσσια δίκτυα οπτικών ινών όπως φαίνεται στο σχήμα 1.2 [21]. Το συνολικό δε μήκος του παγκόσμιου υποθαλάσσιου δικτύου οπτικών ινών που ξεπερνά σήμερα τα 400 εκατομμύρια km, (σχήμα 1.3 [22]), έχει επιφέρει σημαντική αύξηση στον ρυθμό μετάδοσης δεδομένων, κατά πολλαπλές τάξεις μεγέθους, αξιοποιώντας το τεράστιο εύρος ζώνης της οπτικής ίνας μέσω της παράλληλης λειτουργίας πολλαπλών καναλιών και της ενίσχυσης του σήματος με οπτικούς ενισχυτές.



Σχήμα 1.3: Χάρτης παγκόσμιου υποθαλάσσιου δικτύου οπτικών ινών [22]

Τα σύγχρονα συστήματα οπτικών επικοινωνιών παρουσιάζουν ραγδαία ανάπτυξη εξαιτίας της εμφάνισης αρχιτεκτονικών δικτύωσης υπολογιστικού νέφους/άκρου που τροφοδοτούν το μέλλον του 5G/B5G και καθορίζουν πώς θα εξελιχθεί το διαδίκτυο τις επόμενες δεκαετίες (σχήμα 1.4 [21]). Οι διασυνδέσεις αυτές απαιτούν οπτικούς πομποδέκτες υψηλής ταχύτητας και χαμηλής κατανάλωσης και η απαίτηση αυτή αντικατοπτρίζεται στο κίνητρο της βιομηχανίας να τυποποιήσει άμεσα και να λανσάρει την τεχνολογία 800G-1.6T στην αγορά [23]. Αυτή η τάση επηρεάζει επίσης τις απαιτήσεις στα συστήματα οπτικής μετάδοσης μεσαίων και μεγάλων αποστάσεων όπου ο κύριος

στόχος είναι η πλήρης εκμετάλλευση της χωρητικότητας των εγκατεστημένων οπτικών ινών. Τόσο στα συστήματα μικρών αποστάσεων όσο και σε αυτά των μεγάλων αποστάσεων, ο στόχος της εξαιρετικά υψηλής χωρητικότητας πρέπει να εξυπηρετηθεί χωρίς ουσιαστική αύξηση της κατανάλωσης ενέργεια. Το αποτύπωμα άνθρακα των τηλεπικοινωνιών μεγαθύνεται συνεχώς τα τελευταία δέκα χρόνια και οι πράσινες τεχνολογίες είναι απαραίτητες για την αύξηση χωρητικότητας με βιώσιμο τρόπο.



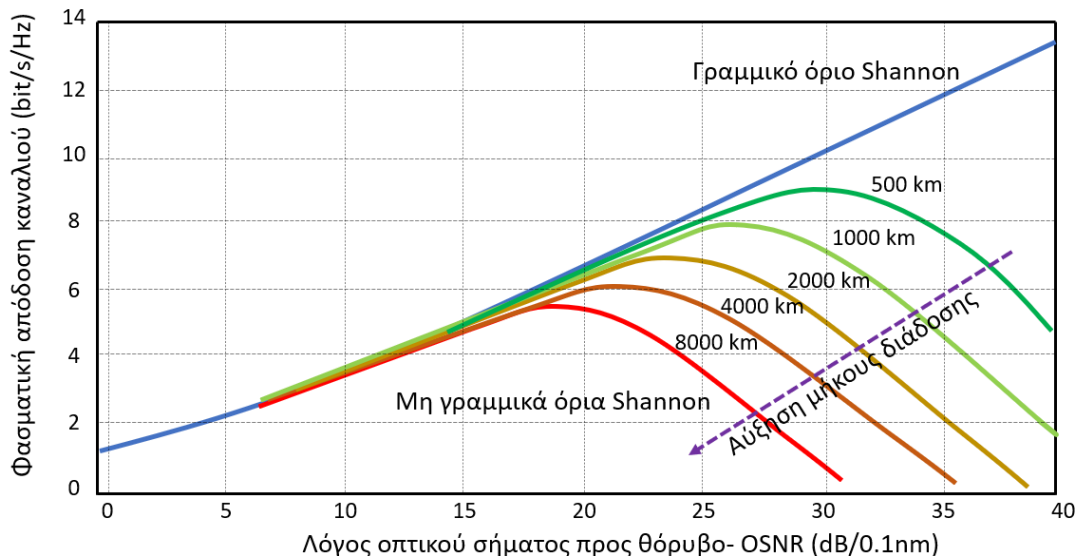
Σχήμα 1.4: Εξέλιξη των οπτικών συστημάτων υψηλής χωρητικότητας [21]

1.2 Συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων

Ο όρος συστήματα επικοινωνιών μεγάλων αποστάσεων αναφέρεται στις οπτικές ζεύξεις που καλύπτουν αποστάσεις μεγαλύτερες από 100 km και στις οποίες επικρατούν οι λύσεις των σύμφωνων (coherent) συστημάτων πομποδεκτών. Τα συστήματα αυτά αντιμετωπίζουν σημαντικές προκλήσεις ως αποτέλεσμα της συνεχούς εξέλιξης και αναβάθμισης των μητροπολιτικών δικτύων και των δικτύων πρόσβασης, εξαιτίας της

αυξανόμενης ζήτησης απαιτητικών ευρυζωνικών εφαρμογών όπως το 5G και το παθητικό οπτικό δίκτυο (passive optical network, PON).

Προκειμένου να ικανοποιηθούν οι διαρκώς αυξανόμενες απαιτήσεις σε χωρητικότητα, χρησιμοποιούνται προηγμένα σχήματα διαμόρφωσης όπως η ορθογώνια διαμόρφωση πλάτους (Quadrature Amplitude Modulation - QAM) τάξης $M \geq 16$ και οι υψηλοί ρυθμοί μετάδοσης συμβόλων που ξεπερνούν τα 64 Gbaud, ενώ υιοθετείται η χρήση οπτικών καναλιών πολύ πυκνής παράθεσης και εξαιρετικά υψηλού εύρους ζώνης, σε συστήματα πολυπλεξίας διαίρεσης μήκους κύματος (Wavelength Division Multiplexing – WDM). Η εκθετική αύξηση των απαιτήσεων σε χωρητικότητα υποχρεώνει την ερευνητική κοινότητα να εξετάσει δύο κύριες προσεγγίσεις: την πολυπλεξία χωρικής διαίρεσης [24] και την επέκταση της μετάδοσης σε νέες ζώνες συχνοτήτων, συμπεριλαμβανομένων των ζωνών O, E και S [25].



Σχήμα 1.5: Γραμμικό και μη γραμμικό όριο Shannon

Καθώς οι ρυθμοί μετάδοσης μεγαλώνουν διαρκώς πλησιάζοντας τα φυσικά όρια χωρητικότητας της μονότροπης ίνας, (σχήμα 1.4), η δοκιμασμένη λύση της αύξησης του λόγου του οπτικού σήματος προς το θόρυβο (Optical Signal to Noise Ratio- OSNR) δε

μπορεί να μεγεθύνει πλέον τη φασματική απόδοση του καναλιού καθώς εμποδίζεται από τα μη γραμμικά όρια Shannon [26], [27], όπως φαίνεται χαρακτηριστικά στο σχήμα 1.5.

Μία από τις μεγαλύτερες προκλήσεις έγκειται στον τρόπο αντιμετώπισης του αντίκτυπου των μη γραμμικοτήτων που οφείλονται στο φαινόμενο Kerr [28], και της αλληλεπίδρασης τους με τον ενισχυμένο θόρυβο αυθόρμητης εκπομπής. Εμφανίζονται δε με τη μορφή της αυτό-διαμόρφωση φάσης (Self Phase Modulation – SPM), της ετερο-διαμόρφωση φάσης (Cross Phase Modulation – SPM) και του φαινομένου της μίξης τεσσάρων κυμάτων (Four Wave Mixing – FWM) μεταξύ των στενά τοποθετημένων μηκών κύματος και των υποφερουσών που θα δούμε και στο κεφάλαιο 2 [28].

Στο παρελθόν έχουν προταθεί πολλές τεχνικές για την εξισορρόπηση των μη γραμμικών φαινομένων, όπως η οπτική συζυγία φάσης (optical phase conjugation) [29], η ψηφιακή αντίστροφη διάδοση (Digital back-propagation - DBP) [30], ο μη γραμμικός μετασχηματισμός Fourier [31] και η συνάρτηση μεταφοράς της αντίστροφης σειράς Volterra [32]. Όμως τόσο η αυξημένη πολυπλοκότητα υλοποίησης όσο και οι ιδιαίτερες απαιτήσεις στις συνθήκες λειτουργίας, δεν έχουν επιτρέψει την χρήση τους σε εμπορικά σύμφωνα συστήματα.

1.3 Συστήματα οπτικών επικοινωνιών μικρών αποστάσεων

Ο όρος συστήματα οπτικών επικοινωνιών μικρών αποστάσεων αναφέρεται σε ένα ευρύ φάσμα οπτικών ζεύξεων, οι οποίες καλύπτουν αποστάσεις μικρότερες από 100-120 km. Τα βασικά χαρακτηριστικά των οπτικών ζεύξεων μικρών αποστάσεων είναι η απλότητα όσον αφορά τα οπτικά στοιχεία και τα σχήματα διαμόρφωσης, οι απαιτήσεις για χαμηλή κατανάλωση και το χαμηλό τους κόστος.

Παρά την κυριαρχία των σύμφωνων (coherent) τεχνολογιών σε δίκτυα οπτικών επικοινωνιών μεγάλων αποστάσεων και μητροπολιτικών δικτύων, τα συστήματα διαμόρφωσης έντασης - άμεσης φώρασης (Intensity Modulation/Direct detection, IM/DD) έχουν επικρατήσει σε εφαρμογές μικρής εμβέλειας, όπως στα κέντρα δεδομένων και στα τοπικά δίκτυα (LAN) λόγω της πολύ καλής απόδοσης σε σχέση με το κόστος και τις

απαιτήσεις τους. Παρά την περιορισμένη φασματική απόδοση και χωρητικότητα είναι εν γένει πολύ πιο οικονομικά από τα σύμφωνα συστήματα μεγάλων αποστάσεων.

Οι πομποδέκτες IM/DD αποτελούν τις τρέχουσες τυπικές οπτικές μονάδες για επικοινωνίες μικρών αποστάσεων. Μπορούν να αποδώσουν αποτελεσματικά ρυθμούς 100 Gbps (100G) για οπτικές ζεύξεις έως και 80 km χρησιμοποιώντας διαμόρφωση on-off keying (OOK), 400 Gbps για ζεύξεις έως και 40 km (400G) και πρόσφατα 800 Gbps (800G) για ζεύξεις έως και 10 km χρησιμοποιώντας διαμόρφωση πλάτους (PAM-4) [33]. Σε αυτούς τους ρυθμούς και αποστάσεις η χρήση της ζώνης-0 είναι σχεδόν μονόδρομος εξαιτίας της χαμηλότερης χρωματική διασποράς.

Ωστόσο, τα συστήματα IM/DD είναι ευαίσθητα στη μη γραμμική υποβάθμιση του σήματος που προκαλείται από τη χρωματική διασπορά και τον τετραγωνικό κανόνα του φωτοδέκτη και τις μη γραμμικότητες των οπτικών πομποδεκτών [34]. Αυτές οι μη γραμμικές επιβαρύνσεις εισάγουν παρεμβολές μεταξύ των συμβόλων και περιορίζουν τους επιτεύξιμους ρυθμούς δεδομένων. Οι διαλείψεις ισχύος που προκαλούνται από τη χρωματική διασπορά λόγω του τετραγωνικού νόμου της φωτοδιόδου είναι η εγγενής αδυναμία που εμποδίζει την επεκτασιμότητα των συστημάτων IM/DD πέρα από τα σενάρια 800G-10km.

Όπως και στην περίπτωση των συστημάτων μεγάλων αποστάσεων, τα μη γραμμικά όρια Shannon (σχήμα 1.5) δεν επιτρέπουν την επίλυση του προβλήματος με την αύξηση του OSNR. Η εξισορρόπηση των μη γραμμικών φαινομένων είναι απαραίτητη για την επαναφορά του συστήματος σε συνθήκες γραμμικής λειτουργίας που επιτρέπουν την περαιτέρω αύξηση της φασματικής απόδοσης και χωρητικότητας. Οι λύσεις που έχουν προταθεί και αφορούν γραμμικούς αλγόριθμους όπως είναι ο ισοσταθμιστής εμπροσθοδιάδοσης (Feed Forward Equalizer, FFE) και ο ισοσταθμιστής ανατροφοδότησης απόφασης (Decision Feedback Equalizer, DFE) [35], αδυνατούν να καταπολεμήσουν επαρκώς τα προηγούμενα μη γραμμικά φαινόμενα, ενώ οι μη γραμμικοί αλγόριθμοι όπως όπως ο ισοσταθμιστής τύπου Volterra κ.α απαιτούν ιδιαίτερα πολύπλοκη ψηφιακή επεξεργασία που πρακτικά ακυρώνει τα πλεονεκτήματα των IM/DD συστημάτων, δηλαδή την απλότητα και την οικονομία.

1.4 Μηχανική μάθηση στα συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων

Τα τελευταία χρόνια, υπάρχει αυξανόμενο ενδιαφέρον για τη διερεύνηση των δυνατοτήτων των τεχνικών μηχανικής μάθησης για τον μετρίασμό των ανεπιθύμητων φαινομένων της διάδοσης στα συστήματα οπτικών επικοινωνιών [36]. Η μηχανική μάθηση προσφέρει διάφορες επιλογές για τη βελτίωση της απόδοσης των συστημάτων οπτικών επικοινωνιών μεγάλων αποστάσεων που επηρεάζονται από τα μη γραμμικά φαινόμενα της διάδοσης.

Από τις διάφορες προσεγγίσεις που έχουν προταθεί, η επικρατέστερη, παρόμοια αυτών που ακολουθούνται εδώ και πολλές δεκαετίες με τη χρήση συμβατικών ισοσταθμιστών, είναι αυτή που βασίζεται στα δεδομένα που προκύπτουν μετά την οπτική διάδοση και την επεξεργασία τους με τη βοήθεια αλγορίθμων βαθιάς μάθησης. Η ψηφιακή επεξεργασία αναφέρεται στην εκπαίδευση ενός νευρωνικού δικτύου είτε με τη μέθοδο της ταξινόμησης (Classification) είτε με τη αυτή της παλινδρόμησης (Regression) [37]. Έχουν προταθεί διάφορα μοντέλα και έχει δειχθεί η αποτελεσματικότητά τους στην εξισορρόπηση των μη γραμμικοτήτων όπως τα προσωτροφοδοτούμενα τεχνητά νευρωνικά δίκτυα (feed-forward artificial neural networks -FNN) και το perceptron πολλαπλών στρωμάτων [38], [39]. Στην εργασία [40] οι συγγραφείς παρουσίασαν με εκτενή ανάλυση των επιδόσεων και της πολυπλοκότητας διαφόρων επεξεργασιών βαθιάς μάθησης ενώ στην εργασία [41] ανέδειξαν τους σημαντικότερους κινδύνους της ακατάλληλης χρήσης των μοντέλων μηχανικής μάθησης.

Ένα από τα πιο ελπιδοφόρα και πολλά υποσχόμενα μοντέλα νευρωνικά δίκτυα για την ισοστάθμιση των μη γραμμικών φαινομένων της διάδοσης στην οπτική ίνα αλλά και του περιορισμένου εύρους ζώνης των σύμφωνων οπτικών συστημάτων, είναι αυτό των επαναλαμβανόμενων νευρωνικών δικτύων (recurrent neural networks - RNN), ιδιαίτερα στην αμφίδρομη εκδοχή τους [42], με τη χρήση μονάδων-πυλών [43]. Και αυτό διότι μπορούν να μιμηθούν και να αντιστρέψουν τις παραμορφώσεις ενός μη γραμμικού καναλιού που εμφανίζει μνήμη, επιτυγχάνοντας την ίδια απόδοση με την εκτίμηση ακολουθίας μέγιστης ομοιομορφίας (Maximum Likelihood Sequence Estimation - MLSE) [44].

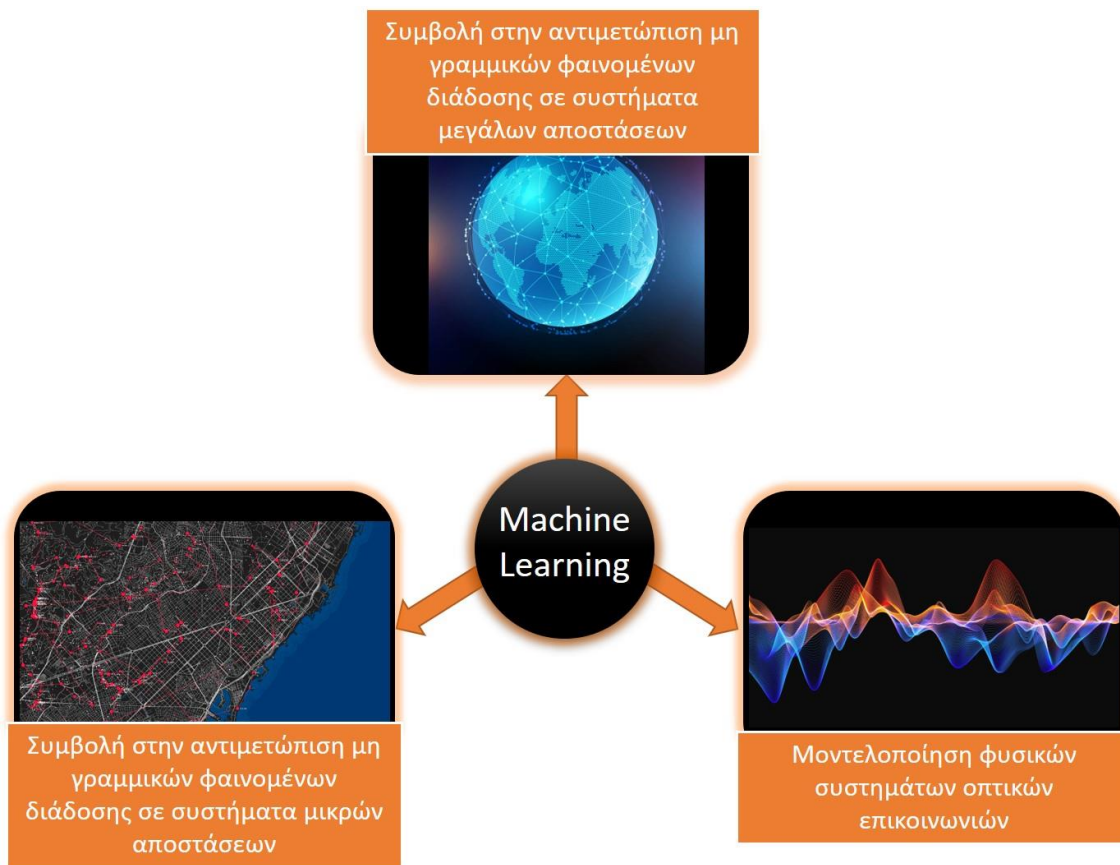
1.5 Μηχανική μάθηση στα συστήματα οπτικών επικοινωνιών μικρών αποστάσεων

Στην προηγούμενη παράγραφο παρουσιάστηκε η πρόταση των μοντέλων μηχανικής μάθησης στη διαχείριση του κύριου περιοριστικού παράγοντα των συστημάτων μετάδοσης μεγάλων αποστάσεων, δηλαδή των μη γραμμικοτήτων της οπτικής ίνας.

Εκτός από τα δίκτυα μεγάλων αποστάσεων, οι λύσεις που προσφέρουν τα δίκτυα βαθιάς μάθησης έχουν προσελκύσει το ενδιαφέρον της κοινότητας των οπτικών επικοινωνιών τα τελευταία δέκα χρόνια για εφαρμογή τους σε συστήματα μικρών αποστάσεων. Διάφορα μοντέλα όπως δίκτυα autoencoders [45] και βαθιά νευρωνικά δίκτυα [46] έχουν δοκιμαστεί σε IM/DD συστήματα και ο κύριος στόχος τους είναι να μειώσουν το περιορισμένο εύρος ζώνης των πομποδεκτών [47], [48], τη χρωματική διασπορά που προκαλεί το φαινόμενο των διαλείψεων ισχύος και σε ορισμένες περιπτώσεις τις μη γραμμικότητες που προέρχονται κυρίως από τις διάφορες συσκευές του οπτικού συστήματος. Άλλες εργασίες αξιοποιούν τα RNN για την παρακολούθηση της ποιότητας του καναλιού [49] ή για τη μοντελοποίηση μη γραμμικών οπτικών συστημάτων [50]. Στην εργασία [51] οι συγγραφείς παρουσίασαν τα RNN μακριάς βραχυπρόθεσμης μνήμης (Long Short Term Memory, LSTM) για την αντιστάθμιση των μη γραμμικών φαινομένων σε IM/DD συστήματα μικρής απόστασης.

Και σε αυτή την περίπτωση, ανάμεσα στις διάφορες αρχιτεκτονικές μοντέλων μηχανικής μάθησης, οι τεχνικές που βασίζονται στο RNN αποτελούν τις πλέον ενδεδειγμένες καθώς η οπτική ίνα παραμένει ένα κανάλι με έντονη διασπορά. Οι περισσότερες εργασίες στη σύγχρονη βιβλιογραφία δείχνουν ξεκάθαρα ότι τα σωστά επιλεγμένα μοντέλα μηχανικής μάθησης έχουν την ιδιότητα να βελτιώνουν την επίδοση BER διατηρώντας παράλληλα την πολυπλοκότητα σε μέτρια επίπεδα σε σύγκριση με τις επικρατέστερες συμβατικές λύσεις μη γραμμικής ισοστάθμισης σε περιβάλλον IM/DD. Όπως για παράδειγμα στην εργασία [52] στην οποία οι συγγραφείς απέδειξαν πώς η έξοδος πολλαπλών συμβόλων LSTM, μπορεί να επιφέρει ανώτερη απόδοση εξισορρόπησης σε συστήματα IM/DD ρυθμού 200 Gb/s σε σύγκριση με τους μη γραμμικούς ισοσταθμιστές Volterra.

1.6 Συμβολή της παρούσας διατριβής



Σχήμα 1.6: Θεματικές περιοχές και πεδία εφαρμογής της διατριβής

Στην διατριβή αυτή μελετήθηκαν εκτενώς τα φαινόμενα που προκαλούνται κατά τη διάδοση υψίρρυθμων σημάτων σε οπτική ίνα και οι αποδοτικές ψηφιακές μέθοδοι επεξεργασίας που μπορούν να εφαρμοστούν στα σύγχρονα συστήματα οπτικών επικοινωνιών.

Τα κύρια αποτελέσματα της έρευνα είναι η επιτυχής αντιμετώπιση των μη γραμμικοτήτων της οπτικής ίνας μέσω της εφαρμογής αμφίδρομων επαναλαμβανόμενων νευρωνικών δικτύων (bi-RNN) σε σύμφωνα συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων και σε IM/DD συστήματα μικρών αποστάσεων. Επίσης η μοντελοποίηση των δυναμικών χαρακτηριστικών φωτονικών διατάξεων που χρησιμοποιούνται σε συστήματα κέντρων δεδομένων, με τη βοήθεια των δικτύων RNN. Το σχήμα 1.6. αναπαριστά τις θεματικές περιοχές και τα πεδία εφαρμογής της διατριβής.

Σε ότι αφορά το πρώτο σκέλος της, η προσθήκη στο δέκτη αυτόνομης μονάδας ψηφιακής επεξεργασίας βασισμένης σε μοντέλο μηχανικής μάθησης και συγκεκριμένα σε αμφίδρομο RNN, επιφέρει τη σημαντική βελτίωση της απόδοσης του συστήματος. Επιπλέον η υιοθέτηση κατάλληλων μεθόδων παράλληλης εξαγωγής συμβόλων, συμβάλει στην αξιοσημείωτη μείωση της πολυπλοκότητας υλοποίησης έναντι άλλων μεθόδων επεξεργασίας, καθιστώντας τους ισοσταθμιστές RNN, μια ιδιαίτερα θελκτική πρόταση για την ψηφιακή επεξεργασία των υψίρρυθμων συστημάτων οπτικών επικοινωνιών μεγάλων αλλά και μικρών αποστάσεων.

Στο δεύτερο σκέλος αυτής της διατριβής, αποδεικνύεται πως οι επεξεργαστές RNN μπορούν να προσομοιώσουν με ιδιαίτερα μεγάλη ακρίβεια συστήματα οπτικών επικοινωνιών μικρών αποστάσεων. Τα δυναμικά χαρακτηριστικά των επιμέρους διατάξεων μπορούν να αναπαραχθούν με αποδοτικό τρόπο, έναντι των καθιερωμένων λύσεων που βασίζονται σε σύνθετα μαθηματικά μοντέλα υψηλής πολυπλοκότητας.

Η διατριβή βασίστηκε σε αναλύσεις, προσομοιώσεις και πειράματα που πραγματοποιήθηκαν σε διάφορα σενάρια εφαρμογής, τόσο σε δίκτυα μεγάλων αποστάσεων (>1000 km) και μητροπολιτικών δικτύων (>120 km) που χρησιμοποιούν σύμφωνο (coherent) δέκτη, όσο και σε οπτικά δίκτυα μικρών αποστάσεων (< 120 km) που βασίζονται σε δέκτες άμεσης φώρασης.

1.7 Δομή της διατριβής

Η διατριβή αυτή αποτελείται από 8 κεφάλαια, τα περιεχόμενα των κεφαλαίων που ακολουθούν περιγράφονται ως εξής:

Κεφάλαιο 2: Συστήματα οπτικών επικοινωνιών

Στο κεφάλαιο 2 παρουσιάζονται οι βασικές έννοιες που διέπουν τα συστήματα οπτικών επικοινωνιών. Αναλύεται η διάδοση στην οπτική ίνα, περιγράφονται τα γραμμικά φαινόμενα όπως οι απώλειες, η διασπορά και τα φαινόμενα πόλωσης, τα μη γραμμικά φαινόμενα και οι συμβατικές τεχνικές αντιμετώπισής τους. Παρουσιάζονται τα IM/DD

συστήματα μικρών αποστάσεων, τα σύμφωνα (coherent) συστήματα μεγάλων αποστάσεων και οι ψηφιακές μέθοδοι επεξεργασίας που έχουν υιοθετηθεί.

Κεφάλαιο 3: Εφαρμογές Μηχανικής Μάθησης, τεχνικές εξισορρόπησης σε Συστήματα Οπτικών Επικοινωνιών

Στο κεφάλαιο 3 γίνεται εισαγωγή στις βασικές έννοιες της μηχανικής μάθησης και της τεχνητής νοημοσύνης. Παρουσιάζονται τα είδη μηχανικής μάθησης ενώ αναλύεται η έννοια της βαθιάς μάθησης, τα κυριότερα μοντέλα νευρωνικών δικτύων και οι βασικές τεχνικές εκπαίδευσής τους. Ιδιαίτερη αναφορά και λεπτομερής ανάλυση γίνεται για τα επαναλαμβανόμενα νευρωνικά δίκτυα, τις διάφορες παραλλαγές και αρχιτεκτονικές τους.

Κεφάλαιο 4: Ισοστάθμιση Βασισμένη σε Επαναλαμβανόμενα Νευρωνικά Δίκτυα στα Συστήματα Οπτικών Επικοινωνιών Μεγάλων Αποστάσεων

Το κεφάλαιο 4 αποτελεί το σημαντικότερο τμήμα της διατριβής καθώς το μεγαλύτερο μέρος της ερευνητικής εργασίας που προηγήθηκε παρουσιάζεται σε αυτό. Προτείνεται η λύση των RNN για την εξισορρόπηση των μη γραμμικών φαινομένων διάδοσης στα συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων και αναλύεται η συμπεριφορά των μοντέλων LSTM, GRU και Vanilla-RNN. Πραγματοποιείται σύγκριση επιδόσεων με συμβατικούς μη γραμμικούς επεξεργαστές όπως τύπου DBP και Volterra και προτείνονται τεχνικές σημαντικής μείωσης της υπολογιστικής πολυπλοκότητας. Τέλος επιβεβαιώνονται πειραματικά οι αναμενόμενες επιδόσεις των RNN ισοσταθμιστών, η περαιτέρω βελτίωση των επιδόσεών τους και η μείωση της υπολογιστικής πολυπλοκότητας, όταν υιοθετείται η πολυκαναλική παραλλαγή τους.

Κεφάλαιο 5: Ισοστάθμιση Βασισμένη σε Επαναλαμβανόμενα Νευρωνικά Δίκτυα στα Συστήματα Οπτικών Επικοινωνιών Μικρών Αποστάσεων

Το κεφάλαιο 5 παρουσιάζει τη χρήση αμφίδρομων ισοσταθμιστών βασισμένων σε GRU και Vanilla-RNN για τη βελτίωση της απόδοσης της WDM διάδοσης στη ζώνη O. Στη βάση

πειραματικής IM/DD διάταξης και σε σενάρια διάδοσης 100 km και 75 km χρησιμοποιώντας διαμόρφωση OOK και PAM4 αντίστοιχα, ο Vanilla-RNN ισοσταθμιστής παρουσιάζει σημαντικά καλύτερη επίδοση BER από το συμβατικό σχήμα DFE, από το γραμμικό και από το μη γραμμικό ισοσταθμιστή Volterra 3^{ης} τάξης.

Κεφάλαιο 6: Προσομοίωση Οπτικών Συστημάτων με τη χρήση Επαναλαμβανόμενων Νευρωνικών Δικτύων

Στο κεφάλαιο 6 παρουσιάζεται η αντίστροφη λειτουργία των RNN. Εξετάζεται η ικανότητά τους να προβλέψουν και να αναπαράγουν τα δυναμικά χαρακτηριστικά ενός πραγματικού οπτικού συστήματος. Χρησιμοποιούνται δύο τύποι RNN, το bi-LSTM και το bi-VRNN για να προσομοιωθεί η δυναμική συμπεριφορά ενός πειραματικά χαρακτηρισμένου οπτικού πομπού που βασίζεται σε VCSEL. Αναλύονται τα χαρακτηριστικά του μοντέλου και προτείνονται τρόποι αξιοποίησής του για την αποδοτικότερη προσομοίωση.

Κεφάλαιο 7: Παράλληλη Επεξεργασία

Στο κεφάλαιο 7 παρουσιάζονται οι τεχνικές παράλληλης επεξεργασίας βασισμένες σε μονάδες επεξεργασίας γραφικών (graphics processing units - GPUs) που χρησιμοποιήθηκαν για την αποδοτικότερη εκπαίδευση των σύνθετων νευρωνικών δικτύων, την επιτάχυνση της εξομοίωσης των μοντέλων οπτικής διάδοσης αλλά και άλλων επιμέρους προσομοιώσεων που πραγματοποιήθηκαν κατά τη διάρκεια εκπόνησης της διατριβής.

Κεφάλαιο 8: Σύνοψη

Αυτό το κεφάλαιο ολοκληρώνει τη διατριβή παρέχοντας μια περίληψη των σημαντικότερων συνεισφορών και παρέχει μια συζήτηση για μελλοντικά σχέδια έρευνας και ανάπτυξης στη θεματική περιοχή που καλύπτει.

2° ΚΕΦΑΛΑΙΟ

Συστήματα οπτικών επικοινωνιών

2.1 Εισαγωγή

Η ιδιότητα της ανάκλασης του φωτός, το βασικότερο φαινόμενο που επιτρέπει τη διάδοση του στην οπτική ίνα, ήταν γνωστή από το 19° αιώνα. Παρά το γεγονός πως οι οπτικές ίνες κατασκευάζονταν ήδη από τη δεκαετία του 1920, χρειάστηκαν 60 έτη εντατικών δοκιμών και βελτιώσεων στα δομικά χαρακτηριστικά τους, όπως προσθήκη διηλεκτρικού μανδύα διαφορετικού δείκτη διάθλασης από τον πυρήνα, στο υλικό τους και στη φασματική περιοχή λειτουργίας τους, ώστε το 1979, έτος ορόσημο για τις οπτικές επικοινωνίες, επιτυγχάνεται η διάδοση φωτός σε μονότροπη οπτική ίνα SiO₂ με απώλειες μόλις 0.2 dB/km στο μήκος κύματος των 1550nm [53].

Η ιδιότητα των χαμηλών απωλειών είναι θεμελιώδους σημασίας καθώς επιτρέπει στα σήματα να διαδίδονται για μεγάλες αποστάσεις, οι οποίες εντούτοις δεν μπορούν να ξεπεράσουν τις μερικές δεκάδες χιλιόμετρα. Όταν οι αποστάσεις μεταφοράς είναι μεγαλύτερες, καθίσταται επιτακτική η ενίσχυση του σήματος. Τη δεκαετία του 1990 δοκιμάστηκε η πρόσμιξη σπάνιων γαιών όπως το Έρβιο στις οπτικές ίνες. Αυτό οδήγησε στην κατασκευή των ενισχυτών Ερβίου, (Erbium-Doped Fiber Amplifiers, EDFA) [54]. Σήμερα οι οπτικοί ενισχυτές που χρησιμοποιούνται είναι ενισχυτές ίνας με προσμίξεις σπάνιων γαιών όπως EDFA, Πρασεοδυμίου, (Praseodymium Doped Fiber Amplifiers – PDFA) κ.ά, τύπου Raman [55] και ημιαγωγικοί οπτικοί ενισχυτές (Semiconductor Optical Amplifiers - SOA) [56]. Η ανακάλυψη του EDFA και η παραγωγή μονότροπης ίνας με απώλειες μικρότερες από 0.2 dB/km σηματοδότησε την ανάδυση και καθιέρωση των συστημάτων οπτικών επικοινωνιών. Έκτοτε αρκετές τεχνολογικές εξελίξεις υιοθετήθηκαν

με σκοπό να ικανοποιηθούν οι διαρκώς αυξανόμενες απαιτήσεις επικοινωνίας υψηλής ταχύτητας και μεγάλου εύρους ζώνης. Μια από τις σημαντικότερες είναι η τεχνική της πολυπλεξίας μήκους κύματος, (Wavelength Division Multiplexing, WDM), με την οποία γίνεται εκμετάλλευση των μεγάλων φασματικών περιοχών χαμηλής εξασθένισης της ίνας πολυπλέκοντας πολλά οπτικά σήματα στην ίδια ίνα, αυξάνοντας με αυτό τον τρόπο δραστικά τη χωρητικότητα του καναλιού.

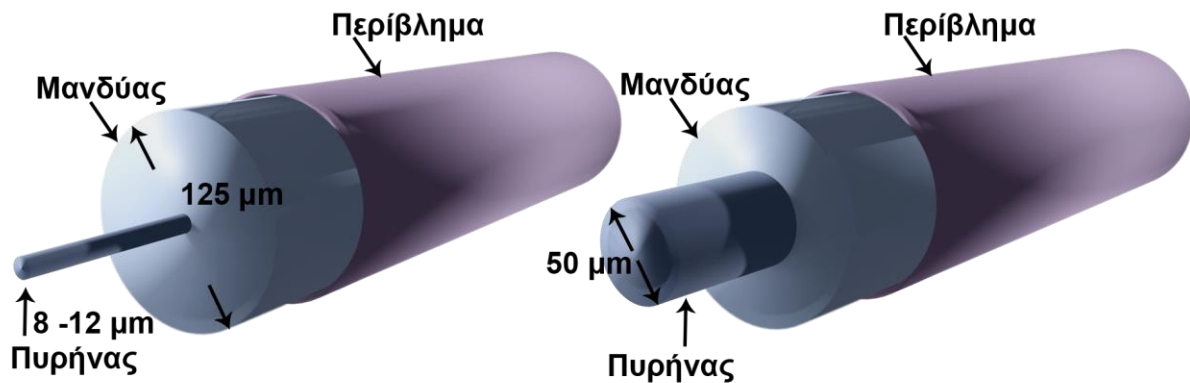
Τον ρόλο του πομπού στις σύγχρονες οπτικές επικοινωνίες, για τη μετατροπή των ηλεκτρικών σημάτων σε οπτικά και κατόπιν τη διάδοσή τους στην οπτική ίνα, έχει κυρίως η πηγή Laser (Light Amplification by Stimulated Emission of Radiation) [57] και λιγότερο η πηγή LED (Light Emitting Diode) [58]. Η πηγή Laser συνήθως συνδυάζεται με κάποιο τύπο οπτικού διαμορφωτή, υπεύθυνου για την αποτύπωση ενός ευρυζωνικού ηλεκτρικού σήματος σε ένα οπτικό πεδίο με επίτευξη μεγάλης γραμμικότητας, υψηλού συντελεστή διαμόρφωσης και δυνατότητα διαμόρφωσης τόσο του πλάτους όσο και της φάσης του σήματος.

Τον ρόλο του δέκτη του διαδιδόμενου σήματος, τη μετατροπή δηλαδή του οπτικού σήματος σε ηλεκτρικό, αναλαμβάνει ο οπτικός ανιχνευτής (photodetector). Συνήθως πρόκειται για μία ή περισσότερες φωτοδιόδους, ανάλογα με το είδος του δέκτη που πρέπει να υλοποιηθεί, συνοδευόμενες από κυκλώματα απόφασης και ανάκτησης παλμών χρονισμού.

Στο κεφάλαιο αυτό παρουσιάζονται τα βασικά φαινόμενα της διάδοσης στην οπτική ίνα [59], δίνοντας μεγάλη έμφαση στα μη γραμμικά φαινόμενα, αναλύονται μέθοδοι διαμόρφωσης και αποδιαμόρφωσης των ψηφιακών σημάτων και περιγράφονται τα επιμέρους στοιχεία που χρησιμοποιούνται στα σύγχρονα συστήματα οπτικών επικοινωνιών.

2.2 Βασικά φαινόμενα διάδοσης του φωτός σε οπτική ίνα

Η οπτική ίνα είναι ένας διηλεκτρικός κυματοδηγός. Στην πιο απλή της μορφή, μια οπτική ίνα αποτελείται από την κεντρική περιοχή του πυρήνα που περικλείεται από την



Σχήμα 2.1 Αναπαράσταση της μονότροπής (α) και πολύτροπης (β) οπτικής ίνας

περιοχή του μανδύα και το προστατευτικό περίβλημα. Βασίζει την ικανότητα κυματοδότησης της στον πλήρη εσωτερικό εγκλωβισμό του φωτός στη διεπαφή μανδύα-πυρήνα. Στο σχήμα 2.1 διακρίνονται οι δύο βασικοί τύποι οπτικών ινών, η μονότροπη και η πολύτροπη.

Η μονότροπη ίνα (single mode fiber, SMF) αποτελείται από πυρήνα διαμέτρου 8-12μm και βηματικό (step index) προφίλ δείκτη διάθλασης, επιτρέποντας με τον τρόπο αυτό τη διάδοση μονάχα ενός τρόπου. Για τη διάδοση υψηλού ρυθμού πληροφορίας σε δίκτυα που καλύπτουν αποστάσεις μεγαλύτερες του 1 km χρησιμοποιούνται αποκλειστικά ίνες SMF.

Όταν η διάμετρος του πυρήνα αυξηθεί περίπου στα 50 μm η οπτική ίνα ονομάζεται πολύτροπη (multimode fiber, MMF) και επιτρέπει τη διέγερση περισσότερων του ενός τρόπων διάδοσης. Οι πολύτροπες οπτικές ίνες δε χρησιμοποιούνται σε επικοινωνίες μεγάλων αποστάσεων ωστόσο είναι ιδανικές για αποστάσεις μικρότερες του 1 km και συγκεκριμένα στα τοπικά δίκτυα (Local Area Networks – LANs) και στα κέντρα δεδομένων (Data Centers).

2.2.1 Απώλειες στην ίνα

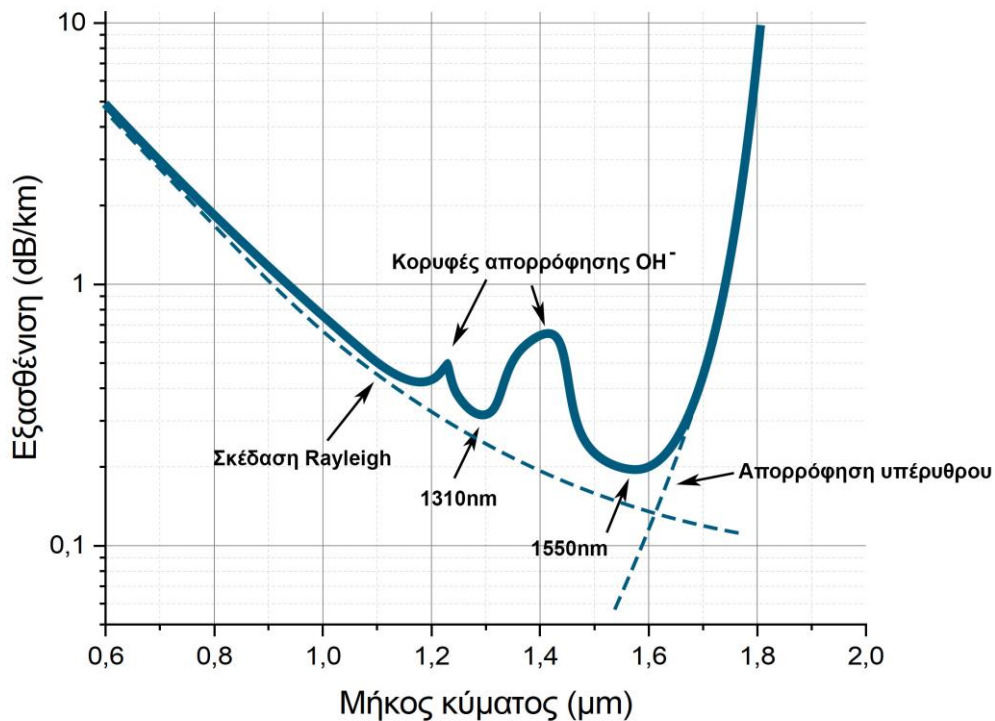
Το φως κατά τη διάδοσή του στην οπτική ίνα υπόκειται σε απώλεια της ισχύος του. Αν P_{in} είναι η οπτική ισχύς εισόδου στην ίνα μήκους L , η ισχύς στην έξοδο P_{out} δίνεται από την εξίσωση:

$$P_{out} = P_{in} e^{-\alpha L} \quad (2.1)$$

Όπου α η σταθερά εξασθένησης εκφρασμένη σε dB/km και υπολογίζεται από τη σχέση :

$$\alpha_{dB} = -\frac{10}{L} \log\left(\frac{P_{out}}{P_{in}}\right) = 4.343\alpha \quad (2.2)$$

Εξαρτάται από το μήκος κύματος του φωτός σύμφωνα με το σχήμα 2.2.



Σχήμα 2.2: Η ολική εξασθένηση ως συνάρτηση του μήκους κύματος. Διακρίνονται τα δύο παράθυρα χαμηλής εξασθένησης

Παρατηρούνται δύο περιοχές χαμηλής εξασθένισης, η περιοχή των 1310nm (ζώνη O, O-band) και η περιοχή των 1550nm (ζώνη C, C-band). Το εύρος ζώνης της καθεμίας έχει μεγάλη έκταση, 1280-1325nm για το πρώτο και 1420-1650nm για το δεύτερο, δίνοντας τη δυνατότητα, όπως θα αναλυθεί στη συνέχεια, της χρήσης του για πολυπλεξία μήκους κύματος.

Διάφοροι παράγοντες συνεισφέρουν στην εξασθένιση του σχήματος 2.1 με τη σκέδαση Rayleigh, την απορρόφηση λόγω ιόντων OH^- και την απορρόφηση στο υπέρυθρο να είναι οι σημαντικότεροι.

Η σκέδαση Rayleigh παρατηρείται όταν οι ατέλειες του υλικού (SiO_2) «παγώνουν» κατά τη διαδικασία της κατασκευής της ίνας και δημιουργούν τοπικές αυξομειώσεις στο δείκτη διάθλασης του φωτός σε όλες τις κατευθύνσεις. Οι απώλειες λόγω σκέδασης Rayleigh είναι ανάλογες του $1/\lambda^4$, τουτέστιν είναι μεγαλύτερες στα μικρότερα μήκη κύματος.

Κατά τη διαδικασία της κατασκευής της ίνας, σχηματίζονται ιόντα OH^- τα οποία έχουν χαρακτηριστικές κορυφές απορρόφησης σε μήκη κύματος $2720 \text{ nm} / (n+1)$, $n=1,2,3,\dots$, όπου n η τάξη της αρμονικής. [60]

Η απορρόφηση στη φασματική περιοχή του υπέρυθρου οφείλεται στο γεγονός ότι μέρος των φωτονίων που προσπίπτουν στο υλικό της ίνας καταναλώνεται σε εξαναγκασμένες ηλεκτρονικές μεταπτώσεις στα μόρια του SiO_2 , ενώ επιπλέον φωτόνια στην περιοχή του υπέρυθρου διεγείρουν με την ενέργειά τους μοριακές δονήσεις. Η απορρόφηση αυτή οδηγεί σε εξασθένιση του φωτός που κυματοδηγείται.

Επιπρόσθετοι παράγοντες απωλειών είναι η κάμψη των οπτικών ινών κατά την εγκατάσταση μίας ζεύξης, οι απώλειες στους ενδιάμεσους συνδετήρες (connectors) και οι απώλειες των συγκολλήσεων (splices).

2.2.2 Διασπορά

Όταν διαφορετικές συνιστώσες του σήματος διαδίδονται με διαφορετικές ταχύτητες στην οπτική ίνα, με αποτέλεσμα την άφιξή τους στο δέκτη σε διαφορετικές χρονικές

στιγμές, αλλοιώνοντας τη χρονική μορφή του παλμού του σήματος. Το φαινόμενο αυτό ονομάζεται διασπορά καθώς ο παλμός συνήθως διευρύνεται (διασπείρεται) χρονικά.

Τα είδη της διασποράς που κυριαρχούν στη μονότροπη οπτική ίνα είναι η χρωματική διασπορά, η διασπορά κυματοδήγησης και η διασπορά τρόπων πόλωσης.

Στις πολύτροπες οπτικές ίνες επικρατεί όλων των προηγούμενων ειδών η διασπορά τρόπων διάδοσης. Λόγω των διαφορετικών διαδρομών που ακολουθούν οι διαφορετικές χωρικές κατανομές του φωτός που ανήκουν σε άλλους τρόπους, η ενέργεια ενός στενού παλμού θα διασπαρεί στον χρόνο.

2.2.2.1 Χρωματική Διασπορά

Οι διαφορετικές φασματικές συνιστώσες ενός σήματος διαδίδονται στην οπτική ίνα με διαφορετικές ταχύτητες περιγράφοντας τη χρωματική διασπορά η οποία έχει δύο προελεύσεις. Η πρώτη και κυρίαρχη σχετίζεται με την εξάρτηση του δείκτη διάθλασης n από τη συχνότητα ω , γεγονός το οποίο προκαλεί τη διασπορά του υλικού. Η δεύτερη είναι η διασπορά κυματοδήγησης κατά την οποία ένα μέρος του φωτός κυματοδηγείται στον πυρήνα και ένα άλλο στον μανδύα, με συνέπεια, η ταχύτητα να εξαρτάται από τον λόγο της ισχύος εντός και εκτός πυρήνα.

Στην περίπτωση των στενών οπτικών παλμών που χρησιμοποιούνται στις οπτικές επικοινωνίες η διασπορά του υλικού έχει βαρύνουσα σημασία διότι οι οπτικοί παλμοί έχουν πολλές φασματικές συνιστώσες. Κάθε μία από αυτές ταξιδεύει στην ίνα με τη δική της ταχύτητα που καθορίζεται από τη σχέση $c/n(\omega)$ όπου c η ταχύτητα του φωτός, με αποτέλεσμα τη διεύρυνση του παλμού.

Η επίδραση της διασποράς μπορεί να υπολογιστεί ποσοτικά αναλύοντας τη σταθερά του τρόπου διάδοσης β σε μια σειρά Taylor συναρτήσεως της κεντρικής συχνότητας ω_0 , σύμφωνα με την εξίσωση (2.3),

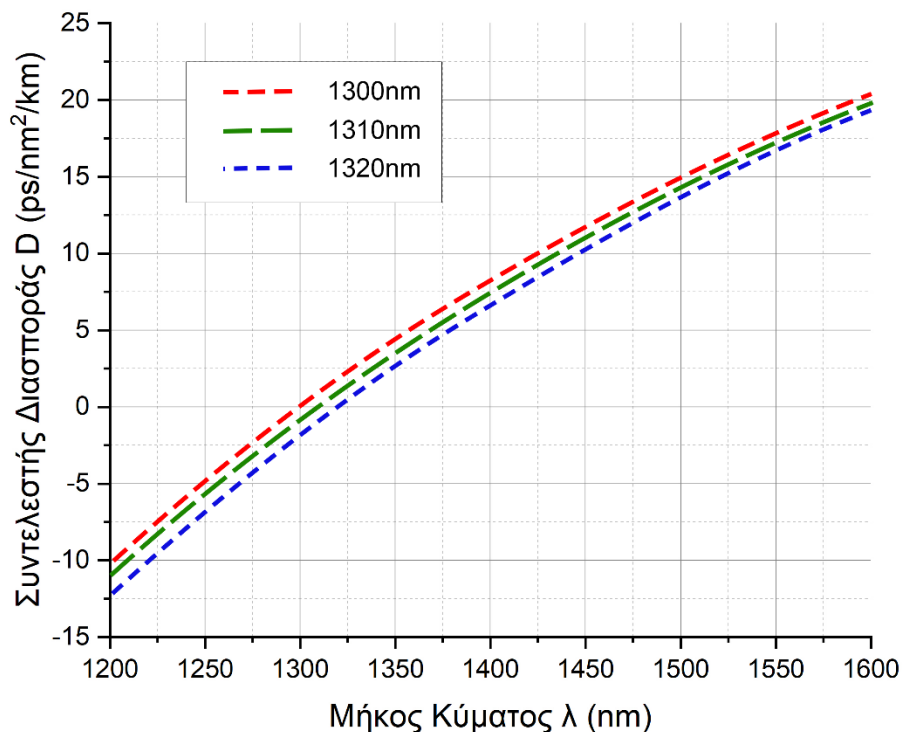
$$\beta(\omega) = n(\omega) \frac{\omega}{c} = \beta_0 + \beta_1(\omega - \omega_0) + \frac{1}{2} \beta_2(\omega - \omega_0)^2 + \dots \quad (2.3)$$

$$\text{όπου } \beta_m = \left(\frac{d^m \beta}{d\omega^m} \right) \Big|_{\omega=\omega_0}, m = 1, 2, \dots, \infty \quad (2.4)$$

Από την εξίσωση (2.3) ο 2^{ος} όρος β_2 μπορεί να εκφραστεί μέσω της σχέσης:

$$\beta_2 = \frac{1}{c} \left(2 \frac{dn}{d\omega} + \omega \frac{d^2n}{d\omega^2} \right) \quad (2.5)$$

Η παράμετρος β_2 ονομάζεται παράμετρος της διασποράς της ταχύτητας ομάδας (group velocity dispersion parameter, GVD parameter). Η ορολογία αυτή δικαιολογείται αν αναλογιστούμε ότι $\beta_1 = \frac{d\beta}{d\omega}$ και $1/\beta_1$ είναι η ταχύτητα ομάδας. Εφόσον λοιπόν η παράμετρος β_2 σχετίζεται με τον ρυθμό της αλλαγής της ταχύτητας ομάδας σε συνάρτηση με τη συχνότητα, η χρωματική διασπορά ονομάζεται επίσης και διασπορά ταχύτητας ομάδας (GVD). Γενικά η τιμή β_2 είναι διάφορη του μηδενός και μάλιστα η τιμή της εξαρτάται από τη συχνότητα. Για τις οπτικές ίνες υπάρχει ένα μήκος κύματος, το ονομαζόμενο μήκος κύματος μηδενικής διασποράς λ_0 για το οποίο ισχύει ότι $\beta_2(\omega_0)=0$.



Σχήμα 2.3: Η εξάρτηση της D από το μήκος κύματος για τις συνήθειες μονότροπες οπτικές ίνες. Παίρνει πολύ μικρές τιμές, σχεδόν 0 γύρω από τα 1310nm.

Αν $\beta_2 > 0$, τότε η οπτική διάδοση γίνεται στο καθεστώς ομαλής χρωματικής διασποράς, ενώ αντίθετα γίνεται στο καθεστώς ανώμαλης χρωματικής διασποράς. Στην περίπτωση της ομαλής διασποράς, τα μεγαλύτερα μήκη κύματος ταξιδεύουν με μεγαλύτερη ταχύτητα. Για την περιγραφή της χρωματικής διασποράς συνηθίζεται η παράμετρος D που συνδέεται με την παράμετρο β_2 με βάση τη σχέση:

$$D = -\frac{2\pi c}{\lambda^2} \beta_2 \quad (2.6)$$

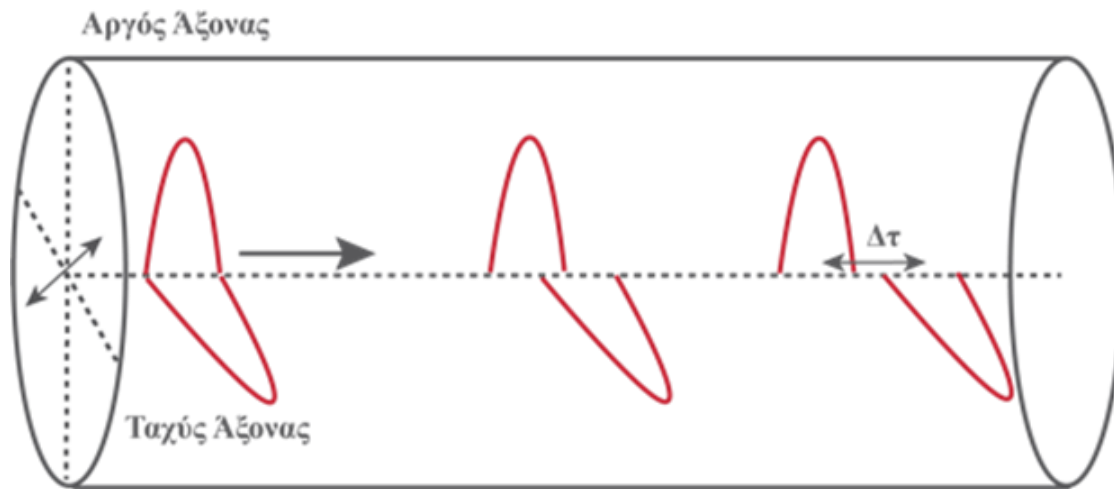
Η χρωματική διασπορά μπορεί να προσδιοριστεί χρησιμοποιώντας το συντελεστή χρωματικής διασποράς αναφοράς. Για παράδειγμα η διασπορά που παρατηρείται στην περιοχή των 1550 nm σε συνάρτηση με το μήκος κύματος μπορεί να υπολογιστεί από τη σχέση:

$$D_L = L [D_{1550} + S_{1550}(\lambda - 1550)] \quad (2.7)$$

Όπου L το μήκος της οπτικής ζεύξης, D_{1550} η χρωματική διασπορά στα 1550nm με τιμή 17 ps/nm × km και S_{1550} η κλίση της καμπύλης της χρωματικής διασποράς στα 1550nm με τιμή 0.056 ps/nm² × km [61].

Όπως αναφέρθηκε παραπάνω ένα μέρος του οπτικού σήματος διαδίδεται διαμέσου του μανδύα της ίνας, φαινόμενο που οδηγεί στη διασπορά κυματοδήγησης. Σε γενικές γραμμές η συνεισφορά της διασποράς κυματοδήγησης σε σχέση με αυτή του υλικού είναι αρκετά μικρότερη εκτός από τη στενή περιοχή κοντά στο μήκος κύματος μηδενικής διασποράς λ_0 (zero dispersion wavelength, ZDWL), για το οποίο η παράμετρος β_2 είναι μηδέν. Στις συνηθισμένες μονότροπες ίνες η διασπορά είναι μηδέν στην περιοχή κοντά στο μήκος των 1310nm (σχήμα 2.3). Θετικές τιμές του β_2 ορίζουν οπτική διάδοση σε καθεστώς ομαλής χρωματικής διασποράς ενώ αρνητικές σε καθεστώς ανώμαλης χρωματικής διασποράς.

2.2.3 Φαινόμενα Πόλωσης



Σχήμα 2.4: Η διασπορά τρόπων πόλωσης ενός οπτικού παλμού κατά τη διάδοσή του σε τμήμα ανομοιομόρφης ίνας

Η τυπική μονότροπη ίνα, κατ' ουσίαν υποστηρίζει δύο κάθετα πολωμένους τρόπους διάδοσης με την ίδια χωρική κατανομή. Σε μία ιδανική ίνα, (με απόλυτη κυλινδρική συμμετρία σε όλο της το μήκος) οι δύο τρόποι είναι ισοδύναμοι, με όμοιους δείκτες διάθλασης n_x και n_y . Αυτό έχει ως αποτέλεσμα το κύμα που εισέρχεται στην ίνα να διεγείρει και τους δύο τρόπους διάδοσης (δύο πολώσεις), οι οποίοι λόγω ίδιας ταχύτητας καταφθάνουν μαζί στον δέκτη. Στις πραγματικές, μη ιδανικές οπτικές ίνες, δε συναντάμε απόλυτη κυλινδρική συμμετρία. Οι διαφορές κατασκευαστικές ατέλειες στη γεωμετρία αλλά και εξαιτίας των μηχανικών τάσεων, διαταραχών λόγω κάμψεων, θερμοκρασιακών αλλαγών κτλ, έχουν ως αποτέλεσμα την εμφάνιση διαφορετικών σταθερών διάδοσης, οπότε και διαχωρισμό των τρόπων πόλωσης δηλαδή διάδοση με διαφορετική ταχύτητα ομάδας (differential group delay, DGD). Το φαινόμενο αυτό ονομάζεται διπλοθλαστικότητα και σχετίζεται με τη διαφορά των δεικτών διάθλασης των δύο πολωτικών αξόνων $\Delta n = n_x - n_y$. Στην περίπτωση διάδοσης στενών οπτικών παλμών η καθυστέρηση μεταξύ των τρόπων δίνεται από τη σχέση:

$$\Delta\tau = \left| \frac{L}{v_{gx}} - \frac{L}{v_{gy}} \right| = L |\beta_{1x} - \beta_{1y}| = L\Delta\beta_1 \quad (2.8)$$

Η ποσότητα $\Delta\beta_1 = \Delta\tau/L$ είναι η παράμετρος DGD. Σε κλασσικές μονότροπες οπτικές ίνες η διπλοθλαστικότητα είναι της τάξης των ps/km και εκδηλώνεται με τυχαία σύζευξη της πόλωσης μεταξύ των δύο πολωτικών αξόνων, που έχει σαν αποτέλεσμα το φαινόμενο της διασποράς τρόπων πόλωσης (polarization mode dispersion, PMD), όπως φαίνεται στο σχήμα 2.4. Το φαινόμενο PMD σχετίζεται με τις τυχαίες διακυμάνσεις στην διπλοθλαστικότητα που προκαλούν αλλοίωση των παλμών, καθώς οι δύο πολωτικές συνιστώσες δε διαδίδονται ταυτόχρονα αλλά η μία προηγείται ή έπεται της άλλης με τυχαίο τρόπο. Τελικά, η τελική χρονική καθυστέρηση Δt - που είναι αντίστοιχη της χρονικής εξάπλωσης που παρατηρούμε σε άλλα είδη διασποράς - είναι δύσκολο να προβλεφθεί επειδή η διπλοθλαστικότητα επηρεάζεται ακόμα και από τις περιβαλλοντικές συνθήκες. Η χρονική διεύρυνση $\Delta\tau$ της σχέσης 2.8 μεταβάλλεται τυχαία κατά μήκος της ίνας λόγω της διασποράς τρόπων πόλωσης PMD και υπολογίζεται στατιστικά από την RMS τιμή της σύμφωνα με τη σχέση:

$$\Delta\tau_{RMS}^2(z) = 2\Delta\beta_1^2 l_c^2 \left[\exp\left(-\frac{z}{l_c}\right) + \frac{z}{l_c} - 1 \right] \quad (2.9)$$

όπου l_c το μήκος συσχέτισης της διπλοθλαστικότητας δηλαδή το μήκος του κάθε τμήματος κατά το οποίο θεωρούμε τη διπλοθλαστικότητα σταθερή) και τις δύο πολωτικές συνιστώσες συσχετισμένες (~ 10 m). Για μικρές αποστάσεις $z \ll l_c$, η διπλοθλαστικότητα είναι σταθερή και ίση με: $\Delta\tau_{RMS}(z) = (\Delta\beta_1)z$. Στην περίπτωση που $z \gg 1$ km, τότε καταλήγουμε στη σχέση:

$$\Delta\tau_{RMS}(z) = (\Delta\beta_1)\sqrt{2l_c z} = D_p\sqrt{z} \quad (2.10)$$

Όπου D_p ο συντελεστής διασποράς τρόπων πόλωσης PMD που μετριέται σε ps/km^{1/2} ενώ οι τιμές του ποικίλουν σε ένα εύρος τιμών 0.01-10 ps/km^{1/2}. Οι οπτικές ίνες που πρωτο-εγκαταστάθηκαν τη δεκαετία του 1980 είχαν πολύ υψηλό συντελεστή PMD. Τα πρότυπα της ITU-T [61] επιβάλλουν οι σύγχρονες οπτικές ίνες να έχουν έναν συντελεστή PMD κατά μέγιστο 0.5 ps/km^{1/2}. Η διασπορά τρόπων πόλωσης PMD υποβαθμίζει τις επιδόσεις ενός υψίρρυθμου συστήματος σε συνδυασμό με τη χρωματική διασπορά αν και με σαφώς μικρότερη επίδραση.

2.2.4 Αντιστάθμιση Γραμμικών Φαινομένων

Τα σύγχρονα συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων βασίζονται σε τεχνολογίες σύμφωνης φώρασης που αξιοποιούν τεχνικές ψηφιακής επεξεργασίας σήματος για την αντιστάθμιση όλων των γραμμικών φαινομένων. Το σημαντικότερο εξ αυτών όπως αναλύθηκε προηγουμένως είναι η χρωματική διασπορά. Σε γενικές γραμμές η αντιστάθμιση της χρωματικής διασποράς επιτυγχάνεται με δύο βασικά σχήματα, την εξισορρόπηση στο πεδίο των συχνοτήτων (Frequency-Domain Equalization – FDE) [62] και την εξισορρόπηση στο πεδίο του χρόνου (Time-Domain Equalization – TDE) [63]. Στην δεύτερη περίπτωση και όταν η χρωματική διασπορά είναι σχετικά μικρή χρησιμοποιούνται απλά φίλτρα πεπερασμένης κρουστικής απόκρισης (Finite impulse response - FIR) [64]. Στην περίπτωση υψηλής συσσωρευμένης χρωματικής διασποράς σε συνδέσμους χωρίς αντιστάθμιση, η εξισορρόπηση της αλλά και η άρση του φαινομένου του PMD μπορεί να πραγματοποιηθεί σε δύο βήματα, αντισταθμίζοντας πρώτα τη στατική διασπορά και στη συνέχεια χρησιμοποιώντας προσαρμοστικούς ισοσταθμιστές (Adaptive Equalizer) [65] επιτυγχάνεται η πλήρης άρση των γραμμικών φαινομένων [66]. Στη συνέχεια του κεφαλαίου θα αναλυθούν οι βασικές DSP τεχνικές που χρησιμοποιούνται στα σύγχρονα συστήματα οπτικών επικοινωνιών.

2.3 Μη γραμμικά φαινόμενα

Το κανάλι της οπτικής ίνας είναι θεμελιωδώς μη γραμμικό. Οι εμφανιζόμενες μη γραμμικότητες προέρχονται από ελαστική σκέδαση του διαδιδόμενου κύματος στο διηλεκτρικό μέσο κατά την οποία δεν ανταλλάσσεται ενέργεια μεταξύ του κύματος και του μέσου ή από ανελαστική σκέδαση όπου μεταφέρεται ενέργεια από φωτόνια του διαδιδόμενου κύματος σε φωνόνια του μέσου. Στην ελαστική σκέδαση τα μη γραμμικά φαινόμενα περιγράφονται από την εξάρτηση του δείκτη διάθλασης του μέσου από την ένταση του κύματος. Στην ανελαστική σκέδαση προκύπτουν φωτόνια χαμηλότερης συχνότητας δηλαδή μεγαλύτερου μήκους κύματος εξαιτίας της διαφοράς ενεργείας.

Η απόκριση του μέσου στο εξωτερικά εφαρμοζόμενο πεδίο \vec{E} περιγράφεται από το διάνυσμα της διηλεκτρικής πόλωσης \vec{P} το οποίο σε γραμμικά, ομογενή και ιστροπικά μέσα είναι γραμμικά ανάλογο με το διάνυσμα του εφαρμοζόμενου πεδίου. Η σταθερά αναλογίας σε αυτή την περίπτωση είναι βαθμωτό μέγεθος και σχετίζεται με τη διηλεκτρική σταθερά του κενού (ϵ_0) και την επιδεκτικότητα (χ_{medium}) του μέσου.

Η μη γραμμική απόκριση του μέσου έχει την παρακάτω εξάρτηση από το ηλεκτρικό πεδίο:

$$\vec{P} = \epsilon_0(\chi^{(1)} \cdot \vec{E} + \chi^{(2)} : \vec{E}\vec{E} + \chi^{(3)} : \vec{E}\vec{E}\vec{E} + \dots) \quad (2.11)$$

Ο όρος $\chi^{(j)}$ καλείται επιδεκτικότητα j τάξης και είναι ένας πίνακας διαστάσεων $(j+1) \times (j+1)$. Εξαιτίας της συμμετρίας του μορίου SiO_2 ο όρος $\chi^{(2)}$ είναι μηδενικός οπότε στις οπτικές ίνες ο κύριος μη γραμμικός όρος που απομένει για την περιγραφή των μη γραμμικών φαινομένων είναι ο $\chi^{(3)}$. [59]

Στην περίπτωση των κλασικών οπτικών ινών που ο όρος $\chi^{(3)}$ γίνεται σημαντικός, η έκφραση του δείκτη διάθλασης τροποποιείται. Έτσι αν ένα κύμα γραμμικά πολωμένο στον x -άξονα προσπίπτει στην ίνα αποδεικνύεται ότι ο δείκτης διάθλασης είναι:

$$n(\omega, |E|^2) = n(\omega) + \frac{3\chi^{(3)}}{8n(\omega)} |E|^2 \quad (2.12)$$

όπου $n(\omega)$ είναι η συνήθης γραμμική συνιστώσα του δείκτη διάθλασης που εξαρτάται από τη συχνότητα του οπτικού κύματος. Η ποσότητα που δίνεται από το δεύτερο κλάσμα της σχέσης (2.12) καλείται συντελεστής μη γραμμικότητας του δείκτη διάθλασης n_2 και δίνεται από την έκφραση:

$$n_2 = \frac{3\chi^{(3)}}{8n(\omega)} \quad (2.13)$$

Η εξάρτηση του δείκτη διάθλασης του μέσου διάδοσης από την ισχύ $|E|^2$ του διαδιδόμενου κύματος περιγράφει το φαινόμενο Kerr.

Έπειτα από διάδοση μήκους L στην ίνα, το μη γραμμικό οπτικό φαινόμενο Kerr θα έχει ως αποτέλεσμα την μετατόπιση (ολίσθηση) φάσης, που αν αθροιστεί με την μετατόπιση λόγω γραμμικής διάδοσης, οδηγεί σε μια συνολική ολίσθηση φάσης που δίνεται από τον τύπο:

$$\varphi = k_0 n_0 L + \varphi_{NL} \quad (2.14)$$

όπου $k_0 = 2\pi/\lambda_0$ είναι ο κυματαριθμός στον ελεύθερο χώρο, λ_0 το οπτικό μήκος κύματος και

$$\varphi_{NL}(t, z) = j\gamma \int_0^z |E(t, z')|^2 dz' = j\gamma |E(t, 0)|^2 L_{eff} \quad (2.15)$$

το $L_{eff} = \frac{1-e^{-\alpha z}}{\alpha}$ το ενεργό μήκος της οπτικής ίνας και α οι απώλειές της.

Ο συντελεστής γ αποτελεί το μέτρο της μη γραμμικότητας και δίνεται από την έκφραση :

$$\gamma = \frac{k_0 n_0}{A_{eff}} \quad (2.16)$$

όπου A_{eff} η ενεργός περιοχή της οπτικής ίνας.

Η διάδοση σημάτων πεπερασμένου εύρους ζώνης, ειδικά σε μεγάλες αποστάσεις, απαιτεί μεγαλύτερα επίπεδα οπτική ισχύος $|E(t, z)|^2$, γεγονός που προσδίδει στα σήματα μη γραμμική παραμόρφωση εξαρτώμενη από την ισχύ. Αυτό καθορίζει ένα όριο μη γραμμικότητας Kerr, που μερικές φορές αναφέρεται ως μη γραμμικό όριο Shannon (Shannon limit) που περιορίζει τη χωρητικότητα του καναλιού, μιας και η μη γραμμική παρεμβολή nonlinear interference (NLI) είναι θόρυβος. [59]

Η διάδοση σημάτων κατά μήκος μιας οπτικής ίνας περιγράφεται από τη μη γραμμική εξίσωση του Schrödinger, (Non-Linear Schrödinger Equation – NLSE), η οποία στην περίπτωση της μίας μόνο πόλωσης, δίνεται από τη σχέση :

$$\frac{\partial E(t, z)}{\partial z} = -\frac{\alpha}{2} E(t, z) - \frac{i}{2} \beta_2 \frac{\partial^2 E(t, z)}{\partial t^2} + i\gamma |E(t, z)|^2 E(t, z) \quad (2.17)$$

όπου $E(t, z)$ είναι το οπτικό πεδίο συναρτήσει του χρόνου t και της απόστασης z , α , β_2 και γ είναι η διασπορά, η εξασθένιση και ο συντελεστής μη γραμμικότητας της ίνας αντίστοιχα.

Στην περίπτωση των δύο οπτικών πολώσεων η NLSE μπορεί να γενικευθεί διαμορφώνοντας την εξίσωση Manakov [67]:

$$\frac{\partial E(t,z)}{\partial z} = -\frac{a}{2} E(t,z) - i\frac{\beta_2}{2} \frac{\partial^2 E(t,z)}{\partial^2 t} + i\frac{8}{9} \gamma |E(t,z)|^2 E(t,z) \quad (2.18)$$

στην οποία τώρα το οπτικό πεδίο $E(t,z)$ περιλαμβάνει και τις δύο πολώσεις.

Για την ερμηνεία των φαινομένων διάδοσης μπορούμε να διαχωρίσουμε την εξίσωση Manakov μηδενίζοντας διαδοχικά τις παραμέτρους β_2 και γ .

$$\text{Για } \beta_2=0 \quad \frac{\partial E(t,z)}{\partial z} = -\frac{a}{2} E(t,z) - i\frac{8}{9} \gamma |E(t,z)|^2 E(t,z) \quad (2.19)$$

της οποίας η λύση στο πεδίο του χρόνου δίνεται από τη σχέση :

$$E(t,z) = E(t,0) \exp(j\varphi_{NL}(t,z)) \quad (2.20)$$

Όπου φ_{NL} η μη γραμμική ολίσθηση φάσης.

$$\text{Για } \gamma=0 \quad \frac{\partial E(t,z)}{\partial z} = -\frac{a}{2} E(t,z) - i\frac{\beta_2}{2} \frac{\partial^2 E(t,z)}{\partial^2 t} \quad (2.21)$$

της οποίας η λύση στο πεδίο των συχνοτήτων δίνεται από τη σχέση :

$$E(\omega,z) = E(\omega,0) \times \exp\left(-\left(\frac{a}{2} z - \frac{i\beta_2}{2} \omega^2\right) z\right) \quad (2.22)$$

ενώ αντιστοιχεί σε διεύρυνση του οπτικού παλμού στο πεδίο του χρόνου ανάλογη του τετραγώνου του οπτικού εύρους ζώνης.

Ο μη γραμμικός όρος των εξισώσεων 2.17 και 2.18 είναι υπεύθυνος για το φαινόμενο Kerr. Οι εξισώσεις 2.20 και 2.22 εξηγούν τη δημιουργία μη γραμμικής παραμόρφωσης και παρεμβολής, εντείνονται δε καθώς η ισχύς του οπτικού πεδίου αυξάνεται. Το φαινόμενο Kerr θεωρείται το σημαντικότερο εμπόδιο στην αξιόπιστη μεταφορά δεδομένων σε συστήματα οπτικών ινών, ιδιαίτερα με την παρουσία θορύβου των οπτικών ενισχυτών που αθροίζεται με τη μη γραμμική παραμόρφωση.

Όπως είδαμε στο κεφάλαιο 1, τα υψίρρυθμα συστήματα οπτικών επικοινωνιών χρησιμοποιούν τεχνικές διαίρεσης μήκους κύματος (Wavelength Division Multiplexing – WDM), μεταφέροντας πολλά κανάλια πληροφορίας στην ίδια οπτική ίνα, με μία απόσταση μεταξύ τους ώστε να αποφεύγεται παρεμβολή (crosstalk). Στην περίπτωση που η μη γραμμικότητα του φαινομένου Kerr οφείλεται στην αλληλεπίδραση μεταξύ γειτονικών

καναλιών, ονομάζεται διακαναλική (interchannel) ενώ αν είναι ενδογενούς φύσης, ενδοκαναλική (intrachannel).

Με σκοπό να ταυτοποιήσουμε τα διαφορετικά αυτά φαινόμενα υποθέτουμε πως το πεδίο τριών γειτονικών καναλιών διαφορετικού μήκους κύματος που διαδίδονται ταυτόχρονα σε μία πόλωση δίνεται από τη σχέση :

$$E(t, z) = E_1(t, z) + E_2(t, z) + E_3(t, z) \quad (2.23)$$

όπου E_1, E_2, E_3 τα επιμέρους πεδία των τριών μηκών κύματος. Θεωρώντας πως το E_1 είναι το κανάλι ενδιαφέροντος, η 2.18 αφαιρώντας τον όρο της εξασθένισης, γίνεται:

$$\frac{\partial(E_1+E_2+E_3)}{\partial z} = i \frac{\beta_2}{2} \frac{\partial^2(E_1+E_2+E_3)}{\partial^2 t} + j\gamma|E_1 + E_2 + E_3|^2(E_1 + E_2 + E_3) \quad (2.24)$$

και θεωρώντας μικρή διαταραχή (perturbation method) [59] μπορεί ως προς E_1 να γραφεί:

$$\frac{\partial(E_1)}{\partial z} = i \frac{\beta_2}{2} \frac{\partial^2(E_1)}{\partial^2 t} + j\gamma|E_1|^2 E_1 + 2j\gamma(|E_2|^2 + |E_3|^2)E_1 + j\gamma E_2^2 E_3^* \quad (2.25)$$

Η τελευταία σχέση αποτελείται από τρεις όρους που περιέχουν τον συντελεστή μη γραμμικότητας γ καθένας από τους οποίους υποδηλώνει μια επίδραση στη μεταβολή της φάσης :

1. Ο πρώτος όρος εξαρτάται μόνο από την ισχύ (τετράγωνο του ηλεκτρικού πεδίου) του καναλιού ενδιαφέροντος, συνεπώς περιγράφει την αυτό-διαμόρφωση φάσης (Self Phase Modulation – SPM)
2. Ο δεύτερος όρος εξαρτάται μόνο από την ισχύ των γειτονικών καναλιών, συνεπώς περιγράφει την ετερο-διαμόρφωση φάσης (Cross Phase Modulation – XPM)
3. Ο τρίτος όρος εξαρτάται από την αλληλεπίδραση του (συζυγούς ή παράγωγου) πεδίου των γειτονικών καναλιών και περιγράφει το φαινόμενο της μίξης τεσσάρων κυμάτων (Four wave mixing, FWM).

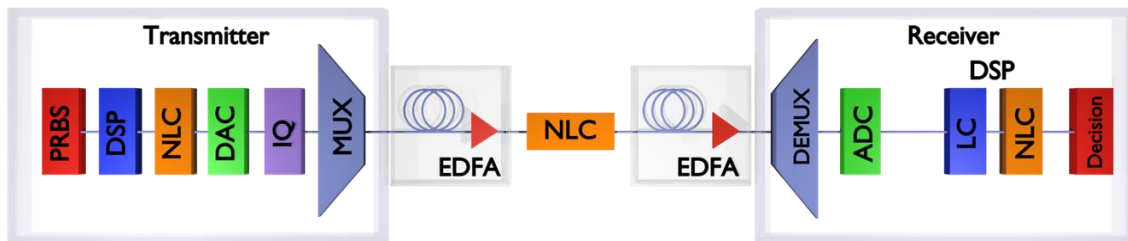
Συνοπτικά μπορούμε να πούμε ότι κατά τη διαδικασία FWM, τρία φωτόνια από κύματα σε συχνότητες $\omega_i, \omega_j, \omega_k$ κάτω από ορισμένες συνθήκες συμβάλλουν, γεννώντας ένα φωτόνιο συχνότητας $\omega_i + \omega_j \pm \omega_k$. Οι ειδικές συνθήκες που απαιτούνται ονομάζονται

συνθήκες προσαρμογής φάσης και είναι το αντίστοιχο της διατήρησης της ορμής. Πάντως για το σύνολο των φωτονίων πριν και μετά τη διαδικασία FWM ισχύει τόσο η διατήρηση της ενέργειας όσο και η διατήρηση της ορμής κύματος. Η διαδικασία FWM είναι επιβλαβής σε συστήματα διάδοσης WDM, διότι υπό κατάλληλες συνθήκες η μη-γραμμική αλληλεπίδραση άνω των τριών μηκών κύματος οδηγεί στη γένεση δικών τους παραγώγων σε μήκη κύματος άλλων σημάτων, όπου η αλληλεπίδραση των ω_2 , ω_3 , ω_4 οδηγεί στην παραγωγή φωτονίων που συμπίπτουν συχνοτικά με το ω_1 , και άρα λειτουργούν ως θόρυβος πάνω σε αυτό το σήμα.

Η πολυπλεξία διαίρεσης πόλωσης υιοθετείται στα συστήματα οπτικών επικοινωνιών λόγω της βελτίωσης της φασματικής απόδοσης. Συνίσταται στη διάδοση του σήματος και στις δύο ορθογώνιες καταστάσεις πόλωσης του μήκους κύματος (States of Polarization - SOP). Στο WDM σύστημα, εμφανίζεται το φαινόμενο της διαμόρφωσης διασταυρωμένης πόλωσης (Cross-polarization modulation - XPoIM) όταν το SOP ενός εκπεμπόμενου καναλιού εξαρτάται από το SOP άλλων καναλιών που έχουν τυχαία διάδοση εντός της οπτικής ίνας λόγω της PMD. Η XPoIM έχει ως αποτέλεσμα την αποπόλωση (depolarization) του μεταδιδόμενου σήματος, το οποίο προκαλεί εξασθένιση και αλληλεπιδράσεις καναλιών για συστήματα διπλής πόλωσης. Η XPoIM μπορεί να κυριαρχεί του XPM και μπορεί να θεωρηθεί ως ένας πρόσθετος Gaussian θόρυβος [68].

2.3.1 Τεχνικές αντιμετώπισης των μη γραμμικών φαινομένων

Η εξάλειψη των μη γραμμικών φαινομένων αποτελεί πεδίο έντονου ερευνητικού ενδιαφέροντος με σκοπό την αύξηση της χωρητικότητας η οποία σε ένα γραμμικό σύστημα αυξάνεται με την αύξηση της εκπεμπόμενης ισχύος. Οι διάφορες τεχνικές μπορούν να υλοποιηθούν στο οπτικό ή στο ψηφιακό πεδίο. Κάποιες από αυτές στον πομπό (transmitter), κάποιες άλλες στο οπτικό link οι περισσότερες όμως υλοποιούνται ψηφιακά στον δέκτη (receiver). Στο σχήμα 2.5 απεικονίζεται ένα ενδεικτικό διάγραμμα συστήματος διάδοσης με τις πιθανές θέσεις διατάξεων αντιστάθμισης μη γραμμικών φαινομένων (nonlinearity compensation - NLC).



Σχήμα 2.5: Διάγραμμα βαθμίδων διάδοσης: PRBS: Ψευδοτυχαίες δυαδικές ακολουθίες, DSP: Ψηφιακή επεξεργασία σήματος, NLC: αντιστάθμιση μη γραμμικών φαινομένων, DAC: μετατροπείας ψηφιακού σε αναλογικό σήμα, IQ: Διαμορφωτής σε φάση και τετραγωνισμού, Mux: Πολυπλέκτης, Demux: Αποπολυπλέκτης, CR: σύμφωνος δέκτης, ADC: Μετατροπείας αναλογικού σε ψηφιακό, Transmitter: πομπός, Receiver: δέκτης

2.3.1.2 Ψηφιακή αντίστροφη διάδοση (Digital back-propagation - DBP)

Η προσέγγιση της αντίστροφης διάδοσης αποτελεί μια από τις πιο αποτελεσματικές τεχνικές αντιστάθμισης τόσο των γραμμικών όσο και των μη γραμμικών φαινομένων καθώς προσομοιώνει σχεδόν τέλεια το οπτικό κανάλι, με εξαίρεση τις αλληλεπιδράσεις σήματος-θορύβου και τη διασπορά τρόπων πόλωσης. Υλοποιείται δε στο ψηφιακό πεδίο, μπορεί να εφαρμοστεί στον πομπό [69], συνήθως όμως εφαρμόζεται στον δέκτη [30].

Ο σκοπός της DBP είναι η επίλυση της αντίστροφης εξίσωσης Manakov, όπως παρουσιάστηκε στις προηγούμενες παραγράφους, εφαρμόζοντας αντίθετες οπτικές παραμέτρους από αυτές της πραγματικής διάδοσης. Με αυτό τον τρόπο προσομοιώνεται

η αντίστροφη διάδοση συνεπώς η αντιστάθμιση της διασποράς και η άρση των μη γραμμικών φαινομένων.

Η επίλυσή της πραγματοποιείται με τη βοήθεια της μεθόδου Split-Step Fourier (SSFM) [70]. Η μέθοδος SSFM είναι αριθμητική τεχνική επίλυσης της μη γραμμικής εξίσωσης Schrödinger (NLSE). Διαχωρίζει τα γραμμικά και μη γραμμικά μέρη της NLSE και χρησιμοποιεί τον γρήγορο μετασχηματισμό Fourier (Fast Fourier transform - FFT) για την επίλυση του γραμμικού μέρους στον πεδίο της συχνότητας, ενώ το μη γραμμικό μέρος στο πεδίο του χρόνου μέσω ενός μη γραμμικού τελεστή. Η NLSE για το πεδίο μπορεί να γραφτεί ως εξής:

$$\frac{\partial E}{\partial z} = (\hat{D} + \hat{N})E \quad (2.26)$$

Όπου \hat{D} είναι ένας διαφορικός τελεστής που λαμβάνει υπόψη τη διασπορά και τις απώλειες και \hat{N} ένας μη γραμμικός τελεστής που περιλαμβάνει τα μη γραμμικά φαινόμενα της οπτικής διάδοσης. Δίνονται από τις σχέσεις :

$$\hat{D} = -\frac{i\beta_2}{2} \frac{\partial^2}{\partial t^2} + \frac{\beta_3}{6} \frac{\partial^3}{\partial t^3} - \frac{\alpha}{2} \quad (2.27)$$

$$\hat{N} = i\gamma \left(|E|^2 + \frac{i}{\omega_0} \frac{1}{E} \frac{\partial}{\partial t} (|E|^2 E) - t \frac{\partial |E|^2}{\partial t} \right) \quad (2.28)$$

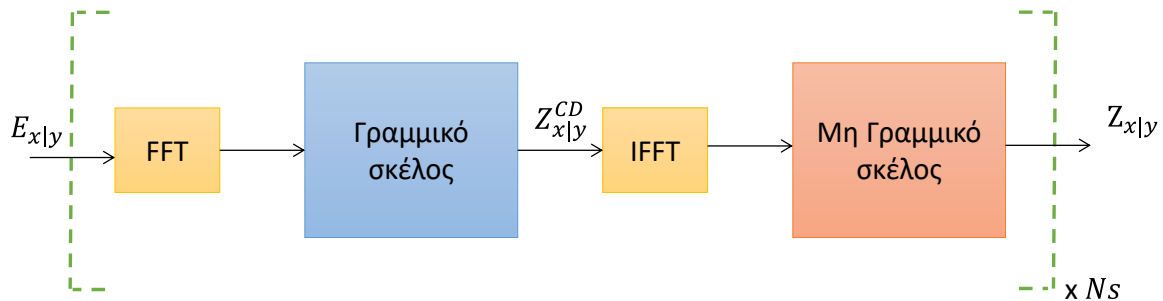
Όπου α ο συντελεστής εξασθένισης και β_2 ο συντελεστής διασποράς 2^{ης} τάξης, t ο χρόνος και ω_0 η συχνότητα. Γενικά, η διασπορά και η μη γραμμικότητα δρουν μαζί κατά μήκος μιας ίνας. Η μέθοδος SSFM λαμβάνει μια κατά προσέγγιση λύση υποθέτοντας ότι η διασπορά και τα μη γραμμικά φαινόμενα μπορούν να θεωρηθούν ότι δρουν ανεξάρτητα για τη διάδοση του οπτικού πεδίου σε μια μικρή απόσταση h . Πιο συγκεκριμένα, η διάδοση από το z στο $z + h$ πραγματοποιείται σε δύο βήματα. Στο πρώτο βήμα, η μη γραμμικότητα ενεργεί μόνη της, και $\hat{D} = 0$ στην εξίσωση (2.26). Στο δεύτερο βήμα, η διασπορά δρα μόνη της, και $\hat{N} = 0$ στην εξίσωση (2.26). Ισχύει προσεγγιστικά η σχέση :

$$E(z + h, T) \approx \exp(h\hat{D}) \exp(h\hat{N})E(z, T) \quad (2.29)$$

Ο εκθετικός όρος $\exp(h\hat{D})$ μπορεί να υπολογιστεί στο επίπεδο του Fourier σύμφωνα με τη σχέση:

$$\exp(h\hat{D})E(z, T) = F_T^{-1} \exp[hD(\omega)]F_T E(z, T) \quad (2.30)$$

όπου F_T υποδηλώνει το μετασχηματισμό Fourier και ω η συχνότητα. [59]



Σχήμα 2.6: Αρχή λειτουργίας DBP

Για τον υπολογισμό της DBP με τη βοήθεια της μεθόδου SSFT, το μήκος της ίνας διαιρείται σε μικρά βήματα καθένα από τα οποία μοντελοποιείται ως η συνένωση των γραμμικών και των μη γραμμικών τμημάτων σύμφωνα με το σχήμα 2.6

Η αντιστάθμιση της διασποράς υλοποιείται στο πεδίο των συχνοτήτων σύμφωνα με τη σχέση :

$$Z_{x|y}^{CD}(\omega, z) = E_{x|y}(\omega, z) e^{-jh\left(\frac{\alpha}{2} + \frac{\beta_2}{z}\omega^2\right)} \quad (2.31)$$

όπου h το μήκος της σχέσης 2.29, Z^{CD} η έξοδος του γραμμικού σκέλους για τις x και y πολώσεις,. Ο εκθετικός όρος αντιπροσωπεύει την αναστροφή του σήματος λόγω της διασποράς. Έπειτα λαμβάνει χώρα η αντιστάθμιση των μη γραμμικών φαινομένων. Ο γρήγορος μετασχηματισμός Fourier (FFT) και ο αντίστροφος FFT (Inverse FFT- IFFT) χρησιμοποιούνται για την εναλλαγή από το πεδίο του χρόνου στο πεδίο των συχνοτήτων και αντίστροφα. Η έξοδος του DBP δίνεται από τη σχέση :

$$Z_{x|y}(t, z) = Z_{x|y}^{CD}(t, z) e^{-j\varphi\gamma'h(|z_x^{CD}|^2 + |z_y^{CD}|^2)} \quad (2.32)$$

όπου $0 < \varphi < 1$ είναι παράγοντας βελτιστοποίησης, $\gamma' = 8/9\gamma$ ο συντελεστής μη γραμμικότητας όπως ισχύει στην εξίσωση Manakov. Από την παραπάνω σχέση φαίνεται πως εκτός της αλλαγής στη φάση λόγω της αυτοδιαμόρφωσης της πόλωσης, προκύπτει

επιπλέον μια μη γραμμική αλλαγή της φάσης στην πόλωση y εξαιτίας του σήματος στην πόλωση x και αντίστροφα.

Όπως φαίνεται και στο σχήμα 2.6 η εκτέλεση της DBP μπορεί να πραγματοποιηθεί σε ακέραια βήματα $Ns \geq 1$. Η ακρίβεια της τεχνικής μπορεί να γίνει αρκετά υψηλή ιδιαίτερα για μεγάλο αριθμό βημάτων. Ωστόσο, η εφαρμογή της σε συστήματα πραγματικού χρόνου εξακολουθεί να παραμένει μη πρακτική λόγω της υψηλής πολυπλοκότητάς της, ειδικά όταν η τεχνική DBP προσπαθεί να προσομοιώσει ένα σενάριο διάδοσης πολλαπλών καναλιών, όπως σε συστήματα WDM [71].

2.3.1.2 Μη γραμμικός ισοσταθμιστής βασισμένος στη σειρά Volterra

Τα μη γραμμικά φαινόμενα που αναπτύσσονται κατά τη διάδοση στην ίνα, μπορούν να μοντελοποιηθούν έχοντας ως βάση τη συνάρτηση μεταφοράς των σειρών Volterra, (Volterra series transfer function - VSFT), επιλύοντας τη NLSE (ή αντίστοιχα την εξίσωση Manakov στην περίπτωση των δύο πολώσεων). Με τη μέθοδο p -οστής τάξης του Schetzen [72] μπορεί να υπολογιστεί ο πυρήνας αντίστροφης συνάρτησης μεταφοράς των σειρών Volterra (IVSFT) συναρτήσει της VSFT. Οι πυρήνες IVSFT χαρακτηρίζουν το μη γραμμικό ισοσταθμιστή βασιζόμενο στις σειρές Volterra (Volterra series based nonlinear equalizer - VNLE) που καταστέλλει τόσο τις μη γραμμικότητες διάδοσης όσο και τη διασπορά. Κατά αντιστοιχία με την τεχνική DBP και η VNLE προσπαθεί να επιτύχει την αντιστροφή του καναλιού.

Εφαρμόζοντας τη μέθοδο p -οστής τάξης ως τον τρίτο βαθμό, οι αντίστροφοι τελεστές Volterra K_1 και K_3 υπολογίζονται από τους τελεστές Volterra H_1 και H_3 από τις σχέσεις :

$$K_1 = H_1^{-1} \quad (2.33)$$

$$K_3 = K_1 H_3 K_1 \quad (2.34)$$

Οι πυρήνες IVSFT περιπτώσεως βαθμού (οι πυρήνες άρτιου βαθμού μηδενίζονται εξαιτίας της των ιστροπικών ιδιοτήτων του διοξειδίου του πυριτίου) υπολογίζονται από τις σχέσεις :

$$k_1(\omega) = e^{\frac{j\omega^2 \beta_2 NL}{2}} \quad (2.35)$$

$$k_3(\omega_1, \omega_2, \omega - \omega_1 + \omega_2) = \frac{jk_1(\omega)}{4\pi^2} \sum_{k=1}^N e^{jk\beta_2\Delta\Omega L} \quad (2.36)$$

Όπου k_1 και k_3 οι IVSFT πυρήνες πρώτου και τρίτου βαθμού, ω η οπτική συχνότητα, L το μήκος της ίνας και $\Delta\Omega$ το βήμα μεταξύ διακριτών συχνοτήτων στο φάσμα. Η παράμετρος c δίνεται από τη σχέση $c=\gamma'L_{eff}$, όπου L_{eff} το ενεργό μήκος της ίνας.

Η έξοδος του ισοσταθμιστή VNLE μπορεί να γραφτεί ως εξής :

$$Z_{x|y}(\omega) = Z_{x|y}^0(\omega) + \sum_{k=1}^N Z_{x|y}^k(\omega) \quad (2.37)$$

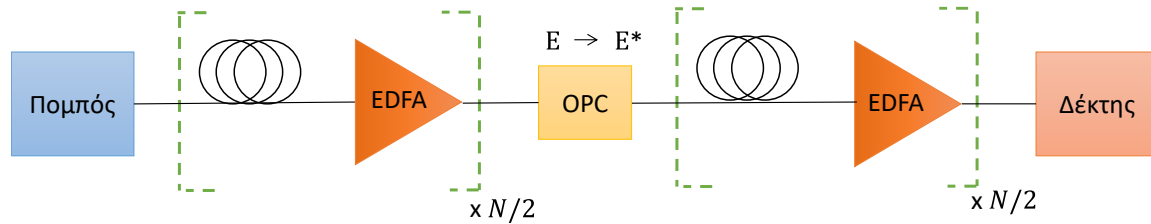
όπου ο πρώτος όρος αναφέρεται στο γραμμικό μέρος και ο δεύτερος στο μη γραμμικό.

Ο VNLE έχει δείξει υψηλή απόδοση στην καταπολέμηση μη γραμμικών επιδράσεων για συστήματα μετάδοσης ενός καναλιού και απαιτεί περίπου το ήμισυ του υπολογιστικού φόρτου της μεθόδου DBP [73], ενώ έχουν προταθεί νέες προσεγγίσεις για την περαιτέρω μείωση της πολυπλοκότητας του VNLE [74].

Η συνηθέστερη εφαρμογή του βασίζεται στη σειρά Volterra τρίτης τάξης ενώ έχουν προταθεί ισοσταθμιστές με σειρά Volterra πέμπτης τάξης [75]. Ενώ η πρόταση αυτή παρουσιάζει καλύτερη απόδοση σε σύστημα μονού καναλιού, αυξάνει την πολυπλοκότητα της υλοποίησης σε σύγκριση με την περίπτωση τρίτης τάξης.

2.3.1.3 Οπτική συζυγία φάσης (Optical Phase Conjugation -OPC)

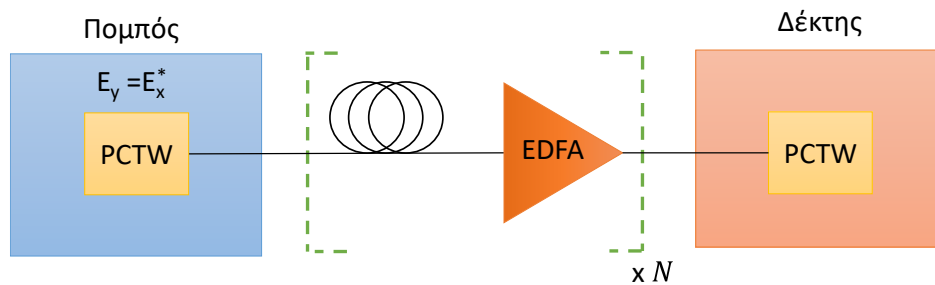
Η τεχνική OPC είναι μια καθαρά αναλογική και επομένως εξαιρετικά γρήγορη τεχνική και υλοποιείται με την αντιστροφή του φάσματος του διαδιδόμενου σήματος στο μέσο της οπτικής ζεύξης όπως φαίνεται στο σχήμα 2.8. [76], [77]. Η κεντρική ιδέα είναι η ακύρωση όλων των αλλαγών που προκλήθηκαν στη φάση του σήματος (γραμμικών λόγω διασποράς και μη-γραμμικών λόγω φαινομένου Kerr) στο πρώτο τμήμα της ζεύξης μέσω της διάδοσης στο δεύτερο τμήμα αφού πρώτα έχει προκληθεί η συζυγία της φάσης. Συνεπώς όποια αλλαγή στη φάση λαμβάνει χώρα στο πρώτο τμήμα, μέσω της συζυγίας εξανεμίζεται στο δεύτερο, αρκεί τα δύο τμήματα να επάγουν τις ίδιες γραμμικές και μη-γραμμικές μεταβολές στη φάση του σήματος.



Σχήμα 2.7: Αρχή λειτουργίας του OPC

Η τεχνική OPC αντισταθμίζει τη ντετερμινιστική μη γραμμική μετατόπιση φάσης και το μη γραμμικό θόρυβο φάσης που προκαλείται από την αλληλεπίδραση μεταξύ σήματος και θορύβου. Το κύριο πλεονέκτημα της τεχνικής είναι η χαμηλή πολυπλοκότητα και η αυξημένη φασματική απόδοση, όπως και η εξαιρετική επίδοση στην ταυτόχρονη αντιστάθμιση διασποράς και μη-γραμμικότητας [78]. Από την άλλη πλευρά, η ανάγκη της συζυγίας φάσης να πραγματοποιείται σε συγκεκριμένα σημεία της οπτικής ζεύξης θέτει περιορισμούς στην τοπολογία. Επίσης απαιτεί επιπλέον υλικό, όπως μετατροπείς μήκους χαμηλού θορύβου και ευρείας ζώνης [76].

Η ψηφιακή εκδοχή της OPC είναι η μέθοδος δίδυμων κυμάτων συζευγμένων στη φάση (Phase conjugated twin waves - PCTW) [79]. Πρόκειται για μία τεχνική ψηφιακής επεξεργασίας σήματος που εκτελείται στο δέκτη. Στα συστήματα δύο πολώσεων η PCTW πραγματοποιείται με τη διάδοση του σήματος ενδιαφέροντος στη x πόλωση και του συζυγούς του στη y . Η αρχή λειτουργίας της παρουσιάζεται στο σχήμα 2.8.



Σχήμα 2.8: Αρχή λειτουργίας του OPC

Η μη γραμμική ολίσθηση φάσης μπορεί να ακυρωθεί από την υπέρθεση των δύο σημάτων στον δέκτη. Το σήμα $V_{x|y}$ που φτάνει στο δέκτη μπορεί προσεγγιστικά να εκφραστεί από τη σχέση :

$$V_{x|y} = E_{x|y} + \delta V_{x|y} \quad (2.38)$$

Όπου $\delta V_{x|y}$ είναι ο όρος της μη γραμμικής παραμόρφωσης του διαδιδόμενου σήματος $E_{x|y}$. Δεδομένου του ότι $E_y = E_x^*$, ο όρος της μη γραμμικής παραμόρφωσης μπορεί να εκφραστεί ως

$$\delta V_y = -[\delta V_x]^* \quad (2.39)$$

Συνεπώς η υπέρθεση του σήματος στον δέκτη V_x και του συζυγούς του V_y ακυρώνει τη μη γραμμική ολίσθηση φάσης και το αρχικό σήμα μπορεί να ανακτηθεί ως εξής :

$$\frac{V_x + V_y^*}{2} = E_x \quad (2.40)$$

Η PCTW προϋποθέτει την αντιστάθμιση της διασποράς πριν την εφαρμογή της.

Οι τεχνικές που βασίζονται στη συζυγία φάσης αντισταθμίζουν τη ντετερμινιστική μη γραμμική ολίσθηση φάσης και εν μέρει τον μη γραμμικό θόρυβο φάσης που προκαλείται από την αλληλεπίδραση του σήματος με το θόρυβο. Το κύριο πλεονέκτημά τους είναι η χαμηλή υπολογιστική πολυπλοκότητα. Από την άλλη πλευρά το κύριο μειονέκτημα της PCTW είναι η απώλεια της μισής φασματικής απόδοσης λόγω της διάδοσης του συζυγούς του σήματος στην πόλωση y . Αυτός ο περιορισμός καθιστά στην πράξη μη αποδοτική τη συγκεκριμένη μέθοδο καθώς οι σύγχρονες οπτικές επικοινωνίες βασίζονται στην πλήρη εκμετάλλευση του διαθέσιμου φάσματος. Το πρόβλημα αυτό αντιμετωπίζεται με κάποιες εναλλακτικές προτάσεις εφαρμογής του PCTW σε OFDM συστήματα [80].

2.3.1.4 Μη γραμμική αντιστάθμιση βασισμένη σε Διαταραχή (Perturbation-based Non Linear Compensation, PB-NLC)

Η προσέγγιση των διαταραχών έχει διερευνηθεί για την αντιστάθμιση των μη γραμμικών φαινομένων όπως και για τη μοντελοποίηση της οπτικής ίνας. Μπορεί να εφαρμοστεί στον πομπό ως προ-παραμόρφωση (predistortion) ή στον δέκτη. Παρέχει μια προσεγγιστική αριθμητική λύση της εξίσωσης Manakov (ή αντίστοιχα της NLSE στην περίπτωση της διάδοσης σε μια πόλωση) [81], [82], [83].

Η κεντρική ιδέα της PB-NLC είναι η χρήση της μη γραμμικής παραμόρφωσης ως διαταραγμένης διόρθωσης στη μη διαταραγμένη λύση. Η μη διαταραγμένη λύση λαμβάνει υπόψη μόνο τη γραμμική παραμόρφωση η οποία οφείλεται σύμφωνα με τα προηγούμενα στη διασπορά και τις απώλειες.

Το οπτικό πεδίο, βασιζόμενο στην διαταραχή πρώτου βαθμού μπορεί να γραφτεί ως εξής:

$$E_{x|y} = E_{x|y,0} + \gamma \delta E_{x|y} \quad (2.41)$$

Όπου $E_{x|y,0}$ αντιστοιχεί στη λύση της γραμμικής διάδοσης, $\delta E_{x|y}$ η διαταραχή πρώτου βαθμού, που για κάθε πόλωση μπορεί να αναπαρασταθεί στο πεδίο των συχνοτήτων από τη σχέση :

$$\delta E_{x|y}(\omega, L) = h_{cd}(\omega) \int_0^L F_{x|y}(\omega, z) e^{-\frac{i\omega^2 z}{2}} dz \quad (2.42)$$

Η h_{cd} είναι η συνάρτηση μεταφοράς της αντιστάθμισης της διασποράς, L το μήκος του διαστήματος και η συνάρτηση $F_{x|y}$ δίνεται από τον τύπο:

$$F_{x|y}(\omega, z) = j \frac{8}{9} \int e^{-j\omega t} E_{x|y,0}(t, z) \times [E_{x,0}(t, z)E_{x,0}^*(t, z) + E_{y,0}(t, z)E_{y,0}^*(t, z)] dt \quad (2.43)$$

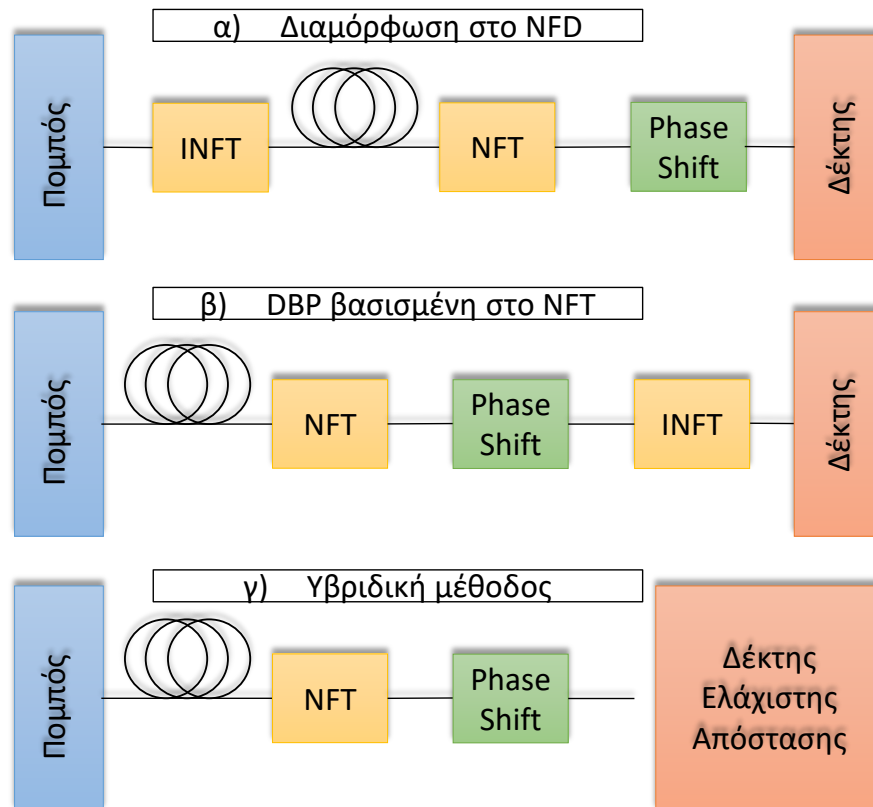
Το βασικό πλεονέκτημα της PB-NLC είναι η δυνατότητα να εφαρμοστεί για όλη τη ζεύξη. Αυτό ισοδυναμεί με την ελαχιστοποίηση της υπολογιστικής πολυπλοκότητας σε σύγκριση με τις τεχνικές DBP και VNLE. Μπορεί επιπλέον να εφαρμοστεί με ένα δείγμα ανά σύμβολο, γεγονός που μειώνει δραστικά τις απαιτήσεις σε ταχύτητα του DAC/ADC.

Από την άλλη μεριά η μέθοδος PB-NLC απαιτεί τον υπολογισμό μεγάλου αριθμού παραμέτρων διαταραχής γεγονός που αυξάνει τη δυσκολία στην πρακτική εφαρμογή της [83].

2.3.1.5 Μη γραμμικός μετασχηματισμός Fourier

Μια μέθοδος που έχει απασχολήσει την ερευνητική κοινότητα τελευταία και θεωρείται ως μια πολλά υποσχόμενη τεχνική που θα χρησιμοποιηθεί για τα μελλοντικά συστήματα

οπτικών επικοινωνιών, είναι η ιδέα της διαμόρφωσης της πληροφορίας στον μη γραμμικό μετασχηματισμό Fourier (Nonlinear Fourier transform - NFT) του σήματος [84].



Σχήμα 2.9: Βασικά σχήματα βασισμένα στην NFT

Η διάδοση του μη γραμμικού φάσματος εντός του πεδίου του μη γραμμικού μετασχηματισμού Fourier γίνεται γραμμική, δίνοντας έτσι τη δυνατότητα της εκμετάλλευσης του φάσματος του NFT για την διαμόρφωση, διάδοση, ανίχνευση και επεξεργασία του σήματος. Υπάρχουν τρία βασικά σχήματα που χρησιμοποιούν NFT και φαίνονται στο σχήμα 2.9.

Στην πρώτη περίπτωση (α) η πληροφορία διαμορφώνεται κατευθείαν στο φάσμα του σήματος NF μέσω του αντίστροφου μη γραμμικού μετασχηματισμού Fourier (Inverse Nonlinear Fourier Transform – INFT). Εδώ μπορούν να διαμορφωθούν διακριτά και συνεχή τμήματα του φάσματος NF ξεχωριστά αλλά και ταυτόχρονα.

Στη δεύτερη περίπτωση της DBP βασιζόμενης στο NFT (β) ο NFT χρησιμοποιείται για να ακυρώσει τη μη γραμμική παραμόρφωση στον δέκτη. Η διαμόρφωση γίνεται όπως τα παραδοσιακά συστήματα οπτικών επικοινωνιών. Η μέθοδος έχει κοινά χαρακτηριστικά, παρουσιάζει παρόμοια απόδοση και έχει συγκρίσιμη πολυπλοκότητα υλοποίησης με το DBP, γεγονός που δικαιολογεί την ονομασία της. Η απαίτηση ακριβών υπολογισμών ενός άγνωστου αριθμού διακριτών ιδιοτιμών στο τυχαία διαμορφωμένο σήμα καθιστά τη λύση αυτή προσωρινά λιγότερο θελκτική.

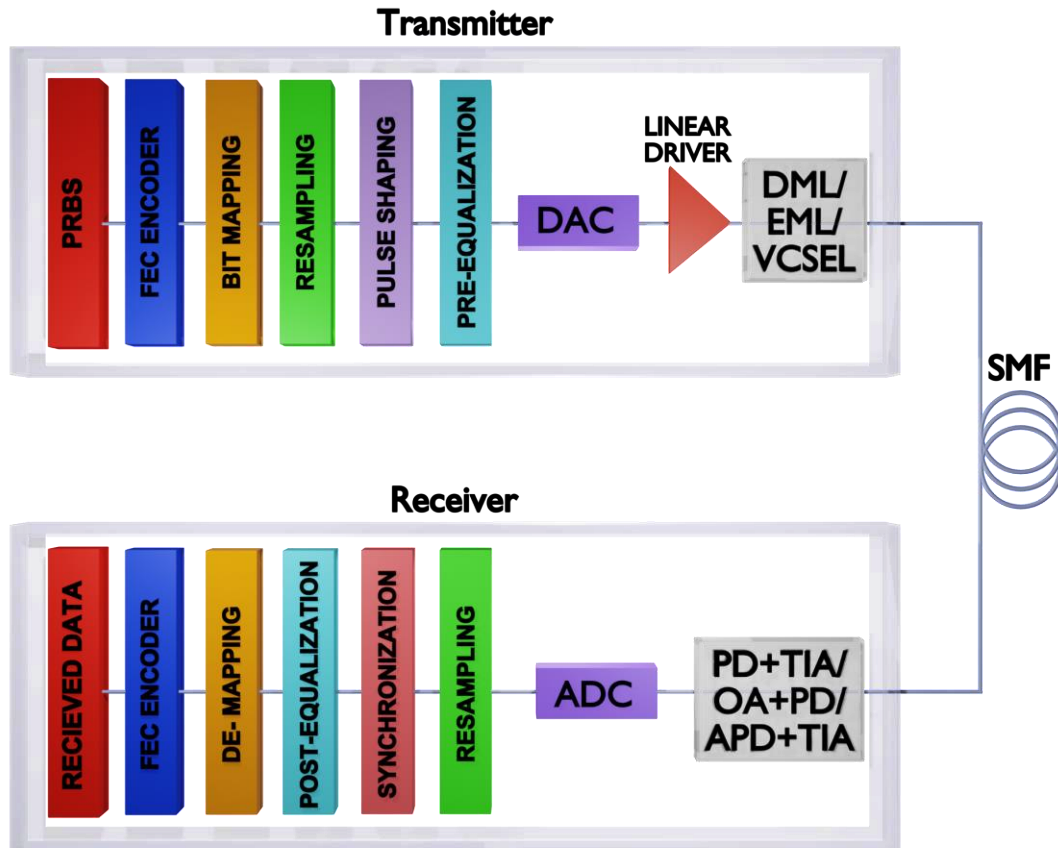
Στην τρίτη περίπτωση που ονομάζεται υβριδική (γ) η διαμόρφωση της πληροφορίας γίνεται στο πεδίο του χρόνου, όμως η ανίχνευση στο δέκτη γίνεται χρησιμοποιώντας τα δεδομένα του NF φάσματος. Καθώς αυτή η προσέγγιση δεν προσφέρει την ευελιξία της βελτιστοποίησης των θέσεων των διακριτών ιδιοτιμών στο μιγαδικό επίπεδο, απαιτεί συνήθως έναν δέκτη ελάχιστης απόστασης. Επιπλέον, εάν οι μεταδιδόμενες κυματομορφές δεν είναι βελτιστοποιημένες για εκπομπές πολλαπλών παλμών, θα πρέπει επίσης να ληφθεί υπόψη το συνεχές τμήμα του μη γραμμικού φάσματος του σήματος. Οι επικοινωνίες που βασίζονται στο NFT δεν επηρεάζονται από κάποια γραμμικά και μη γραμμικά ντετερμινιστικά φαινόμενα, συμπεριλαμβανομένης της διακαναλικής ή ενδοκαναλικής διαφωνίας.

Ωστόσο, τα μοντέλα που βασίζονται σε NFT εξαρτώνται σε κρίσιμο βαθμό από το οπτικό κανάλι, το οποίο μπορεί να διαταραχθεί από τις απώλειες τις ίνας και τις παραμορφώσεις που σχετίζονται με το υλικό, ενώ η απόδοσή τους περιορίζεται από την αλληλεπίδραση μεταξύ σήματος και θορύβου που εισάγουν οι ενισχυτές EDFA. Η προσέγγιση εξακολουθεί να απαιτεί σημαντική προσπάθεια για τη βελτιστοποίησή της όσον αφορά την επίτευξη υψηλότερης απόδοσης και ποιότητας διάδοσης ως ότου χρησιμοποιηθεί σε πραγματικά συστήματα οπτικών επικοινωνιών.

2.4 Συστήματα Οπτικών Επικοινωνιών Διαμόρφωσης Έντασης – Άμεσης Φώρασης

Τα πρώτα συστήματα οπτικών επικοινωνιών εντοπίζονται στις αρχές της δεκαετίας του 1970. Βασίζονταν σε μια απλή τεχνική διάδοσης κατά την οποία χρησιμοποιείται διαμόρφωση της έντασης (Intensity Modulation – IM) του ημιαγωγικού laser. Μετά τη διάδοση στην οπτική ίνα το σήμα ανιχνεύεται από μια φωτοδίοδο. Τα συστήματα που συνδυάζουν τη διαμόρφωση έντασης στον πομπό και την άμεση φώραση (Direct Detection – DD) στο δέκτη ονομάζονται IM/DD συστήματα. Μεγάλο πλεονεκτήματα των IM/DD συστημάτων είναι η απλή κατασκευή τους και η συνεπαγόμενη οικονομία κλίμακας που επιφέρει σε πραγματικά συστήματα μικρών αποστάσεων, αλλά και η μη ευαισθησία του δέκτη στις διακυμάνσεις της φάσης και της πόλωσης [85]. Αρχικά οι οπτικές ζεύξεις επιτυγχάνονταν με διαμόρφωση έντασης με μη επιστροφή στο μηδέν (non-return-to-zero on-off-keying - NRZ-OOK) διαμόρφωση. Με τις απαιτήσεις όμως σε ρυθμούς μεταφοράς να αυξάνονται διαρκώς καθίσταται αναγκαία η υιοθέτηση προηγμένων μορφών διαμόρφωσης που αφενός βελτιώνουν τη φασματικής απόδοσης (spectral efficiency, SE) και αφετέρου μειώνουν τους περιορισμούς του εύρους ζώνης των ηλεκτρονικών και οπτικών στοιχείων [86]. Έτσι, σχήματα διαμόρφωσης όπως πλάτους παλμού τεσσάρων επιπέδων (4 level Pulse-Amplitude Modulation – PAM4), 8 επιπέδων, PAM8, διαμόρφωση πλάτους και φάσης χωρίς φέρουσα (carrier-less amplitude and phase modulation - CAP) και διακριτού πολυτόνου (Discrete Multi-Tone - DMT) δοκιμάστηκαν. Λαμβάνοντας υπόψη την κατανάλωση ενέργειας και την πολυπλοκότητα υλοποίησης, η πιο ελκυστική μορφή σε οπτικές συνδέσεις μικρής απόστασης είναι το PAM4, το οποίο θεωρείται πολλά υποσχόμενο υποψήφιο για συνδέσεις σε κέντρα δεδομένων και για το παθητικό οπτικό δίκτυο (Passive Optical Network- PON) επόμενης γενιάς και έχει επικυρωθεί από το IEEE 802.3 bs για διάδοση στα 400 Gbps [87].

Ένα τυπικό σύστημα για οπτικές συνδέσεις μικρής απόστασης που βασίζεται στη διαμόρφωση PAM4 παρουσιάζεται στο σχήμα 2.10. Τα αρχικά δεδομένα που δημιουργούνται από μια ψευδοτυχαία δυαδική ακολουθία (pseudorandom binary sequence - PRBS) κωδικοποιούνται πρώτα από έναν κωδικοποιητή διόρθωσης



Σχήμα 2.10: Βασικό σχήμα IM-DD συστήματος.

σφαλμάτων (FEC) και στη συνέχεια αντιστοιχίζονται στα σύμβολα PAM4. Μετά την δειγματοληψία και τη σχηματοποίηση παλμού (συνήθως Root Raised Cosine – RRC), το ψηφιακό σήμα υπόκειται σε προ-ισοστάθμιση με τη βοήθεια ενός φίλτρου πεπερασμένης κρουστικής απόκρισης (Finite impulse response - FIR) για να αντισταθμιστεί το περιορισμένο εύρος ζώνης και η μη γραμμικότητα του παλμού. Το επεξεργασμένο σήμα οδηγείται σε ένα DAC προκειμένου να ληφθεί το ηλεκτρικό σήμα στη βασική ζώνη. Η έξοδος του DAC ενισχύεται από ένα γραμμικό ενισχυτή και στη συνέχεια τροφοδοτείται σε έναν οπτικό διαμορφωτή για την παραγωγή οπτικού PAM4 σήματος.

Για να επιτευχθεί η διαμόρφωση έντασης μπορεί είτε να διαμορφωθεί απευθείας το laser με το ρεύμα οδήγησης είτε να χρησιμοποιηθεί ένα laser συνεχούς κύματος (Continuous Wave - CW) που ακολουθείται από έναν εξωτερικό διαμορφωτή έντασης για να

αποσυνδέσει τη δημιουργία φωτός και τους μηχανισμούς διαμόρφωσης. Τα πιο συνηθισμένα laser άμεσης διαμόρφωσης (Directly Modulated Lasers - DMLs) χαμηλού κόστους είναι τα laser κατανεμημένης ανάδρασης (Distributed-feedback laser - DFB), τα laser κατανεμημένου ανακλαστήρα Bragg, (Distributed Bragg Reflector - DBR) και τα, laser επιφανειακής εκπομπής κατακόρυφης κοιλότητας (Vertical-cavity surface-emitting - VCSEL). Στην περίπτωση της εξωτερικής διαμόρφωσης των laser δύο είναι οι συνηθέστεροι τύποι διαμορφωτών έντασης: ο διαμορφωτής ηλεκτροαπορρόφησης (electro-absorption modulator - EAM) και διαμορφωτής Mach-Zehnder (MZM). Συνήθως ο EAM και το laser είναι συχνά ενσωματωμένα στο ίδιο τσιπ, σε μια συσκευή που ονομάζεται διαμορφωμένο laser ηλεκτροαπορρόφησης (EML). Γενικά, οι λύσεις που βασίζονται σε DML και EML είναι φθηνότερες και πιο συμπαγείς από τις λύσεις που βασίζονται σε εξωτερική διαμόρφωση με MZM. Δεδομένου ότι τα ηλεκτρικά σήματα του VCSEL και του DML εφαρμόζονται απευθείας στην κοιλότητα του laser εμφανίζουν πιο ισχυρή διακύμανση της συχνότητας ως αποτέλεσμα της διαμόρφωσης του laser – γνωστό και ως laser chirp -έναντι του EML, η οποία μπορεί να προκαλέσει σοβαρές παραμορφώσεις και να οδηγήσει σε σημαντική υποβάθμιση της απόδοσης [88]. Ο πίνακας 2.1 συνοψίζει τα συστήματα πομπών και τα χαρακτηριστικά τους. Σε αποστάσεις μικρότερες από τα 300 μέτρα ο συνδυασμός VCSEL και πολύτροπης ίνας αποτελεί την επικρατέστερη επιλογή, κυρίως εξαιτίας του χαμηλού κόστους. Όταν οι αποστάσεις μεγαλώνουν η διασπορά τρόπων της πολύτροπης ίνας γίνεται καταστροφική για τη διάδοση του σήματος, οπότε η μονότροπη οπτική ίνα αποτελεί τη βέλτιστη λύση [88].

Πομπός	Εύρος Ζώνης	Κόστος	Απόσταση	Ίνα	Εξασθένιση ισχύος	Μήκος κύματος (nm)
VCSEL	Χαμηλό	Χαμηλό	<300μ	Πολύτροπη	Δ/Υ	850
DML	Χαμηλό	Μέσο	<80χλμ	Μονότροπη	Χαμηλή	1310/1550
EML	Υψηλό	Υψηλό	<80χλμ	Μονότροπη	Υψηλή	1310/1550

Πίνακας 2.1: Σύγκριση τυπικών οπτικών πομπών

Στη μεριά του δέκτη, το σήμα μπορεί να ανιχνευθεί από μία φωτοδίοδο PIN (PIN Photodiode – PIN-PD) ή από μία φωτοδίοδο χιονοστιβάδας (Avalanche photodiode - APD). Για την αύξηση της ευαισθησίας του δέκτη χρησιμοποιείται συνήθως ένας ενισχυτής σύνθετης αντίστασης (transimpedance amplifier – TIA) μετά τη φωτοδίοδο ο οποίος ενισχύει το ηλεκτρικό σήμα. Εναλλακτικά, χρησιμοποιείται ένας οπτικό ενισχυτής και ζωνοπερατό οπτικό φίλτρο πριν τη φωτοδίοδο για να ενισχυθεί το οπτικό σήμα και να αφαιρεθεί ο θόρυβος εκτός της φασματικής περιοχής του αντίστοιχα. Ο πίνακας 2.2 συνοψίζει τα συστήματα δεκτών και τα χαρακτηριστικά τους.

Δέκτης	Κόστος	Απόσταση	Ευαισθησία
PD +TIA	Χαμηλό	Μικρή	Χαμηλή
APD +TIA	Μέσο	Μέση	Μέση
OA + PD	Υψηλό	Μεγάλη	Υψηλή

Πίνακας 2.2: Σύγκριση τυπικών οπτικών δεκτών

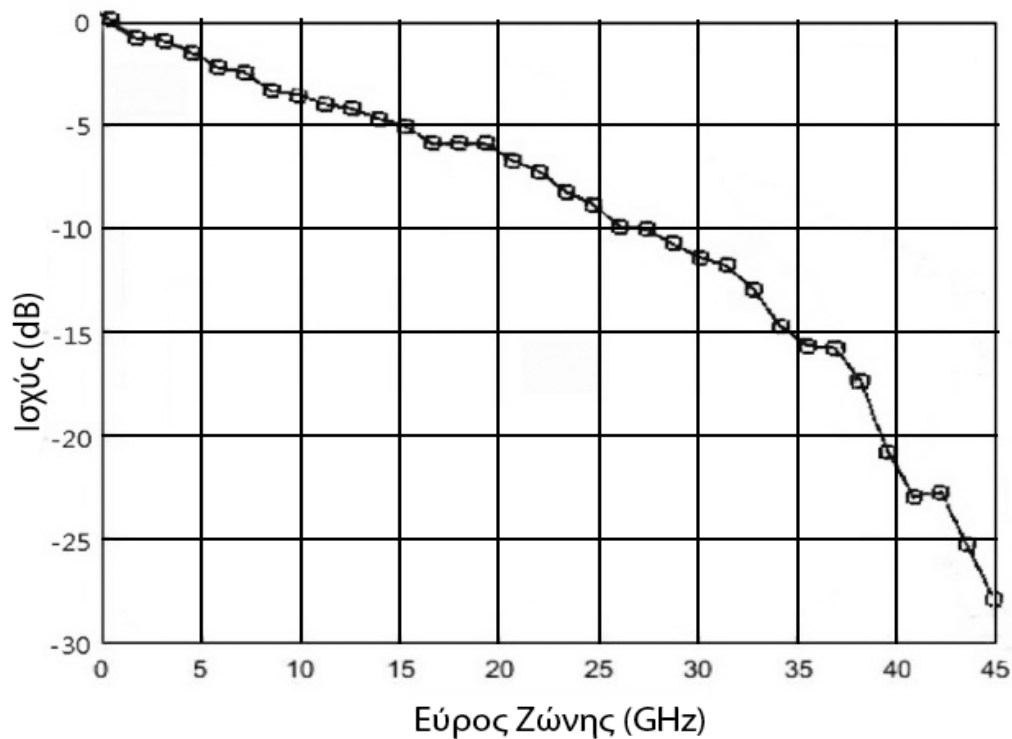
Είναι φανερό πως η λύση των οπτικών ενισχυτών προσφέρεται για συστήματα μεγαλύτερων αποστάσεων και ευαισθησίας, όμως με μεγαλύτερο κόστος. Στις λύσεις των TIA ενισχυτών η περίπτωση της APD προσφέρει μεγαλύτερη ευαισθησία από την απλή φωτοδίοδο [89].

Αφού υποβληθεί σε επεξεργασία από έναν μετατροπέα αναλογικού σε ψηφιακό (analog to digital converter, ADC), το σήμα επαναδειγματοληπτείται και συγχρονίζεται. Στη συνέχεια, χρησιμοποιούνται διάφορες τεχνικές DSP για την εξίσωση του λαμβανόμενου σήματος και τη βελτίωση της απόδοσης του συστήματος. Τέλος, μετά την αποχαρτογράφηση και την αποκωδικοποίηση FEC, λαμβάνονται τα δεδομένα και μπορεί να υπολογιστεί η πιθανότητα σφάλματος (BER).

2.4.1 Επιβαρύνσεις στα IM/DD συστήματα

Όπως είδαμε στα προηγούμενα, τα συστήματα IM/DD αποτελούν τα βέλτιστη επιλογή για επικοινωνίες μικρών αποστάσεων, μικρότερες των 100km κυρίως εξαιτίας της απλότητάς τους και του χαμηλού τους κόστους. Το τελευταίο παρά το γεγονός πως αποτελεί το σημαντικότερο παράγοντα που λαμβάνεται υπόψη σε εμπορικές εφαρμογές μεγάλης

κλίμακας, έχει ως αποτέλεσμα τη χρήση στοιχείων/συσκευών που εισάγουν αυξημένες παραμορφώσεις τόσο γραμμικής όσο και μη γραμμικής φύσης.



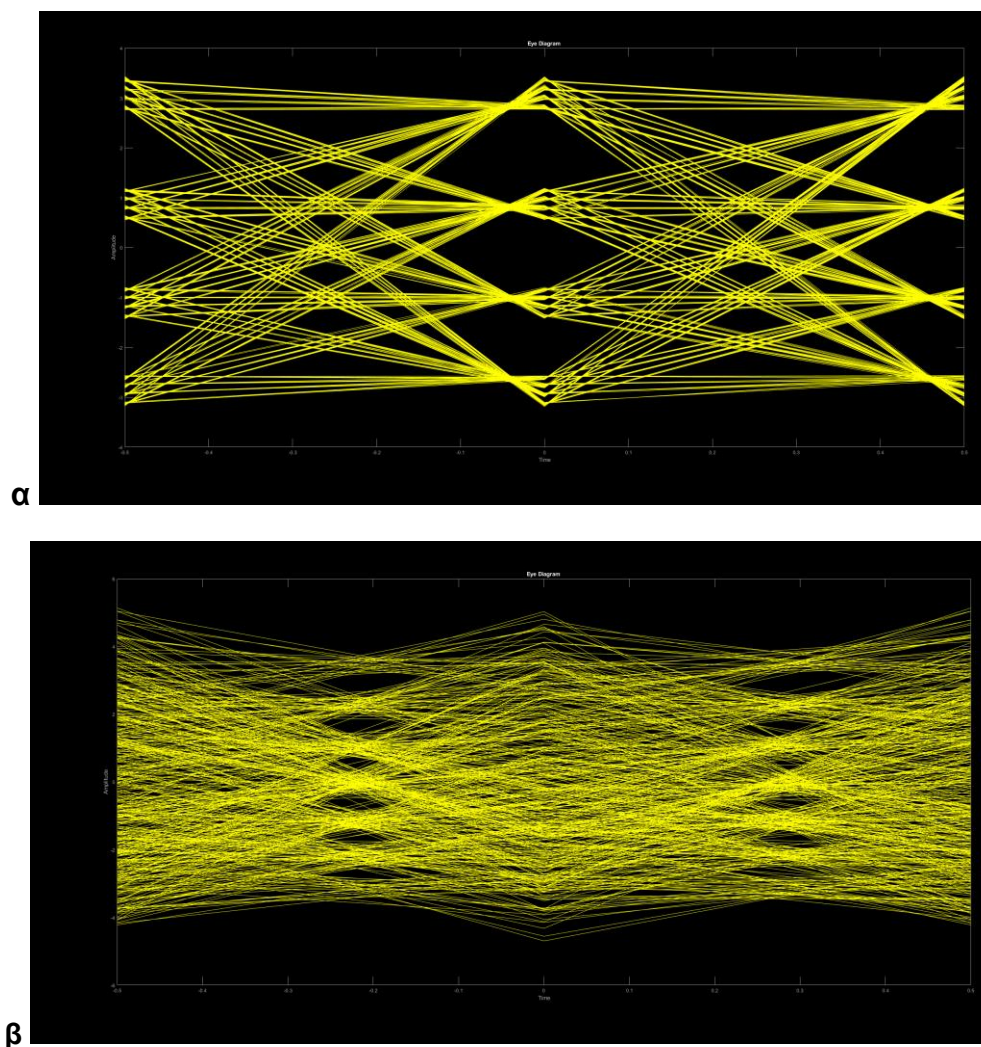
Σχήμα 2.11: Φασματική απόκριση σε σύστημα περιορισμένου εύρους ζώνης

Θεωρώντας ένα τυπικό IM/DD σύστημα που διαμορφώνεται με PAM4 και λαμβάνοντας υπόψη τους περιορισμούς σε εύρος ζώνης εξαιτίας της επιδίωξης του χαμηλού κόστους, μπορούμε να εκφράσουμε το σήμα στον δέκτη από τη σχέση:

$$y_k = I_k + \sum_{n=0}^{\infty} I_k x_{k-n} + v_k \quad (2.44)$$

Όπου x_k, y_k το σήμα στην είσοδο και στην έξοδο του οπτικού συστήματος αντίστοιχα. Το I_k εκφράζει την πληροφορία σε επίπεδο συμβόλων, ο δεύτερος όρος εκφράζει την

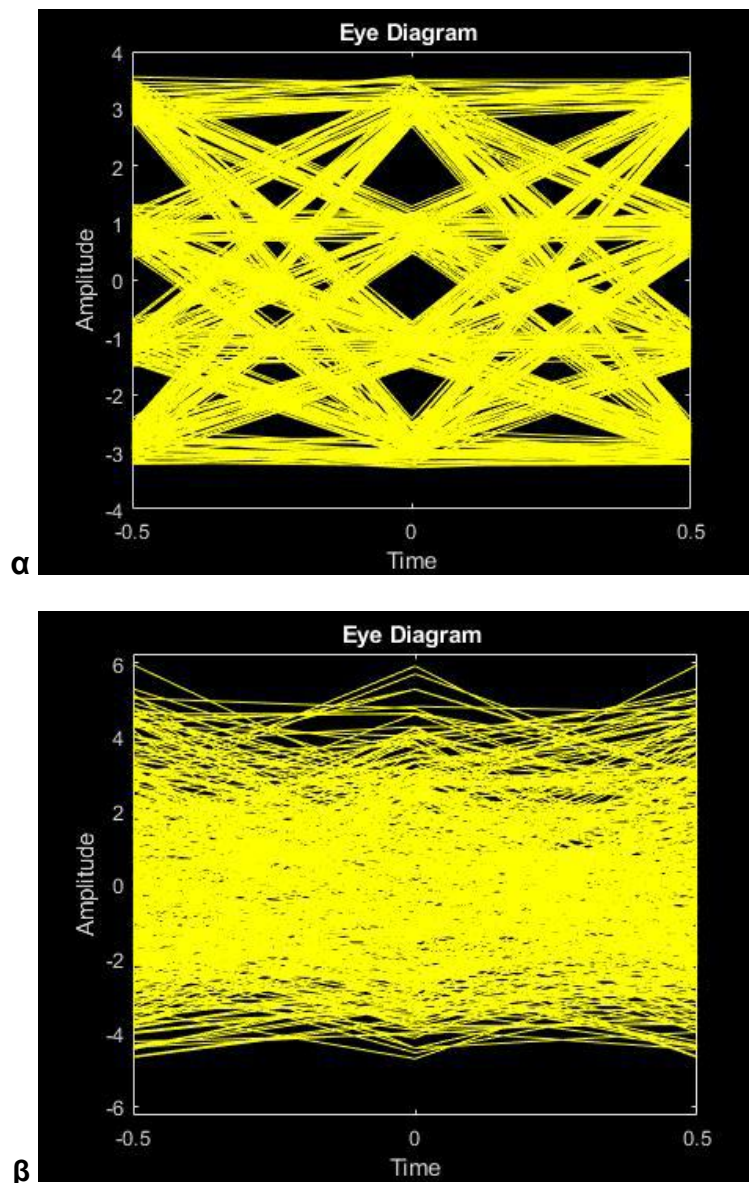
αλληλοπαρεμβολή των συμβόλων (intersymbol interference, ISI) και ο τρίτος όρος v_k εκφράζει το γκαουσιανό θόρυβο.



Σχήμα 2.12: Διάγραμμα οφθαλμού για σύστημα χωρίς (α) και με (β) περιορισμένο εύρος ζώνης

Το περιορισμένο εύρος ζώνης έχει επίδραση στην απόδοση του συστήματος. Στο σχήμα 2.11 απεικονίζεται η απόκριση ενός PAM4, 64GBaud πραγματικού συστήματος [90]. Φαίνεται πως το εύρος ζώνης των 3dB (δηλαδή το φασματικό εύρος υποδιπλασιασμού της ισχύος) είναι μόλις 7.5GHz, τιμή σημαντικά μικρότερη από τις απαιτήσεις του συστήματος σε εύρος ζώνης. Η απομείωση της ισχύος είναι 15 dB στη συχνότητα Nyquist.

Προσομοιώνοντας το ίδιο σύστημα για συνθήκες χωρίς και με περιορισμό του εύρους ζώνης απεικονίζεται το διάγραμμα οφθαλμού στο σχήμα 2.12. Στην πρώτη περίπτωση, όπου δεν υπάρχουν περιορισμοί στο εύρος ζώνης, οι τέσσερις στάθμες (PAM-4) της ισχύος είναι ευδιάκριτες. Αντιθέτως σε σύστημα με περιορισμένο εύρος ζώνης οι στάθμες είναι δυσδιάκριτες, επιδεινώνοντας τις επιδόσεις του.



Σχήμα 2.13: Διάγραμμα οφθαλμού για διάδοση 5km (α) και 10km (β) στη C-band

Σημαντικός παράγοντας που καθορίζει τις επιδόσεις ενός συστήματος IM/DD είναι η διασπορά. Ιδιαίτερα στη περιοχή του μήκους κύματος των 1550nm (C-band) όπου η διασπορά είναι ιδιαίτερα αυξημένη, παρατηρείται σημαντική αλληλοπαρεμβολή συμβόλων ISI όσο το μήκος της οπτικής διάδοσης αυξάνεται. Στο σχήμα 2.13 απεικονίζεται το διάγραμμα οφθαλμού για οπτικό σύστημα PAM4 / 64GBaud στη C-band, για 5km και 10km διάδοσης. Είναι φανερό η σημαντική υποβάθμιση της ποιότητας του συστήματος, απόρροια της διασποράς, όταν το μήκος διάδοσης αυξάνεται.

2.5 Σύμφωνα και μη Σύμφωνα Συστήματα Οπτικών Επικοινωνιών

2.5.1 Δέκτες

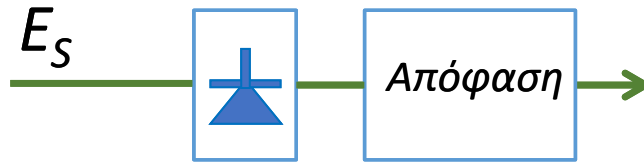
Βασικός στόχος των συστημάτων οπτικών ινών μεγάλων αποστάσεων είναι η μεταφορά δεδομένων χωρίς την ανάγκη αναγέννησης τους μέσω της μετατροπής τους από οπτική σε ηλεκτρική μορφή και το αντίστροφο. Με δεδομένους τους περιορισμούς στο εύρος ζώνης που θέτει η αρχή λειτουργίας των οπτικών ενισχυτών αλλά και η ίδια η οπτική ίνα, είναι σημαντική η μεγιστοποίηση της φασματικής απόδοσης (spectral efficiency, SE), που υπολογίζεται σε bit/s/Hz. Δεδομένων όμως των περιορισμών στην ισχύ του σήματος που επιβάλλονται από τη μη γραμμικότητα της ίνας, είναι επίσης σημαντικό να μεγιστοποιηθεί η απόδοση της ισχύος, δηλαδή να ελαχιστοποιηθεί η απαιτούμενη μέση μεταδιδόμενη ενέργεια ανά bit ή η απαιτούμενη αναλογία σήματος προς θόρυβο (signal-to-noise ratio, SNR), ανά bit.

Τα συστήματα που χρησιμοποιούν απλή δυαδική μορφή διαμόρφωση, όπως On-Off Keying (OOK) κωδικοποιούν ένα bit ανά σύμβολο. Λαμβάνοντας υπόψη τους φυσικούς περιορισμούς των φίλτρων που χρησιμοποιούνται στην περίπτωση των WDM συστημάτων, η φασματική απόδοση της OOK διαμόρφωσης είναι περίπου 0.8 bit/s/Hz ανά πόλωση.

Η πιο πολλά υποσχόμενη τεχνική ανίχνευσης για την επίτευξη υψηλού SE με την ταυτόχρονη διατήρηση υψηλού SNR, είναι η σύμφωνη ανίχνευση (coherent detection) με την πολυπλεξία πόλωσης. Σε αντίθεση με τα IM/DD συστήματα η επιλογή των συμβόλων γίνεται χρησιμοποιώντας τα σήματα εντός φάσης (In phase – I) και ορθογώνια διαμόρφωση πλάτους (Quadrature – Q) στις δύο οπτικές πολώσεις, επιτρέποντας την κωδικοποίηση των συμβόλων σε όλους τους διαθέσιμους βαθμούς ελευθερίας. [91]

Στα μη σύμφωνα συστήματα ανίχνευσης, όπως είναι τα IM/DD στα οποία όπως είδαμε στην προηγούμενη παράγραφο η ένταση του οπτικού σήματος που μεταδίδεται μέσω μιας οπτικής ίνας ανιχνεύεται από μια φωτοδίοδο, εκτός της απλότητάς τους, η ευαισθησία του δέκτη δεν εξαρτάται ούτε από τη φάση του φέροντος ούτε από την

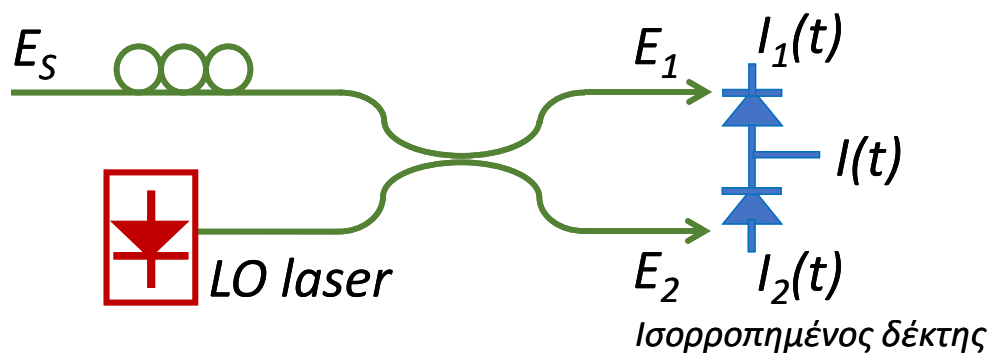
πόλωση του οπτικού σήματος, οι οποίες αμφότερες έχουν τυχαία διακύμανση στα πραγματικά συστήματα. (σχήμα 2.14)



Σχήμα 2.14: Δέκτης απλής φωτοδιόδου

Στην περίπτωση τη διαμόρφωσης της έντασης (amplitude-shift keying - ASK) το επικρατέστερο σχήμα διαμόρφωσης για αρκετές δεκαετίες ήταν το OOK, αποτελώντας τον απλούστερο τρόπο ASK διαμόρφωσης. Στα σύγχρονα συστήματα IM/DD, με σκοπό την επίτευξη υψηλού SE, κωδικοποιώντας περισσότερα από 1 bit ανά σύμβολο, γίνεται χρήση πιο σύνθετων σχημάτων διαμόρφωσης παλμών πολλών επιπέδων (multi level pulse-amplitude modulation) όπως το PAM-4 και το PAM-8 που αναφέραμε προηγουμένως. Ένα σημαντικό μειονέκτημα των μη σύμφωνων συστημάτων ανίχνευσης είναι η μη αναστρέψιμη απώλεια της πληροφορίας της φάσης, γεγονός που αποτρέπει την πλήρη εξισορρόπηση των γραμμικών φαινομένων όπως η χρωματική διασπορά και η διασπορά τρόπων πόλωσης με τη χρήση γραμμικών φίλτρων.

Στα συστήματα σύμφωνης ανίχνευσης ο δέκτης χρησιμοποιεί όλες τις παραμέτρους του ηλεκτρικού πεδίου, δηλαδή περιέχει τόσο την πληροφορία του πλάτους όσο και της



Σχήμα 2.15: Σύμφωνος δέκτης

φάσης αυτού. Επιτρέπει έτσι τη μεγαλύτερη ευελιξία στις μορφές διαμόρφωσης, καθώς η πληροφορία μπορεί να κωδικοποιηθεί σε πλάτος και φάση, ή εναλλακτικά σε συνιστώσες εντός φάσης (I) και ορθογώνιο στη φάση (Q) ενός φέροντος. Η σύμφωνη ανίχνευση απαιτεί από τον δέκτη να έχει γνώση της φάσης του φέροντος, καθώς το λαμβανόμενο σήμα αποδιαμορφώνεται από έναν τοπικό ταλαντωτή (Local Oscillator - LO) η φάση του οποίου χρησιμεύει ως φάση αναφοράς. Το σχήμα 2.15 δείχνει την αρχή λειτουργίας ενός σύμφωνου δέκτη που συγκρίνει το σήμα με το LO θεωρώντας πως η πόλωση του σήματος και του laser του LO είναι ευθυγραμμισμένες. Ο ισορροπημένος δέκτης (balanced receiver) εξαλείφει τη DC συνιστώσα και μεγιστοποιεί το σήμα λόγω υπέρθεσης.

Η αρχή λειτουργίας του σύμφωνου δέκτη μπορεί να εξηγηθεί λαμβάνοντας υπόψη τα εξής:

Έστω το μιγαδικό πεδίο του οπτικού σήματος E_s δίνεται από τη σχέση :

$$E_s(t) = A_s(t) \exp(i\omega_s t) \quad (2.45)$$

Όπου το A_s είναι το πλάτος στο μιγαδικό επίπεδο και ω_s η γωνιακή συχνότητα. Αντίστοιχα το πεδίο του LO δίνεται από τη σχέση :

$$E_{lo}(t) = A_{lo}(t) \exp(i\omega_{lo} t) \quad (2.46)$$

Η ισχύς του σήματος και του LO αντίστοιχα δίνονται από τις σχέσεις:

$$P_s = \frac{|A_s|^2}{2} \quad (2.47)$$

Η διάταξη της ισορροπημένης ανίχνευσης συνήθως χρησιμοποιείται στους σύμφωνους δέκτες κυρίως για να απαλείψει τη DC συνιστώσα και να μεγιστοποιήσει την αλληλεπίδραση του σήματος με τον LO. Ο μηχανισμός που φαίνεται στο σχήμα 2.15 βασίζεται στη χρήση ενός οπτικού συζεύκτη 3-dB που μετατοπίζει κατά 180° τη φάση του σήματος ή του LO μεταξύ των δύο εξόδων. Στη συνέχεια τα πεδία που προσπίπτουν στις φωτοδιόδους του ισορροπημένου ανιχνευτή δίνονται από τις σχέσεις :

$$E_1 = \frac{1}{\sqrt{2}} (E_s + E_{lo}) \quad (2.48)$$

$$E_2 = \frac{1}{\sqrt{2}}(E_s - E_{lo}) \quad (2.49)$$

Και η έξοδοι των φωτοδιόδων δίνουν (φωτό)ρεύμα που εκφράζεται από τις σχέσεις :

$$\begin{aligned} I_1(t) &= kR \left[\operatorname{Re} \left\{ \frac{A_s(t) \exp(i\omega_s t) + A_{lo} \exp(i\omega_{lo} t)}{\sqrt{2}} \right\} \right]^2 \\ &= \frac{R}{2} [P_s(t) + P_{lo} + 2\sqrt{P_s(t)P_{lo}} \cos\{|\omega_s - \omega_{lo}| + \theta_s(t) - \theta_{lo}(t)\}] \end{aligned} \quad (2.50)$$

$$\begin{aligned} I_2(t) &= kR \left[\operatorname{Re} \left\{ \frac{A_s(t) \exp(i\omega_s t) - A_{lo} \exp(i\omega_{lo} t)}{\sqrt{2}} \right\} \right]^2 \\ &= \frac{R}{2} [P_s(t) + P_{lo} - 2\sqrt{P_s(t)P_{lo}} \cos\{|\omega_s - \omega_{lo}| + \theta_s(t) - \theta_{lo}(t)\}] \end{aligned} \quad (2.51)$$

Όπου $\theta_s(t)$, $\theta_{lo}(t)$ οι φάσεις του σήματος και του LO αντίστοιχα. R η απόκριση της φωτοδιόδου.

Η έξοδος του ισοροπημένο ανιχνευτή δίνεται από τη σχέση

$$I(t) = I_1(t) - I_2(t) = 2R \sqrt{P_s(t)P_{lo}} \cos\{\omega_{IF}t + \theta_s(t) - \theta_{lo}(t)\} \quad (2.52)$$

όπου $\omega_{IF} = |\omega_s - \omega_{lo}|$

Από την προηγούμενη σχέση φαίνεται πως η ισχύς του LO είναι σταθερή ενώ η θ_{lo} περιέχει τη συνιστώσα του θορύβου φάσης που μεταβάλλεται τυχαία στον χρόνο. Προκύπτει συνεπώς πως με τον ισοροπημένο ανιχνευτή εξάγεται μόνο η αλληλεπίδραση μεταξύ του σήματος και του LO.

Επίσης εξάγεται το συμπέρασμα πως το φωτόρεύμα είναι ανάλογο του $\sqrt{P_{lo}}$ συνεπώς μπορούμε να αυξήσουμε την απολαβή του αυξάνοντας την ισχύ του LO. Με αυτό τον τρόπο μπορεί να υπερκεραστεί ο θερμικός θόρυβος και να αυξηθεί ο λόγος σήματος προς θόρυβο, ο οποίος με τη σειρά του εξαρτάται από τον ενδογενή θόρυβο του σήματος και μπορεί φτάσει στο όριο του κβαντικού θορύβου.

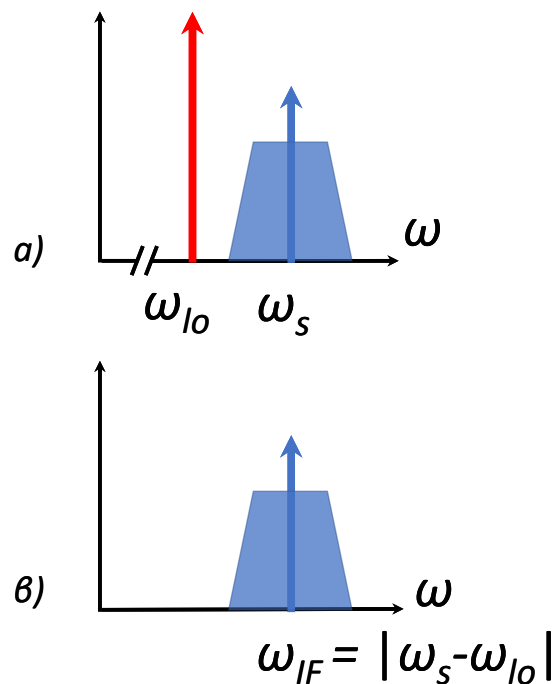
2.5.1.1 Ετερόδυνη ανίχνευση

Η ετερόδυνη ανίχνευση αναφέρεται στην περίπτωση όπου $|\omega_{IF}| \gg \frac{2\pi}{T}$, με την τιμή $1/T$ να υποδηλώνει το ρυθμό των συμβόλων και να καθορίζει το εύρος ζώνης της διαμόρφωσης του οπτικού φέροντος. Σε αυτή την περίπτωση η εξίσωση 2.42 γίνεται :

$$I(t) = 2R \sqrt{P_s(t) P_{lo}} \cos\{\omega_{IF}t + \theta_{sm}(t) + \theta_n(t)\} \quad (2.53)$$

Όπου $\theta_n(t) = \theta_{sn}(t) - \theta_{lo}(t)$, ο συνολικός θόρυβος φάσης, $\theta_s(t)$ η φάση του σήματος, $\theta_{sn}(t)$ ο θόρυβος φάσης του σήματος και $\theta_{sm}(t)$ η φάση του σήματος η διαμόρφωση φάσης.

Από τη σχέση 2.53 προκύπτει πως το ηλεκτρικό πεδίο του οπτικού σήματος μεταπίπτει στο σήμα IF περιέχοντας την πληροφορία του πλάτους και της φάσης, όπως φαίνεται στο σχήμα 2.16α.



Σχήμα 2.16: Ετερόδυνη ανίχνευση. α) Μετάπτωση οπτικού σήματος στο IF, β) επιλογή συχνότητας του LO κοντά στο φάσμα του σήματος

Επιλέγοντας τη συχνότητα LO κοντά σε αυτή του σήματος προκύπτει το φάσμα που φαίνεται στο 2.16β.

Η σχέση 2.53 μπορεί ισοδύναμα να γραφτεί ως :

$$I(t) = 2R \sqrt{P_s(t) P_{lo}} \exp[i\{\theta_{sm}(t) + \theta_n(t)\}] \quad (2.54)$$

από την οποία φαίνεται πως το $I(t)$ είναι ισοδύναμο του μιγαδικού πλάτους του οπτικού σήματος με την εξαίρεση του θόρυβου φάσης που πηγάζει από τον LO.

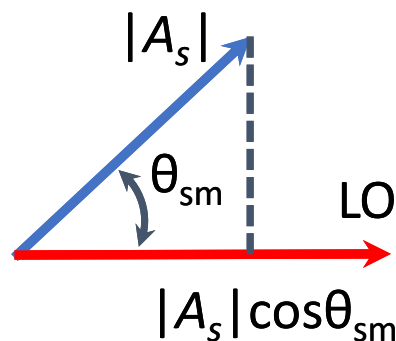
Για την αποκωδικοποίηση του σήματος εξόδου του ετερόδυνου δέκτη πρέπει να εφαρμοστεί ένα κύκλωμα βρόχου κλειδώματος φάσης (Phase-locked loop – PLL). Παρά το γεγονός πως ο συνολικός θόρυβος φάσης αλλάζει στον χρόνο, το ηλεκτρικό PLL μπορεί να τον εκτιμήσει και να αποκωδικοποιήσει τα σύμβολα του σήματος από την εξίσωση 2.54.

2.5.1.2 Ομόδυνη ανίχνευση

Η ομόδυνη ανίχνευση αναφέρεται στην περίπτωση όπου $\omega_{IF} = 0$. Το φωτόρευμα των φωτοδιόδων δίνεται από τη σχέση :

$$I(t) = 2R \sqrt{P_s(t) P_{lo}} \cos\{\theta_{sm}(t) - \theta_{lo}(t)\} \quad (2.55)$$

Το διάγραμμα των φασόρων μεταξύ του σήματος και του LO του ομόδυνου δέκτη φαίνεται στο σχήμα 2.17

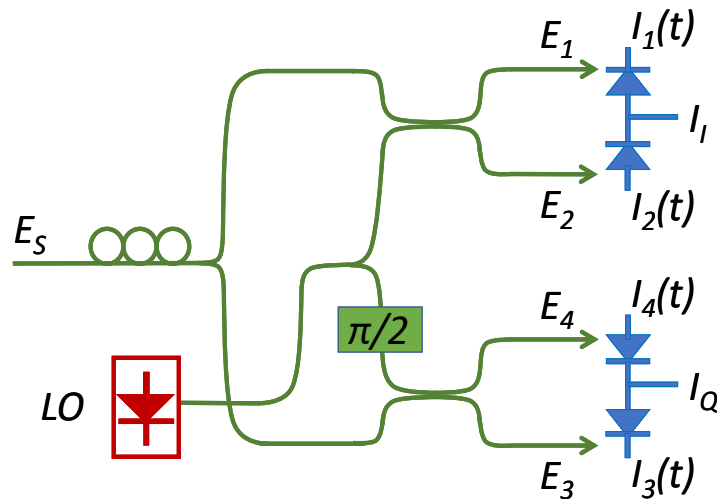


Σχήμα 2.17: Διάγραμμα φασόρων, ομόδυνη ανίχνευση

Για την ορθή αποκωδικοποίηση των συμβόλων του σήματος, η φάση του LO θ_{lo} πρέπει να είναι τέτοια ώστε να μηδενίζει τον θόρυβο φάσης του σήματος. Αυτό επιτυγχάνεται στην πράξη με οπτικό PLL (Optical Phase Locked Loop – OPLL) η υλοποίηση όμως δεν είναι εύκολη ενώ αυξάνει την πολυπλοκότητα του δέκτη. Επιπρόσθετα, όπως προκύπτει και από τη σχέση 2.55, ο ομόδυνος ανιχνευτής υπολογίζει μόνο το συνημίτονο της -συμφασικής- αλληλεπίδρασης σήματος και θορύβου, ενώ το ημίτονο δε μπορεί να ανιχνευτεί. Συνεπώς ο ομόδυνος αυτός δέκτης δεν μπορεί να εξάγει το σύνολο της πληροφορίας που περιέχει το πλάτος του μιγαδικού σήματος.

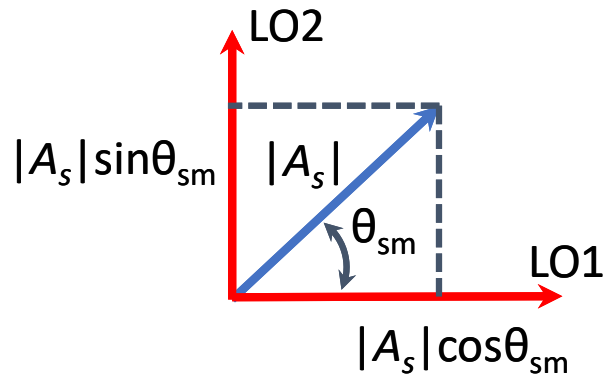
2.5.1.3 Ομόδυνη ανίχνευση – 90° οπτικό υβρίδιο

Για να επιλυθεί το προηγούμενο πρόβλημα, στον ομόδυνο δέκτη εισάγουμε βρόχο που εισάγει μια μετατόπιση της φάσης του LO κατά 90°, όπως φαίνεται στο σχήμα 2.18. Το δικτύωμα αυτό ονομάζεται 90° οπτικό υβρίδιο (optical hybrid).



Σχήμα 2.18: Ομόδυνη ανίχνευση – 90° οπτικό υβρίδιο

Με αυτό τον τρόπο μπορούν να ανιχνευτούν και οι δύο συνιστώσες I και Q . Το διάγραμμα των φασόρων του σχήματος 2.19 δείχνει πως αυτό επιτυγχάνεται με τη χρήση του $LO1$ και $LO2$ αντίστοιχα.



Σχήμα 2.19: Δάγραμμα φασόρων, ομόδونه ανίχνευση - 90ο οπτικό υβρίδιο

Οι έξοδοι E_1, E_2, E_3, E_4 του ομόδουου ανιχνευτή υπολογίζονται από τις σχέσεις :

$$E_1 = \frac{1}{2} (E_s + E_{lo})$$

$$E_2 = \frac{1}{2} (E_s - E_{lo})$$

$$E_3 = \frac{1}{2} (E_s + iE_{lo})$$

$$E_4 = \frac{1}{2} (E_s - iE_{lo}) \quad (2.56)$$

ενώ το φωτόρευμα των φωτοδιόδων δίνεται από τις σχέσεις :

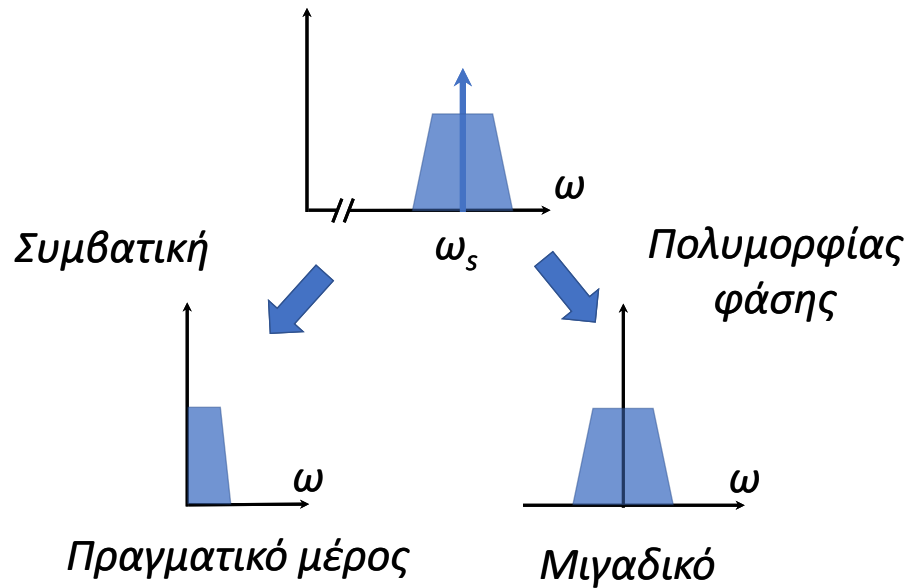
$$I_I(t) = I_{I1}(t) - I_{I2}(t) = R \sqrt{P_s(t) P_{lo}} \cos[\theta_{sm}(t) - \theta_{lo}(t)] \quad (2.57)$$

$$I_Q(t) = I_{Q1}(t) - I_{Q2}(t) = R \sqrt{P_s(t) P_{lo}} \sin[\theta_{sm}(t) - \theta_{lo}(t)] \quad (2.58)$$

και

$$I(t) = I_I(t) + iI_Q(t) = R \sqrt{P_s(t) P_{lo}} \exp\{i [\theta_{sm}(t) + \theta_n(t)]\} \quad (2.59)$$

Η σχέση 2.59 δείχνει πως η μιγαδική έξοδος του ανιχνευτή ισοδυναμεί με το μιγαδικό πλάτος του οπτικού σήματος με την εξαίρεση του θορύβου φάσης που πηγάζει από τον LO. Το ηλεκτρικό πεδίο του οπτικού σήματος μεταπίπτει στη βασική ζώνη (baseband)



Σχήμα 2.20: Ομόδυνη ανίχνευση

Υπάρχουν δύο τύποι ομόδυνων ανιχνευτών, ο συμβατικός, και ο ανιχνευτής πολυμορφίας φάσης (phase diversity). Όπως φαίνεται στο σχήμα 2.20, ο πρώτος προσμετρά μονάχα το πραγματικό μέρος του μιγαδικού σήματος ενώ ο δεύτερος που είναι και ο επικρατέστερος προσμετρά το πραγματικό και το φανταστικό μέρος, στη θετική και αρνητική περιοχή του φάσματος αντίστοιχα.

Από τα σχήματα 2.16 και 2.20 προκύπτει πως ο ετερόδυνος ανιχνευτής και ο ομόδυνος με πολυμορφία φάσης μπορούν να αποδώσουν το σύνολο της πληροφορίας του μιγαδικού σήματος οπότε στα σύγχρονα συστήματα οπτικών επικοινωνιών αυτές οι διατάξεις έχουν επικρατήσει. Επιπροσθέτως, δεδομένης της απευθείας λειτουργίας στη βασική ζώνη του ομόδυνου ανιχνευτή με πολυμορφία ζώνης, άρα καλύτερης καλύτερη αξιοποίησης του εύρους ζώνης, έναντι της λειτουργίας στην υψηλή IF περιοχή του ετερόδυνου, ο πρώτος διαθέτει σαφές πλεονέκτημα.

2.5.2 Σχήματα διαμόρφωσης

Στο σχήμα 2.21 φαίνονται οι αστερισμοί των συνηθέστερων σχημάτων διαμόρφωσης που χρησιμοποιούνται στις οπτικές επικοινωνίες. Όσα βρίσκονται σε μια διάσταση, όπως τα OOK, PAM4, PAM8 κτλ (σχ. 2.21α) χρησιμοποιούνται στα IM/DD συστήματα.

Όπως είδαμε στα προηγούμενα, με τη βοήθεια του σύμφωνου ανιχνευτή μπορούμε να αποδιαμορφώσουμε σχήματα που βρίσκονται σε παραπάνω από μία διαστάσεις. Με τη βοήθεια διαμορφωτών IQ μπορούμε να δημιουργήσουμε σχήματα που κινούνται σε δύο διαστάσεις (μιγαδικό επίπεδο) και πολλά επίπεδα τιμών. Με αυτό τον τρόπο, είτε διαμορφώνοντας μόνο τη φάση διατηρώντας το πλάτος σταθερό όπως πχ στα BPSK, QPSK, 8-PSK κτλ (σχ. 2.21β), είτε διαμορφώνοντας τη φάση και το πλάτος πχ στα 4-QAM, 16-QAM, 32QAM κτλ (σχ. 2.21γ) μπορούμε να βελτιώσουμε τη φασματική απόδοση του συστήματος.

Τα τελευταία χρόνια έχουν κερδίσει έδαφος οι γεωμετρικές [92] και οι πιθανοκρατικές [93] διαμορφώσεις (σχήμα 2.21δ). Οι πρώτες προτείνουν σχήματα διαμόρφωσης που χαρακτηρίζονται από τη μη ίση απόσταση μεταξύ των σημείων του αστερισμού ενώ οι δεύτερες αποδίδουν διαφορετική πιθανότητα στα σύμβολα ανάλογα με την ισχύ τους. Αμφότερες εφαρμόζονται για τη βελτίωση της χωρητικότητας του συστήματος [94].

2.5.3 Φασματική απόδοση

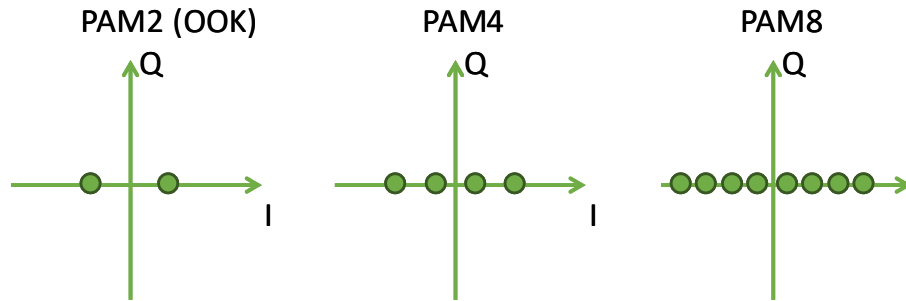
Η χωρητικότητα ενός καναλιού δίνεται από τη σχέση :

$$C = \frac{WD}{2} \log_2 \left(1 + \frac{S}{N} \right) \quad (2.60)$$

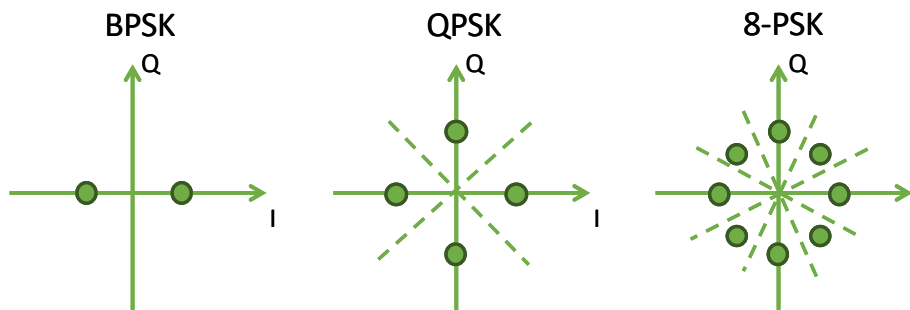
όπου W το εύρος ζώνης του οπτικού σήματος, D η διάσταση του σχήματος διαμόρφωσης, S και N η μέση τιμή ισχύος του σήματος και του θορύβου αντίστοιχα [95]. Στα σύμφωνα συστήματα οπτικών επικοινωνιών χρησιμοποιούνται πάντοτε και οι δύο συνιστώσες I και Q, συνεπώς η διάσταση $D=2$ αν γίνεται διάδοση σε μία πόλωση ή $D=4$ αν γίνεται διάδοση και στις δύο πολώσεις. Στα συστήματα WDM η διάσταση παίρνει μεγαλύτερες τιμές, ανάλογες του αριθμού των διαδιδόμενων καναλιών. Η σχέση 2.60 μπορεί να γραφτεί ως :

$$SE_{lim} = \log_2 \left(1 + SE_{lim} \frac{E_b}{N_0} \right) \quad (2.61)$$

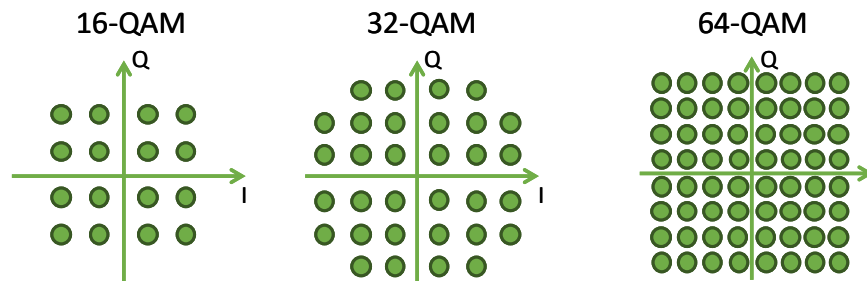
α)



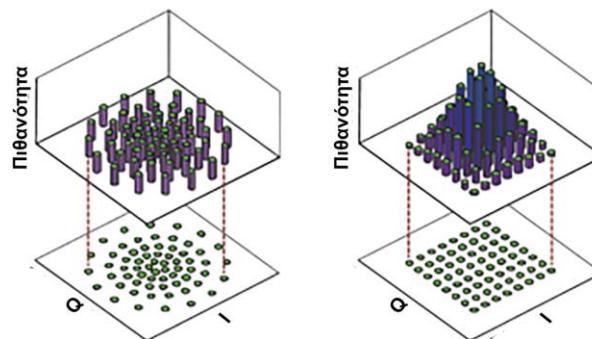
β)



γ)



δ)



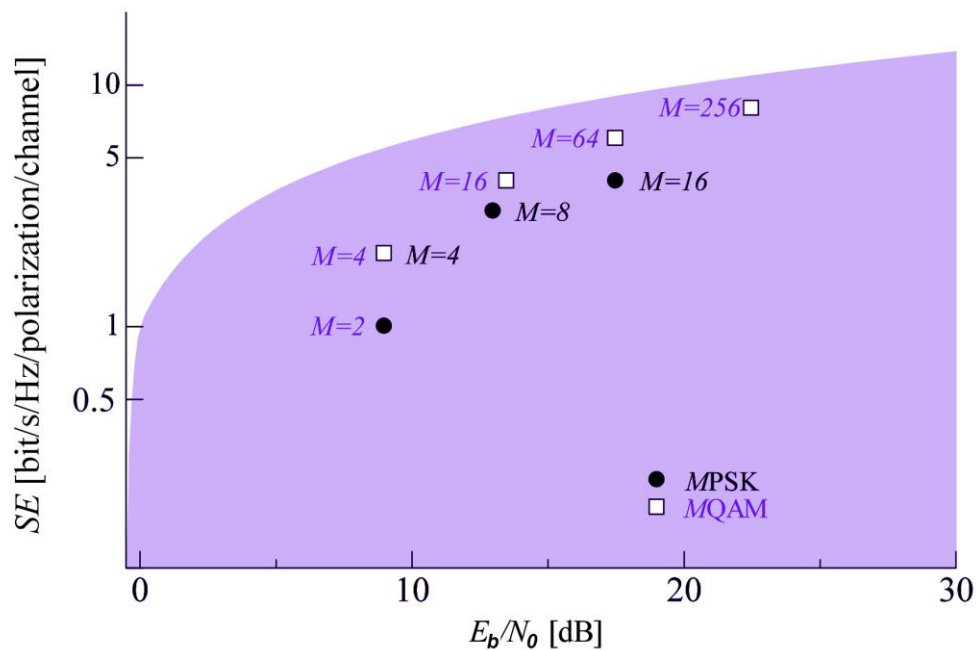
Σχήμα 2.21: Σχήματα διαμόρφωσης α) μίας διάστασης και β-δ) ανώτερης τάξης

ενώ $SE_{lim} = \frac{C}{W(\frac{D}{2})}$ είναι το όριο Shannon της φασματικής απόδοσης ανά πόλωση και κανάλι.

Από τη σχέση 2.61 μπορούμε να υπολογίσουμε τον λόγο $\frac{E_b}{N_0}$ συναρτήσει του ορίου Shannon σύμφωνα με τη:

$$\frac{E_b}{N_0} = \frac{2^{SE_{lim}} - 1}{SE_{lim}} \quad (2.62)$$

Αντίστοιχα μπορούμε να υπολογίσουμε τη φασματική απόδοση ενός σχήματος διαμόρφωσης M συμβόλων από τη σχέση :



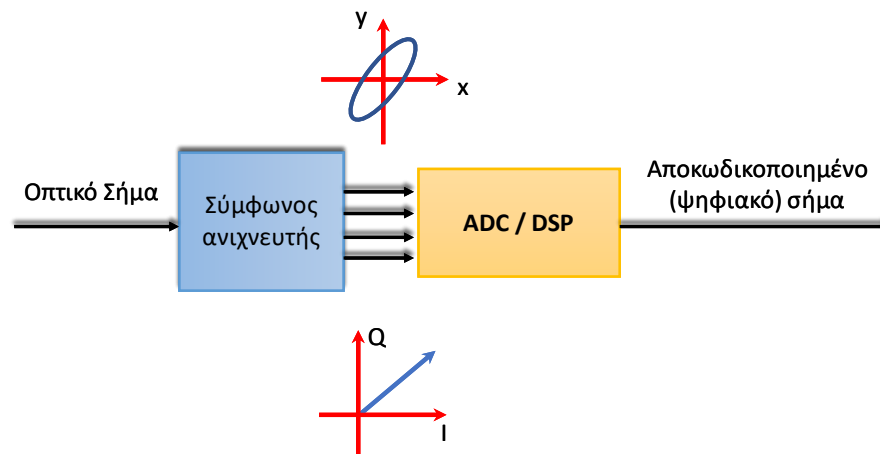
Σχήμα 2.22: Φασματική απόδοση συναρτήσει του λόγου E_b/N_0

$$SE = \frac{2 \log_2 M}{D} \quad (2.63)$$

Στο σχήμα 2.22 αποτυπώνεται η φασματική απόδοση συναρτήσει του λόγου $\frac{E_b}{N_0}$ για τις διαμορφώσεις PSK και QAM. Είναι εμφανές πως η φασματική απόδοση αυξάνεται με το μέγεθος της τάξης της διαμόρφωσης, όμως αυτό γίνεται με τίμημα την αύξηση του απαιτούμενου λόγου $\frac{E_b}{N_0}$.

2.6 Ψηφιακή Επεξεργασία Σήματος σε Σύμφωνα Συστήματα

Οι έξοδοι του σύμφωνου δέκτη περιέχουν ανεπιθύμητες διακυμάνσεις τόσο στη φάση όσο και στην πόλωση που εμποδίζουν την αναγνώριση του διαμορφωμένου σήματος. Οι οπτικές μέθοδοι εξίσωσης των φαινομένων διάδοσης δεν επαρκούν. Αντιθέτως η υψήρρυθμη ψηφιακή επεξεργασία των εισερχομένων σημάτων έχει αποδειχτεί στην πράξη ιδιαίτερα αποδοτική και ωφέλιμη.



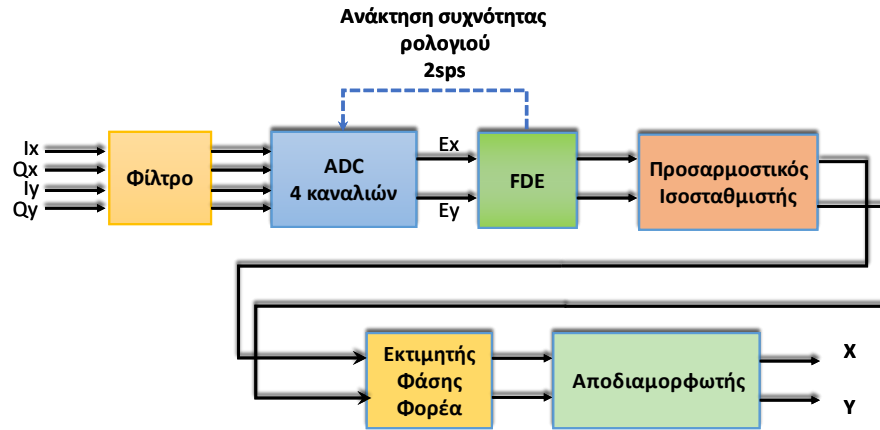
Σχήμα 2.23: Τυπική διάταξη ψηφιακού σύμφωνου δέκτη

Στο σχήμα 2.23 φαίνεται η βασική διάταξη ενός σύμφωνου δέκτη. Το οπτικό σήμα ανιχνεύεται από τον ομόδουνο (ή ετερόδουνο) ανιχνευτή, η έξοδος του οποίου όπως αναλύθηκε στα προηγούμενα περιέχει τις δύο συνιστώσες I και Q των δύο πολώσεων, μετατρέπεται σε ψηφιακή μορφή και υπόκειται σε επεξεργασία από διατάξεις DSP.

Στο σχήμα 2.24 απεικονίζεται μια τυπική αλληλουχία DSP διαδικασιών που ακολουθούνται για την αποκωδικοποίηση των συμβόλων.

2.6.1 ADC

Τα τέσσερα αναλογικά σήματα που, I_x , Q_x , I_y και Q_y φιλτράρονται και οδηγούνται σε έναν ADC τεσσάρων καναλιών για τη μετατροπή τους σε ψηφιακή μορφή. Στις περισσότερες περιπτώσεις απαιτείται μια δειγματοληψία διπλάσια του ρυθμού των συμβόλων.



Σχήμα 2.24: DSP και επιμέρους διαδικασίες

Συγκεκριμένα, το εύρος ζώνης Nyquist του οπτικού σήματος B δίνεται από τη σχέση $B=1/T$ όπου T η περίοδος του συμβόλου. Οι έξοδοι I και Q του ομόδουνου ανιχνευτή έχουν εύρος ζώνης $B/2 = 1/(2T)$. Σύμφωνα με τη θεωρία της δειγματοληψίας, ο ρυθμός δειγματοληψίας πρέπει να είναι διπλάσιος της μέγιστης συχνότητας του σήματος, ενώ ελάχιστος ρυθμός δειγματοληψίας που επιβάλλεται είναι $R_o=1/T$.

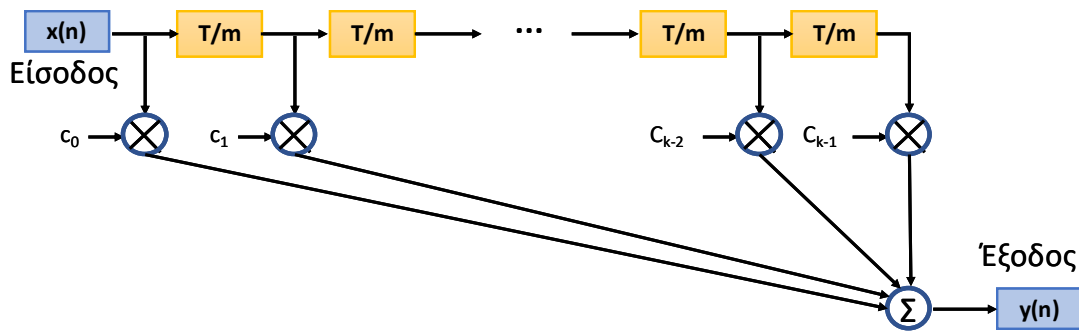
Πχ σε ένα σύστημα 16-QAM διπλής πόλωσης συνολικού ρυθμού 200 Gb/s, ο ρυθμός των συμβόλων είναι $1/T = 25\text{GBd}$. Σύμφωνα με τα παραπάνω ο ρυθμός δειγματοληψίας πρέπει να είναι $R = 50\text{GS/s}$, ο οποίος ισοδυναμεί με το διπλάσιο του ρυθμού [95].

2.6.2 Προσαρμοστικός Ισοσταθμιστής (Adaptive Equalizer)

Η αρχή λειτουργίας ενός προσαρμοστικού ισοσταθμιστή βασίζεται στη διάταξη των FIR φίλτρων [65]. Στο σχήμα 2.25 φαίνεται μια διάταξη FIR κατά την οποία χρησιμοποιώντας έναν επαρκή αριθμό διακριτών συντελεστών (taps) μπορεί να αναπαρασταθεί μια οποιαδήποτε συνάρτηση μεταφοράς.

Το σήμα εισόδου $x(n)$ καθυστερεί κατά ακέραια πολλαπλάσια του T/m όπου το m εκφράζει τον λόγο της δειγματοληψίας. Η έξοδος του FIR δίνεται από τη σχέση:

$$y(n) = \vec{c}(n)^T \vec{x}(n) = \sum_{j=0}^{k-1} c_j(n)x(n-j) \quad (2.64)$$



Σχήμα 2.25: FIR φίλτρο

όπου το $\vec{x}(n) = [x(n), x(n-1), \dots, x(n-k+1)]^T$ το διάνυσμα που περιέχει όλα τα χρονικά στοιχεία του σήματος, με T να υποδηλώνει τον ανάστροφο πίνακα και $\vec{c}(n) = [c_0(n), c_1(n), \dots, c_{k-1}(n)]^T$ το αντίστοιχο διάνυσμα για τους συντελεστές (βάρη) των taps.

Όταν το διάνυσμα $\vec{c}(n)$ ανταποκρίνεται στη φασματική απόκριση του συστήματος, τότε ο διακριτός μετασχηματισμός Fourier δίνει τη συνάρτηση μεταφοράς του. Για να συμβεί αυτό, ο αριθμός των συντελεστών του διανύσματος $\vec{c}(n)$ πρέπει να καλύπτει το εύρος ζώνης του συστήματος.

Τα FIR φίλτρα βρίσκουν εφαρμογή στους ψηφιακούς σύμφωνους δέκτες για την εξισορρόπηση των γραμμικών αλλοιώσεων. Στις περιπτώσεις όπου η ισχύς του σήματος στην είσοδο της οπτικής ίνας είναι αρκετά χαμηλή οπότε και επικρατούν τα γραμμικά φαινόμενα, τα φίλτρα FIR επιτυγχάνουν επαρκή εξισορρόπηση.

Συγκεκριμένα, ο μετασχηματισμός Fourier του μιγαδικού πλάτους στο δέκτη δίνεται από τη σχέση

$$\begin{bmatrix} E_x(\omega) \\ E_y(\omega) \end{bmatrix} = H_0(\omega) \begin{bmatrix} E_{in,x}(\omega) \\ E_{in,y}(\omega) \end{bmatrix} \quad (2.65)$$

Όπου $H_0(\omega)$ είναι η συνάρτηση μεταφοράς του καναλιού και ω η γωνιακή συχνότητα. Όπως είδαμε στα προηγούμενα, το διαδιδόμενο σήμα παραμορφώνεται εξαιτίας της εξάρτησης της συνάρτησης μεταφοράς από τη συχνότητα δημιουργούν τη διασπορά ταχύτητας ομάδας (GVD). Επιπλέον εξαιτίας της διπλοθλαστικότητας της οπτικής ίνας οι

δύο πολώσεις αναμειγνύονται δημιουργώντας το φαινόμενο της διασποράς τρόπων πόλωσης (polarization mode dispersion, PMD).

Λαμβάνοντας τα προηγούμενα υπόψη μπορούμε να υπολογίσουμε τη συνάρτηση μεταφοράς συναρτήσει της GVD, $D(\omega)$, της PMD, $U(\omega)$, τις απώλειες εξαιτίας πόλωσης (polarization-dependent loss - PDL), T και της διπλοθλαστικότητας η οποία εκφράζεται από τον μοναδιαίο πίνακα Jones, K , από τη σχέση :

$$H_0(\omega) = D(\omega)U(\omega)TK \quad (2.66)$$

Το $D(\omega)$ δίνεται από τη σχέση:

$$D(\omega) = e^{-i\frac{\lambda_c^2}{4\pi c}\beta_2 L\omega^2} \quad (2.67)$$

όπου το β_2 είναι παράμετρος της διασποράς και L το μήκος της διάδοσης, c η ταχύτητα του φωτός και λ_c το μήκος κύματος αναφοράς και ω η γωνιακή συχνότητα.

Το $U(\omega)$, δίνεται από τη σχέση:

$$U(\omega) = R_1^{-1} \begin{bmatrix} \exp\left(i\frac{\omega\Delta\tau}{2}\right) & 0 \\ 0 & \exp\left(-i\frac{\omega\Delta\tau}{2}\right) \end{bmatrix} R_1 \quad (2.68)$$

όπου R_1 είναι ο μοναδιαίος πίνακας που μετατρέπει κάθε κατάσταση πόλωσης (State of Polarization – SOP) στις επιμέρους x, y πολώσεις, ενώ το $\Delta\tau$ είναι η καθυστέρηση ομάδας (differential group delay – DGD).

Το T είναι ένας πίνακας Hermite που δίνεται από τη σχέση:

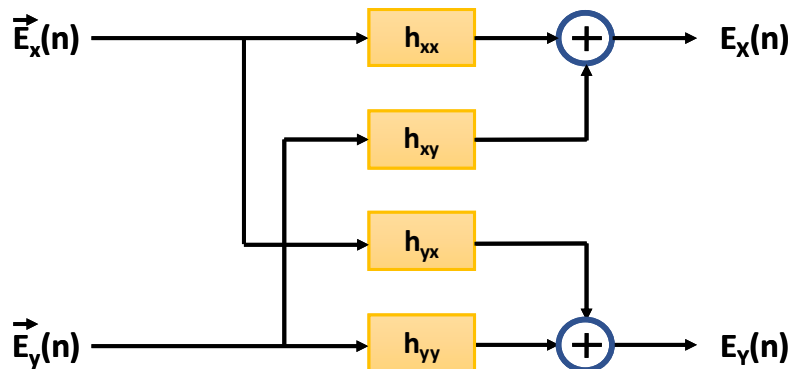
$$T = R_2^{-1} \begin{bmatrix} \sqrt{\Gamma_{max}} & 0 \\ 0 & \sqrt{\Gamma_{min}} \end{bmatrix} R_2 \quad (2.69)$$

όπου R_2 είναι ο μοναδιαίος πίνακας που μετατρέπει κάθε ιδιοτιμή της PDL στις επιμέρους x, y πολώσεις, ενώ τα $\Gamma_{max}, \Gamma_{min}$ είναι οι συντελεστές ισχύος των ιδιοτιμών της PDL [96].

Στις περισσότερες περιπτώσεις στις οποίες το οπτικό κανάλι είναι σταθερό η άρση του μεγαλύτερου μέρους της GVD πραγματοποιείται με σταθερά FIR φίλτρα. Το υπόλοιπο μέρος απαιτεί προσαρμοστικά FIR φίλτρα. Το σύνολο των γραμμικών φαινομένων που εγείρονται κατά τη διάδοση των οπτικών σημάτων στην ίνα εκφράζονται από τη σχέση 2.66. Ο σκοπός λοιπόν των γραμμικών φίλτρων είναι να εξισορροπήσουν τα φαινόμενα αντιστρέφοντας τη συνάρτηση μεταφοράς του καναλιού :

$$H_{eq}(\omega) \cong H_0^{-1}(\omega) \begin{bmatrix} h_{xx}(\omega) & h_{xy}(\omega) \\ h_{yx}(\omega) & h_{yy}(\omega) \end{bmatrix} \quad (2.70)$$

Ο προσαρμοστικός ισοσταθμιστής που απαιτείται για την αντιμετώπιση των δυναμικών φαινομένων όπως το SOP και το PMD είναι μια διάταξη πολλών εισόδων, πολλών εξόδων (multiple-input multiple-output - MIMO) τεσσάρων FIR φίλτρων, όπως φαίνεται στο σχήμα 2.26 [97]



Σχήμα 2.26: Ισοσταθμιστής MIMO

Αν k είναι ο αριθμός των taps των FIR φίλτρων και τα διανύσματα του πεδίου εισόδου είναι

$$\vec{E}_x(n) = [E_x(n), E_x(n-1), \dots, E_x(n-k+2), E_x(n-k+1)]^T \quad (2.71)$$

$$\vec{E}_y(n) = [E_y(n), E_y(n-1), \dots, E_y(n-k+2), E_y(n-k+1)]^T \quad (2.72)$$

Η έξοδος του φίλτρου του σχήματος 2.26 δίνεται από τη σχέση

$$E_X = \overrightarrow{h_{xx}}(n)^T \overrightarrow{E_x}(n) + \overrightarrow{h_{xy}}(n)^T \overrightarrow{E_y}(n) \quad (2.73)$$

$$E_Y = \overrightarrow{h_{yx}}(n)^T \overrightarrow{E_x}(n) + \overrightarrow{h_{yy}}(n)^T \overrightarrow{E_y}(n) \quad (2.74)$$

Οι αλγόριθμοι που χρησιμοποιούνται στα προσαρμοστικά FIR φίλτρα είναι ο decision-directed least-mean-square (DD-LMS) και ο constant-modulus algorithm (CMA).

Και οι δύο αλγόριθμοι βασίζονται στην εκπαίδευση, ενημερώνοντας τους συντελεστές των taps (τα βάρη) χρησιμοποιώντας σύμβολα αναφοράς, σύμφωνα με τις σχέσεις :

$$\begin{aligned} \overrightarrow{h_{xx}}(n+1) &= \overrightarrow{h_{xx}}(n) + le_X(n) \overrightarrow{E_x}(n)^* \\ \overrightarrow{h_{xy}}(n+1) &= \overrightarrow{h_{xy}}(n) + le_X(n) \overrightarrow{E_y}(n)^* \\ \overrightarrow{h_{yx}}(n+1) &= \overrightarrow{h_{yx}}(n) + le_Y(n) \overrightarrow{E_x}(n)^* \\ \overrightarrow{h_{yy}}(n+1) &= \overrightarrow{h_{yy}}(n) + le_Y(n) \overrightarrow{E_y}(n)^* \end{aligned} \quad (2.75)$$

τα όπου το l είναι ο ρυθμός ανανέωσης των βαρών, $e_{X,Y}$ είναι το σφάλμα της εκπαίδευσης που προσπαθεί να ο εκάστοτε αλγόριθμος να μηδενίσει.

Στον DD-LMS το σφάλμα υπολογίζεται από τη σχέση :

$$e_{X,Y}(n) = t_{X,Y} - E_{X,Y}(n) \quad (2.76)$$

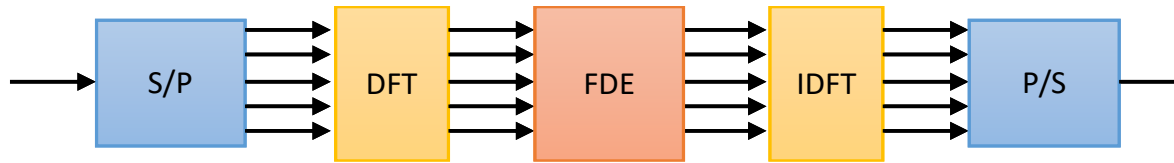
όπου $t_{X,Y}$ είναι τα σύμβολα αναφοράς

Στον αλγόριθμο CMA το σφάλμα υπολογίζεται από τη σχέση :

$$e_{X,Y}(n) = (1 - |E_{X,Y}(n)|^2) E_{X,Y}(n) \quad (2.77)$$

σύμφωνα με την οποία δεν υπάρχει ανάγκη εκπαίδευσης με σύμβολα αναφοράς, αλλά τα βάρη ανανεώνονται έτσι ώστε να το $|E_{X,Y}(n)|^2$ να προσεγγίσει τη μονάδα, καθιστώντας τον CMA λιγότερο πολύπλοκο υπολογιστικά από το DD-LMS.

2.6.3 Εξισορρόπηση στο πεδίο των συχνοτήτων (Frequency-Domain Equalization – FDE)



Σχήμα 2.27: Διάγραμμα συστήματος FDE

Για την εξισορρόπηση οπτικών συστημάτων μεγάλων αποστάσεων απαιτούνται FIR φίλτρα μεγάλου αριθμού taps. Αυτό επιφέρει αύξηση της υπολογιστικής πολυπλοκότητας στα ολοκληρωμένα κυκλώματα συγκεκριμένης εφαρμογής (Application-specific integrated circuit – ASIC). Οι διατάξεις που βασίζουν την ικανότητα εξισορρόπησης των σημάτων στο πεδίο της συχνότητας FDE, υλοποιώντας το διακριτό μετασχηματισμό Fourier, Discrete Fourier Transform - DFT), αποτελούν μια ιδιαίτερα αποδοτική λύση που ταυτόχρονα μειώνει σημαντικά την υπολογιστική πολυπλοκότητα.

Η τυπική διάταξη ενός FDE όπως περιγράφεται στη [62], φαίνεται στο σχήμα 2.27. Το εισερχόμενο μιγαδικό πεδίο αρχικά διαιρείται σε παράλληλα τμήματα μέσω του μετατροπέα Serial to Parallel – S/P. Όλα τα τμήματα παράλληλα μετατρέπονται στο πεδίο των συχνοτήτων μέσω του μετασχηματισμού Fourier (DFT). Κάθε φασματική συνιστώσα ελέγχεται από το FDE, όπου τα βάρη των taps προσαρμόζονται ανάλογα ώστε να σχηματιστεί η κατάλληλη συνάρτηση μεταφοράς. Στη συνέχεια τα τμήματα επιστρέφουν στο πεδίο του χρόνου μέσω του αντίστροφου μετασχηματισμού Fourier (IDFT) και μετατρέπονται σε σειριακά μέσω του μετατροπέα Parallel to Serial (P/S).

Η σημαντικότερη συνεισφορά του FDE σε ένα DSP σύστημα είναι η άρση της χρωματικής διασποράς (CD) που δίνεται από τη σχέση 2.67.

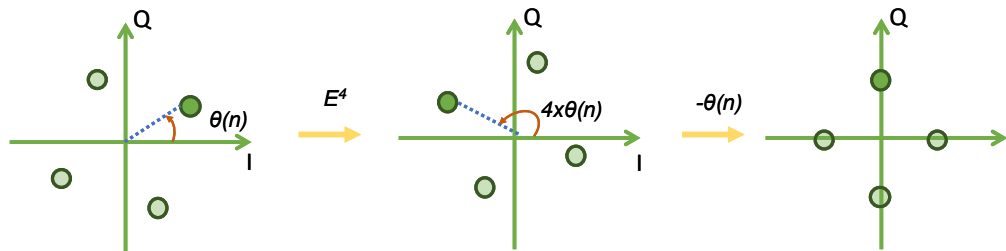
Αυτό επιτυγχάνεται με τον πολλαπλασιασμό της εισόδου του FDE με την αντίστροφη συνάρτηση μεταφοράς της CD:

$$C'(e^{j2\pi\phi T_0}) = \frac{1}{T_0} \sum_k e^{-j \frac{\lambda_c^2}{4\pi c} \beta_2 L \left(\frac{\omega}{2\pi} - \frac{k}{T_0} \right)^2} \quad (2.78)$$

Όπου $T_0 = 2\pi/\omega_0$ η περίοδος που αντιστοιχεί στη συχνότητα δειγματοληψίας.

2.6.3 Εκτίμηση Φάσης Φέρουσας (Carrier-Phase Estimation – CPE)

Στα τυπικά οπτικά σύμφωνα συστήματα ο θόρυβος φάσης $\theta_n(t)$ αυξομειώνεται με πολύ πιο αργούς ρυθμούς από τη διαμόρφωση φάσης του σήματος $\theta_{sm}(t)$. Η εκτίμηση λοιπόν της φάσης της φέρουσας μπορεί να υπολογιστεί με ακρίβεια υπολογίζοντας τον μέσο όρο της φάσης του σε ένα σύνολο συμβόλων.



Σχήμα 2.28: Υπολογισμός φάσης σε QPSK διαμόρφωση

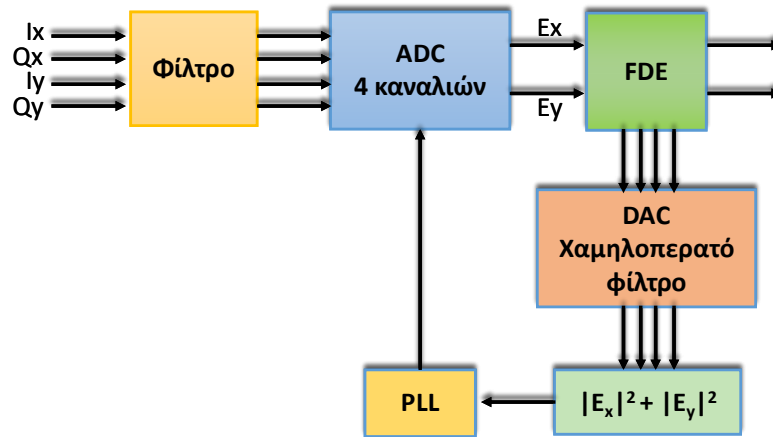
Συγκεκριμένα, αν στην εξίσωση 2.59 υπολογίσουμε το θόρυβο φάσης και διαμόρφωσης σήματος σε ένα σύνολο l τιμών/συμβόλων, $\theta_{sm}(l), \theta_n(l)$ μπορούμε να εκτιμήσουμε τη φάση της φέρουσας. Στην περίπτωση των QPSK αυτό επιτυγχάνεται υψώνοντας στην τέταρτη δύναμη το μιγαδικό πεδίο και αφαιρώντας το θόρυβο φάσης, όπως φαίνεται στο σχήμα 2.28. Αντιστοίχως στα M-PSK σχήματα διαμόρφωσης υψώνουμε στη M-οστή δύναμη. Η μέθοδος αυτή βασίζεται στον αλγόριθμο Viterbi-Viterbi 4^{ου} βαθμού [98].

Στην περίπτωση των QAM διαμορφώσεων ακολουθείται παρόμοια διαδικασία με τη διαφορά πως τα σύμβολα του αστερισμού χωρίζονται σε δύο κατηγορίες. Τα σύμβολα που ανήκουν στην κατηγορία 1 αντιμετωπίζονται όπως την PSK περίπτωση καθώς έχουν την ίδια φάση παρά τις δύο στάθμες ισχύος τους. Τα σύμβολα της κατηγορίας 2 δε συμμετέχουν στη διαδικασία υπολογισμού της φάσης.

Για τη βελτίωση του SNR συνεπώς και την ακριβέστερη εκτίμηση της φάσης υπολογίζεται ο μέσος όρος $2k+1$ δειγμάτων σύμφωνα με τη σχέση :

$$\theta_c(l) = \frac{\arg\left(\sum_{j=-k}^k E(l+j)^M\right)}{M} \quad (2.79)$$

2.6.4 Ανάκτηση συχνότητας ρολογιού



Σχήμα 2.29: DSP και επιμέρους διαδικασίες

Η ανάκτηση της συχνότητας του ρολογιού πραγματοποιείται με διατάξεις παρόμοιες με αυτή που απεικονίζεται στο σχήμα 2.29. Το οπτικό σήμα φιλτράρεται, μετατρέπεται σε ψηφιακό, αφαιρείται η διασπορά με τη διαδικασία του FDE όπως είδαμε στα προηγούμενα και έπειτα μετατρέπεται ξανά σε αναλογική μορφή με τη χρήση DAC.

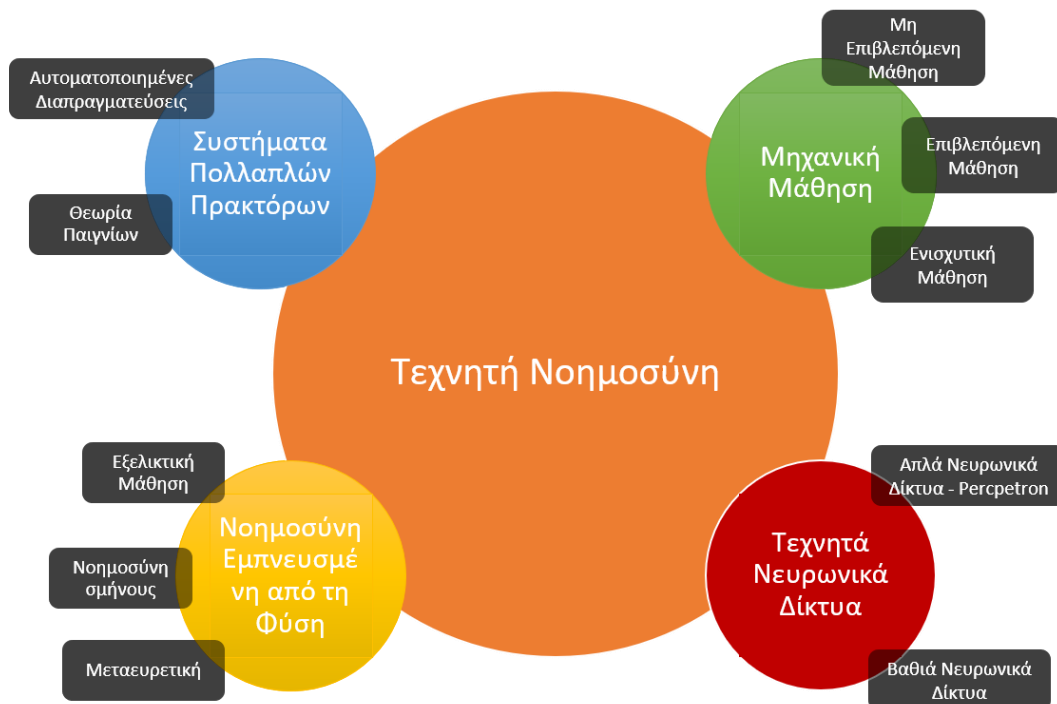
Το εύρος ζώνης του ψηφιακού σήματος στην έξοδο του ADC είναι $1/T$. Ο DAC όμως με τον τετραγωνισμό του σήματος επιτυγχάνει διπλασιασμό του εύρους ζώνης ($2/T$). Τέλος το phase lock loop (PLL) αναλαμβάνει να ανακτήσει την πραγματική συχνότητα του ρολογιού $1/T$, συγχρονίζοντας τον ADC.

3^ο ΚΕΦΑΛΑΙΟ

Εφαρμογές Μηχανικής Μάθησης, τεχνικές εξισορρόπησης σε Συστήματα Οπτικών Επικοινωνιών

3.1 Εισαγωγή

Η μηχανική μάθηση (Machine learning - ML), είναι ένας τύπος τεχνητής νοημοσύνης (Artificial Intelligence-AI) που επιτρέπει στα συστήματα υπολογιστών να μαθαίνουν και να βελτιώνονται από την εμπειρία χωρίς να είναι προγραμματισμένα με έναν ειδικό αλγόριθμο. Με άλλα λόγια, είναι η διαδικασία εκπαίδευσης ενός αλγόριθμου υπολογιστή ώστε να κάνει προβλέψεις ή αποφάσεις με βάση δεδομένα, αναγνωρίζοντας μοτίβα και κάνοντας προσαρμογές για τη βελτίωση της ακρίβειας με την πάροδο του χρόνου.



Σχήμα 3.1: Σχηματική ομαδοποίηση οικοσυστήματος της Τεχνητής Νοημοσύνης

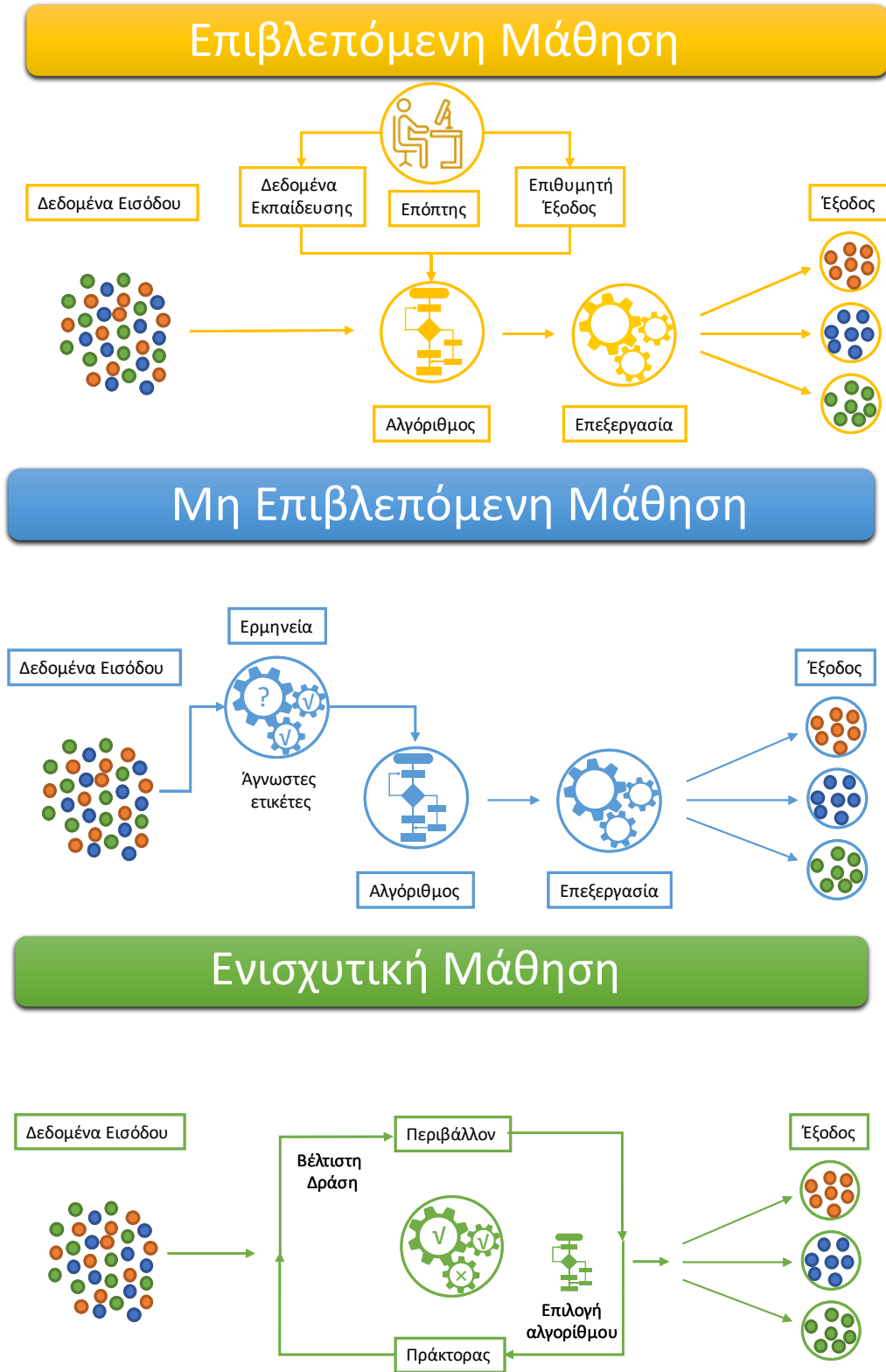
Βάσει του ορισμού αυτού, η Μηχανική Μάθηση έχει ως σκοπό τη δημιουργία μηχανών ικανών να μαθαίνουν, να βελτιώνουν την απόδοσή τους σε κάποια προβλήματα μέσω της αξιοποίησης προηγούμενης γνώσης και εμπειρίας. Ένας σχετικός γενικός ορισμός δίνεται από τον Tom Mitchell, σύμφωνα με τον οποίο η Μηχανική Μάθηση είναι η διαδικασία κατά την οποία ένα πρόγραμμα υπολογιστή μαθαίνει από την εμπειρία E ως προς μια κλάση εργασιών T και ένα μέτρο επίδοσης P , αν η επίδοσή του σε εργασίες της κλάσης T , όπως αποτιμάται από το μέτρο P , βελτιώνεται με την εμπειρία E [99].

Το πεδίο του ML ανήκει στην ευρύτερη οικογένεια της τεχνητής νοημοσύνης όπως χαρακτηριστικά αποτυπώνεται στο σχήμα 3.1. Σε αυτό περιλαμβάνονται οι πιο κλασσικές τεχνικές και αλγόριθμοι, όπως Principal Component Analysis (PCA), k-means, Support Vector Machine (SVM), k-nearest neighbors Algorithm (KNN), Bayesian networks κ.α.[100]. Σημαντική θέση στο οικοσύστημα της AI αποτελούν τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN). Εμπνευσμένα από τα φυσικά νευρωνικά δίκτυα του ανθρώπινου εγκεφάλου, καθένα από τα οποία περιέχει δισεκατομμύρια νευρώνες με μέσο αριθμό συνάψεων ανά νευρώνα 1000 με 10.000, τα ANN επεξεργάζονται πληροφορίες ανταποκρινόμενα δυναμικά σε εξωτερικά ερεθίσματα – εισόδους. Η έντονη ερευνητική δραστηριότητα των τελευταίων δεκαετιών τους έχει προσδώσει μια σχεδόν αυθύπαρκτη υπόσταση, διαχωρίζοντάς τα από τις κλασσικές τεχνικές του ML, όπως φαίνεται και στο σχήμα 3.1. Εντούτοις και στη συνέχεια της εργασίας, ακολουθώντας την πρακτική των σημαντικότερων εργασιών, αναφερόμαστε στα ANN ως υπαγόμενα στην ευρεία έννοια του ML, καθώς η αρχή λειτουργία τους εντάσσεται στο ορισμό που δόθηκε παραπάνω.

3.2 Είδη Μηχανικής Μάθησης

Μιμούμενος τον τρόπο με τον οποίο μαθαίνει ένας άνθρωπος, ο τομέας του ML αναπτύσσει τρεις τρόπους μάθησης : επιβλεπόμενη μάθηση, μη επιβλεπόμενη μάθηση και ενισχυτική μάθηση όπως στο σχήμα 3.2.

α) Επιβλεπόμενη Μάθηση (Supervised Learning) είναι ένας τύπος μηχανικής μάθησης όπου ο αλγόριθμος μαθαίνει από ένα επισημασμένο σύνολο δεδομένων.



Σχήμα 3.2: α) Επιβλεπόμενη μάθηση, β) Μη επιβλεπόμενη μάθηση και γ) Ενισχυτική μάθηση

Στην επιβλεπόμενη μάθηση, τα δεδομένα εισόδου (γνωστά και ως χαρακτηριστικά) και τα αντίστοιχα αποτελέσματα (γνωστά και ως ετικέτες) παρέχονται στον αλγόριθμο κατά τη διάρκεια της εκπαίδευσης. Ο αλγόριθμος μαθαίνει από αυτό το επισημασμένο σύνολο δεδομένων και χρησιμοποιεί αυτή τη γνώση για να κάνει προβλέψεις ή αποφάσεις όταν συναντά νέα δεδομένα. Χρησιμοποιείται σε προβλήματα Ταξινόμησης (Classification) και Παλινδρόμησης (Regression) με πλήθος εφαρμογών όπως ταξινόμηση εικόνας, διάγνωση, πρόβλεψη, βελτιστοποίηση διαδικασίας κ.ά.

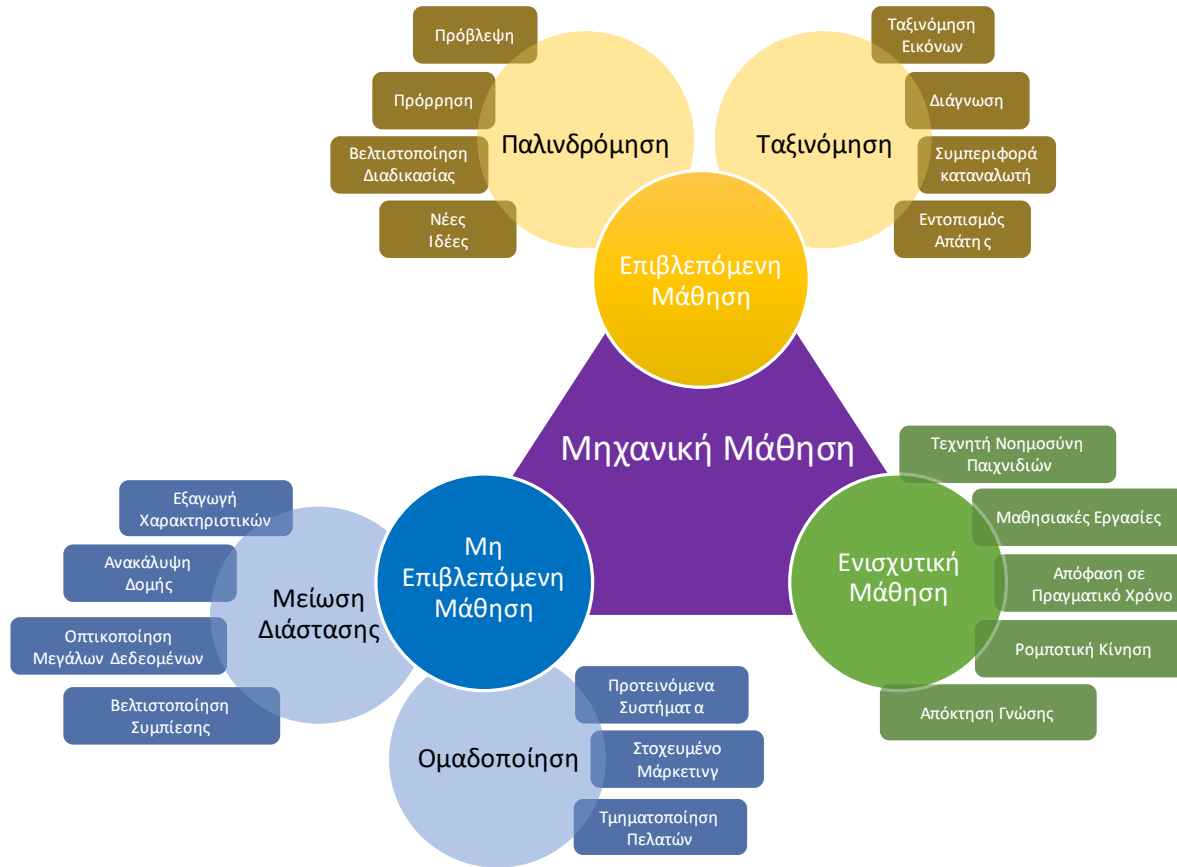
β) Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning), όπου ο αλγόριθμος κατασκευάζει ένα μοντέλο για κάποιο σύνολο εισόδων υπό μορφή παρατηρήσεων χωρίς να γνωρίζει τις επιθυμητές εξόδους. Χρησιμοποιείται σε προβλήματα ομαδοποίησης (Clustering) και μείωσης διάστασης. Συναντάται σε πλήθος εφαρμογών όπως στοχευμένο μάρκετινγκ, βελτιστοποιημένη συμπίεση δεδομένων, οπτικοποίηση μεγάλων δεδομένων (Big Data) κ.α

γ) Ενισχυτική Μάθηση (Reinforcement Learning), όπου ο αλγόριθμος μαθαίνει μια στρατηγική ενεργειών μέσα από άμεση αλληλεπίδραση με το περιβάλλον. Χρησιμοποιείται κυρίως σε προβλήματα Σχεδιασμού (Planning), όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ, λήψη απόφασης σε πραγματικό χρόνο, χρήση τεχνητής νοημοσύνης σε παιχνίδια υπολογιστή κ.α.

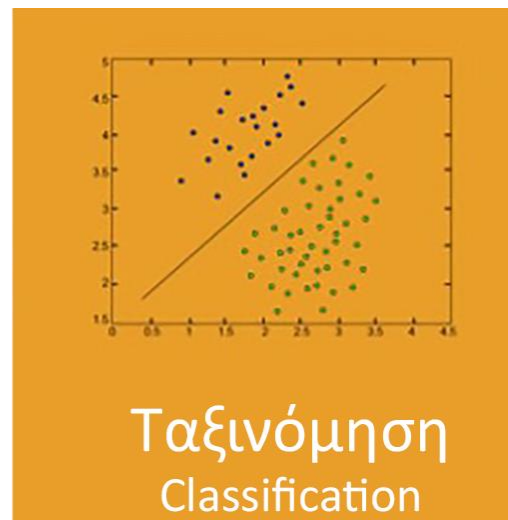
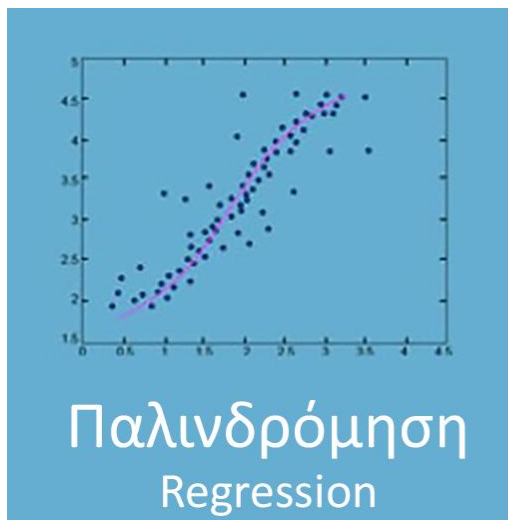
Στο σχήμα 3.3 οπτικοποιείται ο τομέας της Μηχανικής Μάθησης απαριθμώντας τις σημαντικότερες γνωστές εφαρμογές της σε πραγματικά σύγχρονα προβλήματα.

Έχοντας ως κριτήριο τη μορφή των δεδομένων που επεξεργάζεται ο αλγόριθμος οι δύο βασικές μέθοδοι που ακολουθούνται για την εκπαίδευση μοντέλων μηχανικής μάθησης είναι η παλινδρόμηση (regression) και η ταξινόμηση (classification), (σχήμα 3.4).

Η παλινδρόμηση είναι η διαδικασία που ακολουθούμε για την εύρεση μιας βέλτιστης συνάρτησης που θα προσδιορίζει τα δεδομένα ή θα πραγματοποιεί πρόβλεψη μελλοντικών συνεχών πραγματικών τιμών. Η παλινδρόμηση με πολλές μεταβλητές ή χαρακτηριστικά στην είσοδο για την εκπαίδευση του αλγόριθμου, είναι γνωστή ως παλινδρόμηση πολλαπλών μεταβλητών (multivariate regression problem).



Σχήμα 3.3: Σχηματική ομαδοποίηση οικοσυστήματος της Μηχανικής Μάθησης



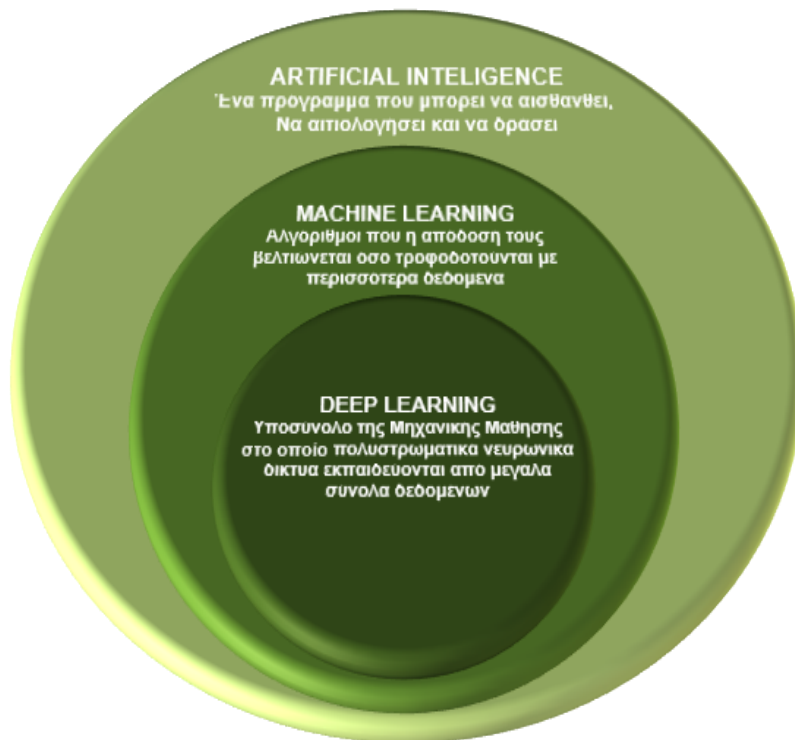
Σχήμα 3.4: Παλινδρόμηση (regression) και Ταξινόμηση (classification)

Εάν στην παλινδρόμηση, οι τιμές εισόδου εξαρτώνται ή ταξινομούνται με βάση τον χρόνο, τότε είναι γνωστό ως πρόβλημα πρόβλεψης χρονοσειρών (time series forecasting problem).

Από την άλλη μεριά, ένας αλγόριθμος ταξινόμησης ταξινομεί το απαιτούμενο σύνολο δεδομένων σε μία από δύο ή περισσότερες ετικέτες. Ένας αλγόριθμος που ασχολείται με δύο κλάσεις ή κατηγορίες είναι γνωστός ως δυαδικός ταξινομητής (binary classifier) και εάν υπάρχουν περισσότερες από δύο κλάσεις τότε μπορεί να ονομαστεί αλγόριθμος ταξινόμησης πολλαπλών κλάσεων (multi-class classification algorithm).

3.3 Βαθιά Μάθηση – Deep Learning

Οι απλοί αλγόριθμοι της μηχανικής που αναφέρθηκαν προηγουμένως δουλεύουν αρκετά ικανοποιητικά σε μεγάλου εύρους εφαρμογές. Όμως δεν μπορούν να επιλύσουν κομβικά προβλήματα της τεχνητής νοημοσύνης, όπως αναγνώριση εικόνας ή ομιλίας. Όταν η διάσταση των δεδομένων αυξάνεται σημαντικά, οι απλοί ML αλγόριθμοι συναντούν



Σχήμα 3.5: Βαθιά Μάθηση

ανυπέρβλητες δυσκολίες στην επίλυση αυτών των προβλημάτων. Τα τελευταία χρόνια ο όρος Βαθιά Μάθηση (Deep Learning - DL) συναντάται όλο και περισσότερο. Η βαθιά μάθηση αποτελεί ένα υποπεδίο του ML που πραγματεύεται μεγάλης διάστασης, πολυστρωματικά (multi-layer) ANNs, ενώ επεξεργάζεται και εκπαιδεύεται από μεγάλο αριθμό δεδομένων πολλών διαστάσεων (σχήμα 3.5). Το DL συνεπώς καλείται να υπερκεράσει τα εμπόδια των πολυδιάστατων και σύνθετων προβλημάτων που αδυνατούν να επεξεργαστούν επαρκώς τα απλά ML μοντέλα. Σε αντίθεση πχ με το Perceptron [101] που αποτελεί την πλέον απλή δομή νευρωνικού δικτύου με τη δυνατότητα να μάθει τα βάρη που καθορίζουν τις κατηγορίες όταν τροφοδοτηθεί με δεδομένες εισόδους σε γνωστές επιθυμητές εξόδους/κατηγορίες, (σχήμα 3.6α), το Βαθύ Νευρωνικό Δίκτυο (Deep Neural Network DNN) προσιδιάζει καλύτερα στη δομή η καλύτερα στην απόδοση του νευρωνικού δικτύου του ανθρώπινου εγκεφάλου (σχήμα 3.6β)

3.4 Προσωτροφοδοτούμενα Νευρωνικά Δίκτυα– Βασικές τεχνικές εκπαίδευσης

Τα DNNs ονομάζονται επίσης προσωτροφοδοτούμενα (feedforward neural networks - FNN) ή πολυστρωματικά Perceptrons (multilayer perceptrons -MLPs).

Ας υποθέσουμε ως παράδειγμα ένα πρόβλημα ταξινόμησης, στο οποίο η είσοδος x αντιστοιχίζεται σε μία κατηγορία y μέσω της συνάρτησης f σύμφωνα με τη σχέση :

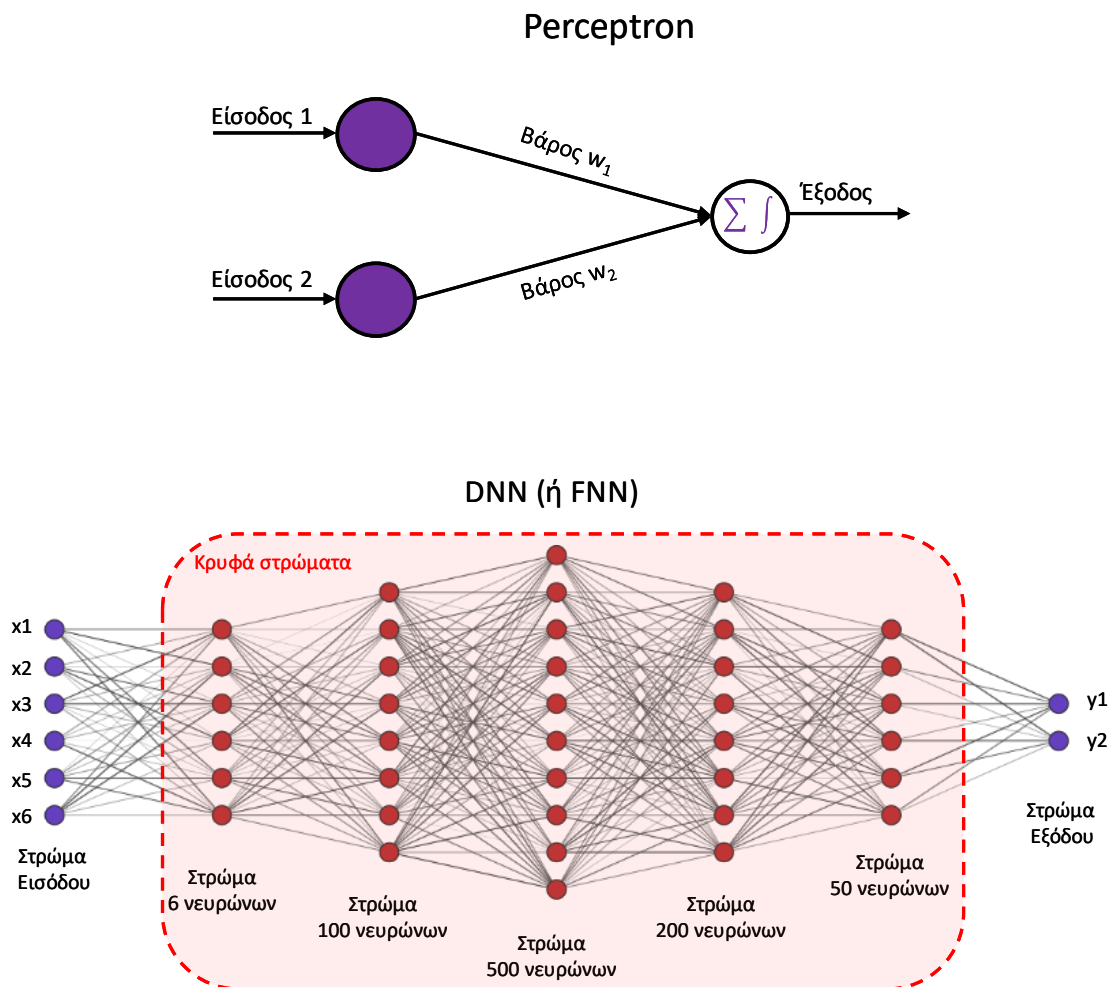
$$y = f(x) \quad (3.1)$$

Ο σκοπός ενός FNN είναι να προσεγγίσει μία συνάρτηση f^* όσο το δυνατόν πιο κοντινή, ιδανικά ταυτόσημη, της f . Για την επίτευξη του σκοπού το FNN ορίζει μία απεικόνιση $y=f^*(\mathbf{x};\boldsymbol{\theta})$ και μαθαίνει την τιμή των παραμέτρων $\boldsymbol{\theta}$ που οδηγούν στη βέλτιστη προσέγγιση της f [102].

Η ονομασία feedforward προκύπτει από το γεγονός ότι η πληροφορία ρέει από την είσοδο x προς την έξοδο y χωρίς ανατροφοδότηση. Στο σχήμα 3.6β απεικονίζεται ένα τυπικό FNN. Αποτελείται από το στρώμα εισόδου (Input Layer), το μέγεθος του οποίου καθορίζεται από τον αριθμό των χαρακτηριστικών του προβλήματος (6 εισοδοι στο σχήμα: $x_1, x_2, x_3, x_4, x_5, x_6$), το στρώμα εξόδου (Output Layer), το μέγεθος του οποίου

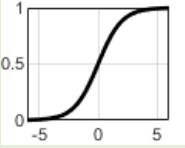
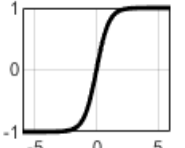
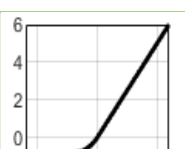

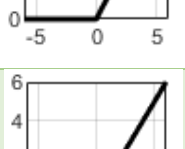

καθορίζεται από τον αριθμό των κατηγοριών (classes) στην περίπτωση των προβλημάτων ταξινόμησης, (2 έξοδοι στο σχήμα: y_1, y_2). Η ειδοποιός διαφορά του από το perceptron (σχ 3.5α) όμως είναι τα επάλληλα κρυφά στρώματα (Hidden Layers) ενδιάμεσα των επιπέδων εισόδου-εξόδου, (5 κρυφά στρώματα στο σχήμα). Ο αριθμός τους δίνει και το βάθος του δικτύου ερμηνεύοντας με αυτό τον τρόπο την ονομασία βαθιά νευρωνικά δίκτυα. Καθένα από αυτά επιδρά στα δεδομένα εισόδου με μια επιμέρους συνάρτηση f^1, f^2, \dots, f^k αντίστοιχα έτσι ώστε η συνάρτηση του συνολικού νευρωνικού δικτύου να δίνεται από τη σχέση:

$$f(x) = f^k \dots (f^2(f^1(x))) \quad (3.2)$$



Σχήμα 3.6: α) Perceptron, απλό νευρωνικό δίκτυο, β) DNN ή FNN, σύνθετο νευρωνικό δίκτυο με επάλληλα κρυφά στρώματα νευρώνων.

Ονομάζονται κρυφά γιατί τα εσωτερικά αποτελέσματα για τα δεδομένα εισόδου των επιμέρους συναρτήσεων δεν είναι εμφανή, παρά μόνο το αποτέλεσμα της $f(x)$ στο στρώμα εξόδου.

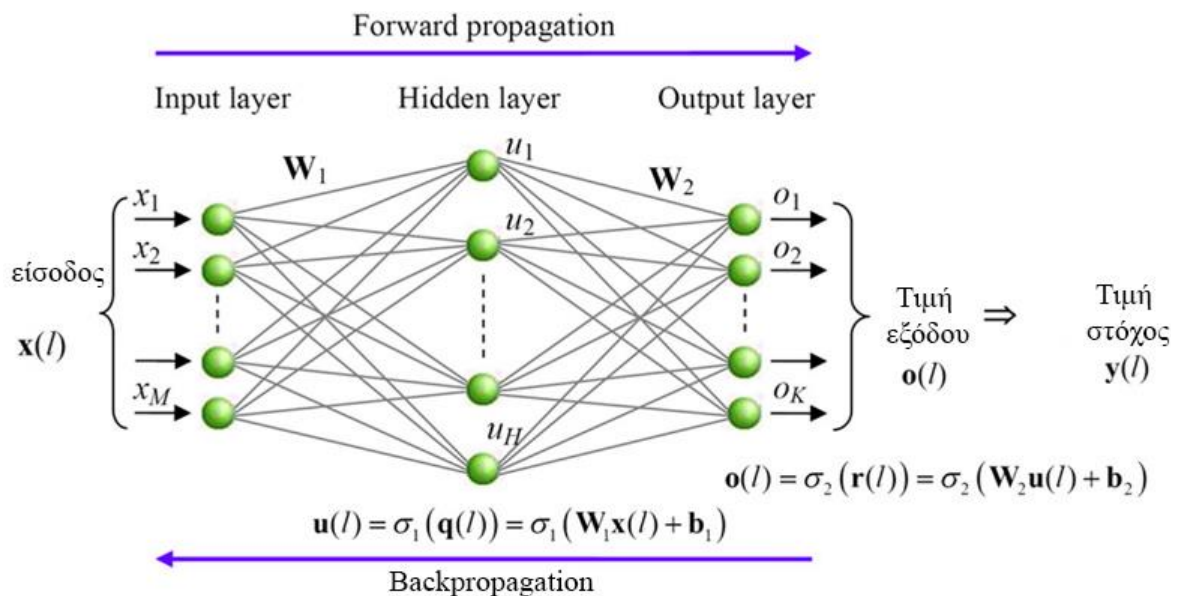
Όνομα	Συνάρτηση	Μορφή
Sigmoid	$\frac{1}{1 + 1e^{-x}}$	
Hyperbolic tangent (tanh)	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$	
Exponential linear unit (ELU)	$\begin{cases} a(e^x - 1) & x \leq 0 \\ x & x > 0 \end{cases}$	
Rectified linear unit (ReLU)	$\begin{cases} 0 & x \leq 0 \\ x & x > 0 \end{cases}$	
Leaky rectified linear unit (Leaky ReLU)	$\begin{cases} ax & x \leq 0 \\ x & x > 0 \end{cases}$	
Gaussian	e^{-x^2}	

Πίνακας 3.1: Συναρτήσεις ενεργοποίησης

Κάθε κρυφό στρώμα, κατά αντιστοιχία με νευρωνικά δίκτυα του εγκεφάλου διαθέτει μεγάλο αριθμό νευρώνων (6, 100, 500, 200, 50 στο σχήμα). Ο αριθμός αυτό διαμορφώνει και το πλάτος (width) του στρώματος και του νευρωνικού δικτύου συνολικά. Η ιδέα πίσω

από αυτή την αρχιτεκτονική είναι πως κάθε νευρώνας καθενός στρώματος θα λαμβάνει είσοδο όχι μόνο από ένα αλλά από όλους τους νευρώνες του προηγούμενου στρώματος και ανάλογα τα χαρακτηριστικά του θα ενεργοποιείται ή όχι και ακολούθως θα συνεισφέρει σε μεγάλο ή σε μικρό βαθμό στους νευρώνες των επόμενων σταδίων και ούτω καθεξής.

Καθένας από τους νευρώνες συνδέεται με αυτούς του επόμενου στρώματος με μία συνάρτηση ενεργοποίησης (activation function), προσδίδοντας στον αλγόριθμο εκπαίδευσης του νευρωνικού δικτύου μη γραμμικά χαρακτηριστικά. Οι συνηθέστερες συναρτήσεις ενεργοποίησης παρουσιάζονται στον πίνακα 3.1



Σχήμα 3.7: Διαδικασία εκπαίδευσης (αλγόριθμος Gradient Descent) τυπικού νευρωνικού δικτύου

Ο βασικότερος αλγόριθμος εκπαίδευσης των νευρωνικών δικτύων, που ονομάζεται αλγόριθμος καθόδου βασισμένης στην κλίση (Gradient Descent, GD), βασίζεται στον μηχανισμό διόρθωσης των βαρών των συνάψεων (συνδέσεων) μεταξύ των νευρώνων [103]. Αρχικά ακολουθώντας την πρόσω διάδοση υπολογίζεται η έξοδος του δικτύου λαμβάνοντας υπόψη την συνάρτηση ενεργοποίησης σ και το bias b όλων των στρωμάτων που το αποτελούν. Στη συνέχεια εφαρμόζεται η αντίστροφη διάδοση (backpropagation)

κατά την οποία υπολογίζονται οι μερικές παράγωγοι, διαδίδοντας το σφάλμα (διαφορά από την τιμή στόχο) από το τέλος προς την αρχή του δικτύου. Μέσω της ελαχιστοποίησης της συνάρτησης απώλειας πραγματοποιείται η ενημέρωση των παραμέτρων (βαρών) του δικτύου [104]. Ο υπολογισμός των μερικών παραγώγων επαναλαμβάνεται μέχρι να επιτευχθεί σύγκλιση (τοπικό ή ολικό ελάχιστο) ή να ολοκληρωθεί ένας προκαθορισμένος αριθμός επαναλήψεων με στόχο τη βελτίωση των προβλέψεων του δικτύου. Μια περιγραφική σχηματική αναπαράσταση του μηχανισμού εκπαίδευσης αποτυπώνεται στο σχήμα 3.7.

Ουσιαστικά το νευρωνικό δίκτυο παίρνει τα δεδομένα εισόδου $x(l)$, ορίζει τα βάρη W αρχικά τυχαία και στη συνέχεια συγκρίνει την εκτίμηση του $o(l)$ με την τιμή στόχο $y(l)$. Το αποτέλεσμα που προκύπτει σύμφωνα με την επιλεγμένη συνάρτηση απώλειας, χρησιμοποιείται για να πραγματοποιηθεί η διόρθωση χρησιμοποιώντας την αντίστροφη φορά (backpropagation) $u(l)$. Η ανανέωση των τιμών των βαρών υπολογίζεται από τον τύπο:

$$W_l^{<t+1>} = W_l^{<t>} + a_{lr} \nabla \hat{L} (W_l^{<t>}) \quad (3.3)$$

Όπου $W_l^{<t>}$ ο πίνακας βαρών του στρώματος l στη χρονική στιγμή t , \hat{L} η τιμή της απώλειας που προκύπτει από τα συγκεκριμένα βάρη, $\nabla \hat{L}$ το ανάδελτα του \hat{L} και $a_{lr} > 0$ ο ρυθμός εκπαίδευσης, παράμετρος που καθορίζει το εύρος με τον οποίο θα αυξομειώνονται τα βάρη.

Οι συναρτήσεις απώλειας (loss functions) που χρησιμοποιούνται πιο συχνά στα μοντέλα μηχανικής μάθησης συνοψίζονται στον Πίνακα 3.2.

Όνομα	Συνάρτηση
Mean Square Error	$\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$
Mean Absolute Error	$\frac{1}{N} \sum_{i=1}^N y_i - \hat{y}_i $
Mean Bias Error	$\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)$

Binary Cross-entropy	$-\frac{1}{N} \sum_{i=1}^N [y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)]$
Multiclass Cross-entropy	$-\frac{1}{N} \sum_{i=1}^N (y_i \log \hat{y}_i)$ *για κάθε τάξη/κατηγορία

Πίνακας 3.2 Συναρτήσεις απώλειας

Η μέθοδος μέσου τετραγωνικού σφάλματος (Mean Square Error, MSE) είναι η πλέον διαδεδομένη συνάρτηση απώλειας σε προβλήματα παλινδρόμησης, ενώ οι Binary Cross-entropy και Multiclass Cross-entropy είναι οι συνηθέστερες σε δυαδικά ή πολλών τάξεων προβλήματα ταξινόμησης αντίστοιχα.

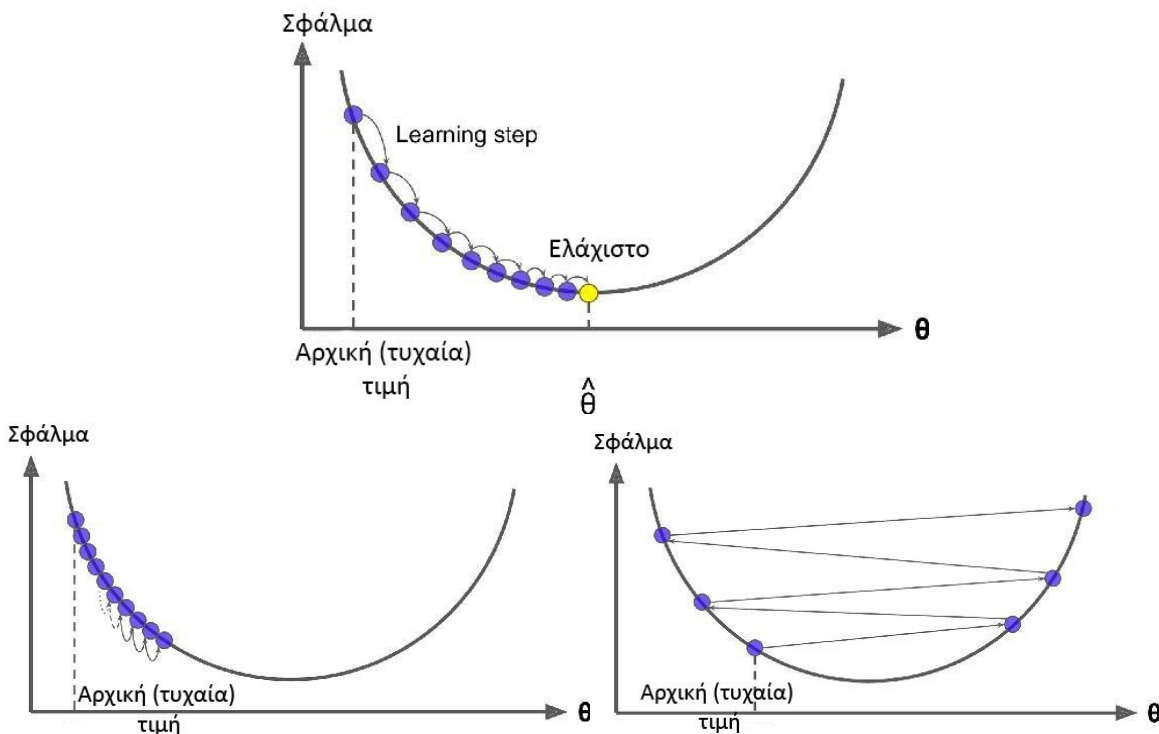
Το στρώμα εξόδου του σχήματος 3.7 αποτελείται από ένα διάνυσμα της μίας ή περισσότερων εξόδων, ίσου με το μέγεθος των χαρακτηριστικών που υπολογίζει το νευρωνικό δίκτυο. Στην περίπτωση των προβλημάτων παλινδρόμησης, ο αριθμός των εξόδων είναι ίσος με τη διάσταση του προβλήματος, πχ μία μόνο έξοδο για μοναδιάστατα προβλήματα ταξινόμησης όπως η πρόβλεψη θερμοκρασίας κτλ. Στην περίπτωση των προβλημάτων ταξινόμησης το διάνυσμα περιέχει το σύνολο των πιθανών κατηγοριών, πχ στην περίπτωση προβλήματος αναγνώρισης χειρόγραφων δεκαδικών αριθμών, όπως το σετ MNIST [105], το πλήθος των εξόδων είναι 10 κάθε μία από τις οποίες μέσω του μηχανισμού εκπαίδευσης αποκτά και μια τιμή. Η τελική επιλογή της εξόδου γίνεται με τη συνάρτηση softmax :

$$\text{softmax}(z) = \frac{1}{\sum_{k=1}^K e^{z_k}} [e^{z_1} \ e^{z_2} \ \dots \ e^{z_k}] \quad (3.4)$$

Όπως είδαμε παραπάνω ο αλγόριθμος GD είναι ένας αρκετά γενικός αλγόριθμος ικανός να βρει τις ιδανικές λύσεις σε αρκετά προβλήματα. Ο μηχανισμός του βασίζεται στην απλή ιδέα του υπολογισμού της μερικής παράγωγου της συνάρτησης απώλειας και της μετατόπισης αυτής στην κατεύθυνση του ελάχιστου. Πολύ σημαντική παράμετρος του GD είναι ο ρυθμός των βημάτων υπολογισμού ή διαφορετικά ο ρυθμός εκπαίδευσης (learning rate). Πιθανή επιλογή πολύ μικρού ρυθμού εκπαίδευσης οδηγεί σε πολύ μεγάλες

καθυστερήσεις στη σύγκλιση του μοντέλου ενώ πολύ μεγάλες τιμές, σε πιθανή απόκλιση και αδυναμία προσέγγισης του τοπικού ελάχιστου (σχήμα 3.8).

Για την επίτευξη της πιο αποδοτικής λειτουργίας του ρυθμού εκπαίδευσης έχουν αναπτυχθεί αρκετοί αλγόριθμοι βελτιστοποίησης, με σημαντικότερο τον Adam [106]. Πρόκειται για ένα προσαρμοστικό αλγόριθμο που ενσωματώνει τον μηχανισμό της ορμής (momentum), δηλαδή αυξάνει το μήκος του ρυθμού εκπαίδευσης όταν απέχει σημαντικά από την ελάχιστη τιμή του σφάλματος ενώ μικραίνει όσο πλησιάζει σε αυτήν καθιστώντας την προσέγγιση του σημείου σύγκλισης αποδοτικότερη. Άλλοι σημαντικοί αλγόριθμοι είναι ο RMSProp [107] και ο AdaGrad [108].



Σχήμα 3.8: GD με κανονικό (α), μικρό (β) και μεγάλο (γ) ρυθμό εκπαίδευσης.

Οι προσεγγίσεις του GD που ακολουθούνται είναι α) Batch Gradient Decent, β) Stochastic Gradient Decent και γ) Mini-batch Gradient Decent. Η σημαντικότερη διαφορά τους είναι η επιλογή των δειγμάτων των δεδομένων εκπαίδευσης (training set) που χρησιμοποιούν.

Η προσέγγιση Batch Gradient Decent χρησιμοποιεί το σύνολο των δεδομένων εκπαίδευσης και αυτό τον καθιστά πολύ γρήγορο στην εκτέλεση. Παρόλα αυτά στις σύγχρονες διαθέσιμες μονάδες επεξεργασίας και προσωρινής αποθήκευσης μνήμης η υλοποίηση αυτής της τεχνικής δεν είναι αποδοτική ιδιαίτερα όταν τα δεδομένα εκπαίδευσης είναι αρκετά εκτεταμένα.

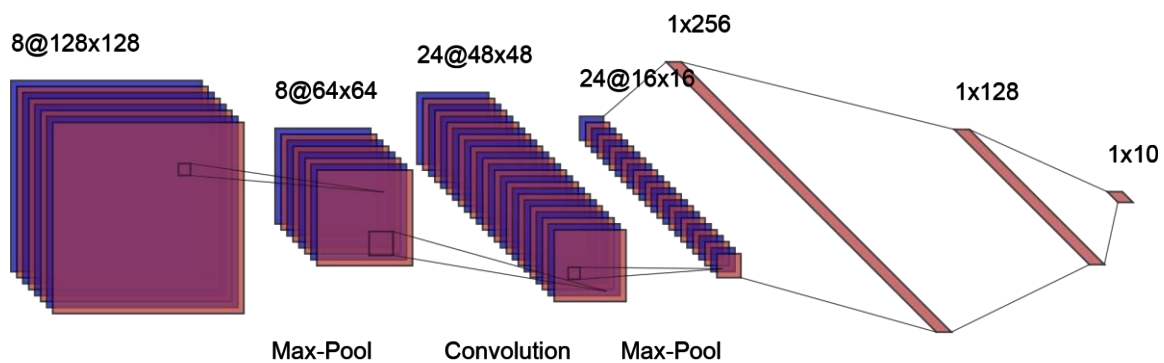
Από την άλλη μεριά η προσέγγιση Stochastic Gradient Decent παίρνει τυχαία (στοχαστικά) δείγματα και τα προσθέτει στην εκπαίδευση. Μια παραλλαγή του αποτελεί η προοδευτική εκπαίδευση (online / incremental training) δηλαδή η προσθήκη νέων δεδομένων εκπαίδευσης για τη βελτιστοποίηση των ήδη εκπαιδευμένων αλγορίθμων μηχανικής μάθησης. Η παραλλαγή αυτή βρίσκει εφαρμογή στα σύγχρονα μοντέλα βαθιάς μάθησης και είναι ιδιαίτερα αποδοτική, όμως είναι αρκετά αργή για εκπαίδευση νέων μοντέλων από μηδενική βάση.

Η ενδιάμεση λύση, που χρησιμοποιείται ευρέως στις σύγχρονες εφαρμογές βαθιάς μάθησης, είναι η Mini-batch Gradient Decent. Σε αυτή, το στρώμα εισόδου τροφοδοτείται διαδοχικά με ένα κλάσμα του συνόλου των δεδομένων εκπαίδευσης, τα πακέτα (batches). Όταν το νευρωνικό δίκτυο εκπαιδευτεί για όλα τα batches, δηλαδή με όλα τα διαθέσιμα δεδομένα εκπαίδευσης, εκτελώντας και τη πρόσω διάδοση και την οπισθοδιάδοση, λέμε πως έχει διανύσει μια εποχή (epoch).

Η ρύθμιση του μέγεθος του κάθε batch σε συνδυασμό με το ρυθμό εκπαίδευσης και τον αριθμό των εποχών (epochs), δίνει τη δυνατότητα της παραμετροποίησης και βελτιστοποίησης του μοντέλου που καλείται να επιλύσει το συγκεκριμένο πρόβλημα. Δεν υπάρχει κάποιος σαφής κανόνας επιλογής της τιμής των παραμέτρων αυτών, όπως δεν υπάρχει για τον ακριβή αριθμό των κρυφών στρωμάτων και των νευρώνων που τα αποτελούν. Η διαδικασία κατά την οποία αναζητούνται οι βέλτιστες τιμές για την επίλυση του συγκεκριμένου προβλήματος ονομάζεται βελτίωση υπερπαραμέτρων (hyperparameter tuning) και συνήθως εκτελείται με τη βοήθεια διαδοχικών δοκιμών με διάφορα σενάρια χρησιμοποιώντας αλγορίθμους βελτιστοποίησης (όχι αποκλειστικά) κατά Bayes [109].

Στα νευρωνικά δίκτυα που μελετήθηκαν σε αυτή τη διατριβή επιλέχθηκε η λύση του Mini-batch Gradient Decent, καθώς οι βαθμοί ελευθερίας τους οποίους προσδίδει η δυνατότητα κλιμάκωσης της εκπαίδευσης των απαιτητικών μοντέλων που χρησιμοποιήθηκαν, κρίθηκε ιδιαίτερα ωφέλιμη για τη βελτιστοποίησή τους αλλά και αναγκαία για την ομαλή εκτέλεση τους στα διαθέσιμα υπολογιστικά περιβάλλοντα.

3.5 Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks – CNN)



Σχήμα 3.9: Τυπικό Συνελικτικό Νευρωνικό Δίκτυο (CNN), για την ταξινόμηση εικόνων

Τα συνελικτικά νευρωνικά δίκτυα είναι μια κατηγορία FNN που αρχικά είχαν προταθεί για τη βελτιστοποίηση της αναγνώρισης εικόνων. Η τεράστια επιτυχία τους σε αυτό τον τομέα οδήγησε την υιοθέτησή τους σε προβλήματα όπως ανάλυση βίντεο και υπολογιστική όραση. Σύντομα όμως επεκτάθηκε και σε μονοδιάστατα προβλήματα όπως την επεξεργασία σήματος [102].

Τα CNN εμπνέονται ιστορικά από τη δομή και λειτουργία του οπτικού φλοιού, στον οποίο οι μεμονωμένοι φλοιώδεις νευρώνες αντιδρούν μόνο σε ερεθίσματα σε ένα στενό τμήμα του οπτικού πεδίου, γνωστό ως δεκτικό πεδίο και τα δεκτικά πεδία των διαφορετικών νευρώνων εν μέρει συντίθενται, καλύπτοντας ολόκληρο το οπτικό πεδίο.

Ο ρόλος του CNN είναι να μειώνει τη διάσταση των εικόνων σε μια μορφή που είναι πιο εύκολη στην επεξεργασία, χωρίς να χάσει χαρακτηριστικά που είναι κρίσιμα για την επίτευξη καλής πρόβλεψης, καθώς είναι σε θέση να καταγράψει με επιτυχία τις χωρικές και χρονικές εξαρτήσεις σε μια εικόνα ή ένα πίνακα εισόδου.

Αποτελούνται από ένα ή περισσότερα στρώματα συνέλιξης (convolution layers) μέσω της εφαρμογής σχετικών φίλτρων που ονομάζονται πυρήνες (kernels), από στρώματα συγκέντρωσης (pooling layers) ακολουθούμενα από πλήρως συνδεδεμένα στρώματα (Fully Connected Layers, FCL) συνοδευόμενα από συναρτήσεις ενεργοποίησης, (σχήμα 3.9)

Οι πυρήνες είναι πίνακες βαρών που εκπαιδεύονται με την επαναλαμβανόμενη εκτέλεση του αλγορίθμου GD παράγοντας με αυτό τον τρόπο πίνακες χαρακτηριστικών που οδηγούνται με τη σειρά τους στο επόμενο στρώμα.

Στην απλή περίπτωση επεξεργασίας μονοδιάστατων δεδομένων εισόδου με N_{ch} κανάλια, η διάσταση του πυρήνα δίνεται από τον τύπο $L_{ker} \times N_{ch}$, όπου L_{ker} το μήκος του πυρήνα. Εφαρμόζοντας την πράξη της συνέλιξης των δεδομένων εισόδου με τον πυρήνα, το m -οστό στοιχείο του πίνακα χαρακτηριστικών στο l στρώμα δίνεται από τον τύπο:

$$x_l[m] = \sigma_l(z_l[m] + b_l[m]) \quad (3.5)$$

όπου

$$z_l[m] = \sum_{k=1}^{N_c} \sum_{i=1}^{L_{ker}} W[i, k] x_{l-1} \left[m - \frac{L_{ker}}{2} + i, k \right] \quad (3.6)$$

W , ο πίνακας βαρών του πυρήνα και $W[i, k]$ το βάρος της i σειράς του k καναλιού.

Τα στρώματα συγκέντρωσης αποτελούν το μηχανισμός μείωσης της διάστασης που συναντάμε συνήθως μετά από ένα στρώμα συνέλιξης. Σε αντίθεση με τα συνελκτικά στρώματα, τα στρώματα συγκέντρωσης λειτουργούν χωριστά σε κάθε κανάλι, εκτελώντας κάποιο είδος χωρικής αναλλοίωσης. Οι τύποι μέγιστης συγκέντρωσης (max-pooling, MP) και μέσης συγκέντρωσης (average-pooling, AP) αποτελούν τους πιο συνηθισμένους, σύμφωνα με τους οποίους εξάγονται οι μέγιστες και μέσες τιμές του προηγούμενου πυρήνα, αντίστοιχα [102].

3.6 Επαναλαμβανόμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks – RNN)

Τα Επαναλαμβανόμενα Νευρωνικά Δίκτυα - RNN είναι μια κατηγορία νευρωνικών δικτύων που έχουν σχεδιαστεί για να λειτουργούν σε χρονικές ακολουθίες δεδομένων με συσχετισμένα δείγματα. Βρίσκουν εφαρμογή σε εφαρμογές αναγνώρισης ομιλίας, πρόβλεψης χρονοσειρών κ.ά. Στη βιβλιογραφία και ιδιαίτερα στην ελληνική ορολογία συναντώνται και ως Αναδρομικά Δίκτυα λόγω της σχέσης παρελθόντος μέλλοντος.

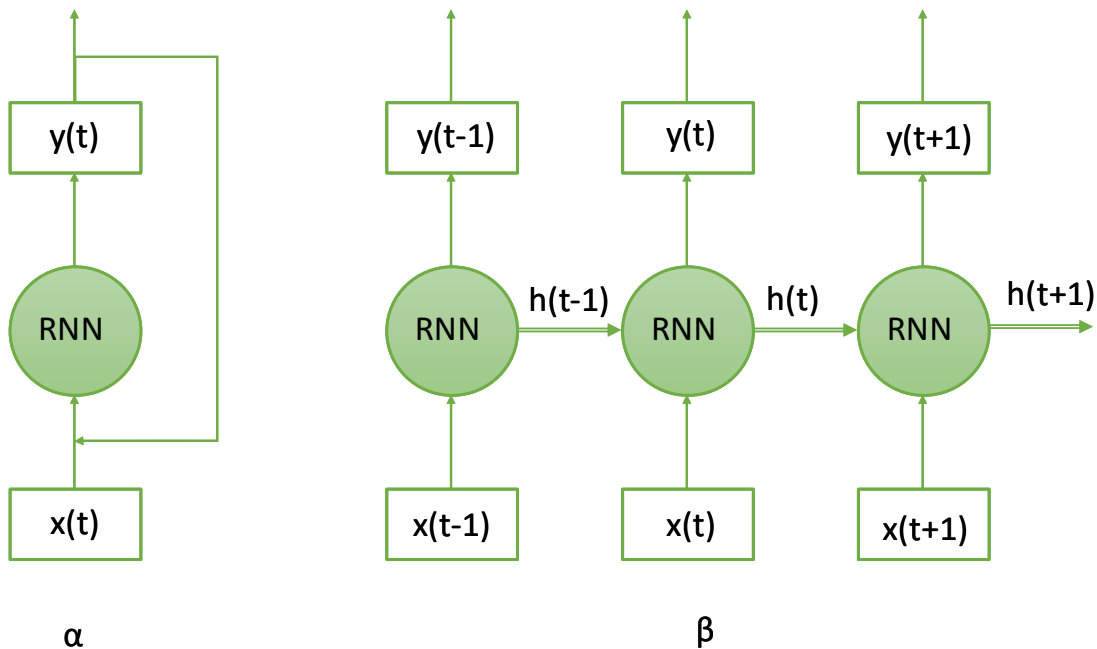
Εμπνέονται και μιμούνται τη φύση των χρονοεξαρτώμενων φαινομένων και τη μνήμη που αυτά εμπεριέχουν. Αποτελούνται από επαναλαμβανόμενες μονάδες (units) ή κύτταρα (cells) των οποίων οι καταστάσεις επηρεάζονται όχι μόνο από την τρέχουσα είσοδο αλλά και από παρελθοντικά ή ακόμα και μελλοντικά, (στην περίπτωση των αμφίδρομων RNN) χρονικά βήματα. Σε ένα RNN η «μνήμη» αναφέρεται στην ικανότητα του δικτύου να διατηρεί και να χρησιμοποιεί πληροφορίες από προηγούμενα χρονικά βήματα. Επιτρέπει στο δίκτυο να διατηρεί το περιβάλλον και τις εξαρτήσεις μεταξύ διαδοχικών εισόδων. Η μνήμη υλοποιείται μέσω της έννοιας μιας κρυφής κατάστασης (hidden state).

Η κρυφή κατάσταση, γνωστή και ως κρυφή αναπαράσταση, είναι ένα διάνυσμα τιμών που καταγράφει την εσωτερική κατάσταση του δικτύου σε ένα συγκεκριμένο χρονικό βήμα. Συνοψίζει τις πληροφορίες που έχει δει το δίκτυο μέχρι στιγμής και κωδικοποιεί σχετικές πληροφορίες από προηγούμενα χρονικά βήματα. Η κρυφή κατάσταση ενημερώνεται αναδρομικά καθώς το δίκτυο επεξεργάζεται κάθε νέα είσοδο, ενσωματώνοντας τόσο την τρέχουσα είσοδο όσο και την προηγούμενη κρυφή κατάσταση. Χρησιμεύει ως μορφή μνήμης για το RNN, επιτρέποντάς του να μοντελοποιεί χρονικές εξαρτήσεις και να κάνει προβλέψεις με βάση ιστορικές πληροφορίες.

Συγκεκριμένα το RNN επεξεργάζεται τις χρονικές ακολουθίες σειριακά, επεξεργαζόμενο ένα προς ένα τα εισερχόμενα στοιχεία της και η-οι κρυφή-ές μονάδα-ες (hidden units) συσσωρεύει-ουν την πληροφορία για όλα τα προηγούμενα στοιχεία της ακολουθίας. Οι διαδοχικές πληροφορίες διατηρούνται στην κρυφή κατάσταση του RNN, η οποία χρησιμοποιείται για να συνεκτιμηθεί στην επεξεργασία κάθε νέου χρονικού βήματος [110].

Στο σχήμα 3.10α αναπαρίσταται η αρχή λειτουργίας ενός απλού RNN, το οποίο στην είσοδό του εκτός των νέων (χρονικών) στοιχείων δέχεται και τα δεδομένα των

προηγούμενων. Παρά το γεγονός ότι τα RNNs έχουν τη δυνατότητα να επεξεργάζονται δεδομένα από πολλά χρονικά βήματα, καθώς ο αριθμός των χρονικών βημάτων αυξάνεται, το απλό αυτό διάγραμμα σύνδεσης αρχίζει να χάνει το νόημα του. Αντί αυτού ένας χρήσιμος τρόπος οπτικοποίησης της λειτουργίας του είναι το «ξεδίπλωμα» (unfolding) σε μεγάλες χρονικές ακολουθίες (σχήμα 3.10β), όπου x , y , h , η είσοδος, έξοδος και το περιεχόμενο της κρυφής κατάστασης στις χρονικές στιγμές $t-1$, t και $t+1$ αντίστοιχα [111].



Σχήμα 3.10: α) Απλή αναπαράσταση RNN, β) RNN «ξεδίπλωμένο» στο χρόνο

Είναι σημαντικό να επισημανθεί πως το «ξεδιπλωμένο» RNN δεν αλλάζει κατά τη διάρκεια των διαδοχικών χρονικών σημείων. Αυτό που μεταβάλλεται είναι η έξοδος και το περιεχόμενο της κρυφής κατάστασης. Υπ' αυτή την έννοια το όλο δίκτυο, τοπολογία και βάρη, λογίζεται ως πιστό αντίγραφο του εαυτού του για κάθε χρονικό σημείο της ακολουθίας εισόδου. Μπορούμε συνεπώς να θεωρήσουμε τις εξόδους των κρυφών μονάδων στα διαφορετικά χρονικά σημεία ως εξόδους διαφορετικών νευρώνων ενός FNN [110].

Η ιδέα του «ξεδιπλωμένου» RNN έχει σημαντικότερο ρόλο στην αντίστροφη διαδρομή, εφαρμόζοντας δηλαδή τη μέθοδο backpropagation για την εκπαίδευση του δικτύου,

ενημερώνοντας τις τιμές των επιμέρους βαρών ώστε να μειωθεί η τιμή της συνάρτησης σφάλματος. Στην περίπτωση των RNNs το backpropagation περιλαμβάνει και τον χρόνο και έτσι ο αλγόριθμος που εφαρμόζεται ονομάζεται Backpropagation Through Time, (BPTT). Ο μηχανισμός του BPTT μπορεί να ερμηνευθεί υιοθετώντας ξανά την ιδέα του ξεδιπλώματος του RNN: Κάθε χρονικό στάδιο έχει τη δική του είσοδο, ένα αντίγραφο της τοπολογίας και των βαρών του δικτύου και μία έξοδο. Στη συνέχεια το σφάλμα υπολογίζεται και συσσωρεύεται σε κάθε χρονικό στάδιο και τέλος, το δίκτυο «διπλώνει» ξανά και ανανεώνονται τα βάρη βάσει του συνολικού συσσωρευμένου σφάλματος.

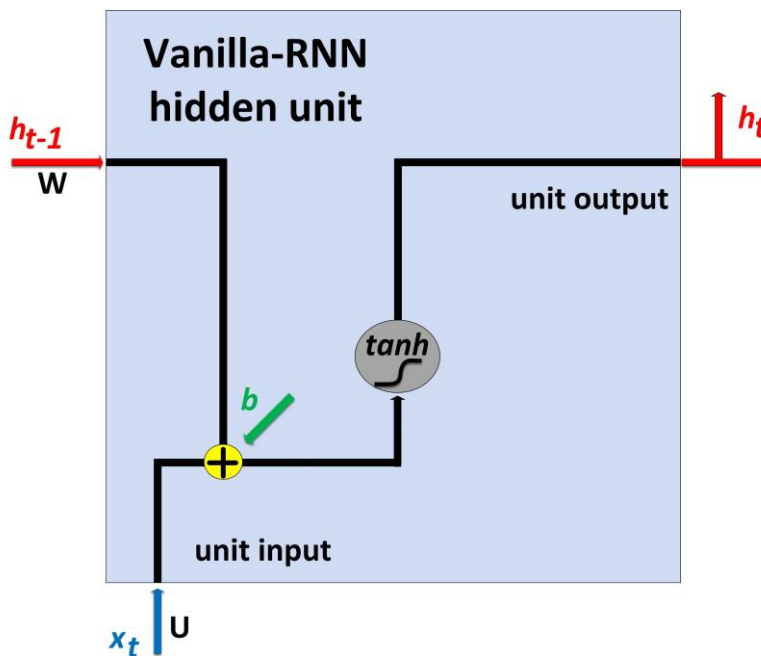
Το BPTT γίνεται ιδιαίτερα απαιτητικό υπολογιστικά όσο ο αριθμός των χρονικών σημείων αυξάνεται. Στην περίπτωση που οι χρονικές ακολουθίες αποτελούνται από χιλιάδες χρονικά σημεία, αυτός είναι και ο αριθμός των μερικών παραγώγων που πρέπει να υπολογιστούν για την ανανέωση μιας και μόνο τιμή βάρους, γεγονός που καθιστά αδύνατη τη χρήση μεγάλου αριθμού επαναλήψεων. Επιπλέον μπορεί να προκαλέσει την εξαφάνιση ή μηδενισμό (vanish) ή την έκρηξη ή απειρισμό (explode) των μερικών παραγώγων. Το πρόβλημα αυτό είναι γνωστό ως Vanishing Gradient Problem και κάνει την μέθοδο backpropagation προβληματική, όχι μονάχα εξαιτίας της μεγάλης διαφοροποίησης στην τάξη μεγέθους των παραγώγων αλλά κυρίως επειδή υποβαθμίζει εκθετικά τις εξαρτήσεις που εκτείνονται σε μεγάλο χρόνο σε σχέση με τις κοντινές. Το γεγονός αυτό οδηγεί στη θορυβώδη και ασταθή εκπαίδευση του μοντέλου ή και συνολικά την αδυναμία σύγκλισης της συνάρτησης σφάλματος δηλαδή την αδυναμία εκπαίδευσής του. Ο πιο απλός τρόπος να ξεπεραστεί αυτό το πρόβλημα, ο οποίος ωστόσο δεν εγγυάται την επίλυσή του είναι η επιλογή πεπερασμένου αριθμού χρονικών βημάτων για την εφαρμογή του BPTT. Μια παραλλαγή του είναι η μέθοδος Truncated BPTT, στην οποία ο μηχανισμός του BPTT ελέγχεται από τον αριθμό των βημάτων της εμπρόσθιας κίνησης, k_1 , και τον αριθμό των βημάτων που εφαρμόζεται η BPTT, k_2 . Με κατάλληλη επιλογή των k_1 και k_2 , είναι δυνατή η αποφυγή του μηδενισμού ή απειρισμού των βαρών [112].

3.6.1 Vanilla – RNN (VRNN)

Το Vanilla-RNN (VRNN) αποτελεί την πιο απλή περίπτωση επαναλαμβανόμενων νευρωνικών δικτύων. Η αρχιτεκτονική του ακολουθεί τη γενική μορφολογία του σχήματος 3.10. Αποτελείται από μία ή περισσότερες κρυφές μονάδες (hidden units) που στη βιβλιογραφία πολλές φορές αναφέρεται ως κύτταρο (cell) η δομή του οποίου απεικονίζεται στο σχήμα 3.11. Η εξίσωση που περιγράφει το VRNN είναι η εξής :

$$h_t = \sigma (W h_{t-1} + U x_t + b) \quad (3.7)$$

όπου x_t , h_t , h_{t-1} , η είσοδος, η έξοδος της κρυφής μονάδας του τωρινού και του προηγούμενου χρονικού βήματος αντίστοιχα, W και U οι πίνακες των βαρών των συγκεκριμένων συνδέσεων ενώ το b αναφέρεται στο διάνυσμα του bias. Η συνάρτηση ενεργοποίησης σ στη γενικότερη περίπτωση, όπως αυτή του σχήματος 3.11, είναι η υπερβολική εφαπτομένη (hyperbolic tangent, \tanh). Ο τελεστής $+$ του σχήματος υποδηλώνει τον πολλαπλασιασμό ανά στοιχείο μεταξύ των πινάκων.



Σχήμα 3.11: Κρυφή μονάδα (ή κύτταρο) VRNN

Για την εκπαίδευση του VRNN χρησιμοποιείται η μέθοδος του Truncated BPTT με κατάλληλη επιλογή των k_1 και k_2 ή απλά η BPTT με πεπερασμένο αριθμό χρονικών βημάτων, ούτως ώστε να αποφευχθεί ο μηδενισμός ή ο απειρισμός των βαρών.

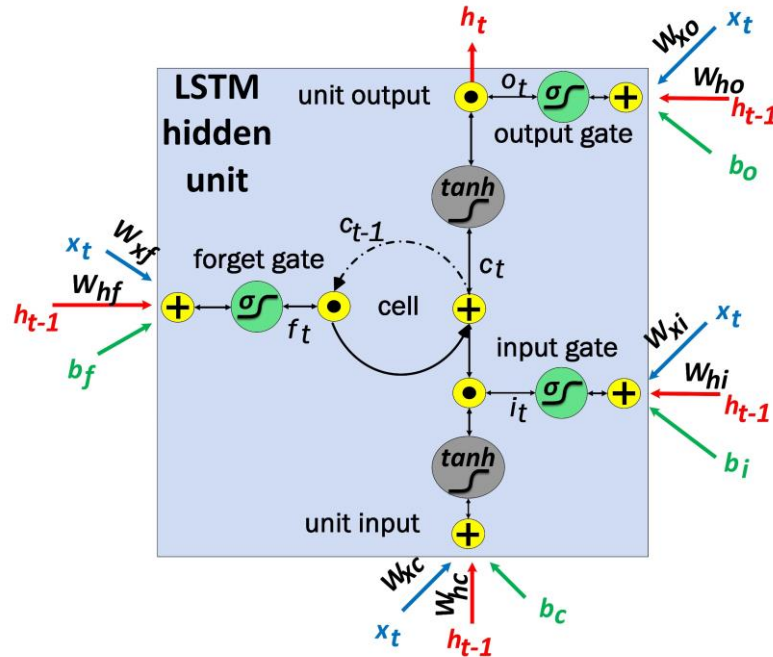
Το πλεονέκτημα του VRNN είναι ότι εκμεταλλεύεται την ιδιότητα της μνήμης του RNN και μπορεί με αυτό τον τρόπο να αντιληφθεί τις χρονικές εξαρτήσεις ακόμη και σε μεγάλης έκτασης χρονικές ακολουθίες. Επίσης αποτελεί την πιο απλή, δομικά αλλά και υπολογιστικά, λύση, γεγονός που την καθιστά θελκτική σε εφαρμογές στις οποίες είναι κρίσιμη η μείωση της απαιτούμενης υπολογιστικής πολυπλοκότητας.

Το σημαντικό μειονέκτημά της είναι πως δεν είναι πάντοτε δυνατή η αποφυγή του μηδενισμού ή του απειρισμού των βαρών κατά τη διάρκεια εφαρμογής του BPTT οπότε και η αστάθεια ή αδυναμία εκπαίδευσής του δικτύου.

3.6.2 Long-Short Term Memory Networks (LSTM)

Διάφορες λύσεις έχουν προταθεί για την αποφυγή του μηδενισμού ή του απειρισμού των βαρών με σημαντικότερη αυτή των δικτύων Long-Short Term Memory Networks (LSTM). Προτάθηκε αρχικά από τους Hochreiter and Schmidhuber το 1997 με σκοπό, όπως υποδηλώνει και το όνομά του, να διευρύνει τη μνήμη των μεγάλων αλλά και μικρών χρονικών ακολουθιών, εντοπίζοντας το χρήσιμο περιεχόμενο τους. Ήταν τόσο καθοριστική η συμβολή της πρότασης αυτής που πλέον αποτελεί την βασικότερη αρχιτεκτονική RNN [113].

Το LSTM ενσωματώνει τρεις μονάδες-πύλες εκτός του κυττάρου: την input gate i_t , output gate o_t και την forget gate f_t με σκοπό να επιτύχει τη διαχείριση της μνήμης των μεγάλων χρονοσειρών. Το κύτταρο (cell) λειτουργεί ως υποστηρικτική μονάδα μνήμης που ελέγχει τη ροή των επιμέρους διεργασιών, καθορίζοντας σε κάθε κύκλο εκτέλεσης μια κατάσταση την cell state (σχήμα 3.12).



Σχήμα 3.12: Κρυφή μονάδα LSTM

Συγκεκριμένα η forget gate f_t καθορίζει ποια πληροφορία από την προηγούμενη κατάσταση του κυττάρου της πρέπει να απορριφθεί και ποια να παραμείνει και σε ποιο βαθμό. Το κατορθώνει αυτό με τη συνδυαστική επεξεργασία της κρυφής εξόδου του προηγούμενου βήματος h_{t-1} με την τρέχουσα είσοδο x_t και το αποτέλεσμα το περνάει διαμέσου μιας σιγμοειδούς συνάρτησης ενεργοποίησης στο cell. Οι τιμές εξόδου της σιγμοειδούς βρίσκονται στο διάστημα $(0,1)$. Αν η τιμή της εξόδου f_t είναι κοντά στο 0 σημαίνει πως δεν είναι χρήσιμη και μπορεί να ξεχαστεί ενώ αντίθετα αν η τιμή της εξόδου είναι κοντά στο 1 σημαίνει πως είναι χρήσιμη και διατηρείται.

Ομοίως, οι πύλες εισόδου και εξόδου μαθαίνουν, αντίστοιχα, ποιες τιμές του cell state πρέπει να ενημερωθούν και ποιες πληροφορίες στην τρέχουσα cell state πρέπει να διατηρηθούν ως έξοδοι της κρυφής κατάστασης h_t .

Συνολικά οι εξισώσεις που περιγράφουν σύμφωνα με τα προηγούμενα το μηχανισμό λειτουργίας και τις επιμέρους εξόδους του LSTM για τη χρονική στιγμή t δίνεται από τις σχέσεις :

$$\begin{aligned}
i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\
f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\
o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \\
c_t &= f_t * c_{t-1} + i_t * \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\
h_t &= o_t * \tanh(c_t)
\end{aligned}$$

(3.8)

Όπου W οι πίνακες των βαρών των επιμέρους συνδέσεων: f , i , o και c που αντιστοιχούν στις πύλες forget, input, output και στο cell state. Τα x_t , h_t , h_{t-1} αποτελούν τα διανύσματα εισόδου, τρέχουσας κρυφής εξόδου και προηγούμενης κρυφής εξόδου αντίστοιχα, ενώ τα b αποτελούν τα διανύσματα bias. Ο τελεστής $*$ υποδηλώνει το γινόμενο στοιχείου με στοιχείο, σ είναι η λογιστική σιγμοειδής (logistic sigmoid) και \tanh η υπερβολική εφαπτομένη (hyperbolic tangent) συνάρτηση μεταφοράς αντίστοιχα [114].

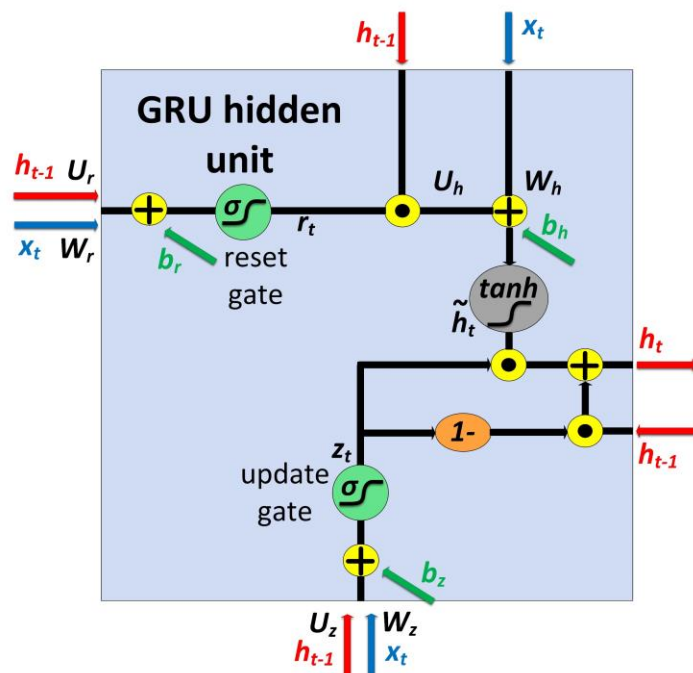
Η αρχιτεκτονική του LSTM του προσδίδει αρκετά πλεονεκτήματα όπως :

- Δυνατότητα να αντιληφθεί χρονικές συσχετίσεις σε πολύ μακριές χρονικές ακολουθίες. Χαρακτηριστικό παράδειγμα αποτελεί η μετάφραση κειμένων, που εξαιτίας της διαφορετικής σύνταξης των λέξεων σε κάθε γλώσσα, ενδεχομένως π.χ. να βρίσκεται το ρήμα σε μεγάλη απόσταση από το υποκείμενο. Ο μηχανισμός του LSTM κατορθώνει να διατηρήσει τη χρήσιμη μνήμη, να εντοπίσει το χρήσιμο περιεχόμενο ακόμη και αν αυτό είναι απομακρυσμένο χρονικά και να απορρίψει το λιγότερο απαραίτητο. Μπορεί συνεπώς να επιτύχει καλύτερα αποτελέσματα από τις συμβατικές αρχιτεκτονικές RNN τόσο σε μεγάλες όσο και μικρές χρονικές ακολουθίες [114].
- Έχει τη δυνατότητα να ελέγξει τον εγγενή θόρυβο των χρονικών ακολουθιών και τις ακολουθίες συνεχών και επαναλαμβανόμενων τιμών.
- Ο μηχανισμός της εκπαίδευσης είναι ελαστικός και σχετικά ανεξάρτητος από τις τιμές των παραμέτρων όπως ο ρυθμός εκπαίδευσης (learning rate) συνεπώς δεν υπάρχει η ανάγκη εξαντλητικών ρυθμίσεων τους [113].

3.6.3 Gated Recurrent Unit Networks (GRU)

Το δίκτυο GRU αποτελεί μια πρόταση της ίδιας λογικής με το LSTM ως προς το στόχο του εντοπισμού χρήσιμου περιεχομένου σε μεγάλες χρονικές ακολουθίες και ως προς την αρχιτεκτονική η οποία βασίζεται, όπως υποδηλώνει και το όνομά του, στην παρουσία πυλών [115].

Το GRU ενσωματώνει δύο πύλες αντί τριών του LSTM, την update gate z_t και τη reset gate r_t , που έχουν σχεδόν ταυτόσημη λειτουργία με τις input και forget gate του LSTM, με σκοπό να επιτύχει τη διαχείριση της μνήμης των μεγάλων χρονοσειρών.



Σχήμα 3.13: Κρυφή μονάδα GRU

Η κρυφή έξοδος h_t του GRU τη χρονική στιγμή t αποτελεί μια γραμμική υπέρθεση της προηγούμενης κρυφής κατάστασης h_{t-1} και της υποψήφιας κρυφής εξόδου \tilde{h}_t . Η update gate αποφασίζει σε ποιο βαθμός η κρυφή μονάδα GRU θα μεταβάλει τις τιμές του περιεχομένου του. Η διαδικασία του γραμμικού αθροίσματος της προηγούμενης με την τωρινή κατάσταση είναι παρόμοια με το LSTM (σχήμα 3.13).

Όταν η τιμή της reset gate r_t είναι κοντά στο μηδέν, η κρυφή μονάδα του GRU συμπεριφέρεται σα να διαβάζει το πρώτο σύμβολο μιας χρονοσειράς επιτρέποντας, μέσω της update gate, να «ξεχάσει» τις προηγούμενες καταστάσεις [116].

Συνολικά οι εξισώσεις που περιγράφουν σύμφωνα με τα προηγούμενα το μηχανισμό λειτουργίας και τις επιμέρους εξόδους του GRU για τη χρονική στιγμή t δίνεται από τις σχέσεις :

$$\begin{aligned} z_t &= \sigma(W_z x_t + U_z h_{t-1} + b_z) \\ r_t &= \sigma(W_r x_t + U_r h_{t-1} + b_r) \\ \tilde{h} &= \tanh[W_h x_t + U_h(r_t * h_{t-1}) + b_h] \\ h_t &= (1 - z_t) * h_{t-1} + z_t * \tilde{h} \end{aligned} \tag{3.9}$$

όπου W και U οι πίνακες των βαρών των επιμέρους συνδέσεων: z και r που αντιστοιχούν στις πύλες update και reset. Τα x_t , h_t , h_{t-1} , \tilde{h} αποτελούν τα διανύσματα εισόδου, τρέχουσας κρυφής εξόδου, προηγούμενης κρυφής εξόδου και υποψήφιας κρυφής εξόδου αντίστοιχα, ενώ τα b αποτελούν τα διανύσματα bias.

Όπως φαίνεται από τα σχήματα 3.12 και 3.13 αλλά και τις εξισώσεις 3.8 και 3.9 ο μηχανισμός του LSTM και του GRU είναι πολύ κοντινός. Σε αντίθεση με τα απλά RNN όπως το VRNN, διαθέτουν και οι δύο αρχιτεκτονικές στοιχεία που διαμορφώνουν το περιεχόμενο τους συνυπολογίζοντας την ενημέρωση που προέκυψε στη μετάβαση μεταξύ της χρονικής στιγμής t και $t+1$. Τα απλά RNN πάντοτε αντικαθιστούν το περιεχόμενο της κρυφής μονάδας με την τιμή που προκύπτει από την τωρινή και την προηγούμενη κρυφή κατάσταση (hidden state). Αντιθέτως, τα LSTM και GRU κρατούν το υπάρχον περιεχόμενο και προσθέτουν πάνω του την όποια πιθανή ενημέρωση προκύψει. Αυτή η επιπρόσθετη λειτουργία επιτρέπει σε κάθε μονάδα να θυμάται το χρήσιμο περιεχόμενο, το οποίο έχει αναδειχθεί από τη forget gate του LSTM ή την reset gate του GRU και να το διατηρεί ακόμη και για αρκετά απομακρυσμένα χρονικά βήματα.

Εξίσου σημαντικό είναι πως ο παραπάνω μηχανισμός δημιουργεί διαδρομές συντομεύσεων που παρακάμπτουν πολλαπλά χρονικά βήματα. Αυτές οι συντομεύσεις

επιτρέπουν στο σφάλμα να επαναδιαδίδεται εύκολα χωρίς να εξαφανίζεται (vanish) πολύ γρήγορα (αν η μονάδα πύλης είναι σχεδόν κορεσμένη στο 1) μειώνοντας με αυτό τον τρόπο το πρόβλημα Vanishing Gradient [117].

Ωστόσο, αυτές οι δύο μονάδες έχουν επίσης αρκετές διαφορές. Ένα χαρακτηριστικό της μονάδας LSTM που λείπει από το GRU είναι η ελεγχόμενη έκθεση του περιεχομένου της μνήμης. Στη μονάδα LSTM, η ποσότητα του περιεχομένου μνήμης που εμφανίζεται ή χρησιμοποιείται από άλλες μονάδες στο δίκτυο ελέγχεται από την πύλη εξόδου. Από την άλλη πλευρά, η GRU εκθέτει το πλήρες περιεχόμενό της χωρίς κανέναν έλεγχο.

Μια άλλη διαφορά είναι στη θέση της input gate ή στην αντίστοιχη reset gate. Η μονάδα LSTM υπολογίζει το νέο περιεχόμενο μνήμης χωρίς χωριστό έλεγχο του όγκου των πληροφοριών που ρέουν από το προηγούμενο χρονικό βήμα. Ελέγχει το νέο περιεχόμενο μνήμης που προστίθεται στο cell ανεξάρτητα από τη forget gate. Από την άλλη πλευρά, το GRU ελέγχει τη ροή πληροφοριών από την προηγούμενη hidden state κατά τον υπολογισμό της νέας, υποψήφιας hidden state, αλλά δεν ελέγχει ανεξάρτητα τη τιμή αυτής. Ο έλεγχος πραγματοποιείται από την update gate [118].

Διάφορες συγκριτικές μελέτες μεταξύ των τριών σχημάτων έχουν δείξει το εμφανές πλεονέκτημα των gating αρχιτεκτονικών έναντι των απλών RNN σε απαιτητικά προβλήματα όπως εκμάθηση και τεχνητή δημιουργία μουσικών κομματιών, αναγνώριση κειμένου, μετάφραση κ.α. Κοινός τόπος αυτών των προβλημάτων είναι, όπως αναφέρθηκε στα προηγούμενα, η ύπαρξη σημαντικών πληροφοριών σε μεγάλες χρονοσειρές που ο μηχανισμός των LSTM και GRU μπορεί να ανιχνεύσει, όχι όμως και το απλό RNN. Εντούτοις μεταξύ των LSTM και GRU δεν έχει αναφερθεί σαφές πλεονέκτημα στις επιδόσεις [116].

Τέλος αξίζει να αναφερθεί πως υπάρχει σημαντική διαφορά στην υπολογιστική πολυπλοκότητα μεταξύ των τριών σχημάτων, με το LSTM να θεωρείται το πιο σύνθετο και το VRNN το πιο ελαφρύ. Εκτενής ανάλυση της υπολογιστικής πολυπλοκότητας θα παρουσιαστεί στα επόμενα καθώς αποτελεί σημαντικό κομμάτι της παρούσας διατριβής.

3.6.4 Αρχιτεκτονικές βασισμένες σε RNN

Οι βασικές δομές RNN που αναφέρθηκαν στα προηγούμενα αποτελούν τα βασικά δομικά στοιχεία ολοκληρωμένων νευρωνικών δικτύων. Οποιοδήποτε από τα VRNN, LSTM και GRU μπορεί να χρησιμοποιηθεί στις αρχιτεκτονικές που θα αναφερθούν στα επόμενα, οπότε επιλέγουμε το γενικό τίτλο RNN.

Βασική αρχιτεκτονική RNN

Η απλούστερη διάταξη αναφέρεται ως βασική αρχιτεκτονική RNN, αποτελείται από το στρώμα (layer) εισόδου, το RNN η έξοδος του οποίου οδηγείται σε ένα πλήρως συνδεδεμένο στρώμα (Fully Connected Layer, FCL) και τέλος το στρώμα εξόδου όπως απεικονίζεται στο σχήμα 3.14.



Σχήμα 3.14: Βασική αρχιτεκτονική RNN μοντέλου

Η είσοδος αποτελείται από χρονικές ακολουθίες ή γενικότερα σειρές δεδομένων των οποίων οι τιμές έχουν μια μεταξύ τους συσχέτιση την οποία το δίκτυο καλείται να αναγνωρίσει. Το RNN αποτελείται από μία ή (συνήθως) περισσότερες κρυφές μονάδων, οι οποίες όπως είδαμε επεξεργάζονται παράλληλα τα (ίδια) δεδομένα εισόδου. Η έξοδος του οδηγείται σε ένα FCL του οποίου ο αριθμός των νευρώνων είναι ίδιος με τις πιθανές τιμές των κατηγοριών στην περίπτωση των προβλημάτων ταξινόμησης (classification) ή της διάστασης του πεδίου τιμών στην περίπτωση των προβλημάτων παλινδρόμησης (regression). Στην πρώτη περίπτωση η έξοδος του FCL οδηγείται στην έξοδο του συνολικού δικτύου μέσω της Softmax συνάρτησης ενεργοποίησης για τον υπολογισμό της κατηγορίας με την μεγαλύτερη πιθανότητα ενώ στη δεύτερη περίπτωση οδηγείται η έξοδος του FCL αποτελεί και την έξοδο του συνολικού μοντέλου. Αξίζει να αναφέρουμε

πως το στρώμα FCL διαθέτει τα δικά του βάρη και τιμές bias. Κατά τη διάρκεια της εκπαίδευσης εκτελείται ο μηχανισμός backpropagation τόσο για το RNN όσο και για το FCL. Στην περίπτωση των προβλημάτων ταξινόμησης χρησιμοποιείται η συνάρτηση απώλειας categorical crossentropy ή η binary crossentropy αν η ταξινόμηση αποτελείται από δύο μόνο κατηγορίες ενώ στα προβλήματα παλινδρόμησης χρησιμοποιείται η Mean Square Error (MSE).

Συστοιχία RNN

Υιοθετώντας την ιδέα των βαθιών νευρωνικών δικτύων σύμφωνα με την οποία επάλληλα κρυφά στρώματα μπορούν να προστεθούν σε ένα πολυστρωματικό νευρωνικό δίκτυο δημιουργώντας νέες αναπαραστάσεις σε υψηλότερα επίπεδα αφαίρεσης, είναι δυνατή η



Σχήμα 3.15: Αρχιτεκτονική μοντέλου συστοιχίας RNN

προσθήκη στη βασική αρχιτεκτονική RNN επιπλέον δομών RNN σε συστοιχία, όπως απεικονίζεται στο σχήμα 3.15. Το καθένα RNN διαθέτει το δικό του αυτόνομο αριθμό κρυφών μονάδων. Είναι σημαντικό να παρατηρήσουμε πως κάθε RNN δέχεται σειριακά τα δεδομένα εισόδου και τα επεξεργάζεται στις δικές του κρυφές μονάδες, όπως και πριν, παράλληλα.

Κατ' αναλογία των FNN στα οποία η εμβάθυνση πραγματοποιείται με την προσθήκη νέων στρωμάτων, τα μοντέλα συστοιχίας RNN έχει παρατηρηθεί πως έχουν τη δυνατότητα να επεξεργαστούν σύνθετα ακολουθιακά προβλήματα [119]. Σαφώς οι επιπλέον δυνατότητες αναπαραστάσης οδηγούν σε σημαντικά αυξημένη υπολογιστική πολυπλοκότητα. Συνεπώς θα πρέπει να σταθμιστεί η διάσταση του προβλήματος και το κατά πόσον η προσθήκη σε σειρά μονάδων RNN υπερτερεί της ενίσχυσης της μίας με επιπλέον κρυφές μονάδες.

Κατά τα άλλα ισχύουν όλα τα χαρακτηριστικά που αναφέρθηκαν στην βασική RNN αρχιτεκτονική.

Αρχιτεκτονική CNN - RNN



Σχήμα 3.16: Αρχιτεκτονική βασικού μοντέλου CNN-RNN

Συνδυάζοντας τα ιδιαίτερα χαρακτηριστικά των συνελκτικών νευρωνικών δικτύων (CNN), όπως αυτά παρουσιάστηκαν στα προηγούμενα και των επαναλαμβανόμενων νευρωνικών δικτύων μπορούμε να σχεδιάσουμε υβριδικές αρχιτεκτονικές. Τα μοντέλα αυτού του τύπου είναι τόσο χωρικά όσο και χρονικά βαθιά και έχουν την ευελιξία να εφαρμοστούν σε μια ποικιλία εργασιών μηχανικής όρασης που περιλαμβάνουν διαδοχικές εισόδους και εξόδους, παρουσιάζοντας καλές επιδόσεις σε σύνθετα προβλήματα οπτικών χρονοσειρών, όπως video κά. [120]

Στη βασική του μορφή το CNN-RNN μοντέλο αποτελείται από το στρώμα εισόδου, ένα στρώμα CNN, ένα στρώμα RNN και το στρώμα της εξόδου, όπως φαίνεται στο σχήμα 3.16. Με αυτό τον τρόπο η κάθε δομή αναλαμβάνει να επεξεργαστεί το χαρακτηριστικό αυτό που εγγενώς έχει καλύτερες επιδόσεις. Πχ στην περίπτωση του video το CNN αναλαμβάνει την επεξεργασία και εκπαίδευση στη βάση των οπτικών χαρακτηριστικών ενώ το RNN των χρονικών.

Έχει επιτυχώς εφαρμοστεί σε προβλήματα αναγνώρισης ομιλίας και επεξεργασίας γλώσσας όπου τα CNN χρησιμοποιούνται ως εξαγωγείς χαρακτηριστικών και υποβοηθούν τα RNN να επεξεργαστούν αποδοτικότερα τα δεδομένα εισόδου ήχου και κειμένου.

Αρχιτεκτονική Encoder- Decoder

Μια ενδιαφέρουσα αρχιτεκτονική που έχει αποδειχθεί ιδιαίτερα αποτελεσματική αποτελεί αυτή των μοντέλων Encoder- Decoder. Εφαρμόζεται σε προβλήματα στα οποία το μοντέλο καλείται να εκπαιδευτεί έχοντας στη διάθεσή του στην είσοδο μια ακολουθία δεδομένων και παράγει στην έξοδο μια ακολουθία δεδομένων κυρίως διαφορετικού μήκους. Στην απλή περίπτωση του σχήματος 3.17 αποτελείται από δυο διακριτά RNN τα οποία αποτελούν τον κωδικοποιητή (Encoder) και τον αποκωδικοποιητή (Decoder) αντίστοιχα [121].



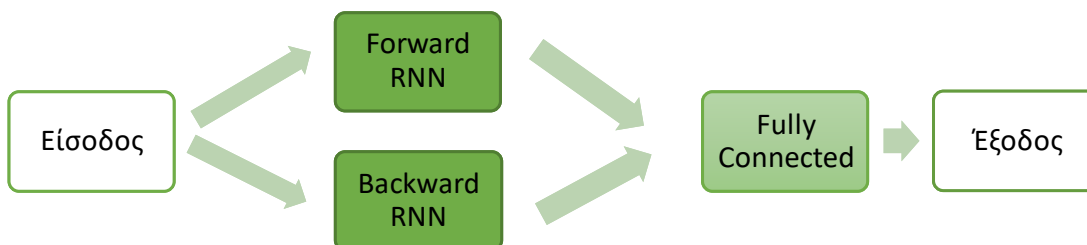
Σχήμα 3.17: Αρχιτεκτονική βασικού μοντέλου Encoder- Decoder

Τα πιο συνηθισμένα προβλήματα στα οποία έχει εφαρμοστεί με πολύ καλές επιδόσεις είναι η μετάφραση κειμένων. Δεδομένης της διαφοράς στη σύνταξη των προτάσεων από γλώσσα σε γλώσσα, η λέξη προς λέξη απόδοση στις περισσότερες περιπτώσεις οδηγεί σε λανθασμένη μετάφραση. Τη λύση δίνει η συγκεκριμένη αρχιτεκτονική σύμφωνα με την οποία, το RNN του Encoder αναλαμβάνει να κωδικοποιήσει την εισερχόμενη πληροφορία (την πρωτότυπη πρόταση) ενώ το RNN του Decoder να την αποκωδικοποιήσει, δηλαδή να την αποδώσει στην επιθυμητή γλώσσα.

Αξίζει να επισημανθεί πως η αρχιτεκτονική του Encoder- Decoder δεν εφαρμόζεται μονάχα με δίκτυα RNN. Αντιθέτως πρόκειται για μια γενικότερη αρχιτεκτονική σύμφωνα με την οποία, καθένας από τους Encoder και Decoder διαθέτει το δικό του αυτόνομο νευρωνικό δίκτυο. Όπως είναι κατανοητό, η ύπαρξη δύο αυτόνομων νευρωνικών δικτύων επιφέρει μια αύξηση στην πολυπλοκότητα που δε συνοδεύεται αυτονόητα με αύξηση της απόδοσης, δεδομένα που πρέπει να σταθμιστούν πριν εφαρμοστεί αυτή η σύνθετη λύση.

Αμφίδρομη (Bi-directional) αρχιτεκτονική

Όπως έχουμε δει το μοντέλο του RNN με τις διάφορες εκδοχές του (VRNN, LSTM και GRU), οφείλει τις καλές του επιδόσεις στο γεγονός ότι χρησιμοποιεί τη χρήσιμη πληροφορία που υπάρχει στα προηγούμενα βήματα. Ως τώρα η παράμετρος του χρόνου αφορούσε αποκλειστικά το παρελθόν, τα προηγούμενα δηλαδή χρονικά στιγμιότυπα διαμόρφωναν την εξέλιξη της τρέχουσας χρονικής κατάστασης. Εντούτοις, σε αρκετά προβλήματα τα οποία επεξεργάζονται χρονικές ακολουθίες, η πληροφορία που υπάρχει στις μελλοντικά γεγονότα, όταν αυτά είναι διαθέσιμα, είναι εξίσου σημαντική και ο συνυπολογισμός της μπορεί να βελτιώσει περαιτέρω τις επιδόσεις του νευρωνικού μοντέλου. Η πιο χαρακτηριστική αρχιτεκτονική που χρησιμοποιεί τη λογική αυτή είναι η λεγόμενη Αμφίδρομη (Bi-directional), σχήμα 3.18.



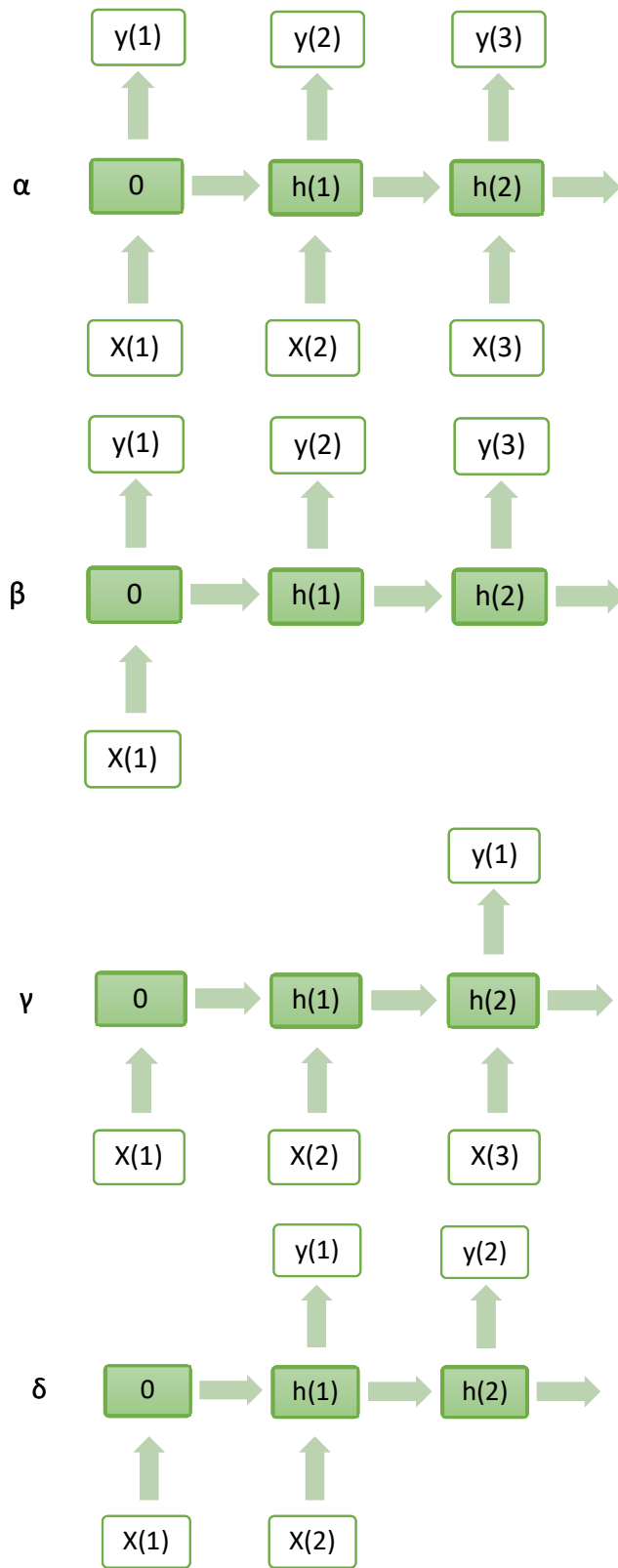
Σχήμα 3.18: Αμφίδρομη αρχιτεκτονική

Τα αμφίδρομα RNN αξιοποιούν στο έπακρο τα χαρακτηριστικά της ακολουθίας εισόδου περνώντας τα βήματα του χρόνου τόσο προς την εμπρόσθια όσο και προς την οπίσθια κατεύθυνση. Στην πράξη, αυτή η αρχιτεκτονική περιλαμβάνει την αντιγραφή του πρώτου RNN, δημιουργώντας πρακτικά το εμπρόσθιο και το οπίσθιο RNN, στη συνέχεια να τροφοδοτεί την ακολουθία εισόδου ως έχει ως είσοδο στο Forward και να παρέχει ένα αντίστροφο αντίγραφο της ακολουθίας εισόδου στο Backward RNN [122]. Η προσέγγιση της αμφίδρομης αρχιτεκτονικής χρησιμοποιήθηκε με μεγάλη επιτυχία με LSTM, επιλύοντας προβλήματα αναγνώρισης ομιλίας [123].

Η διοχέτευση της ακολουθίας στο δίκτυο με την ορθή και την ανάστροφη φορά βασίζεται στην υπόθεση πως ολόκληρη η ακολουθία της εισόδου είναι διαθέσιμη, τόσο για την χρονοσειρά εκπαίδευσης όσο και για την αντίστοιχη της εκτίμησης. Με άλλα λόγια η εκτίμηση ενός τωρινού γεγονότος βασίζεται στη γνώση μελλοντικών συμβάντων, γεγονός που εκ πρώτης όψεως δείχνει να παραβιάζει την αρχή της αιτιότητας. Στο πρόβλημα της αναγνώρισης ομιλίας παρότι φαίνεται παράδοξο η αναγνώριση να γίνεται λαμβάνοντας υπόψη μελλοντικά γεγονότα, εντούτοις στην πραγματικότητα η ανθρώπινος εγκέφαλος με αυτό τον τρόπο αντιλαμβάνεται την ομιλία. Η γραμμική παράθεση λέξεων, δηλαδή η εκτίμηση που βασίζεται μόνο στις προηγούμενες λέξεις δεν είναι επαρκής για την ολοκληρωμένη αναγνώριση μιας φράσης. Η πληροφορία που έπεται, συχνά είναι εξίσου χρήσιμη για την ακριβέστερη κατανόηση της. Το γεγονός αυτό εκμεταλλεύεται η αρχιτεκτονική των αμφίδρομων LSTM δικτύων, επιδεικνύοντας σημαντική αναβάθμιση των επιδόσεων που έκτοτε παγιώθηκε και γενικεύθηκε η χρήση τους στα RNN δίκτυα [123].

3.6.5 Τοπολογίες εκπαίδευσης μοντέλων βασισμένων σε RNN

Η διάταξη των επιμέρους κρυφών μονάδων και το πλήθος των εισόδων που μπορεί να επεξεργαστεί ή των εξόδων που αριθμεί, διαμορφώνει κάποιες αρκετά ενδιαφέρουσες τοπολογίες. Αναλόγως των εκάστοτε ιδιαίτερων χαρακτηριστικών του προβλήματος που καλείται το νευρωνικό δίκτυο να επιλύσει, καθεμία από αυτές επιτυγχάνει διαφορετική απόδοση, συνήθως εκτιμώμενη με την ακρίβεια πρόβλεψης στα προβλήματα ταξινόμησης ή την απόκλιση από την ιδανική τιμή στα προβλήματα παλινδρόμησης, αλλά και διαφορετική υπολογιστική πολυπλοκότητα. Ορίζοντας με X , h και y την ακολουθία εισόδου, την κρυφή κατάσταση και την έξοδο του δικτύου αντίστοιχα, στο σχήμα 3.19 παρουσιάζονται οι διαφορετικές τοπολογίες και ο τρόπος που η καθεμία εισάγει, επεξεργάζεται και εξάγει τα δεδομένα στις διαφορετικές χρονικές στιγμές.



Σχήμα 3.19: Τοπολογία: α) Ένα προς Ένα, β) Ένα προς Πολλά, γ) Πολλά προς ένα, δ) Πολλά προς πολλά

Έτσι, η *ένα προς ένα* τοπολογία (σχ. 3.19α) παράγει μία έξοδο $y(t)$ για κάθε μία είσοδο $X(t)$. Παρά την προφανή απλότητα της, στην πραγματικότητα η τοπολογία αυτή κάνει μηδαμινή χρήση των ιδιοτήτων της μνήμης των RNN καθώς λαμβάνει ως είσοδο ένα μεμονωμένο χρονικό στιγμιότυπο, αδιαφορώντας για τα γειτονικά, συνεπώς χάνει την πληροφορία που βρίσκεται στα χρονικά εξελισσόμενα γεγονότα. Η δυναμική της περιορίζεται στην εσωτερική δυναμική των κρυφών μονάδων. Οι περιορισμένες εμβέλειας χρήση της είναι στις περιπτώσεις που διερευνάται η εκπαίδευση μεταξύ των διάφορων χρονικών σημείων.

Η *ένα προς πολλά* τοπολογία (σχ. 3.19β) παράγει πολλές εξόδους ($y(t), y(t+1), \dots$) για μία μεμονωμένη είσοδο $X(t)$. Μια πρόχειρη εκτίμηση θα έβγαζε το συμπέρασμα πως λειτουργεί παρόμοια με την *ένα προς ένα* τοπολογία, απλά παράγοντας πολλές εξόδους. Εντούτοις σε προβλήματα στα οποία η -μεμονωμένη χρονικά- είσοδος περιέχει πληροφορία που μπορεί να εκφραστεί στην έξοδο με ακολουθιακά δεδομένα η τοπολογία αυτή αποτελεί μια πολύ ισχυρή πρόταση. Χαρακτηριστικό παράδειγμα αποτελεί η δημιουργία περιγραφικών υποτίτλων για μια εικόνα, συνδυάζοντας την *ένα προς πολλά* τοπολογία με το μοντέλο του CNN-RNN, όπου στο CNN ανατίθεται ο ρόλος της εξαγωγής των χαρακτηριστικών της εικόνας και το RNN αναλαμβάνει να εκφράσει τα χαρακτηριστικά αυτά με μία ολοκληρωμένη πρόταση.

Η *πολλά προς ένα* τοπολογία (σχ. 3.19γ) παράγει μία έξοδο $y(t)$ αφού έχει λάβει αρκετές εισόδους ($X(t), X(t+1), \dots$). Πρόκειται στην πράξη για την πλέον αντιπροσωπευτική τοπολογία και την πιο συχνά χρησιμοποιημένη. Λαμβάνοντας υπόψιν πολλές διαδοχικές τιμές εισόδου, οι κρυφές μονάδες του RNN τις επεξεργάζονται κάνοντας χρήση του μηχανισμού που αναλύθηκε στα προηγούμενα, παράγεται μια τελική τιμή εξόδου.

Η *πολλά προς ένα* τοπολογία είναι κατάλληλη για προβλήματα πρόβλεψης ακολουθίας όπου απαιτούνται πολλαπλά χρονικά βήματα εισόδου για να γίνει μια πρόβλεψη ενός βήματος. Χαρακτηριστικά παραδείγματα αποτελούν η πρόβλεψη της επόμενης πραγματικής τιμής σε μια χρονοσειρά με δεδομένη μια ακολουθία εισερχόμενων παρατηρήσεων (πχ πρόβλεψη χρηματιστηριακών δεδομένων), η ταξινόμησης μίας ακολουθίας (πχ αναγνώριση μουσικής τονικότητας) κα. Ο αριθμός των διαδοχικών τιμών

εισόδων που πρέπει να ληφθεί εξαρτάται αποκλειστικά από το εκάστοτε πρόβλημα και καθορίζει την μνήμη του.

Τέλος η *πολλά προς πολλά* τοπολογία (σχ. 3.19δ) παράγει πολλές εξόδους ($y(t)$, $y(t+1)$, ...) λαμβάνοντας αρκετές εισόδους ($X(t)$, $X(t+1)$, ...). Η τοπολογία αυτή συνδυάζει τα χαρακτηριστικά των δύο προηγούμενων. Όπως στην περίπτωση της *πολλά προς ένα*, λαμβάνοντας υπόψιν πολλές διαδοχικές τιμές εισόδου, εξάγονται πολλαπλά χρονικά βήματα. Είναι σημαντικό να αναφερθεί πως ο αριθμός των εξόδων δεν είναι κατ' ανάγκη ίσος με αυτό των εισόδων. Χαρακτηριστικά, η περίπτωση της Encoder-Decoder αρχιτεκτονικής, όπως παρουσιάστηκε προηγουμένως, χρησιμοποιεί σχεδόν αποκλειστικά τη συγκεκριμένη τοπολογία και ο αριθμός των εισόδων συνήθως διαφέρει από αυτόν των εξόδων.

Η *πολλά προς πολλά* τοπολογία είναι κατάλληλη για προβλέψεις ακολουθίας όπου απαιτούνται πολλαπλά χρονικά βήματα εισόδου προκειμένου να προβλεφθεί μια ακολουθία χρονικών βημάτων εξόδου. Η τοπολογία αυτή έχει ιδιαίτερο ενδιαφέρον, αποτελεί πεδίο έντονης έρευνας στο χώρο της μηχανικής μάθησης και θα παρουσιαστεί αναλυτικά στην παρούσα διατριβή καθώς η υιοθέτησή της της επιτρέπει τη βελτιστοποίηση της πολυπλοκότητας όπως θα δούμε στα επόμενα κεφάλαια.

4^ο ΚΕΦΑΛΑΙΟ

Ισοστάθμιση Βασισμένη σε Επαναλαμβανόμενα Νευρωνικά Δίκτυα στα Συστήματα Οπτικών Επικοινωνιών Μεγάλων Αποστάσεων

4.1 Εισαγωγή

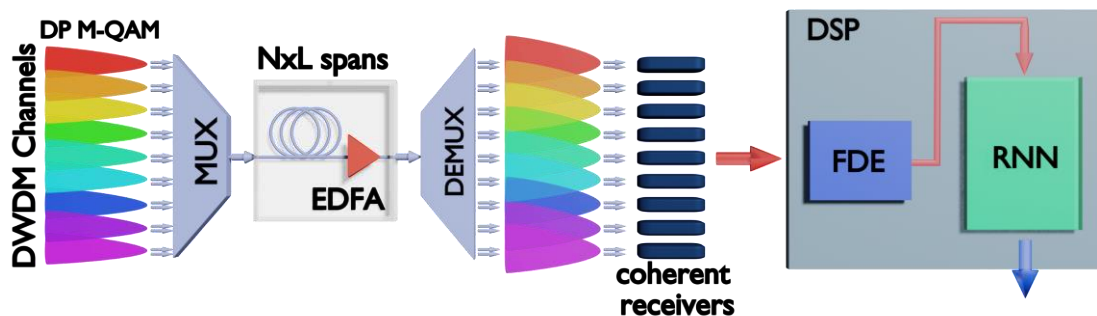
Τα σημαντικά άλματα που έχουν γίνει τα τελευταία χρόνια από την ερευνητική κοινότητα στον χώρο της μηχανικής μάθησης και ιδιαίτερα στην ανάπτυξη των νευρωνικών δικτύων, μας δίνουν τη δυνατότητα αξιοποίησης τους για την άρση των μη γραμμικών φαινομένων που εκδηλώνονται κατά τη διάδοση του σήματος στην οπτική ίνα, όπως παρουσιάστηκαν αναλυτικά στα προηγούμενα κεφάλαια.

Το σημαντικότερο μέρος της διατριβής αποτελεί την πρόταση, ανάλυση και εφαρμογή μοντέλων βασιζόμενων σε επαναλαμβανόμενα νευρωνικά δίκτυα (RNN) για την ισοστάθμιση του σήματος καναλιών που λαμβάνεται στον δέκτη ενός συστήματος οπτικών επικοινωνιών μεγάλων αποστάσεων (long haul), στον οποίο απαιτούνται σύνθετες τεχνικές ψηφιακής επεξεργασίας σήματος για την απαλοιφή των επιδράσεων των φαινομένων διάδοσης.

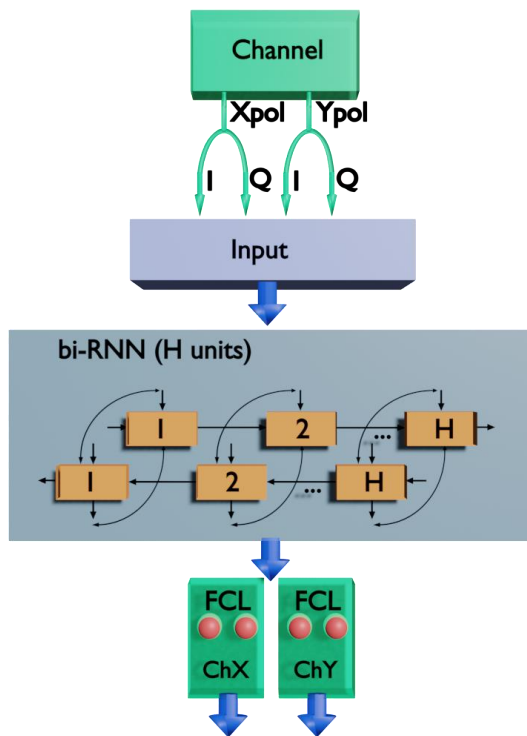
Το σύστημα της οπτικής διάδοσης που χρησιμοποιείται αποτυπώνεται στο σχήμα 4.1α. Συγκεκριμένα αναφερόμαστε σε ένα σύστημα WDM καναλιών πολυπλεξίας πόλωσης διαμορφωμένων με M -QAM, όπου M η τάξη της διαμόρφωσης. Η διάδοση πραγματοποιείται σε N αριθμό τμημάτων οπτικής ίνας μήκους L και ενισχύεται, επίσης περιοδικά, με διάταξη EDFA. Η έξοδος του σύμφωνου δέκτη οδηγείται στο τμήμα της ψηφιακής επεξεργασίας (DSP) στο πρώτο στάδιο του οποίου χρησιμοποιείται ισοστάθμιση τύπου FDE για την ιδανική αντιστάθμιση της χρωματικής διασποράς. Έπειτα

χρησιμοποιείται ο προτεινόμενος bi-RNN ισοσταθμιστής, η λεπτομερής δομή του οποίου αποτυπώνεται στο σχήμα 4.1β.

α



β



Σχήμα 4.1: α) Το σύστημα οπτικής διάδοσης μεγάλων αποστάσεων βασιζόμενο σε σύμφωνο δέκτη, ακολουθούμενο από το τμήμα της ψηφιακής επεξεργασίας (DSP) β) Λεπτομερής αποτύπωση του προτεινόμενου RNN δικτύου.

Στο κεφάλαιο αυτό θα παρουσιαστούν οι πρόσφατες και πλέον σημαντικές τεχνικές εξισορρόπησης που χρησιμοποιούνται στα συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων και ιδιαίτερα αυτές που βασίζονται σε αλγορίθμους νευρωνικών δικτύων. Θα παρουσιαστεί αναλυτικά η πρόταση της χρήσης RNN και συγκεκριμένα LSTM, ως αποδοτικότερης τεχνικής αντιστάθμισης της καθιερωμένης στις οπτικές επικοινωνίες, DBP. Αξίζει να αναφέρουμε πως η χρήση LSTM για την αντιστάθμιση των μη γραμμικών φαινομένων σε σύμφωνα ψηφιακά συστήματα μεγάλων αποστάσεων προτάθηκε στην ερευνητική κοινότητα για πρώτη φορά στην εργασία της αναφοράς [1].

Στη συνέχεια θα παρουσιαστούν οι προεκτάσεις και συγκρίσεις μεταξύ διαφορετικών μοντέλων RNN ως προς τις επιδόσεις και την πολυπλοκότητα [6], οι ενδελεχείς συγκρίσεις με την καθιερωμένη τεχνική αντιστάθμισης Volterra και λεπτομερής αποτύπωση της συμπεριφοράς σε περιβάλλον διάδοσης με διαφορετικές τιμές οπτικής διασποράς και προτάσεις για την αποδοτικότερη εκμετάλλευση της αρχιτεκτονικής των RNN για τη μείωση της υπολογιστικής πολυπλοκότητας [2].

Τέλος θα παρουσιαστεί η πρόταση της πολυκαναλικής εξισορρόπησης των μη γραμμικών φαινομένων (multi-channel equalization) σε WDM κανάλια οπτικών επικοινωνιών δείχνοντας την περαιτέρω βελτίωση των επιδόσεων και ταυτόχρονα τη μείωση της υπολογιστικής πολυπλοκότητας [5].

4.2 Αντιστάθμιση μη γραμμικότητας οπτικών ινών σε ψηφιακά σύμφωνα συστήματα που αξιοποιούν νευρωνικά δίκτυα (LSTM)

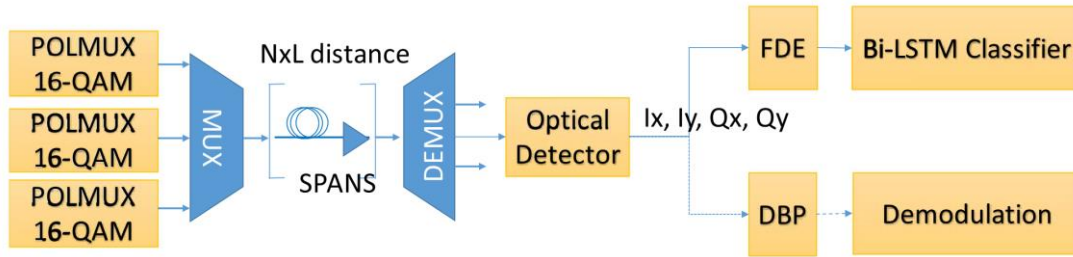
Τα τελευταία χρόνια υπάρχει μια διαρκής τάση υιοθέτησης τεχνικών μηχανικής μάθησης είτε για τη μείωση των φαινομένων υποβάθμισης της ποιότητας της διάδοσης [124] είτε την εκτίμηση της ποιότητας της διάδοσης στα σύγχρονα συστήματα οπτικών επικοινωνιών [125]. Διάφορες εκδοχές βασιζόμενες σε τεχνητά νευρωνικά δίκτυα (ANNs) [126], συνελικτικά νευρωνικά δίκτυα (CNNs) [127] και επαναλαμβανόμενα νευρωνικά δίκτυα [128] έχουν εφαρμοστεί κυρίως σε σχήματα διαμόρφωσης έντασης – άμεσης φώρασης (IM/DD) και πολυπλεξίας ορθογώνιας διαίρεσης συχνότητας (OFDM) [124].

Πρόσφατα, προτάθηκε η χρήση αμφίδρομων LSTM για την μείωση των φαινομένων υποβάθμισης σε συστήματα βασιζόμενα σε PAM-4 διαμόρφωση [51], [45], δείχνοντας με τον τρόπο αυτό πως τα LSTM μοντέλα μπορούν να αποτελέσουν μια σημαντική εναλλακτική πρόταση εξομάλυνσης σε IM/DD συστήματα.

Τα σημαντικότερα πλεονεκτήματα των bidirectional LSTM – και εν γένει των bi-RNN- είναι η δυνατότητά τους να διαχειριστούν την αλληλοπαρεμβολή συμβόλων (ISI) μεταξύ των προηγούμενων και των επόμενων συμβόλων που προκαλούνται από χρωματική διασπορά.

Στην παράγραφο αυτή εξετάζουμε την αποτελεσματικότητα του bi-LTSM στην αντιστάθμιση της μη γραμμικότητας των ινών σε ψηφιακά σύμφωνα συστήματα οπτικών επικοινωνιών. Αν και η διασπορά αντισταθμίζεται πολύ αποτελεσματικά με τη χρήση γραμμικής επεξεργασίας σήματος στον δέκτη, η αλληλεπίδρασή της με τη μη γραμμικότητα κατά μήκος της διάδοσης καθιστά την οπτική ίνα ένα μη γραμμικό κανάλι με μνήμη. Το μέγεθος της μνήμης αυξάνεται με την ποσότητα της συσσωρευμένης διασποράς από άκρο σε άκρο. Προκειμένου να μελετήσουμε την αποτελεσματικότητα στη μείωση της μη γραμμικότητας του bi-LSTM ισοσταθμιστή, πραγματοποιούμε αριθμητικές προσομοιώσεις στα 1550 nm (ζώνη C) και 1310 nm (ζώνη O) λαμβάνοντας υπόψη τόσο μονοκάναλη όσο και πολυκάναλη διάδοση. Ως τεχνική αναφοράς για τη σύγκριση χρησιμοποιούμε την τεχνική DBP. Αποδεικνύεται ότι το bi-LSTM ξεπερνά το DBP ιδιαίτερα σε WDM διάδοση καθώς μπορεί να αναγνωρίσει και να εξάγει τις ιδιότητες του μη γραμμικού καναλιού, οι οποίες αποτυπώνονται ως αλλοιώσεις στο διαδιδόμενο οπτικό σήμα. Η ανάλυση περιλαμβάνει την εκτίμηση του ρυθμού του σφάλματος ως συνάρτηση του αριθμού των κρυφών μονάδων και του μήκους της συμβολοσειράς που χρησιμοποιείται για την εκπαίδευση του νευρωνικού δικτύου. Η ανάλυση περιέχει επίσης εκτίμηση της πολυπλοκότητας σε σύγκριση με αυτή της DBP τεχνικής.

4.2.1 Ανάλυση Μοντέλου Οπτικής Διάδοσης



Σχήμα 4.2: Το προσομοιωμένο σύστημα οπτικής διάδοσης

Το σύστημα διάδοσης που απεικονίζεται στο σχήμα 4.2 προσομοιώνεται αριθμητικά με τη βοήθεια της μη γραμμικής εξίσωσης του Schrödinger (NLSE) η οποία επιλύεται με τη μέθοδο της Split Step Fourier. Η απόφαση μας να μελετήσουμε την οπτική διάδοση στις δύο διαφορετικές ζώνες, ζώνη C και ζώνη O, βασίζεται στο γεγονός πως η μεν ζώνη C παραμένει η κυρίαρχη για τα συστήματα οπτικών επικοινωνιών, ενώ η ζώνη O αναμένεται στο εγγύς μέλλον να αποτελέσει μια ελπιδοφόρα λύση στην επέκταση του εύρους ζώνης όταν οι τρέχουσες ζώνες C και L θα έχουν κορεστεί.

Οι δύο αυτές ζώνες έχουν σημαντικές διαφορές: Η ζώνη C παρουσιάζει τη χαμηλότερη εξασθένιση και τη μεγαλύτερη διασπορά, 0.2 dB/km and 17 ps/nm/km αντίστοιχα, ενώ αντιθέτως η ζώνη O παρουσιάζει σχεδόν μηδενική διασπορά συνοδευόμενη από σημαντικά μεγαλύτερες απώλειες, περίπου 0.34 dB/km, καθιστώντας την πιο ευάλωτη στις μη γραμμικότητες και τον θόρυβο.

Για τους παραπάνω λόγους θεωρήσαμε σημαντικό να διερευνήσουμε το κατά πόσο η προτεινόμενη μας τεχνική μη γραμμικής εξισορρόπησης μπορεί να είναι αποτελεσματική στα δύο αυτά διαφορετικά περιβάλλοντα.

Σύμβολο	Παράμετρος	Τιμή
G	Απολαβή του ενισχυτή	10dB @ 1550nm, 17dB @1310nm

a	εξασθένιση	0.2dB/km @ 1550nm, 0.34dB/km @ 1310nm
β_2	Χρωματική Διασπορά 2 ^{ης} τάξης	-21,5 ps ² /km @ 1550 nm, -0.82 ps ² /km @ 1310 nm
γ	Συντελεστής μη γραμμικότητας	1.3 W ⁻¹ km ⁻¹ και στα δύο μήκη κύματος
R	Ρυθμός συμβόλων	25 Gbaud/κανάλι
M	Τύπος διαμόρφωσης	Dual-polarization 16-QAM
L	Απόσταση ενισχυτών	50 km
Δf	Απόσταση καναλιών	50 GHz

Πίνακας 4.1 : Παράμετροι αριθμητικού μοντέλου

Θεωρούμε μονοκάναλη και πολυκάναλη διάδοση. Κάθε κανάλι είναι ένα σήμα διπλής πόλωσης 16-QAM που λειτουργεί στον ρυθμό των 25 Gbaud. Το σήμα ενισχύεται μετά από κάθε διάδοση των 50 km (50km απόσταση των ενισχυτών ή αλλιώς span) ενώ ο ενισχυτής έχει εικόνα θορύβου ίση με 5 dB. Όλες οι παράμετροι συνοψίζονται στον Πίνακα 4.1. Ο οπτικός δέκτης αποτελείται από ένα οπτικό υβρίδιο (optical hybrid), ισορροπημένους φωτοανιχνευτές, ηλεκτρικά χαμηλοπερατά φίλτρα με συχνότητα αποκοπής προσαρμοσμένη στον ρυθμό διάδοσης.

Όπως φαίνεται και στο σχήμα 4.2α, εξετάζουμε δύο πιθανές μεθόδους επεξεργασίας, μία βασισμένη στο bi-LSTM και μία στο DBP. Όταν χρησιμοποιείται bi-LSTM, πρώτα δειγματοληπτούμε το σήμα και εκτελούμε την ιδανική αντιστάθμιση χρωματικής διασποράς με τη χρήση ενός ιδανικού (FDE) ισοσταθμιστή. Όταν χρησιμοποιείται το DBP, δειγματοληπτούμε το σήμα και το μεταφέρουμε στη DBP μονάδα. Σε αυτή την περίπτωση δεν υπάρχει ανάγκη για επιπλέον επεξεργασία για την αντιστάθμιση της χρωματικής διασποράς καθώς το DBP αναλαμβάνει να επεξεργαστεί συνολικά το ψηφιακό σήμα. Υποθέσαμε την ιδανική εκτίμηση φάσης φέρουσας (CPE) και συχνότητας (δεν λήφθηκε υπόψη ο θόρυβος φάσης του laser) καθώς και ιδανική αποπολυπλεξία πόλωσης και στις

δύο περιπτώσεις, καθώς θέλουμε να εστιάσουμε αποκλειστικά στα μη γραμμικά φαινόμενα. Η διάδοση του σήματος στο μοντέλο μας περιγράφεται από τις εξισώσεις Manakov [67].

$$\frac{\partial E_{x,y}}{\partial z} = -\frac{\alpha}{2} E_x + \frac{j\beta_2}{2} \frac{\partial^2 E_{x,y}}{\partial t^2} - j\gamma \frac{8}{9} (|E_x|^2 + |E_y|^2) E_{x,y} \quad (4.1)$$

όπου E_x , E_y , οι δύο συνιστώσες του ηλεκτρικού πεδίου E που αντιστοιχούν στις κάθετες πολώσεις x και y . Η προσομοίωση της NLSE πραγματοποιείται χρησιμοποιώντας 64 δείγματα για κάθε σύμβολο. Στο δέκτη, η FDE απαιτεί δύο (2) δείγματα ανά σύμβολο ούτως ώστε να παράσχει σχεδόν ιδανική αντιστάθμιση της διασποράς. Η μέθοδος DBP λειτουργεί ικανοποιητικά όταν χρησιμοποιεί 4 δείγματα ανά σύμβολο. Από την άλλη μεριά για την επαρκή εκπαίδευση του LSTM απαιτείται μόνο ένα δείγμα ανά σύμβολο, δεδομένου όμως πως έχει προηγηθεί η αντιστάθμιση της διασποράς από την FDE μονάδα. Η DBP εφαρμόζεται σε ένα σήμα, αυτό του καναλιού ενδιαφέροντος, συνεπώς στη συγκεκριμένη μελέτη το DPB μπορεί να χειρίζεται μόνο ενδοκαναλικά φαινόμενα. Το πολυκαναλικό DBP είναι ικανό να αντιμετωπίζει και διακαναλικά φαινόμενα, ωστόσο μια τέτοια προσέγγιση ενισχύει σημαντικά την πολυπλοκότητα του DSP στον δέκτη, επομένως δεν μελετήθηκε καθόλου [129]. Ομοίως, το bi-LSTM εκπαιδεύεται και δοκιμάζεται αποκλειστικά με βάση τις πληροφορίες που μεταφέρει το κανάλι ενδιαφέροντος που ανιχνεύεται. Όταν μελετάται η πολυκαναλική διάδοση, αξιολογούμε την απόδοση του κεντρικού WDM καναλιού, θεωρώντας ότι είναι αυτό που επηρεάζεται περισσότερο από τα ενδοκαναλικά φαινόμενα.

4.2.1 Βασικά Χαρακτηριστικά του Μοντέλου Bi- LSTM

Το LSTM ως ο πλέον αντιπροσωπευτικός τύπος RNN αναλύθηκε διεξοδικά στο κεφάλαιο 3. Όταν αναφερόμαστε στο LSTM, θεωρούμε πως κάθε μπλοκ του αποτελείται από παράλληλα συνδεδεμένες κρυφές μονάδες/ κύτταρα της μορφής του σχήματος 3.11 των οποίων οι επιμέρους έξοδοι εξισώσεις δίνεται από τις σχέσεις :

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i)$$

$$\begin{aligned}
f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\
o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \\
c_t &= f_t * c_{t-1} + i_t * \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\
h_t &= o_t * \tanh(c_t)
\end{aligned}
\tag{4.2}$$

Όπου W οι πίνακες των βαρών των επιμέρους συνδέσεων: f , i , o και c που αντιστοιχούν στις πύλες forget, input, output και στο cell state. Τα x_t , h_t , h_{t-1} αποτελούν τα διανύσματα εισόδου, τρέχουσας κρυφής εξόδου και προηγούμενης κρυφής εξόδου αντίστοιχα, ενώ τα b αποτελούν τα διανύσματα bias. Ο τελεστής $*$ υποδηλώνει το γινόμενο στοιχείου με στοιχείο, σ είναι η logistic sigmoid και \tanh η συνάρτηση μεταφοράς της υπερβολικής εφαπτομένης αντίστοιχα [114].

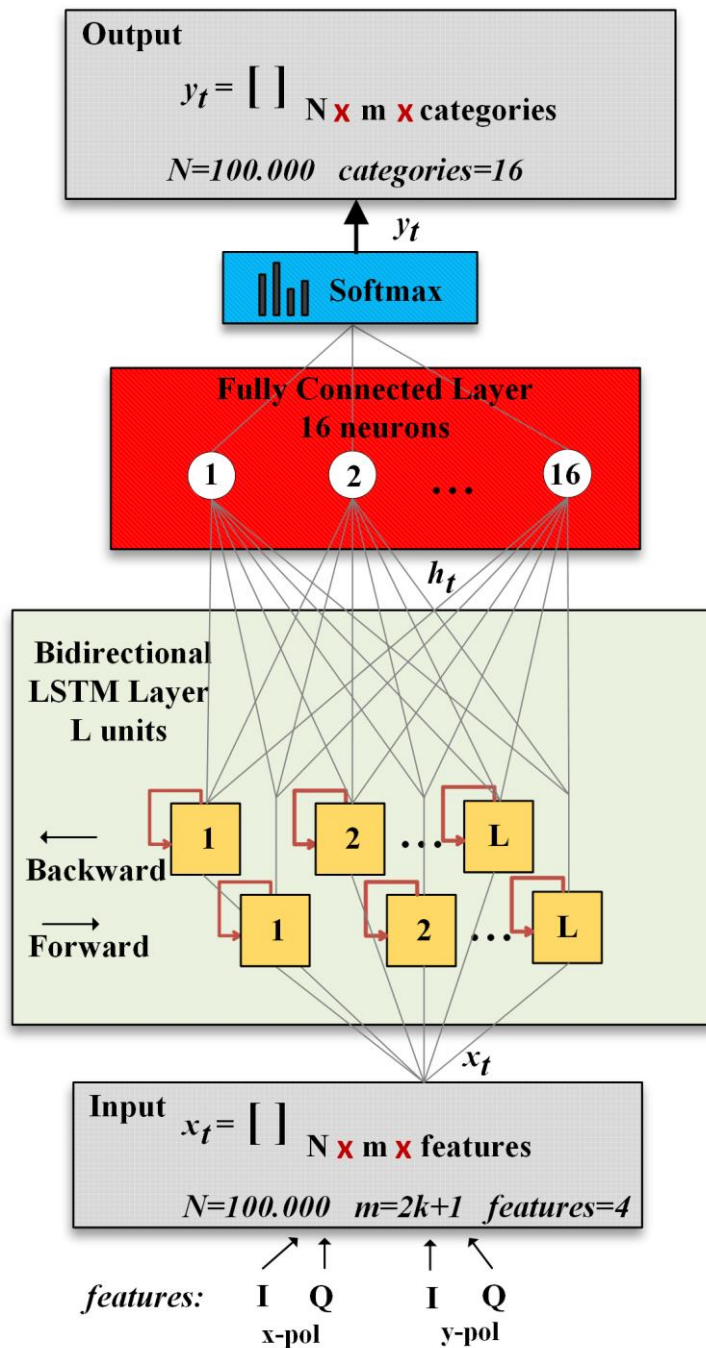
Στην παράγραφο αυτή χρησιμοποιούμε ένα ακολουθιακό δίκτυο βασισμένο στο αμφίδρομο LSTM η αρχιτεκτονική του οποίου φαίνεται στο σχήμα 4.3α, ενώ ο τρόπος που τα διαδοχικά LSTM τμήματα συνδέονται μεταξύ τους στο χρόνο φαίνεται στο σχήμα 4.3β.

Η είσοδος x_t είναι η αλλοιωμένη από τη διάδοση συμβολοσειρά και έχει την ακόλουθη μορφή:

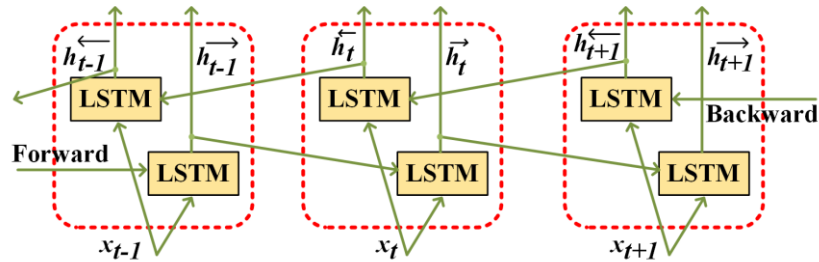
$$X_{t,m} = [X_{t-k}, \dots, X_{t-1}, X_t, X_{t+1}, \dots, X_{t+k}] \tag{4.3}$$

όπου το m συμβολίζει το συνολικό μήκος της «λέξης», δηλαδή τον αριθμό των διαδοχικών συμβόλων που διαδίδονται και υπολογίζεται από τη σχέση

$$m = 2k + 1 \tag{4.4}$$



Σχήμα 4.3α: Αρχιτεκτονική δικτύου bi-LSTM



Σχήμα 4.3β: Διασύνδεση bi-LSTM στον χρόνο

Η σχέση 4.4 δείχνει πως τη χρονική στιγμή t τροφοδοτούμε το δίκτυο μας με συμβολοσειρά μήκους m η οποία περιέχει εκτός από το κεντρικό σύμβολο ενδιαφέροντος, k προηγούμενα και k επόμενα σύμβολα, εφαρμόζοντας δηλαδή την τοπολογία πολλά σε ένα ή πολλά σε πολλά ($m=3, 5, 7, \dots, 201$). Το βασικό πλεονέκτημα του bi-LSTM είναι ότι παρακολουθεί τις εξαρτήσεις μεταξύ των διαδοχικών συμβόλων καθώς λαμβάνει υπόψη τη μη γραμμική τους αλληλεπίδραση μέσω της αντίστοιχης αλληλεπίδρασης των φαινομένων της χρωματικής διασποράς και του μη-γραμμικού φαινομένου Kerr που συναντήσαμε στο κεφάλαιο 2. Για μεγαλύτερες τιμές χρωματικής διασποράς (ζώνη C), το μήκος της λέξης πρέπει να αυξηθεί ούτως ώστε το bi-LSTM δίκτυο να διακρίνει και να «μάθει» τα πολυάριθμα μοτίβα που δημιουργούνται ως αποτέλεσμα της αλληλεπίδρασης διασποράς και μη γραμμικότητας.

Αντιθέτως, από τη σχέση 4.4 επίσης φαίνεται πως αν $k=0$, εφαρμόζουμε ουσιαστικά τις τοπολογίες ένα προς ένα ή ένα προς πολλά, οι οποίες όμως στο πρόβλημα που εξετάζουμε δεν έχουν κανένα πρακτικό ενδιαφέρον, μιας και το κανάλι έχει μνήμη, η οποία αν δε ληφθεί υπόψη, τότε το νευρωνικό δίκτυο θα αποτύχει στην εξίσωση του μη-γραμμικού καναλιού.

Από το σχήμα 4.2 βλέπουμε επίσης πως κάθε τιμή εισόδου $Xx-pol$ and $Xy-pol$ περιέχει τέσσερις τιμές/ χαρακτηριστικά (features), τα I και Q της κάθε πόλωσης.

Για να υπολογίσουμε το BER των προβλεπόμενων συμβόλων οδηγούμε την έξοδο του δικτύου LSTM σε ένα πλήρως συνδεδεμένο στρώμα (FCL) των 16 κόμβων και στη συνέχεια σε ένα στρώμα Softmax που πραγματοποιεί την ταξινόμηση μεταξύ 16

συμβόλων κατηγοριών/QAM συμβόλων και των δύο πολώσεων για όλα τα σύμβολα στην έξοδο Y_{x-pol} .

Η έξοδος τη χρονική στιγμή t μπορεί να εκφραστεί από τη σχέση :

$$y_t = \text{Softmax}[W_{\text{Softmax}}h_t + b_{\text{Softmax}}] \quad (4.4)$$

Όπου $h_t = \frac{1}{2}(\vec{h}_t + \overleftarrow{h}_t)$ με τα \vec{h}_t και \overleftarrow{h}_t να είναι οι έξοδοι του LSTM στην ορθή και αντίστροφη φορά, W_{Softmax} και b_{Softmax} τα βάρη και η τιμή του bias αντίστοιχα.

Η Softmax είναι συνάρτηση ενεργοποίησης που ορίζεται ως $y = \text{Softmax}(z)$ που εκχωρεί πιθανότητες σε κάθε κατηγορία. Στη συγκεκριμένη περίπτωση

$$y_i = \frac{e^{z_i}}{\sum_{16} e^{z_i}} \quad (4.5)$$

Το BER υπολογίζεται για το κεντρικό σύμβολο κάθε λέξης $y_{t,m}$ σε σχέση με το αντίστοιχο $\hat{y}_{t,m}$, εφαρμόζοντας την τοπολογία πολλά σε ένα.

Το bi-LSTM δίκτυο σχεδιάζεται, εκπαιδεύεται και αξιολογείται στο *Keras* [130] με τη βοήθεια της βιβλιοθήκης *Tensorflow GPU* [131]. Δεδομένου του ότι η εκπαίδευση βασίζεται στην ταξινόμηση επιλέγεται η συνάρτηση σφάλματος categorical cross-entropy. Ως αλγόριθμος βελτιστοποίησης για την ανανέωση των βαρών του συνολικού δικτύου επιλέγεται ο Adam ενώ ο ρυθμός εκπαίδευσης ορίζεται στη τιμή 10^{-4} . Οι τιμές των υπολοίπων παραμέτρων του αλγορίθμου ορίζονται στις τιμές που αυτός παρουσιάζει τη βέλτιστη συμπεριφορά ήτοι, $b_1 = 0.9$, $b_2 = 0.999$ και $e = 10^{-7}$ σύμφωνα με την [106].

Από τις $N=100.000$ διαθέσιμες λέξεις των συμβόλων μήκους m χρησιμοποιούμε τις 60.000 για εκπαίδευση, τις 20.000 για επικύρωση και τις υπόλοιπες 20.000 για εκτίμηση της ακρίβειας της εκπαίδευσης σε δεδομένα που το δίκτυο δεν έχει ξαναδεί. Η διαδικασία της εκπαίδευσης γίνεται με την προσέγγιση του Mini-batch Gradient Decent χρησιμοποιώντας πακέτα (batches) των 512 λέξεων. Η τιμή αυτή επιλέχθηκε ως η χρυσή τομή μεταξύ δέσμωσης της διαθέσιμης μνήμης του υπολογιστικού συστήματος που είχαμε στη διάθεσή μας (βλ. κεφ. 7) και του χρόνου εκτέλεσης.

Ο μέγιστος αριθμός των εποχών, δηλαδή των πλήρων κύκλων εκπαίδευσης με εφαρμογή του BPTT σε όλο το διαθέσιμο σετ δεδομένων ορίζεται στις 400. Ωστόσο, χρησιμοποιούμε

το μηχανισμό early-stopping, τη διακοπή δηλαδή της εκπαίδευσης όταν η τιμή του σφάλματος παραμένει αμετάβλητη, για να αποφύγουμε την υπερπροσαρμογή (overfitting) για το πολύ 20 εποχές διατηρώντας τα βάρη με τη μικρότερη τιμή της categorical cross-entropy.

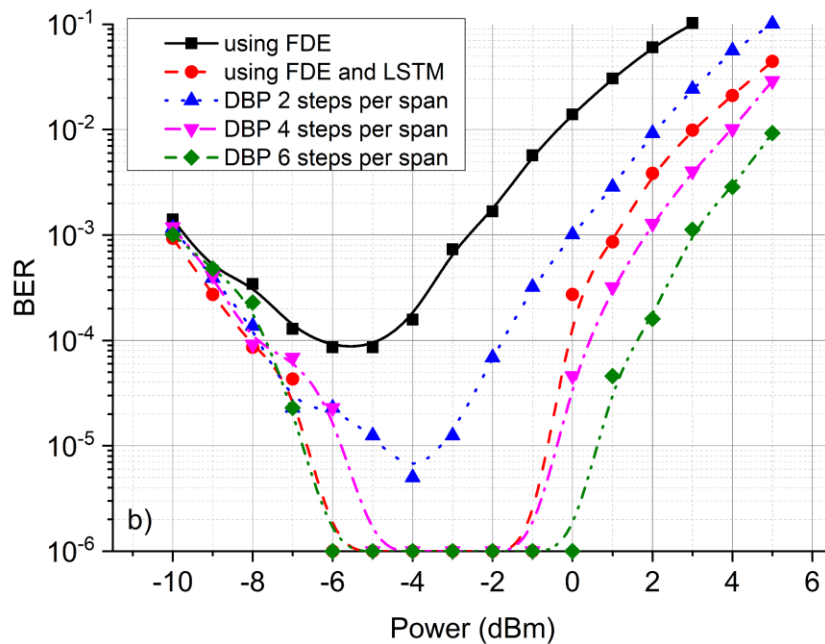
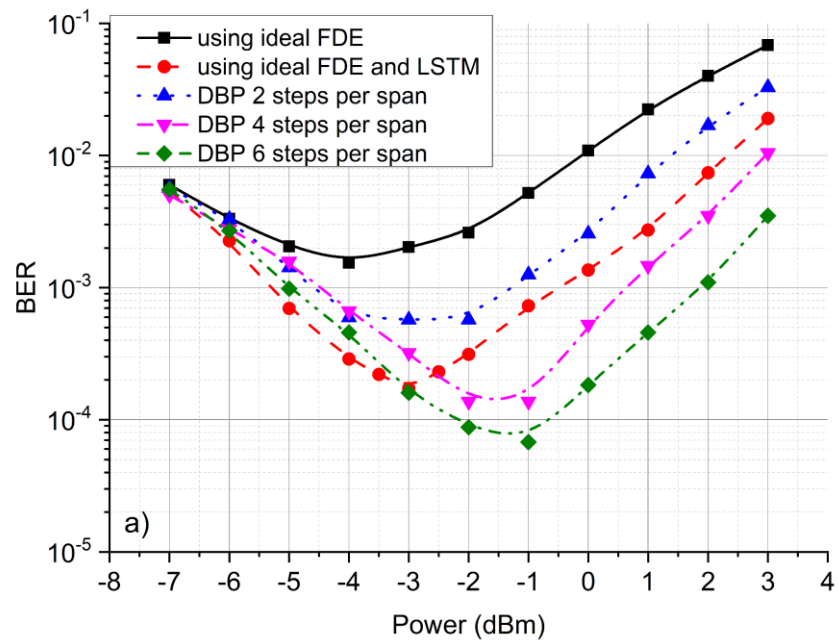
4.2.3 Παρουσίαση Αποτελεσμάτων για διάδοση σε ένα ή πολλά κανάλια

Η προσομοίωση της διάδοσης πραγματοποιήθηκε στο λογισμικό Matlab τόσο στην περίπτωση μονοκάναλης όσο και πολυκάναλης διάδοσης, θεωρώντας στη δεύτερη περίπτωση 10 WDM κανάλια με 50GHz τη μεταξύ τους απόσταση. Ιδιαίτερη μέριμνα δόθηκε στη διαδικασία παραγωγής των ψευδοτυχαίων ακολουθιών. Συγκεκριμένα, επιλέχθηκε η rng('shuffle') και η γεννήτρια τυχαίων αριθμών Mersenne Twister με περίοδο ($2^{19937} - 1$). Η επιλογή αυτή γίνεται για να αποφευχθεί το φαινόμενο το bi-LSTM να μαθαίνει την ψευδοτυχαία ακολουθία και έτσι να προβλέπει τις επόμενες τιμές [132].

Με σκοπό τον υπολογισμό αρκετά μικρών τιμών BER με μεγαλύτερη ακρίβεια, επαναλαμβάνουμε με την ίδια διαδικασία την προσομοίωση παράγοντας επιπλέον 20.000 σύμβολα για την εκτίμηση. Με 40.000 πλέον 16QAM σύμβολα η ελάχιστη τιμή BER που μπορεί να υπολογιστεί είναι 10^{-6} .

Παρά το γεγονός πως η διάδοση ενός μόνο καναλιού (μονοχρωματική διάδοση) δε συναντάται στα πραγματικά συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων, το σενάριο αυτό είναι πολύ χρήσιμο για την εκτίμηση σε πρώτο βαθμό της συμπεριφοράς του προτεινόμενου bi-LSTM ισοσταθμιστή, σε περιβάλλον στο οποίο το κυρίαρχο φαινόμενο είναι το SPM (ενδοκαναλική μη γραμμικότητα).

Στο σχήμα 4.4 αποτυπώνεται η τιμή BER σε σχέση με την οπτική ισχύ για διάφορα σενάρια τα οποία αναφέρονται ως ιδανική FDE, ιδανική FDE και bi-LSTM και DBP των 2, 4 και 6 βημάτων ανά την απόσταση των ενισχυτών για α) 700km διάδοση στη ζώνη O (1310nm) και β) 1500km διάδοση στη ζώνη C (1550nm). Το bi-LSTM σε αυτή τη μελέτη διαθέτει $L=32$ κρυφές μονάδες ενώ το μήκος των λέξεων ορίζεται σε $m=71$, κοινό και για τις δύο ζώνες O και C.



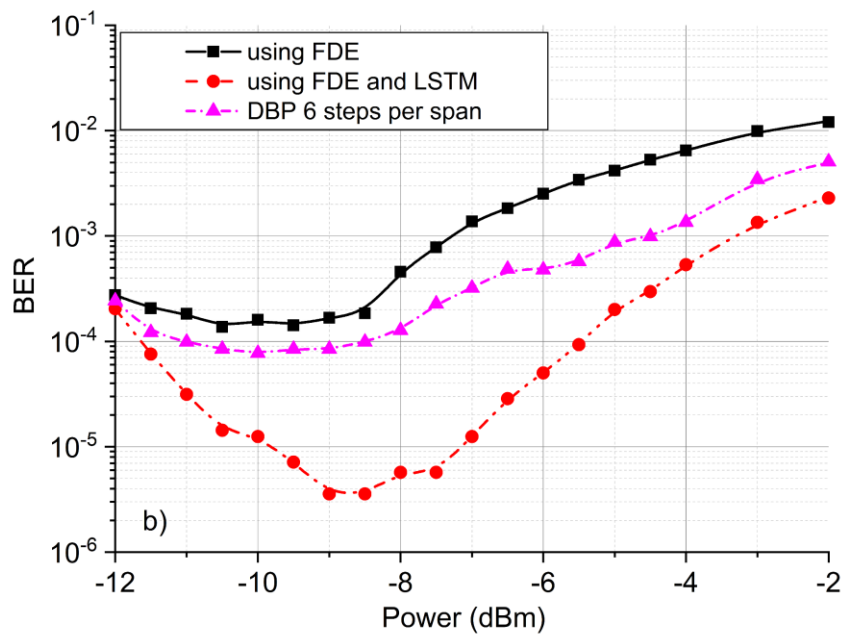
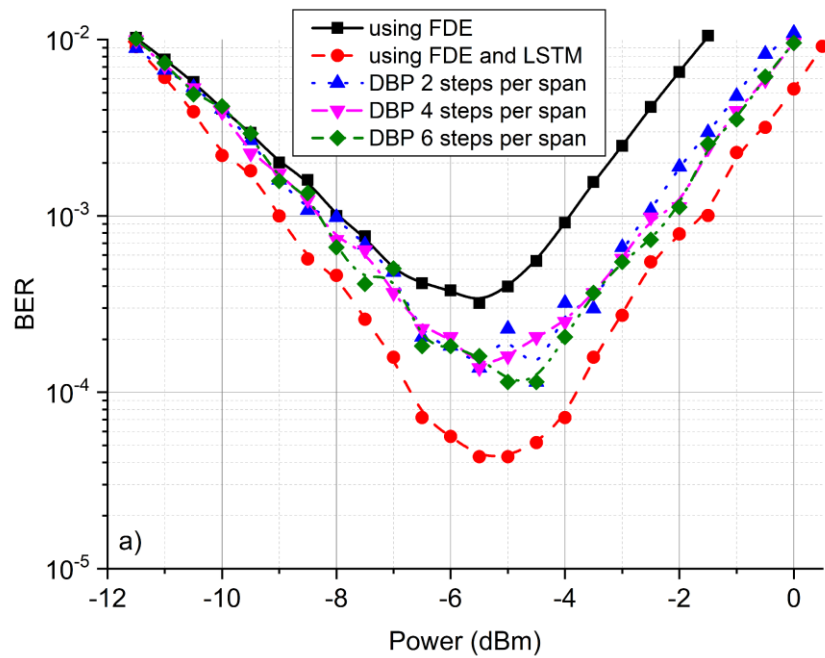
Σχήμα 4.4: Διάδοση ενός καναλιού α) 1310nm για 700km β) 1550nm για 1500km

Από το σχήμα 4.4 είναι φανερό πως όσο αυξάνεται ο αριθμός των βημάτων ανά απόσταση ενισχυτών του DBP, αυτό αποκτά καλύτερες επιδόσεις BER με το τίμημα της αναπόφευκτης αύξησης της υπολογιστικής πολυπλοκότητας. Η σύγκριση των 4.4α και 4.4β επιβεβαιώνει το συμπέρασμα πως η διάδοση στη ζώνη C είναι πιο ανεκτική στα μη γραμμικά φαινόμενα εξαιτίας της χαμηλότερης εξασθένισης και μεγαλύτερης διασποράς, όπως θα εξηγήσουμε αναλυτικά παρακάτω.

Από την άλλη μεριά, το bi-LSTM κατορθώνει να βελτιώσει την τιμή BER περισσότερο από μια τάξη μεγέθους συγκρινόμενο με τον γραμμικό αλγόριθμο του FDE ή ακόμη και δύο τάξεις μεγέθους στην περίπτωση της ζώνης C. Υπερνικά την τεχνική DBP των 2 βημάτων ανά span, παρουσιάζει ισάξιες επιδόσεις με την τεχνική DBP των 4 βημάτων ανά span, όχι όμως και την τεχνική DBP των 6 βημάτων ανά την απόσταση ενισχυτών το οποίο με τη σειρά του παρουσιάζει τις συγκριτικά καλύτερες επιδόσεις. Θα πρέπει όμως να θυμίσουμε πως η τεχνική DPB στις προσομοιώσεις απαιτεί 4 δείγματα ανά σύμβολο ενώ το bi-LSTM κατορθώνει να ταξινομήσει επαρκώς τα σύμβολα στη σωστή τους κλάση χρησιμοποιώντας μονάχα 1 δείγμα ανά σύμβολο.

Συνεχίζοντας την ανάλυση στην περιοχή της πολυκάναλης διάδοσης, επικεντρωνόμαστε στο κεντρικό WDM κανάλι. Το σύνολο των μεθόδων επεξεργασίας βασίζονται στο κεντρικό κανάλι το οποίο είναι και το μόνο που ανιχνεύεται στον δέκτη. Τα αποτελέσματα αποτυπώνονται στο σχήμα 4.5, στο οποίο αποδεικνύεται emphaticά η υπεροχή του bi-LSTM συγκριτικά με το DPB καθώς το δεύτερο λαμβάνει υπόψη μόνο τα ενδοκαναλικά φαινόμενα και αγνοεί τα διακαναλικά. Παρότι το bi-LSTM εκπαιδεύεται βασίζόμενο μονάχα στην πληροφορία που υπάρχει στο κεντρικό κανάλι, εντούτοις καταφέρνει να αναγνωρίσει σε σημαντικό βαθμό τις αλληλοεπιδράσεις μεταξύ καναλιών και κατορθώνει να ταξινομήσει τις άγνωστες σε αυτό συμβολοσειρές με σημαντικά μεγαλύτερη ακρίβεια σε σχέση με την FDE, αλλά και με την τεχνική DBP.

Η δυνατότητα της αναγνώρισης των πολυκαναλικών επιδράσεων ακόμη και στην περίπτωση της μονοκαναλικής φώρασης έχει συζητηθεί στο παρελθόν και βασίζεται στη λογαριθμική διαταραχή ανάλυσης συχνότητας η οποία δείχνει ότι τα συστήματα οπτικών επικοινωνιών μπορούν να μοντελοποιηθούν ως γραμμικά χρονικά μεταβαλλόμενα συστήματα, όπου η επίδραση της διακαναλικής μη γραμμικότητας εμφανίζεται ως



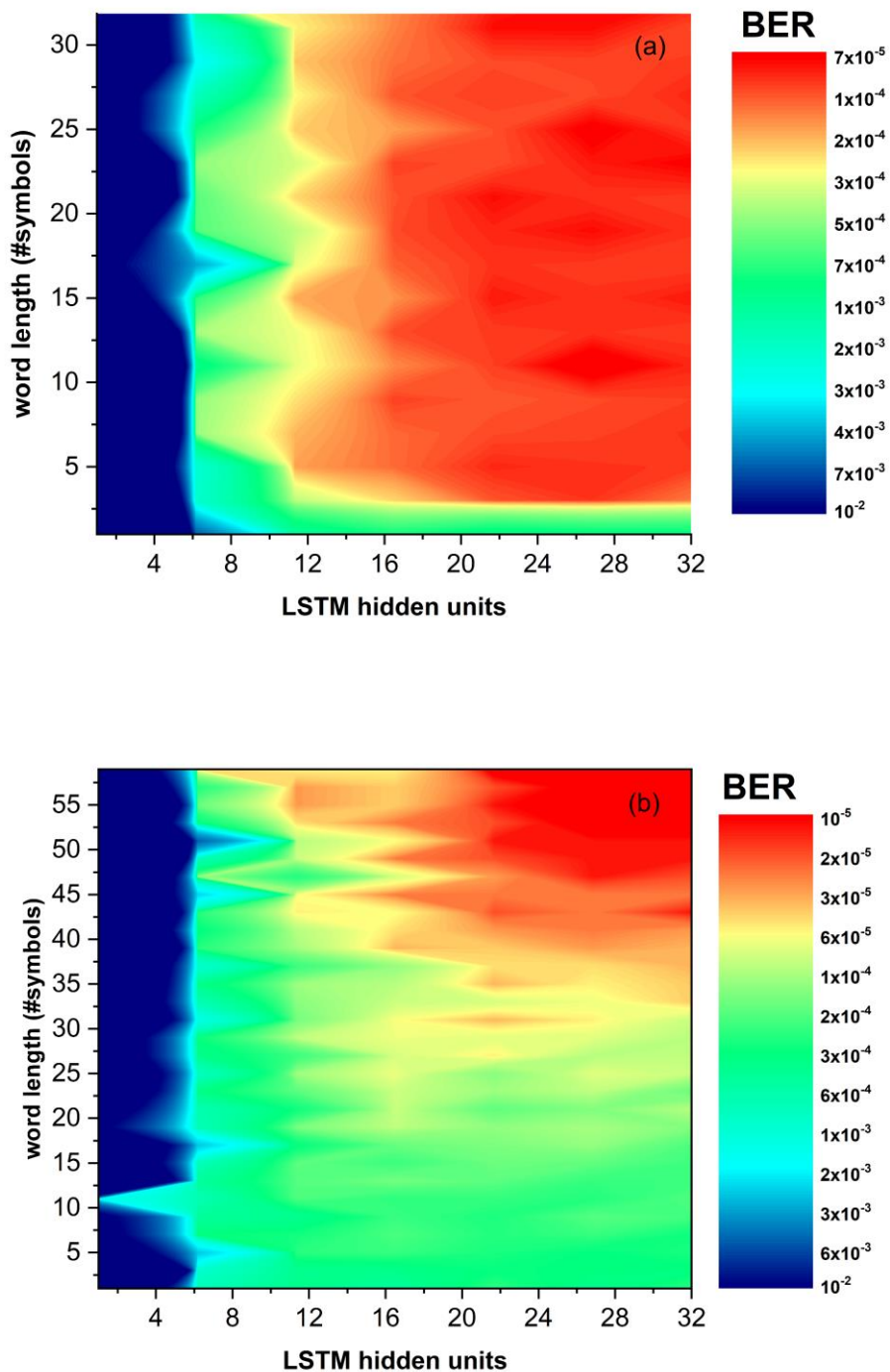
Σχήμα 4.5: Διάδοση 10 WDM καναλιών α) 1310nm για 300km β) 1550nm για 500km

γραμμική χρονικά μεταβαλλόμενη παρεμβολή μεταξύ συμβόλων. Εάν ο χρόνος συμφωνίας του καναλιού είναι πολύ μεγαλύτερος από την περίοδο των συμβόλων, το φαινόμενο XPM είναι αρκετά αργό παρέχοντας στον bi-LSTM ισοσταθμιστή τη δυνατότητα να παρακολουθεί σχεδόν ντετερμινιστικά τις επιδράσεις του [133]. Αυτό όπως θα δούμε παρακάτω επιτυγχάνεται καλύτερα όταν η χρωματική διασπορά μεγαλώνει, οπότε και η αλληλοπαρεμβολή των συμβόλων εκτείνεται σε μεγαλύτερο χρονικό διάστημα επηρεάζοντας και την τιμή του χρόνου συμφωνίας.

Στη ζώνη O, το bi-LSTM καταφέρνει να βελτιώσει την τιμή BER κατά μία τάξη μεγέθους σε σύγκριση με την εξίσωση FDE και κατά μισή τάξη μεγέθους σε σύγκριση με την εξίσωση DBP (σχήμα 4.5α). Η αποτελεσματικότητά του είναι σημαντική και στη ζώνη C, όπου η βελτίωση του BER ξεπερνά τη μια τάξη μεγέθους σε σύγκριση με FDE και DBP (σχήμα 4.5β). Ο κύριος λόγος των καλύτερων αποτελεσμάτων στη ζώνη C είναι ότι τα φαινόμενα μεταξύ των καναλιών σε αυτήν την περίπτωση κυριαρχούνται από τις αργές διακυμάνσεις XPM λόγω της σημαντικής διαφοράς ταχύτητας μεταξύ των γειτονικών καναλιών (φαινόμενο walk-off), ενώ στη ζώνη O, η ισχύς των διακαναλικών φαινομένων αυξάνεται ως αποτέλεσμα χαμηλότερης διασποράς. Η ζώνη O ενισχύει την παρουσία μίξης τεσσάρων κυμάτων (FWM) μεταξύ σημάτων WDM και μεταξύ σήματος και θορύβου [134].

Παρότι έχουμε πιστοποιήσει τη δυνατότητα του bi-LSTM να ελέγξει και εξομαλύνει τα μη γραμμικά φαινόμενα δεν έχουμε ακόμη αναλύσει την επίδραση των εσωτερικών χαρακτηριστικών του LSTM στις επιδόσεις της ταξινόμησης. Ο ρόλος του αριθμού των κρυφών μονάδων και του μήκους της λέξης των συμβόλων απεικονίζεται στο σχήμα 4.6. Είναι φανερό πως όσο αυξάνεται ο αριθμός των κρυφών μονάδων, βελτιώνονται τα επίπεδα της τιμής BER. Μπορούμε με ασφάλεια να ισχυριστούμε πως όταν ο αριθμός των κρυφών μονάδων ξεπεράσει τις 20 στο συγκεκριμένο πρόβλημα, το BER προσεγγίζει την βέλτιστη (χαμηλότερη) τιμή του.

Επιπρόσθετα το μήκος της λέξης των συμβόλων είναι εξίσου σημαντικό και σχετίζεται με τη μνήμη του καναλιού η οποία καθορίζεται από τη συσσωρευμένη διασπορά. Προκειμένου να τονιστούν καλύτερα οι βαθύτεροι λόγοι αυτής της συμπεριφοράς, πρέπει



Σχήμα 4.6: Το BER συναρτήσεσι των κρυφών μονάδων του bi-LSTM και του μήκους της λέξης των συμβόλων στη διάδοση 10 WDM καναλιών α) 1310nm για 300km β) 1550nm για 500km

να σημειωθεί ότι επιτυγχάνουμε ταξινόμηση με τη χρήση της Softmax που αντιστοιχίζει άμεσα την έξοδο του δικτύου LSTM y_t με το αρχικό 16-QAM σύμβολο x_t , όταν η εκπαίδευση του LSTM πραγματοποιείται με τη λέξη που περιέχει k προηγούμενα και k διαδοχικά σύμβολα του x_t . Όσο μεγαλύτερη είναι η ακολουθία που χρησιμοποιείται για την εκπαίδευση του δικτύου, τόσο μεγαλύτερη είναι η μνήμη του καναλιού που λαμβάνεται υπόψη.

Είναι ευρέως γνωστό ότι η μνήμη καναλιών των οπτικών ινών συνδέεται με την καθυστέρηση διεύρυνσης που προκαλείται από τη χρωματική διασπορά. Όταν ένα σήμα οπτικού εύρους ζώνης ίσο με B διαδίδεται σε κανάλι μήκους L , αντιμετωπίζει μια καθυστέρηση διασποράς ίση με

$$\Delta t_{CD} = 2\pi|\beta_2|LB \quad (4.6)$$

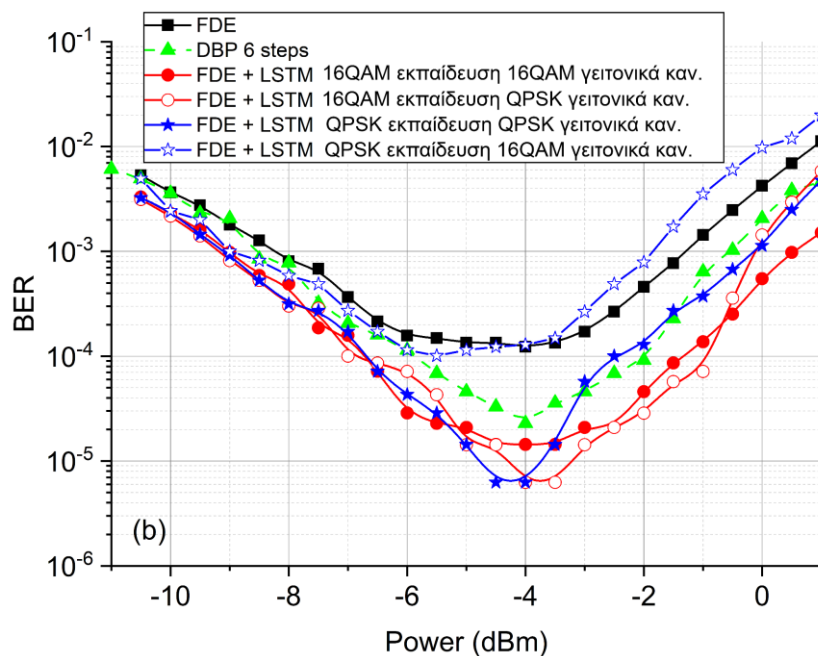
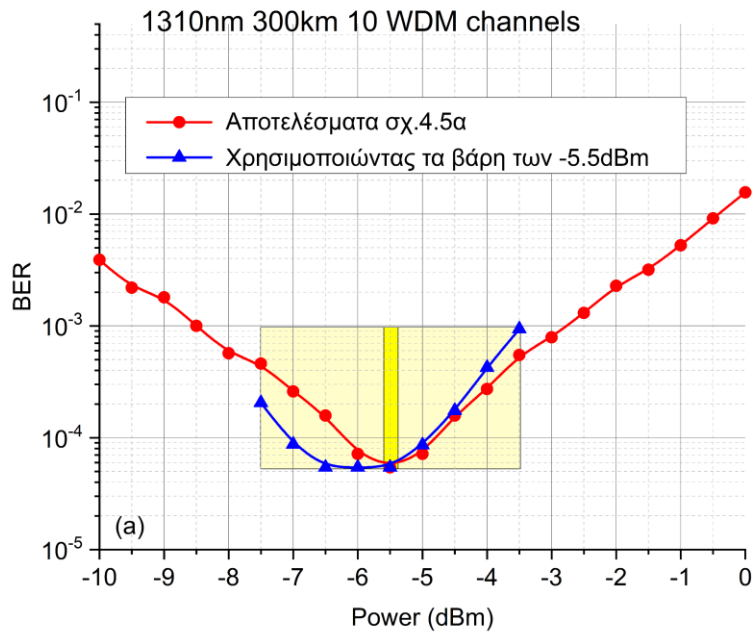
Στην περιοχή μήκων κύματος 1310 nm, έχοντας θεωρήσει τη χρωματική διασπορά $\beta_2 = -0.82 \text{ ps}^2/\text{km}$, αυτός ο απλός τύπος δίνει μια καθυστέρηση διασποράς 77 ps, επομένως η ενδοκαναλική αλληλοπαρεμβολή συμβόλων (ISI) εκτείνεται σε διάρκεια δύο παλμών (25 Gbaud αντιστοιχούν σε διάρκεια παλμού 40 ps). Αντίθετα, στα 1550 nm η καθυστέρηση διασποράς ανέρχεται σε 3-4 ns για 500 km, πράγμα που σημαίνει ότι τα ενδοκαναλικά μη γραμμικά φαινόμενα θα μπορούσαν να περιλαμβάνουν ISI μεταξύ 85 γειτονικών συμβόλων. Λαμβάνοντας υπόψη τη διάδοση WDM, η μνήμη του καναλιού αυξάνεται εξαιτίας του φαινομένου του walk-off μεταξύ των συν-διαδιδόμενων κυμάτων, ωστόσο η μη γραμμική αλληλεπίδραση μειώνεται ως αποτέλεσμα του ίδιου φαινομένου. Το σχ. 4.6 δείχνει ξεκάθαρα ότι στη ζώνη O, όπου η μνήμη του καναλιού είναι μικρότερη (2 σύμβολα στην περίπτωση ενός καναλιού), οι επιδόσεις BER φτάνουν τη βέλτιστη τιμή τους για μήκος λέξης πάνω από 3 σύμβολα. Αντίθετα, στη ζώνη C όπου η πεπερασμένη μνήμη του καναλιού αυξάνεται λόγω μεγαλύτερης διασποράς, επιτυγχάνεται η βέλτιστη τιμή για μήκος λέξης που ξεπερνά τα 50 σύμβολα.

4.2.3 Ανθεκτικότητα της εκπαιδευμένου μοντέλου στις διακυμάνσεις του καναλιού

Το επόμενο θέμα αυτής της παραγράφου είναι να προσδιορίσει την ανθεκτικότητα του εκπαιδευμένου μοντέλου bi-LSTM στις διάφορες διακυμάνσεις του καναλιού. Γενικά, το κανάλι οπτικών ινών είναι λιγότερο χρονικά μεταβαλλόμενο σε σύγκριση π.χ. με το ασύρματο κανάλι. Οι κύριες διακυμάνσεις που θα μπορούσαν να επηρεάσουν τη σταθερότητα της διάδοσης είναι οι διακυμάνσεις ισχύος, οι περιστροφές της πόλωσης και η αλλαγή των μορφών διαμόρφωσης γειτονικών WDM καναλιών στα δυναμικά οπτικά δίκτυα. Μια άλλη κρίσιμη παράμετρος είναι η αλλαγή του μήκους διάδοσης, ωστόσο αυτό μπορεί να αντιμετωπιστεί εάν ο δέκτης είναι εκ των προτέρων εκπαιδευμένος για διαφορετικές αποστάσεις διάδοσης. Οι περιστροφές πόλωσης συμβαίνουν σε σχετικά χαμηλούς ρυθμούς και αντιμετωπίζονται με κατάλληλους αλγόριθμους στη μονάδα DSP του οπτικού δέκτη. Σε αυτή την παράγραφο, θεωρούμε ως πιο σημαντικές αλλαγές αυτές που αναφέρονται σε διακυμάνσεις ισχύος και μορφές μεταβλητής διαμόρφωσης γειτονικών καναλιών καθώς επηρεάζουν άμεσα τις μη γραμμικές ιδιότητες του καναλιού.

Η ισχύς επηρεάζει τη μη γραμμικότητα μέσω του φαινομένου Kerr και οι διακαναλικές αλλοιώσεις εξαρτώνται από την ισχύ των γειτονικών καναλιών και τη μορφή διαμόρφωσης. Στο σχήμα 4.7 μελετάμε την ανθεκτικότητα της εκπαίδευσης σε σχέση με τις παραλλαγές μορφής ισχύος και διαμόρφωσης. Στο σχ. 4.7α, η τιμή BER στην έξοδο της διάδοσης WDM της ζώνης O απεικονίζεται για το κεντρικό κανάλι. Η κόκκινη γραμμή με τους κύκλους δείχνει το αποτέλεσμα του σχ. 4.5 και αναφέρεται στην απόδοση που επιτυγχάνεται με τη χρήση του bi-LSTM μετά την εκπαίδευση για κάθε τιμή ισχύος που φαίνεται στο σχήμα.

Παρατηρούμε ότι η βέλτιστη απόδοση επιτυγχάνεται στα -5.5 dBm (μέση ισχύς ανά κανάλι). Έτσι, εάν θεωρήσουμε ότι το κανάλι παραμένει αναλλοίωτο, ο δέκτης θα πρέπει να λειτουργεί με βάση την εκπαίδευση που πραγματοποιείται στα -5.5 dBm. Στη συνέχεια, το ερώτημα είναι ποια θα είναι η απόδοση εάν η μέση ισχύς αλλάξει ελαφρά λόγω τυχαίων γεγονότων στη ζεύξη (πχ αλλάξει το κέρδος των ενισχυτών, μειωθεί η ισχύς του πομπού κ.λπ.). Η μπλε γραμμή στο σχήμα 4.7α απαντά σε αυτό το ερώτημα δείχνοντας πώς θα συμπεριφερθεί ο δέκτης που έχει εκπαιδευτεί στα -5.5 dBm.



Σχήμα 4.7: Αξιολόγηση της ανθεκτικότητας του μοντέλου σε παραλλαγές καναλιών: α) Ευαισθησία της εκπαίδευσης σε διακυμάνσεις ισχύος (1310 nm, 16-QAM, 300 km), β) Ευαισθησία της εκπαίδευσης στην τροποποίηση της μορφής διαμόρφωσης γειτονικών καναλιών (1310 nm, γειτονικές QPSK για ένα κεντρικό μήκος κύματος 16-QAM, 300 km).

Για τιμές ισχύος που κυμαίνονται από -7.5 dBm έως -3.5 dBm (± 2 dB διακύμανση γύρω από την ονομαστική τιμή των -5.5 dBm) το BER δεν επηρεάζεται ουσιαστικά. Αντίθετα, φαίνεται ότι εάν κάποιος εφαρμόσει τις τιμές των βαρών που υπολογίστηκαν από την εκπαίδευση στα -5.5 dBm σε χαμηλότερες τιμές ισχύος, το BER θα βελτιωθεί ελαφρώς, επομένως η εκπαίδευση που αποκτάται στο βέλτιστο σημείο δρα ευεργετικά σε χαμηλότερες τιμές ισχύος. Η κύρια εξήγηση για αυτή τη συμπεριφορά είναι ότι στο σημείο που το μοντέλο παρουσιάζει την μεγαλύτερη ακρίβεια ταξινόμησης (ελάχιστη τιμή BER), υπάρχει ισοδυναμία μεταξύ γραμμικού και μη γραμμικού θορύβου. Οι τιμές των βαρών που προέκυψαν από την εκπαίδευση μπορούν να εφαρμοστούν σε χαμηλότερες τιμές ισχύος, καθώς ο μη γραμμικός θόρυβος θα είναι μικρότερος, επομένως ένα σύστημα που έχει μάθει να λειτουργεί σε πιο εχθρικό περιβάλλον όσον αφορά τις μη γραμμικότητες μπορεί να προσαρμοστεί σε ένα λιγότερο μη γραμμικό οικοσύστημα.

Ο δεύτερος τρόπος για να ελεγχθεί η ανθεκτικότητα της εκπαίδευσης έναντι των διακυμάνσεων του καναλιού διάδοσης είναι η αλλαγή του σχήματος διαμόρφωσης των γειτονικών WDM καναλιών και η εκ νέου αξιολόγηση της απόδοσης.

Έτσι, εκπαιδεύουμε το bi-LSTM λαμβάνοντας υπόψη τα ακόλουθα σενάρια του σχήματος 4.7β):

- i) (γεμάτοι κόκκινοι κύκλοι) Δίκτυο εκπαιδευμένο με γειτονικά κανάλια 16-QAM και αξιολογούμενο σε γειτονικά κανάλια 16-QAM,
- ii) (κενοί κόκκινοι κύκλοι) Δίκτυο εκπαιδευμένο με γειτονικά κανάλια 16-QAM αλλά αξιολογούμενο σε γειτονικά κανάλια QPSK,
- iii) (γεμάτα μπλε αστέρια) Δίκτυο εκπαιδευμένο με γειτονικά κανάλια QPSK και αξιολογούμενο σε γειτονικά κανάλια QPSK και
- iv) (κενά μπλε αστέρια) Δίκτυο εκπαιδευμένο με γειτονικά κανάλια QPSK αλλά αξιολογούμενο σε γειτονικά κανάλια 16-QAM.

Ο στόχος μας ήταν να δούμε εάν τα βάρη του νευρωνικού δικτύου που ελήφθησαν για μια διάδοση αμιγώς 16-QAM καναλιών θα μπορούσαν να ισχύουν για ένα σύστημα

διάδοσης όπου το κεντρικό κανάλι παραμένει 16-QAM και τα γειτονικά κανάλια μεταπίπτουν σε QPSK και αντίστροφα.

Με βάση τα αποτελέσματα του σχ. 4.7β, φαίνεται ότι ο δέκτης που έχει εκπαιδευτεί στην περίπτωση διάδοσης αμιγώς 16-QAM παρέχει λίγο πολύ την ίδια επίδοση BER με τον δέκτη που ήταν εκπαιδευμένος σε ένα σύστημα διάδοσης όπου το κεντρικό κανάλι είναι 16-QAM και όλα τα άλλα μήκη κύματος μεταφέρουν σήματα QPSK. Αντίθετα, όταν το δίκτυο εκπαιδεύεται με γειτονικά κανάλια QPSK αλλά αξιολογείται σε γειτονικά κανάλια 16-QAM, η υποβάθμιση της απόδοσης είναι εμφανής, δείχνοντας την αδυναμία του δικτύου LSTM να εξισορροπήσει το σήμα σε πιο εχθρικά περιβάλλοντα (γειτονικά 16-QAM) από αυτό στο οποίο έχει εκπαιδευτεί (QPSK γείτονες). Επομένως, η εκπαίδευση του bi-LSTM είναι ισχυρή έναντι των αλλαγών του σχήματος διαμόρφωσης των γειτονικών καναλιών, υπό την προϋπόθεση ότι η εκπαίδευση έχει πραγματοποιηθεί με την πιο σύνθετη μορφή διαμόρφωσης που προβλέπεται να εφαρμοστεί στο σύστημα.

Συνοπτικά, ο ισοσταθμιστής bi-LSTM είναι ανθεκτικός στις διακυμάνσεις των καναλιών, υπό την προϋπόθεση ότι έχει εκπαιδευτεί στο πιο εχθρικό, μη γραμμικό περιβάλλον λειτουργίας ενός συγκεκριμένου συστήματος. Υπό αυτό το πρίσμα, ο προτεινόμενος ισοσταθμιστής είναι συμβατός με τα οπτικά δίκτυα επόμενης γενιάς, των οποίων οι πομποδέκτες, μέσω κατάλληλου λογισμικού, επιλέγουν συχνά το σχήμα διαμόρφωσης με σκοπό τη βέλτιστη εκμετάλλευση της διαθέσιμης χωρητικότητας του δικτύου.

4.2.4 Ανάλυση πολυπλοκότητας bi-LSTM

Στο εδάφιο αυτό θα γίνει η σύγκριση της πολυπλοκότητας των συγκρινόμενων μοντέλων ισοστάθμισης, bi-LSTM και DBP. Για το σκοπό αυτό θα εκτιμηθεί ο αριθμός των πολλαπλασιασμών που απαιτούνται σε κάθε περίπτωση.

Για το μεν DPB σύμφωνα με την αναφορά [135] μπορεί να υπολογιστεί από τον τύπο :

$$C_{DBP} = 4N_{Span}N_{StpSt} \left[\frac{N(\log_2 N + 1)n_s}{(N - N_D + 1)\log_2 M} + n_s \right] \quad (4.7)$$

Όπου N_{span} είναι ο αριθμός των spans, N_{stpSt} ο αριθμός των βημάτων ανά span, N το μέγεθος του FFT το οποίο εξαρτάται από τη συσσωρευμένη διασπορά, n_s ο συντελεστής υπερδειγματοληψίας, M η τάξη της διαμόρφωσης, και $N_D = \frac{n_s \tau_D}{T}$, όπου το τ_D αντιστοιχεί στην παλμική απόκριση του καναλιού και T η χρονική διάρκεια του παλμού. [135]

Πολλαπλασιάζουμε την πολυπλοκότητα επί 4 καθώς στην τεχνική DBP χρησιμοποιούμε μιγαδικούς αριθμούς και ένας μιγαδικός πολλαπλασιασμός ισούται με τέσσερις πραγματικούς πολλαπλασιασμούς.

Για την υλοποίηση bi-LSTM πρέπει να προσθέσουμε την πολυπλοκότητα της εξίσωσης FDE :

$$C_{FDE} = 4 \frac{N(\log_2 N + 1)n_s}{(N - N_D + 1)\log_2 M} \quad (4.8)$$

Εδώ το N αντιστοιχεί στο μέγεθος FFT λαμβάνοντας υπόψη τη συσσωρευμένη διασπορά σε ολόκληρο το link της διάδοσης [135].

Η πολυπλοκότητα της αρχιτεκτονικής bi-LSTM αποτελείται από δύο μέρη: την εκπαίδευση (4.9) και την πολυπλοκότητα πρόβλεψης (4.10). Η υπολογιστική πολυπλοκότητα του προτεινόμενου δικτύου bi-LSTM εξαρτάται από τον αριθμό των κρυφών μονάδων και τον αριθμό των εισόδων (μήκος της ακολουθίας συμβόλων)

$$C_{train} = N_{ep} N_{TS} C_{pred} \quad (4.9)$$

$$C_{pred} = 2 \times 4L[(FH + H^2) + HM] \quad (4.10)$$

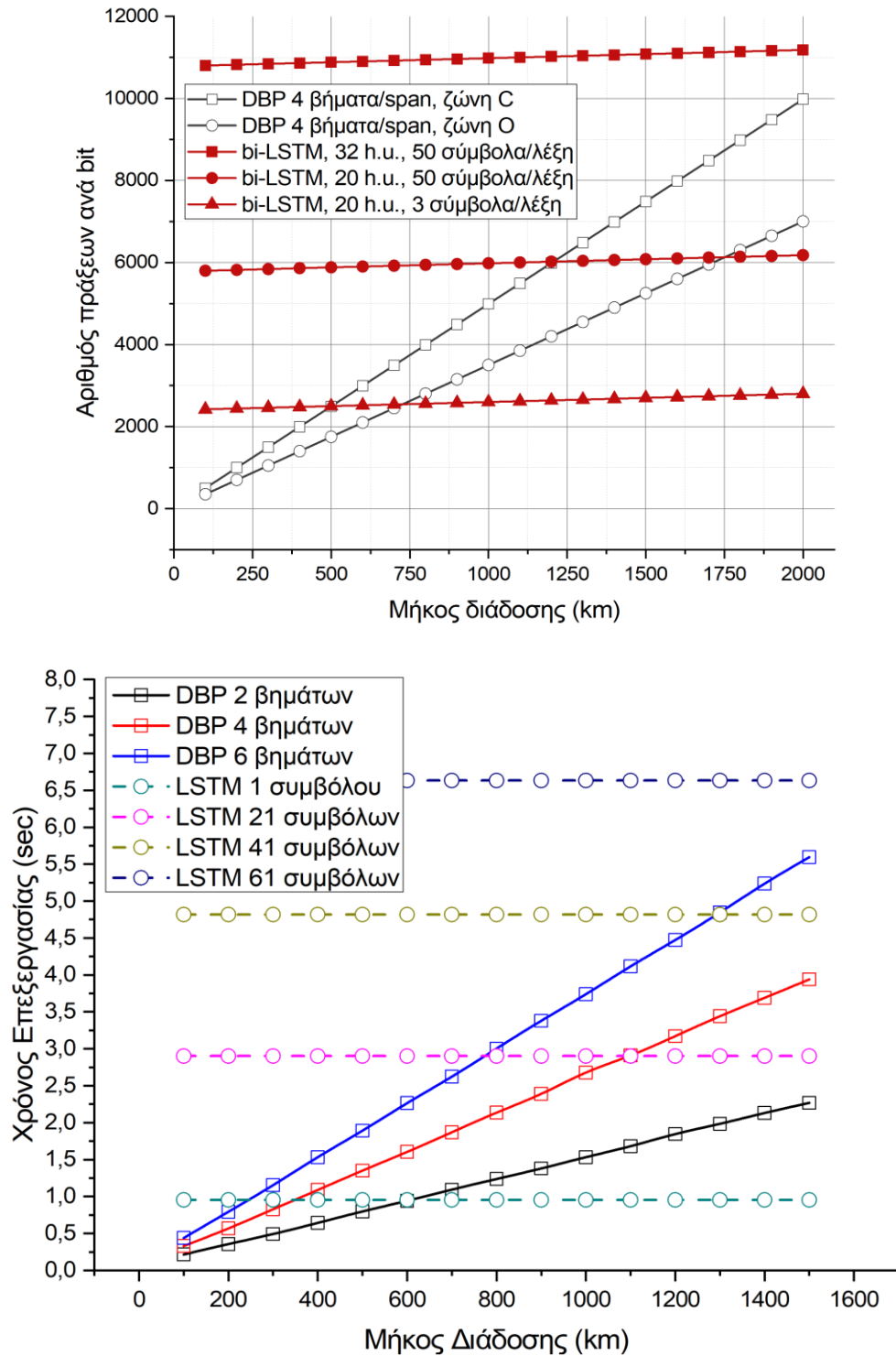
Όπου N_{ep} είναι ο αριθμός των εποχών εκπαίδευσης, N_{ts} ο αριθμός των συμβόλων που χρησιμοποιούνται για την εκπαίδευση, H ο αριθμός των κρυφών μονάδων LSTM, F ο αριθμός των χαρακτηριστικών εισόδου, M οι διαφορετικές κλάσεις της ταξινόμησης (16 σε αυτή την περίπτωση) και L το μήκος της λέξης των συμβόλων. Ο συντελεστής 2 υποδηλώνει την αμφίδρομη αρχιτεκτονική του μοντέλου ενώ ο συντελεστής 4 το πλήθος των πυλών του LSTM. Προκειμένου να υπολογιστεί η πολυπλοκότητα για το bi-LSTM πρέπει να προσδιοριστεί πόσο συχνά αναμένεται να λαμβάνει χώρα η εκπαίδευση.

Θα μπορούσε κανείς να υποθέσει ότι δεδομένου ότι το bi-LSTM είναι ανθεκτικό στις αλλαγές της ισχύος και του σχήματος διαμόρφωσης, το μόνο χρονικά μεταβαλλόμενο

στοιχείο που αναμένεται να διαταράξει το κανάλι είναι η τυχαία περιστροφή πόλωσης. Οι χρονικές κλίμακες για τα δυναμικά χαρακτηριστικά της πόλωσης ξεπερνούν τα λίγα ms. Μέσα σε ένα ms, σχεδόν 5×10^7 σύμβολα θα έχουν μεταδοθεί ανά πόλωση, ενώ η εκπαίδευση απαιτεί περίπου 20000 σύμβολα για 100 εποχές, που ισοδυναμεί με μια χρονική περίοδο 2×10^6 συμβόλων. Επομένως, ακόμα κι αν το οπτικό κανάλι μεταβάλλεται κάθε ms, η πολυπλοκότητα της εκπαίδευσης είναι μικρότερη από το 4% της πολυπλοκότητας της πρόβλεψης και έτσι μπορεί να αγνοηθεί με ασφάλεια.

Έχοντας αυτό υπόψη, υπολογίζουμε τον αριθμό των πράξεων ανά bit ως συνάρτηση της απόστασης διάδοσης για την περίπτωση του DBP, ενώ για το μοντέλο LSTM λαμβάνουμε υπόψη μόνο την πολυπλοκότητα πρόβλεψης bi-LSTM και αθροίζουμε αυτή του FDE. Όπως φαίνεται στο σχ. 4.8, η πολυπλοκότητα του DBP αυξάνεται γραμμικά ως αποτέλεσμα του αριθμού των spans του τύπου 4.7. Επιπλέον, το DBP δεν εξαρτάται τόσο πολύ από τη χρωματική διασπορά, δηλαδή η πολυπλοκότητα για μια συγκεκριμένη απόσταση δεν είναι τόσο διαφορετική στη ζώνη O και στη ζώνη C. Πρέπει να υπενθυμίσουμε ότι όσο μειώνεται η διασπορά, η παράμετρος N , δηλαδή το μέγεθος FFT ποικίλλει αναλογικά (τυπικές τιμές 64, 128, 256 για κάθε διάστημα) και εκεί οφείλεται η διακύμανση της πολυπλοκότητας.

Αντίθετα, το LSTM έχει υψηλή πολυπλοκότητα, η οποία ωστόσο δεν εξαρτάται τόσο πολύ από την απόσταση διάδοσης αν λάβουμε υπόψη μας σταθερό μήκος λέξης στη διαδικασία εκπαίδευσης (με εξαίρεση την πολυπλοκότητα FDE που είναι ωστόσο μικρή). Επιπλέον, η πολυπλοκότητά του μειώνεται πολύ όταν μειώνεται ο αριθμός των συμβόλων στη λέξη εισόδου/εξόδου και αυτό εξαρτάται από την αναμενόμενη συσσωρευμένη διασπορά που εξαρτάται από την απόσταση διάδοσης. Αυτό σημαίνει ότι το LSTM θα μπορούσε να είναι πολύ αποτελεσματικό σε ένα μη γραμμικό κανάλι με μικρή μνήμη, όπως στην περίπτωση της ζώνης O ή της ζώνης C με τη χρήση της οπτικής διαχείρισης διασποράς. Για 20 κρυφές μονάδες και διάδοση της ζώνης C, το LSTM με λέξεις των 50 συμβόλων γίνεται λιγότερο περίπλοκο από το DBP για αποστάσεις που ξεπερνούν τα 1200 km, ενώ η ίδια τάση παρατηρείται κάτω από 500 km όταν η διάδοση πραγματοποιείται στη ζώνη O όπου θεωρούμε πως η λέξη αποτελείται από 3 σύμβολα (σχήμα 4.8α).



Σχήμα 4.8: α) Υπολογιστική πολυπλοκότητα για bi-LSTM με FDE και DBP. (β) Χρόνος επεξεργασίας με την κάρτα γραφικών για διαφορετικές συνθήκες λειτουργίας, $H = 32$ σε αυτήν την περίπτωση.

Προκειμένου να αποκτήσουμε μια πιο ρεαλιστική εικόνα αυτών των τάσεων πραγματοποιήσαμε αριθμητικές προσομοιώσεις και των δύο αλγορίθμων σε μια GPU GTX 1070. Το DBP προσομοιώθηκε στο λογισμικό Matlab με GPU Coder χρησιμοποιώντας περιβάλλον ανάπτυξης εργαλειοθήκης CUDA 10.1 που περιλαμβάνει βιβλιοθήκες επιτάχυνση με τη βοήθεια της GPU, εργαλεία εντοπισμού σφαλμάτων και βελτιστοποίησης, έναν μεταγλωττιστή C/C++ και μια βιβλιοθήκη χρόνου εκτέλεσης. Το Bi-LSTM εκπαιδεύτηκε και δοκιμάστηκε στο Tensorflow 2.1.0 με την ίδια GPU χρησιμοποιώντας το ίδιο CUDA 10.1. Για τον ίδιο αριθμό συμβόλων παρατηρήσαμε παρόμοιες τάσεις που δείχνουν ξεκάθαρα την εξάρτηση της πολυπλοκότητας LSTM από τον αριθμό των κρυφών μονάδων και συμβόλων και την τάση του να λειτουργεί πιο αποτελεσματικά από το DBP σε αποστάσεις που ξεπερνούν τα 1000 km (σχήμα 4.8β).

4.2.5 Συμπέρασμα

Αυτή η μελέτη έδειξε τη δυνατότητα χρήσης των νευρωνικών δικτύων βασιζόμενων σε bi-LSTM ως μηχανών μετα-επεξεργασίας σε ψηφιακά σύμφωνα συστήματα. Το LSTM έχει παρόμοια αποτελεσματικότητα με το DBP στον μετριάσμο των ενδοκαναλικών επιδράσεων και ανώτερη απόδοση όταν υπάρχουν ενδοκαναλικά φαινόμενα. Η ανάλυση έδειξε πώς η απόδοση του αλγορίθμου επηρεάζεται από σημαντικές παραμέτρους (αριθμός κρυφών μονάδων, αριθμός συμβόλων στην είσοδο/έξοδο) και ότι η εκπαίδευση είναι ανεκτική στις διακυμάνσεις ισχύος και τις τροποποιήσεις της διαμόρφωσης γειτονικών καναλιών υπό την προϋπόθεση ότι η εκπαίδευση έχει πραγματοποιηθεί σε δυσκολότερο, μη γραμμικό περιβάλλον.

Τέλος, η ανάλυση πολυπλοκότητας έδειξε ότι το LSTM μπορούσε να ανταγωνιστεί το DBP ειδικά σε μεγάλες αποστάσεις και σε οπτικά κανάλια που παρουσιάζουν μικρή συσσωρευμένη διασπορά όπου η πολυπλοκότητα του LSTM μπορεί να μειωθεί.

4.3 Σύγκριση πολυπλοκότητας και απόδοσης διαφόρων επαναλαμβανόμενων νευρωνικών δικτύων για την εξισορρόπηση των μη γραμμικών φαινομένων της ίνας σε ψηφιακά σύμφωνα συστήματα

4.3.1 Εισαγωγή

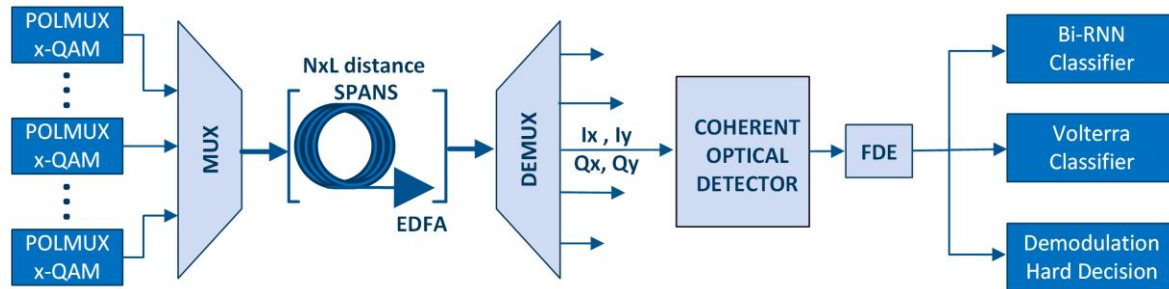
Στην προηγούμενη ενότητα μελετήσαμε το LSTM ως μια ανταγωνιστική της DBP μέθοδο εξομάλυνσης των μη γραμμικών φαινομένων της οπτικής ίνας σε ψηφιακά σύμφωνα δίκτυα μεγάλων αποστάσεων. Σε αυτή την ενότητα, επεκτείνουμε την ανάλυση εξετάζοντας δύο πρόσθετα μοντέλα RNN που είναι κατ' αρχήν λιγότερο πολύπλοκα από το LSTM, δηλαδή το Gate Recurrent Unit (GRU) και το VRNN, προκειμένου να διερευνήσουμε τη δυνατότητα υιοθέτησης αμφίδρομων μοντέλων RNN σε ψηφιακά σύμφωνα συστήματα οπτικών επικοινωνιών επόμενης γενιάς και μέτριας πολυπλοκότητας.

Παρέχουμε μια συγκριτική ανάλυση των τριών μοντέλων RNN, δηλαδή των LSTM, GRU και VRNN σε συστήματα WDM με μικρές αποστάσεις καναλιών και σχήματα διαμόρφωσης υψηλής τάξης. Επιπλέον, συγκρίνουμε τα μοντέλα bi-RNN με έναν μη γραμμικό ισοσταθμιστή Volterra 3ης τάξης που αποτελεί τεχνική αναφοράς με δυνατότητα αντιμετώπισης σύνθετων μη γραμμικών επιδράσεων διασποράς χρόνου [136], [74] τόσο από άποψη απόδοσης όσο και από άποψη πολυπλοκότητας. Δείχνουμε ξεκάθαρα ότι με τη χρήση της τοπολογίας πολλά σε πολλά, μειώνουμε δραστικά την πολυπλοκότητα των μοντέλων RNN έναντι του Volterra, διατηρώντας παράλληλα την υπεροχή της απόδοσής τους όσον αφορά την επίδοση BER του αποκωδικοποιημένου σήματος.

Τα αποτελέσματα της παρούσας παραγράφου αποκαλύπτουν ξεκάθαρα τις δυνατότητες εξισορρόπησης του μη γραμμικού σήματος από τα μοντέλα bi-RNN και τη χρήση τους σε σύμφωνα συστήματα οπτικών επικοινωνιών επόμενης γενιάς με την ιδιότητα των αποδοτικών και χαμηλής πολυπλοκότητας μετα-επεξεργασιών.

Η επόμενη ενότητα περιγράφει λεπτομερώς τη μοντελοποίηση του συστήματος.

4.3.2 Μοντελοποίηση συστήματος



Σχήμα 4.9: Η προσομοιωμένη ζεύξη διάδοσης

Το σύστημα οπτικής διάδοσης μεγάλων αποστάσεων, ως 1000km που φαίνεται στο σχήμα 4.9, προσομοιώνεται με την ολοκλήρωση της μη γραμμικής εξίσωσης Schrodinger (NLSE). Η οπτική διάδοση μοντελοποιείται με βάση τις εξισώσεις του Manakov με τη χρήση της μεθόδου Split Step Fourier, όπως έχει αναλυθεί στα προηγούμενα.

$$\frac{\partial E_{x,y}}{\partial z} = -\frac{\alpha}{2} E_{x,y} + \frac{j\beta_2}{2} \frac{\partial^2 E_{x,y}}{\partial t^2} - j\gamma \frac{8}{9} (|E_x|^2 + |E_y|^2) E_{x,y} \quad (4.11)$$

Θεωρούμε διάδοση 9 καναλιών πυκνής πολυπλεξίας διαίρεσης μήκους κύματος (DWDM). Η σχέση 4.11 τα $E_{x,y}$ περιέχει το συνολικό πεδίο των 9 κυμάτων ανά πόλωση που διαδίδονται παράλληλα, λαμβάνοντας με αυτό τον τρόπο υπόψη το σύνολο των διακαναλικών μη γραμμικών φαινομένων.

Υπολογίζουμε το ρυθμό σφάλματος BER σε όλα τα σενάρια που μελετήθηκαν και στοχεύουμε στην επίτευξη BER μικρότερη από την τιμή 3.8×10^{-3} όπως υπαγορεύεται από τον αλγόριθμο διόρθωσης σφαλμάτων hard decision (HD-FEC) το οποίο προσθέτει μια επιβάρυνση 7% [137].

Εξετάζουμε αρκετά διαφορετικά σενάρια, με το πρώτο να αποτελεί τη συνέχεια της μελέτης της προηγούμενης ενότητας, ήτοι : 25 Gbaud ρυθμός συμβόλων ανά πόλωση, 16-QAM διαμόρφωση, 50 GHz απόσταση καναλιών, ενώ το σήμα ενισχύεται κάθε 50 km με την εικόνα θορύβου ίση με 5 dB. Οι τιμές των παραμέτρων είναι αυτές του πίνακα 4.1.

Στα επόμενα σενάρια, πιο απαιτητικά και σύμφωνα με τα πρότυπα λειτουργίας των οπτικών συστημάτων αιχμής, προσομοιώνουμε τη λειτουργία αρχικά ενός συστήματος 16-QAM, με 32 Gbaud ανά πόλωση που αντιστοιχεί σε ρυθμό γραμμής 256 Gb/s (238 Gb/s καθαρός ρυθμός δεδομένων) ανά μήκος κύματος. Επεκτείνουμε την ανάλυση στη διαμόρφωση των 32-QAM, στα 25 Gbaud (ρυθμός γραμμής 250 Gb/s και καθαρός ρυθμός δεδομένων 232,5 Gb/s) και 16QAM στα 64 Gbaud (ρυθμός γραμμής 512 Gb/s, καθαρός ρυθμός δεδομένων 476 Gb/s), ώστε να υπάρχει ποικιλία σχημάτων διαμόρφωσης και ρυθμών δεδομένων. Ο καθαρός ρυθμός δεδομένων προκύπτει από την υπόθεση πως λαμβάνεται υπόψη μόνο η επιβάρυνση που θέτει το HD-FEC. Χρησιμοποιείται οπτική ενίσχυση με span των 50 ή των 80km και εικόνας θορύβου 5dB. Οι τιμές των παραμέτρων των τελευταίων σεναρίων παρουσιάζονται στον πίνακα 4.2

Σύμβολο	Παράμετρος	Τιμή
G	Απολαβή του ενισχυτή	10dB @ 50 km ή 16 dB @ 80 km
a	εξασθένιση	0.2dB/km
β_2	Διασπορά 2 ^{ης} τάξης	-21ps ² /km
γ	Μη γραμμικός όρος ίνας	1.3 W ⁻¹ km ⁻¹
R	Ρυθμός συμβόλων	25 Gbaud, 32 Gbaud, 64 Gbaud
M	Τύπος διαμόρφωσης	32-QAM ή 16-QAM με πολυπλεξία πόλωσης
L	Μήκος span	50 km ή 80 km
Δf	Απόσταση καναλιών	35 GHz @ 32 Gbaud ή 75 GHz @ 64 Gbaud

Πίνακας 4.2 : Παράμετροι αριθμητικών μοντέλων

Ο οπτικός δέκτης που απεικονίζεται στο σχήμα 4.9 αποτελείται από έναν οπτικό υβρίδιο (optical hybrid), ισορροπημένο δέκτη, ηλεκτρικά χαμηλοπερατά φίλτρα με συχνότητα αποκοπής προσαρμοσμένη στον ρυθμό των συμβόλων. Υποθέσαμε την ιδανική εκτίμηση

φάσης και συχνότητας φέροντος, καθώς και την ιδανική αποπολυπλεξία πόλωσης, καθώς θέλουμε να εστιάσουμε αποκλειστικά στις μη γραμμικές αλλοιώσεις και στον μετριάσμό τους. Πριν από οποιαδήποτε επεξεργασία ή αποδιαμόρφωση, δειγματοληπτούμε το σήμα και στη συνέχεια αντισταθμίζουμε τη χρωματική διασπορά με τη χρήση ενός ιδανικού εξισωτή FDE. Επίσης χρησιμοποιήθηκε προωθημένη διαμόρφωση παλμού και συγκεκριμένα ψηφιακού φίλτρου Root-Raised-Cosine (RRC), με συντελεστή προοδευτικής αποκοπής (roll-off) ίσο με 0.2.

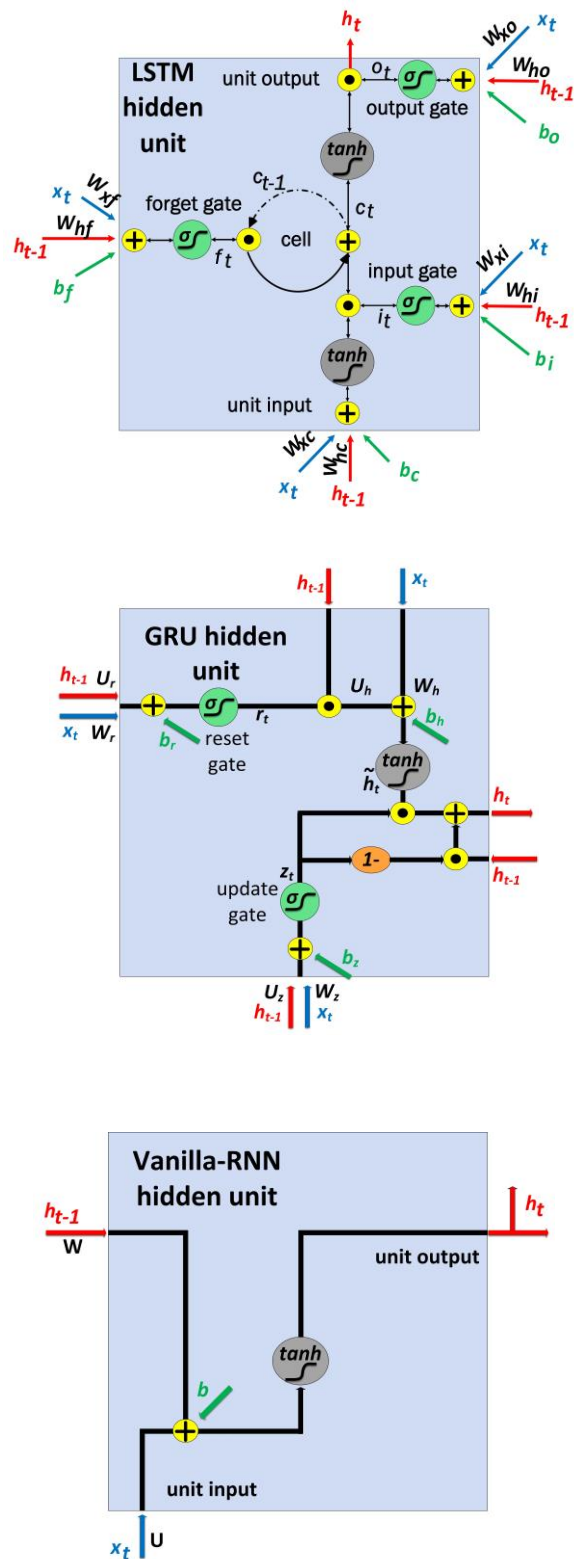
Διερευνούμε αριθμητικά την αποτελεσματικότητα τριών τύπων αμφίδρομων επαναλαμβανόμενων νευρωνικών δικτύων (bi-RNN) σε συστήματα αντιστάθμισης επικοινωνιών οπτικών ινών, και συγκεκριμένα τα LSTM, GRU και VRNN όπως ήδη αναφέρθηκε στην εισαγωγή. Το σχήμα 4.10 απεικονίζει τις μονάδες RNN που χρησιμοποιούμε ενώ η εξίσωση (4.12)-(4.14) υποδεικνύουν πώς υπολογίζεται η έξοδος \underline{h}_t σε κάθε περίπτωση :

$$\begin{aligned} i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\ f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\ o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \\ c_t &= f_t * c_{t-1} + i_t * \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\ h_t &= o_t * \tanh(c_t) \end{aligned} \quad (4.12, \text{LSTM})$$

$$\begin{aligned} z_t &= \sigma(W_z x_t + U_z h_{t-1} + b_z) \\ r_t &= \sigma(W_r x_t + U_r h_{t-1} + b_r) \\ \tilde{h} &= \tanh[W_h x_t + U_h (r_t * h_{t-1}) + b_h] \\ h_t &= (1 - z_t) * h_{t-1} + z_t * \tilde{h} \end{aligned} \quad (4.13, \text{GRU})$$

$$h_t = \tanh(W h_{t-1} + U x_t + b) \quad (4.14, \text{VRNN})$$

Όπου W , U οι πίνακες των βαρών των επιμέρους συνδέσεων: f , i , o , c , z και r , που αντιστοιχούν στις πύλες forget, input, output, cell state, update και reset. Τα x_t , h_t , h_{t-1} αποτελούν τα διανύσματα εισόδου, τρέχουσας κρυφής εξόδου και προηγούμενης κρυφής εξόδου αντίστοιχα, ενώ τα b αποτελούν τα διανύσματα bias.



Σχήμα 4.10: Κρυφή μονάδα (ή κύτταρο): α) LSTM, β) GRU, γ) VRNN

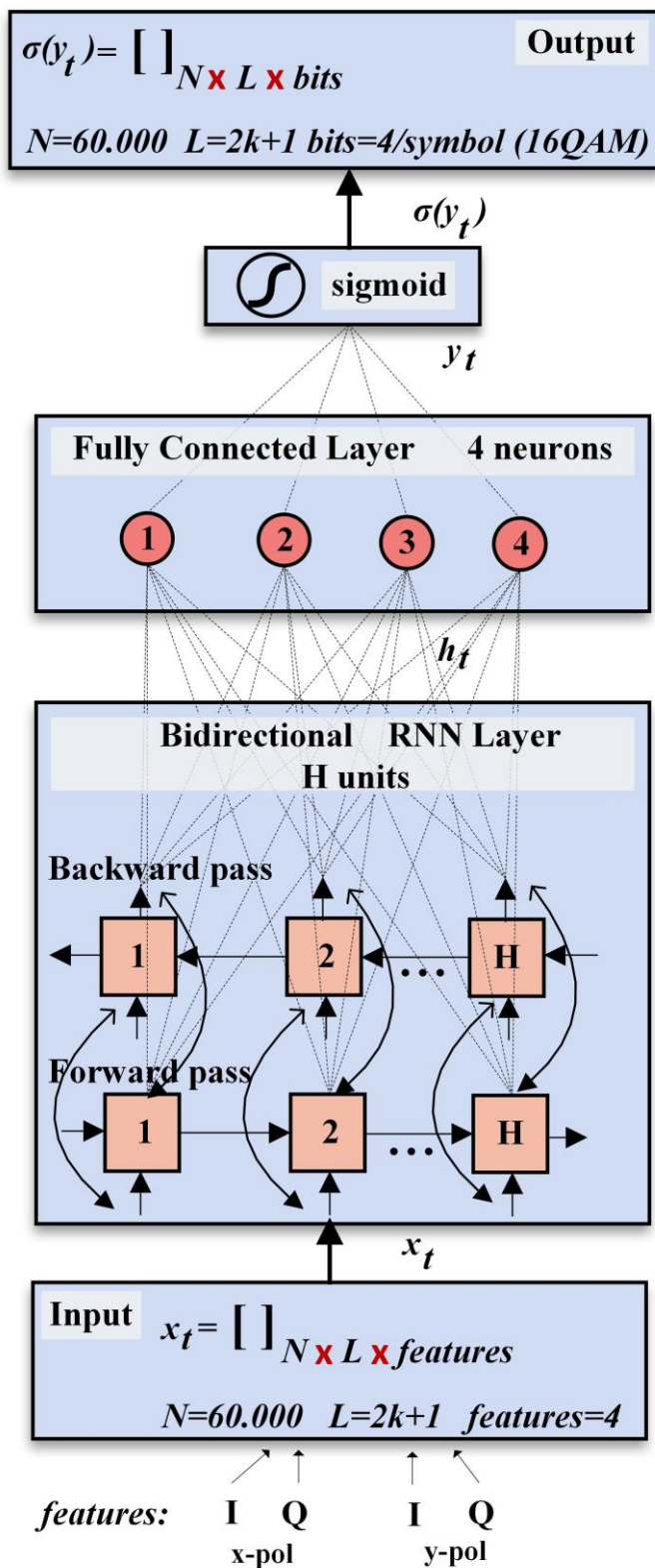
Το ακολουθιακό μοντέλο νευρωνικού δικτύου bi-RNN στην περίπτωση του σχήματος διαμόρφωσης 16-QAM απεικονίζεται στο σχήμα 4.11 ενώ αντίστοιχο, με τις επιμέρους τροποποιήσεις στον αριθμό των νευρώνων του FCL και της εξόδου, είναι αυτό για το 32-QAM.

Η είσοδος x_t είναι η αλλοιωμένη από τη διάδοση συμβολοσειρά και έχει την ακόλουθη μορφή $x_{t,m}=[x_{t-k}, \dots, x_{t-1}, x_t, x_{t+1}, \dots, x_{t+k}]$ όπου το L αντιπροσωπεύει το συνολικό μήκος της λέξης και ισούται με $L=2k+1$. Το μήκος L εξαρτάται από τη μνήμη καναλιού που σχετίζεται με τη συσσωρευμένη χρωματική διασπορά. Κάθε σύμβολο σε κάθε χρονικό παράθυρο περιέχει τέσσερις τιμές/χαρακτηριστικά (τα I και Q των δύο πολώσεων) οπότε οι είσοδοι X_{x-pol} και X_{y-pol} τροφοδοτούν το bi-RNN στρώμα των H κρυφών μονάδων.

Για να υπολογίσουμε την τιμή BER οδηγούμε την κρυφή έξοδο του bi-RNN h_t σε ένα στρώμα FC των 4 νευρώνων στην περίπτωση του 16-QAM ή 5 νευρώνων στην περίπτωση του 32-QAM και έπειτα την έξοδό του $y_{t,L}=[y_{t-k}, \dots, y_{t-1}, y_t, y_{t+1}, \dots, y_{t+k}]$ σε μία σιγμοειδή συνάρτηση που εκτελεί την εκτίμηση $\sigma(y_t)$ για όλα τα σύμβολα στην έξοδο. Η προσέγγιση αυτή που επικεντρώνεται στα επιμέρους bit των συμβόλων είναι ελαφρώς διαφορετική σε σύγκριση με την προσέγγιση της καθαρής ταξινόμησης των συμβόλων που χρησιμοποιήσαμε στην προηγούμενη μελέτη μας [1] και βελτιώνει οριακά την πολυπλοκότητα του μοντέλου χωρίς να θυσιάζει τις επιδόσεις BER [138].

Εκπαιδεύουμε το μοντέλο χρησιμοποιώντας πολλά σύμβολα εισόδου και πολλά σύμβολα εξόδου, (προσέγγιση πολλά σε πολλά) που προβλέπει ταυτόχρονα τον ίδιο αριθμό συμβόλων y_t με αυτά της εισόδου x_t . Τα μοντέλα RNN κατασκευάζονται, σχεδιάζονται και αξιολογούνται στο *Keras* [130] με τη βοήθεια της βιβλιοθήκης του *Tensorflow GPU 2.3* [131]. Ως συνάρτηση απώλειας επιλέγεται η *binary crossentropy* και ο *Adam* ως αλγόριθμος βελτιστοποίησης με τις τυπικές τιμές των παραμέτρων [106].

Χρησιμοποιούμε 40.000 σύμβολα για την εκπαίδευση, 20.000 για επικύρωση και 60.000



Σχήμα 4.11: Αρχιτεκτονική μοντέλου bi-RNN στην περίπτωση του 16-QAM

για αξιολόγηση με άγνωστα δεδομένα. Το στάδιο της εκπαίδευσης εκτελείται σε πακέτα των 512 λέξεων συμβόλων για βέλτιστη ισορροπία μεταξύ του μεγέθους δέσμησης μνήμης και του χρόνου εκτέλεσης. Ο μέγιστος αριθμός των εποχών είναι 300 ωστόσο χρησιμοποιούμε το μηχανισμό *early-stopping*, για να αποφύγουμε την υπερπροσαρμογή για το πολύ 20 εποχές διατηρώντας τα βάρη που επιτυγχάνουν τη μικρότερη τιμή της *binary cross-entropy*.

Ο *bi-RNN* εφαρμόζεται στο κεντρικό κανάλι *WDM* που είναι καταρχήν αυτό με τη μεγαλύτερη παραμόρφωση και απαιτεί μόνο ένα δείγμα ανά σύμβολο για τη μη-γραμμική επεξεργασία.

Ο μη γραμμικός ισοσταθμιστής *Volterra* τρίτης τάξης που χρησιμοποιήθηκε ως εναλλακτικός του *bi-RNN*, προσομοιώθηκε ως εξής:

Η εξισορροπημένη έξοδος $y(n)$ του σήματος δειγματοληψίας $x(n)$ από μια σειρά *Volterra* 3ης τάξης μπορεί να περιγραφεί σύμφωνα με τις εξισώσεις 4.15.

$$\begin{aligned}
 y(n) = & \sum_{i_1=-k_1}^{k_1} w_{i_1} x(n + i_1) + \sum_{i_1=-k_2}^{k_2} \sum_{i_2=i_1}^{k_2} w_{i_1, i_2} x(n + i_1) x(n + i_2) \\
 & + \sum_{i_1=-k_3}^{k_3} \sum_{i_2=i_1}^{k_3} \sum_{i_3=i_2}^{k_3} w_{i_1, i_2, i_3} x(n + i_1) x(n + i_2) x(n + i_3)
 \end{aligned}
 \tag{4.15}$$

όπου $k_j = (L_j - 1)/2$, στην οποία L_j είναι το μήκος της εισόδου της j -ής τάξης. w_{i_1, \dots, i_j} τα βάρη για κάθε δείγμα της j -ής τάξης [74]. Εκπαιδεύουμε το μοντέλο *Volterra* προσπαθώντας να ελαχιστοποιήσουμε το σφάλμα χρησιμοποιώντας τον αλγόριθμο του ελάχιστου μέσου τετραγωνικού σφάλματος (*Least Mean Square*). Θεωρούμε 60.000 σύμβολα για εκπαίδευση και 60.000 για δοκιμή με άγνωστα δεδομένα. Κάθε σύμβολο εισόδου είναι ένα διάνυσμα τεσσάρων χαρακτηριστικών (I και Q και για τις δύο πολώσεις).

Στη συνέχεια παρουσιάζονται τα αποτελέσματα της αριθμητικής μας ανάλυσης ξεκινώντας με τη σύγκριση μεταξύ των διαφορετικών μοντέλων *bi-RNN*.

4.3.2 Σύγκριση απόδοσης μοντέλων bi-RNN

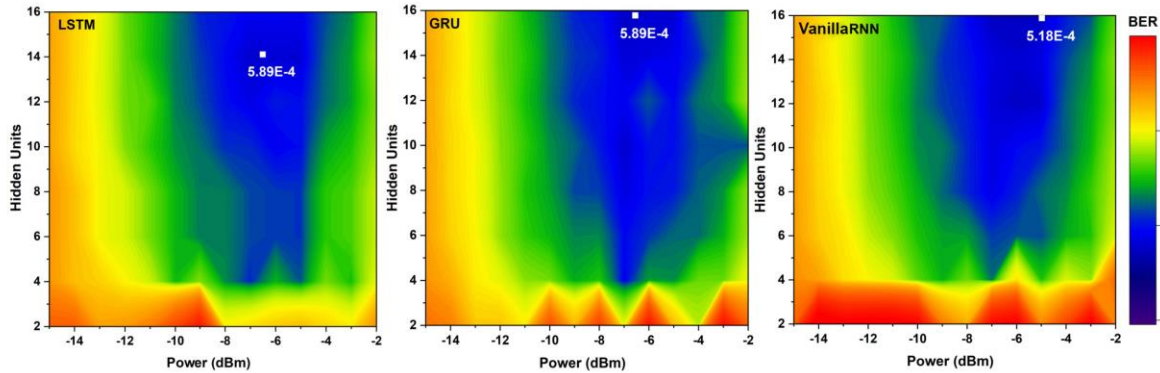
Αρχικά, συγκρίνουμε τα τρία μοντέλα bi-RNN ως προς τις επιδόσεις BER. Δοκιμάζουμε την ικανότητά τους να αναγνωρίζουν διαμορφωμένες σειρές συμβόλων 16-QAM των 25 ή των 32 Gbaud, που έχουν διαδοθεί κατά μήκος 1000 km (με 50km span) ή 960 km (με 80 km span) στα 1550nm (διασπορά $\beta_2 = -21$ ps²/km). Στην περίπτωση των 25 GBaud η απόσταση των καναλιών είναι 50 GHz και δεν έχει εφαρμοστεί διαμόρφωση παλμού ενώ στην περίπτωση των 32 GBaud η απόσταση των καναλιών είναι 35GHz και έχει εφαρμοστεί διαμόρφωση παλμού RRC με συντελεστή roll-off $\alpha = 0.2$. Οι τιμές των υπολοίπων παραμέτρων διάδοσης είναι αυτές του πίνακα 4.2.

Εκπαιδεύουμε τα διαφορετικά bi-RNN με μια λέξη-σύμβολο 201 ή 151 συμβόλων που προσεγγίζει τη μνήμη καναλιού ανάλογα με τη συσσωρευμένη διασπορά. Στα σχ. 4.12 και 4.13 μπορεί να φανεί ότι όλα τα μοντέλα αποδίδουν επαρκώς και ισοδύναμα καθώς βελτιώνουν σημαντικά την επίδοση BER σε σύγκριση με ένα σύστημα διάδοσης που χρησιμοποιεί μόνο γραμμική εξισορρόπηση (FDE).

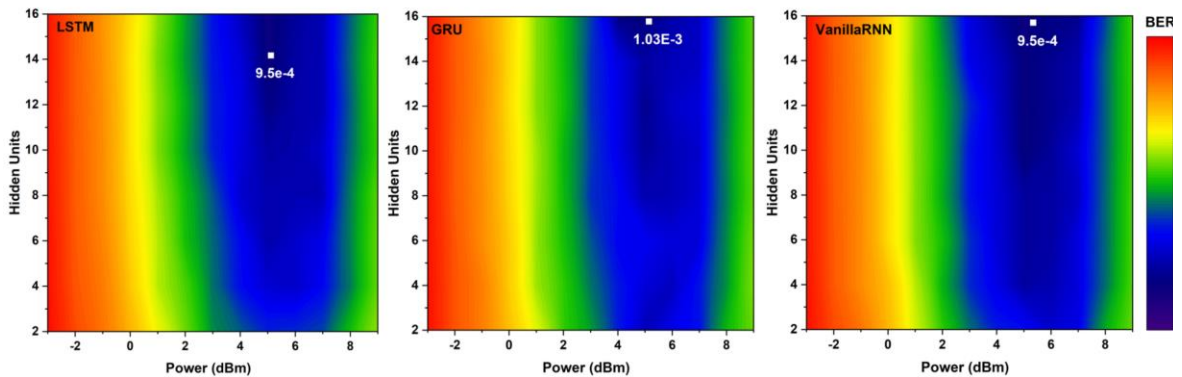
Συγκεκριμένα στην περίπτωση των 25 Gbaud παρατηρούμε βελτίωση του BER κατά σχεδόν μία τάξη μεγέθους, περίπου 5×10^{-4} για το VRNN έναντι περίπου 4×10^{-3} για το FDE. Αντίστοιχα στην περίπτωση των 32 Gbaud έχουμε βελτίωση του BER κατά σχεδόν μισή τάξη μεγέθους, περίπου 1×10^{-3} για το VRNN έναντι περίπου 4×10^{-3} για το FDE.

Μπορεί να φανεί ότι το bi-LSTM προσφέρει την καλύτερη απόδοση για τουλάχιστον 14 μονάδες, ενώ το bi-GRU και το bi-VRNN για 16 μονάδες. Είναι ενδιαφέρον ότι όλα τα μοντέλα λειτουργούν ισοδύναμα, επομένως δεν υπάρχει ουσιαστική ανάγκη χρήσης απαιτητικών και πολύπλοκων μοντέλων bi-RNN που βασίζονται σε μονάδες LSTM ή GRU. Το bi-VRNN με σχεδόν την ίδια χωρητικότητα, όπως υπαγορεύεται από τον αριθμό των κρυφών μονάδων, αλλά σημαντικά απλούστερη αρχιτεκτονική μπορεί να παρουσιάσει παρόμοια απόδοση με βάση τα ευρήματα των σχημάτων 4.12 και 4.13.

Όπως αναλύθηκε στο κεφάλαιο 3 τα μοντέλα LSTM χρησιμοποιούν την input gate i_t , output gate o_t και την forget gate f_t και τα μοντέλα GRU χρησιμοποιούν update gate z_t και τη reset gate r_t προκειμένου να αποτρέψουν ταυτόχρονα το πρόβλημα της



Σχήμα 4.12: Το BER ως συνάρτηση της οπτικής ισχύος και του αριθμού των κρυφών μονάδων για το αμφίδρομο LSTM (α), το GRU(β) και το VRNN (γ) Σε 1000km οπτική μετάδοση με διασπορά -21 ps²/km, 25GBaud χωρίς διαμόρφωση παλμού



Σχήμα 4.13: Το BER ως συνάρτηση της οπτικής ισχύος και του αριθμού των κρυφών μονάδων για το αμφίδρομο LSTM (α), το GRU(β) και το VRNN (γ) Σε 960km οπτική μετάδοση με διασπορά -21 ps²/km, 32GBaud με RRC διαμόρφωση παλμού

εξαφανιζόμενης κλίσης (vanishing gradient) [116] και να χειριστούν μακροπρόθεσμες ή βραχυπρόθεσμες εξαρτήσεις [139], δηλαδή καθυστερήσεις άγνωστης διάρκειας μεταξύ σημαντικών γεγονότων σε μια χρονολογική σειρά. Η αλληλοπαρεμβολή συμβόλων ISI που προκαλείται κατά τη διάδοση οπτικών σημάτων στην ίνα, ως αποτέλεσμα της συσσωρευμένης διασποράς, καθορίζει τη μνήμη του καναλιού, η οποία παραμένει

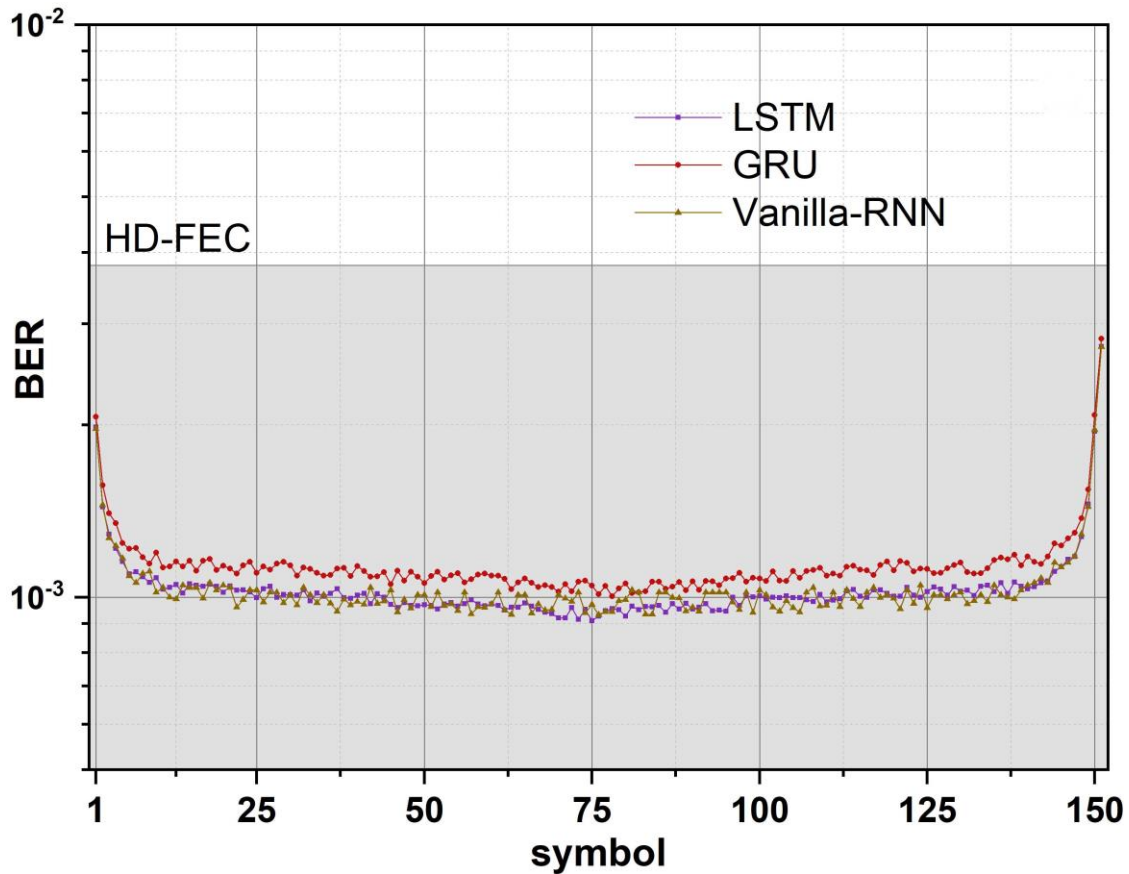
σταθερή όσο οι παράμετροι διάδοσης παραμένουν αμετάβλητες. Μόλις το δίκτυο μάθει τη μνήμη των καναλιών, αποκτά αντίληψη της θέσης των συμβόλων στην ακολουθία και δεν υπάρχουν απροσδόκητες ή χρονικά μεταβαλλόμενες μακροπρόθεσμες ή βραχυπρόθεσμες εξαρτήσεις, όπως για παράδειγμα υπάρχουν στα ταχέως μεταβαλλόμενα χρονικά κανάλια των ασύρματων επικοινωνιών.

Επομένως, ο πολύπλοκος προαναφερόμενος μηχανισμός των LSTM και GRU καθίσταται περιττός. Με βάση το προαναφερθέν εύρημα, η εκπαίδευση γίνεται μια διαδικασία που μπορεί εύκολα να πραγματοποιηθεί από ένα μοντέλο bi-VRNN, υπό την προϋπόθεση ότι το μήκος της λέξης συμβόλου είναι αρκετά μεγάλο ώστε να περιλαμβάνει τη μνήμη καναλιού και τη μη γραμμική ISI επίδραση.

Όσον αφορά την αποτελεσματικότητα της εκπαίδευσης, σε όλες τις προσομοιώσεις που πραγματοποιήσαμε, παρατηρήσαμε ότι όλα τα RNN συγκλίνουν λίγο πολύ στην ίδια τιμή απώλειας, όμως το VRNN χρειάζεται περίπου 30% περισσότερες εποχές από το LSTM και το GRU, με το τελευταίο να παρουσιάζει την ταχύτερη εκπαίδευση. Λαμβάνοντας υπόψη ότι το οπτικό κανάλι δεν απαιτεί συχνή εκπαίδευση, η ταχύτητα εκπαίδευσης δεν είναι κρίσιμη για την υιοθέτηση ή την απόρριψη ενός συγκεκριμένου μοντέλου RNN. Όπως έχει ήδη δειχθεί στο [1], τα μοντέλα RNN είναι ανθεκτικά σε διακυμάνσεις ισχύος ή ακόμα και στο σχήμα διαμόρφωσης γειτονικών καναλιών. Ως εκ τούτου, η πιθανή περίοδος επανεκπαίδευσης σχετίζεται με τις χρονικές κλίμακες των φαινομένων της πόλωσης, δηλαδή μερικά ms που αντιστοιχεί στη διάρκεια εκατοντάδων εκατομμυρίων συμβόλων. Δεδομένου ότι το σύστημα θα έχει αρχικά εκπαιδευτεί, η επανεκπαίδευση θα χρειαστεί λιγότερα σύμβολα (μερικές χιλιάδες) και λιγότερο χρόνο, καθώς ο σκοπός του θα είναι η απλή ρύθμιση και όχι η εκπαίδευση από την αρχή. Αυτό μπορεί να επιτευχθεί ακόμη και με τη χρήση ενός παράλληλου συστήματος παρακολούθησης που θα τροφοδοτεί την εκπαιδευτική μονάδα με ένα μικρό μέρος (< 1%) των δεδομένων που μεταδίδονται σε πραγματικό χρόνο, ώστε να ρυθμίζονται συνεχώς τα βάρη και να διατηρείται το σύστημα σταθερό.

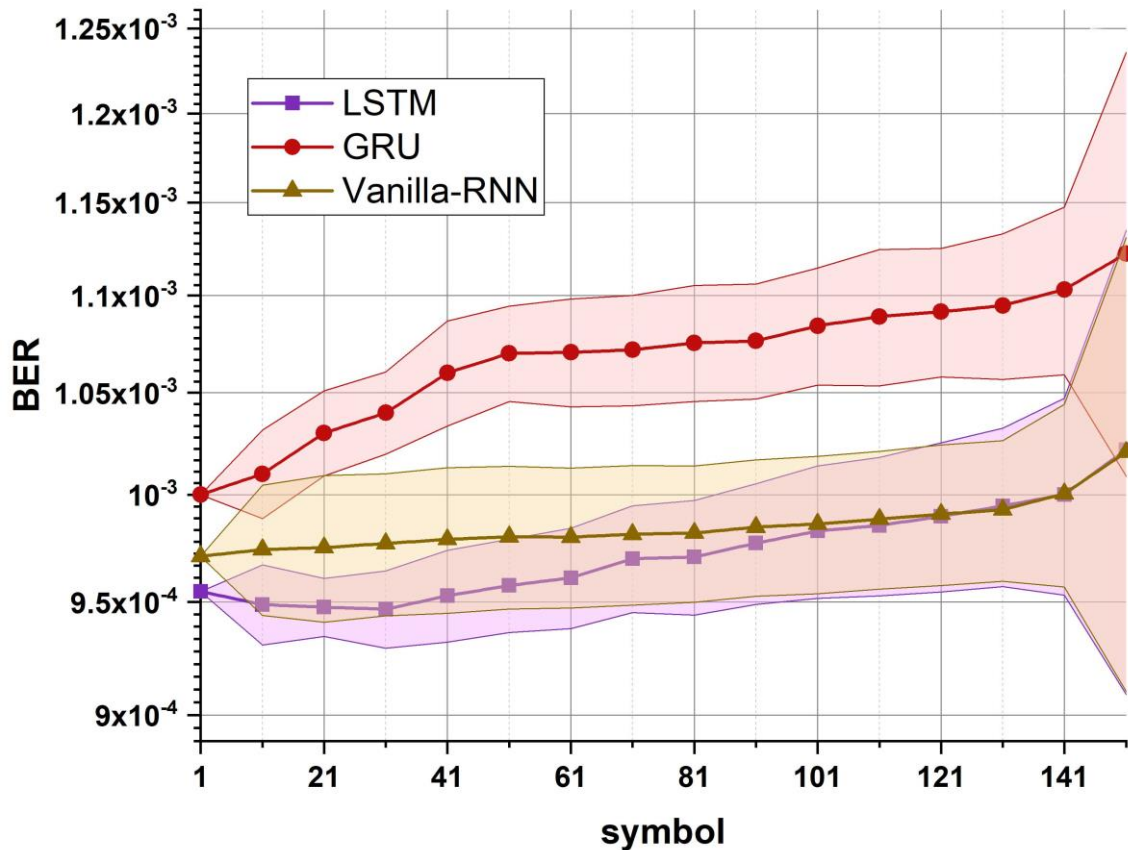
Πρέπει να αναφερθεί ότι η τιμή BER που εμφανίζεται στα σχήματα 4.12 και 4.13 είναι αυτή που λαμβάνεται για το κεντρικό σύμβολο του y_t . Δεδομένου ότι τα μοντέλα RNN μας

εκπαιδεύονται χρησιμοποιώντας την προσέγγιση πολλά σε πολλά, μπορεί κανείς να εξάγει ταυτόχρονα όλα τα σύμβολα στη λέξη.



Σχήμα 4.14: Η τιμή του BER για καθένα από τα 151 σύμβολα της λέξης, για όλα τα μοντέλα RNN. 120 σύμβολα έχουν περίπου την ίση ελάχιστη τιμή BER, ενώ συνολικά 100 σύμβολα εμφανίζουν $BER < 10^{-3}$.

Το σχήμα. 4.14 δείχνει την τιμή BER για όλα τα σύμβολα στο y_i . Γίνεται προφανές ότι τουλάχιστον 100 σύμβολα απολαμβάνουν περίπου την ελάχιστη τιμή BER ($\sim 1 \times 10^{-3} \pm 0,5 \times 10^{-4}$) που έχει και το κεντρικό σύμβολο. Όταν πλησιάζουμε στα άκρα της λέξης, η απόδοση σταδιακά υποβαθμίζεται, γεγονός που επιβεβαιώνει την ανάγκη για εκπαίδευση με αρκετά γειτονικά προηγούμενα και επόμενα σύμβολα, έτσι ώστε να αποτυπώνονται επαρκώς τα μη γραμμικά φαινόμενα παρεμβολής μεταξύ συμβόλων. Στο αριστερό άκρο, τα σύμβολα του παρελθόντος δεν είναι αρκετά, ενώ στη δεξιά άκρη, τα μελλοντικά σύμβολα είναι αυτά που λείπουν.



Σχήμα 4.15: Ο μέσος όρος του BER και η τυπική του απόκλιση υπολογίζεται σε μια υποομάδα συμβόλων στη λέξη των 151 συμβόλων. Η τυπική απόκλιση εκρήγνυται όταν ανιχνεύονται περισσότερα από 140 σύμβολα ταυτόχρονα

Το σχήμα 4.15 δείχνει με μεγαλύτερη έμφαση αυτή τη συμπεριφορά με τη μορφή του μέσου BER και την τυπική του απόκλιση όταν αυτό υπολογίζεται για ένα τμήμα των συμβόλων της λέξης (151 στην περίπτωση μας). Αυτές οι υποομάδες συμβόλων σχηματίζονται, ξεκινώντας από το κεντρικό σύμβολο της λέξης και συμπεριλαμβάνοντας τον ίδιο αριθμό προηγούμενων και επόμενων συμβόλων. Για παράδειγμα, όταν η υποομάδα συμβόλων έχει μήκος 21, περιέχει το κεντρικό σύμβολο, 10 σύμβολα που προηγούνται και 10 που έπονται. Με βάση το σχήμα 4.15 γίνεται προφανές ότι ο μέσος

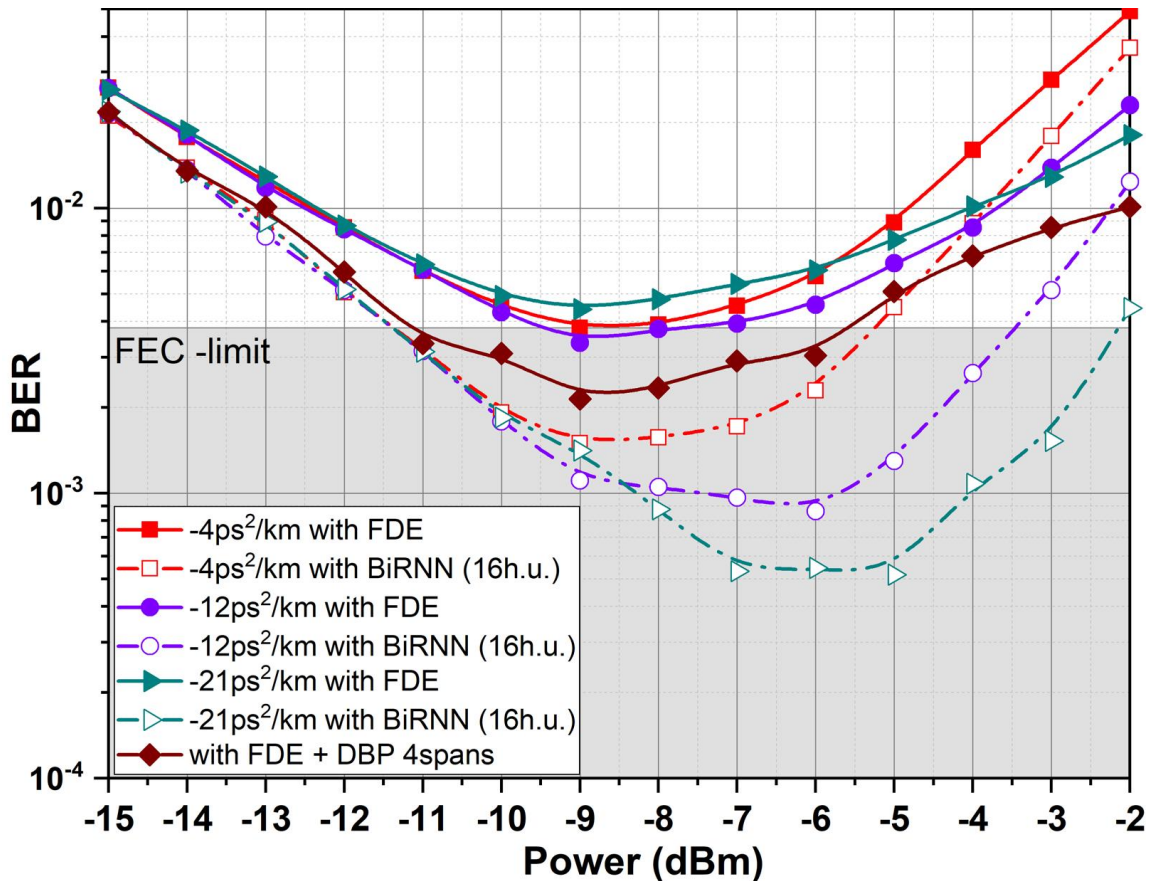
όρος της τιμής BER παραμένει σταθερός ακόμη και αν 120 από τα 150 σύμβολα ανιχνεύονται ταυτόχρονα και η τυπική του απόκλιση εκρήγνυται όταν οι υποομάδες περιέχουν περισσότερα από 140 σύμβολα. Έτσι, μπορεί κανείς εύκολα να συμπεράνει ότι η ταυτόχρονη ανίχνευση 100 συμβόλων σε μια λέξη των 151 συμβόλων είναι εφικτή και αποτελεσματική αφού δεν θα προκαλέσει σοβαρή υποβάθμιση του BER σε σχέση με την περίπτωση που ανιχνεύεται μόνο το κεντρικό σύμβολο.

4.3.3 Σύγκριση ισοσταθμιστών βασισμένων σε bi-RNN και Volterra

Ως τώρα έχουμε δείξει πως το bi-VRNN είναι εξίσου αποτελεσματικό με τα εναλλακτικά και πιο πολύπλοκα σχήματα των LSTM και RNN. Έπειτα δείξαμε τα οφέλη της προσέγγιση πολλά σε πολλά, τη ταυτόχρονη δηλαδή ανίχνευση και αξιολόγηση πολλών -περισσότερων του ενός- συμβόλων, που έχουν προκύψει από την εκπαίδευση πολλών επίσης συμβόλων.

Στη συνέχεια, προκειμένου να μελετήσουμε την αποτελεσματικότητα του bi-VRNN για διαφορετικά σενάρια μνήμης καναλιών, πραγματοποιούμε αριθμητικές προσομοιώσεις τιμές διασποράς δεύτερης τάξης και συγκεκριμένα για $\beta_2 = -4, -12$ και -21 ps²/km. Αν και σε πραγματικά συστήματα, όταν συναντάμε τιμές διασποράς διαφορετικές από αυτή ενός τυπικού SMF (~-21 ps²/km) συνεπάγεται πως έχει χρησιμοποιηθεί μονάδα αντιστάθμισης διασποράς ή άλλου τύπου οπτική ίνα που συνακόλουθα διαφοροποιούν άλλες κρίσιμες παραμέτρους του συστήματος, εδώ υποθέτουμε ότι όλες οι υπόλοιπες παράμετροι παραμένουν ανεπηρέαστες, προκειμένου να αποσαφηνιστεί πώς συμπεριφέρεται ο RNN ισοσταθμιστής για διαφορετικές μνήμες καναλιών. Υποθέτουμε επίσης πως ο λόγος σήματος προς θόρυβο και η μη γραμμικότητα διατηρούνται σταθερές. Το μήκος της λέξης έχει προσαρμοστεί έτσι ώστε να καλύπτει αποτελεσματικά τη μνήμη που ισοδυναμεί στην υψηλότερη διασπορά των -21ps²/km.

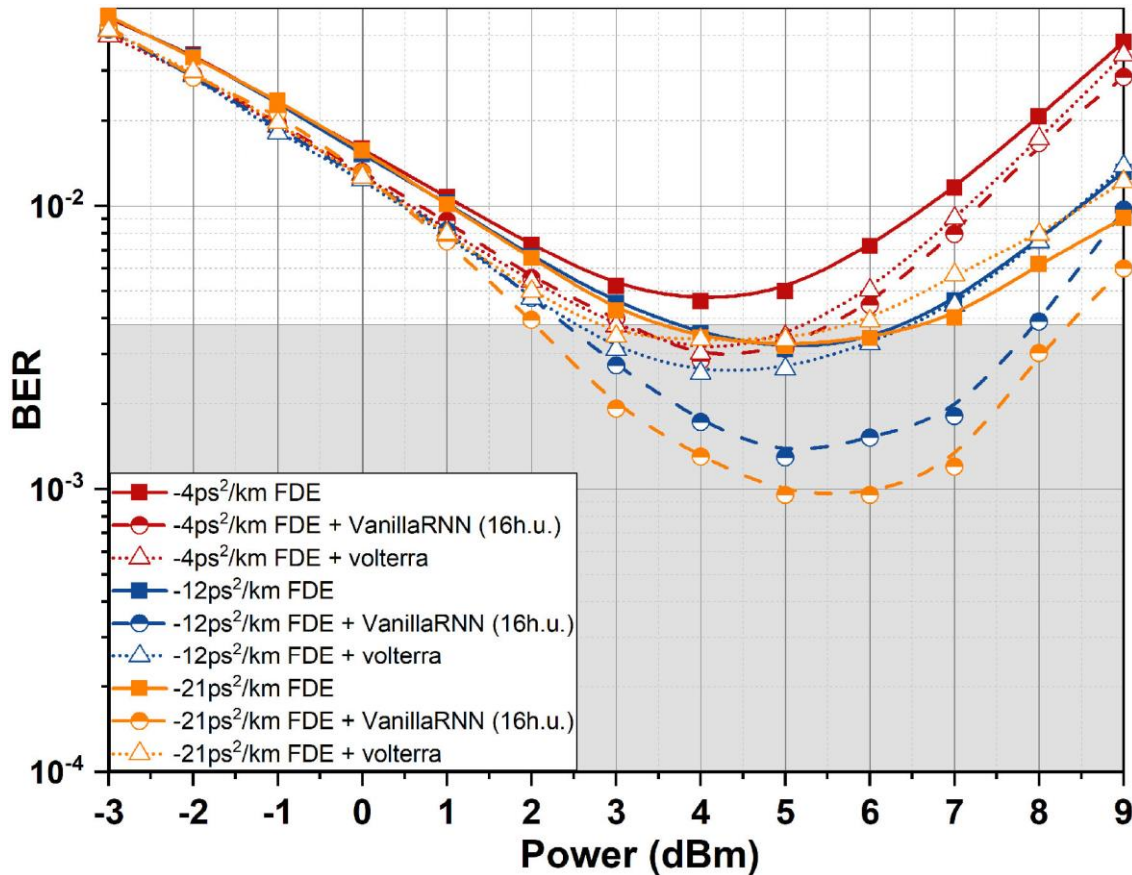
Στο σχήμα 4.16, απεικονίζεται η συμπεριφορά του συστήματος των 25 Gbaud με διαμόρφωση 16-QAM, χωρίς διαμόρφωση παλμού. Μπορεί κανείς να δει ότι τα συστήματα επεξεργασίας στον δέκτη που βασίζονται σε FDE παρουσιάζουν σχεδόν



Σχήμα 4.16: Η τιμή BER ως συνάρτηση της οπτικής ισχύος για διασπορά των -4, -12 και -21 ps²/km, με γραμμική εξίσωση (FDE) και με ένα Bi-VRNN των 16 κρυφών μονάδων για την περίπτωση των 25 GBaud, 16-QAM

πανομοιότυπη συμπεριφορά με την ελάχιστη τιμή BER να βρίσκεται κοντά στο 4×10^{-3} παρά τις διαφορές στη συσσωρευμένη χρωματική διασπορά. Αντίθετα, η αντιστάθμιση με το bi-VRNN εμφανίζει καλύτερη επίδοση BER σε μεγαλύτερη συσσωρευμένη διασπορά. Αυτή η συμπεριφορά επαληθεύτηκε και για το bi-GRU και το bi-LSTM και σχετίζεται με τον χρόνο συμφωνίας του καναλιού που είναι πολύ μεγαλύτερος από την περίοδο συμβόλων όσο η διασπορά αυξάνεται. Ως εκ τούτου, τα διακαναλικά φαινόμενα γίνονται πολύ αργά και παρακολουθούνται εύκολα από τον ισοσταθμιστή bi-RNN [133].

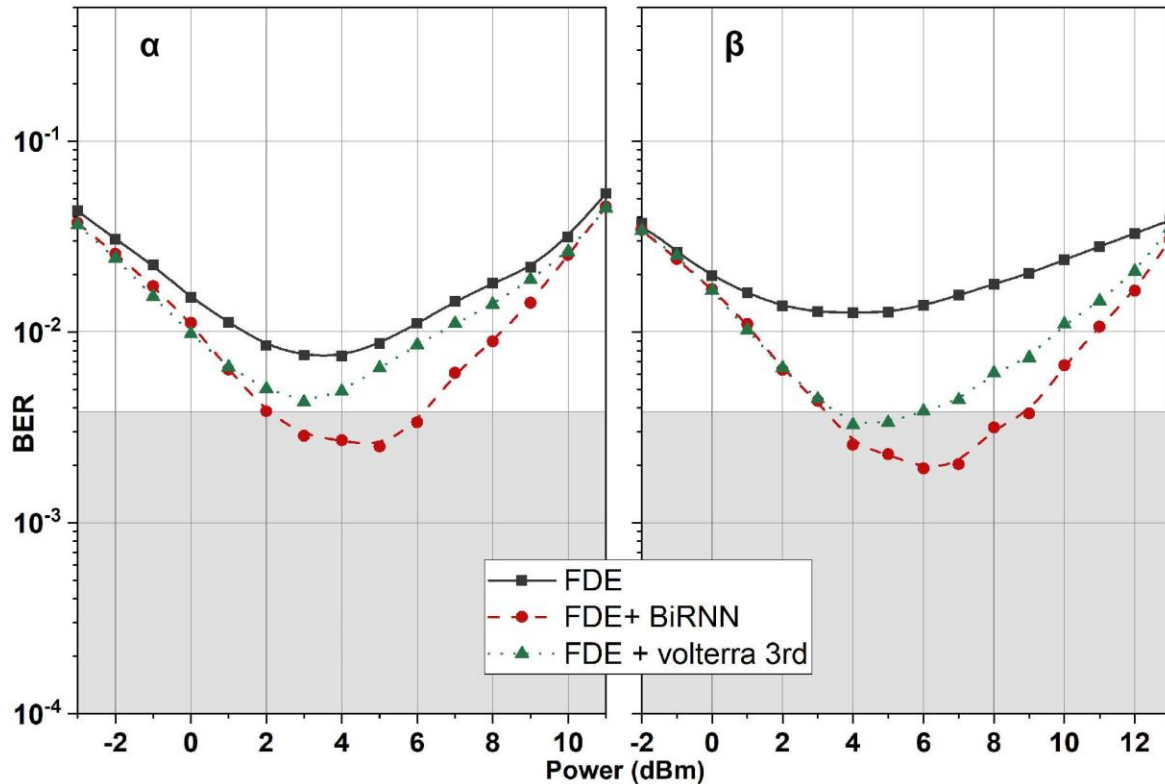
Το bi-RNN αποδίδει καλύτερα από το DBP καθώς το τελευταίο ισοσταθμίζει αποκλειστικά τις ενδοκαναλικές επιδράσεις και αγνοεί τις διακαναλικές όπως έχουμε εξηγήσει [1].



Σχήμα 4.17: Οι επιδόσεις BER ως συνάρτηση της οπτικής ισχύος για διασπορά των -4 , -12 και $-21 \text{ ps}^2/\text{km}$, με γραμμική εξίσωση (FDE) και με ένα Bi-VRNN των 16 κρυφών μονάδων ή Volterra 3^{ης} τάξης, για την περίπτωση των 32 GBaud, 16-QAM

Εκτός από τη γραμμική εξισορρόπηση με τη βοήθεια του FDE, πραγματοποιήσαμε επίσης αριθμητικές προσομοιώσεις χρησιμοποιώντας τον μη γραμμικό ισοσταθμιστή Volterra. Για την ισότιμη σύγκριση των δύο ισοσταθμιστών χρησιμοποιούμε το ίδιο μήκος μνήμης $L=151$ για το RNN και $L_1=151$ για την 1η τάξη του Volterra (γραμμικός όρος). Όσον αφορά τη Volterra, εξετάσαμε το σενάριο των τιμών $L_2=51$ και $L_3=11$ για τους πυρήνες 2ης και 3ης τάξης, δύο τιμές που επαρκούν για την εξίσωση των μη γραμμικών φαινομένων εντός του καναλιού και προσφέρουν την καλύτερη απόδοση με τη μικρότερη δυνατή πολυπλοκότητα, όπως αποδεικνύεται από εκτενείς προσομοιώσεις.

Στην περίπτωση όπου ο ρυθμός των διαμορφωμένων με 16-QAM συμβόλων είναι στα 32 GBaud, μπορεί κανείς να δει στο σχήμα 4.17 ότι τα συστήματα διάδοσης με αντιστάθμιση FDE εμφανίζουν επίδοση BER που ποικίλλει από 4×10^{-3} έως 2×10^{-3} με



Σχήμα 4.18: Το BER ως συνάρτηση της οπτικής ισχύος, α) 25 GBaud 32-QAM (960 km, απόσταση ενισχυτών 80 km) (β) 64 GBaud, 16-QAM (640 km, απόσταση ενισχυτών 80 km).

καλύτερες τιμές να εμφανίζονται στις υψηλότερες συσσωρευμένες τιμές διασποράς λόγω του γεγονότος ότι τα μη γραμμικά φαινόμενα μεταξύ των καναλιών γίνονται λιγότερο έντονα καθώς αυξάνεται η διασπορά. Η αντιστάθμιση bi-VRNN εμφανίζει σημαντικά καλύτερη επίδοση BER σε μεγαλύτερη συσσωρευμένη διασπορά. Όπως πριν, πανομοιότυπη συμπεριφορά BER επαληθεύτηκε και για τις μονάδες GRU και LSTM. Ο ισοσταθμιστής Volterra βελτιώνει ελαφρώς το BER σε σύγκριση με τη γραμμική ισοστάθμιση, ωστόσο η απόδοσή του δεν βελτιώνεται σε μεγαλύτερες τιμές διασποράς. Το VRNN αποδίδει πολύ καλύτερα από τον ισοσταθμιστή Volterra, καθώς το πρώτο φαίνεται να παρακολουθεί επαρκώς τα διακαναλικά φαινόμενα, ενώ το δεύτερο διακρίνει και επιλύει μόνο τα ενδοκαναλικά [73].

Στη συνέχεια δοκιμάζουμε τη συμπεριφορά του bi-VRNN σε δυσκολότερες συνθήκες, είτε αυξάνοντας την τάξη της διαμόρφωσης είτε τον ρυθμό των συμβόλων. Δουλεύοντας με

τυπικές μονότροπες ίνες στη ζώνη C (1550nm, $\beta_2 = -21$ ps²/km διασπορά) αυξήσαμε την τάξη της διαμόρφωσης σε 32-QAM και μειώσαμε τον ρυθμό των συμβόλων στα 25 Gbaud, με 12 τμήματα των 80km, συνολικού μήκους διάδοσης 960km. Στο σχήμα 4.18α, παρατηρούμε ότι μόνο ο ισοσταθμιστής bi-VRNN μπορεί να αποδώσει κάτω από το όριο FEC των 3.8×10^{-3} , δείχνοντας έτσι τη συστηματική υπεροχή του και σε πιο σύνθετα σχήματα διαμόρφωσης. Αναφέρεται ότι τα μοντέλα GRU και LSTM σε αυτό το σκληρό περιβάλλον δεν εμφάνισαν καλύτερη επίδοση BER από το VRNN, και έτσι αποδείχθηκε για άλλη μια φορά η ισοδυναμία τους.

Προκειμένου να ευθυγραμμίσουμε τη μελέτη μας με τα τρέχοντα συστήματα μετάδοσης αιχμής, αυξήσαμε τον ρυθμό συμβόλων ανά κανάλι στα 64 GBaud με απόσταση καναλιών 75GHz, χρησιμοποιώντας 8 spans των 80 km για συνολική διάδοση 640 km (σχήμα 4.18β). Τα αποτελέσματα της προσομοίωσης, ακόμη και για αυξημένο baudrate ανά κανάλι (~ 500 Gb/s χωρητικότητα ανά μήκος κύματος) επαληθεύουν ότι ο προτεινόμενος ισοσταθμιστής bi-VRNN είναι ικανός να μετριάσει τις μη γραμμικότητες στα κανάλια διασποράς και να δείξει για άλλη μια φορά την υπεροχή του έναντι των μη γραμμικών ισοσταθμιστών Volterra.

4.3.4 Ανάλυση υπολογιστικής πολυπλοκότητας

Στο τελευταίο μέρος αυτής της μελέτης διερευνούμε την πολυπλοκότητα των μοντέλων bi-RNN και Volterra. Η πολυπλοκότητα του φίλτρου FDE έχει αναλυθεί επαρκώς στο προηγούμενο εδάφιο. Γενικά, η συνολική πολυπλοκότητα του ισοσταθμιστή bi-RNN εξαρτάται από τον αριθμό των παραμέτρων (βαρών) που πρέπει να υπολογίσει κάθε δίκτυο, από τον αριθμό των κρυφών μονάδων και από το μήκος της λέξης εισαγωγής.

Σύμφωνα με τα σχήματα 4.12 και 4.13 το LSTM χρειάζεται τουλάχιστον 14 μονάδες ενώ το GRU και το VRNN χρειάζονται 16 μονάδες για να επιτύχουν το βέλτιστο BER, επομένως τα τρία μοντέλα είναι λίγο-πολύ ισοδύναμα. Όσον αφορά το βέλτιστο μήκος λέξης, αυτό δεν εξαρτάται από την επιλεγμένη αρχιτεκτονική μοντέλου, σχετίζεται αυστηρά με τη μνήμη καναλιών. Η μόνη παράμετρος που διαφοροποιεί τα τρία μοντέλα

είναι η πολυπλοκότητα κάθε μονάδας. Με βάση τις σχέσεις 4.12-4.14, μπορεί κανείς εύκολα να υπολογίσει τον αριθμό των παραμέτρων για κάθε μοντέλο RNN ως εξής:

$$bi - RNN_{param} = 2B[H(H + F) + H] + (2H + 1)b \quad (4.16)$$

όπου $B=4,3,1$ για το LSTM, το GRU και το VRNN αντίστοιχα, υποδεικνύοντας τον αριθμό των πυλών που περιέχονται σε κάθε μονάδα RNN (σχήμα. 4.10), H ο αριθμός των κρυφών μονάδων, $F=4$ ο αριθμός των χαρακτηριστικών εισόδου (σχήμα 4.11) και b ο αριθμός των bits ($b=4$ στην περίπτωση του 16-QAM). Τέλος ο πολλαπλασιαστής 2 του πρώτου όρου της 4.16 υποδηλώνει τη λειτουργία της αμφίδρομης αρχιτεκτονικής.

Προκειμένου να συγκριθούν απευθείας τα bi-RNN με τον μη γραμμικό ισοσταθμιστή Volterra, είναι απαραίτητο να υπολογιστεί η υπολογιστική πολυπλοκότητα ως προς τον αριθμό των πραγματικών πολλαπλασιασμών. Για τα bi-RNN μοντέλα μπορούμε να υπολογίσουμε τον αριθμό των πολλαπλασιασμών, με τη χρήση της σχέσης 4.17. Σημειώνουμε ότι η τιμή του bias είναι μηδενική για όλα τα μοντέλα bi-RNN, επομένως δεν θα πρέπει να λαμβάνεται υπόψη στην εκτίμηση της πολυπλοκότητας.

$$bi - RNN_{mult} = 2B(FH + H^2)L + 2HbL \quad (4.17)$$

όπου L το μήκος της λέξης των συμβόλων.

Η υπολογιστική πολυπλοκότητα ως προς τον αριθμό των πραγματικών πολλαπλασιασμών του μη γραμμικού ισοσταθμιστή Volterra που περιγράφεται από τη σχέση 4.15, δίνεται από την σχέση :

$$CC_{Volterra}(K, L) = \sum_{k=1}^K \frac{(L_k - 1 + k)!}{(k - 1)! (L_k - 1)!} \quad (4.18)$$

Είναι εμφανές από τη σχέση 4.18 πως ο αριθμός των πολλαπλασιασμών εκρήγνυται καθώς αυξάνεται η τάξη των σειρών Volterra. Αυτός είναι ο σημαντικότερος λόγος που επιλέγουμε τις τιμές L_2 και L_3 μικρότερες του L_1 [74]. Ο δεύτερος λόγος είναι πως πρακτικά οι μη γραμμικότητες που θα μπορούσαν να προσομοιωθούν από τη δεύτερη τάξη των σειρών Volterra, έχουν σημαντικά μικρότερη έκταση από τα γραμμικά φαινόμενα που

A. ΥΠΟΛΟΓΙΣΤΙΚΗ ΠΟΛΥΠΛΟΚΤΟΤΗ ΩΣ ΣΥΝΑΡΤΗΣΗ ΤΟΥ ΑΡΙΘΜΟΥ ΤΩΝ ΠΟΛΛΑΠΛΑΣΙΑΣΜΩΝ
– ΜΟΝΟ ΚΕΝΤΡΙΚΟ ΣΥΜΒΟΛΟ ΣΤΗΝ ΕΞΟΔΟ (16QAM)

Ισοσταθμιστή ς	Μήκος μνήμης (σύμβολα)	Κρυφές μονάδες	Πολλαπλασιασμοί
<i>Bi-LSTM</i>	151	16	386688
<i>Bi-GRU</i>	151	16	290048
<i>Bi-VRNN</i>	151	16	96768
Volterra	151/51/11	-	14644

B. ΥΠΟΛΟΓΙΣΤΙΚΗ ΠΟΛΥΠΛΟΚΤΟΤΗ ΩΣ ΣΥΝΑΡΤΗΣΗ ΤΟΥ ΑΡΙΘΜΟΥ ΤΩΝ ΠΟΛΛΑΠΛΑΣΙΑΣΜΩΝ
– ΠΟΛΛΑ ΣΥΜΒΟΛΑ ΣΤΗΝ ΕΞΟΔΟ (16QAM)

Ισοσταθμιστής	Μήκος μνήμης (σύμβολα)	Κρυφές μονάδες	Σύμβολα αξιολόγησης	Πολλαπλ ασιασμοί ανά σύμβολο
<i>Bi-LSTM</i>	151	16	80	5074
<i>Bi-LSTM</i>	151	16	120	3382
<i>Bi-GRU</i>	151	16	80	3866
<i>Bi-GRU</i>	151	16	120	2577
<i>Bi-VRNN</i>	151	16	80	1450
<i>Bi-VRNN</i>	151	16	120	966
<i>Volterra</i>	151/51/11	-	1	14644

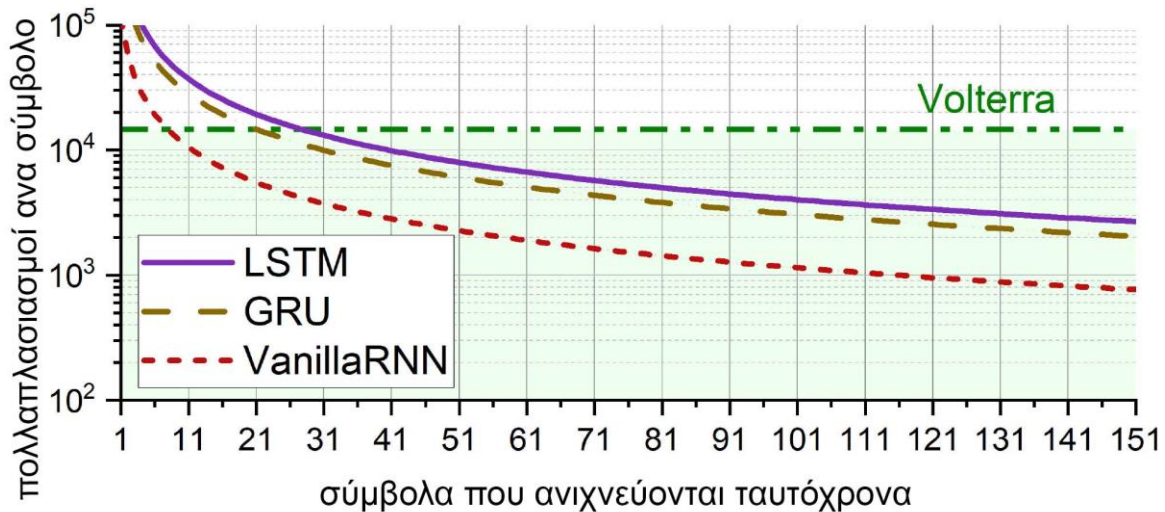
Πίνακας 4.3 : Υπολογιστική πολυπλοκότητα

περιγράφονται από την 1^η τάξη. Αντίστοιχα, τα μη γραμμικά φαινόμενα που περιγράφονται από την 3^η τάξη έχουν ακόμη μικρότερη έκταση.

Με τα παραπάνω δεδομένα και τους αντίστοιχους υπολογισμούς από τους τύπους 4.17 και 4.18 συγκρίνουμε τα τρία μοντέλα bi-RNN με το Volterra ως προς τον αριθμό των πολλαπλασιασμών και τα αποτελέσματα συνοψίζονται στον πίνακα 4.3.

Το μήκος της λέξης διατηρείται στα 151 σύμβολα για όλους τους ισοσταθμιστές ($L_1=151$ για το Volterra), ενώ για τη δεύτερη και τρίτη τάξη έχουμε χρησιμοποιήσει $L_2=51$ και $L_3=11$ όπως εξηγείται στην ανάλυση επίδοσης BER. Μπορεί να φανεί ότι μεταξύ των τριών μοντέλων bi-RNN, το bi-VRNN είναι το λιγότερο πολύπλοκο λόγω της παραμέτρου B . Στην περίπτωση που το RNN εντοπίζει και προβλέπει μόνο το κεντρικό σύμβολο (πολλά σε ένα), το Volterra είναι πολύ λιγότερο πολύπλοκο από όλα τα σχήματα RNN, καθιστώντας έτσι τη χρήση τους μη ελκυστική, (πίνακας 4.3A). Ωστόσο, αυτό το εύρημα είναι παραπλανητικό με την έννοια ότι το μεγάλο πλεονέκτημα της μεθόδου εκπαίδευσης RNN, η προσέγγιση «πολλά σε πολλά», δεν αξιοποιείται εδώ. Από τη σχέση 4.17, μπορεί κανείς εύκολα να διαπιστώσει ότι ο αριθμός των πολλαπλασιασμών ανά σύμβολο στην τοπολογία πολλών σε πολλά μειώνεται με τον αριθμό των συμβόλων που μπορούν να ανιχνευθούν ταυτόχρονα, όπως φαίνεται στο σχήμα 4.14.

Ο Πίνακας 4.3B δείχνει τους πολλαπλασιασμούς ανά σύμβολο που απαιτούνται για την προσέγγιση «πολλά προς πολλά» που χρησιμοποιήθηκε σε αυτήν την εργασία όταν αποκωδικοποιούνται ταυτόχρονα πολλά σύμβολα. Είναι προφανές ότι λαμβάνοντας υπόψη ότι όλα τα σχήματα RNN προσφέρουν ταυτόχρονη ανίχνευση πολλαπλών συμβόλων, αποδεικνύεται ότι είναι λιγότερο πολύπλοκα από τον μη γραμμικό ισοσταθμιστή Volterra που δεν μπορεί να υποστηρίξει παρόμοιο μηχανισμό για ανίχνευση πολλαπλών συμβόλων. Το bi-VRNN αποδεικνύεται ότι είναι το λιγότερο πολύπλοκο όταν εξάγονται 120 σύμβολα εκ των 151 της εισόδου σε βάρος μιας ελαφράς υποβάθμισης στην επίδοση BER όπως απεικονίζεται στο σχήμα 4.14. Επομένως, το μοντέλο bi-RNN μπορεί να είναι σχεδόν 90% ελαφρύτερο από τον μη γραμμικό ισοσταθμιστή Volterra. Πιο αναλυτικά, ο αριθμός των πολλαπλασιασμών που απαιτούνται για την εξαγωγή 1 συμβόλου με τη μέθοδο Volterra που έχει εκπαιδευτεί με λέξη μήκους 151 συμβόλων



Σχήμα 4.19: Αριθμός πολλαπλασιασμών ανά σύμβολο ως συνάρτηση των συμβόλων που ανιχνεύονται ταυτόχρονα.

ισοδυναμεί με την εξαγωγή περίπου 8 συμβόλων με το μοντέλο του bi-VRNN, 22 με το μοντέλο του bi-GRU και 27 με αυτό του bi-LSTM, όπως αναλυτικά αποτυπώνεται στο σχήμα 4.19. Όταν ο αριθμός των εξαγόμενων συμβόλων ξεπερνά τις παραπάνω τιμές, το αντίστοιχο μοντέλο καθίσταται λιγότερο πολύπλοκο από το Volterra.

Αξίζει να σημειωθεί ότι αν και υπάρχουν πολλές μελέτες για τη μείωση της πολυπλοκότητας τόσο για τους εξισωτές Volterra [74] όσο και για τα RNN, όπως η μέθοδος του κλαδέματος (pruning) [140] αποφασίσαμε να προσδιορίσουμε την πολυπλοκότητα των συμβατικών μοντέλων. Ακόμη και αν ληφθεί υπόψη η μείωση της πολυπλοκότητας της τάξης του 70% για τον Volterra με βάση το [74], το μοντέλο bi-VRNN εξακολουθεί να είναι προτιμότερο όσον αφορά την πολυπλοκότητα.

4.3.5 Συμπέρασμα

Σε αυτή την παράγραφο μελετήσαμε αριθμητικά τρία μοντέλα bi-RNN, (LSTM, GRU και VRNN), ως πιθανούς ισοσταθμιστές μη γραμμικότητας οπτικών ινών σε ψηφιακά σύμφωνα συστήματα επικοινωνιών υψηλής χωρητικότητας. Σε αποστάσεις 1000 km όλα

τα μοντέλα εμφάνισαν βελτίωση του BER άνω της μισής τάξης μεγέθους σε σύγκριση με συστήματα που χρησιμοποιούν αποκλειστικά γραμμική εξισορρόπηση και τρεις φορές χαμηλότερη σε σύγκριση με τον μη γραμμικό ισοσταθμιστή Volterra. Η αποτελεσματικότητά τους γίνεται ισχυρότερη σε συστήματα υψηλής διασποράς και η υπεροχή τους έναντι της γραμμικής και μη γραμμικής εξισορρόπησης Volterra επαληθεύτηκε σε πολλά σενάρια διάδοσης. Μεταξύ των τριών μοντέλων, το bi-VRNN παρουσιάζει τη χαμηλότερη πολυπλοκότητα χωρίς να υστερεί σε επιδόσεις BER. Τέλος, υπολογίσαμε την πολυπλοκότητα ως προς τον αριθμό των πολλαπλασιασμών ανά σύμβολο εφαρμόζοντας την τοπολογία «πολλά σε πολλά» και αποδείξαμε ότι το bi-RNN μπορεί να είναι σχεδόν 90% λιγότερο σύνθετο από έναν μη γραμμικό ισοσταθμιστή Volterra 3ης τάξης.

4.4 Πολυκαναλική ανίχνευση βασισμένη σε αμφίδρομα επαναλαμβανόμενα νευρωνικά δίκτυα για την εξομάλυνση διακαναλικών μη γραμμικών φαινομένων σε ψηφιακά σύμφωνα συστήματα

4.4.1 Εισαγωγή

Στις προηγούμενες παραγράφους, προτάθηκε η χρήση διαφόρων τύπων RNN για την εξομάλυνση των μη γραμμικών φαινομένων της διάδοσης στην οπτική ίνα. Παρά το γεγονός της ύπαρξης έντονης, μη γραμμικής διακαναλικής αλληλεπίδρασης, εκπαιδεύσαμε και αξιολογήσαμε τα RNN μόνο χρησιμοποιώντας το κανάλι ενδιαφέροντος.

Σε αυτή την παράγραφο χρησιμοποιούμε την πολυκαναλική ανίχνευση WDM καναλιών προκειμένου να εξισορροπηθούν τα μη γραμμικά φαινόμενα μεταξύ των καναλιών, τροποποιώντας κατάλληλα τους επεξεργαστές bi-RNN των προηγούμενων παραγράφων [1], [2]. Ο ισοσταθμιστής χρησιμοποιεί την αρχιτεκτονική επεξεργασίας πολλαπλών εισόδων πολλαπλών εξόδων (multi-input multi-output, MIMO) για την αποκωδικοποίηση των WDM καναλιών. Με αυτόν τον τρόπο, εκμεταλλεύεται τη χρήσιμη πληροφορία από τα γειτονικά κανάλια για την καλύτερη αναγνώριση και άρση των διακαναλικών αλληλεπιδράσεων. Η MIMO εξισορρόπηση που βασίζεται στη βαθιά μάθηση έχει προταθεί στο παρελθόν [141], αλλά κυρίως στόχευε στην αντιστάθμιση του φαινομένου FWM σε σήματα που υπόκεινται σε πολυπλεξία ορθογώνιας διαίρεσης συχνότητας (orthogonal frequency division multiplexing, OFDM), και έδειξε μικρή βελτίωση της πολυπλοκότητας σε σχέση με την τεχνική DBP. Επιπλέον, η πολυπλοκότητα αυξάνεται γραμμικά με το πλήθος των καναλιών προς επεξεργασία. Στο πλαίσιο της μηχανικής μάθησης έχει προταθεί μια λύση DBP πολλαπλών καναλιών με τη χρήση CNN [142] που παρέχει ισοδύναμη ή ακόμα καλύτερη απόδοση από την DBP και πολύ λιγότερη πολυπλοκότητα. Ωστόσο, αυτή η εργασία δεν αποφεύγει το γεγονός ότι η πολυπλοκότητα κλιμακώνεται με τον αριθμό των spans. Αντιθέτως, η πρότασή μας επικεντρώνεται αποκλειστικά στη μη γραμμική εξισορρόπηση καθώς τα γραμμικά φαινόμενα μπορούν εύκολα να αντιμετωπιστούν από γραμμικούς ισοσταθμιστές στον τομέα του χρόνου ή της συχνότητας, πράγμα που σημαίνει ότι η γραμμική αντιστάθμιση εφαρμόζεται μόνο μία

φορά και όχι για κάθε διάστημα, όπως στην περίπτωση DBP. Στοχεύει τόσο στη βελτίωση της απόδοσης μέσω της αύξησης της χρήσιμης πληροφορίας των γειτονικών καναλιών WDM όσο και στη μείωση της πολυπλοκότητας, όταν χρησιμοποιείται ένας ισοσταθμιστής RNN πολλαπλών καναλιών αντί για πολλούς μονοκαναλικούς. Το RNN είναι πιο δύσκολο να παραλληλιστεί στα FPGA όπως φαίνεται στο [143], ωστόσο πιστεύουμε ότι αξίζει να χρησιμοποιηθεί το bi-RNN που προτείνεται εδώ καθώς απαιτεί έναν αρκετά χαμηλό αριθμό κρυφών μονάδων που μπορούν να υλοποιηθούν από το DSP [143]. Επιπλέον, μια πρόσφατη εργασία έδειξε ότι χρησιμοποιώντας την απόσταση γνώσης μπορεί κανείς να επιλύσει το πρόβλημα της παραλληλοποίησης των επαναλαμβανόμενων νευρωνικών δικτύων [144].

Αποδεικνύουμε μέσω εκτεταμένων αριθμητικών προσομοιώσεων ότι η προτεινόμενη προσέγγιση ξεπερνά τόσο την τυπική εξισορρόπηση πολλαπλών καναλιών με τη μορφή προσαρμοστικών ισοσταθμιστών, όσο και τους ισοσταθμιστές τύπου DBP αλλά και τον μονοκαναλικό ισοσταθμιστή bi-RNN. Επαληθεύουμε πειραματικά αυτή την πρόταση με ένα σύστημα διάδοσης 3 καναλιών 11,25 Gbaud QPSK μήκους 180 km που μιμείται τις μη γραμμικότητες ενός συστήματος διάδοσης μεγάλων αποστάσεων αυξάνοντας την ισχύ. Τέλος, παρουσιάζεται μείωση της υπολογιστικής πολυπλοκότητας περισσότερη από 550% σε σύγκριση με το μονοκαναλικό DBP σε σενάρια που εξετάζονται σε μεγάλες αποστάσεις (>1000 km).

Η ερευνητική δουλειά στο πλαίσιο αυτής της συνιστώσας της διδακτορικής διατριβής αποτελεί αποτέλεσμα συνεργασίας με το Optoelectronics Research Centre (ORC) του Πανεπιστημίου του Southampton και συγκεκριμένα με τους Kyle R. H. Bottrill και τον Καθηγητή Π. Πετρόπουλο.

4.4.2 Προτεινόμενο σύστημα πολυκαναλικής ισοστάθμισης

Το προτεινόμενο σύστημα που προσομοιώνουμε πραγματοποιεί διάδοση τυπικής μονότροπης οπτικής ίνας στη ζώνη C. Λαμβάνοντας υπόψη το γεγονός ότι η οπτική ίνα είναι ένα μη γραμμικό κανάλι με έντονο το χαρακτηριστικό της διασποράς,

προσομοιώνουμε επιπλέον το σενάριο στο οποίο η διασπορά είναι μικρότερη από την τυπική της ζώνης C. Παρά το γεγονός πως η περίπτωση αυτή δεν συναντάται συχνά στις υπάρχουσες οπτικές ζεύξης και θα μπορούσε να έχει εφαρμογή μόνο στην περίπτωση οπτικών ινών μη μηδενικής μετατοπισμένης διασποράς (Non-zero dispersion-shifted fiber, NZDSF), είναι επιστημονικά σημαντικό να διερευνηθεί η επίδραση των διαφορετικών τιμών της διασποράς στις επιδόσεις της πολυκαναλικής ισοστάθμισης. Το τελευταίο αποκτά ιδιαίτερη σημασία αν λάβει κανείς υπόψη την προοπτική της επέκτασης των οπτικών συστημάτων διάδοσης μεγάλων αποστάσεων σε άλλες ζώνες, όπως πχ η ζώνη – O, η οποία χαρακτηρίζεται από μικρότερη διασπορά σε σχέση με τη ζώνη C.

Η απόδοση του συστήματος αξιολογείται με την τιμή του BER που υπολογίζεται μέσω της μέτρησης των σφαλμάτων στον δέκτη.

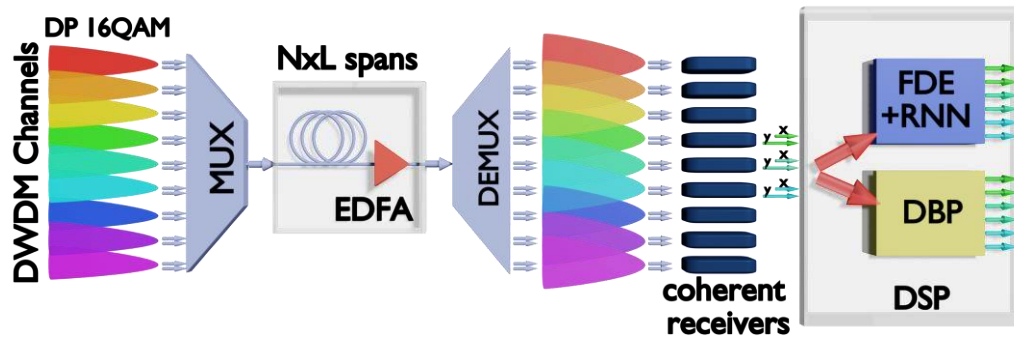
Η σχηματική απεικόνιση του προτεινόμενου μοντέλου παρουσιάζεται στο σχήμα 4.20. Αντιστοιχεί σε διάδοση 1.200 km στη ζώνη C, προσομοιώνεται με την ολοκλήρωση της μη γραμμικής εξίσωσης Schrödinger (NLSE), θεωρώντας οπτική ίνα με συντελεστή εξασθένησης $\alpha=0.2$ dB/Km, συντελεστή διασποράς 2^{ης} τάξης $\beta_2=-20$ ps²/km και μη γραμμικό συντελεστή Kerr $\gamma=1.3$ W⁻¹km⁻¹. Η οπτική διάδοση μοντελοποιήθηκε βασιζόμενη στις εξισώσεις του Manakov, χρησιμοποιώντας τη μέθοδο split-step Fourier [145]. Θεωρούμε διάδοση 9 WDM καναλιών, με τη μεταξύ τους απόσταση στα 75 GHz, διαμόρφωση 16-QAM διπλής πόλωσης και ρυθμό 64 Gbaud. Πραγματοποιείται ενίσχυση ανά 50 km με εικόνα θορύβου 5 dB. Για τη σε βάθος μελέτη της συμπεριφοράς των προτεινόμενων ισοσταθμιστών, μοντελοποιήθηκαν επιπλέον σενάρια, και συγκεκριμένα οπτική διάδοση συνολικά 800 km (16 τμήματα των 50 km) με τις προηγούμενες παραμέτρους και ένα σύστημα που προσιδιάζει στα πλέον σύγχρονα που έχουν προταθεί [146] στο οποίο προσομοιώνουμε τη διάδοση 1200 km σε ρυθμό 100 GBaud ανά πόλωση, 21 WDM καναλιών με τη μεταξύ τους απόσταση στα 125 GHz και τμηματική ενίσχυση ανά 80 km.

Η αποπολυπλεξία πόλωσης, η εκτίμηση της συχνότητας και της φάσης της φέρουσας πραγματοποιούνται ιδανικά ενώ δεν λήφθηκε υπόψη ο θόρυβος φάσης του laser. Οι πομποδέκτες θεωρούνται επίσης ιδανικοί (χωρίς ατέλειες) καθώς ο κύριος στόχος είναι η

εστίαση στην εξισορρόπηση των μη γραμμικών φαινομένων που προκαλούνται από το φαινόμενο Kerr.

Οι προσομοιώσεις διεξήχθησαν χρησιμοποιώντας RRC διαμόρφωση παλμού, με συντελεστή roll-off 0.1 και προσαρμοσμένο φίλτρο στον δέκτη. Η ισοστάθμιση υλοποιείται με έναν ισοσταθμιστή bi-VRNN, που δέχεται στην είσοδό του μονάχα 1 δείγμα ανά σύμβολο. Πριν από οποιαδήποτε μετα-επεξεργασία ή αποδιαμόρφωση, εκτελούμε αντιστάθμιση της χρωματικής διασποράς με τη χρήση ενός ιδανικού ισοσταθμιστή τύπου FDE.

Ο ADC στον δέκτη λειτουργεί με ρυθμό συμβόλων με ανάλυση 8 bit.



Σχήμα 4.20: Προσομοιωμένο σύστημα διάδοσης και πολυκαναλικής ισοστάθμισης

Το σύνολο των παραμέτρων της προσομοίωσης συνοψίζεται στον πίνακα 4.4

Σύμβολο	Παράμετρος	Τιμή
a	εξασθένιση	0.2 dB/km
β_2	Διασπορά 2 ^{ης} τάξης	-21 ps ² /km (τυπική διασπορά στη ζώνη C), -0.82 ps ² /km (χαμηλή διασπορά)
γ	Μη γραμμικός όρος ίνας	1.3 W ⁻¹ km ⁻¹

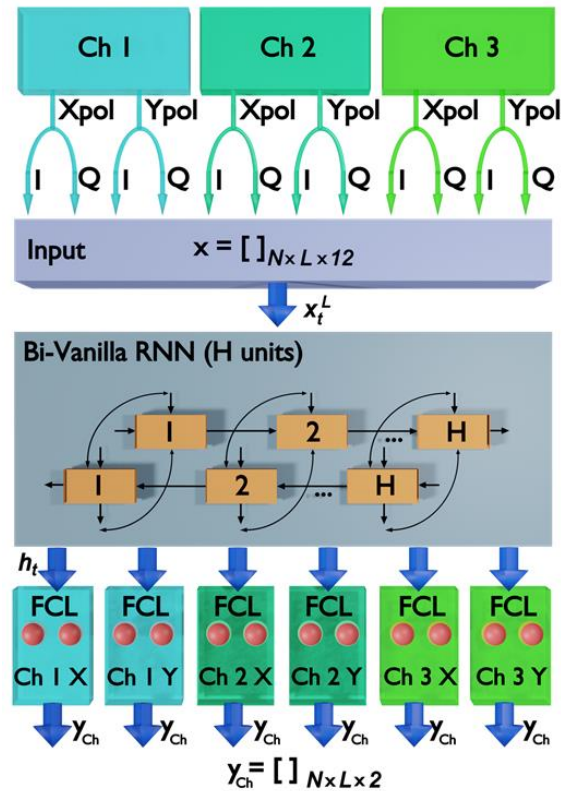
R	Ρυθμός συμβόλων	64 Gbaud/πόλωση
M	Τύπος διαμόρφωσης	16-QAM με πολυπλεξία πόλωσης
L	Περιοδικό μήκος διάδοσης (span)	50 km (ή 80 km)
Δf	Απόσταση καναλιών	75 GHz
L_{tot}	Συνολικό Μήκος διάδοσης	1200 km (24 x 50 km), 800 km (16 x 50 km), 1200 km (15 x 80 km)

Πίνακας 4.4 : Παράμετροι αριθμητικού μοντέλου

Συγκεκριμένα, χρησιμοποιούμε και τις δύο πολώσεις τριών (3), πέντε (5) ή και m γενικότερα γειτονικών καναλιών ως είσοδο στο νευρωνικό μας δίκτυο προκειμένου να εφαρμόσουμε πολυκαναλική εξίσωση. Στην αναφορά [2] χρησιμοποιήσαμε και τις δύο πολώσεις του καναλιού ενδιαφέροντος, ως είσοδο στο μοντέλο νευρωνικού δικτύου, το οποίο παρέχει λιγότερες πληροφορίες στη διαδικασία εξισορρόπησης αφού δε λαμβάνει υπόψη τα διακαναλικά μη-γραμμικά φαινόμενα επαρκώς (XPM, FWM).

Σημειώνουμε ότι υιοθετήσαμε το μοντέλο bi-VRNN για την υλοποίηση της εξισορρόπησης, επειδή προσφέρει παρόμοια απόδοση σε σύγκριση με άλλα μοντέλα RNN (π.χ. LSTM και GRU), ενώ παρουσιάζει χαμηλότερη υπολογιστική πολυπλοκότητα, όπως διεξοδικά αναλύθηκε στις προηγούμενες παραγράφους και δημοσιεύτηκε στις [2] και [6].

Το διάγραμμα στρωμάτων του ισοσταθμιστή bi-VRNN φαίνεται στο σχήμα 4.21. Απεικονίζεται η περίπτωση της κοινής επεξεργασίας 3 καναλιών των δύο πολώσεων.



Σχήμα 4.21: Διάγραμμα του Bi-VRNN ισοσταθμιστή. Απεικονίζεται η περίπτωση πολυκαναλικής επεξεργασίας 3 καναλιών (2 πολώσεις ανά κανάλι)

Η είσοδος x είναι ένα διάνυσμα $N \times L \times 12$ διαμορφωμένων με σχήμα 16-QAM καναλιών, όπου το N υποδηλώνει τον συνολικό αριθμό των συμβόλων εισόδου και το $L = 2k + 1$ δείχνει το μήκος της λέξης συμβόλων της εισόδου. Η είσοδος x έχει συνολικά 12 χαρακτηριστικά, αυτά του I και του Q των πολώσεων καθενός από τα 3 κανάλια. Στην περίπτωση των 5 καναλιών η είσοδος διαθέτει 20 χαρακτηριστικά και η περίπτωση των m καναλιών αντίστοιχα $m \times 4$ χαρακτηριστικά.

Τη χρονική στιγμή t η είσοδος μπορεί να εκφραστεί από τη σχέση :

$$x_t^L = [x_{t-k}, \dots, x_{t-1}, x_t, x_{t+1}, \dots, x_{t+k}] \quad (4.19)$$

υποδηλώνοντας ότι χρησιμοποιούνται τα k προηγούμενα και k επόμενα σύμβολα του τρέχοντος διανύσματος συμβόλων x_t λαμβάνοντας υπόψη τη χρονική εξέλιξη του

σήματος, τις ενδοκαναλικές και διακαναλικές αλληλεπιδράσεις μεταξύ των συμβόλων, όπως άλλωστε ακολουθήθηκε σε όλα τα προηγούμενα μοντέλα στην μονοκαναλική εξίσωση. Το μήκος L εξαρτάται από τη μνήμη του καναλιού, η οποία σχετίζεται με τη συσσωρευμένη χρωματική διασπορά στη μονότροπη οπτική ίνα και το περιορισμένο εύρος ζώνης του πομποδέκτη.

Καθεμία από τις κρυφές μονάδες παράγει μία έξοδο h_t που δίνεται από τον τύπο:

$$h_t = \tanh(W \cdot h_{t-1} + U \cdot x_t) \quad (4.20)$$

όπου $x_t \in \mathbb{R}^L$ είναι το διάνυσμα εισόδου διάστασης L τη χρονική στιγμή t , H ο αριθμός των κρυφών μονάδων, $U \in \mathbb{R}^{H \times L}$ and $W \in \mathbb{R}^{H \times H}$ αντιπροσωπεύουν τους πίνακες των βαρών εκπαίδευσης, h_t , h_{t-1} είναι η είσοδος, η κρυφή έξοδος και η προηγούμενη κρυφή έξοδος αντίστοιχα. Οι κρυφές μονάδες συνδέονται μεταξύ τους μέσω μιας επαναλαμβανόμενης δομής. Το στρώμα bi-VRNN αποτελείται από δύο πίνακες βάρους $H \times H$, έναν για την προς την πρόσω και έναν για την οπίσθια κατεύθυνση. Οι δύο κατευθύνσεις συνδέονται στην έξοδο h_t του στρώματος bi-VRNN με ένα πίνακα βαρών ($2H \times 2$). Η ίδια έξοδος του επιπέδου bi-VRNN, h_t , αποστέλλεται στη συνέχεια παράλληλα σε έξι (6) πλήρως συνδεδεμένα στρώματα (FCL), τα οποία αντιστοιχούν και στις δύο πολώσεις καθενός από τα 3 κανάλια που ανιχνεύονται από κοινού. Κάθε FCL αποτελείται από δύο νευρώνες, που αντιπροσωπεύουν τα I και Q του 16-QAM και η έξοδος του y_{ch} είναι ένα διάνυσμα $N \times L \times 2$.

Σε αντίθεση με τις προηγούμενες παραγράφους, όπου χρησιμοποιήθηκε η μέθοδος της ταξινόμησης στη βάση του διαθέσιμου αριθμού των συμβόλων (16 στην περίπτωση του 16-QAM) [1], ή η ταξινόμηση στη βάση των επιμέρους bits των συμβόλων (4 στην περίπτωση του 16-QAM) [2], στη μελέτη αυτή χρησιμοποιούμε τη μέθοδο της παλινδρόμησης για την εκπαίδευση του νευρωνικού μοντέλου. Η μέθοδος αυτή μειώνει σημαντικά την πολυπλοκότητα υλοποίησης σε σύγκριση με τις προηγούμενες μεθόδους ταξινόμησης και σε πολλές περιπτώσεις δείχνει πιο σταθερή συμπεριφορά και εξίσου, αν όχι καλύτερα αποτελέσματα ως προς την επίδοση BER [37].

Η εκπαίδευση πραγματοποιείται χρησιμοποιώντας την προσέγγιση πολλά προς πολλά που, όπως αναλυτικά παρουσιάστηκε στα προηγούμενα, παράγει ταυτόχρονα τον ίδιο

αριθμό συμβόλων y με αυτά στην είσοδο x και είναι ικανό να εξάγει πολλαπλά σύμβολα ταυτόχρονα [2]. Επιλέξαμε προσεκτικά τη διαδικασία για την δημιουργία ψευδοτυχαίων μη επαναλαμβανόμενων ακολουθιών χρησιμοποιώντας την πολύ μεγάλη περίοδο της γεννήτριας Mersenne Twister $2^{19937} - 1$, ώστε να αποφευχθεί η πρόβλεψη από το bi-RNN του επόμενου συμβόλου της ψευδοτυχαίας ακολουθίας που θα οδηγούσε στην υπερεκτίμηση των επιδόσεών του.

Το μοντέλο σχεδιάστηκε, εκπαιδεύτηκε και αξιολογήθηκε στο Keras με υποστήριξη της πλατφόρμας Tensorflow-GPU 2.10 GPU. Ως συνάρτηση απώλειας χρησιμοποιήθηκε το μέσο τετραγωνικό σφάλμα (MSE) καθώς αποτελεί τη βέλτιστη επιλογή στα προβλήματα παλινδρόμησης [37]. Ο αλγόριθμος βελτιστοποίησης που χρησιμοποιούμε είναι ο Adam με βήμα εκπαίδευση 5×10^{-4} . Για να αποφύγουμε την υπερβολική προσαρμογή κατά τη διάρκεια της εκπαίδευσης, χρησιμοποιούμε πρόωρη διακοπή όταν η συνάρτηση απώλειας δεν μειώνεται για 50 συνεχόμενες εποχές. Τα βάρη τόσο του bi-Vanilla-RNN όσο και του FCL αρχικοποιούνται χρησιμοποιώντας το Glorot uniform initializer [147].

Θεωρήσαμε 100.000 λέξεις συμβόλων για εκπαίδευση, 50.000 λέξεις για επικύρωση και 100.000 λέξεις σύμβολα για αξιολόγηση με άγνωστα δεδομένα ενώ η $(i)^{\text{η}}$ λέξη είναι μετατόπιση της $(i-1)^{\text{η}}$ λέξης κατά ένα σύμβολο. Με αυτόν τον τρόπο αυξάνουμε τον αριθμό των λέξεων που μπορούν να χρησιμοποιηθούν για την εκπαίδευση του μοντέλου. Στο στάδιο συμπερασμάτων η μετατόπιση είναι 40 σύμβολα μεταξύ διαδοχικών λέξεων με αποτέλεσμα περίπου 2.5 k λέξεις. Επομένως, το BER υπολογίζεται λαμβάνοντας υπόψη 2.5 k λέξεις με άγνωστα δεδομένα με βάση την αξιολόγηση των 41 κεντρικών συμβόλων κάθε λέξης.

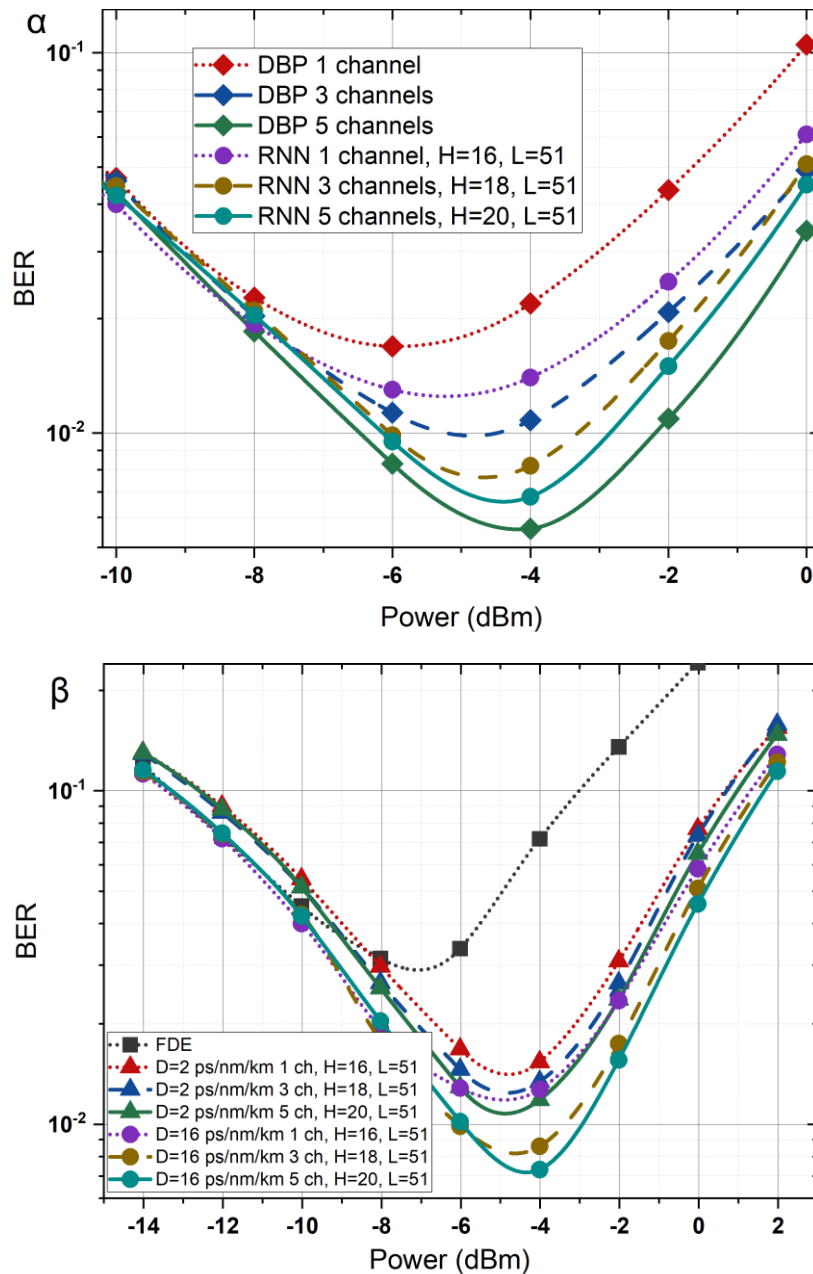
Το στάδιο εκπαίδευσης εκτελέστηκε σε πακέτα (batches) των 512 λέξεων συμβόλων, εφαρμόζοντας τη μέθοδο Mini-batch Gradient Decent για μέγιστο 1000 εποχές για την περίπτωση του ενός καναλιού ή 2000 για τις περιπτώσεις των πολλαπλών καναλιών. Η μέθοδος του πρόωρου τερματισμού (early stopping) σταματά την εκπαίδευση όταν το μέσο τετραγωνικό σφάλμα παραμένει σταθερό για περισσότερες από 20 εποχές εμποδίζοντας την υπερπροσαρμογή (overfitting).

Για τη συγκριτική μελέτη των επιδόσεων και της πολυπλοκότητας του προτεινόμενου ισοσταθμιστή, χρησιμοποιούμε τον πολυκαναλικό ισοσταθμιστή τύπου DBP μαζί με ένα προσαρμοστικό MIMO ισοσταθμιστή. Ο τελευταίος αναλαμβάνει να εξισορροπήσει τα εναπομείναντα γραμμικά φαινόμενα μετά το στάδιο της ισοστάθμισης τύπου FDE. Όπως έχει φανεί ο προσαρμοστικός MIMO ισοσταθμιστής μπορεί να βελτιώσει περαιτέρω τις επιδόσεις του συστήματος ψηφιακής επεξεργασίας τύπου DBP [148]. Ο ισοσταθμιστής τύπου DBP πλήρους πεδίου θεωρείται μη εφικτός για πραγματικές υλοποιήσεις, καθώς απαιτεί έναν δέκτη με το εύρος ζώνης ολόκληρου του πολυκαναλικού φάσματος που υπερβαίνει κατά πολύ τα 500 GHz σε τυπικά συστήματα μετάδοσης WDM [149]. Ωστόσο, αυτό είναι το βέλτιστο σενάριο συγκριτικής αξιολόγησης για τα αποτελέσματα της προσομοίωσης και οποιαδήποτε άλλη παραλλαγή ισοσταθμιστή τύπου DBP θα ήταν υποδεέστερη από άποψη απόδοσης. Στην περίπτωση μας, για την προσαρμογή του αριθμού των καναλιών στο πεδίο της οπισθοδιάδοσης, επιλέγουμε τα κανάλια που μας ενδιαφέρουν μέσω μιας διαδικασίας οπτικού φιλτραρίσματος διαδικασία που απορρίπτει επίσης τον ανεπιθύμητο ASE θόρυβο εκτός ζώνης. Έπειτα δειγματοληπτούμε το ληφθέν πεδίο στα 4 δείγματα ανά σύμβολο εφαρμόζοντας τη μέθοδο DBP με 20 βήματα ανά τμήμα μήκους. Παρότι τα 2 δείγματα ανά σύμβολο με 2 ή 4 βήματα ανά τμήμα μήκους θα ήταν πιο ρεαλιστικό σενάριο, εφαρμόζουμε αυτό το μαξιμαλιστικό και μη εφαρμόσιμο σενάριο με σκοπό να συγκρίνουμε το bi-RNN με ένα σχεδόν ιδανικό σχήμα μη-γραμμικής, πολυκαναλικής ισοστάθμισης. Βελτιστοποιούμε τη μη γραμμική παράμετρο σε κάθε προσομοίωση DBP προκειμένου να αποκτήσουμε το χαμηλότερο δυνατό BER. Μετά την εφαρμογή του DBP τα κανάλια διαχωρίζονται με ψηφιακά φίλτρα, επιστρέφουν στη βασική ζώνη και δειγματοληπτούνται στο 1 δείγμα ανά σύμβολο. Τέλος, ένας γραμμικός προσαρμοστικός αλγόριθμος με 21 taps υλοποιείται ξεχωριστά για το πραγματικό και το φανταστικό μέρος κάθε καναλιού, ώστε να εξισωθούν όλα τα εναπομείναντα γραμμικά φαινόμενα.

4.4.3 Παρουσίαση αποτελεσμάτων εξομοίωσης της πολυκαναλικής επεξεργασίας

Προχωρώντας στην εκπαίδευση των μοντέλων bi-VRNN με τα προαναφερθέντα χαρακτηριστικά, εξάγουμε τα αποτελέσματα με τη μορφή της τιμής BER συναρτήσεως της οπτικής ισχύος που φαίνονται στο σχήμα 4.22. Λαμβάνοντας υπόψη τη διάδοση 9 πυκνών WDM καναλιών με τη μεταξύ τους απόσταση στα 75GHz τόσο για την περίπτωση της τυπικής ζώνης C (ο συντελεστής χρωματικής διασποράς $D=16$ ps/nm/km) όσο και για τη διάδοση σε περιβάλλον χαμηλής διασποράς ($D=2$ ps/nm/km), για συνολικό μήκος διάδοσης 1200 km αποτελούμενο από 24 τμήματα οπτικής ίνας μήκους 50 km. Συγκρίνουμε τη βελτίωση του BER που προσφέρει το σύστημα πολυκαναλικής ανίχνευσης για 3 και 5 γειτονικά κανάλια έναντι της ανίχνευσης ενός καναλιού. Στο 4.22. απεικονίζεται μόνο το BER για την πόλωση x του κεντρικού καναλιού. Τόσο τα γειτονικά κανάλια όσο και η y πόλωση παρουσιάζουν παρόμοια αποτελέσματα οπότε η απεικόνιση τους παραλείπεται. Παρά το γεγονός πως η μνήμη του καναλιού στην περίπτωση της χαμηλής διασποράς είναι αρκετά μικρότερη, για την ισότιμη σύγκριση των δύο ζωνών το μήκος της λέξης των συμβόλων L έχει και στις δύο περιπτώσεις την τιμή 51 (σχήμα 4.22β).

Εύκολα μπορεί κανείς να συμπεράνει ότι όταν χρησιμοποιείται πολυκαναλική εκπαίδευση, το bi-VRNN κατορθώνει να ανιχνεύσει καλύτερα τα χαρακτηριστικά της μη γραμμικής αλληλεπίδρασης μεταξύ των καναλιών, που αποδίδονται στην ετεροδιαμόρφωση φάσης (XPM) και μίξης τεσσάρων κυμάτων (FWM). Τα αποτελέσματα δείχνουν πως στην περίπτωση της C-band με την τυπική διασπορά η ελάχιστη τιμή σφάλματος BER για την αναγνώριση ενός καναλιού είναι 1.3×10^{-2} στα -4 dBm ισχύος ενώ

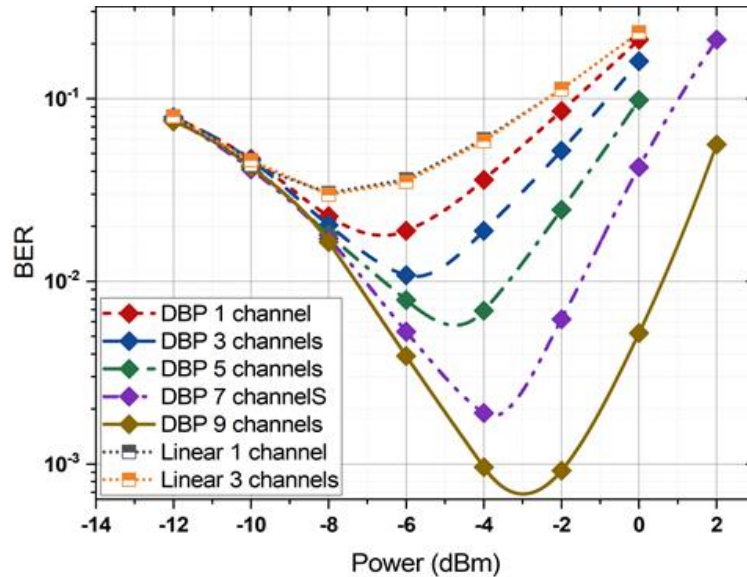


Σχήμα 4.22: Το BER ως συνάρτηση της οπτικής ισχύος για μονοκαναλική και πολυκαναλική ανίχνευσης, για το RNN σε σχέση με το DBP (α) και το σενάριο μεγάλης (τυπικής για τη ζώνη C) ή χαμηλής διασποράς (β) αντίστοιχα. Αποτυπώνεται μόνο το κεντρικό κανάλι ενώ τα γειτονικά κανάλια δείχνουν παρόμοιες επιδόσεις. Η και L ο αριθμός των κρυφών μονάδων και το μήκος της λέξης των συμβόλων αντίστοιχα

για τα τις περιπτώσεις της αναγνώρισης πολλών καναλιών ταυτόχρονα παρατηρούμε 8.6×10^{-3} στα -4 dBm και 7.3×10^{-3} στα -3.5 dBm ισχύος για τα 3 και τα 5 κανάλια αντίστοιχα, όπως φαίνεται στο σχήμα 4.22α. Συγκρινόμενο με τους ισοσταθμιστές τύπου DBP, οι οποίοι παρουσιάζουν με τη σειρά τους βελτίωση στις επιδόσεις τους όσο αυξάνονται τα κανάλια που επεξεργάζονται ταυτόχρονα και συγκεκριμένα αποδίδουν ελάχιστες τιμές BER 1.9×10^{-2} στα -6 dBm, 1.1×10^{-2} στα -6 dBm και 6.9×10^{-3} στα -4 dBm για 1, 3 και 5 κανάλια αντίστοιχα, το bi-VRNN των 5 καναλιών παρουσιάζει ελαφρώς καλύτερες επιδόσεις από τις αντίστοιχες του DBP των 3 καναλιών το οποίο όμως, όπως αναλύθηκε στις προηγούμενες παραγράφους, αποτελεί μία εξιδανικευμένη περίπτωση εξισωτή που δεν μπορεί να εφαρμοστεί σε πρακτικό επίπεδο, αλλά μελετάται ως τεχνική αναφοράς. Συνεπώς το πολυκαναλικό bi-RNN μπορεί να επιτύχει επιδόσεις εφάμιλλες ενός πολυκαναλικού ιδανικού, και ιδιαίτερα πολύπλοκου DBP ισοσταθμιστή.

Από την άλλη μεριά, στην περίπτωση που η διασπορά είναι μικρή ($D=2$ ps/nm/km) η συνεισφορά των γειτονικών καναλιών δείχνει να είναι λιγότερο σημαντική, όπως αποτυπώνεται στο σχήμα 4.22β. Αποδεικνύεται πως εκτός από την επίδραση των γειτονικών καναλιών, καθώς μειώνεται η διασπορά, ο χρόνος συμφωνίας του καναλιού μειώνεται επίσης με αποτέλεσμα τη δημιουργία διακαναλικής παρεμβολής με περιεχόμενο υψηλής συχνότητας συγκρίσιμο με το ρυθμό των συμβόλων. Σε αυτό το σενάριο το μοντέλο RNN δε μπορεί να αφομοιώσει επαρκώς την διακαναλική παρεμβολή και να την ενσωματώσει στην διαδικασία της ισοστάθμισης. Αυτό το εύρημα είναι παρόμοιο με το αντίστοιχο της μονακαναλικής RNN ισοστάθμισης που παρουσιάστηκε στο κεφ.4.3.3 (σχ. 4.17). Καθίσταται πλέον προφανές ότι η απόδοση της μη γραμμικής εξισορρόπησης του bi-RNN είναι βελτιωμένη όταν η μνήμη του καναλιού αυξάνεται εξαιτίας της (αυξημένης) διασποράς. Κατά συνέπεια η επιβάρυνση της εκπαίδευσης του μοντέλου με τα γειτονικά κανάλια σε περιβάλλον χαμηλής διασποράς σε σύγκριση με το σενάριο της τυπικής διασποράς στη ζώνη C, οδηγεί σε οριακή βελτίωση ως προς το BER. Συγκεκριμένα, παρατηρήθηκε ελάχιστο BER 1.5×10^{-2} στα -4 dBm ισχύος, 1.3×10^{-2} στα -4 dBm και 1.2×10^{-2} επίσης στα -4 dBm για 1, 3 και 5 κανάλια αντίστοιχα.

Για λόγους πληρότητας στο σχήμα 4.23 παρουσιάζονται οι επιδόσεις του DBP ισοσταθμιστή για επεξεργασία ενός πλήρους πεδίου για ένα έως εννέα κανάλια, μαζί με

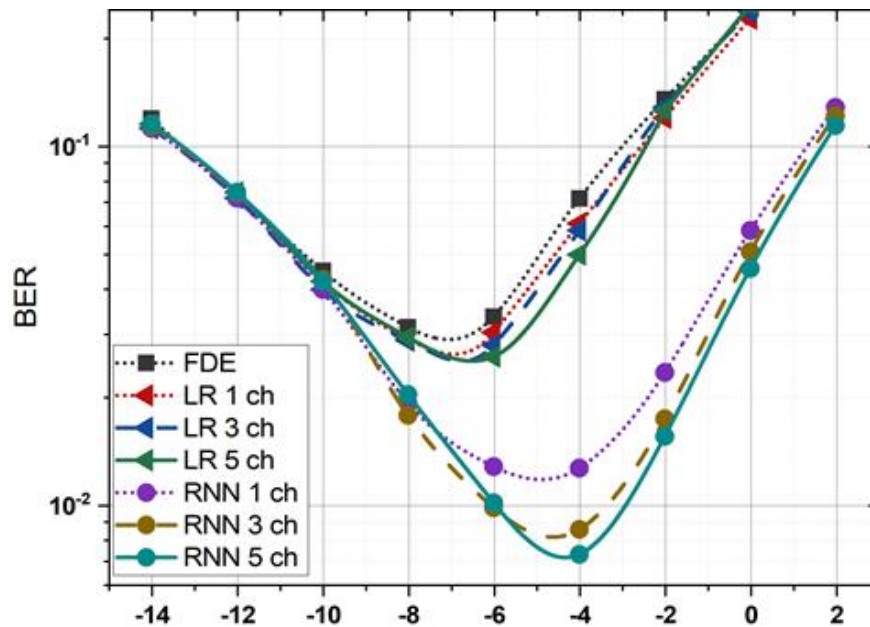


Σχήμα 4.23: Το BER ως συνάρτηση της οπτικής ισχύος για το DBP κλιμακώνοντας τον αριθμό των καναλιών που διαδίδονται αντίστροφα από ένα σε εννέα

τον γραμμικό προσαρμοστικό MIMO ισοσταθμιστή στο 1 κανάλι ή στην έκδοση των 3 καναλιών. Όπως μπορεί κανείς να παρατηρήσει ο γραμμικός αλγόριθμος παρέχει τη βάση αναφοράς ενώ κάθε κανάλι που προστίθεται στο DBP ισοσταθμιστή αποφέρει τμηματικά βελτίωση στις επιδόσεις. Αντιθέτως η προσθήκη της επιπλέον πληροφορίας των γειτονικών καναλιών στο γραμμικό αλγόριθμο δεν προσφέρει καμία βελτίωση. Από την άλλη πλευρά, οι επιδόσεις BER βελτιώνονται σημαντικά όταν 5, 7 ή τα πλήρη 9 κανάλια συμπεριλαμβάνονται στο πεδίο της αντίστροφης διάδοσης. Ωστόσο είναι ιδιαίτερα σημαντικό να αναφέρουμε πως το bi-RNN των 5 καναλιών δεν μπορεί να συγκριθεί με το DBP των 7 ή 9 καναλιών λόγω του δεκαπλάσιου εύρους ζώνης που χρησιμοποιεί το DBP πλήρους πεδίου.

Δοκιμάζουμε επίσης την περίπτωση της αποκλειστικά γραμμικής εξισορρόπησης μέσω του αλγορίθμου της γραμμικής παλινδρόμησης, ενδεικτικά στο σενάριο της ζώνης C ($D=16$ ps/nm/km). Για τον σκοπό αυτό θεωρούμε το δίκτυο του σχ. 4.21, αφαιρώντας εντελώς το επίπεδο bi-VRNN, με τη γραμμική παλινδρόμηση να εφαρμόζεται στην έξοδο του FDE. Η επίδραση του γραμμικού αλγορίθμου, εφαρμοζόμενου σε 1, 3 ή 5 κανάλια σε μια διάταξη MIMO είναι αμελητέα, όπως φαίνεται στο σχήμα 4.24, αποδεικνύοντας ότι το

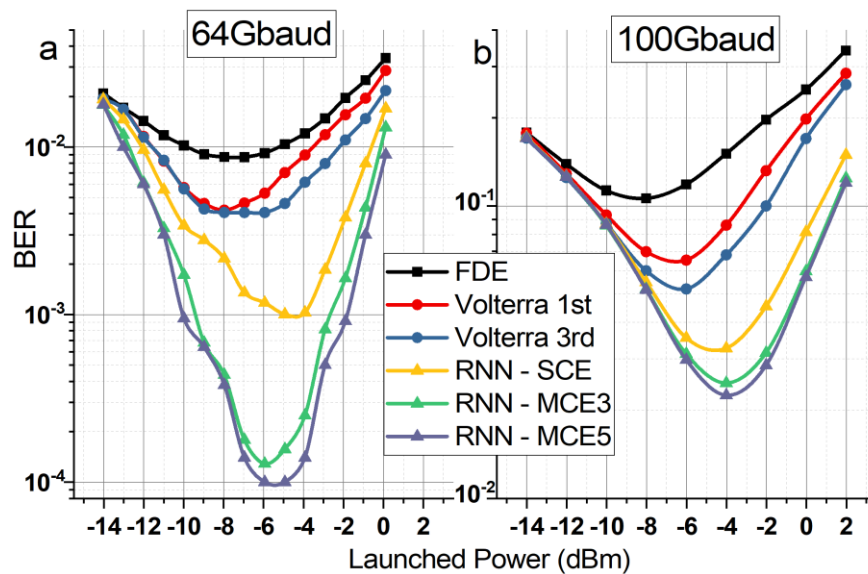
bi-VRNN όχι μόνο βελτιώνει την απόδοση του BER στην ανίχνευση ενός καναλιού, αλλά προσφέρει επίσης περαιτέρω βελτίωση εάν εκπαιδευτεί ταυτόχρονα με τις ακολουθίες πολλαπλών καναλιών. Το υψηλότερο κέρδος παρατηρείται όταν μετακινηθούμε από 1 σε 3 κανάλια, παρά από τα 3 στα 5 κανάλια στη διαδικασία κοινής εξισορρόπησης. Αυτό το εύρημα είναι λογικό, καθώς ένα κανάλι WDM επηρεάζεται κυρίως από τα δύο γειτονικά του κανάλια.



Σχήμα 4.24: Το BER ως συνάρτηση της οπτικής ισχύος για μονακαναλική και πολυκαναλική εξισορρόπηση με τη βοήθεια ή ισοσταθμιστή RNN ή γραμμικής παλινδρόμησης (LR), εφαρμοζόμενου στην έξοδο του FDE.

Πραγματοποιούμε επιπλέον σύγκριση με ισοσταθμιστή βασισμένο σε σειρές Volterra 1^{ης} τάξης (γραμμικό) και 3^{ης} τάξης, όπως φαίνεται στο σχήμα 4.25, για τα σενάρια των 9 WDM καναλιών, σε απόσταση 75GHz ρυθμού 64 Gbaud, 800 km συνολικής οπτικής διάδοσης (16 × 50 km), (σχ. 4.25α) και των 21 WDM καναλιών, σε απόσταση 125GHz ρυθμού 100 Gbaud, 1200 km συνολικής οπτικής διάδοσης (15 × 80 km), (σχ. 4.25β). Στα σχήματα αυτά απεικονίζονται οι επιδόσεις BER μετά από καταμέτρηση σφαλμάτων για

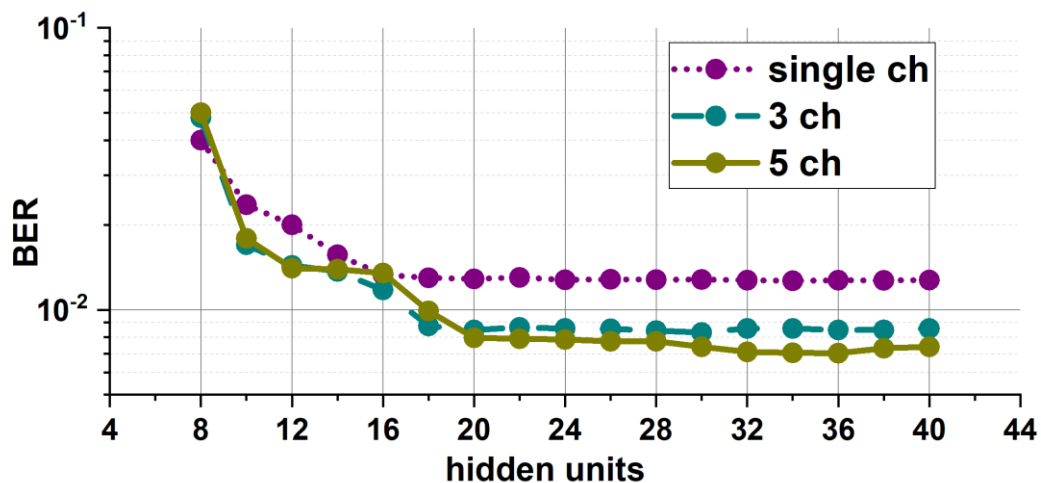
εξισορρόπηση ενός καναλιού χρησιμοποιώντας μόνο FDE και FDE ακολουθούμενο από μη γραμμικό ισοσταθμιστή Volterra (VNLE) (1ης τάξης και 3ης τάξης), μονοκαναλικού bi-RNN και πολλαπλών καναλιών (3 ή 5 WDM) εξισορρόπηση bi-RNN. Η επίδοση BER απεικονίζεται για το κεντρικό κανάλι WDM και στα δύο σενάρια προσομοίωσης διάδοσης (64 Gbaud και 100 Gbaud). Παρόμοια συμπεριφορά έχει καταγραφεί και για τα άλλα κανάλια του WDM σήματος. Αποδεικνύεται η συστηματική βελτίωση των επιδόσεων στις περιπτώσεις της πολυκαναλικής bi-RNN εξισορρόπησης η οποία αγγίζει τη μια τάξη μεγέθους σε σχέση με τη μονοκαναλική εκδοχή του στην περίπτωση των 64Gbaud. Στην προσομοίωση του ιδιαίτερα απαιτητικού σενάριο των 100Gbaud, η συνολική βελτίωση είναι μεν αρκετά μικρότερη από τα προηγούμενα, παρόλα αυτά το πολυκαναλικό bi-RNN διατηρεί ακόμη και εδώ την υπεροχή του έναντι του μονοκαναλικού. Τέλος αξιοσημείωτη είναι η αδυναμία του FDE ισοσταθμιστή αλλά και των Volterra 1^{ου} και 3^{ου} βαθμού να ανταποκριθούν επαρκώς στο περιβάλλον υψηλής μη-γραμμικότητας.



Σχήμα 4.25: Το BER ως συνάρτηση της οπτικής ισχύος για ένα κανάλι (περιπτώσεις FDE, VLNE, bi-RNN) και εξισορρόπησης πολλαπλών καναλιών (bi-RNN) στα 64 Gbaud (α) και 100 Gbaud (β) αντίστοιχα. Αποτυπώνεται μόνο το κεντρικό κανάλι ενώ τα γειτονικά κανάλια δείχνουν παρόμοιες επιδόσεις. (Το SCE υποδηλώνει μονοκαναλική εξισορρόπηση και το MCE πολυκαναλική).

4.4.3.2 Εξάρτηση του bi-RNN από τον αριθμό των κρυφών μονάδων και των εποχών εκπαίδευση

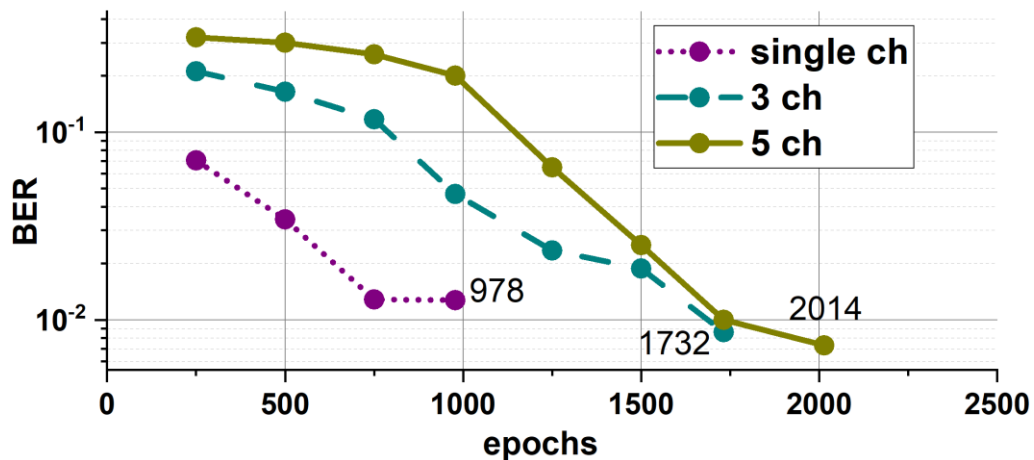
Με σκοπό τη βελτιστοποίηση της ικανότητας εξισορρόπησης του RNN όπως αυτή εκφράζεται με τον αριθμό των κρυφών μονάδων (h.u) αποδεικνύεται το εύλογο γεγονός πως η σταδιακή προσθήκη h.u είναι απαραίτητη καθώς αυξάνεται ο αριθμός των καναλιών προς επεξεργασία. Αυτό αποτυπώνεται στο σχήμα 4.26 από το οποίο προκύπτει το συμπέρασμα πως απαιτούνται 16, 18 και 20 h.u για την επίτευξη του ελάχιστου σφάλματος BER στις περιπτώσεις του 1^{ος}, των 3^{ών} και των 5 καναλιών αντίστοιχα. Αξιοσημείωτο είναι πως οι επιδόσεις των δικτύων δεν βελτιώνονται με την περαιτέρω αύξηση των h.u. Πρέπει να σημειωθεί ότι η ελαφρά αύξηση των h.u. δεν αποτελεί σοβαρό μειονέκτημα όσον αφορά την υπολογιστική πολυπλοκότητα. Επιπλέον, για στατικά συστήματα μετάδοσης WDM όπου τα κανάλια διάδοσης παραμένουν σταθερά με την πάροδο του χρόνου, η ουσιαστική αύξηση της πολυπλοκότητας της εκπαίδευσης δεν αποτελεί μείζον ζήτημα, καθώς η εκπαίδευση δεν αναμένεται να λαμβάνει χώρα πολύ συχνά σε ένα πραγματικό σενάριο [1].



Σχήμα 4.26: Το BER ως συνάρτηση των κρυφών μονάδων (hidden units - h.u.) του bi-RNN

Ομοίως, η εκπαίδευση του πολυκαναλικού δικτύου απαιτεί μεγαλύτερο αριθμό εποχών σε σύγκριση με το σενάριο μιας περίπτωσης, (βλ. σχήμα 4.27). Αυτό μπορεί να εξηγηθεί από το γεγονός ότι τα πρόσθετα κανάλια αυξάνουν τις πληροφορίες στην είσοδο και

επομένως πρέπει να χρησιμοποιηθεί μεγαλύτερος αριθμός βαρών για να πραγματοποιηθεί επαρκής εξισορρόπηση. Αυτός είναι ο λόγος που το μονοκαναλικό σενάριο απαιτεί 978 εποχές εκπαίδευσης ενώ το πολυκαναλικό 1732 και 2014 εποχές για 3 και 5 κανάλια αντίστοιχα όταν η εκπαίδευση πραγματοποιείται με το μηχανισμό του πρόωρου τερματισμού.

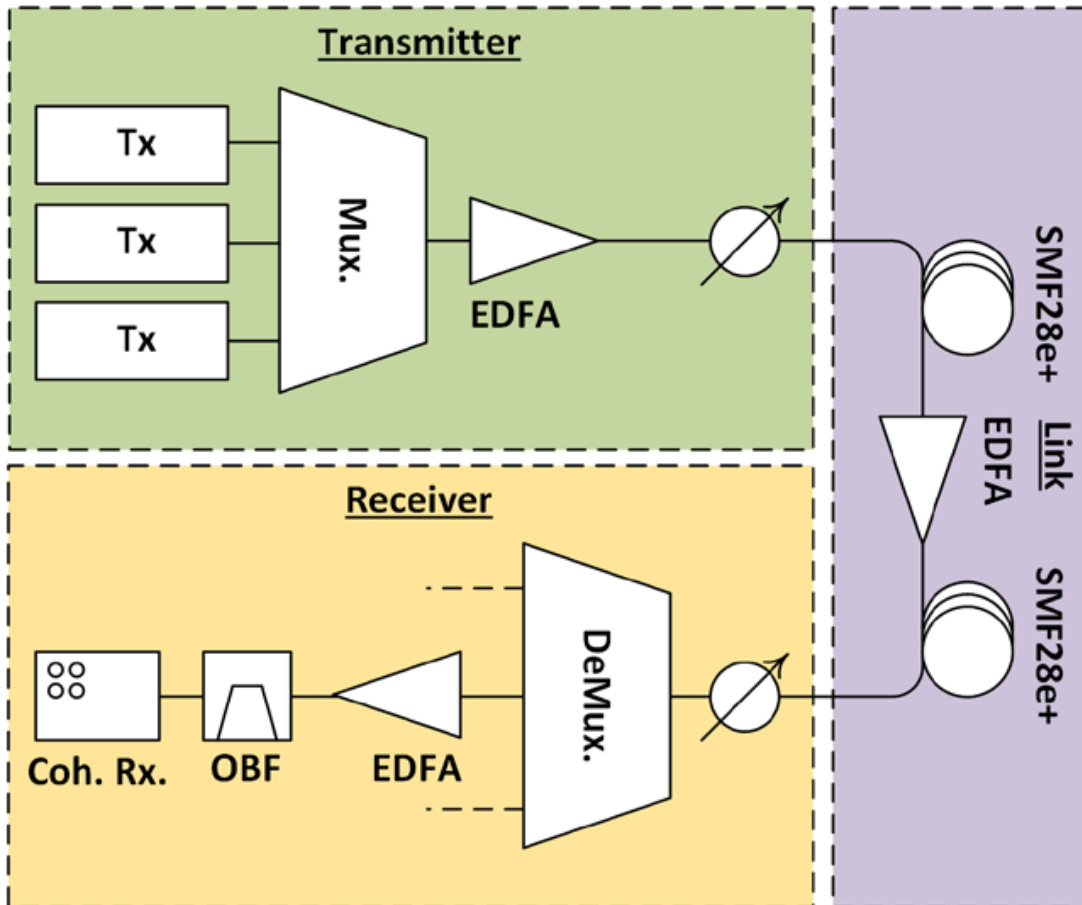


Σχήμα 4.27: Το BER ως συνάρτηση των εποχών εκπαίδευσης (epochs). Απαιτούνται 978, 1732 και 2014 εποχές εκπαίδευσης για 1, 3 και 5 κανάλια αντίστοιχα

4.4.4 Πειραματική επιβεβαίωση αποτελεσμάτων πολυκαναλικής επεξεργασίας βασισμένης σε RNN

4.4.4.1 Πειραματική διάταξη

Η πειραματική διάταξη του συστήματος διάδοσης που χρησιμοποιούμε απεικονίζεται στο σχήμα 4.28 και υλοποιήθηκε από το Πανεπιστήμιο του Southampton στο πλαίσιο της συνεργασίας μας τόσο στο ερευνητικό έργο NOOK όσο και ως αποτέλεσμα του κοινού ενδιαφέροντος μας για την εφαρμογή σύγχρονων τεχνικών επεξεργασίας για την αντιμετώπιση των μη-γραμμικών φαινομένων διάδοσης. Αποτελείται από τρεις



Σχήμα 4.28: Το πειραματικό σύστημα διάδοσης που χρησιμοποιεί δύο τμήματα μήκους 90 km του NDFP και υλοποιήθηκε στις υποδομές σκοτεινής ίνας που διαθέτει το Πανεπιστήμιο του Southampton

ανεξάρτητους πομπούς, με τον καθένα να χρησιμοποιεί ένα laser εύρους γραμμής περίπου 25 kHz, ακολουθούμενο από διαμορφωτές IQ που οδηγούνται από γεννήτριες τυχαίων κυματομορφών (Arbitrary Waveform Generator - AWG). Το τελικό αποτέλεσμα είναι τρία κανάλια, CH1, CH2 και CH3, που βρίσκονται σε συχνότητες 193,5 THz, 193,6 THz και 193,7 THz, αντίστοιχα. Λόγω του περιορισμού της μνήμης της γεννήτριας αυθαίρετης κυματομορφής (DAC) και της χρήσης του radix-2 FFT για τη μορφοποίηση των παλμών, το κάθε κανάλι παράγει 2^{16} σύμβολα ασυσχέτιστων μεταξύ τους, τυχαίων δεδομένων, (χρησιμοποιώντας όπως και στην προσομοίωση τη γεννήτρια τυχαίων αριθμών Mersenne Twister), σε ρυθμό 22.5Gb, χρησιμοποιώντας διαμόρφωση QPSK μίας πόλωσης.

Με τη βοήθεια ελεγκτών πόλωσης τα τρία κανάλια ευθυγραμμίζονται ως προς την πολωτική τους κατάσταση προτού πολυπλεχθούν και ενισχυθούν με οπτικό ενισχυτή ερβίου (Erbium Doped Fibre Amplifier - EDFA). Ένας εξασθενητής επιτρέπει τον έλεγχο της ισχύος που διοχετεύεται στην οπτική ζεύξη συνολικού μήκους 180 km. Η διάδοση πραγματοποιείται σε ένα τμήμα της Εθνικής Εγκατάστασης Σκοτεινών Οπτικών Ινών του Ηνωμένου Βασιλείου (National Dark Fibre Facility, NDFF). Η οπτική ίνα τύπου SMF28e+ του NDFF που χρησιμοποιείται, αποτελείται από δύο τμήματα των 90 km που αναπτύσσονται στο πεδίο στη διαδρομή Southampton – Reading - Southampton. Τα σήματα ενισχύονται με έναν ενδιάμεσο EDFA, έτσι ώστε η συνολική ισχύς διάδοσης των δύο τμημάτων της SMF28e+ να είναι ίδια.

Μετά τη διάδοση, τα σήματα οδηγούνται στον δέκτη, ο οποίος περιέχει έναν ακόμη εξασθενητή για να διευκολύνει τον χαρακτηρισμό, έναν αποπολυπλέκτη για την επιλογή του υπό μέτρηση καναλιού και έναν EDFA για την οπτική προενίσχυση του σήματος. Στη συνέχεια ένα ζωνοπερατό οπτικό φίλτρο (Optical Bandpass Filter - OBF) χρησιμοποιείται για την απόρριψη του θορύβου εκτός της επιθυμητής ζώνης προτού ληφθεί τελικά το σήμα χρησιμοποιώντας έναν σύμφωνο δέκτη.

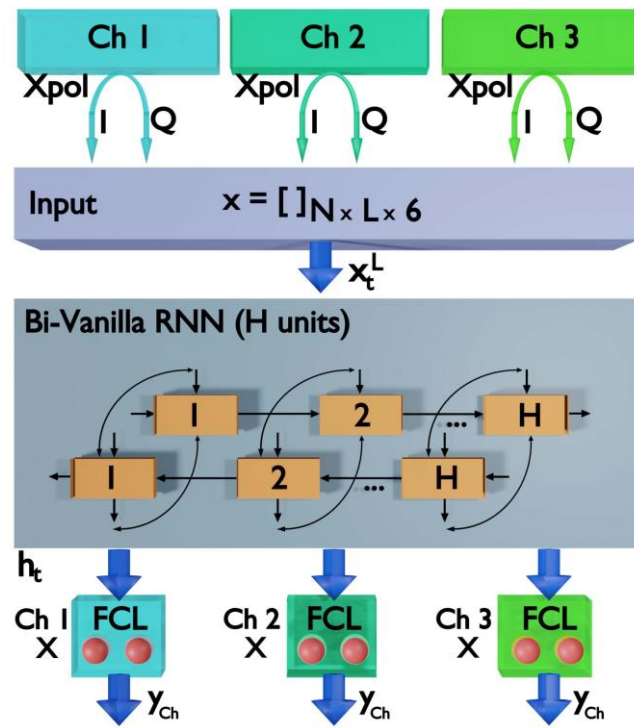
Ο σύμφωνος δέκτης είναι τυπικής κατασκευής, αποτελούμενος από ένα οπτικό υβρίδιο ακολουθούμενο από 4 ισορροπημένους φωτοδέκτες των οποίων η έξοδος αποθηκεύεται σε έναν ψηφιακό παλμογράφο με ρυθμό 40 GSa/s.

Ο τοπικός ταλαντωτής και η φέρουσα για κάθε κανάλι προέρχονται από τα laser του πομπού. Μόλις καταγραφούν οι κυματομορφές των καναλιών (ένα κανάλι κάθε φορά) για διάφορες τιμές οπτικής ισχύος διάδοσης, υποβάλλονται σε επεξεργασία εκτός σύνδεσης.

4.4.4.2 Ψηφιακή Επεξεργασία Σήματος

Πριν από οποιαδήποτε μη γραμμική μετα-επεξεργασία ή αποδιαμόρφωση, πραγματοποιείται αποπολυπλεξία πόλωσης, αντιστάθμιση της χρωματικής διασποράς με τη χρήση FDE και ισοστάθμιση με τη βοήθεια προσαρμοστικού φίλτρου constant-

modulus algorithm (CMA) των 15 taps. Στη συνέχεια εκτελούμε συγχρονισμό φέρουσας, εξαγωγή ρολογιού και τέλος δειγματοληψία.



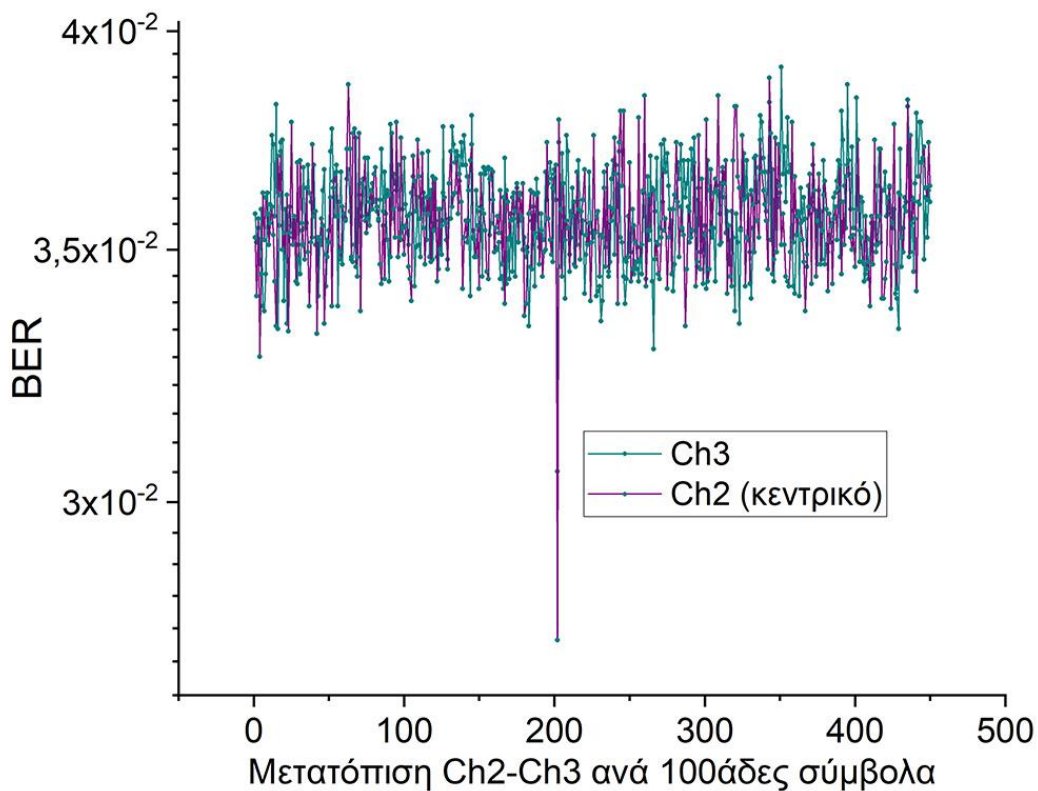
Σχήμα 4.29: Διάγραμμα του bi-VRNN ισοσταθμιστή επεξεργασίας 3 καναλιών

Τελικά, ένα δείγμα ανά σύμβολο αποστέλλεται στη μονάδα επεξεργασίας bi-RNN. Η μόνη διαφορά στην περίπτωση της επεξεργασίας με τον ισοσταθμιστή τύπου DBP είναι η αφαίρεση του μπλοκ τύπου FDE, το οποίο αντίθετα ενσωματώνεται σε κάθε βήμα της διαδικασίας του DBP. Το DBP εκτελείται με 2 δείγματα ανά σύμβολο και 20 βήματα ανά τμήμα διάδοσης, πριν από την εφαρμογή του CMA αλγορίθμου.

Χρησιμοποιούμε τα 3 γειτονικά κανάλια μίας πόλωσης ως είσοδο στο νευρωνικό μας δίκτυο, τροποποιώντας κατάλληλα τη δομή του σχήματος 4.21 από 16-QAM διπλής πόλωσης σε QPSK μονής πόλωσης, όπως φαίνεται στο σχήμα 4.29. Η είσοδος του x είναι ένα διάνυσμα $N \times L \times 6$ συμβόλων QPSK όπου το N υποδηλώνει τον συνολικό αριθμό των συμβόλων εισόδου και το $L = 2k + 1$ αντιπροσωπεύει το μήκος της εισόδου. Η είσοδος x έχει συνολικά 6 χαρακτηριστικά: το I και το Q καθενός από τα 3 κανάλια. Η είσοδος x οδηγείται στο bi-VRNN των H κρυφών μονάδων. Η έξοδος του x_t^L αποστέλλεται

στη συνέχεια παράλληλα σε 3 πλήρως συνδεδεμένα στρώματα (FCL), τα οποία αντιστοιχούν σε καθένα από τα 3 κανάλια. Κάθε FCL αποτελείται από δύο νευρώνες, που αντιπροσωπεύουν τα I και Q του QPSK με την έξοδο του y να είναι ένα διάνυσμα $N \times L \times 2$. Το bi-VRNN εκπαιδεύεται ως πρόβλημα παλινδρόμησης με τα ίδια χαρακτηριστικά των προσομοιώσεων που παρουσιάστηκαν στις προηγούμενες παραγράφους. Χρησιμοποιήθηκαν 50.000 σύμβολα για εκπαίδευση, 5.000 σύμβολα για επικύρωση και 10.000 σύμβολα για εκτίμηση αποτελεσμάτων σε άγνωστα δεδομένα.

Η πειραματική διαδικασία παρείχε δεδομένα για ένα ευρύ φάσμα τιμών ισχύος εισόδου και δέκτη ανά κανάλι. Αυτά τα δεδομένα επιτρέπουν την άμεση σχεδίαση της καμπύλης BER για κάθε κανάλι όταν λαμβάνει χώρα μόνο γραμμική εξισορρόπηση ή ακόμα και μετά από μη-γραμμική εξισορρόπηση ενός καναλιού με τη βοήθεια μονοκαναλικού bi-RNN. Αντίθετα, η εξισορρόπηση πολλαπλών καναλιών δεν είναι μια τετριμμένη εργασία, αφού

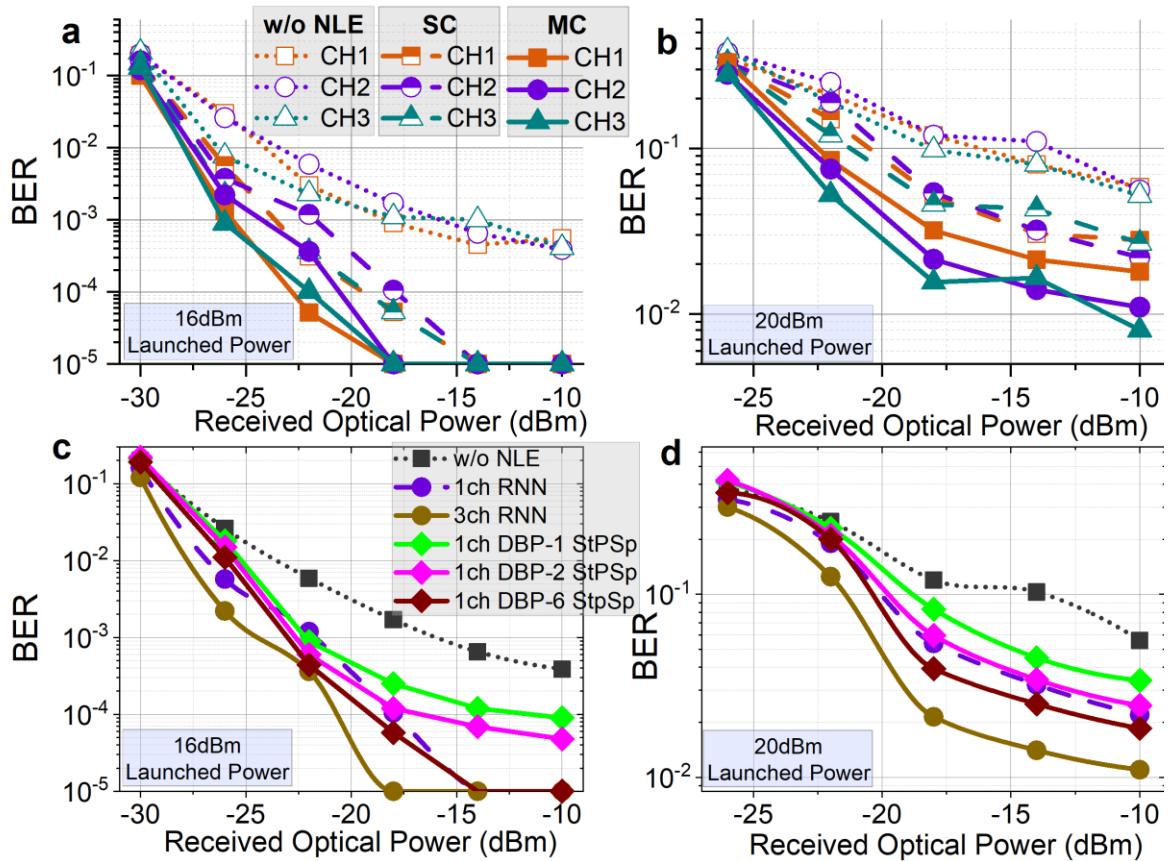


Σχήμα 4.30: Brute force τεχνική χρονικής ευθυγράμμισης με κοινή εξισορρόπηση bi-VRNN δύο καναλιών

τα δεδομένα που εισέρχονται στον αλγόριθμο bi-RNN θα πρέπει να συλλέγονται με σύγχρονο τρόπο από τον παλμογράφο κάτι που δεν ήταν εφικτό να συμβεί στην πράξη. Προκειμένου να ευθυγραμμίσουμε τα τρία ανεξάρτητα κανάλια, τα οποία έχουν ληφθεί με ασύγχρονο τρόπο, πρέπει να χρησιμοποιήσουμε μια χρονοβόρα τεχνική εξαντλητικής αναζήτησης brute-force για να προσδιορίσουμε τον ακριβή χρονισμό τους. Για το λόγο αυτό εφαρμόζουμε πουλκαναλική εκπαίδευση αρχικά ανά δύο κανάλια, για όλους τους πιθανούς συνδυασμούς της χρονικής τους θέσης. Όπως απεικονίζεται στο παράδειγμα του σχήματος 4.30, εστιάζοντας στις διακαναλικές αλληλεπιδράσεις μεταξύ των καναλιών Ch2 (κεντρικό) και Ch3 και χρησιμοποιώντας δικαναλικό bi-VRNN ισοσταθμιστή για διαφορετικές χρονικές μετατοπίσεις, κατορθώνουμε να εντοπίσουμε ένα σημείο όπου το σφάλμα BER είναι προφανώς βελτιωμένο, ως αποτέλεσμα της χρονικής ευθυγράμμισης των δύο καναλιών. Ενδεικτικά, σε ισχύ εκπομπής 20 dBm και ισχύ λήψης -10 dBm, παρατηρούμε αισθητή μείωση από μέση τιμή 3.6×10^{-2} του BER σε χρονικά ασυσχέιστα κανάλια, στην τιμή 2.5×10^{-2} . Αυτό επιβεβαιώνει πειραματικά την ικανότητα του μοντέλου bi-RNN να αναγνωρίζει τις συσχετίσεις μεταξύ των καναλιών. Πρέπει να σημειωθεί ότι σε ένα πραγματικό σύστημα πολυκαναλικής ισοστάθμισης, η ανίχνευση των καναλιών θα είναι σύγχρονη, επομένως τα δεδομένα θα αποστέλλονται απευθείας στο bi-RNN για εκπαίδευση και εξαγωγή συμπερασμάτων χωρίς την ανάγκη χρονικής ευθυγράμμισης.

4.4.4.3 Πολυκαναλική και μονακαναλική bi-RNN ισοστάθμιση έναντι DBP

Στο σχήμα 4.31, μπορεί κανείς να μελετήσει την απόδοση του bi-VRNN ισοσταθμιστή στη μονοκαναλική και πολυκαναλική εκδοχή του με βάση τα πειραματικά δεδομένα. Η απόδοση του πολυκαναλικού bi-VRNN παρουσιάζει βελτίωση της τιμής BER έως και μισής τάξης μεγέθους σε σύγκριση με την επεξεργασία ενός καναλιού. Για παράδειγμα, λαμβάνοντας υπόψη το κανάλι 2 (ch2, κεντρικό) σε ισχύ 16 dBm (σχ. 4.31α), ο ρυθμός σφαλμάτων BER βελτιώνεται από 1×10^{-3} σε 4×10^{-4} . Ακόμη και στο σημείο λειτουργίας της ισχυρότερης μη γραμμικής αλληλεπίδρασης (ισχύς 20 dBm, σχ.4.31β), ο πολυκαναλικός ισοσταθμιστής μπορεί να προσφέρει ικανοποιητική βελτίωση στην απόδοση του BER σε όλες τις τιμές της λαμβανόμενης ισχύος. Στα σχήματα 4.30α και



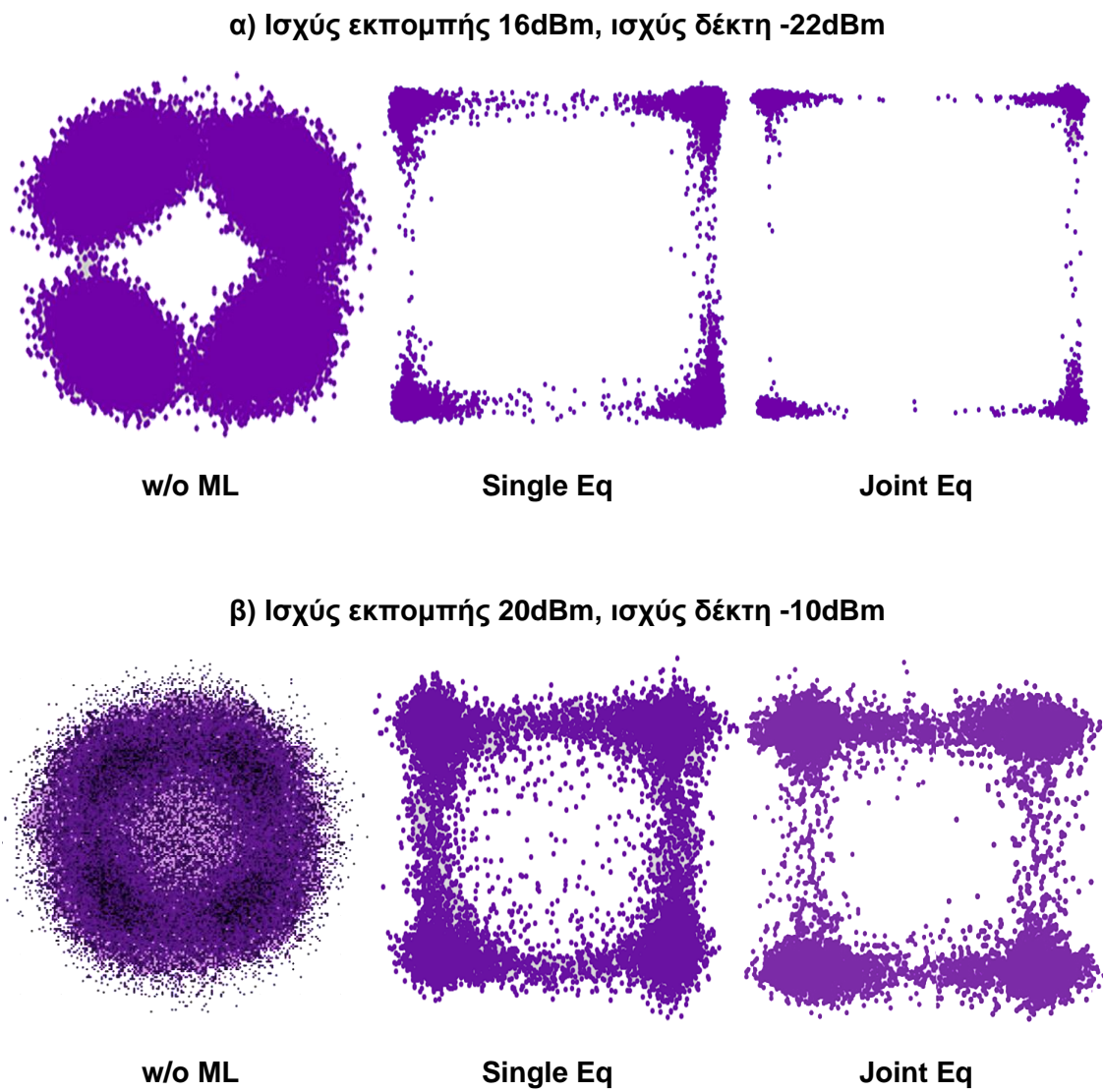
Σχήμα 4.31: Οι επιδόσεις BER ως συνάρτηση της οπτικής ισχύος του δέκτη για μονοκαναλική και πολυκαναλική ανίχνευση όταν η ισχύς εκπομπής ανά κανάλι είναι 16dBm (α),(γ) και 20dBm (β),(δ) αντίστοιχα. Στα (α) και (β) απεικονίζονται οι επιδόσεις όλων των διαδιδόμενων και ανιχνευόμενων καναλιών στην εφαρμογή του μεμονωμένου ή κοινού RNN, ενώ στα (γ) και (δ) τα RNN, DBP ή χωρίς καμία μη γραμμική εξίσωση (w/o NLE) μόνο στο κεντρικό κανάλι

4.31β παρατηρούμε ότι και τα 3 κανάλια επωφελούνται σχεδόν εξίσου από την κοινή επεξεργασία. Στο σενάριο της ισχύος εκπομπής των 20 dBm και λαμβανόμενης ισχύος των -10 dBm, παρατηρούμε μείωση της τιμής BER από 2.8×10^{-2} σε 1.8×10^{-2} για το κανάλι 1 και από 2.7×10^{-2} σε 8×10^{-3} για το κανάλι 3, με την εφαρμογή μονοκαναλικής και πολυκαναλικής ισοστάθμισης αντίστοιχα.

Σε αυτό το έντονα μη γραμμικό περιβάλλον που προκαλείται από την υψηλή ισχύ, ο μη γραμμικός θόρυβος φάσης είναι πολύ ισχυρός και ο ισοσταθμιστής δεν μπορεί να προσφέρει βελτίωση BER μιας τάξης μεγέθους. Αντίθετα, όταν η ισχύς εκκίνησης είναι 16

dBm, η βελτίωση είναι πολύ σημαντική και η απόδοση όσον αφορά τη σχετική βελτίωση του BER είναι σημαντικό για όλα τα κανάλια. Αν και οι πειραματικές συνθήκες μοιάζουν με αυτές ενός συστήματος χαμηλής διασποράς, καθώς χρησιμοποιούνται μόνο δύο spans, ο ισοσταθμιστής έχει καλύτερη απόδοση από αυτή που αξιολογήθηκε στις προσομοιώσεις, καθώς όλα τα κανάλια συν-διάδοσης εισέρχονται στον ισοσταθμιστή και όχι μόνο ένα τμήμα τους (3 από 9 στο παράδειγμα προσομοίωσης). Το γεγονός ότι το μονοκαναλικό RNN είναι πολύ κοντά στον ισοσταθμιστή τύπου DBP με πολλαπλά βήματα ανά διάστημα (σχήμα 4.31γ και 4.31δ), δείχνει ότι αποτελεί έναν σχεδόν βέλτιστο μη γραμμικό ισοσταθμιστή ενδοκαναλικών μη-γραμμικών παρεμβολών. Η περαιτέρω βελτίωση που προσφέρει το μοντέλο RNN των 3 καναλιών μπορεί επομένως να αποδοθεί στον αποτελεσματικό μετριασμό των XPM φαινομένων, παρά στην υπολειπόμενη αντιστάθμιση SPM. Στο σχ. 4.31δ, για την ισχύ εκπομπής των 20 dBm, η αναλογία οπτικού σήματος προς θόρυβο (OSNR) σε σύγκριση με το bi-RNN ενός καναλιού είναι σχεδόν 7.5 dB όπου το BER βρίσκεται στο 2×10^{-2} , ενώ για την ισχύ εκπομπής των 16 dBm (σχήμα 4.31γ), το κέρδος OSNR είναι κοντά στα 2.5 dB όπου το BER βρίσκεται στο 10^{-3} , αποδεικνύοντας ότι ο προτεινόμενος ισοσταθμιστής αποδίδει καλύτερα σε εξαιρετικά μη γραμμικά περιβάλλοντα.

Στο σχήμα 4.32 φαίνονται οι QPSK αστερισμοί για ενδεικτικές τιμές ισχύος από τους οποίους μπορεί να φανεί η σημαντική βελτίωση που παρέχει η ισοστάθμιση του μονοκαναλικού bi-VRNN και αντίστοιχα η περαιτέρω βελτίωσή της όταν εφαρμόζεται ο πολυκαναλικός bi-VRNN. Χαρακτηριστική είναι επίσης η ελάττωση του φαινομένου του “παραθύρου κελιού φυλακής” στην περίπτωση της πολυκαναλικής ισοστάθμισης, γεγονός που οφείλεται στη μείωση της αβεβαιότητας του αλγορίθμου MSE όταν είναι διαθέσιμη η επιπλέον πληροφορία των γειτονικών καναλιών.



Σχήμα 4.32: QPSK αστερισμοί

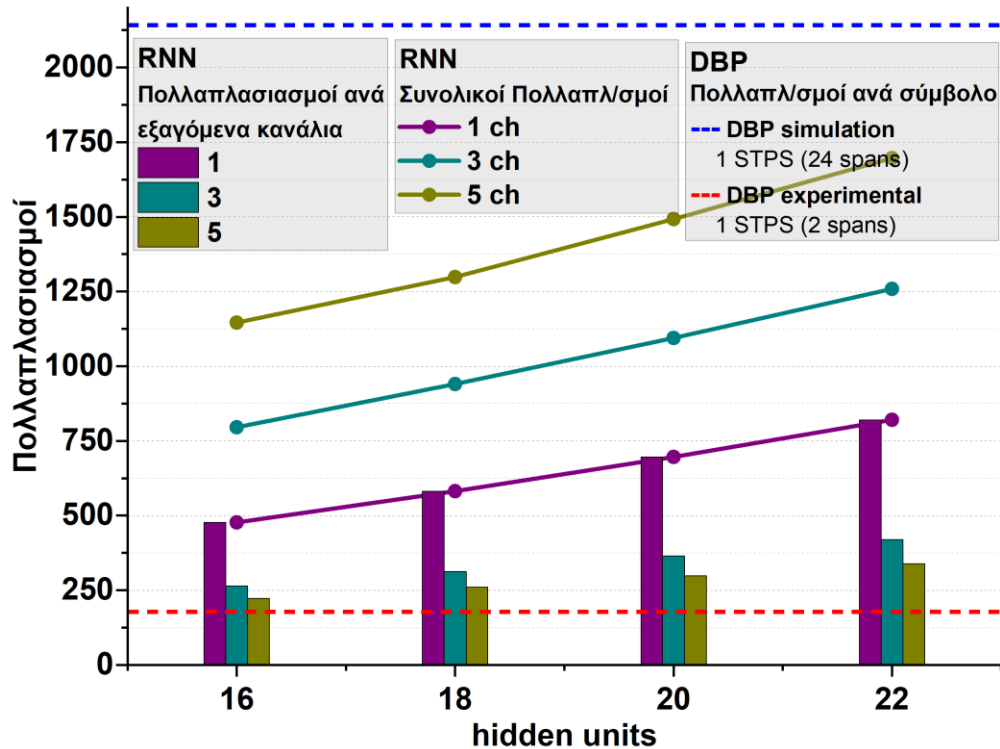
4.4.5 Ανάλυση Υπολογιστικής Πολυπλοκότητας

Για να αξιολογήσουμε την υπολογιστική πολυπλοκότητα της πολυκαναλικής ισοστάθμισης, υπολογίζουμε τον αριθμό των απαραίτητων πολλαπλασιασμών για την αξιολόγηση του νευρωνικού δικτύου, δεδομένων των βαρών και όχι αυτών της εκπαίδευσης. Ο αριθμός των πολλαπλασιασμών που ακολουθούν την προσέγγιση πολλών σε πολλά, εκφράζεται από τη σχέση 4.17, η οποία στην περίπτωση του bi-VRNN ($B=1$) και της μεθόδου της παλινδρόμησης $b=1$ που χρησιμοποιούμε έναντι της μεθόδου ταξινόμησης που χρησιμοποιήθηκε στις προηγούμενες παραγράφους, απλοποιείται ως εξής :

$$bi - VRNN_{mult} = 2(FH + H^2)L + 2HLy \quad (4.21)$$

Όπου F , y είναι τα χαρακτηριστικά εισόδου και τα διανύσματα εξόδου αντίστοιχα. Οι παράμετροι F και y είναι ίσες με $2m$, αφού έχουμε 2 εισόδους και εξόδους (στοιχεία I/Q) για κάθε κανάλι, όπου m είναι ο αριθμός των καναλιών που ανιχνεύονται από κοινού. Χρησιμοποιούμε $H=16$, 18 και 20 κρυφές μονάδες RNN για απλή και κοινή ανίχνευση 3 και 5 γειτονικών καναλιών. Λαμβάνοντας υπόψη το εύρημα ότι τουλάχιστον το 80% των γειτονικών συμβόλων εμφανίζουν παρόμοιο BER στην προσέγγιση πολλά προς πολλά [2] και δεδομένου ότι έχουμε χρησιμοποιήσει στην είσοδο του δικτύου μια λέξη μήκους 51 συμβόλων, εξάγουμε τα 41 κεντρικά για να υπολογίσουμε τον αριθμό των πολλαπλασιασμών. Στο σχήμα 4.33 επιδεικνύεται ο συνολικός αριθμός πολλαπλασιασμών για τις περιπτώσεις μονακαναλικού και πολυκαναλικών bi-VRNN ισοσταθμιστών συναρτήσει των κρυφών μονάδων του bi-VRNN. Μπορεί κανείς να δει ότι παρόλο που ο αριθμός των πολλαπλασιασμών αυξάνεται με την προσθήκη περισσότερων καναλιών εισόδου/εξόδου (χαρακτηριστικών), τελικά ο αριθμός των πολλαπλασιασμών ανά σύμβολο μειώνεται δραστικά. Διαιρώντας την εξίσωση 4.21 με τον αριθμό των τρόπων πόλωσης και πολλαπλασιάζοντας με τον αριθμό των WDM καναλιών, υπολογίζουμε 478 πολλαπλασιασμούς ανά ανιχνευμένο σύμβολο (mps) για τον μονοκαναλικό ισοσταθμιστή των 16 h.u, 313 mps για πολυκαναλικό των 3 καναλιών με 18 h.u. (34,38% μείωση πολυπλοκότητας σε σύγκριση με ένα κανάλι) και 299 mps για πολυκαναλικό των 5 καναλιών με 20 h.u. (37,5% μείωση πολυπλοκότητας σε σύγκριση

με το μονοκαναλικό). Η μείωση της πολυπλοκότητας του ισοσταθμιστή πολλαπλών καναλιών σε σχέση με αυτή του αντίστοιχου μονοκαναλικού στα σενάρια που εξετάσαμε μπορεί να φτάσει έως και 58,7% στην περίπτωση των 22 h.u. Συγκεκριμένα υπολογίζονται 821mpcs για το 1 κανάλι σε σχέση με 339 mpcs για την ανίχνευση 5 καναλιών (σχήμα 4.33).



Σχήμα 4.33: Αριθμός Πολλαπλασιασμών σε σχέση με τον αριθμό των κρυφών μονάδων για το μονοκαναλικό και πολυκαναλικό ισοσταθμιστή

Η ερμηνεία της μείωσης της υπολογιστικής πολυπλοκότητας στην περίπτωση του πολυκαναλικού ισοσταθμιστή RNN βρίσκεται στη σχέση 4.21 που αποκαλύπτει ότι η υπολογιστική πολυπλοκότητα καθορίζεται κατά κύριο λόγο από τον αριθμό των h.u. Στην πραγματικότητα, εκμεταλλευόμενοι την ήδη πολύπλοκη δομή του δικτύου και προσθέτοντας μια εύλογη επιβάρυνση (πρόσθετες εισοδοι και έξοδοι των καναλιών), καταφέρνουμε να μειώσουμε την ψηφιακή επεξεργασία ανά σύμβολο, βελτιώνοντας παράλληλα σημαντικά τις επιδόσεις (μείωση της τιμής BER). Σε αυτό το σημείο θα πρέπει

να τονιστεί ότι τόσο σε επίπεδο απόδοσης όσο και σε εφαρμογή της σε πραγματικό υλικό, η πρόταση της πολυκαναλικής επεξεργασίας RNN έχει συγκεκριμένους περιορισμούς. Τα σχήματα 4.22 και 4.24 δείχνουν ότι η χρήση περισσότερων των 3 καναλιών συμβάλλει λιγότερο στη βελτίωση του BER. Αν και είναι εκτός του πεδίου αυτής της μελέτης, η σχεδίαση των ισοσταθμιστών RNN σε συστήματα οπτικών επικοινωνιών είναι μια πρόκληση που ήδη απασχολεί την ερευνητική κοινότητα [150]. Πιστεύουμε ότι η σχεδίαση και υλοποίηση σε πραγματικό υλικό, πχ Field Programmable Gate Array, FPGA ενός ισοσταθμιστή bi-RNN τριών καναλιών που απαιτεί έναν λογικό αριθμό κρυφών μονάδων είναι εφικτή.

Όσον αφορά την υπολογιστική πολυπλοκότητα του RNN στην πειραματική διάταξη που παρουσιάστηκε, υπολογίζουμε 796 mps για τον μονοκαναλικό ισοσταθμιστή των 16 h.u. και 448 mps για τον πολυκαναλικό ισοσταθμιστή των 3 καναλιών και των 18 h.u. Αυτό ισοδυναμεί με 43,75% μείωση της πολυπλοκότητας σε σύγκριση με το RNN ενός καναλιού.

Για τον αλγόριθμο DBP η πολυπλοκότητα ανά σύμβολο μπορεί να εκτιμηθεί σύμφωνα με τον τύπο 4.7. Λαμβάνοντας υπόψη $N=256$ και $N_s=30$, με 1, 2 και 6 βήματα ανά τμήμα για τα 2 spans υπολογίζουμε 178, 357 και 1070 πολλαπλασιασμούς αντίστοιχα. Ακόμα κι αν μειώσουμε τα βήματα ανά τμήμα σε 4 και το μέγεθος του πίνακα στο μετασχηματισμού FFT σε 128, θυσιάζοντας σημαντικά την απόδοση, ο DBP ισοσταθμιστής θα απαιτούσε τουλάχιστον 918 mps για ένα μόνο κανάλι.

Προκειμένου να επιτευχθεί μια δίκαιη σύγκριση μεταξύ της προσέγγισης bi-RNN και DBP, είναι επιτακτική ανάγκη να ενσωματωθεί στην πρώτη η μέτρηση των πολλαπλασιασμών ανά σύμβολο που απαιτείται από τον ισοσταθμιστή τύπου FDE, σύμφωνα με τη σχέση 4.8.

Έτσι, η προσέγγιση του πολυκαναλικού bi-RNN ισοσταθμιστή με 18 h.u (απαιτεί 529 πολλαπλασιασμούς) θα προσφέρει πάνω από 42% μείωση της πολυπλοκότητας ανά κανάλι σε σχέση με τον ισοσταθμιστή DBP (918 πολλαπλασιασμούς), όταν ενσωματώνονται 3 κανάλια, και 88% μείωση από την εφαρμογή μας εδώ 20 step-per-span DBP (4528 πολλαπλασιασμούς).

Πίνακας 4.5 : Ανάλυση υπολογιστικής πολυπλοκότητας

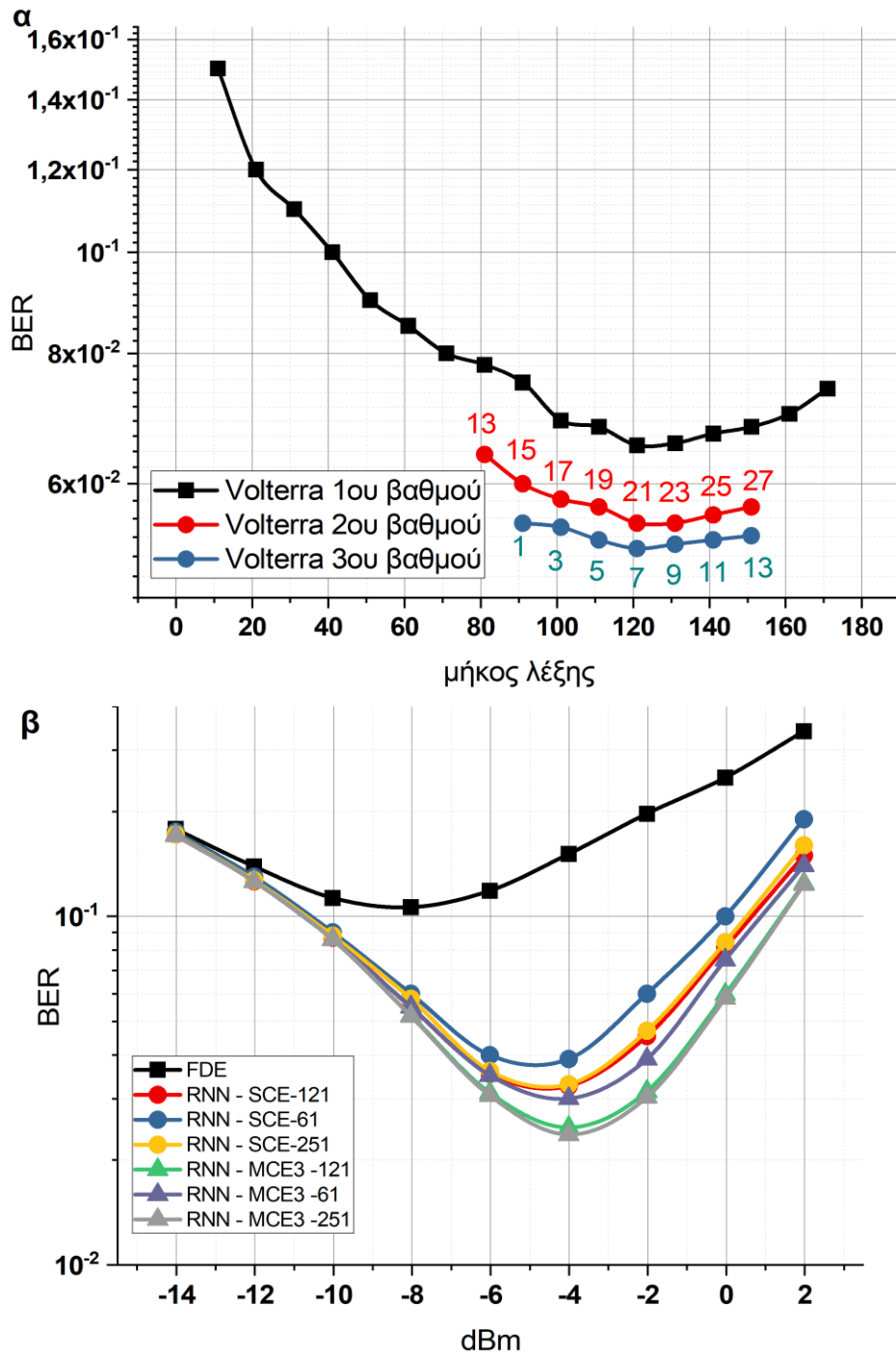
Πειραματικό (2 spans, QPSK μιας πάλωσης)	Πολλαπλασιασμοί ανά σύμβολο
1 κανάλι DBP 1 step per span	178
1 κανάλι DBP 2 steps per span	357
1 κανάλι DBP 6 steps per span	1070
1 κανάλι bi-VRNN, 16 h.u	877 (796 RNN+81 FDE)
3 κανάλια bi-VRNN, 18 h.u	529 (447 RNN + 81 FDE)
Προσομοίωση (24 spans, 16-QAM pol-mux)	Πολλαπλασιασμοί ανά σύμβολο
1 κανάλι DBP 1 step per span	2141
1 κανάλι DBP 2 steps per span	4282
1 κανάλι DBP 6 steps per span	12845
1 κανάλι bi-VRNN, 16 h.u	558 (477 RNN + 81 FDE)
3 κανάλια bi-VRNN, 18 h.u	394 (313 RNN + 81 FDE)
5 κανάλια bi-VRNN, 20 h.u	380 (299 RNN + 81 FDE)

Ο πίνακας 4.5 δείχνει ξεκάθαρα πως ακόμη και για δύο spans, ο bi-VRNN επεξεργαστής είναι λιγότερο περίπλοκος από το DBP των 6 STPS (σχεδόν 200% πιο αποδοτικός) και το πλεονέκτημα πολυπλοκότητάς του είναι ακόμη μεγαλύτερο (~ 550%) όταν ο αριθμός των spans αυξάνεται σε 24 όπως στην περίπτωση προσομοίωσης που έχουμε ερευνήσει σε αυτό το κεφάλαιο, ακόμη και αν λάβουμε υπόψη DBP του 1^{ος} STPS.

Είναι σημαντικό να τονιστεί ότι η υπολογιστική πολυπλοκότητα του bi-RNN δεν επηρεάζεται απόλυτα από το μήκος διάδοσης ενώ το DBP παρουσιάζει σημαντική κλιμάκωση της πολυπλοκότητας με την αύξηση του αριθμού των τμημάτων διάδοσης. Σίγουρα, για σενάρια διάδοσης μεγάλων αποστάσεων που αποτελούνται από πολλά spans, η προσέγγισή bi-RNN θα είναι πολύ λιγότερο περίπλοκη από την DBP όπως φαίνεται στον Πίνακα 4.5. Επιπλέον, τόσο ο ισοσταθμιστής DBP όσο και ο ισοσταθμιστής FDE απαιτούν τουλάχιστον 2 sps, σε αντίθεση με το bi-RNN, το οποίο δείχνει ικανοποιητική απόδοση με μόνο 1sps.

Συνεπώς τα πειραματικά αποτελέσματα και τα αποτελέσματα της προσομοίωσης έδειξαν υπεροχή στη βελτίωση του BER σε σύγκριση με το DBP πλήρους πεδίου 3 καναλιών και ένα σημαντικό πλεονέκτημα πολυπλοκότητας που μπορεί να είναι μια μείωση της πολυπλοκότητας της τάξης του 500% σε συστήματα μετάδοσης μεγάλων αποστάσεων.

Τέλος για τα επιπρόσθετα σενάρια προσομοίωσης και τη σύγκριση με τους ισοσταθμιστές Volterra, για λόγους ισότιμης σύγκρισης επιδόσεων και πολυπλοκότητας χρησιμοποιείται για την εκπαίδευση του bi-RNN μήκος λέξης των 121 συμβόλων. Ομοίως στο Volterra



Σχήμα 4.34: Επιλογή μήκους λέξης συμβόλων για τον ισοσταθμιστή τύπου Volterra (α) και bi-RNN (β)

χρησιμοποιούνται λέξη των 121 συμβόλων για την πρώτη τάξη και λέξεις των 21 και 7 συμβόλων για τη 2^η και 3^η αντίστοιχα. Η επιλογή των παραμέτρων προκύπτει από την δοκιμή που αποτυπώνεται στα σχήματα 4.34α και 4.34β, με βάση το απαιτητικό σενάριο των 100GBaud. Φαίνεται πως στο μεν Volterra το μήκος της λέξης των 121 συμβόλων για την 1^η τάξη δίνει τη χαμηλότερη τιμή BER και ακολούθως διερευνώνται οι ανώτερες τάξεις που δίνουν τις τιμές 21 και 7 αντίστοιχα. Η εκπαίδευση του bi-VRNN πραγματοποιήθηκε με 3 διαφορετικά μήκη λέξεων, των 61, 121 και 251 συμβόλων. Διακρίνεται στο σχήμα 4.34β πως η βέλτιστη επίδοση του ισοσταθμιστή δίνεται για το μήκος της λέξης των 121 συμβόλων επιβεβαιώνοντας την ίδια στιγμή πως αυτό είναι και το μήκος της μνήμης του. Για λέξεις λιγότερων συμβόλων (61) τόσο ο μονοκαναλικός όσο και ο πολυκαναλικός bi-RNN ισοσταθμιστής παρουσιάζουν χειρότερες επιδόσεις ενώ δεν φαίνεται να επωφελούνται από λέξεις περισσότερων συμβόλων (251).

Στις περιπτώσεις των ισοσταθμιστών τύπου Volterra υπολογίζουμε 242 πολλαπλασιασμούς ανά εξαγόμενο σύμβολο για το μεν 1^{ης} τάξης (γραμμικό) και 1670 για αυτόν της 3^{ης} τάξης (μη γραμμικό). Παρόλα αυτά όπως φαίνεται και στο σχήμα 4.25α οι επιδόσεις ακόμη και του ισχυρού μη γραμμικού ισοσταθμιστή τύπου Volterra 3^{ης} τάξης, δεν είναι ικανοποιητικές αφού δεν καταφέρνει να αποδώσει BER χαμηλότερο από το όριο του HD-FEC ($BER = 3.4 \times 10^{-3}$).

Αντιθέτως, οι ισοσταθμιστές βασιζόμενοι στο bi-VRNN αποδίδουν αρκετά καλύτερα. Κατά τα πρότυπα της προσέγγισης πολλά σε πολλά, αν εξαχθούν 101 σύμβολα από τα 121 της εισόδου, το μέσο (βέλτιστο) BER είναι κάτω από το όριο του HD-FEC και συγκεκριμένα 10^{-3} και 10^{-4} για τις περιπτώσεις της μονοκαναλικής και της πολυκαναλικής (3 και 5 καναλιών) RNN ισοστάθμισης.

Στον πίνακα 4.6 συγκεντρώνονται οι πολλαπλασιασμοί και η βέλτιστη επίδοση BER των συγκρινόμενων σχημάτων ισοστάθμισης του σχήματος 4.25α. Το σημαντικότερο εύρημα του πίνακα 4.6 είναι πως ο αριθμός των πολλαπλασιασμών ανά εξαγόμενο σύμβολο είναι διαρκώς μειούμενος, όσο αυξάνεται η τάξη της πολυκαναλικής ισοστάθμισης, παρά την αύξηση του αριθμού των κρυφών μονάδων που απαιτούνται για επαρκή εκπαίδευση. Συγκεκριμένα η μονοκαναλική RNN ισοστάθμιση απαιτεί 460 πολλαπλασιασμούς ανά

εξαγόμενο σύμβολο ενώ μειώνεται στους 351 και 288 για τον πολυκαναλικό ισοσταθμιστή 3^{ης} και 5^{ης} τάξης αντίστοιχα. Αυτό ισοδυναμεί με 37% μείωση της υπολογιστικής πολυπλοκότητας γεγονός που επαληθεύει την αποδοτικότερη συμπεριφορά της προτεινόμενης λύσης.

Πρέπει να τονιστεί ωστόσο ότι αυτή η πολυπλοκότητα είναι ένα πρόσθετο φορτίο σε αυτό που έχει ήδη δημιουργηθεί από την τυπική ψηφιακή επεξεργασία που συναντάται σε σύμφωνα συστήματα για τη γραμμική ισοστάθμιση (FDE, αποπολυπλεξία πόλωσης) και τον συγχρονισμό φέρουσας. Ένα τυπικό μπλοκ FDE που συνοδεύεται από προσαρμοστικούς ισοσταθμιστές θα καταναλώνει περίπου 200-360 πολλαπλασιασμούς ανά σύμβολο (40 πολλαπλασιασμοί για το FDE και $16 \times N$ για προσαρμοστικούς ισοσταθμιστές όπου N είναι ο αριθμός των taps, συνήθως στην περιοχή από 10 έως 20).

Από τα προηγούμενα γίνεται σαφές ότι η πολυκαναλική bi-RNN ισοστάθμιση μπορεί να είναι ακόμη λιγότερο πολύπλοκη από τις τεχνικές ενός καναλιού χάρη στην προσέγγιση εκπαίδευσης πολλά προς πολλά, ενώ βελτιώνει ταυτόχρονα τις επιδόσεις μειώνοντας το BER.

Χωρίς αμφιβολία, η διαδικασία της εκπαίδευσης των πολυκαναλικών bi-RNN είναι πιο απαιτητική διαδικασία και συναρτάται με τη διάσταση του νευρωνικού. Ωστόσο, στο περιβάλλον ενός συστήματος οπτικών επικοινωνιών, όπου η χρονική μεταβλητότητα του καναλιού είναι μικρή, δεν αναμένεται να επηρεάσει σημαντικά τη σταθερότητα του ισοσταθμιστή.

Πίνακας 4.6 : Σύγκριση αριθμού πολλαπλασιασμών και BER για τα διάφορα συγκρινόμενα σχήματα εξισορρόπησης

ΣΥΓΚΡΙΣΗ ΑΡΙΘΜΟΥ ΠΟΛΛΑΠΛΑΣΙΑΣΜΩΝ ΚΑΙ BER		
Τύπος Ισοσταθμιστή	Πολλαπλασιασμοί ανά εξαγόμενο σύμβολο	Βέλτιστη Επίδοση (ελάχιστο BER)
Volterra 1 ^{ης} τάξης	242 (121 taps)	4.2×10^{-3}
Volterra 3 ^{ης} τάξης	1670 (121-21-7 taps)	4×10^{-3}
1 κανάλι bi-RNN	460 (16h.u - 101/121 πολλά σε πολλά)	10^{-3}
3 κανάλια bi-RNN	351 (18h.u - 101/121 πολλά σε πολλά)	10^{-4}
5 κανάλια bi-RNN	288 (20h.u - 101/121 πολλά σε πολλά)	10^{-4}

4.5 Συμπεράσματα

Στο κεφάλαιο αυτό προτάθηκε η λύση των RNN για την εξισορρόπηση των μη γραμμικών φαινομένων διάδοσης στα συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων. Στη βάση αριθμητικών μοντέλων προσομοίωσης ψηφιακών σύμφωνών συστημάτων σε δύο διαφορετικές ζώνες, την C και την O αναλύθηκε η συμπεριφορά του μοντέλου LSTM, πραγματοποιήθηκε σύγκριση με τον ισοσταθμιστή DBP και συζητήθηκε εκτενώς η σταθερότητά του στις αλλαγές των συνθηκών διάδοσης.

Στη συνέχεια πραγματοποιήθηκε σύγκριση των επιδόσεων και της υπολογιστικής πολυπλοκότητα τριών διαφορετικών τύπων bi-RNN, των LSTM, GRU και Vanilla-RNN. Ο τελευταίος αποδείχθηκε εξίσου αποδοτικός με τους προηγούμενους και ως ο λιγότερο σύνθετος από άποψη υπολογιστικής πολυπλοκότητας επιλέχθηκε για τη συνέχεια της ανάλυσης, η οποία περιλαμβάνει σύγκριση επιδόσεων με τον ισοσταθμιστή Volterra και τεχνική σημαντικής μείωσης της υπολογιστικής πολυπλοκότητας με βάση την τεχνική εκπαίδευσης «πολλά σε πολλά».

Αναλύθηκε η ιδέα της πολυκαναλικής ισοστάθμισης με τη βοήθεια επεξεργαστών bi-VRNN στη βάση αριθμητικών μοντέλων προσομοίωσης και πώς αυτοί υποβοηθούν τη βελτίωση των επιδόσεων της εξισορρόπησης των μη γραμμικών φαινομένων της οπτικής διάδοσης και ταυτόχρονα την περεταίρω μείωση της υπολογιστικής πολυπλοκότητας.

Τέλος επιβεβαιώθηκε το σύνολο της προαναφερθείσας μελέτης με τη βοήθεια πειραματικής οπτικής διάδοσης 180km, 3 WDM καναλιών, QPSK διαμόρφωσης και ρυθμού 22,5 Gb/s. Με την εφαρμογή πολυκαναλικού bi-VRNN ισοσταθμιστή αποδείχθηκε η αναμενόμενη εξισορρόπηση των μη γραμμικών φαινομένων, η επίτευξη βελτιωμένων επιδόσεων BER χάρη στην από κοινού εκπαίδευση, όπως και η μείωση της πολυπλοκότητας.

5° ΚΕΦΑΛΑΙΟ

Ισοστάθμιση Βασισμένη σε Επαναλαμβανόμενα Νευρωνικά Δίκτυα στα Συστήματα Οπτικών Επικοινωνιών Μικρών Αποστάσεων

5.1 Εισαγωγή

Τα επαναλαμβανόμενα νευρωνικά δίκτυα (RNN) που αναλύθηκαν και εφαρμόστηκαν για την ισοστάθμιση του σήματος καναλιών οπτικών επικοινωνιών μεγάλων αποστάσεων (long haul), μπορούν να εφαρμοστούν, με την ανάλογη προσαρμογή και στα συστήματα επικοινωνιών μικρών αποστάσεων που λειτουργούν σε πολύ υψηλούς ρυθμούς.

Στο κεφάλαιο αυτό θα παρουσιαστούν οι πρόσφατες και πλέον σημαντικές τεχνικές εξισορρόπησης που βασίζονται σε αλγόριθμους νευρωνικών δικτύων και χρησιμοποιούνται στα συστήματα οπτικών επικοινωνιών μικρών αποστάσεων (short reach).

Θα παρουσιαστούν αναλυτικές συγκρίσεις μεταξύ της απλοποιημένης και βελτιστοποιημένης εκδοχής RNN, του αμφίδρομου VRNN και διαφορετικών ισοσταθμιστών ως προς τις επιδόσεις και την πολυπλοκότητα πάνω σε πειραματικό IM/DD CWDM σύστημα διάδοσης 100km στη ζώνη O [3].

5.2 Ισοστάθμιση υποβοηθούμενη από bi-RNN για CWDM 100km διάδοση 50-Gb/s/λ στη ζώνη O

5.2.1 Εισαγωγή

Η οπτική διάδοση στη ζώνη O του οπτικού φάσματος είναι ελκυστική λόγω της διαθεσιμότητας πομποδεκτών χαμηλού κόστους και της χαμηλής χρωματικής διασποράς (CD) της μονότροπης οπτικής ίνας σε αυτά τα μήκη κύματος [151].

Εξαιτίας των μεγάλων απωλειών που χαρακτηρίζουν τα συστήματα της ζώνης O και της έλλειψης αποτελεσματικών ενισχυτών, η εφαρμογή τους περιορίζεται σε συστήματα οπτικών επικοινωνιών μικρών αποστάσεων [152]. Αν και οι οπτικοί ενισχυτές ημιαγωγών (semiconductor optical amplifier - SOA) έχουν υιοθετηθεί για την ενίσχυση της ζώνης O, οι υψηλές τιμές εικόνας θορύβου και η μη γραμμική παραμόρφωσή τους περιορίζουν την απόδοση της διάδοσης. Σε αντίθεση με τους SOA, οι ενισχυτές ινών με πρόσμειξη βισμούθιου (BDFFA) που εμφανίστηκαν πρόσφατα, παρουσιάζουν σαφώς βελτιωμένες επιδόσεις και διευκολύνουν την επέκταση των οπτικών δικτύων διευρύνοντας το εύρος ζώνης λειτουργίας [153]. Ωστόσο, καθώς η εμβέλεια στα συστήματα της ζώνης O επεκτείνεται, η μη γραμμικότητα της οπτικής ίνας γίνεται πιο σημαντική από ό,τι στη ζώνη C. Επιπλέον, η συσσωρευμένη διασπορά μπορεί ακόμα να επιφέρει σημαντικές διαλείψεις ισχύος σε συστήματα διαμόρφωσης έντασης και άμεσης ανίχνευσης (IM/DD), παρά τις χαμηλές καθαρές τιμές CD [152].

Για τον σκοπό αυτό, απαιτούνται ιδιαίτερα αποτελεσματικά συστήματα εξισορρόπησης για την καταπολέμηση αυτών των επιβαρύνσεων [154]. Μεταξύ των διαφόρων προτάσεων, ενδιαφέρον παρουσιάζει η τεχνική εξισορρόπησης στη μετάδοση NRZ-OOK σημάτων ρυθμού 4x25-Gb/s σε αποστάσεις έως 20 km στη ζώνη O [155]. Ωστόσο, λαμβάνοντας υπόψη τη συνεισφορά των φαινομένων της ανίχνευσης στη φωτοδίοδο (square-law detection), τις αστάθειες του laser και του διαμορφωτή, τη διασπορά των οπτικών ινών και τον πρόσθετο θόρυβο στα συστήματα IM/DD, η ακριβής εκτίμηση της

μη γραμμικής απόκρισης του καναλιού αποτελούν για σχήματα εξισορρόπησης όπως FFE ή DFE, σημαντική πρόκληση [35].

Η υιοθέτηση τεχνικών που βασίζονται στη μηχανική μάθηση συμπεριλαμβανομένων των ισοσταθμιστών νευρωνικών δικτύων, έδειξε τη δυνατότητα ανάκτησης στοχευμένων σημάτων από μη γραμμικές παραμορφώσεις τόσο σε σύμφωνα ψηφιακά συστήματα (coherent), όπως αναλυτικά παρουσιάστηκαν στο προηγούμενο κεφάλαιο, όσο και στα συστήματα IM/DD [35], [156], [45].

Οι περισσότερες από τις προηγούμενες εργασίες επικεντρώθηκαν στη ζώνη C, ωστόσο η χρήση της εξισορρόπησης που βασίζεται στη μηχανική μάθηση είναι ακόμη πιο επιθυμητή στη ζώνη O, ειδικά σε συστήματα διαίρεσης μήκους κύματος (WDM), όπου οι μη γραμμικότητες που προκύπτουν και από την οπτική ίνα αλλά και τα υπόλοιπα στοιχεία του συστήματος διάδοσης γίνονται πιο έντονες λόγω της χαμηλής διασποράς.

Χρησιμοποιώντας τον αμφίδρομο μοντέλο Vanilla-RNN (bi-VRNN), σε σύστημα CWDM διάδοσης στη ζώνη O, μέσω ενός 4x50-Gb/s IM/DD σχήματος, δείχνουμε πως η RNN εξισορρόπηση μπορεί να επιτύχει σημαντικά καλύτερη επίδοση BER σε σύγκριση με το συμβατικό σχήμα DFE τόσο για τα σήματα Nyquist OOK όσο και για τα σήματα PAM4 των 200 Gb/s οπτικής διάδοσης 100 km και 75 km, αντίστοιχα. Επιπλέον, οι επιδόσεις του bi-VRNN ως συνάρτηση της επίδοσης BER αλλά και η πολυπλοκότητα υλοποίησης αξιολογούνται περαιτέρω σε σύγκριση με άλλα γραμμικά και μη γραμμικά σχήματα, συμπεριλαμβανομένου του γραμμικού ισοσταθμιστή, του ισοσταθμιστή τύπου Volterra 3ης τάξης και του σύνθετους σχήματος Volterra και DFE.

Η ερευνητική δουλειά στο πλαίσιο αυτής της συνιστώσας της διδακτορικής διατριβής αποτελεί αποτέλεσμα συνεργασίας με το Optoelectronics Research Centre (ORC) του Πανεπιστημίου του Southampton και συγκεκριμένα με τους Dr. Yang Hong, Dr. Natsupa Taengnoi, Kyle R. H. Bottrill, Naresh Kumar Thipparapu, Yu Wang, Jayanta Sahu, Prof. David J. Richardson και Prof. P. Petropoulos.

5.2.2 Αρχή λειτουργίας του προτεινόμενου συστήματος εξισορρόπησης

Η δομή της κρυφής μονάδας του Vanilla-RNN αποτυπώνεται στο σχήμα 5.1. Η έξοδος του h_t δίνεται από τη σχέση:

$$h_t = \tanh(W h_{t-1} + U x_t) \quad (5.1)$$

όπου οι πίνακες W και U περιέχουν τα βάρη των επιμέρους συνδέσεων, x_t , h_t και h_{t-1} η είσοδος, κρυφή έξοδος και προηγούμενη κρυφή έξοδος αντίστοιχα.

Η διαφορά της από το σχήμα 3.10 και τη σχέση 3.5 είναι η απαλοιφή του bias. Στη μελέτη που ακολουθεί, τόσο στη σχεδίαση του μοντέλου και την εκπαίδευσή του όσο και στον υπολογισμό των βαρών δεν έχει χρησιμοποιηθεί ο όρος του bias.

Το σχήμα 5.1β απεικονίζει το διάγραμμα ροής του μοντέλου bi-VRNN. Η είσοδος του x είναι ένα διάνυσμα $N \times L$ όπου N ο συνολικός αριθμός των βαρών της εισόδου και $L=2k+1$ το μήκος της λέξης των συμβόλων εκπαίδευσης. Τη χρονική στιγμή t η είσοδος εκφράζεται από τη σχέση:

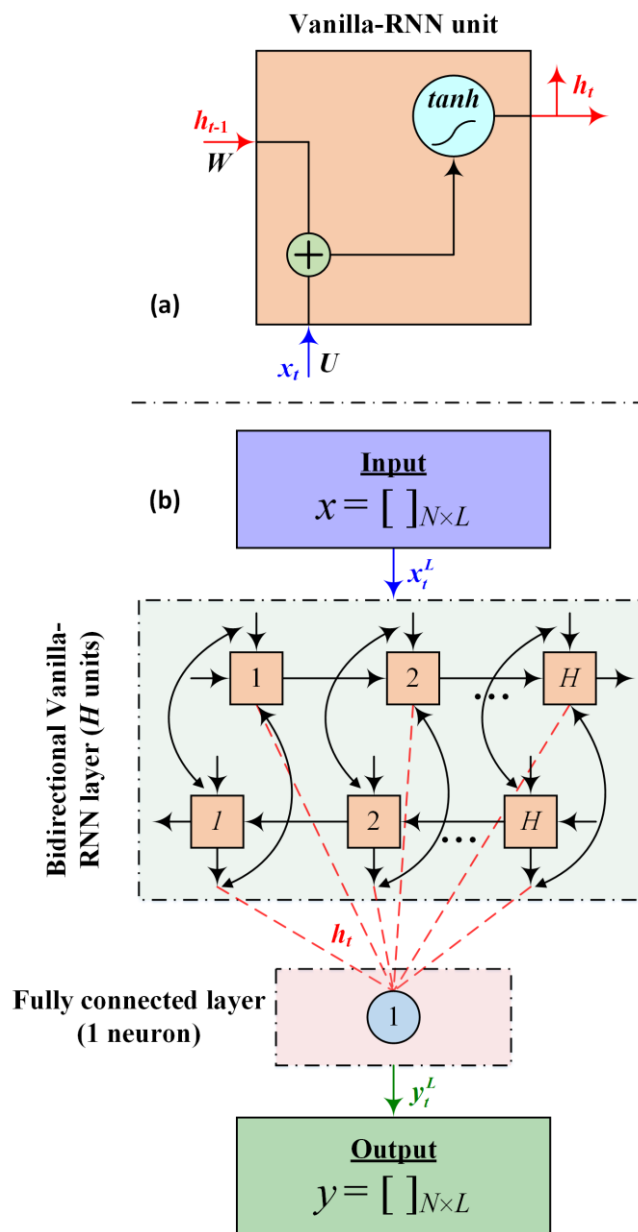
$$x_{t,m} = [x_{t-k}, \dots, x_{t-1}, x_t, x_{t+1}, \dots, x_{t+k}] \quad (5.2)$$

όπου $m=2k+1$ και k δηλώνει τον αριθμό των προηγούμενων και των επόμενων συμβόλων που λαμβάνονται υπόψη για την επεξεργασία του συμβόλου x_t .

Η επιλογή του μήκους L εξαρτάται από την προβλεπόμενη μνήμη του καναλιού, η οποία σχετίζεται με τη συσσωρευμένη χρωματική διασπορά στην ίνα και το περιορισμένο εύρος ζώνης του πομποδέκτη. Η είσοδος x έχει F χαρακτηριστικά στη γενική της περίπτωση (όπως στις περιπτώσεις του κεφαλαίου 4), όμως στη μελέτη αυτή η ίδια η ακολουθία συμβόλων εισόδου ορίζεται ως το μοναδικό χαρακτηριστικό, επομένως $F=1$ τόσο για τη διαμόρφωση OOK όσο και για τη PAM4.

Η είσοδος x οδηγείται στο bi-VRNN το οποίο διαθέτει H κρυφές μονάδες. Η επεξεργασία των δεδομένων της εισόδου πραγματοποιείται αμφίδρομα, όπως υποδηλώνεται από τα καμπυλοειδή τόξα του σχήματος 5.1β. Αντιθέτως τα μικρά τόξα που δείχνουν τις κρυφές μονάδες υποδηλώνουν την παράλληλη επεξεργασία του ίδιου συμβόλου τη χρονική στιγμή t , δηλαδή του x_t . Οι επιμέρους έξοδοι h_t των κρυφών μονάδων συνενώνονται,

(κόκκινα διακεκομμένα βέλη στο σχήμα 5.1β) και οδηγούνται στο FC επίπεδο του ενός και μοναδικού νευρώνα ($n=1$) γεγονός που ισχύει στην περίπτωση της διαμόρφωσης OOK και της PAM4.



Σχήμα 5.1: (α) Εννοιολογική απεικόνιση της μονάδας Vanilla-RNN και (β) διάγραμμα του αμφίδρομου ισοσταθμιστή Vanilla-RNN.

Το πλήρως συνδεδεμένο επίπεδο υιοθετεί την προσέγγιση της παλινδρόμησης για την αναγνώριση του συμβόλου, η οποία μειώνει σημαντικά την πολυπλοκότητα υλοποίησης, ιδιαίτερα στις διαμορφώσεις ανώτερης τάξης όπως πχ την PAM4, χωρίς να θυσιάζει την επίδοση BER, σε σύγκριση με την προσέγγιση της ταξινόμησης, όπως δείξαμε στην ενότητα 4.3.

Τέλος η έξοδος y_i ακολουθεί την προσέγγιση πολλά σε πολλά και συγκεκριμένα εξάγει τον ίδιο αριθμό συμβόλων με την είσοδο x_i ενώ περιγράφεται από τον τύπο:

$$y_i^L = [y_{t-k}, \dots, y_{t-1}, y_t, y_{t+1}, \dots, y_{t+k}] \quad (5.3)$$

Στη συγκεκριμένη περίπτωση που μελετάμε, χρησιμοποιούμε $L=21$ σύμβολα για εκπαίδευση και εξαγωγή. Λαμβάνοντας υπόψη την ανάλυση που προηγήθηκε και συγκεκριμένα τις επιδόσεις BER των γειτονικών συμβόλων του σχήματος 4.14, διαπιστώνουμε πως τα 10 κεντρικά σύμβολα διαθέτουν εφάμιλλές επιδόσεις BER οπότε και τα επιλέγουμε για να την προσμέτρηση των επιδόσεων και τον υπολογισμό της πολυπλοκότητας.

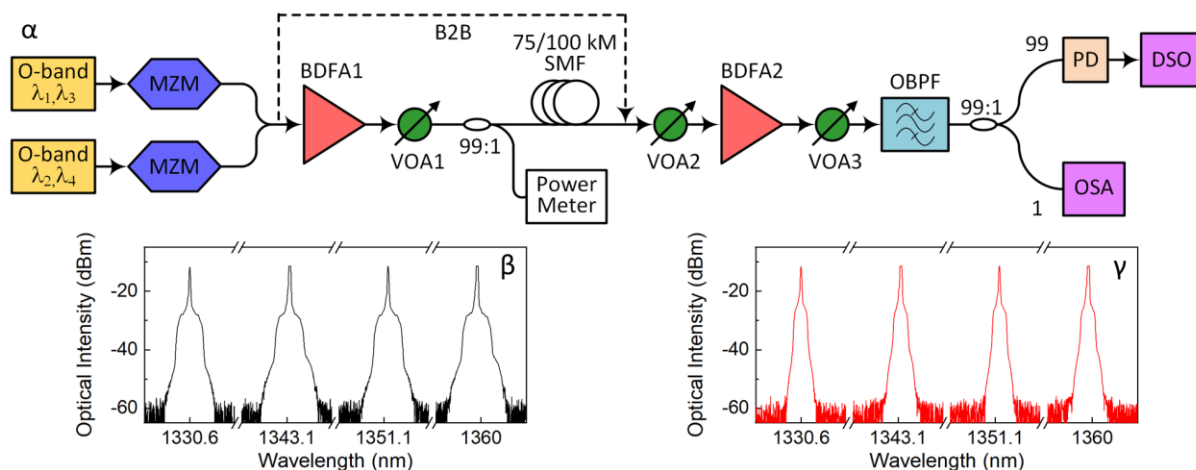
Ο αριθμός των κρυφών μονάδων H που απαιτούνται για την επαρκή εκπαίδευση του bi-VRNN είναι 8 για την OOK και 10 για την PAM4 διαμόρφωση αντίστοιχα. Η μικρή αυτή επαύξηση δικαιολογείται από την μεγαλύτερη ευαισθησία του PAM4 στις μη γραμμικότητες, ιδιαίτερα στις στάθμες υψηλότερης ισχύος. Το μοντέλο του σχήματος 5.1β σχεδιάστηκε και εκπαιδεύτηκε στο Keras με τη βοήθεια του *Tensorflow* 2.3 GPU. Το μέσο τετραγωνικό σφάλμα (MSE) και ο Adam επιλέχθηκαν ως συνάρτηση απώλειας και αλγόριθμος βελτιστοποίησης της εκπαίδευσης, αντίστοιχα. Για τη λεπτομερή αξιολόγηση το διαθέσιμα πειραματικά δεδομένα χωρίστηκαν σε σύνολα, το καθένα από τα οποία αποτελείται από 40.000 σύμβολα για εκπαίδευση, 20.000 σύμβολα για επικύρωση και 60.000 σύμβολα για δοκιμή με άγνωστα δεδομένα, ενώ συνολικά πέντε σύνολα δεδομένων χρησιμοποιήθηκαν για τον υπολογισμό του BER.

Το στάδιο εκπαίδευσης εκτελέστηκε με πακέτα (batches) των 1000 λέξεων συμβόλων με σκοπό βέλτιστη ισορροπία μεταξύ του μεγέθους εκχώρησης μνήμης και του χρόνου εκτέλεσης. Όλες οι λέξεις των συμβόλων σε κάθε πακέτο επεξεργάζονται παράλληλα χρησιμοποιώντας τη διαθέσιμη GPU στο στάδιο της εκπαίδευσης. Ενώ η ποσότητα της

μνήμης κλιμακώνεται με το μέγεθος του πακέτου, το πλεονέκτημα των πολυπύρηνων αρχιτεκτονικών των GPU συνήθως δεν χρησιμοποιείται αν το μέγεθος των πακέτων είναι σχετικά μικρό [102] (βλ κεφ. 7). Επομένως, βελτιστοποιήσαμε την ισορροπία μεταξύ της εκχώρησης μνήμης και του χρόνου εκτέλεσης, η οποία είχε ως αποτέλεσμα το μέγεθος των πακέτων στις 1000 λέξεις συμβόλων. Σημειώνουμε πως το μέγεθος των πακέτων δεν επηρεάζει την απόδοση εξισορρόπησης, καθώς προσαρμόζει μόνο τον ρυθμό με τον οποίο αξιολογούνται οι μερικές παράγωγοι του SGD αλγορίθμου. Ο μέγιστος αριθμός των εποχών εκπαίδευσης είναι 500 ωστόσο χρησιμοποιείται ο μηχανισμός πρόωρης διακοπής της εκπαίδευσης όταν η συνάρτηση σφάλματος παραμένει σταθερή για 20 εποχές, με σκοπό την αποφυγή του υπερπροσαρμογής.

5.2.3 Πειραματική διάταξη

Για την πειραματική επιβεβαίωση των επιδόσεων του προτεινόμενου bi-VRNN χρησιμοποιείται διάταξη οπτικής διάδοσης στη ζώνη O, τεσσάρων καναλιών σε CWDM πολυπλεξία και ρυθμό 4x50-Gb/s, όπως φαίνεται στο σχήμα 5.2α



Σχήμα 5.2: (α) Πειραματική διάταξη της CWDM διάδοσης στην O-band και οπτικά φάσματα των διαμορφωμένων σημάτων μετά τους MZM στην περίπτωση (β) της διαμόρφωσης OOK και (γ) της PAM4.

Ο πομπός αποτελείται από 4 laser συνεχούς κύματος και 2 διαμορφωτές τύπου Mach-Zender (MZM). Τα μονά κανάλια ($\lambda_1=1330.6$ nm και $\lambda_3=1351.1$ nm) και τα ζυγά ($\lambda_2=1343.1$ nm και $\lambda_4=1360.0$ nm) τροφοδοτούνται σε διαφορετικούς MZM. Με αυτόν τον τρόπο, τα μονά και ζυγά WDM κανάλια μπορούν να αποσυσχετιστούν πλήρως. Αναμένουμε αμελητέα επίδραση από τη συσχέτιση καναλιών στα πειράματά μας. Μετά τους διαμορφωτές MZM, τα οπτικά σήματα συνδυάζονται μέσω ενός οπτικού συζεύκτη και στη συνέχεια ενισχύονται από έναν οπτικό ενισχυτή τύπου BDFA. Η εισερχόμενη ισχύς στην οπτική ίνα ελέγχεται από ένα μεταβλητό οπτικό εξασθενητή (Variable Optical Attenuator, VOA), τον VOA₁, το 1% της ισχύς του οποίου χρησιμοποιείται για έλεγχο της ισχύος. Μετά τη διάδοση στη μονότροπη οπτική ίνα χρησιμοποιείται ο VOA₂ για την ρύθμιση της εισερχόμενης ισχύος στον προενισχυτή BDFA₂ με σκοπό τον έλεγχο του OSNR στον δέκτη. Ακολούθως χρησιμοποιείται ένας ακόμα οπτικός εξασθενητής, ο VOA₃. Για τις διάφορες τιμές OSNR και όλα τα WDM κανάλια καταλήγουν στη φωτοδίοδο (PD) η οποία είναι τύπου Finisar XPRV2022A. Σημειώνουμε ότι η φωτοδίοδος προσφέρει τυπικό κέρδος μετατροπής 500 V/W χάρη στην ενσωμάτωση ενός ενισχυτή σύνθετης αντίστασης, ο οποίος εξασφαλίζει ότι η λαμβανόμενη οπτική ισχύς που χρησιμοποιείται στα πειράματα είναι πολύ πάνω από το όριο ευαισθησίας της PD.

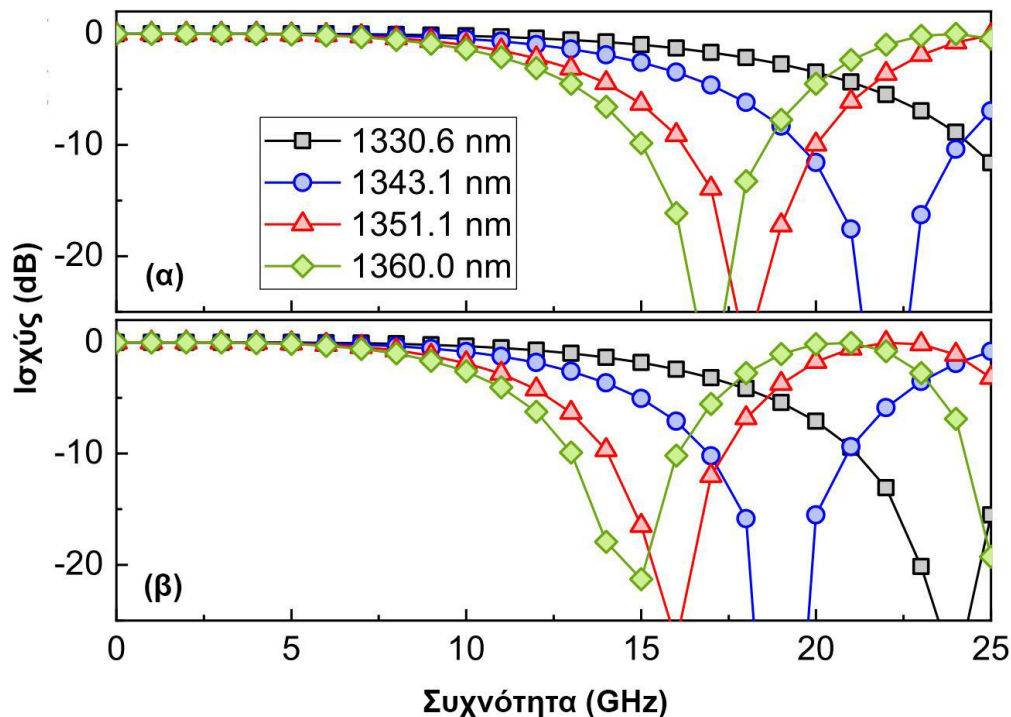
Χρησιμοποιείται ένα οπτικό φίλτρο διέλευσης ζώνης (optical bandpass filter, OBPF) με εύρος ζώνης 1.2 nm για την επιλογή του καναλιού WDM του οποίου θα αξιολογηθεί η απόδοση. Το 1% του οπτικού σήματος μετά το OBPF τροφοδοτείται σε έναν αναλυτή οπτικού φάσματος (optical spectrum analyzer, OSA) για την παρακολούθηση του λόγου σήματος προς θόρυβο σε εύρος ζώνης ανάλυσης 0.1 nm. Αξίζει επίσης να αναφερθεί πως η επιλογή του μήκους κύματος των CWDM καναλιών (~1330nm έως 1360 nm) και η μεταξύ τους απόσταση (10 nm) έγινε ούτως ώστε να ταιριάζουν στο προφίλ απολαβής των διαθέσιμων BDFA.

Παρά το γεγονός πως τα επιμέρους χαρακτηριστικά της πειραματικής οπτικής διάταξης δεν ευθυγραμμίζονται απόλυτα με τα σύγχρονα συστήματα αιχμής (πχ το πρότυπο 200 GBASE-FR4, το οποίο υποστηρίζει 4 οπτικές γραμμές των 50 Gbps), εντούτοις το πείραμά μας προσφέρει μια χρήσιμη διερεύνηση της απόδοσης των ισοσταθμιστών μηχανικής μάθησης σε πραγματικά WDM συστήματα στη ζώνη O.

Χάρη στη χρήση οπτικών ενισχυτών επιτυγχάνεται διάδοση 75 km και 100 km στη μονότροπη ίνα του πειράματος για τη διαμόρφωση PAM4 και OOK αντίστοιχα. Στα μήκη κύματος ενδιαφέροντος οι απώλειες εισαγωγής ήταν περίπου 23.5 dB και 31.3 dB για τα 75 km και τα 100 km αντίστοιχα.

Οι τιμές της χρωματικής διασποράς της οπτικής ίνας στα 4 μήκη κύματος υπολογίζονται στα 1.5, 2.3, 3.3 και 3.8 ps/nm/km αντίστοιχα. Ο BDFA₁ αποδίδει κέρδος περίπου 20 dB όταν η ισχύς εισόδου είναι -6 dBm ενώ ο BDFA₂ αποδίδει κέρδος μεγαλύτερο των 25 dB όταν η τιμή της ισχύος εισόδου είναι -20 dBm.

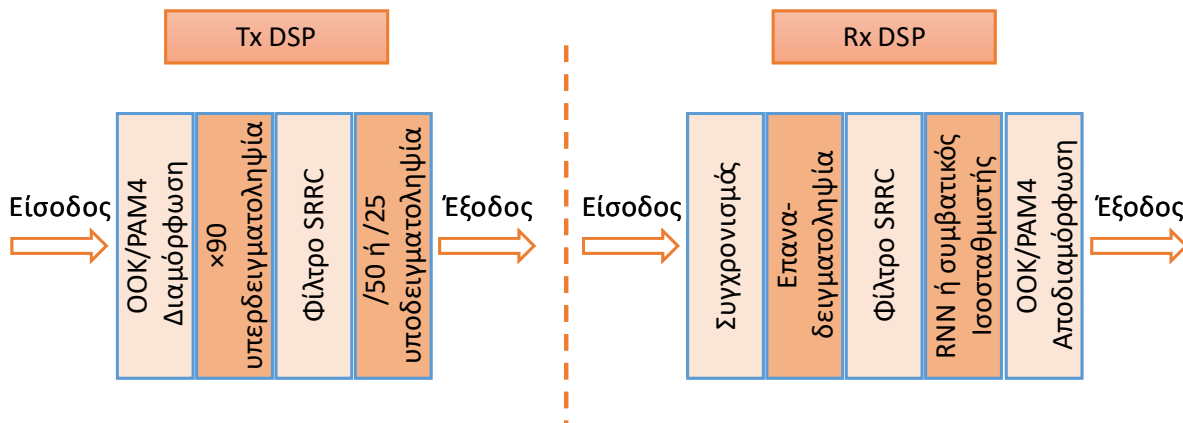
Εξετάζουμε επιπλέον τις επιδόσεις του συστήματος και όταν ο πομπός συνδέεται απευθείας (back-to-back, B2B) με τον δέκτη ως επίδοση αναφοράς. Στην περίπτωση αυτή αφαιρείται από τη διάταξη ο BDFA₁, ο VOA₁ και φυσικά η οπτική ίνα των 75 ή των 100km, όπως φαίνεται με διακεκομμένη γραμμή στο σχήμα 5.2α.



Σχήμα 5.3: (α) Αριθμητικά αποτελέσματα των διαλείψεων ισχύος που προκαλείται από τη χρωματική διασπορά στα τέσσερα CWDM κανάλια της ζώνης O μετά από μετάδοση σε μονότροπη οπτική ίνα (α) 75 km και (β) 100 km.

Τα σχήματα 5.2β και 5.2γ δείχνουν τα οπτικά φάσματα των 4x50 Gb/s/λ WDM σημάτων στην οπτική διάδοση με PAM4 και OOK διαμόρφωση αντίστοιχα.

Παρά το γεγονός πως οι αριθμητικές τιμές της χρωματικής διασποράς των τεσσάρων WDM καναλιών είναι σχετικά χαμηλές, η διασπορά συσσωρεύεται κατά το μήκος της διάδοσης. Αυτό έχει ως αποτέλεσμα, όταν το μήκος της διάδοσης είναι σχετικά μεγάλο, όπως στις περιπτώσεις της διάδοσης στη ζώνη O του πειράματος, τελικά η συσσωρευμένη χρωματική διασπορά να δημιουργεί το φαινόμενο των διαλείψεων ισχύος (power fading) που έχει ως αποτέλεσμα σημαντικές παρεμβολές μεταξύ των συμβόλων. Χρησιμοποιώντας τις παραπάνω αριθμητικές τιμές για τα τέσσερα WDM κανάλια του πειράματος, προσομοιώνουμε το φαινόμενο των διαλείψεων ισχύος για τα 75 km και τα 100 km οπτικής διάδοσης και τα αποτελέσματα απεικονίζονται στα σχήματα 5.3α και 5.3β αντίστοιχα.



Σχήμα 5.4: Τα στάδια της ψηφιακής επεξεργασίας του σήματος στον πομπό (Tx) και το δέκτη (Rx) αντίστοιχα

Είναι φανερή η σημαντική ένταση του φαινομένου των διαλείψεων ισχύος ιδιαίτερα στα μεγαλύτερα μήκη κύματος τα οποία χαρακτηρίζονται από μεγαλύτερες τιμές χρωματικής διασποράς. Διαπιστώνουμε λοιπόν εμφατικά την ανάγκη για την υιοθέτηση αξιόπιστων τεχνικών ισοστάθμισης, όπως αυτές που προτείνουμε σε αυτή την εργασία, που να μπορούν να αντιμετωπίσουν με επάρκεια το φαινόμενο των διαλείψεων της ισχύος και

της συνεπαγόμενης παρεμβολής μεταξύ των συμβόλων. Η αξιοποίηση τους θα ήταν ιδιαίτερα ευεργετική για την επέκταση της WDM διάδοσης στη ζώνη O.

Τα σήματα με διαμόρφωση PAM4 και OOK παράγονται και γίνονται αντικείμενο επεξεργασίας εκτός σύνδεσης από υπολογιστή χρησιμοποιώντας τα βήματα ψηφιακής επεξεργασίας που φαίνονται στο σχήμα 5.4. Συγκεκριμένα, στον πομπό (Tx DSP) η δυαδική είσοδος αρχικά διαμορφώνεται σε PAM4 ή OOK σύμβολα. Η ψευδοτυχαία είσοδος δημιουργείται με τη γεννήτρια Mersenne Twister με περίοδο $2^{19937}-1$ για αποφυγή της επανάληψης των συμβόλων. Στη συνέχεια υπερδειγματοληπτούνται 90 φορές και φιλτράρονται με τη βοήθεια φίλτρου τετραγωνικής ρίζας συνημιτόνου (square-root raised cosine, SRRC). Το φίλτρο SRRC χρησιμοποιεί συντελεστή roll-off 0.1 και έχει εύρος 40 συμβόλων. Η έξοδος του υποδειγματοληπτείται 50 φορές στην περίπτωση της OOK και 25 φορές στην περίπτωση της PAM4 διαμόρφωσης. Τέλος το σήμα οδηγείται σε μία γεννήτρια τυχαίων κυματομορφών (Arbitrary Waveform Generator, AWG) ρυθμού 90 GSa/s για τη δημιουργία των PAM4 ή OOK σημάτων ρυθμού 50 Gb/s. Σημειώνουμε πως για ευκολία, η υπερδειγματοληψία των 90 φορές έγινε για να ευθυγραμμιστεί με τον ρυθμό δειγματοληψίας του διαθέσιμου AWG (ρυθμού 90 GSa/s). Για τη δημιουργία των σημάτων ρυθμού 50 Gb/s οι συντελεστές υπερδειγματοληψίας και υποδειγματοληψίας μπορούν να μειωθούν τόσο ώστε ο λόγος τους να είναι ίσος με 18/5 για την περίπτωση της διαμόρφωσης PAM4 ή 9/5 για την περίπτωση της διαμόρφωσης OOK.

Στο δέκτη τα δεδομένα που λαμβάνονται από τον ψηφιακό παλμογράφο αποθήκευσης (Digital storage oscilloscope, DSO) αρχικά συγχρονίζονται και στη συνέχεια επαναδειγματοληπτούνται πριν την εφαρμογή του SRRC φίλτρου. Ακολούθως, εφαρμόζεται ένας τυπικός ισοσταθμιστής ή ο προτεινόμενος Vanilla-RNN (VRNN). Η έξοδος του ισοσταθμιστή αποδιαμορφώνεται και στην προκύπτουσα δυαδική χρονοσειρά υπολογίζεται η τιμή BER για την αξιολόγηση του σφάλματος.

Η σύγκριση του VRNN γίνεται αρχικά με έναν τυπικό ισοσταθμιστή τύπου DFE ο οποίος μπορεί να ανταπεξέλθει επαρκώς σε συνθήκες έντονων φαινομένων διαλείψεων ισχύος ως αποτέλεσμα της συσσωρευμένης διασποράς [157]. Ο ισοσταθμιστής τύπου DFE επαναδειγματοληπτείται σε 2 δείγματα/ σύμβολο, βασίζεται στον προσαρμοστικό αλγόριθμο ελαχίστων τετραγώνων και ο αριθμός των βημάτων προσωτροφοδότησης και

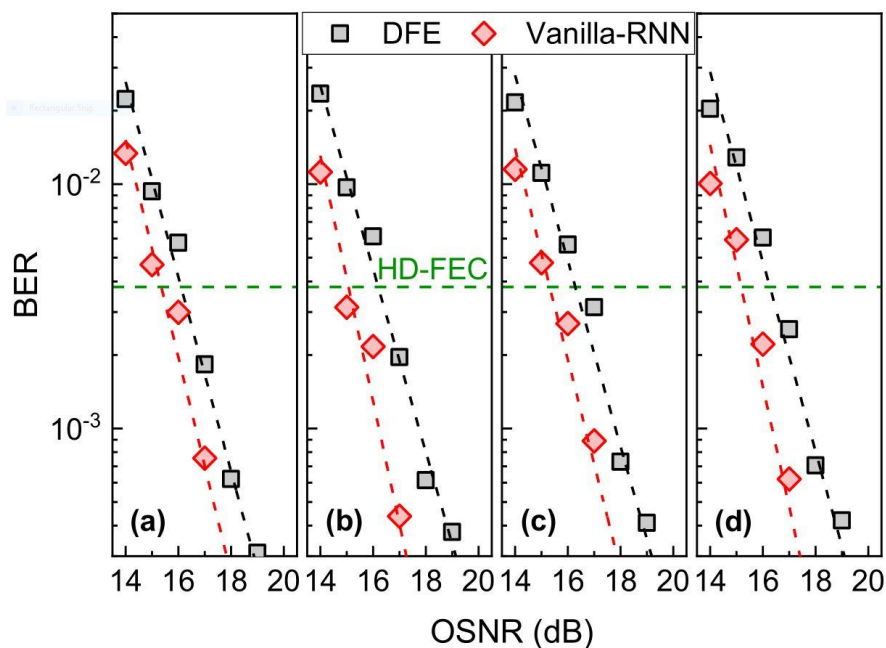
ανάδρασης καθορίζονται στα 17 και 7 αντίστοιχα, έπειτα από αρκετές δοκιμές για τη βελτιστοποίηση της απόδοσής του. Σημειώνουμε πως η λειτουργία του σε 2 δείγματα / σύμβολο οδηγεί σε βελτιωμένη απόδοση και ταυτόχρονα προσφέρει μικρότερη ευαισθησία στην καθυστέρηση δειγματοληψίας φάσης [158]. Αντιθέτως, ο VRNN ισοσταθμιστής λειτουργεί μονάχα με 1 δείγμα / σύμβολο για την ελαχιστοποίηση της υπολογιστικής πολυπλοκότητας, εντούτοις μπορεί να προσφέρει εξαιρετική απόδοση (ελάχιστο BER) όπως θα φανεί στα επόμενα.

Επιπλέον στις επόμενες παραγράφους θα παρουσιαστούν αναλυτικές συγκρίσεις και με άλλου τύπου ισοσταθμιστών όπως τον γραμμικό τύπου FFE και τους μη γραμμικούς τύπου Volterra ή μεικτά σχήματα Volterra και DFE. Τα αποτελέσματα παρουσιάζονται ενδεικτικά για το κανάλι μήκους κύματος 1330.6 nm χωρίς βλάβη της γενικότητας.

5.2.4 Πειραματικά αποτελέσματα για την CWDM διάδοση στη ζώνη O

5.2.4.1 Διάδοση B2B

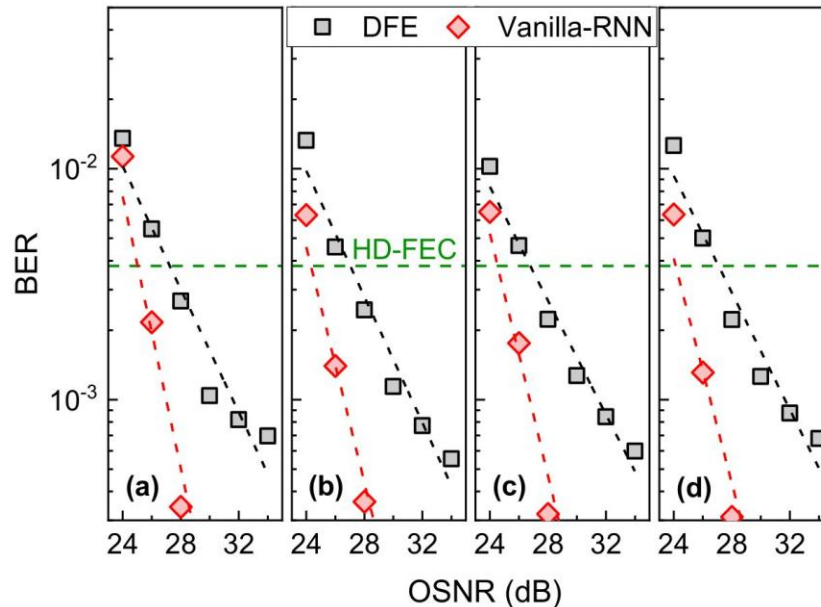
Η απόδοση του VRNN εξετάζεται αρχικά στο σενάριο διάδοσης B2B. Τα αποτελέσματα αποτυπώνονται στο σχήμα 5.5 και 5.6 για την OOK και PAM4 διαμόρφωση αντίστοιχα. Είναι εμφανές πως και στις δύο διαμορφώσεις τα 4 WDM κανάλια παρουσιάζουν παρόμοιες επιδόσεις BER σε σχέση με το OSNR. Στη διαμόρφωση OOK επιτυγχάνεται επίδοση BER κάτω από το όριο της διόρθωσης σφαλμάτων σκληρής απόφασης (HD-FEC, 3.8×10^{-3}). Εντωμεταξύ σε σχέση με τον τυπικό DFE ισοσταθμιστή, ο προτεινόμενος VRNN χρειάζεται περίπου 1 dB χαμηλότερο OSNR για να βρεθεί κάτω από το όριο HD-FEC.



Σχήμα 5.5: BER συναρτήσεϊ του OSNR για τη διάδοσης B2B OOK διαμόρφωσης σε: (α) 1330,6 nm, (β) 1343,1 nm, (γ) 1351,1 nm και (δ) 1360,0 nm.

Αντίστοιχα, όπως φαίνεται στο σχήμα 5.6, στην περίπτωση της διαμόρφωσης PAM4 ο ισοσταθμιστής VRNN προσφέρει περίπου 2 dB χαμηλότερο OSNR για να βρεθεί κάτω από το όριο HD-FEC. Η σημαντικότερη βελτίωση που παρατηρείται στην περίπτωση της

PAM4 διαμόρφωσης αποδίδεται στο γεγονός ότι η PAM4 διαμόρφωση είναι πιο ευαίσθητη στις μη γραμμικότητες του συστήματος που προέρχονται από τους πομποδέκτες, σε σύγκριση με την OOK [159].

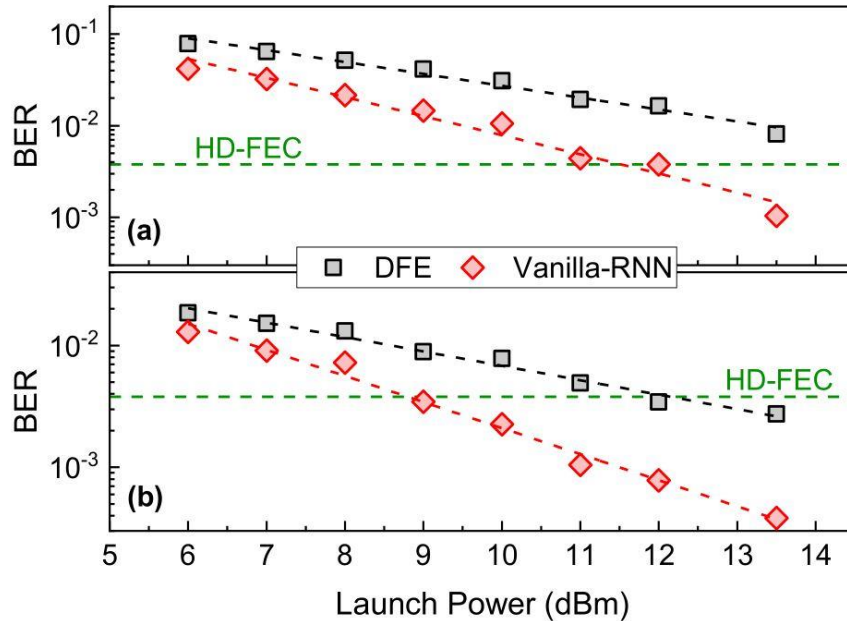


Σχήμα 5.6: BER συναρτήσεϊ του OSNR για τη διάδοσης B2B PAM4 διαμόρφωσης σε: (α) 1330,6 nm, (β) 1343,1 nm, (γ) 1351,1 nm και (δ) 1360,0 nm.

5.2.4.2 Διάδοση 100 km για διαμόρφωση OOK και 75 km για διαμόρφωση PAM4

Έχοντας τα αποτελέσματα της προηγούμενης παραγράφου ως αναφορά πραγματοποιούμε στη συνέχεια πειράματα μετάδοσης τεσσάρων σημάτων ρυθμού 50 Gb/s/λ διαμόρφωσης OOK και PAM4 για διάδοση σε μονότροπη οπτική ίνα μήκους 100 km και 75 km αντίστοιχα. Το κανάλι των 1330.6 nm λήφθηκε ως παράδειγμα για τη διερεύνηση της επίδοσης BER έναντι της συνολικής ισχύος και τα αντίστοιχα αποτελέσματα παρουσιάζονται στο σχήμα 5.7 τόσο για τις περιπτώσεις της OOK διαμόρφωσης όσο και για τις περιπτώσεις της PAM4. Παρατηρούμε ότι και στις δύο περιπτώσεις η τιμή BER μειώνεται σταθερά με την αύξηση της οπτικής ισχύος. Ακόμη και στην μέγιστη τιμή των 13.5 dBm δε διακρίνεται υποβάθμιση της απόδοσης των ισοσταθμιστών DFE και VRNN. Αυτό οφείλεται στη σχετικά αραιή διάταξη των καναλιών

μεταξύ τους (απόσταση περίπου ίση με 10 nm) το οποίο επιφέρει αυξημένη ανθεκτικότητα στις μη γραμμικότητες.

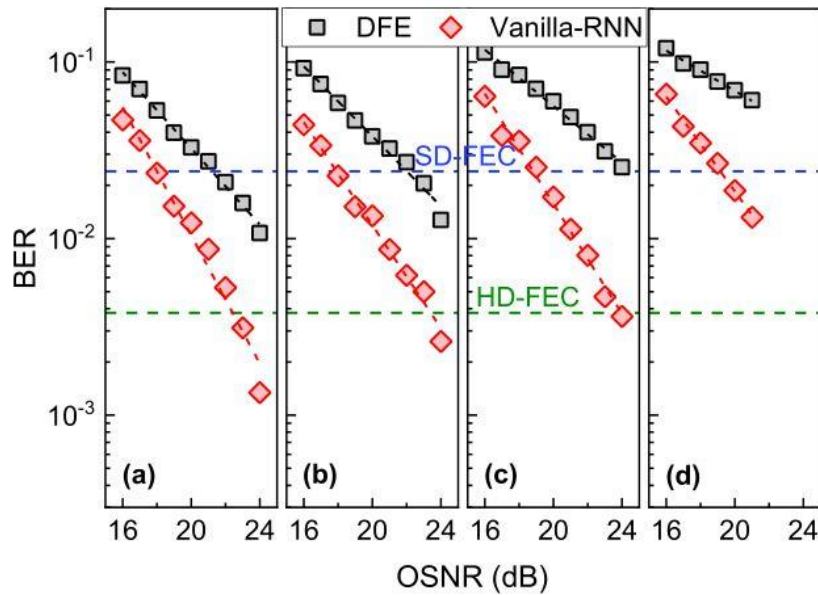


Σχήμα 5.7: BER έναντι συνολικής ισχύος του καναλιού 1330,6 nm: (α) Διάδοση OOK μετά από 100 km μήκους οπτική ίνα και (β) Διάδοση PAM4 μετά από 75 χιλιόμετρα μήκους οπτική ίνα

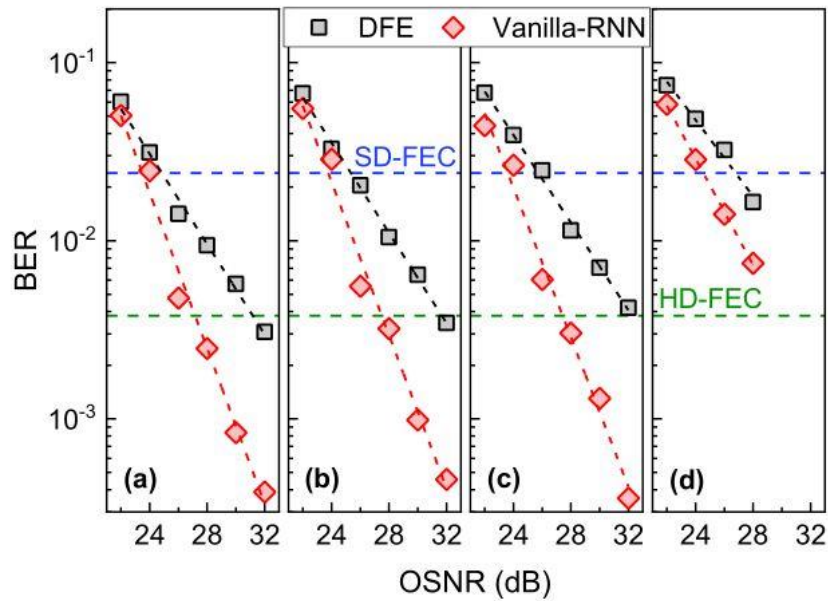
Στο σχήμα 5.7α φαίνεται πως με τη χρήση του DFE ισοσταθμιστή η τιμή BER δεν καταφέρνει να περάσει κάτω από το όριο HD-FEC ακόμη και όταν χρησιμοποιήθηκε η μέγιστη διαθέσιμη συνολική ισχύς των 13.5 dBm. Αντιθέτως το VRNN αγγίζει το όριο HD-FEC για τιμή ισχύος 12 dBm και κατορθώνει να το περάσει για μεγαλύτερες τιμές ισχύος. Για επίτευξη ίδιων επιδόσεων με το DFE το VRNN απαιτεί περίπου 2.5 dB λιγότερη ισχύ.

Στην περίπτωση της διαμόρφωσης PAM4 και διάδοσης 75 km του σχήματος 5.7β φαίνεται πως ο ισοσταθμιστής DFE κατορθώνει να ξεπεράσει το όριο HD-FEC εντούτοις το VRNN σαφώς υπερτερεί καθώς το καταφέρνει άνετα απαιτώντας περίπου 3 dB χαμηλότερη συνολική ισχύ σε σχέση με το DFE. Στη συνέχεια, ρυθμίζοντας την τιμή ισχύος στα 13.5 dBm διερευνούμε την απόδοση και των τεσσάρων καναλιών συναρτήσει της τιμής OSNR για τα ίδια δεδομένα διάδοσης και (100 km για την OOK διαμόρφωση,

75 km για την PAM4). Τα αποτελέσματα αποτυπώνονται στα σχήματα 5.8 και 5.9 αντίστοιχα.



Σχήμα 5.8: BER έναντι OSNR για τη διάδοση 100 km OOK διαμόρφωσης, για το κανάλι: (α) 1330,6 nm, (β) 1343,1 nm, (γ) 1351,1 nm, και (δ) 1360,0 nm



Σχήμα 5.9: BER έναντι OSNR για τη διάδοση 75 km PAM4 διαμόρφωσης, για το κανάλι: (α) 1330,6 nm, (β) 1343,1 nm, (γ) 1351,1 nm, και (δ) 1360,0 nm

Από τα παραπάνω σχήματα φαίνεται πως και στις δύο περιπτώσεις μετάδοσης / διαμόρφωσης οι επιδόσεις BER υποβαθμίζονται όσο το μήκος κύματος του καναλιού μεγαλώνει. Αυτό οφείλεται στην υψηλότερη συσσωρευμένη χρωματική διασπορά των μεγαλύτερων μηκών κύματος που με τη σειρά της προκαλεί το φαινόμενο των διαλείψεων ισχύος με μεγαλύτερη ένταση [160].

Παρατηρούμε λοιπόν πως στην περίπτωση της διάδοσης των 100 km για OOK διαμόρφωση, σχήμα 5.8, οι επιδόσεις BER του ισοσταθμιστή DFE δεν κατορθώνουν να ξεπεράσει το όριο HD-FEC για κανένα κανάλι και καμία τιμή OSNR. Ακόμη και στην περίπτωση που λαμβάνουμε υπόψη το όριο της διόρθωσης σφαλμάτων Soft Decision Forward Error Correction, SD-FEC, 2.4×10^{-2} , μονάχα τα δύο πρώτα κανάλια των μικρότερων μηκών κύματος, 1330.6 nm και 1343.1 nm κατορθώνουν να το ξεπεράσουν. Τουναντίον, ο προτεινόμενος VRNN ισοσταθμιστής κατορθώνει με μεγάλη ευκολία να ξεπεράσει το όριο SD-FEC για όλα τα κανάλια και να προσεγγίσει ή και ξεπεράσει το όριο HD-FEC για όλα τα κανάλια πλην του τελευταίου, με μήκος κύματος 1360.0 nm. Παρατηρούμε ότι οι επιδόσεις του VRNN έναντι του DFE αντιστοιχούν σε τουλάχιστον 4 dB βελτίωση της ευαισθησίας OSNR. Παρόλο που καταδεικνύεται σημαντική μείωση BER στο κανάλι 1360.0 nm σε σχέση με το σχήμα DFE, η επιτεύξιμη επίδοση BER περιορίζεται από το διαθέσιμο OSNR, το οποίο οφείλεται στο χαμηλότερο κέρδος των BDFA σε αυτό το μήκος κύματος.

Επιπλέον, η επίδοση BER μεταξύ των τεσσάρων καναλιών παρουσίασε μικρές διαφορές σε σύγκριση με εκείνη της περίπτωσης του ισοσταθμιστή DFE. Αυτό υποδεικνύει ότι ο VRNN ισοσταθμιστής είναι σχεδόν διαφανής ως προς το μήκος κύματος, αποδεικνύοντας την ικανότητά του να αντισταθμίζει το φαινόμενο των διαλείψεων ισχύος που προκαλείται από τη χρωματική διασπορά. Θα πρέπει να αναφερθεί πως μπορεί να επιτευχθεί περαιτέρω βελτίωση στην επίδοση του BER, σε κανάλια μεγαλύτερου μήκους κύματος (δηλαδή, με μεγαλύτερη διασπορά) μέσω της χρήσης μεγαλύτερου αριθμού κρυφών μονάδων ή/και μεγαλύτερου μήκους λέξης συμβόλων. Ωστόσο, αυτό αναπόφευκτα θα οδηγήσει σε αύξηση της πολυπλοκότητας υλοποίησης του σχήματος VRNN.

Σε σύγκριση με τη διαμόρφωση OOK, η διαμόρφωση PAM4 καταλαμβάνει το μισό εύρος ζώνης για τον ίδιο ρυθμό δεδομένων, αν και αυτό έρχεται σε βάρος της αυξημένης

ευαισθησίας στις μη γραμμικές παραμορφώσεις. Το μισό εύρος ζώνης, ωστόσο, παρέχει τέσσερις φορές υψηλότερη ανοχή στο φαινόμενο διαλείψεων ισχύος που προκαλείται από τη χρωματική διασπορά. Ως αποτέλεσμα, επιτυγχάνεται παρόμοια επίδοση BER για τα τρία πρώτα κανάλια WDM στην περίπτωση διάδοσης 75 km διαμόρφωσης PAM4 και το κανάλι με το μεγαλύτερο μήκος κύματος παρουσιάζει μια μικρή μόνο υποβάθμιση της επίδοσης BER.

Όπως φαίνεται στο σχήμα 5.9, ο ισοσταθμιστής VRNN προσφέρει βελτιώσεις στην ευαισθησία OSNR περίπου 1 dB και 4 dB στα όρια SD-FEC και HD-FEC αντίστοιχα. Όμοια με την περίπτωση της διαμόρφωσης OOK, η επιτεύξιμη επίδοση BER του καναλιού 1360.0 nm παραμένει ανεπαρκής επίσης από το διαθέσιμο OSNR λόγω της χαμηλότερης απολαβής των BDFAs σε αυτό το κανάλι.

Από τα παραπάνω σχήματα είναι σαφές ότι ο προτεινόμενος ισοσταθμιστής VRNN προσφέρει σημαντικά καλύτερη ανοχή στο φαινόμενο διαλείψεων ισχύος που προκαλείται από τη χρωματική διασπορά, σε σύγκριση με το σχήμα DFE. Επιπλέον, παρουσιάζει ανώτερη ικανότητα ισοστάθμισης σε περιβάλλον οπτικής διάδοσης, όπως χαρακτηριστικά αποτυπώνεται στα σχήματα 5.6 και 5.9 όπου η βελτίωση αυξήθηκε από περίπου 2 dB στο σενάριο διάδοσης B2B σε περίπου 4 dB στην περίπτωση διάδοσης 75 km και διαμόρφωσης PAM4.

5.2.5 Σύγκριση επιδόσεων μεταξύ διαφόρων σχημάτων ισοστάθμισης

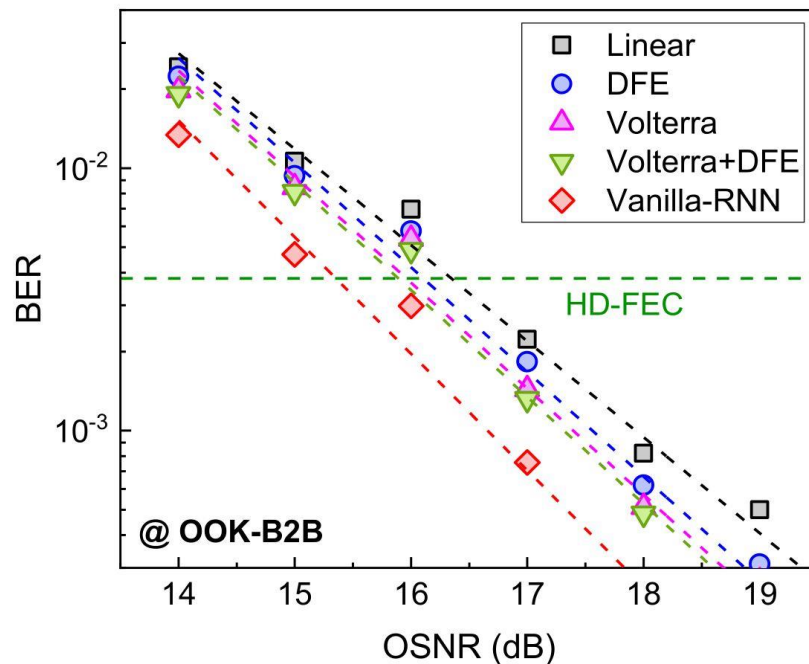
Στις επόμενες παραγράφους ο ισοσταθμιστής Vanilla-RNN συγκρίνεται με διάφορα σχήματα γραμμικών και μη γραμμικών ισοσταθμιστών. Εκτός του ισοσταθμιστή τύπου DFE που αναλύθηκε προηγουμένως, συγκρίσεις πραγματοποιούνται με γραμμικό ισοσταθμιστή, μη γραμμικό Volterra και σχήμα που συνδυάζει Volterra και DFE. Προφανώς ο γραμμικός ισοσταθμιστής δεν είναι ικανός να εξισορροπήσει επαρκώς το φαινόμενο των διαλείψεων ισχύος σε αντίθεση με τους μη γραμμικούς αλγορίθμους, εντούτοις περιλαμβάνεται στη σύγκριση τόσο για λόγους αναφοράς σε πραγματικά συστήματα οπτικών επικοινωνιών όσο και για τη σημαντικά μικρή πολυπλοκότητα

υλοποίησης του [90]. Από την άλλη πλευρά ο μη γραμμικός αλγόριθμος Volterra όπως φάνηκε και στη μελέτη του προηγούμενου κεφαλαίου αποτελεί έναν ισχυρό ισοσταθμιστή που μπορεί να ανταπεξέλθει σε έντονα μη γραμμικά περιβάλλοντα οπτικής διάδοσης. Εδώ χρησιμοποιείται εξαιτίας της αποτελεσματικότητάς του απέναντι στις μη γραμμικότητες τόσο του πομπού όσο και του δέκτη, παρά το γεγονός ότι η πολυπλοκότητά υλοποίησης του είναι σημαντικά μεγαλύτερη από το γραμμικό ισοσταθμιστή [161].

Ο γραμμικός ισοσταθμιστής εφαρμόζεται για 25 βήματα (taps) και χρησιμοποιείται ο ίδιος προσαρμοστικός αλγόριθμος με αυτόν του σχήματος DFE. Ο μη γραμμικός Volterra χρησιμοποιεί τις 3 πρώτες τάξεις και υλοποιείται με 21, 11 και 5 βήματα για την 1^η, 2^η και την 3^η τάξη αντίστοιχα. Για να διασφαλιστεί η ισότιμη σύγκριση των επιμέρους ισοσταθμιστών τα προηγούμενα βήματα επιλέγονται στη βάση των καλύτερων τιμών επίδοσης BER για σχετικά χαμηλή υπολογιστική πολυπλοκότητα. Για τη σύγκριση τόσο των επιδόσεων όσο και της υπολογιστικής πολυπλοκότητας χρησιμοποιείται ως παράδειγμα το κανάλι των 1330.6 nm.

5.2.5.1 Σύγκριση επιδόσεων BER

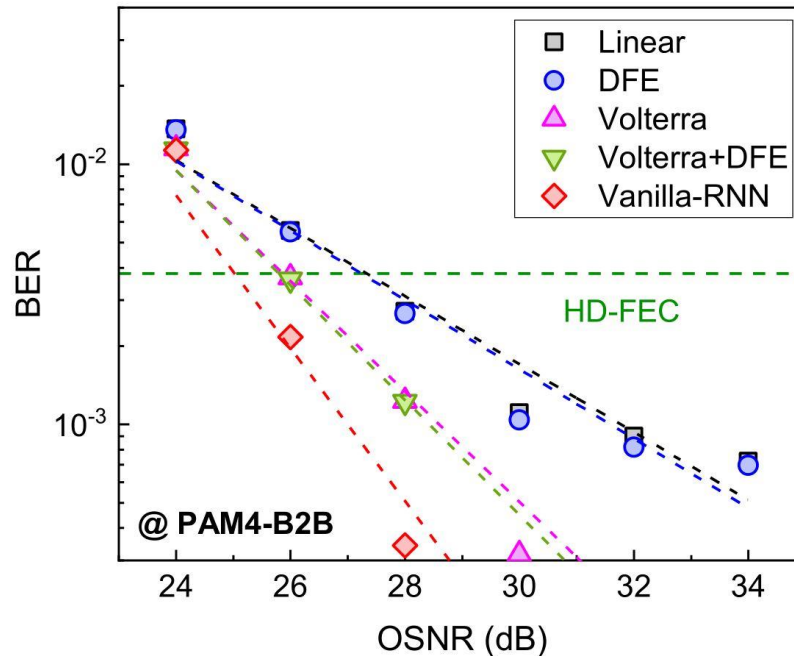
Στο σχήμα 5.10 αποτυπώνεται η επίδοση BER των σχημάτων ισοστάθμισης στη περίπτωση της διαμόρφωσης OOK και διάδοσης B2B. Φαίνεται πως τα σχήματα DFE, Volterra και Volterra+DFE παρουσιάζουν παρόμοιες επιδόσεις και σχετικά καλύτερες από το γραμμικό. Ο προτεινόμενος Vanilla-RNN ισοσταθμιστής είναι ο αποτελεσματικότερος όλων για όλες τις τιμές OSNR, ενώ για να κατορθώσει να περάσει κάτω από το κατώφλι HD-FEC ($BER = 3.4 \times 10^{-3}$) απαιτεί περίπου 0.5 dB λιγότερο OSNR.



Σχήμα 5.10: BER έναντι OSNR διαφόρων σχημάτων ισοστάθμισης. Αφορά διάδοση B2B, OOK διαμόρφωσης, για το κανάλι των 1330,6 nm

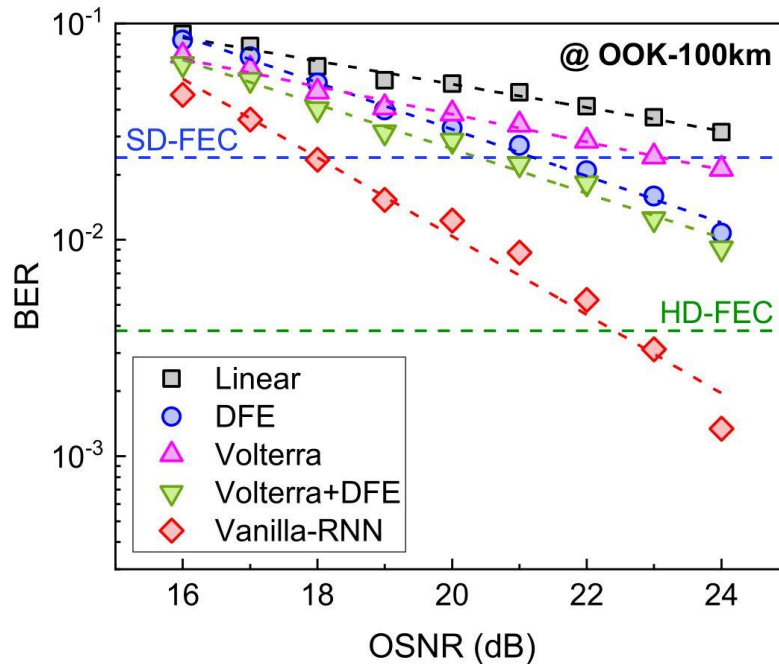
Παρόμοια όπως αποτυπώνεται στο σχήμα 5.11, όταν χρησιμοποιηθεί η διαμόρφωση PAM4 οι επιδόσεις του γραμμικού και DFE ισοσταθμιστή εξισώνονται γεγονός που φανερώνει τις πεπερασμένες δυνατότητες του προσαρμοστικού αλγορίθμου στην αντιμετώπιση των μη γραμμικών φαινομένων του πομποδέκτη. Αντιθέτως ο Volterra παρουσιάζει αξιοσημείωτη βελτίωση των επιδόσεων BER, ενώ η προσθήκη του DFE δείχνει μάλλον ανώφελη αφού το μεικτό σχήμα κατορθώνει να δείξει μόνο μια αμυδρή βελτίωση. Επιπλέον ο VRNN ισοσταθμιστής δείχνει για ακόμη μια φορά την υπεροχή του παρουσιάζοντας τις καλύτερες επιδόσεις και τη βέλτιστη ευαισθησία OSNR μεταξύ των συγκρινόμενων σχημάτων.

Στη συνέχεια διερευνώνται τα σενάρια της διάδοσης σε οπτική ίνα των 100 και 75 km για τις διαμορφώσεις OOK και PAM4 ενώ τα αποτελέσματα φαίνονται στα σχήματα 5.12 και 5.13 αντίστοιχα.



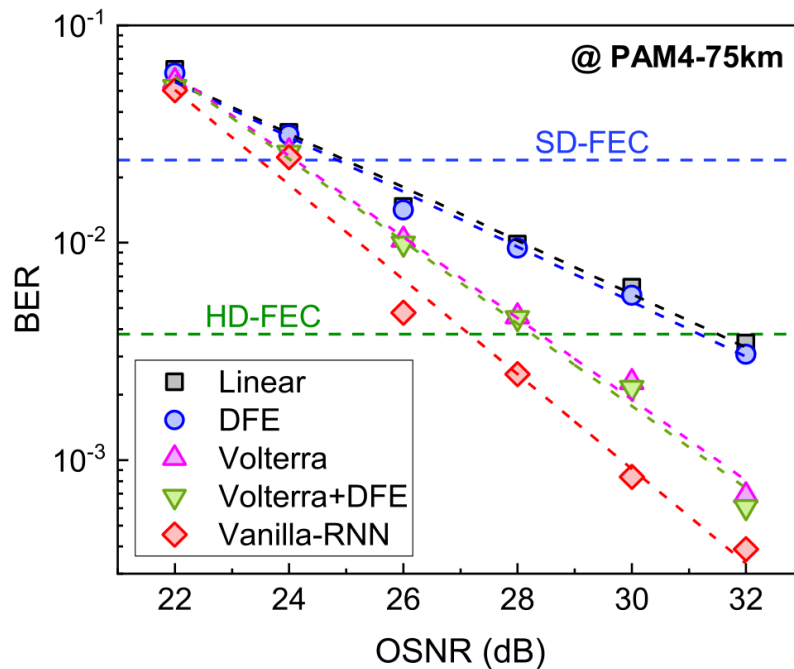
Σχήμα 5.11: BER έναντι OSNR διαφόρων σχημάτων ισοστάθμισης. Αφορά διάδοση B2B, PAM4 διαμόρφωσης, για το κανάλι των 1330,6 nm

Για τη διαμόρφωση OOK (σχήμα 5.12), παρόμοια με τη διάδοση B2B ο γραμμικός αλγόριθμος επιτυγχάνει τη χειρότερη επίδοση BER. Αντιθέτως, τώρα ο ισοσταθμιστής Volterra μη κατορθώνοντας να αντιμετωπίσει με επάρκεια το φαινόμενο των διαλείψεων ισχύος που προκαλείται από τη χρωματική διασπορά, δείχνει χειρότερη επίδοση από τον ισοσταθμιστή DFE. Το πλεονέκτημα του DFE αποτυπώνεται και στο μεικτό σχήμα Volterra+DFE που βελτιώνει ικανοποιητικά τις επιδόσεις BER. Ωστόσο κανένα από τα προηγούμενα σχήματα δεν κατορθώνει να ξεπεράσει το όριο HD-FEC για καμία από τις τιμές OSNR του πειράματος σε αντίθεση με το VRNN που ξεκάθαρα το ξεπερνά με χαρακτηριστική άνεση.



Σχήμα 5.12: BER έναντι OSNR διαφόρων σχημάτων ισοστάθμισης. Αφορά διάδοση 100 km, OOK διαμόρφωσης, για το κανάλι των 1330,6 nm

Τέλος, στην περίπτωση της διαμόρφωσης PAM-4 (σχήμα 5.13) και σε αντίθεση με τη διαμόρφωση OOK όπου υπερίσχυε το φαινόμενο των διαλείψεων ισχύος, κατά τη διάδοση σε μονότροπη οπτική ίνα των 75 km το φαινόμενο είναι σχετικά ανίσχυρο. Αντιθέτως εδώ κυριαρχούν τα μη γραμμικά φαινόμενα που πηγάζουν από τις επιμέρους διατάξεις του συστήματος και την οπτική διάδοση. Ως αποτέλεσμα αυτού, οι επιδόσεις των διαφόρων ισοσταθμιστών προσιδιάζουν με τα αποτελέσματα του σχήματος 5.11, λαμβάνοντας φυσικά υπόψη τις σημαντικές επιβαρύνσεις της διάδοσης. Συγκεκριμένα ο γραμμικός FFE και ο ισοσταθμιστής τύπου DFE δείχνουν εφάμιλλες επιδόσεις και απαιτούν περίπου 3 dB μεγαλύτερο OSNR συγκρινόμενοι με τους Volterra/Volterra+DFE για να επιτύχουν BER στο όριο HD-FEC. Και σε αυτή την περίπτωση οι επιδόσεις BER του προτεινόμενου VRNN είναι οι καλύτερες παρέχοντας περίπου 1 dB υψηλότερη ευαισθησία OSNR σε σχέση με τα σχήματα Volterra.



Σχήμα 5.13: BER έναντι OSNR διαφόρων σχημάτων ισοστάθμισης. Αφορά διάδοση 75 km, PAM4 διαμόρφωσης, για το κανάλι των 1330,6 nm

5.2.6 Σύγκριση της πολυπλοκότητας υλοποίησης

Στη συνέχεια πραγματοποιείται η σύγκριση των προηγούμενων ισοσταθμιστών ως προς την υπολογιστική πολυπλοκότητα. Τα σχήματα του γραμμικού και του DFE ισοσταθμιστή απαιτούν έναν αριθμό πραγματικών πολλαπλασιασμών ανά σύμβολο που δίνεται από τη σχέση :

$$N_{sample} \times N_{tap} \quad (5.4)$$

όπου N_{sample} είναι ο αριθμός των δειγμάτων ανά σύμβολο και N_{tap} είναι ο συνολικός αριθμός των βημάτων εμπροσθοδιάδοσης και οπισθοδιάδοσης (στην περίπτωση του DFE αλγορίθμου).

Αντίστοιχα για τον ισοσταθμιστή Volterra ο αριθμός πραγματικών πολλαπλασιασμών ανά σύμβολο δίνεται από τη σχέση :

$$N_{volterra} = 2 \times \sum_{j=1}^3 \frac{(L_{j-1}+j)!}{(j-1)!(L_{j-1})!} \quad (5.5)$$

Όπου το ! υποδηλώνει το παραγοντικό, $L_1 = 21$, $L_2 = 11$ και $L_3 = 5$ ο αριθμός των βημάτων της 1^{ης}, 2^{ης} και 3^{ης} τάξης Volterra αντίστοιχα [74].

Στην περίπτωση του VRNN ο αριθμός των πραγματικών πολλαπλασιασμών δίνεται από τη σχέση [2]

$$N_{VRNN} = 2 \frac{(FH + H^2)L + 2HnL}{K} \quad (5.6)$$

Όπου $F=1$ το πλήθος των χαρακτηριστικών εισόδου του νευρωνικού δικτύου, $H = 8$ ή 10 ο αριθμός των κρυφών μονάδων, $n=1$ ο αριθμός των νευρώνων της εξόδου, $L=21$ το μήκος της λέξης των συμβόλων στην είσοδο, και $K=10$ τα σύμβολα που εξάγονται ταυτόχρονα χρησιμοποιώντας την προσέγγιση πολλά σε πολλά.

Πρέπει να αναφερθεί πως σε αντίθεση με τους υπόλοιπους ισοσταθμιστές που λαμβάνουν υπόψη δύο δείγματα ανά σύμβολο για να προσφέρουν τις επιδόσεις που αναφέραμε, ο VRNN αρκείται στο ένα δείγμα ανά σύμβολο τόσο στο στάδιο της εκπαίδευσης όσο και της αξιολόγησης. Επίσης παρά το γεγονός πως η πολυπλοκότητα υλοποίησης της συνάρτησης *tanh* που χρησιμοποιείται στο VRNN είναι σχετικά πολύπλοκη, είναι εφικτή η αποδοτική υλοποίηση σε FPGA μη γραμμικών συναρτήσεων ενεργοποίησης με μικρές προσεγγίσεις που δεν επηρεάζουν σημαντικά τις επιδόσεις, ενώ έχουν μικρό αποτύπωμα στο υλικό, ελάχιστες μονάδες *Look Up Tables (LUT)* [162]. Συνεπώς το συντριπτικά μεγαλύτερο μέρος της συνολικής υπολογιστικής πολυπλοκότητας του VRNN οφείλεται στον αριθμό πολλαπλασιασμών της σχέσης 5.6.

Λαμβάνοντας υπόψη τις παραπάνω σχέσεις ο αριθμός πολλαπλασιασμών των 3 συγκρινόμενων ισοσταθμιστών παρουσιάζεται στον πίνακα 5.2. Σε αυτόν φαίνεται πως η πολυπλοκότητα του Volterra είναι σταθερή και ανεξάρτητη του σχήματος διαμόρφωσης (515 πραγματικοί πολλαπλασιασμοί) ενώ του VRNN αυξάνεται όταν η τάξη της διαμόρφωσης μεγαλώνει (336 στη διαμόρφωση OKK και 504 στη διαμόρφωση PAM4).

Το τελευταίο οφείλεται κυρίως στην απαίτηση για ελαφρώς πιο ισχυρό νευρωνικό δίκτυο στην περίπτωση της PAM4 διαμόρφωσης (10 κρυφές μονάδες έναντι 8 στην περίπτωση της OOK). Παρόλα αυτά η υπολογιστική πολυπλοκότητα του VRNN παραμένει σημαντικά χαμηλότερη στη διαμόρφωση OOK και οριακά μικρότερη στην PAM4 από αυτή του ισοσταθμιστή Volterra ενώ, όπως αναλύθηκε στα προηγούμενα, παρουσιάζει σε όλες τις περιπτώσεις σημαντικά καλύτερες επιδόσεις. Τέλος για λόγους αναφοράς παρουσιάζεται και η υπολογιστική πολυπλοκότητα του DFE η οποία είναι ελάχιστη (μόνο 48 πραγματικοί πολλαπλασιασμοί ανά σύμβολο) και ανεξάρτητη της τάξης της διαμόρφωσης. Παρ' όλα αυτά η επίδοση BER του DFE ισοσταθμιστή είναι κατώτερη και δεν μπορεί να ικανοποιήσει τις απαιτήσεις των σύγχρονων συστημάτων CWDM διάδοσης στη ζώνη O.

Πίνακας 5.1

Σύγκριση πολυπλοκότητας εκφρασμένης με τον αριθμό των απαραίτητων πραγματικών πολλαπλασιασμών

	DFE	Volterra	VRNN
OOK	48	516	336
PAM4	48	516	504

Τέλος πρέπει να σημειωθεί πως σε σύγκριση με τους ισοσταθμιστές RNN που αναλύθηκαν στο προηγούμενο κεφάλαιο και αναφέρονταν κυρίως σε DWDM σύμφωνα συστήματα μεγάλων αποστάσεων διάδοσης στη ζώνη C, τα προτεινόμενα σε αυτή τη περίπτωση (IM/DD, CWDM διάδοσης στη ζώνη O) απαιτούν αρκετά πιο ελαφριές δομές δηλαδή RNN με πολύ μικρό αριθμό κρυφών μονάδων που επεξεργάζονται λέξεις συμβόλων μικρού μήκους. Στην πιθανή μετάβαση σε συνθήκες διάδοσης στη ζώνη C όπου αυξάνεται σημαντικά η χρωματική διασπορά, προκειμένου να επιτευχθούν αποδεκτές επιδόσεις θα πρέπει να χρησιμοποιηθούν ισχυρότεροι RNN ισοσταθμιστές με αυξημένη πολυπλοκότητα υλοποίησης.

5.3 Συμπεράσματα

Σε αυτό το κεφάλαιο προτάθηκε η χρήση ενός αμφίδρομου Vanilla-RNN για την ισοστάθμιση ενός πειραματικού συστήματος CWDM διάδοσης στη ζώνη O τεσσάρων καναλιών με ρυθμό 50 Gb/s/λ. Αποδείχθηκε ότι ο ισοσταθμιστής Vanilla-RNN είναι ικανός να αντισταθμίσει τόσο τη μη γραμμικότητα του πομποδέκτη όσο και τη μη γραμμικότητα λόγω της χρωματικής διασποράς και άμεσης φώρασης καθώς και τη μη γραμμικότητα της μονότροπης ίνας λόγω φαινομένου Kerr, στα σενάρια διάδοσης B2B, 100 km και 75 km χρησιμοποιώντας διαμόρφωση OOK και PAM4 αντίστοιχα. Δείξαμε ότι ο προτεινόμενος Vanilla-RNN ισοσταθμιστής παρουσιάζει σημαντικά καλύτερη επίδοση BER από το συμβατικό σχήμα DFE, από το γραμμικό και από το μη γραμμικό Volterra 3^{ης} τάξης ισοσταθμιστή, σε όλα τα προηγούμενα σενάρια. Τέλος συγκρίνοντας τον επιμέρους αριθμό των πραγματικών πολλαπλασιασμών αποδείχθηκε πως ο Vanilla-RNN είναι λιγότερο πολύπλοκος από το Volterra ισοσταθμιστή δίνοντας τη δυνατότητα χρήσης του σε πραγματικά υψήρρυθμα συστήματα CWDM οπτικής διάδοσης στη ζώνη O.

6° ΚΕΦΑΛΑΙΟ

Προσομοίωση Οπτικών Συστημάτων με τη χρήση Επαναλαμβανόμενων Νευρωνικών Δικτύων

6.1 Εισαγωγή

Ο σχεδιασμός των φωτονικών και οπτικοηλεκτρονικών δομικών στοιχείων είναι πεδίο έντονης έρευνας εξαιτίας του μεγάλου εύρους των εφαρμογών που βασίζονται σε αυτές τις συσκευές, όπως οι οπτικές επικοινωνίες, εφαρμογές αισθητήρων, απεικόνιση κ.ά. Οι παραδοσιακές μέθοδοι σχεδιασμού που περιλαμβάνουν αναλυτικές αριθμητικές προσομοιώσεις και επαναλαμβανόμενες προσαρμογές – βελτιστοποιήσεις των μοντέλων είναι αρκετά χρονοβόρες ενώ απαιτούν ιδιαίτερα αυξημένους υπολογιστικούς πόρους. Η δυνατότητα σχεδιασμού και προσομοίωσης φωτονικών δομών γρήγορα και αποτελεσματικά θα μπορούσε να φέρει νέες τεχνολογίες στην αγορά με ταχύτερο ρυθμό. Θα μπορούσε επίσης να διευκολύνει τη διαδικασία εξαγωγής κρίσιμων μετρήσεων απόδοσης των σύνθετων φωτονικών διατάξεων και συσκευών.

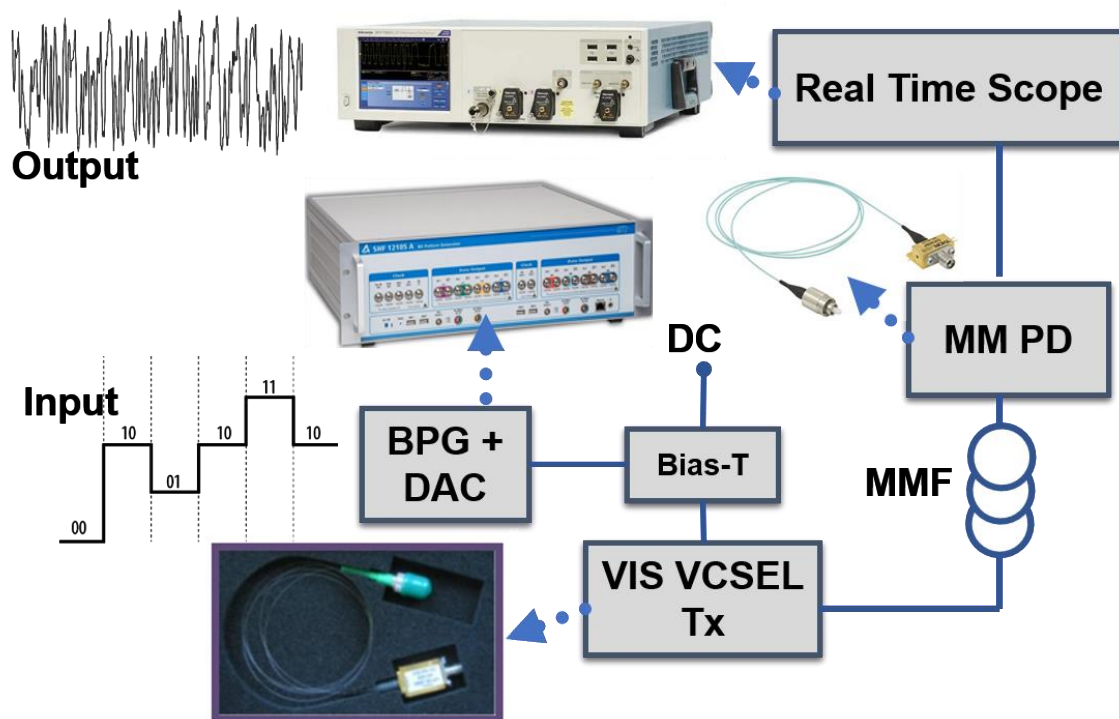
Ο ακριβής χαρακτηρισμός και η προσομοίωση της δυναμικής συμπεριφοράς των οπτικών πομπών διαδραματίζουν κρίσιμο ρόλο στη βελτιστοποίηση της απόδοσής τους και στον σχεδιασμό ισχυρών συστημάτων. Το Laser Εκπομπής Επιφανείας Κατακόρυφης Κοιλότητας (Vertical Cavity Surface Emitting Laser, VCSEL) και η χρήση του ως πομπού που χρησιμοποιείται ευρέως στις διασυνδέσεις κέντρων δεδομένων είναι ένα παράδειγμα ενός πολύπλοκου μη γραμμικού συστήματος που παρουσιάζει επίσης μηχανισμούς περιορισμού του εύρους ζώνης που αποδίδονται στην απόκριση συχνότητας και στις παρασιτικές επιδράσεις της σε εφαρμογές επικοινωνιών. Με τεχνικές αριθμητικής προσομοίωσης που ακολουθούν τις συμβατικές εξισώσεις ρυθμού μπορεί μόνο εν μέρει να προσεγγιστεί η δυναμική συμπεριφορά τέτοιων υποσυστημάτων.

Τα τελευταία χρόνια, η ενσωμάτωση των τεχνικών βαθιάς μάθησης έχει φέρει επανάσταση σε διάφορους τομείς, συμπεριλαμβανομένων των συστημάτων οπτικών επικοινωνιών όπως είδαμε και στα προηγούμενα κεφάλαια, επιτρέποντας επιπλέον την αποτελεσματική μοντελοποίηση και πρόβλεψη πολύπλοκων φαινομένων των χαρακτηριστικών διάδοσης σε πολύτροπη οπτική ίνα, οδηγώντας σε σημαντικές βελτιώσεις στις επιδόσεις [163]. Χρησιμοποιώντας διάφορες αρχιτεκτονικές νευρωνικών δικτύων για τη μοντελοποίηση της μη γραμμικής διάδοσης οπτικών παλμών [164] και φωτονικών δομών [165] καθίσταται δυνατή η ακριβής προσομοίωση και πρόβλεψη της περίπλοκης δυναμικής των οπτικών συστημάτων.

Στα προηγούμενα κεφάλαια παρουσιάστηκε αναλυτικά η δυνατότητα εκμετάλλευσης των εκτεταμένων δυνατοτήτων των βαθιών νευρωνικών δικτύων και ιδιαίτερα των επαναλαμβανόμενων (RNN) για την εξισορρόπηση των μη γραμμικών φαινομένων σε συστήματα οπτικών επικοινωνιών μεγάλων και μεσαίων αποστάσεων. Οι επιδόσεις τους δοκιμάστηκαν και αξιολογήθηκαν τόσο σε μοντέλα προσομοίωσης όσο και σε πειραματικά συστήματα οπτικών επικοινωνιών.

Στο κεφάλαιο αυτό παρουσιάζεται μία άλλη ενδιαφέρουσα λειτουργία των RNN. Εξετάζεται η ικανότητά τους να προβλέψουν και να αναπαράγουν τα δυναμικά χαρακτηριστικά ενός πραγματικού οπτικού συστήματος. Χρησιμοποιούνται δύο τύποι αμφίδρομων RNN, το bi-LSTM και το bi-VRNN για να προσομοιωθεί η δυναμική συμπεριφορά ενός πειραματικά χαρακτηρισμένου οπτικού πομπού VCSEL. Αυτές οι αρχιτεκτονικές επιλέχθηκαν για την ικανότητά τους να αποτυπώνουν αποτελεσματικά τις χρονικές εξαρτήσεις και την πολύπλοκη δυναμική του πομπού. Οι προσομοιώσεις μας πετυχαίνουν μια αξιοσημείωτα υψηλή ακρίβεια πρόβλεψης που πλησιάζει το 100%, αποδεικνύοντας την αποτελεσματικότητα της προτεινόμενης προσέγγισης στην ακριβή μοντελοποίηση της συμπεριφοράς του οπτικού συστήματος. Η ερευνητική δουλειά στο πλαίσιο αυτής της συνιστώσας της διδακτορικής διατριβής αποτελεί αποτέλεσμα συνεργασίας με το ερευνητικό τμήμα της NVIDIA και συγκεκριμένα τους Δρ. Ν. Αργύρη, Δρ. Σ. Δρη, Δρ. Δ. Καλαβρουζιώτη και Δρ. Π. Μπακόπουλο.

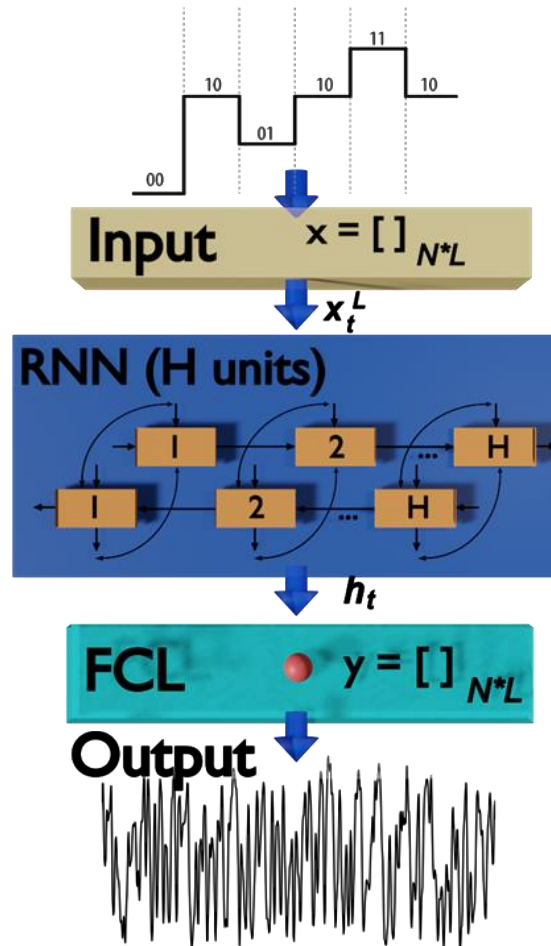
6.2 Πειραματική Διάταξη – Αρχιτεκτονική RNN



Σχήμα 6.1: Πειραματική διάταξη του βασισμένου στο VCSEL οπτικού συστήματος

Η πειραματική οπτική διάταξη που χρησιμοποιήθηκε στα πειράματα απεικονίζεται στο σχήμα 6.1. Οι διαμορφωμένοι με PAM4 ψηφιακοί παλμοί ρυθμού 106.25 Gb/s οδηγούνται στο γεννήτρια προτύπων bit (Bit Pattern Generator, BPG) και μετατρέπονται σε αναλογική μορφή από τον ψηφιακό σε αναλογικό μετατροπέα (DAC). Στη συνέχεια το σήμα οδηγείται στη μονάδα του οπτικού πομπού που βασίζεται σε ένα VCSEL που εκπέμπει στο μήκος κύματος των 850 nm (VIS VCSEL Tx, VM100-850) η οπτική έξοδος του οποίου διαδίδεται σε πολύτροπη οπτική ίνα (Multimode Fiber, MMF) και ανιχνεύεται από φωτοδίοδο (MM PD) εύρους ζώνης 30 GHz. Το ηλεκτρικό σήμα διαβάζεται από παλμογράφο εύρους ζώνης αναλογικού σήματος 33 GHz με ρυθμό 100 GSa/s.

Οι διάφορες περιοχές λειτουργίας του οπτικού συστήματος αντιστοιχούν στις διάφορες τιμές της DC τάσης οδήγησης του VCSEL και ρυθμίζονται από τη Bias-T διάταξη.



Σχήμα 6.2: Διάγραμμα του προσομοιωτή RNN

Για την προσομοίωση του έντονα μη γραμμικού οπτικού συστήματος χρησιμοποιήθηκαν τα αμφίδρομα επαναλαμβανόμενα νευρωνικά δίκτυα bi-LSTM και bi-Vanilla-RNN (bi-VRNN) ως ιδιαίτερα ικανοί μη γραμμικοί επεξεργαστές οι ιδιότητες των οποίων έχουν αναλυθεί εκτενώς στα προηγούμενα κεφάλαια. Το διάγραμμα του RNN προσομοιωτή απεικονίζεται στο σχήμα 6.2. Το RNN με H κρυφές μονάδες τροφοδοτείται στην είσοδο του με τις ψηφιακές PAM4 συμβολοσειρές. Η έξοδος h_t του RNN οδηγείται σε ένα στρώμα FC, το οποίο παράγει την έξοδο του δικτύου y με τη μέθοδο της παλινδρόμησης. Τόσο το bi-LSTM όσο και το bi-VRNN λαμβάνουν τα αρχικά σύμβολα PAM-4 ως είσοδο και

εκπαιδεύονται να προβλέπουν την έξοδο του πομπού VCSEL με βάση τις πειραματικά καταγεγραμμένες από τον παλμογράφο χρονοσειρές.

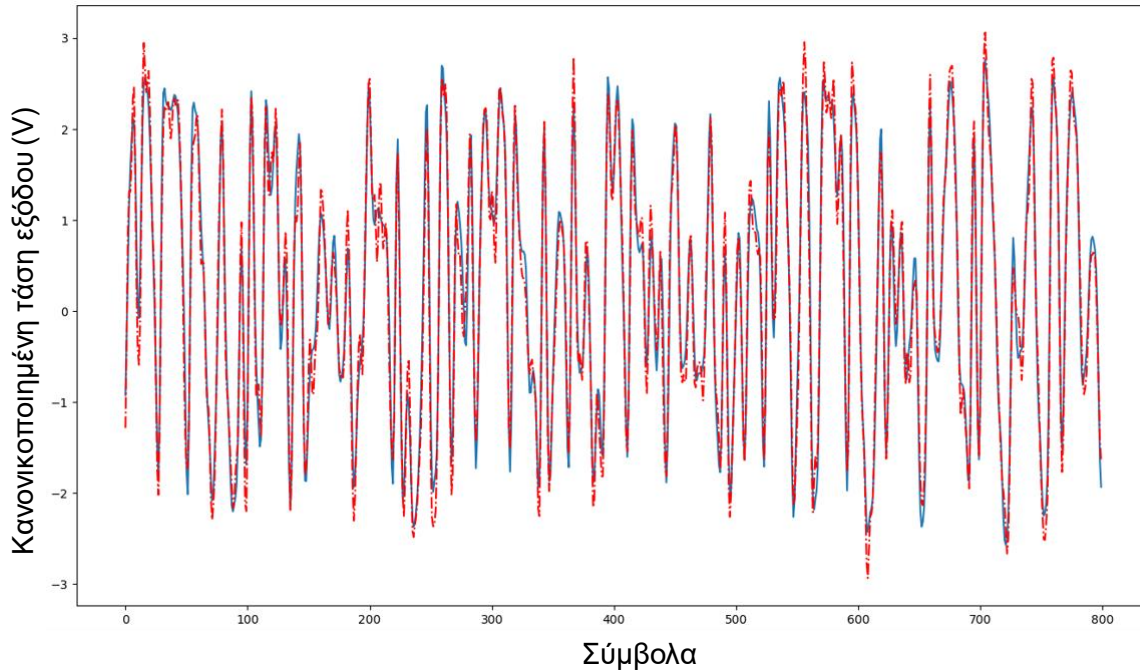
Για κάθε μια από τις περιοχές λειτουργίας του οπτικού συστήματος χρησιμοποιούνται δύο διαφορετικές συμβολοσειρές ενός εκατομμυρίου διαμορφωμένων PAM4 συμβόλων. Από την πρώτη χρησιμοποιούμε το 80% των συμβόλων για εκπαίδευση και το υπόλοιπο 20% για επαλήθευση, ενώ χρησιμοποιούμε τη δεύτερη για αξιολόγηση. Η παραγωγή των συμβολοσειρών εισόδου πραγματοποιείται με τη γεννήτρια τυχαίων αριθμών Mersenne Twister με περίοδο $2^{19937}-1$ συνεπώς διασφαλίζεται η τυχαιότητα και η έλλειψη επαναλήψεων και μοτίβων που θα μπορούσε να νοθεύσει την εκπαίδευση.

Η εκπαίδευση πραγματοποιείται σε πακέτα των 1000 λέξεων μήκους $L = 800$ συμβόλων για 1000 εποχές με τη βοήθεια της βιβλιοθήκης Tensorflow GPU. Για την ενημέρωση των βαρών του δικτύου επιλέγεται ο αλγόριθμος βελτιστοποίησης Adam με ρυθμό εκμάθησης 10^{-4} . Η εκπαίδευση βασίζεται στην παλινδρόμηση με το μέσο τετραγωνικό σφάλμα (MSE) ως συνάρτηση απώλειας.

Χρησιμοποιώντας 28 και 30 κρυφές μονάδες για το bi-LSTM και το bi-VRNN αντίστοιχα, τα δύο δίκτυα κατάφεραν να αναπαράγουν τη δυναμική συμπεριφορά του VCSEL συστήματος με πολύ μεγάλη ακρίβεια. Η κανονικοποιημένη τιμή MSE κυμαίνεται μεταξύ του 0.02 και 0.045. Στο σχήμα 6.3 απεικονίζεται χαρακτηριστικά ένα μικρό τμήμα πραγματικής χρονοσειράς όπως καταγράφηκε από τον παλμογράφο και της χρονοσειράς που προέβλεψε ο RNN προσομοιωτής. Είναι φανερό η ακρίβεια της πρόβλεψης ακόμη και έντονων μη γραμμικών διακυμάνσεων.

Δεν παρατηρήθηκαν σημαντικές ή συστηματικές διαφορές στις επιδόσεις των δύο τύπων RNN οδηγώντας σε μια προτίμηση για το ελαφρύτερο μοντέλο bi-VRNN στη συγκεκριμένη περίπτωση. Ωστόσο, για την προσομοίωση πιο σύνθετων συστημάτων με εντονότερα μη γραμμικά φαινόμενα, θα μπορούσε να χρησιμοποιηθεί το LSTM εφόσον οι αυξημένες υπολογιστικές απαιτήσεις δεν αποτελούν πρόβλημα.

Με την εξαγωγή και την αποθήκευση των βαρών σε έναν πίνακα αναζήτησης (LUT), καθίσταται δυνατή η προσομοίωση της συμπεριφοράς του μη γραμμικού οπτοηλεκτρονικού συστήματος με αξιοσημείωτη ακρίβεια.



Σχήμα 6.3: Προβλεφθείσα (μπλε) και πειραματική (κόκκινο) χρονοσειρά οπτικής διάδοσης. Απεικονίζεται η κανονικοποιημένη στις στάθμες αναφοράς των PAM-4 συμβόλων τάση εξόδου για ένα τυχαίο δείγμα λέξης των 800 συμβόλων

Στις περιπτώσεις όπου το επιθυμητό σημείο λειτουργίας δεν περιλαμβάνεται άμεσα στο LUT, μια απλή γραμμική παρεμβολή μπορεί να εφαρμοστεί στα βάρη των πλησιέστερων καταγεγραμμένων σημείων λειτουργίας για κάθε σημαντική παράμετρο (όπως ο ρυθμός διαμόρφωσης, ρεύμα πόλωσης, θερμοκρασία, βάθος διαμόρφωσης, κτλ). Αυτό επιτρέπει την επέκταση της προτεινόμενης προσέγγισης προσομοίωσης σε διάφορες συνθήκες, ενισχύοντας την ευελιξία και την εφαρμογή της σε πρακτικά σενάρια.

6.3 Προοδευτική Εκπαίδευση

Στην προηγούμενη μελέτη, κάθε σημείο λειτουργίας αντιμετωπίζονταν ως νέο και η εκπαίδευση του νευρωνικού δικτύου πραγματοποιούνταν από μηδενική βάση. Παρά το γεγονός ότι η προσέγγιση αυτή είναι η πιο στέρεη καθώς σε κάθε βήμα πιστοποιείται η

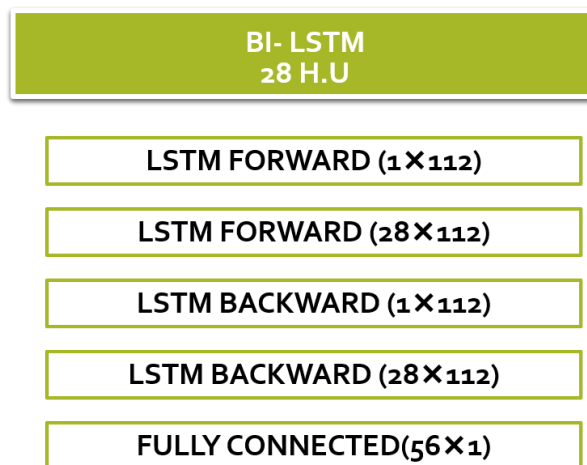
ορθή λειτουργία του μοντέλου και μπορούν να εντοπιστούν πιθανές αστοχίες, είναι αρκετά κοστοβόρα, ιδιαίτερα αν το υπό μοντελοποίηση σύστημα έχει σύνθετη δυναμική συμπεριφορά. Επίσης εφόσον η εκπαίδευση γίνεται από μηδενική βάση, ο SGD αλγόριθμος μπορεί να οδηγήσει τα βάρη σε εντελώς διαφορετικές κατευθύνσεις για κάθε σημείο λειτουργίας, δυσχεραίνοντας σε σημαντικό βαθμό την διαδικασία της γραμμικής παρεμβολής που αναφέρθηκε παραπάνω.

Τη λύση σε αυτό το πρόβλημα δίνει η μέθοδος της προοδευτικής εκπαίδευσης (Online / Incremental Training). Η προοδευτική εκπαίδευση αναφέρεται στη διαδικασία συνεχούς ενημέρωσης και βελτίωσης ενός μοντέλου μηχανικής μάθησης με την πάροδο του χρόνου χρησιμοποιώντας νέα δεδομένα. Αντί να εκπαιδεύεται το μοντέλο από την αρχή με ολόκληρο το σύνολο δεδομένων, η προοδευτική εκπαίδευση επιτρέπει στο μοντέλο να μαθαίνει από τα νέα δεδομένα, διατηρώντας παράλληλα τη γνώση που αποκτήθηκε από την προηγούμενη εκπαίδευση. Αυτή η προσέγγιση είναι ιδιαίτερα χρήσιμη σε σενάρια όπου τα δεδομένα εξελίσσονται συνεχώς και το μοντέλο πρέπει να προσαρμοστεί στις αλλαγές ή τη δυναμική των δεδομένων, όπως συμβαίνει στην περίπτωση που εξετάζουμε, στην οποία η νέα συνθήκη λειτουργίας του οπτικού συστήματος προκύπτει από την αλλαγή σε μικρό βαθμό μιας παραμέτρου, της τάσης του laser.

Με τον τρόπο αυτό κατορθώνουμε να εξοικονομήσουμε υπολογιστικούς πόρους και χρόνο εκπαίδευσης. Συγκεκριμένα, χρησιμοποιώντας τα βάρη που προκύπτουν από τη διαδικασία της εκπαίδευσης των 1000 εποχών βασιζόμενοι στην αρχική συνθήκη λειτουργίας ($V_1=1.1$ V όπου V η τάση του VCSEL), εφαρμόζουμε τη μέθοδο της προοδευτικής εκπαίδευσης στα δεδομένα της επόμενης συνθήκης λειτουργίας $V_2=1.2$ V. Ο αριθμός των εποχών που απαιτούνται τώρα δεν ξεπερνά τις 100. Με τον τρόπο αυτό μειώνονται περίπου στο 10% οι απαιτήσεις σε υπολογιστικούς πόρους. Όσον αφορά τις επιδόσεις της παλινδρόμησης, εφαρμόζοντας διαδοχικά τα βάρη που έχουν προκύψει από μηδενικής βάσης εκπαίδευση και αυτά που προκύπτουν από την προοδευτική εκπαίδευση, σε όλες τις περιπτώσεις προκύπτει πως η τιμή σφάλματος MSE είναι σχεδόν η ίδια, ($MSE \approx 0.02-0.045$). Αυτό υποδηλώνει αφενός πως δεν υπάρχει ένα μόνο σύνολο βαρών που επιλύει το πρόβλημα και αφετέρου πως η μέθοδος της προοδευτικής

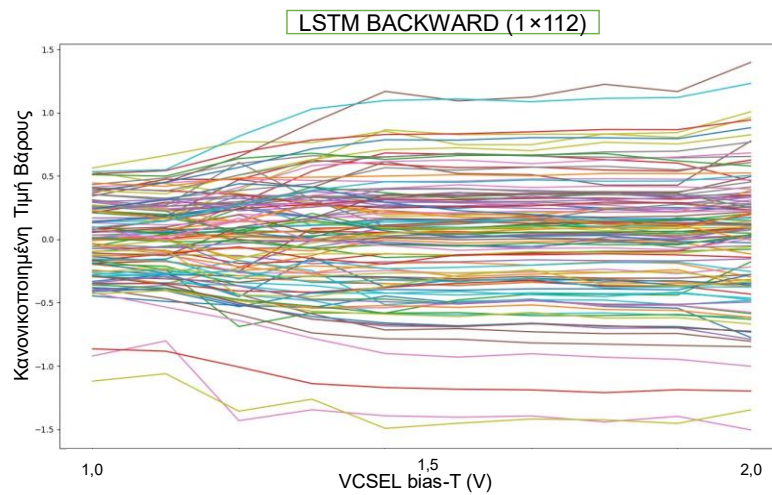
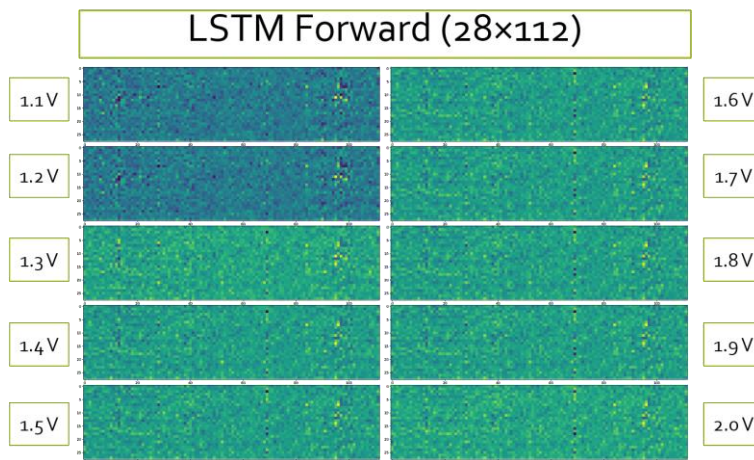
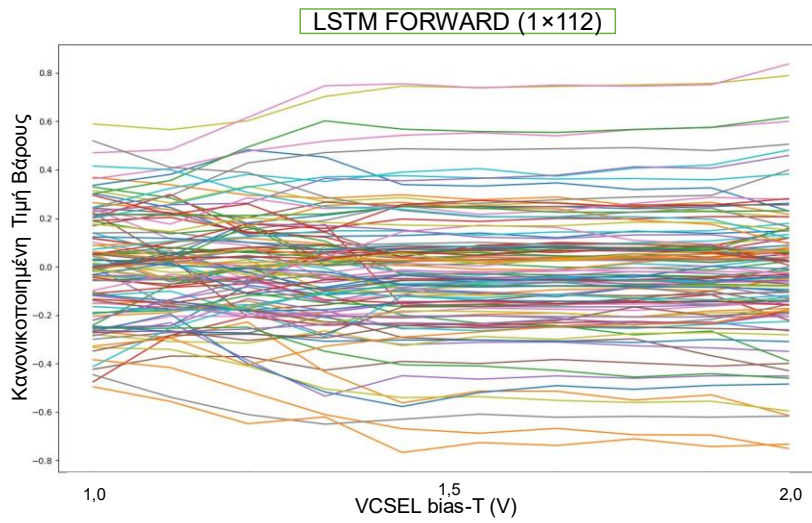
εκπαίδευσης παράσχει ισάξια αποτελέσματα μειώνοντας δραστικά τον χρόνο εκπαίδευσης, ενώ διευκολύνει την παρακολούθηση της εξέλιξης των βαρών.

Με σκοπό να μελετηθεί λεπτομερώς η προηγούμενη διαπίστωση, ακολουθούμε την ίδια διαδικασία για όλες τις διαθέσιμες συνθήκες λειτουργίας και εκπαιδεύουμε ενδεικτικά ένα δίκτυο bi-LSTM των 28 κρυφών μονάδων. Η διάσταση των βαρών των επιμέρους επιπέδων απεικονίζεται στον σχήμα 6.4. Στο συγκεκριμένο νευρωνικό δίκτυο συναντάμε 5 επίπεδα βαρών και συγκεκριμένα πίνακες διαστάσεων 1×112 και 28×112 για το LSTM εμπροσθοδιάδοσης και οπισθοδιάδοσης αντίστοιχα και 56×1 για το πλήρως συνδεδεμένο στρώμα (FC).

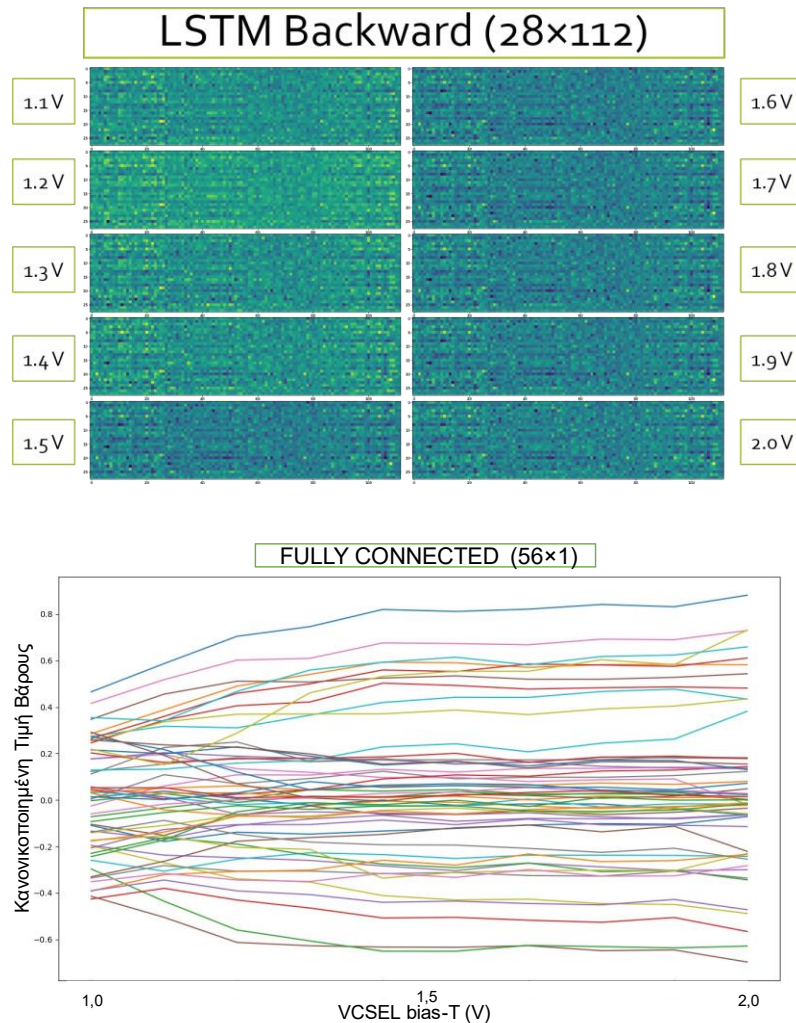


Σχήμα 6.4: Η διάσταση των βαρών των επιμέρους επιπέδων ενός bi-LSTM των 28 h.u

Η εξέλιξη των βαρών του για το σύνολο των διαθέσιμων συνθηκών λειτουργίας κάθε επιπέδου απεικονίζεται στο σχήμα 6.5. Από το τελευταίο φαίνεται πως σε όλα τα επίπεδα η μετάβαση από τη συνθήκη λειτουργίας $V_2=1.2$ V σταδιακά προς την $V_3=1.3$ V και $V_4=1.4$ V η μεταβολή των βαρών είναι έντονη συγκριτικά με τη μετάβαση στις επόμενες συνθήκες λειτουργίας υποδεικνύοντας και την περιοχή με τα πιο έντονα δυναμικά χαρακτηριστικά του VCSEL. Φαίνεται επίσης πως η μεταβολή είναι καθορισμένη και στις περισσότερες περιπτώσεις γραμμική. Το εύρημα αυτό ενισχύει τον ισχυρισμό ότι μπορούν να μοντελοποιηθούν περιοχές λειτουργίας για τις οποίες δεν έχει προηγηθεί εκπαίδευση του μοντέλου.



Σχήμα 6.5α: Η εξέλιξη των βαρών των επιμέρους επιπέδων ενός bi-LSTM των 28 h.u



Σχήμα 6.5β: Η εξέλιξη των βαρών των επιμέρους επιπέδων ενός bi-LSTM των 28 h.u

Στη συνέχεια, με σκοπό να διερευνηθεί η συνεισφορά κάθε επιπέδου στην απεικόνιση των δυναμικών χαρακτηριστικών του οπτικού συστήματος, διαταράσσουμε τα επιμέρους βάρη του, προσθέτοντας σε αυτά θόρυβο κανονικής κατανομής και τυπικής απόκλισης $\sigma=0.5 \cdot \mu$ όπου μ η μέση τιμή των βαρών, διατηρώντας τα υπόλοιπα αναλλοίωτα. Τα αποτελέσματα απεικονίζονται στο σχήμα 6.6. Σε αυτό φαίνεται πως διαταράσσοντας τους πίνακες βαρών των ακραίων στρωμάτων και συγκεκριμένα του πρώτου επιπέδου του εμπροσθοδιαδομένου LSTM που αποτελεί την είσοδο του δικτύου και του πλήρως

συνδεδεμένου στρώματος που αποτελεί την έξοδο του δικτύου, η ακρίβεια πρόβλεψης υποβαθμίζεται σημαντικά ($MSE= 0.14\pm 0.08$ και $MSE= 0.25\pm 0.12$ αντίστοιχα). Αντιθέτως τα ενδιάμεσα στρώματα παρουσιάζουν μια σχετική ανοχή στην απόκλιση των βαρών τους. Ιδιαίτερα το επίπεδο του οπισθοδιαδιδόμενου LSTM διάστασης 28×112 φαίνεται πως είναι ιδιαίτερα ανεκτικό στις διακυμάνσεις των βαρών του καθώς υποβαθμίζει λιγότερο τις επιδόσεις πρόβλεψης ($MSE= 0.043\pm 0.003$).

Το τελευταίο μας οδηγεί στο συμπέρασμα πως τα σημαντικότερα επίπεδα του νευρωνικού δικτύου είναι η είσοδος και η έξοδος, ενώ η συνεισφορά των εσωτερικών επιπέδων μειώνεται σταδιακά, όπως υποδηλώνεται από την ένταση του χρώματος του σχήματος 6.6. Η διαπίστωση αυτή δεν υποβαθμίζει τη συνεισφορά του αμφίδρομου LSTM. Περισσότερο υποδηλώνει την σταδιακή ισχυροποίηση του νευρωνικού δικτύου όσο αυτό οδηγείται προς την έξοδο, με το περισσότερο φορτίο να αναλαμβάνουν τα αρχικά επίπεδα. Μέσω της διαδικασίας της βελτίωσης υπερπαραμέτρων θα μπορούσε να βρεθεί το απαραίτητο βάθος του νευρωνικού δικτύου για τη μέγιστη ακρίβεια προσομοίωσης.

	M.S.E
Αρχικά (ιδανικά) βάρη	0,0415
noisy LSTM Forward 1x112	0,14 ± 0,08
noisy LSTM Forward 28x112	0,10 ± 0,04
noisy LSTM Backward 1x112	0,06 ± 0,02
noisy LSTM Backward 28x112	0,043 ± 0,003
noisy Fully Connected 56x1	0,25 ± 0,12

Σχήμα 6.6: Επίδραση των επιμέρους στρωμάτων στην ακρίβεια της πρόβλεψης, μέσω της προσθήκης θορύβου στον πίνακα των βαρών

6.4 Συμπεράσματα

Στο κεφάλαιο αυτό χρησιμοποιήθηκαν τα δίκτυα RNN για την προσομοίωση της δυναμικής συμπεριφοράς ενός πραγματικού οπτικού συστήματος διάδοσης, βασιζόμενου σε laser επιφανειακής εκπομπής, VCSEL ρυθμού 106.25 Gb/s και PAM4 διαμόρφωσης.

Αποδείχθηκε πως τα RNN δίκτυα είναι ικανά να αναπαράγουν τα δυναμικά χαρακτηριστικά του οπτικού συστήματος με ακρίβεια πρόβλεψης που πλησιάζει το 100%.

Εφαρμόζοντας τη μέθοδο της προοδευτικής εκπαίδευσης επιτεύχθηκε μείωση του υπολογιστικού φόρτου κατά 90% χωρίς υποβάθμιση της ακρίβειας πρόβλεψης ενώ υποβοηθείται η προσομοίωση περιοχών λειτουργίας του οπτικού συστήματος στις οποίες το διαθέσιμο νευρωνικό δίκτυο δεν έχει εκπαιδευτεί, μέσω της γραμμικής παρεμβολής των γειτονικών βαρών.

Τέλος μελετήθηκε η συνεισφορά των επιμέρους επιπέδων του αμφίδρομου LSTM στις επιδόσεις της εξομοίωσης.

7° ΚΕΦΑΛΑΙΟ

Σύνοψη

7.1 Συνεισφορά της Διατριβής στις εφαρμογές των οπτικών επικοινωνιών

Στα προηγούμενα κεφάλαια παρουσιάστηκε λεπτομερώς η πρόταση του ψηφιακού επεξεργαστή βασισμένου σε επαναλαμβανόμενα νευρωνικά δίκτυα RNN και η χρήση του αφενός ως αυτόνομης μονάδας επεξεργασίας για την εξίσωση των μη γραμμικών φαινομένων διάδοσης στα σύμφωνα συστήματα μεγάλων αποστάσεων και στα IM/DD συστήματα ως 100 km, αφετέρου ως μοντέλου προσομοίωσης μη-γραμμικών οπτικών συστημάτων κέντρων δεδομένων.

Στο πλαίσιο εκπόνησης αυτής της διατριβής δημοσιεύτηκαν επιστημονικές εργασίες, τα αποτελέσματα των οποίων έχουν προσελκύσει το ενδιαφέρον άλλων ερευνητικών ομάδων, συνεισφέροντας με τον τρόπο αυτό στην ερευνητική κοινότητα που ασχολείται με την ψηφιακή επεξεργασία σήματος στα συστήματα οπτικών επικοινωνιών.

Συγκεκριμένα το 2020 προτείναμε για πρώτη φορά τη χρήση ενός αμφίδρομου LSTM δικτύου ως υποσύστημα μετα-επεξεργασίας για την αντιστάθμιση των μη γραμμικοτήτων σε ψηφιακά σύμφωνα WDM συστήματα [1], ενώ το 2021 παρείχαμε μια συγκριτική ανάλυση τριών καθιερωμένων μοντέλων RNN, των bi-LSTM, bi-GRU και Vanilla bi-RNN σε σύμφωνα συστήματα οπτικών επικοινωνιών μεγάλων αποστάσεων και δείξαμε ότι με τη χρήση της εκπαίδευσης «πολλά προς πολλά», μπορούμε να μειώσουμε δραστικά την πολυπλοκότητα των μοντέλων RNN [2], [6].

Τα τελευταία αποτελέσματα έχουν υιοθετηθεί για τη μείωση της πολυπλοκότητας αυτού του πολλά υποσχόμενου συστήματος και την πρακτική εφαρμογή του σε υλικό περιορισμένων δυνατοτήτων [166], [167], [168]. Το Bi-RNN που βασίζεται στην

αρχιτεκτονική εκπαίδευσης «πολλά σε πολλά» έχει συγκριθεί πρόσφατα με διάφορα άλλα επαναλαμβανόμενα ή μη δίκτυα όσον αφορά την πολυπλοκότητα και την απόδοση εξισορρόπησης σε συστήματα μετάδοσης μεγάλων αποστάσεων.

Όπως αναφέρθηκε στις εργασίες [2] και [6] και αναπτύχθηκε στο κεφάλαιο 4, οι κύριες παράμετροι που επηρεάζουν την πολυπλοκότητα των μοντέλων bi-RNN είναι ο αριθμός των κρυφών μονάδων, ενώ το bi-VRNN είναι προτιμότερο όσον αφορά την πολυπλοκότητα σε σύγκριση με το bi-LSTM ή το bi-GRU. Μια ολοκληρωμένη εργασία έχει πραγματοποιηθεί στην αναφορά [40] όπου διαφορετικές αρχιτεκτονικές συγκρίνονται μεταξύ τους ως προς την πολυπλοκότητα και την απόδοση εξισορρόπησης. Η εργασία αυτή έδειξε ότι ο συνδυασμός συνελκτικών νευρωνικών δικτύων (CNN) με LSTM μπορεί να προσφέρει την καλύτερη επίδοση BER στην υψηλότερη πολυπλοκότητα, ενώ η πιο πρακτική λύση είναι ένα πολυστρωματικό perceptron (MLP). Το bi-LSTM σε αυτή τη μελέτη πέτυχε σχεδόν πανομοιότυπες επιδόσεις με αυτές του CNN-LSTM σε χαμηλότερη πολυπλοκότητα. Παρόλα αυτά οι συγγραφείς δεν έλαβαν υπόψη ότι η εκπαίδευση πολλών προς πολλά μπορεί να βελτιώσει περαιτέρω την πολυπλοκότητα για τα μοντέλα bi-RNN και να την καταστήσει συγκρίσιμη με αυτή του MLP.

Μια προσέγγιση που κέντρισε το ενδιαφέρον της ερευνητικής κοινότητας είναι η βαθιά μάθηση από άκρο σε άκρο (end to end) όπως αυτή εισήχθη στην εργασία [169]. Αυτή η προσέγγιση συνδέει τον πομπό, το κανάλι επικοινωνίας και τον δέκτη μαζί, αντιμετωπίζοντας το σύστημα επικοινωνίας ως μια ολοκληρωμένη οντότητα. Αξιοποιεί τις δυνατότητες των αυτό-κωδικοποιητών (autoencoders) για τη βελτιστοποίηση του συνολικού οπτικού συστήματος και έχει κερδίσει έδαφος ως πρόταση που μετασχηματίζει τον αστερισμό της γεωμετρικής διαμόρφωσης με σκοπό την ελαχιστοποίηση των μη γραμμικοτήτων [170] ή που ισοσταθμίζει το σύστημα διάδοσης από άκρο σε άκρο [171]. Με την από κοινού βελτιστοποίηση των διαφόρων στοιχείων του συστήματος επικοινωνίας, συμπεριλαμβανομένων των χαρακτηριστικών του πομπού, του δέκτη και του καναλιού, καθίσταται δυνατή η επίτευξη σημαντικών βελτιώσεων όσον αφορά την επίδοση BER. Εντούτοις η προσέγγιση αυτή είναι ιδιαίτερα πολύπλοκη και ευαίσθητη στις διακυμάνσεις των συνθηκών λειτουργίας του συστήματος [45].

Μια σημαντική πρόταση για τη μείωση της πολυπλοκότητας της μετα-επεξεργασίας που βασίζεται σε RNN σε σύμφωνα συστήματα είναι η εργασία των Ming κ.α. [166]. Οι συγγραφείς πρότειναν μια παραλλαγή της αμφίδρομης αρχιτεκτονικής του LSTM και προσδιόρισαν ότι είναι σχεδόν παρόμοιο με το bi-LSTM όταν επιλέγεται η προσέγγιση εκπαίδευσης «πολλά προς πολλά», τόσο σε επίδοση όσο και σε πολυπλοκότητα.

Στην εργασία της αναφοράς [168] οι συγγραφείς παρουσίασαν μια εναλλακτική μέθοδο που αναπτύσσει έναν κωδικοποιητή CNN και λειτουργεί παράλληλα με ένα Vanilla-RNN μονής κατεύθυνσης σχηματίζοντας ένα ενιαίο δίκτυο (CRNN), το οποίο εκπαιδεύεται με την τοπολογία πολλά σε πολλά. Η απόδοση εξισορρόπησης αυτού του σχήματος είναι παρόμοια με αυτή των μοντέλων bi-RNN ενώ η πολυπλοκότητά του είναι σημαντικά βελτιωμένη (50% του bi-RNN). Ωστόσο, η βελτίωση της πολυπλοκότητας εμφανίζεται σε έναν πολύ μεγάλο αριθμό κρυφών μονάδων (>100), κάτι που δεν είναι πρακτικά εφικτό, καθώς η πολυπλοκότητα του μοντέλου εξαρτάται τετραγωνικά από τον αριθμό των κρυφών μονάδων (βλ. παράγραφο 4.2).

Στην εργασία της αναφοράς [172], οι συγγραφείς μελέτησαν μια παραλλαγή του bi-LSTM που χρησιμοποιεί μιγαδικούς πολλαπλασιασμούς αντί για πραγματικούς. Η συγκεκριμένη εργασία δείχνει ότι το μιγαδικό LSTM προσφέρει 28% βελτίωση στην πολυπλοκότητα και 0.2 dB στην απόδοση εξισορρόπησης. Ως εκ τούτου, μια πολύπλοκη υλοποίηση LSTM σε ένα πρακτικό σύστημα φαίνεται να είναι μια πολλά υποσχόμενη αναβάθμιση. Απαιτούνται επίσης τεχνικές συμπίεσης προκειμένου να μειωθεί περαιτέρω η πολυπλοκότητα των μοντέλων bi-RNN.

Μια σημαντική εργασία από τους Freire et al. στην αναφορά [143] δείχνει πώς μπορούν να ακολουθηθούν διαφορετικές διαδρομές προκειμένου να μειωθεί η πολυπλοκότητα του bi-LSTM αξιοποιώντας την εκπαίδευση πολλά-προς-πολλά, την κβάντιση, την ομαδοποίηση των βαρών, το κλάδεμα συνάψεων και τη συμπίεση με τη βοήθεια της κατά Bayes βελτιστοποίησης. Αποδεικνύεται ξεκάθαρα ότι ένας ισοσταθμιστής που βασίζεται σε bi-RNN μπορεί επίσης να επιτύχει καλύτερη επίδοση, διατηρώντας τον ίδιο βαθμό πολυπλοκότητας με το τμήμα της ψηφιακής αντιστάθμισης της χρωματικής διασποράς, το οποίο είναι ένα σημαντικό βήμα προς την εφαρμογή σε πραγματικά συστήματα αυτής της οικογένειας μη γραμμικών ισοσταθμιστών.

Αντίστοιχα, στο πεδίο εφαρμογής των IM/DD συστημάτων, το 2021 δείξαμε την πρώτη πειραματική εφαρμογή ισοστάθμισης βασιζόμενης σε bi-GRU [7], ενώ λίγους μήνες αργότερα, το 2022, σε μια εκτενή μελέτη δημοσιεύσαμε πειραματικά αποτελέσματα διάδοσης ως και 100km CWDM σημάτων στη ζώνη-O με ρυθμό $4 \times 50\text{-Gb/s}$ [3]. Αποδείχθηκε η σημαντικά καλύτερη επίδοση BER του σχήματος Vanilla-RNN σε σχέση με τον συμβατικό ισοσταθμιστή ανάδρασης απόφασης (DFE) και το μη-γραμμικό Volterra 3ης τάξης, για διαμορφώσεις OOK και PAM4.

Όπως στα σύμφωνα συστήματα, έτσι και στα IM/DD οι τεχνικές που βασίζονται στο RNN παραμένουν από τις πιο πολλά υποσχόμενες λύσεις αποτελεσματικής μη γραμμικής ισοστάθμισης, καθώς το οπτικό κανάλι στις μικρές αποστάσεις παραμένει ένα μη γραμμικό κανάλι με μνήμη. Στην εργασία της αναφοράς [173] οι συγγραφείς καταγράφουν τις σημαντικότερες προτάσεις για την εξισορρόπηση IM/DD με συστήματα αναφέροντας ανάμεσα σε αυτές την πρόταση του GRU και του VRNN.

Στην εργασία της αναφοράς [52] οι συγγραφείς απέδειξαν πώς η έξοδος πολλαπλών συμβόλων LSTM, η οποία λειτουργεί με παρόμοιο τρόπο με την εκπαίδευση πολλά σε πολλά, μπορεί να επιφέρει ανώτερη απόδοση εξισορρόπησης σε συστήματα IM/DD ρυθμού 200 Gb/s σε σύγκριση με τους μη γραμμικούς ισοσταθμιστές Volterra.

Οι περισσότερες από τις μελέτες που εμφανίζονται στη βιβλιογραφία δείχνουν ξεκάθαρα ότι τα σωστά επιλεγμένα μοντέλα RNN έχουν την ιδιότητα να βελτιώνουν την επίδοση BER διατηρώντας παράλληλα την πολυπλοκότητα σε μέτρια επίπεδα σε σύγκριση με τους πλέον σύγχρονους μη γραμμικούς ισοσταθμιστές σε περιβάλλον IM/DD.

7.2 Συμπεράσματα και μελλοντικά βήματα

Η προαναφερθείσα ανάλυση έδειξε ότι το bi-RNN βρίσκεται στο επίκεντρο της ερευνητικής κοινότητας ως μη-γραμμικός ισοσταθμιστής μετα-επεξεργασίας σε σύμφωνα συστήματα μεγάλων αποστάσεων. Το αρχικό σχήμα ή οι παραλλαγές του αποτελούν μια ιδιαίτερα ελκυστική πρόταση για την εφαρμογή προηγμένης εξισορρόπησης των μη γραμμικών φαινομένων της διάδοσης σε λογικό κόστος.

Με γνώμονα τις κύριες συνεισφορές που παρουσιάζονται σε αυτή τη διατριβή, μπορούν να ακολουθηθούν αρκετές νέες ερευνητικές κατευθύνσεις. Μερικά από τα ανοιχτά θέματα για μελλοντική εργασία είναι τα εξής :

- Ένα σημαντικό βήμα προς την πρακτική ανάπτυξη των ισοσταθμιστών RNN είναι η εφαρμογή τους σε προγραμματιζόμενα ολοκληρωμένα κυκλώματα FPGA. Πρόσφατη μελέτη [143] έδειξε ότι το bi-LSTM απαιτεί μόνο 2.5 φορές περισσότερους πόρους στο FPGA από τη ψηφιακή μονάδα αντιστάθμισης της χρωματικής διασποράς, ενώ εξακολουθεί να έχει καλύτερη απόδοση από τον ισοσταθμιστή DBP. Ωστόσο, το πιο κρίσιμο ζήτημα παραμένει η δυσκολία παραλληλοποίησης των δομών RNN στο DSP, που αποτελεί σημαντικό εμπόδιο για την ανάπτυξή τους. Η έρευνα πάνω στην αποδοτική ανάπτυξη των RNN ισοσταθμιστών σε FPGA θα δώσει απαντήσεις στη δυνατότητα ή την αδυναμία πρακτικής ενσωμάτωσής τους σε πραγματικά συστήματα οπτικών επικοινωνιών.
- Ακόμη και τα ελαφριά σχήματα bi-RNN απαιτούν αρκετά περισσότερους πολλαπλασιασμούς από αυτόν ενός αλγόριθμου FFE/DFE που μεταφράζεται σε σημαντική αύξηση της κατανάλωσης ενέργειας στην πλευρά του δέκτη (βλ. κεφ 5.3). Ως εκ τούτου, αν και τα μοντέλα RNN αναδεικνύονται ως οι πιο αποτελεσματικοί μη γραμμικοί ισοσταθμιστές από άποψη επίδοσης και πολυπλοκότητας, ακόμη και για άμεση ανίχνευση, εξακολουθούν να υστερούν στην πρακτικότητα και δεν μπορούν να ανταποκριθούν στις αυστηρές απαιτήσεις των οικοσυστημάτων μικρής εμβέλειας όσον αφορά την κατανάλωση ενέργειας και την πολυπλοκότητα. Οι τεχνικές κλαδέματος και κβάντισης [167], μπορούν να

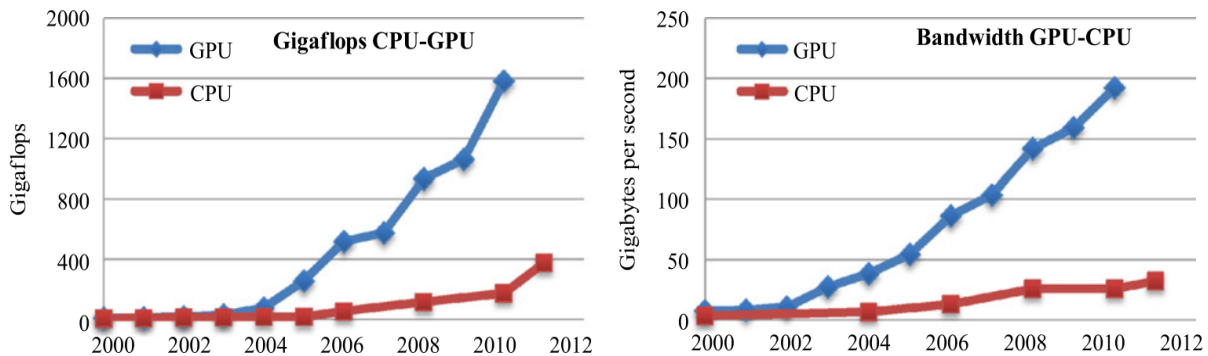
μειώσουν την πολυπλοκότητα, ενώ περαιτέρω έρευνα στην τεχνική της προοδευτικής εκπαίδευσης θα μπορούσε να αποφέρει σημαντικά αποτελέσματα στην πρακτική εφαρμογή των RNN ελαχιστοποιώντας τον χρόνο εκπαίδευσής τους.

- Τα σχετικά συμπεράσματα του κεφαλαίου 6.3 όσον αφορά την διαβάθμιση των εσωτερικών επιπέδων των RNN οδηγούν στην κατεύθυνση εναλλακτικών RNN δομών χαμηλότερης πολυπλοκότητας. Μεταξύ αυτών, το πιο διερευνημένο στην περιοχή των τηλεπικοινωνιών είναι το Reservoir Computing (RC) [174]. Εμπνευσμένο από τον τρόπο λειτουργίας της ανθρώπινης παρεγκεφαλίδας, το RC διαχωρίζει την τυπική δομή ενός RNN σε μια δεξαμενή μη γραμμικών, τυχαία διασυνδεδεμένων κόμβων με μη εκπαιδευσιμα βάρη και ένα γραμμικό επίπεδο ανάγνωσης με δυνατότητα εκπαίδευσης. Μια πιθανή μελλοντική εργασία αποτελεί η σχεδίαση ψηφιακών RC συστημάτων και τη χρήση τους ως ισοσταθμιστών μη γραμμικών φαινομένων και ως προσομοιωτών οπτικών IM/DD συστημάτων.
- Μια εναλλακτική των RNN μοντέλων είναι η χρήση του μηχανισμού προσοχής (attention) [175] η οποία έχει επίσης προταθεί για τη μείωση της πολυπλοκότητας των ισοσταθμιστών που βασίζονται σε RNN [176]. Ο μηχανισμός προσοχής και η παραλλαγή της αυτο-προσοχής (self-attention) αποτελεί βασικό συστατικό των σύγχρονων μοντέλων τεχνητής νοημοσύνης, όπως του μοντέλου Transformer [175], που έχει τους τελευταίους μήνες εκτοξεύσει τη δημοτικότητά του εξαιτίας της εφαρμογής του σε μοντέλα Generative Pre-trained Transformer – GPT, όπως το ChatGPT [177]. Η συγκριτική μελέτη των μοντέλων αυτών με τα RNN στο πεδίο εφαρμογής των οπτικών συστημάτων αποτελεί μια πολλά υποσχόμενη μελλοντική εργασία.

ΠΑΡΑΡΤΗΜΑ

Παράλληλη Επεξεργασία

Π.1 Εισαγωγή στην Παράλληλη Επεξεργασία – GPGPU



Σχήμα Π.1: Ιστορική σύγκριση της θεωρητικής κορυφαίας απόδοσης όσον αφορά το ρυθμό εκτέλεσης των πράξεων κινητής υποδιαστολής και το εύρος ζώνης για τις ταχύτερες διαθέσιμες NVIDIA GPU και CPU της Intel.

Τις τελευταίες δεκαετίες παρατηρείται έντονο ενδιαφέρον για τη χρήση μονάδων επεξεργασίας γραφικών (GPUs) για εφαρμογές διαφορετικές από την επεξεργασία γραφικών. Η χρήση των GPUs έχει πλέον ωριμάσει σε σημείο που υπάρχουν αμέτρητες βιομηχανικές εφαρμογές, ενώ έχει επιτελεστεί μια μεγάλη ανάπτυξη στις γλώσσες και τα εργαλεία προγραμματισμού τους και τη χρήση τους για υπολογισμούς γενικής χρήσης, (General-purpose computing on graphics processing units - GPGPU).

Το κλειδί της επιτυχίας του GPGPU είναι η υπερβολικά αυξημένη απόδοση του σε σχέση με τη χρήση κεντρικής μονάδας επεξεργασίας (Central Processing Unit – CPU). Στο σχήμα Π.1 απεικονίζεται η ιστορική σύγκριση της θεωρητικής κορυφαίας απόδοσης όσον αφορά το ρυθμό εκτέλεσης των πράξεων κινητής υποδιαστολής (Floating points per sec -FLOPS) και το εύρος ζώνης για τις τότε ταχύτερες διαθέσιμες GPU και CPU [178]. Σήμερα το χάσμα έχει διευρυνθεί αρκετά περισσότερο. Αυτό το κενό απόδοσης έχει τις

ρίζες του στους φυσικούς περιορισμούς ανά πυρήνα και στις αρχιτεκτονικές διαφορές μεταξύ των δύο επεξεργαστών. Η CPU είναι στην ουσία ένας σειριακός επεξεργαστής von Neumann [179] και είναι εξαιρετικά βελτιστοποιημένος για να εκτελεί σειριακά εντολές. Ένας από τους σημαντικότερους παράγοντες απόδοσης των CPU ήταν παραδοσιακά η σταθερά αυξανόμενη συχνότητά τους. Διπλασιάζοντας τη συχνότητα λειτουργία τους, διπλασιαζόταν η απόδοση. Στη δεκαετία του 2000, ωστόσο, αυτή η αύξηση σταμάτησε απότομα καθώς προσεγγιζόταν το τείχος ισχύος σύμφωνα με το οποίο η κατανάλωση ενέργειας μιας CPU είναι ανάλογη με τη συχνότητα λειτουργίας υψωμένη στον κύβο. Η επαρκής ψύξη ενός τέτοιου ολοκληρωμένου κυκλώματος ήταν εξίσου απαιτητική συνεπώς η τάση της εκθετικής αύξησης της συχνότητας σταμάτησε λίγο κάτω από τα 4 GHz. Ο σειριακός υπολογισμός είχε φτάσει στο ζενίθ της απόδοσης και οι CPU άρχισαν να αυξάνουν την απόδοση μέσω εντολών πολλαπλών πυρήνων και διανυσμάτων [180].

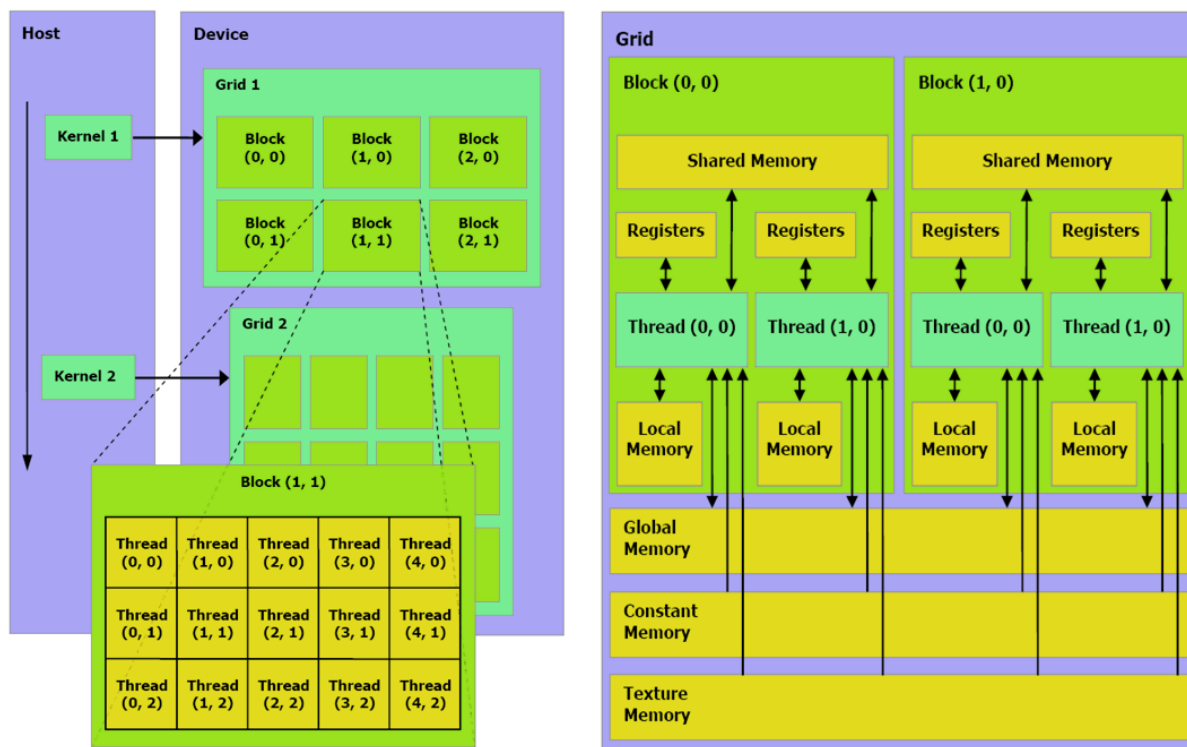
Την ίδια στιγμή η απόδοση των GPUs αυξάνεται εκθετικά εξαιτίας της εκτενούς παραλληλίας.

Οι πολυπύρηνες CPUs αποτελούνται από πολύπλοκους πυρήνες με μεγάλη κρυφή μνήμη (cache memory) οι οποίοι είναι βελτιστοποιημένοι στην εκτέλεση εντολών ενός νήματος (single-thread). Χρησιμοποιώντας την τεχνική hyper-treading μπορούν να ελέγξουν έως δύο νήματα ανά πυρήνα. Αυτό σημαίνει πως ένα μεγάλο τμήμα του ολοκληρωμένου είναι αφιερωμένο στην εκτέλεση σύνθετων εντολών παράλληλης επεξεργασίας, όπως instruction pipelining. Παραμένει ένα μικρό μόνο τμήμα του ολοκληρωμένου για τις μονάδες υπολογισμού πράξεων κινητής υποδιαστολής

Αντιθέτως οι GPUs αποτελούνται από εκατοντάδες απλούστερους πυρήνες που μπορούν να χειριστούν χιλιάδες ταυτόχρονα νήματα υλικού. Έχουν σχεδιαστεί με κατάλληλη αρχιτεκτονική σύμφωνα με την οποία περισσότερα τρανζίστορ σε κάθε πυρήνα είναι αφιερωμένα στον υπολογισμό των πράξεων κινητής υποδιαστολής. Η αρχιτεκτονική σχεδίασης των GPUs σε αντίθεση με τις αντίστοιχες αρχιτεκτονικές των CPUs, στοχεύει στη μεγιστοποίηση της απόδοσης των πράξεων κινητής υποδιαστολής και λιγότερο στην εκτέλεση σύνθετων εντολών.

Ωστόσο, η παραλληλοποίηση αυξάνει μόνο την απόδοση των παράλληλων τμημάτων κώδικα, ενώ το σειριακό μέρος του κώδικα γίνεται σύντομα το σημείο συμφόρησης. Αυτό αναφέρεται συχνά ως νόμος του Amdahl [181]. Έτσι, οι περισσότερες εφαρμογές επωφελούνται από τον ισχυρό συνδυασμό μιας μαζικά παράλληλης GPU και μιας γρήγορης CPU πολλαπλών πυρήνων. Αυτός ο συνδυασμός, γνωστός ως ετερογενής υπολογιστής έχει πρακτικά επικρατήσει στις σύγχρονες εφαρμογές υπολογισμού υψηλής απόδοσης (high-performance computing - HPC).

Ανάμεσα στις διάφορες γλώσσες προγραμματισμού GPU ξεχωρίζουν η CUDA [182] και η OpenCL [183]. Και οι δύο παρουσιάζουν παρόμοια χαρακτηριστικά αλλά μέσω διαφορετικών διεπαφών προγραμματισμού. Η CUDA είναι ένα ιδιόκτητο API και ένα σύνολο επεκτάσεων γλώσσας που λειτουργεί μόνο σε GPU της NVIDIA. Πρόκειται για μια πλατφόρμα παράλληλων υπολογισμών και ένα μοντέλο προγραμματισμού που έχει σχεδιαστεί για να παρέχει τη μεγαλύτερη ευελιξία και απόδοση για εφαρμογές με



Σχήμα Π.2 Το μοντέλο CUDA σε μια GPU. Κάθε μπλοκ αποτελείται από ένα σύνολο νημάτων που μπορούν να επικοινωνούν και να συνεργάζονται και διαχείριση μονάδων μνήμης από τα επιμέρους νήματα των μπλοκς.

επιτάχυνση GPU. Το OpenCL είναι ένα ανοιχτό πρότυπο για παράλληλο προγραμματισμό που χρησιμοποιεί CPUs, GPUs, Επεξεργαστές Ψηφιακού Σήματος (DSP) και άλλους τύπους επεξεργαστών.

Το μοντέλο εκτέλεσης CUDA (Σχήμα Π.2) σε μία GPU βασίζεται στην εκτέλεση ενός πυρήνα (kernel) σε ένα πλέγμα αποτελούμενο από μπλοκ (blocks). Κάθε μπλοκ αποτελείται από νήματα που μπορούν να συγχρονιστούν χρησιμοποιώντας γρήγορη κοινή μνήμη. Αυτό αντιστοιχίζεται στο υλικό έτσι ώστε ένα μπλοκ να εκτελείται σε έναν μόνο πολυεπεξεργαστή και ένας πολυεπεξεργαστής να μπορεί να εκτελέσει πολλαπλά μπλοκ. Οι διαστάσεις του πλέγματος και του μπλοκ μπορεί να είναι μιας, δύο ή τριών διαστάσεων και καθορίζουν τον αριθμό των νημάτων που θα χρησιμοποιηθούν. Κάθε νήμα έχει ένα μοναδικό αναγνωριστικό μέσα στο μπλοκ του και κάθε μπλοκ έχει ένα μοναδικό καθολικό αναγνωριστικό. Αυτά συνδυάζονται για να δημιουργήσουν ένα μοναδικό καθολικό αναγνωριστικό ανά νήμα.

Π.2 Χρήση Προγραμματισμού GPU στην προσομοίωση της Οπτικής Διάδοσης

Για την προσομοίωση συστημάτων διάδοσης σήματος σε οπτική ίνα, τόσο στη ζώνη-Ο όσο και στη ζώνη-C οι εξισώσεις του Manakov περιγράφουν τις δύο συνιστώσες ορθογωνίας πόλωσης του ηλεκτρικού πεδίου E_x και E_y

$$\frac{\partial E_{x,y}}{\partial z} = -\frac{\alpha}{2} E_{x,y} + \frac{j\beta_2}{2} \frac{\partial^2 E_{x,y}}{\partial t^2} - j\gamma \frac{8}{9} (|E_x|^2 + |E_y|^2) E_{x,y} \quad (\text{Π.1})$$

Η πλέον αξιόπιστη μέθοδος επίλυσης της εξίσωσης είναι η ολοκλήρωση της μη γραμμικής εξίσωσης του Schrodinger (NLSE) χρησιμοποιώντας τη μέθοδο Split-step Fourier, όπως αναλύθηκε στο κεφάλαιο 2.

Σημαντικοί παράγοντες για την όσο το δυνατό πιο ακριβή προσομοίωση της διάδοσης είναι το μέγεθος του βήματος της Split-step Fourier, όπως και ο αριθμός των δειγμάτων ανά σύμβολο που διαδίδεται για την αναλογική (οπτική) αναπαράσταση του ψηφιακού σήματος. Τα προηγούμενα, σε συνδυασμό με το γεγονός πως για τις ανάγκες αξιόπιστης εξαγωγής συμπερασμάτων και τη μέτρηση του λόγου σφαλμάτων BER απαιτείται η

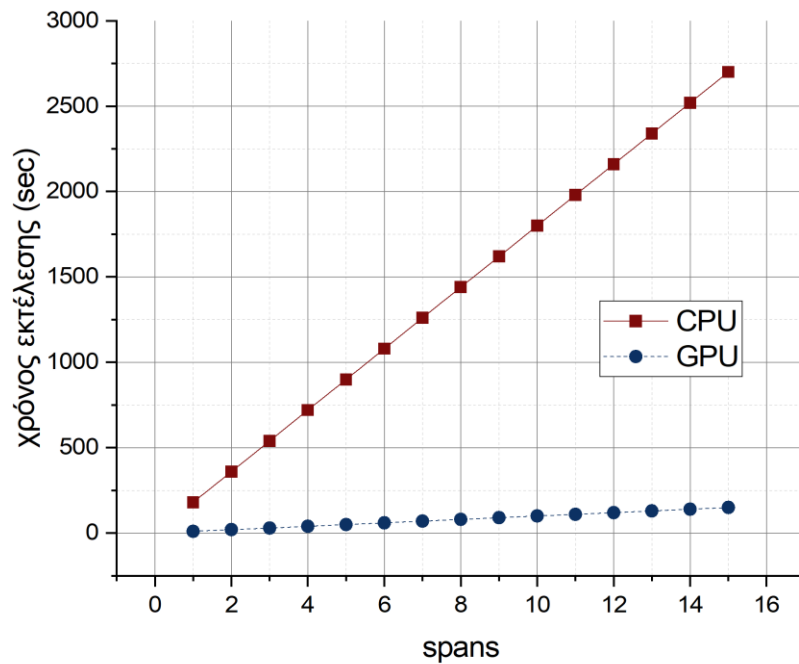
διάδοση αρκετά αυξημένου αριθμού συμβόλων, οδηγούν στην ανάγκη επίλυσης των εξισώσεων NLSE με τεράστιο μέγεθος χρονοσειρών.

Το πρόβλημα μεγεθύνεται ακόμη περισσότερο όταν το μοντέλο διάδοσης αφορά την παράλληλη διάδοση πολλών οπτικών σημάτων με την τεχνολογία της πολυπλεξίας WDM.

Η παράλληλη επεξεργασία με χρήση GPU σε αυτή την περίπτωση επιφέρει σημαντική επιτάχυνση της προσομοίωσης. Το τμήμα της προσομοίωσης που επιδέχεται παραλληλοποίησης είναι η συνάρτηση *span*. Αυτή δέχεται ως εισόδους τα ηλεκτρικά πεδία των δύο οπτικών πολώσεων (E_x, E_y), τον αριθμό των τμημάτων (*spans*) της οπτικής που μεσολαβούν ανά δύο ενισχυτές και τον αριθμό των βημάτων $Nz=L/dz$, όπου L το συνολικό μήκος της ίνας και dz το μήκος του βήματος της μεθόδου Split-Step Fourier. Για κάθε βήμα εφαρμόζεται μετασχηματισμός Fourier, πολλαπλασιάζεται με την εκθετική συνάρτηση διάδοσης η οποία περιέχει τον γραμμικό όρο α των απωλειών ισχύος και β_2 που εκφράζει τη χρωματική διασπορά 2^{ης} τάξης. Έπειτα εφαρμόζεται αντίστροφος μετασχηματισμός Fourier και τέλος αθροίζεται ο θόρυβος ASE ως αποτέλεσμα εφαρμογής οπτικής ενίσχυσης στην απόληξη κάθε τμήματος οπτικής ίνας.

Οι μετασχηματισμοί Fourier μπορούν να επιταχυνθούν χρησιμοποιώντας τη βιβλιοθήκη *cuFFT* (*CUDA Fast Fourier Transform library*). Η μετατροπή της συνάρτησης σε βελτιστοποιημένο κώδικα CUDA kernel γίνεται με τη χρήση του λογισμικού MATLAB - GPU Coder.

Η παραγόμενη συνάρτησης *span.cu* η οποία εκτελείται σε μία κάρτα γραφικών Nvidia 2080-Ti σε σύγκριση με τη συνάρτησης *span.m* η οποία εκτελείται στη κεντρική μονάδα επεξεργασίας (CPU – AMD Ryzen 7 – 3700X – 8cores) αναλύεται στο σχήμα Π.3. Είναι εμφανής η επιτάχυνση της εκτέλεσης του κώδικα με τη χρήση GPU η οποία στο συγκεκριμένο παράδειγμα είναι περίπου 16 – 25 φορές. Ενδεικτικά για την προσομοίωση της διάδοσης 150.000 συμβόλων 10 καναλιών WDM διπλής πόλωσης, σε ίνα μήκους 750 km (15 spans) απαιτούνται 2.700 sec για εκτέλεση στη CPU έναντι 150 sec για εκτέλεση στη GPU.



Σχήμα Π.3 : Χρόνος εκτέλεσης της συνάρτησης span.cu (GPU) σε σχέση με την εκτέλεσης της span.m (CPU)

Π.3 Εκπαίδευση σύνθετων νευρωνικών δικτύων με τη χρήση GPU

Όπως είδαμε στις προηγούμενες παραγράφους τα μοντέλα βαθιάς μάθησης αποτελούνται από εκτεταμένα νευρωνικά δίκτυα. Όσο ο αριθμός των επιπέδων ή των κρυφών μονάδων των RNN αλλά και ο όγκος των δεδομένων εκπαίδευσης αυξάνονται, εκτοξεύεται η υπολογιστική πολυπλοκότητα καθιστώντας επιτακτική την ανάγκη εξεύρεσης αποδοτικότερων μεθόδων εκπαίδευσης [102].

Η παράλληλη επεξεργασία των δεδομένων εισόδου κινείται ακριβώς σε αυτή την κατεύθυνση. Η μονάδα GPU εξαιτίας ακριβώς της αρχιτεκτονικής της είναι ιδανική στην επιτάχυνση της εκπαίδευσης των σύνθετων νευρωνικών δικτύων.

Κοινός παρονομαστής του συνόλου των εργασιών στο πλαίσιο της παρούσας διατριβής, όπως διεξοδικά αναλύθηκαν στα προηγούμενα κεφάλαια, είναι η συνεχής και επαναλαμβανόμενη εκπαίδευση διαφορετικών τοπολογιών νευρωνικών δικτύων με τη βοήθεια του λογισμικού Tensorflow, σε περιβάλλον Python.

Σημαντικός παράγοντας για τη βελτιστοποίηση της ακρίβειας του προς εξέταση μοντέλου είναι το πλήθος των δειγμάτων εκπαίδευσης. Ενδεικτικά χρησιμοποιήθηκαν περίπου 60.000 - 100.000 λέξεις σύμβολων στα μοντέλα RNN των κεφαλαίων 4 και 5. Τα αντίστοιχα δίκτυα που χρησιμοποιήθηκαν στο κεφάλαιο 6 εκπαιδεύτηκαν με 800.000-1.000.000 χρονοσειρές. Εξίσου σημαντικοί παράγοντες είναι το βάθος του RNN μοντέλου εκφρασμένο σε αριθμό κρυφών μονάδων, ο τύπος του (LSTM, GRU, VRNN), η αμφίδρομη αρχιτεκτονική του και η τοπολογία πολλά σε πολλά.

Ο αριθμός των πολλαπλασιασμών που πρέπει να πραγματοποιηθούν τόσο στην εκπαίδευση όσο και στην αξιολόγηση, δεδομένης της αρχιτεκτονικής των διαφόρων μοντέλων RNN των κεφαλαίων 4-6 κυμαίνονται από 50.000 ως 4.000.000. Λαμβάνοντας υπόψη τέλος πως στην περίπτωση της εκπαίδευσης η διαδικασία πρέπει να επαναληφθεί για περίπου 100 ως 2000 εποχές, συμπεραίνουμε πως ο υπολογιστικός φόρτος είναι ιδιαίτερα αυξημένος.

Τη λύση στο συγκεκριμένο πρόβλημα δίνουν οι GPU. Η παράλληλη επεξεργασία που αποδίδεται στο πλήθος των πυρήνων CUDA προσφέρει μια σημαντική συρρίκνωση του απαιτούμενου χρόνου εκπαίδευσης, χρησιμοποιώντας την έκδοση του Tensorflow που υποστηρίζει τη χρήση GPUs.

Στον πίνακα Π.2 αποτυπώνονται ενδεικτικά οι χρόνοι υπολογισμού των συγκρινόμενων τύπων RNN του κεφαλαίου 4.3, bi-LSTM, bi-GRU και bi-VRNN. Το πλήθος των δειγμάτων εκπαίδευσης είναι 200.000 (100.000 για κάθε οπτική πόλωση) λέξεις των 51 συμβόλων. Οι χρόνοι του πίνακα Π.2 αναφέρονται σε μια (1) εποχή ενώ η εκπαίδευση εκτελείται σε πακέτα (batches) των 512 συμβόλων.

Με τη χρήση της/των GPUs παρατηρείται μια χρονική επιτάχυνση που μπορεί να ξεπεράσει τις 50 φορές σε σύγκριση με τον χρόνο επεξεργασίας στη μονάδα CPU.

Αξίζει να σημειωθεί πως με την κατανομημένη επεξεργασία του επεξεργαστικού φόρτου σε δύο GPUs (2xRTX 2080Ti) επιτυγχάνονται ακόμη καλύτερες επιδόσεις καθώς σχεδόν διπλασιάζονται οι διαθέσιμοι CUDA πυρήνες παράλληλης επεξεργασίας.

Τέλος πρέπει να σημειωθεί πως παρότι το VRNN είναι ο λιγότερο πολύπλοκος τύπος RNN, όπως φαίνεται στον χρόνο εκτέλεσης στη μονάδα CPU, η επιτάχυνση με τη βοήθεια της μονάδας GPU δεν είναι το ίδιο αποδοτική με τις υπόλοιπες. Αυτό οφείλεται στο γεγονός πως η ομάδα ανάπτυξης του λογισμικού Tensorflow έχει βελτιστοποιήσει στις τελευταίες διαθέσιμες εκδόσεις την παράλληλη και κατανομημένη εκτέλεση μόνο των μοντέλων LSTM και GRU.

Τύπος RNN	CPU	GPU (Titan RTX 24GB) CUDA cores: 4608	GPUs (2x RTX 2080Ti) CUDA cores: 2x4352
bi-LSTM	184 sec	~5 sec	~4 sec
bi-GRU	168 sec	~4 sec	~3 sec
bi-VRNN	131 sec	~7 sec	~6 sec

Πίνακας Π.2 Απαιτούμενος υπολογιστικός χρόνος για την εκπαίδευση νευρωνικών δικτύων με τη χρήση CPU και GPU(s).

Αναφορές

- [1] S. Deligiannidis, A. Bogris, C. Mesaritakis, and Y. Kopsinis, “Compensation of Fiber Nonlinearities in Digital Coherent Systems Leveraging Long Short-Term Memory Neural Networks,” *J. Light. Technol.*, vol. 38, no. 21, pp. 5991–5999, Nov. 2020, doi: 10.1109/JLT.2020.3007919.
- [2] S. Deligiannidis, C. Mesaritakis, and A. Bogris, “Performance and Complexity Analysis of Bi-Directional Recurrent Neural Network Models Versus Volterra Nonlinear Equalizers in Digital Coherent Systems,” *J. Light. Technol.*, vol. 39, no. 18, pp. 5791–5798, 2021, doi: 10.1109/JLT.2021.3092415.
- [3] Y. Hong *et al.*, “ML-assisted Equalization for 50-Gb/s/ O-band CWDM Transmission over 100-km SMF,” *IEEE J. Sel. Top. Quantum Electron.*, pp. 1–1, 2022, doi: 10.1109/JSTQE.2022.3155990.
- [4] K. Sozos, S. Deligiannidis, G. Sarantoglou, C. Mesaritakis, and A. Bogris, “Recurrent Neural Networks and Recurrent Optical Slicers as Equalizers in High Symbol Rate Optical Transmission Systems,” no. 871330, 2022.
- [5] S. Deligiannidis, K. R. Bottrill, K. Sozos, C. Mesaritakis, P. Petropoulos, and A. Bogris, “Multichannel Nonlinear Equalization in Coherent WDM Systems based on Bi-directional Recurrent Neural Networks,” *J. Light. Technol.*, 2023.
- [6] S. Deligiannidis, C. Mesaritakis, and A. Bogris, “Performance and Complexity Evaluation of Recurrent Neural Network Models for Fibre Nonlinear Equalization in Digital Coherent Systems,” *Ecoc*, no. 1, pp. 1–4, 2020.
- [7] Y. Hong *et al.*, “Performance-enhanced Amplified O-band WDM Transmission using Machine Learning based Equalization,” no. c, pp. 3–4.
- [8] A. Bogris, K. Sozos, S. Deligiannidis, G. Sarantoglou, and C. Mesaritakis, “Machine

- Learning and Neuromorphic Computing Approaches for the mitigation of transmission impairments in high baud rate transmission systems,” in *European Conference and Exhibition on Optical Communication*, 2022, pp. Th2C--5.
- [9] S. Deligiannidis *et al.*, “Deep - Learning – based VCSEL transmitter emulator,” 2023.
- [10] K. Sozos, A. Bogris, P. Bienstman, G. Sarantoglou, S. Deligiannidis, and C. Mesaritakis, “High Speed Photonic Neuromorphic Computing Using Recurrent Optical Spectrum Slicing Neural Networks,” *arXiv Prepr. arXiv2203.15807*, 2022.
- [11] A. Bogris, C. Mesaritakis, S. Deligiannidis, and P. Li, “Fabry-Perot lasers as enablers for parallel reservoir computing,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 27, no. 2, pp. 1–7, 2020.
- [12] M. Skontranis, G. Sarantoglou, S. Deligiannidis, A. Bogris, and C. Mesaritakis, “Time-multiplexed spiking convolutional neural network based on vcsels for unsupervised image classification,” *Appl. Sci.*, vol. 11, no. 4, 2021, doi: 10.3390/app11041383.
- [13] K. Sozos, S. Deligiannidis, C. Mesaritakis, and A. Bogris, “Self-Coherent Receiver Based on a Recurrent Optical Spectrum Slicing Neuromorphic Accelerator,” *J. Light. Technol.*, pp. 1–9, 2023, doi: 10.1109/JLT.2023.3235278.
- [14] A. Tsirigotis *et al.*, “Unconventional Integrated Photonic Accelerators for High Throughput Convolutional Neural Networks,” *Intell. Comput.*, pp. 1–33, Jun. 2023, doi: 10.34133/icomputing.0032.
- [15] M. Skontranis, G. Sarantoglou, S. Deligiannidis, A. Bogris, and C. Mesaritakis, “Unsupervised image classification through time-multiplexed photonic multi-layer spiking convolutional neural network,” *arXiv*, no. 1, 2020, doi: 10.1109/ECOC48923.2020.9333320.
- [16] K. Sozos, S. Deligiannidis, C. Mesaritakis, and A. Bogris, “Unconventional Computing based on Four Wave Mixing in Highly Nonlinear Media,” *CLEO-Europe*, 2023.

- [17] I. Tsilikas, S. Deligiannidis, A. Tsirigotis, G. N. Tsigaridas, and A. Bogris, “Neuromorphic Camera Assisted High-Flow Imaging Cytometry for Particle Classification,” 2023.
- [18] I. Tsilikas, A. Tsirigotis, S. Deligiannidis, G. N. Tsigaridas, and A. Bogris, “Time-Stretched Imaging Flow Cytometry and Photonic Neuromorphic Processing for Particle Classification,” 2023.
- [19] A. Bogris, K. Sozos, G. Sarantoglou, S. Deligiannidis, and C. Mesaritakis, “Neuromorphic computing by means of recurrent spectrum slicing for next generation high baud rate transmission systems,” in *2023 IEEE Photonics Society Summer Topicals Meeting Series (SUM)*, 2023, pp. 1–2. doi: 10.1109/SUM57928.2023.10224454.
- [20] A. Fayad, T. Cinkler, J. Rak, and M. Jha, “Design of Cost-Efficient Optical Fronthaul for 5G/6G Networks: An Optimization Perspective,” *Sensors*, vol. 22, no. 23, 2022, doi: 10.3390/s22239394.
- [21] Y. Miyamoto and R. Kawamura, “Space division multiplexing optical transmission technology to support the evolution of high-capacity optical transport networks,” *NTT Tech. Rev.*, vol. 15, no. 6, 2017.
- [22] “www.submarinecablemap.com.” <https://www.submarinecablemap.com/>
- [23] J. J. Maki, “Evolution of Pluggable Optics and What Is Beyond,” in *Optical Fiber Communication Conference (OFC) 2019*, 2019, p. Th3A.2. doi: 10.1364/OFC.2019.Th3A.2.
- [24] D. J. Richardson, J. M. Fini, and L. E. Nelson, “Space-division multiplexing in optical fibres,” *Nat. Photonics*, vol. 7, no. 5, pp. 354–362, 2013.
- [25] L. Rapp and M. Eiselt, “Optical Amplifiers for Wideband Optical Transmission Systems,” in *Optical Fiber Communication Conference (OFC) 2021*, 2021, p. Th4C.1. doi: 10.1364/OFC.2021.Th4C.1.
- [26] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, “Rate Adaptation and Reach Increase by Probabilistically Shaped 64-QAM: An

- Experimental Demonstration,” *J. Light. Technol.*, vol. 34, no. 7, pp. 1599–1609, 2016, [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-34-7-1599>
- [27] S. Zhang *et al.*, “Capacity-Approaching Transmission Over 6375 km Using Hybrid Quasi-Single-Mode Fiber Spans,” *J. Light. Technol.*, vol. 35, no. 3, pp. 481–487, 2017, doi: 10.1109/JLT.2016.2631151.
- [28] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov, “Digital signal processing for fiber nonlinearities,” *Opt. Express*, vol. 25, no. 3, pp. 1916–1936, Feb. 2017, doi: 10.1364/OE.25.001916.
- [29] A. Sobhanan, L. N. Venkatasubramani, R. D. Koilpillai, and D. Venkitesh, “Dispersion and Nonlinearity Distortion Compensation of the QPSK/16QAM Signals Using Optical Phase Conjugation in Nonlinear SOAs,” *IEEE Photonics J.*, vol. 12, no. 1, pp. 1–7, 2020, doi: 10.1109/JPHOT.2020.2964414.
- [30] E. Ip and J. M. Kahn, “Compensation of dispersion and nonlinear impairments using digital backpropagation,” *J. Light. Technol.*, vol. 26, no. 20, pp. 3416–3425, 2008.
- [31] S. A. Derevyanko, J. E. Prilepsy, and S. K. Turitsyn, “Capacity estimates for optical transmission based on the nonlinear Fourier transform,” *Nat. Commun.*, vol. 7, 2016, doi: 10.1038/ncomms12710.
- [32] A. Bakhshali *et al.*, “Frequency-domain Volterra-based equalization structures for efficient mitigation of intrachannel Kerr nonlinearities,” *J. Light. Technol.*, vol. 34, no. 8, pp. 1770–1777, 2015.
- [33] “<https://www.fs.com/de-en/c/200-400-800g-modules-3859>.”
<https://www.fs.com/de-en/c/200-400-800g-modules-3859>
- [34] A. Ferrari *et al.*, “Assessment on the Achievable Throughput of Multi-Band ITU-T G.652.D Fiber Transmission Systems,” *J. Light. Technol.*, vol. 38, no. 16, pp. 4279–4291, 2020, doi: 10.1109/JLT.2020.2989620.
- [35] R. T. Jones *et al.*, “High Speed PAM-8 Optical Interconnects with Digital Equalization based on Neural Network,” *Opt. InfoBase Conf. Pap.*, p. AS1C.1, 2016, doi: 10.1364/acpc.2016.as1c.1.

- [36] D. Wang and M. Zhang, “Artificial Intelligence in Optical Communications: From Machine Learning to Deep Learning,” *Front. Commun. Networks*, vol. 2, no. March, pp. 1–9, Mar. 2021, doi: 10.3389/frcmn.2021.656786.
- [37] P. J. Freire, J. E. Prilepsy, Y. Osadchuk, S. K. Turitsyn, and V. Aref, “Deep Neural Network-aided Soft-Demapping in Coherent Optical Systems: Regression versus Classification,” *IEEE Trans. Commun.*, vol. PP, p. 1, Sep. 2022, doi: 10.1109/TCOMM.2022.3213284.
- [38] S. Zhang *et al.*, “Field and lab experimental demonstration of nonlinear impairment compensation using neural networks,” *Nat. Commun.*, vol. 10, no. 1, 2019, doi: 10.1038/s41467-019-10911-9.
- [39] D. Wang *et al.*, “System impairment compensation in coherent optical communications by using a bio-inspired detector based on artificial neural network and genetic algorithm,” *Opt. Commun.*, vol. 399, no. November 2016, pp. 1–12, 2017, doi: 10.1016/j.optcom.2017.04.050.
- [40] P. J. Freire *et al.*, “Performance versus Complexity Study of Neural Network Equalizers in Coherent Optical Systems,” *J. Light. Technol.*, vol. 39, no. 19, pp. 1–11, Oct. 2021, doi: 10.1109/JLT.2021.3096286.
- [41] P. J. Freire, A. Napoli, B. Spinnler, N. Costa, S. K. Turitsyn, and J. E. Prilepsy, “Neural Networks-Based Equalizers for Coherent Optical Transmission: Caveats and Pitfalls,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 28, no. 4, pp. 1–23, 2022, doi: 10.1109/JSTQE.2022.3174268.
- [42] M. Schaedler, F. Pittala, G. Bocherer, C. Bluemm, M. Kushnerov, and S. Pachnicke, “Recurrent Neural Network Soft-Demapping for Nonlinear ISI in 800Gbit/s DWDM Coherent Optical Transmissions,” *2020 Eur. Conf. Opt. Commun. ECOC 2020*, no. 1, pp. 31–34, 2020, doi: 10.1109/ECOC48923.2020.9333204.
- [43] X. Liu, Y. Wang, X. Wang, H. Xu, C. Li, and X. Xin, “Bi-directional gated recurrent unit neural network based nonlinear equalizer for coherent optical communication system,” *Opt. Express*, vol. 29, no. 4, p. 5923, Feb. 2021, doi: 10.1364/oe.416672.

- [44] S.-C.-K. Kalla, C. Gagné, M. Zeng, and L. A. Rusch, “Recurrent neural networks achieving MLSE performance for optical channel equalization,” *Opt. Express*, vol. 29, no. 9, p. 13033, 2021, doi: 10.1364/oe.423103.
- [45] B. Karanov *et al.*, “End-to-End Deep Learning of Optical Fiber Communications,” *J. Light. Technol.*, vol. 36, no. 20, pp. 4843–4855, Oct. 2018, doi: 10.1109/JLT.2018.2865109.
- [46] G. S. Yadav, C.-Y. Chuang, K.-M. Feng, J. Chen, and Y.-K. Chen, “Computation efficient sparse DNN nonlinear equalization for IM/DD 112 Gbps PAM4 inter-data center optical interconnects,” *Opt. Lett.*, vol. 46, no. 9, p. 1999, 2021, doi: 10.1364/ol.417834.
- [47] Z. Xu, C. Sun, T. Ji, H. Ji, and W. Shieh, “Cascade Recurrent Neural Network Enabled 100-Gb/s PAM4 Short-Reach Optical Link Based on DML,” in *Optical Fiber Communication Conference (OFC) 2020*, 2020, p. W2A.45. doi: 10.1364/OFC.2020.W2A.45.
- [48] X. Lu *et al.*, “Memory-controlled deep LSTM neural network post-equalizer used in high-speed PAM VLC system,” *Opt. Express*, vol. 27, no. 5, pp. 7822–7833, 2019, doi: 10.1364/OE.27.007822.
- [49] C. Wang, S. Fu, Z. Xiao, M. Tang, and D. Liu, “Long Short-Term Memory Neural Network (LSTM-NN) Enabled Accurate Optical Signal-to-Noise Ratio (OSNR) Monitoring,” *J. Light. Technol.*, vol. 37, no. 16, pp. 4140–4146, 2019, [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-37-16-4140>
- [50] H. Sui, H. Zhu, J. Wu, B. Luo, S. Taccheo, and X. Zou, “Modeling pulse propagation in fiber optical parametric amplifier by a long short-term memory network,” *Optik (Stuttg.)*, vol. 260, p. 169125, 2022, doi: <https://doi.org/10.1016/j.ijleo.2022.169125>.
- [51] X. Dai, X. Li, M. Luo, Q. You, and S. Yu, “LSTM networks enabled nonlinear equalization in 50-Gb/s PAM-4 transmission links,” *Appl. Opt.*, vol. 58, no. 22, p. 6079, 2019, doi: 10.1364/ao.58.006079.
- [52] B. Sang *et al.*, “Low Complexity Neural Network Equalization Based on Multi-

- Symbol Output Technique for 200+ Gbps IM/DD Short Reach Optical System,” *J. Light. Technol.*, vol. 40, no. 9, pp. 2890–2900, 2022, [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-40-9-2890>
- [53] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita, “Ultimate low-loss single-mode fibre at 1.55 μm ,” *Electron. Lett.*, vol. 15, no. 4, p. 106, 1979, doi: 10.1049/el:19790077.
- [54] R. J. Mears, L. Reekie, I. M. Jauncey, and D. N. Payne, “Low-noise erbium-doped fibre amplifier operating at 1.54 μm ,” *Electron. Lett.*, vol. 23, no. 19, pp. 1026–1028, 1987.
- [55] H. Kidorf, K. Rottwitt, M. Nissov, M. Ma, and E. Rabarijaona, “Pump interactions in a 100-nm bandwidth Raman amplifier,” *IEEE Photonics Technol. Lett.*, vol. 11, no. 5, pp. 530–532, May 1999, doi: 10.1109/68.759388.
- [56] S. Kobayashi and T. Kimura, “Semiconductor optical amplifiers: Beyond their role as transmitters in fiber-optic systems, laser diodes can be made to amplify light, perhaps promising simple all-optical repeaters,” *IEEE Spectr.*, vol. 21, no. 5, pp. 26–33, 1984, doi: 10.1109/MSPEC.1984.6370264.
- [57] W. T. Silfvast, *Laser fundamentals*. Cambridge university press, 2004.
- [58] T. Komine and M. Nakagawa, “Fundamental analysis for visible-light communication system using LED lights,” *IEEE Trans. Consum. Electron.*, vol. 50, no. 1, pp. 100–107, Feb. 2004, doi: 10.1109/TCE.2004.1277847.
- [59] G. P. Agrawal, “Nonlinear fiber optics,” in *Nonlinear Science at the Dawn of the 21st Century*, Springer, 2000, pp. 195–211.
- [60] M. M.-K. Liu, *Principles and applications of optical communications*. McGraw-Hill Science, Engineering & Mathematics, 1996.
- [61] Itu-t, “ITU-T Rec. G.652 (11/2009) Characteristics of a single-mode optical fibre and cable,” 2009.
- [62] F. Pancaldi, G. Vitetta, R. Kalbasi, N. Al-Dhahir, M. Uysal, and H. Mheidat, “Single-

- carrier frequency domain equalization,” *IEEE Signal Process. Mag.*, vol. 25, no. 5, pp. 37–56, Sep. 2008, doi: 10.1109/MSP.2008.926657.
- [63] J. Zhang, J. Yu, N. Chi, and H.-C. Chien, “Time-domain digital pre-equalization for band-limited signals based on receiver-side adaptive equalizers,” *Opt. Express*, vol. 22, no. 17, p. 20515, Aug. 2014, doi: 10.1364/oe.22.020515.
- [64] C. R. S. Fludger *et al.*, “Coherent equalization and POLMUX-RZ-DQPSK for robust 100-GE transmission,” in *Journal of Lightwave Technology*, Jan. 2008, vol. 26, no. 1, pp. 64–72. doi: 10.1109/JLT.2007.912128.
- [65] S. S. Haykin, *Introduction to adaptive filters*. Macmillan, 1984.
- [66] M. Kuschnerov *et al.*, “DSP for coherent single-carrier receivers,” *J. Light. Technol.*, vol. 27, no. 16, pp. 3614–3622, Aug. 2009, doi: 10.1109/JLT.2009.2024963.
- [67] G. P. Agrawal, *Lightwave technology: telecommunication systems*. John Wiley & Sons, 2005.
- [68] A. Amari, O. A. Dobre, R. Venkatesan, O. S. Sunish Kumar, P. Ciblat, and Y. Jaouen, “A Survey on Fiber Nonlinearity Compensation for 400 Gb/s and beyond Optical Communication Systems,” *IEEE Commun. Surv. Tutorials*, vol. 19, no. 4, pp. 3097–3113, 2017, doi: 10.1109/COMST.2017.2719958.
- [69] R.-J. Essiambre and P. J. Winzer, “Fibre nonlinearities in electronically pre-distorted transmission,” in *2005 31st European Conference on Optical Communication, ECOC 2005*, 2005, vol. 2, pp. 191–192.
- [70] O. V. Sinkin, R. Holzlöhner, J. Zweck, and C. R. Menyuk, “Optimization of the split-step Fourier method in modeling optical-fiber communications systems,” *J. Light. Technol.*, vol. 21, no. 1, p. 61, 2003.
- [71] D. Rafique, M. Mussolin, M. Forzati, J. Mårtensson, M. N. Chughtai, and A. D. Ellis, “Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm,” *Opt. Express*, vol. 19, no. 10, pp. 9453–9460, 2011.
- [72] M. Schetzen, “Theory of pth-order inverses of nonlinear systems,” *IEEE Trans.*

- Circuits Syst.*, vol. 23, no. 5, pp. 285–291, May 1976, doi: 10.1109/TCS.1976.1084219.
- [73] L. Liu *et al.*, “Intrachannel nonlinearity compensation by inverse Volterra series transfer function,” *J. Light. Technol.*, vol. 30, no. 3, pp. 310–316, 2011.
- [74] N. P. Diamantopoulos, H. Nishi, W. Kobayashi, K. Takeda, T. Kakitsuka, and S. Matsuo, “On the Complexity Reduction of the Second-Order Volterra Nonlinear Equalizer for IM/DD Systems,” *J. Light. Technol.*, vol. 37, no. 4, pp. 1214–1224, 2019, doi: 10.1109/JLT.2018.2890118.
- [75] A. Amari, P. Ciblat, and Y. Jaouën, “Fifth-order Volterra series based nonlinear equalizer for long-haul high data rate optical fiber communications,” in *2014 48th Asilomar Conference on Signals, Systems and Computers*, 2014, pp. 1367–1371. doi: 10.1109/ACSSC.2014.7094684.
- [76] I. D. Phillips *et al.*, “Exceeding the Nonlinear-Shannon Limit using Raman Laser Based Amplification and Optical Phase Conjugation,” in *Optical Fiber Communication Conference*, 2014, p. M3C.1. doi: 10.1364/OFC.2014.M3C.1.
- [77] S. L. Jansen, D. van den Borne, P. M. Krummrich, S. Spalter, G.-D. Khoe, and H. de Waardt, “Long-haul DWDM transmission systems employing optical phase conjugation,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 12, no. 4, pp. 505–520, Jul. 2006, doi: 10.1109/JSTQE.2006.876621.
- [78] A. D. Ellis, M. E. McCarthy, M. A. Z. Al-Khateeb, and S. Sygletos, “Capacity limits of systems employing multiple optical phase conjugators,” *Opt. Express*, vol. 23, no. 16, p. 20381, 2015, doi: 10.1364/oe.23.020381.
- [79] X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, “Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit,” *Nat. Photonics*, vol. 7, no. 7, pp. 560–568, 2013, doi: 10.1038/nphoton.2013.109.
- [80] S. T. Le *et al.*, “Demonstration of Phase-Conjugated Subcarrier Coding for Fiber Nonlinearity Compensation in CO-OFDM Transmission,” *J. Light. Technol.*, vol. 33, no. 11, pp. 2206–2212, 2015, [Online]. Available:

<http://opg.optica.org/jlt/abstract.cfm?URI=jlt-33-11-2206>

- [81] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, “Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate,” *J. Light. Technol.*, vol. 29, no. 17, pp. 2570–2576, 2011, doi: 10.1109/JLT.2011.2160933.
- [82] X. Liang and S. Kumar, “Multi-stage perturbation theory for compensating intrachannel nonlinear impairments in fiber-optic links,” *Opt. Express*, vol. 22, no. 24, p. 29733, Dec. 2014, doi: 10.1364/OE.22.029733.
- [83] Y. Gao *et al.*, “Reducing the complexity of perturbation based nonlinearity pre-compensation using symmetric EDC and pulse shaping,” *Opt. Express*, vol. 22, no. 2, p. 1209, 2014, doi: 10.1364/oe.22.001209.
- [84] S. K. Turitsyn *et al.*, “Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives,” *Optica*, vol. 4, no. 3, p. 307, Mar. 2017, doi: 10.1364/OPTICA.4.000307.
- [85] J. C. Cartledge and A. S. Karar, “100 Gb/s intensity modulation and direct detection,” *J. Light. Technol.*, vol. 32, no. 16, pp. 2809–2814, Aug. 2014, doi: 10.1109/JLT.2014.2314611.
- [86] J. Wei, Q. Cheng, R. V. Penty, I. H. White, and D. G. Cunningham, “400 Gigabit Ethernet using advanced modulation formats: Performance, complexity, and power dissipation,” *IEEE Commun. Mag.*, vol. 53, no. 2, pp. 182–189, Feb. 2015, doi: 10.1109/MCOM.2015.7045407.
- [87] A. P. T. Lau *et al.*, “Experimental study of PAM-4, CAP-16, and DMT for 100 Gb/s Short Reach Optical Transmission Systems,” *Opt. Express*, Vol. 23, Issue 2, pp. 1176–1189, vol. 23, no. 2, pp. 1176–1189, Jan. 2015, doi: 10.1364/OE.23.001176.
- [88] D. V. Plant, H. Xin, K. Zhang, Q. Zhuge, and W. Hu, “Performance comparison of DML, EML and MZM in dispersion-unmanaged short reach transmissions with digital signal processing,” *Opt. Express*, Vol. 26, Issue 26, pp. 34288–34304, vol. 26, no. 26, pp. 34288–34304, Dec. 2018, doi: 10.1364/OE.26.034288.

- [89] K. Zhong *et al.*, “Amplifier-Less Transmission of 56Gbit/s PAM4 over 60km Using 25Gbps EML and APD,” *Opt. Fiber Commun. Conf. (2017), Pap. Tu2D.1*, vol. Part F40-OFC 2017, p. Tu2D.1, Mar. 2017, doi: 10.1364/OFC.2017.TU2D.1.
- [90] H. Zhou *et al.*, “Recent advances in equalization technologies for short-reach optical links based on PAM4 modulation: A review,” *Appl. Sci.*, vol. 9, no. 11, p. 2342, 2019, doi: 10.3390/app9112342.
- [91] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, “Coherent detection in optical fiber systems: erratum,” *Opt. Express*, vol. 16, no. 26, p. 21943, Dec. 2008, doi: 10.1364/oe.16.021943.
- [92] B. Chen, C. Okonkwo, H. Hafermann, and A. Alvarado, “Increasing Achievable Information Rates via Geometric Shaping,” no. 1, pp. 10–12.
- [93] G. Böcherer, P. Schulte, and F. Steiner, “Probabilistic Shaping and Forward Error Correction for Fiber-Optic Communication Systems,” *J. Light. Technol.*, vol. 37, no. 2, pp. 230–244, 2019, doi: 10.1109/JLT.2019.2895770.
- [94] Z. Qu and I. B. Djordjevic, “On the Probabilistic Shaping and Geometric Shaping in Optical Communication Systems,” *IEEE Access*, vol. 7, pp. 21454–21464, 2019, doi: 10.1109/ACCESS.2019.2897381.
- [95] C. E. Shannon, “A Mathematical Theory of Communication,” *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, 1948, doi: 10.1002/j.1538-7305.1948.tb00917.x.
- [96] M. S. Faruk, Y. Mori, C. Zhang, and K. Kikuchi, “Multi-impairments monitoring from the equalizer in a digital coherent optical receiver,” *Eur. Conf. Opt. Commun. ECOC*, vol. 1–2, no. 26, pp. 26929–26936, 2010, doi: 10.1109/ECOC.2010.5621452.
- [97] R. Raheli and G. Picchi, “Synchronous and fractionally-spaced blind equalization in dually-polarized digital radio links,” in *ICC 91 International Conference on Communications Conference Record*, 2002, pp. 156–161. doi: 10.1109/ICC.1991.162352.
- [98] A. Viterbi, A. V.-I. T. on I. theory, and undefined 1983, “Nonlinear estimation of

- PSK-modulated carrier phase with application to burst digital transmission,” *ieeexplore.ieee.org*, Accessed: Oct. 19, 2022. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/1056713/>
- [99] T. M. Mitchell, *Machine learning*, vol. 1, no. 9.
- [100] S. B. Kotsiantis, I. Zaharakis, P. Pintelas, and others, “Supervised machine learning: A review of classification techniques,” *Emerg. Artif. Intell. Appl. Comput. Eng.*, vol. 160, no. 1, pp. 3–24, 2007.
- [101] F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain.,” *Psychol. Rev.*, vol. 65, no. 6, p. 386, 1958.
- [102] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
- [103] N. Qian, “On the momentum term in gradient descent learning algorithms,” *Neural networks*, vol. 12, no. 1, pp. 145–151, 1999.
- [104] R. Hecht-Nielsen, “Theory of the backpropagation neural network,” in *Neural networks for perception*, Elsevier, 1992, pp. 65–93.
- [105] Y. LeCun, C. Cortes, and C. J. Burges, “Mnist handwritten digit database. ATT Labs.” 2010.
- [106] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *arXiv Prepr. arXiv1412.6980*, Dec. 2014, [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [107] G. Hinton, N. Srivastava, and K. Swersky, “Neural networks for machine learning lecture 6a overview of mini-batch gradient descent,” *Cited on*, vol. 14, no. 8, p. 2, 2012.
- [108] J. Duchi, E. Hazan, and Y. Singer, “Adaptive subgradient methods for online learning and stochastic optimization.,” *J. Mach. Learn. Res.*, vol. 12, no. 7, 2011.
- [109] T. O’Malley *et al.*, “Keras Tuner. 2019,” *Available online github. com/keras-team/kerastuner (accessed 2 April 2022)*.
- [110] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553,

- pp. 436–444, 2015, doi: 10.1038/nature14539.
- [111] K. Kawakami, “Supervised sequence labelling with recurrent neural networks,” Technical University of Munich, 2008.
- [112] R. Pascanu, T. Mikolov, and Y. Bengio, “On the difficulty of training recurrent neural networks,” in *Proceedings of the 30th International Conference on Machine Learning*, 2013, vol. 28, no. 3, pp. 1310–1318. [Online]. Available: <https://proceedings.mlr.press/v28/pascanu13.html>
- [113] S. Hochreiter, J. Urgan Schmidhuber, J. Schmidhuber, H. Sepp, and J. Schmidhuber, “LONG SHORT-TERM MEMORY,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [114] A. Graves, “Generating sequences with recurrent neural networks,” *arXiv Prepr. arXiv1308.0850*, 2013.
- [115] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, “On the properties of neural machine translation: Encoder–decoder approaches,” *Proc. SSST 2014 - 8th Work. Syntax. Semant. Struct. Stat. Transl.*, pp. 103–111, 2014, doi: 10.3115/v1/w14-4012.
- [116] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling,” pp. 1–9, 2014, [Online]. Available: <http://arxiv.org/abs/1412.3555>
- [117] Y. Bengio, P. Simard, P. F.-I. transactions on neural, and undefined 1994, “Learning long-term dependencies with gradient descent is difficult,” *ieeexplore.ieee.org*, Accessed: Dec. 08, 2022. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/279181/>
- [118] D. Bahdanau, K. Cho, Y. B. preprint arXiv:1409.0473, and undefined 2014, “Neural machine translation by jointly learning to align and translate,” *arxiv.org*, Accessed: Dec. 08, 2022. [Online]. Available: <https://arxiv.org/abs/1409.0473>
- [119] R. Pascanu, C. Gulcehre, K. Cho, and Y. Bengio, “How to Construct Deep Recurrent Neural Networks,” *2nd Int. Conf. Learn. Represent. ICLR 2014 - Conf.*

- Track Proc.*, pp. 1–13, Dec. 2013, [Online]. Available: <http://arxiv.org/abs/1312.6026>
- [120] J. Donahue *et al.*, “Long-term recurrent convolutional networks for visual recognition and description,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, pp. 2625–2634, 2015, doi: 10.1109/CVPR.2015.7298878.
- [121] K. Cho *et al.*, “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” *EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 1724–1734, 2014, doi: 10.3115/v1/d14-1179.
- [122] M. Schuster and K. K. Paliwal, “Bidirectional recurrent neural networks,” *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, 1997.
- [123] A. Graves and J. Schmidhuber, “Framewise phoneme classification with bidirectional LSTM and other neural network architectures,” *Neural Networks*, vol. 18, no. 5–6, pp. 602–610, 2005, doi: 10.1016/j.neunet.2005.06.042.
- [124] E. Giacomidis, Y. Lin, J. Wei, I. Aldaya, A. Tsokanos, and L. Barry, “Harnessing machine learning for fiber-induced nonlinearity mitigation in long-haul coherent optical OFDM,” *Futur. Internet*, vol. 11, no. 1, p. 2, Dec. 2018, doi: 10.3390/fi11010002.
- [125] D. Wang *et al.*, “Intelligent constellation diagram analyzer using convolutional neural network-based deep learning,” *Opt. Express*, vol. 25, no. 15, p. 17150, Jul. 2017, doi: 10.1364/OE.25.017150.
- [126] C. Hager, H. D. Pfister, and H. Christian, “Nonlinear interference mitigation via deep neural networks,” *2018 Opt. Fiber Commun. Conf. Expo. OFC 2018 - Proc.*, no. 2, pp. 1–3, 2018.
- [127] P. Li, L. Yi, L. Xue, and W. Hu, “56 Gbps IM/DD PON based on 10G-class optical devices with 29 dB loss budget enabled by machine learning,” *2018 Opt. Fiber Commun. Conf. Expo. OFC 2018 - Proc.*, vol. 1, pp. 1–3, 2018.
- [128] C. Ye, D. Zhang, X. Hu, X. Huang, H. Feng, and K. Zhang, “Recurrent Neural

- Network (RNN) Based End-to-End Nonlinear Management for Symmetrical 50Gbps NRZ PON with 29dB+ Loss Budget,” *Eur. Conf. Opt. Commun. ECOC*, vol. 2018-Septe, pp. 29–31, 2018, doi: 10.1109/ECOC.2018.8535428.
- [129] C. Fougstedt, C. Häger, L. Svensson, H. D. Pfister, and P. Larsson-Edefors, “ASIC implementation of time-domain digital backpropagation with deep-learned chromatic dispersion filters,” in *2018 European Conference on Optical Communication (ECOC)*, Jun. 2018, vol. 2018-Septe, no. 1, pp. 1–3. doi: 10.1109/ECOC.2018.8535430.
- [130] F. Chollet and others, “Keras.” GitHub, 2015. [Online]. Available: <http://github.com/fchollet/keras>
- [131] M. Abadi *et al.*, “TensorFlow: A system for large-scale machine learning,” in *OSDI*, May 2016, vol. 101, no. C, pp. 582–598. doi: 10.1016/0076-6879(83)01039-3.
- [132] T. A. Eriksson, H. Bülow, A. Leven, H. Bulow, and A. Leven, “Applying Neural Networks in Optical Communication Systems: Possible Pitfalls,” *IEEE Photonics Technol. Lett.*, vol. 29, no. 23, pp. 2091–2094, 2017, doi: 10.1109/LPT.2017.2755663.
- [133] M. Secondini and E. Forestieri, “On XPM Mitigation in WDM Fiber-Optic Systems,” *IEEE Photonics Technol. Lett.*, vol. 26, no. 22, pp. 2252–2255, 2014, doi: 10.1109/LPT.2014.2353217.
- [134] A. D. E. Llis, M. E. M. C. C. Arthy, M. A. Z. A. L. K. Hateeb, M. S. Orokina, and N. J. D. Oran, “Performance limits in optical communications due to fiber nonlinearity,” vol. 9, no. 3, pp. 429–503, 2017, doi: <https://doi.org/10.1364/AOP.9.000429>.
- [135] B. Spinnler, “Equalizer design and complexity for digital coherent receivers,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no. 5, pp. 1180–1192, 2010, doi: 10.1109/JSTQE.2009.2035931.
- [136] A. Rezania, J. C. Cartledge, A. Bakhshali, and W. Y. Chan, “Compensation Schemes for Transmitter-and Receiver-Based Pattern-Dependent Distortion,” *IEEE Photonics Technol. Lett.*, vol. 28, no. 22, pp. 2641–2644, 2016, doi:

10.1109/LPT.2016.2613401.

- [137] J. Verbist *et al.*, “A 40-GBd QPSK/16-QAM Integrated Silicon Coherent Receiver,” *IEEE Photonics Technol. Lett.*, vol. 28, no. 19, pp. 2070–2073, 2016, doi: 10.1109/LPT.2016.2582799.
- [138] M. Schädler *et al.*, “Recurrent Neural Network Soft-Demapping for Nonlinear ISI in 800Gbit/s DWDM Coherent Optical Transmissions,” *J. Light. Technol.*, vol. 39, no. 16, pp. 5278–5286, 2021, doi: 10.1109/JLT.2021.3102064.
- [139] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, “A Novel Connectionist System for Unconstrained Handwriting Recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 5, pp. 855–868, 2009, doi: 10.1109/TPAMI.2008.137.
- [140] M. H. Zhu and S. Gupta, “To prune, or not to prune: exploring the efficacy of pruning for model compression,” *arXiv*, Oct. 2017, [Online]. Available: <http://arxiv.org/abs/1710.01878>
- [141] E. Giacomidis, J. Wei, I. Aldaya, and L. P. Barry, “Exceeding the Nonlinear Shannon-Limit in Coherent Optical Communications by MIMO Machine Learning,” no. 1, pp. 1–19, Feb. 2018, Accessed: May 03, 2023. [Online]. Available: <http://arxiv.org/abs/1802.09120>
- [142] O. Sidelnikov, A. Redyuk, S. Sygletos, M. Fedoruk, and S. Turitsyn, “Advanced Convolutional Neural Networks for Nonlinearity Mitigation in Long-Haul WDM Transmission Systems,” *J. Light. Technol.*, vol. 39, no. 8, pp. 2397–2406, 2021, doi: 10.1109/JLT.2021.3051609.
- [143] P. J. Freire *et al.*, “Reducing Computational Complexity of Neural Networks in Optical Channel Equalization: From Concepts to Implementation,” *J. Light. Technol.*, pp. 1–26, Aug. 2023, doi: 10.1109/JLT.2023.3234327.
- [144] S. Srivallapanonndh *et al.*, “Knowledge Distillation Applied to Optical Channel Equalization: Solving the Parallelization Problem of Recurrent Connection,” *2023 Opt. Fiber Commun. Conf. Exhib. OFC 2023 - Proc.*, vol. 1, pp. 7–9, 2023, doi:

10.23919/OFC49934.2023.10116725.

- [145] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Light. Technol.*, vol. 15, no. 9, pp. 1735–1746, 1997.
- [146] R. Maher *et al.*, "Real-Time 100.4 GBd PCS-64QAM Transmission of a 1.6 Tb/s Super-Channel Over 1600 km of G.654.E Fiber," in *Optical Fiber Communication Conference (OFC) 2021*, 2021, p. Tu6D.2. doi: 10.1364/OFC.2021.Tu6D.2.
- [147] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249–256.
- [148] R. S. Lu\`is, B. J. Puttnam, G. Rademacher, Y. Awaji, H. Furukawa, and N. Wada, "Digital back propagation in long-haul, MIMO-supported, multicore fiber transmission," *IEEE Photonics Technol. Lett.*, vol. 32, no. 12, pp. 730–732, 2020.
- [149] G. Liga, T. Xu, A. Alvarado, R. I. Killey, and P. Bayvel, "On the performance of multichannel digital backpropagation in high-capacity long-haul optical transmission," *Opt. Express*, vol. 22, no. 24, pp. 30053–30062, 2014.
- [150] P. J. Freire *et al.*, "Towards FPGA Implementation of Neural Network-Based Nonlinearity Mitigation Equalizers in Coherent Optical Transmission Systems," no. 1, pp. 4–7, 2022, [Online]. Available: <http://arxiv.org/abs/2206.12180>
- [151] N. Sambo *et al.*, "Provisioning in Multi-Band Optical Networks," *J. Light. Technol.*, vol. 38, no. 9, pp. 2598–2605, 2020, doi: 10.1109/JLT.2020.2983227.
- [152] Y. Hong *et al.*, "First Investigation on Double- and Single-sideband Formats in BDFFA-enabled O-band Transmission," in *Conference on Lasers and Electro-Optics*, 2020, p. STh3L.3. doi: 10.1364/CLEO_SI.2020.STh3L.3.
- [153] N. Taengnoi *et al.*, "WDM transmission with in-line amplification at 1.3 μm using a Bi-doped fiber amplifier," *J. Light. Technol.*, vol. 37, no. 8, pp. 1826–1830, 2019.
- [154] M. Tao *et al.*, "Improved Dispersion Tolerance for 50G-PON Downstream

- Transmission via Receiver-Side Equalization,” in *Optical Fiber Communication Conference (OFC) 2019*, 2019, p. M2B.3. doi: 10.1364/OFC.2019.M2B.3.
- [155] L. Xue, L. Yi, H. Ji, P. Li, and W. Hu, “Symmetric 100-Gb/s TWDM-PON in O-band based on 10G-class optical devices enabled by dispersion-supported equalization,” *J. Light. Technol.*, vol. 36, no. 2, pp. 580–586, 2017.
- [156] C. Ye, D. Zhang, X. Huang, H. Feng, and K. Zhang, “Demonstration of 50Gbps IM/DD PAM4 PON over 10GHz class optics using neural network based nonlinear equalization,” in *2017 European Conference on Optical Communication (ECOC)*, 2017, pp. 1–3.
- [157] R. Rath, D. Clausen, S. Ohlendorf, S. Pachnicke, and W. Rosenkranz, “Tomlinson-Harashima precoding for dispersion uncompensated PAM-4 transmission with direct-detection,” *J. Light. Technol.*, vol. 35, no. 18, pp. 3909–3917, 2017.
- [158] M. Dalla Santa, C. Antony, G. Talli, and P. D. Townsend, “25Gb/s PAM4 adaptive receiver equalisation requirements for burst-mode transmission systems,” in *ECOC 2016; 42nd European Conference on Optical Communication*, 2016, pp. 1–3.
- [159] N. K. Thipparapu, A. A. Umnikov, P. Barua, and J. K. Sahu, “Bi-doped fiber amplifier with a flat gain of 25 dB operating in the wavelength band 1320–1360 nm,” *Opt. Lett.*, vol. 41, no. 7, pp. 1518–1521, 2016.
- [160] Y. Hong *et al.*, “Numerical and experimental study on the impact of chromatic dispersion on O-band direct-detection transmission,” *Appl. Opt.*, vol. 60, no. 15, pp. 4383–4390, 2021.
- [161] G. Lu, H.-B. Zhang, S. Shinada, J. Hong, Y. Cheng, and S. Yokoyama, “Power-Efficient O-Band 40 Gbit/s PAM4 Transmitter Based on Dual-Drive Cascaded Carrier-Depletion and Carrier-Injection Silicon Mach-Zehnder Modulator With Binary Driving Electronics at CMOS Voltages,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 27, no. 3, pp. 1–8, 2021, doi: 10.1109/JSTQE.2021.3056721.
- [162] A. M. Abdelsalam, J. M. P. Langlois, and F. Cheriet, “A configurable FPGA implementation of the tanh function using DCT interpolation,” in *2017 IEEE 25th*

- annual international symposium on field-programmable custom computing machines (FCCM)*, 2017, pp. 168–171.
- [163] B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, “Multimode optical fiber transmission with a deep learning network,” *Light Sci. & Appl.*, vol. 7, no. 1, p. 69, 2018.
- [164] N. Gautam, A. Choudhary, and B. Lall, “Comparative study of neural network architectures for modelling nonlinear optical pulse propagation,” *Opt. Fiber Technol.*, vol. 64, p. 102540, 2021.
- [165] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, “Deep learning for the design of photonic structures,” *Nat. Photonics*, vol. 15, no. 2, pp. 77–90, 2021.
- [166] H. Ming, X. Chen, X. Fang, L. Zhang, C. Li, and F. Zhang, “Ultralow complexity long short-term memory network for fiber nonlinearity mitigation in coherent optical communication systems,” *J. Light. Technol.*, vol. 40, no. 8, pp. 2427–2434, Aug. 2021, doi: 10.1109/JLT.2022.3141404.
- [167] D. A. Ron, P. J. Freire, J. E. Prilepsky, M. Kamalian-Kopae, A. Napoli, and S. K. Turitsyn, “Experimental implementation of a neural network optical channel equalizer in restricted hardware using pruning and quantization,” *Sci. Rep.*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-12563-0.
- [168] A. Shahkarami, M. Yousefi, and Y. Jaouen, “Complexity Reduction over Bi-RNN-Based Nonlinearity Mitigation in Dual-Pol Fiber-Optic Communications via a CRNN-Based Approach,” pp. 1–24, Jul. 2022, [Online]. Available: <http://arxiv.org/abs/2207.12154>
- [169] T. O’shea and J. Hoydis, “An introduction to deep learning for the physical layer,” *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, 2017.
- [170] R. T. Jones, T. A. Eriksson, M. P. Yankov, and D. Zibar, “Deep Learning of Geometric Constellation Shaping Including Fiber Nonlinearities,” *Eur. Conf. Opt. Commun. ECOC*, vol. 2018-Septe, no. 1, pp. 1–3, 2018, doi:

10.1109/ECOC.2018.8535453.

- [171] S. Li, N. Garcia, H. Wymeersch, C. Häger, N. Garcia, and H. Wymeersch, “Achievable Information Rates for Nonlinear Fiber Communication via End-to-end Autoencoder Learning,” *Eur. Conf. Opt. Commun. ECOC*, vol. 2018-Septe, no. 1, pp. 0–2, 2018, doi: 10.1109/ECOC.2018.8535456.
- [172] K. Peng, L. Xi, X. Li, and X. Zhang, “Complex Long Short-Term Memory Neural Networks for Fiber Nonlinearity Equalization in Long-Haul Transmission Systems,” *ICOON 2022 - 20th Int. Conf. Opt. Commun. Networks*, vol. 1, pp. 6–8, 2022, doi: 10.1109/ICOON55511.2022.9901114.
- [173] F. Da Ros, S. M. Ranzini, Y. Osadchuk, A. Cem, B. J. G. Castro, and D. Zibar, “Reservoir-computing and neural-network-based equalization for short reach communication,” in *Optica Advanced Photonics Congress 2022*, 2022, p. SpTu1J.1. doi: 10.1364/SPPCOM.2022.SpTu1J.1.
- [174] M. Lukoševičius and H. Jaeger, “Reservoir computing approaches to recurrent neural network training,” *Comput. Sci. Rev.*, vol. 3, no. 3, pp. 127–149, Aug. 2009, doi: 10.1016/j.cosrev.2009.03.005.
- [175] A. Vaswani *et al.*, “Attention is all you need,” *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.
- [176] A. Shahkarami, M. I. Yousefi, and Y. Jaouen, “Attention-Based Neural Network Equalization in Fiber-Optic Communications,” in *Asia Communications and Photonics Conference 2021*, 2021, no. 1, p. M5H.3. doi: 10.1364/ACPC.2021.M5H.3.
- [177] “ChatGPT.” <https://openai.com/blog/chatgpt-plus>
- [178] A. R. Brodtkorb, T. R. Hagen, and M. L. Sætra, “Graphics processing unit (GPU) programming strategies and trends in GPU computing,” *J. Parallel Distrib. Comput.*, vol. 73, no. 1, pp. 4–13, 2013, doi: 10.1016/j.jpdc.2012.04.003.
- [179] W. Aspray, *John von Neumann and the origins of modern computing*. Mit Press, 1990.

- [180] K. Asanovic *et al.*, “The Landscape of Parallel Computing Research: A View from Berkeley,” no. UCB/EECS-2006-183, 2006.
- [181] G. M. Amdahl, “Validity of the single processor approach to achieving large scale computing capabilities,” in *Proceedings of the April 18-20, 1967, spring joint computer conference*, 1967, pp. 483–485.
- [182] “CUDA.” <https://developer.nvidia.com/cuda-toolkit> (accessed Mar. 03, 2022).
- [183] “OpenCL.” <https://www.khronos.org/opencv/> (accessed Mar. 03, 2022).

Ακρωνύμια

Συντομογραφία	Αγγλικός Όρος	Ελληνικός Όρος
ADC	Analog-to-digital converter	Μετατροπέας αναλόγικού σε ψηφιακό
ANN	Artificial Neural Network	Τεχνητό Νευρωνικό Δίκτυο
APD	Avalanche Photodiode	Φωτοδίοδος Χιονοστιβάδας
ASE	Amplified Spontaneous Emission	Θόρυβος αυθόρμητης εκπομπής
ASIC	Application-specific integrated circuit	Ολοκληρωμένα κύκλωματα συγκεκριμένης εφαρμογής
AWG	Arbitrary Waveform Generator	Γεννήτρια τυχαίων κυματομορφών
BER	Bit Error Rate	Πιθανότητα σφάλματος
CNN	Convolutional Neural Networks	Συνελκτικά Νευρωνικά Δίκτυα
CR	Coherent reciever	Σύμφωνος δέκτης
CWL	Continuous Wave Laser	Laser συνεχούς κύματος
DAC	Digital-to-analog converter	Μετατροπέας ψηφιακού σε αναλογικό
DBP	Digital back-propagation	Ψηφιακή αντίστροφη διάδοση
DBR	Distributed Bragg Reflector	Laser καταμεμημένου ανακλαστήρα Bragg
DEMUX	De-multiplexer	Αποπολυπλέκτης
DFB	Distributed-feedback laser	laser καταμεμημένης ανάδρασης
DFT	Discrete Fourier Transform	Διακριτός μετασχηματισμός Fourier
DML	Directly Modulated Lasers	laser άμεσης διαμόρφωσης
DNN	Deep Neural Network	Βαθύ Νευρωνικό Δίκτυο
DSP	Digital signal processing	Ψηφιακή επεξεργασία σήματος
DWDM	Dense Wavelength Division Multiplexing	Πυκνή πολυπλεξία διαίρεσης μήκους κύματος
EAM	Electro-absorption modulator	Διαμορφωτής ηλεκτροαπορρόφησης
EDFA	Erbium Doped Fiber Amplifier	Ενισχυτής με ίνα προσμίξεων Ερβίου
EDFA	Bismuth Doped Fiber Amplifier	Ενισχυτής με ίνα προσμίξεων Βισμούθιου
EML	Electro-absorption modulator - Laser	Διαμορφωμένο laser ηλεκτροαπορρόφησης
FCL	Fully Connected Layer	Πλήρως Συνδεδεμένα Στρώμα
FDE	Frequency-Domain Equalization	Εξισορρόπηση στο πεδίο των συχνοτήτων
FFT	Fast Fourier transform	Γρήγορος μετασχηματισμός Fourier
FIR	Finite impulse response	Πεπερασμένης παλμικής απόκρισης
FNN	Feedforward Neural Network	Προσωτροφοδοτούμενο Νευρωνικό Δίκτυο

FWM	Four wave mixing	Μίξη τεσσάρων κυμάτων
IFFT	Inverse Fast Fourier transform	Αντίστροφος γρήγορος μετασχηματισμός Fourier
IM	Intensity Modulation	Διαμόρφωση έντασης
INFT	Inverse Nonlinear Fourier Transform	Αντίστροφος μη γραμμικός μετασχηματισμός Fourier
IQ	In-phase and quadrature modulator	Διαμορφωτής σε φάση και τετραγωνισμού
ISI	Inter-Symbol Interference	Αλληλοπαρεμβολή συμβόλων
IVSFT	Inverse Volterra series transfer function	Αντίστροφη Συνάρτηση μεταφοράς των σειρών Volterra
MLP	Muplti Perceptron	Πολυεπίπεδο Perceptron
MMF	Multi Mode Fiber	Πολύτροπη ίνα
MUX	Multiplexer	Πολυπλέκτης
MZM	Mach-Zehnder Modulator	Διαμορφωτής Mach-Zehnder
NFT	Nonlinear Fourier Transform	Μη γραμμικός μετασχηματισμός Fourier
NLC	Nonlinearity compensation	Αντιστάθμιση μη γραμμικότητας
NLSE	Non Linear Schrödinger Equation	Μη γραμμική εξίσωση του Schrödinger
NRZ-OOK	Non-return-to-zero on-off-keying	Διαμόρφωση έντασης με μη επιστροφή στο μηδέν
OBF	Optical Bandpass Filter	Ζωνοπερατό οπτικό φίλτρο
OBPF	Optical Bandpass Filter	Οπτικό φίλτρο διέλευσης ζώνης
OFDM	Orthogonal Frequency Division Multiplexing	Πολυπλεξία ορθογώνιας διαίρεσης συχνότητας
OOK	On-Off Keying	Διαμόρφωση κατά πλάτος
OPC	Optical Phase Conjugation	Οπτική σύζευξη φάσης
OSA	Optical Spectrum Analyzer	Αναλυτής Οπτικού Φάσματος
OSNR	Optical Signal-to-Noise Ratio	Λόγος οπτικού σήματος προς θόρυβο
PAM	Pulse-Amplitude Modulation	Διαμόρφωση πλάτους παλμού
PCTW	Phase conjugated twin waves	Δίδυμα κύματα συζευγμένα στη φάση
PD	Photodiode	Φωτοδίοδος
PDFA	Praseodymium Doped Fiber Amplifiers	Ενισχυτής με ίνα προσμίξεων Πρασεοδυμίου
PRBS	Pseudo-random binary sequences	Ψευδοτυχαίες δυαδικές ακολουθίες
RNN	Recurrent Neural Network	Επαναλαμβανόμενα Νευρωνικά Δίκτυα
SMF	Single Mode Fiber	Μονότροπη ίνα
SOA	Semiconductor Optical Amplifier	Ημιαγωγικός Οπτικός Ενισχυτής
SOP	State of Polarization	Κατάσταση Πόλωσης
SPM	Self Phase Modulation	Αυτό-διαμόρφωση φάσης
TIA	transimpedance amplifier	Ενισχυτής σύνθετης αντίστασης

VCSEL	Vertical-cavity surface-emitting	Laser επιφανειακής εκπομπής κατακόρυφης κοιλότητας
VNLE	Volterra series based nonlinear equalizer	Μη γραμμικός ισοσταθμιστής βασισμένος στις σειρές Volterra
VOA	Variable Optical Attenuator	Μεταβλητός οπτικός εξασθενητής
VSFT	Volterra series transfer function	Συνάρτηση μεταφοράς των σειρών Volterra
WDM	Wavelength Division Multiplexing	Διαίρεση Μήκους Κύματος
XPM	Cross Phase Modulation	Ετερο-διαμόρφωση φάσης