

Report for Project.

➤ ΘΕΜΑ 1ο :

Ερώτημα 1.1: Παρακάτω εμφανίζονται τα γραφήματα για τα πρώτα 10000 δείγματα των αρχείων Data_Test_# για κάθε ένα από τα αρχεία που δίνονται , κατά αύξοντα αριθμό. Για την υλοποίηση των παρακάτω γραφημάτων χρησιμοποιήθηκε η εντολή plot για τα πρώτα 10000 στοιχεία του πίνακα z στον οποίο είναι αποθηκευμένες οι τιμές του σήματος.

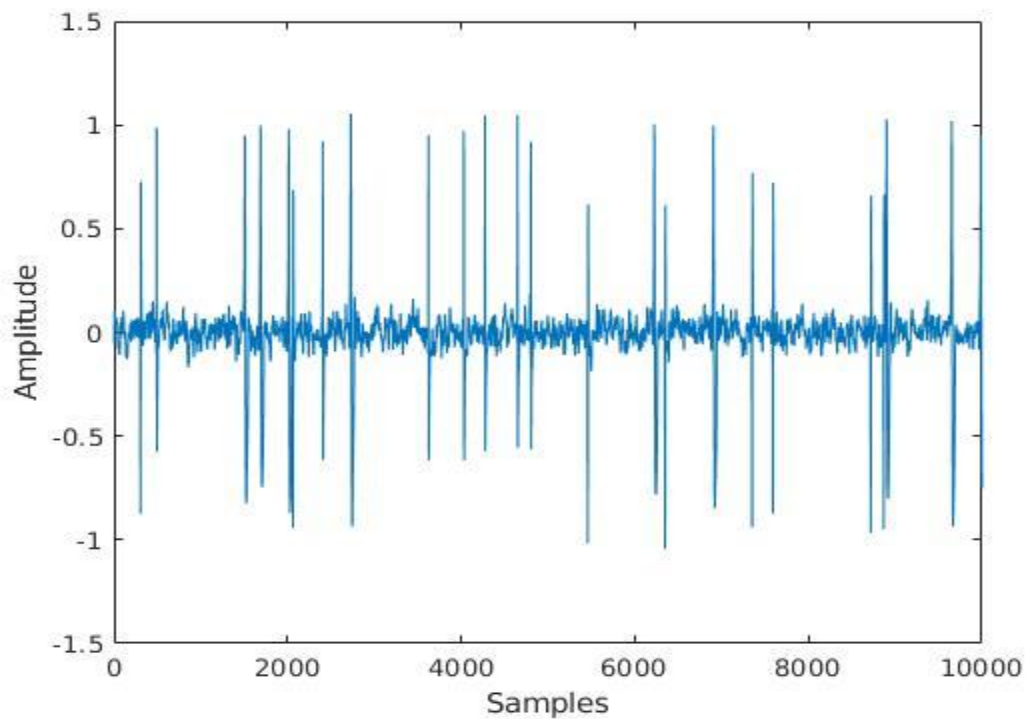


Figure 1.1.1 : Data_Test_1

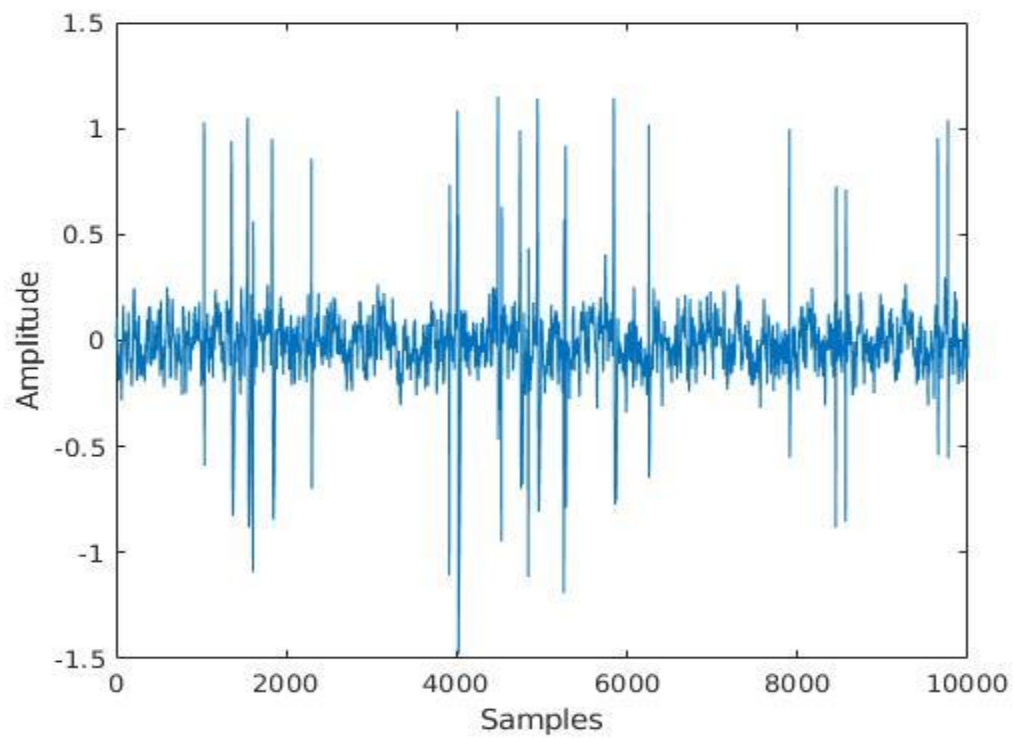


Figure 1.1.2 : Data_Test_2

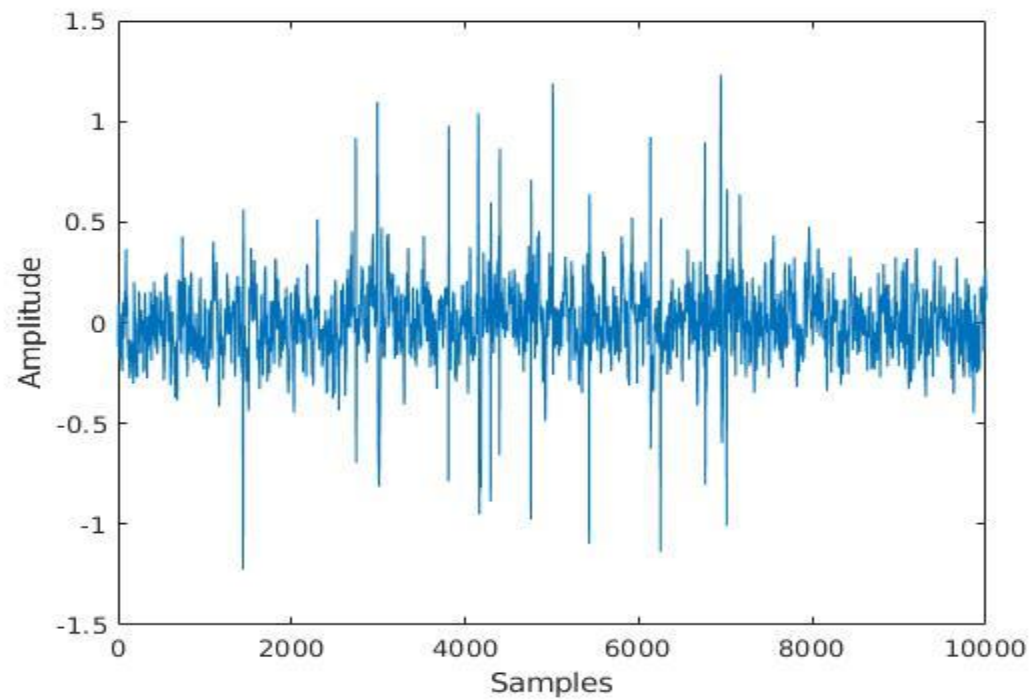


Figure 1.1.3 : Data_Test_3

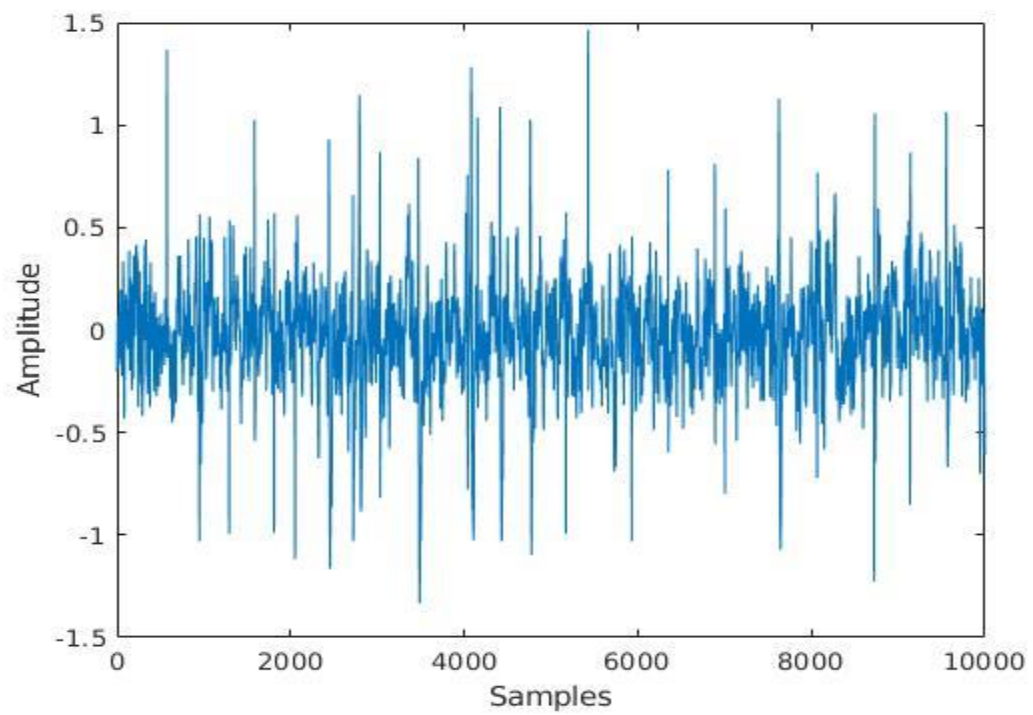


Figure 1.1.4: Data_Test_4

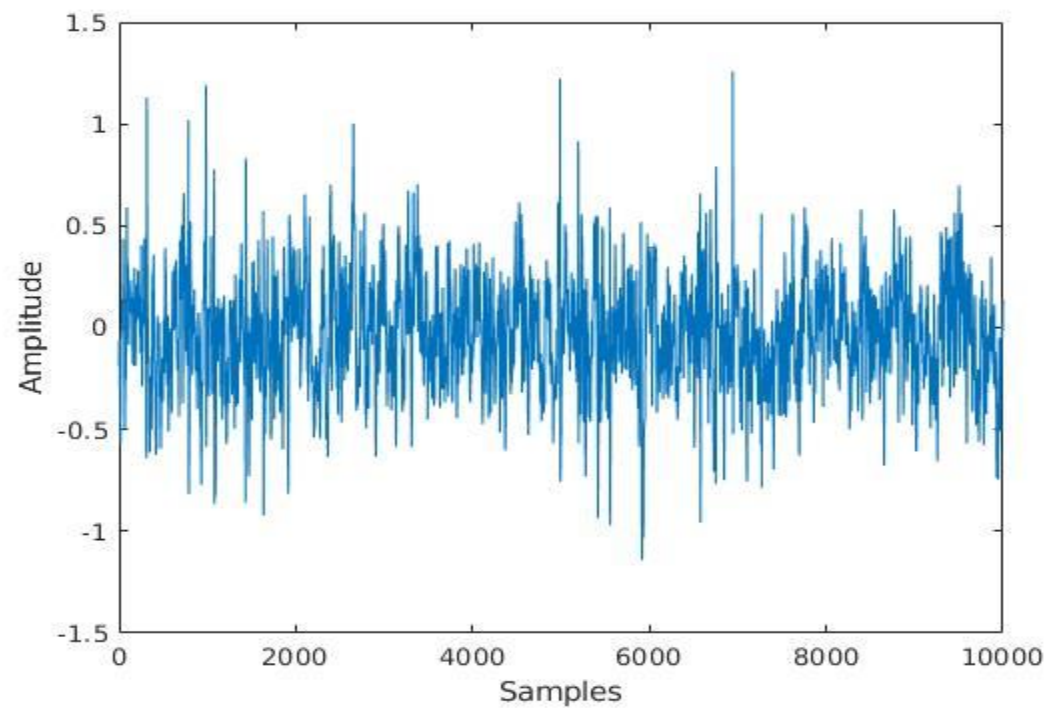


Figure 1.1.5 : Data_Test_5

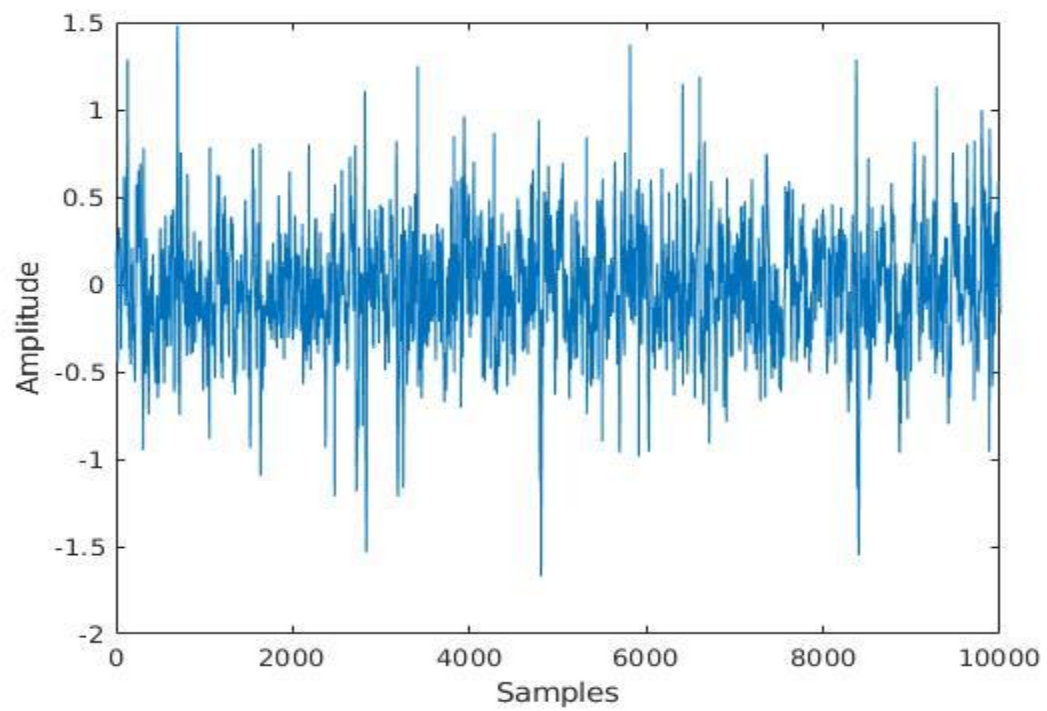


Figure 1.1.6 : Data_Test_6

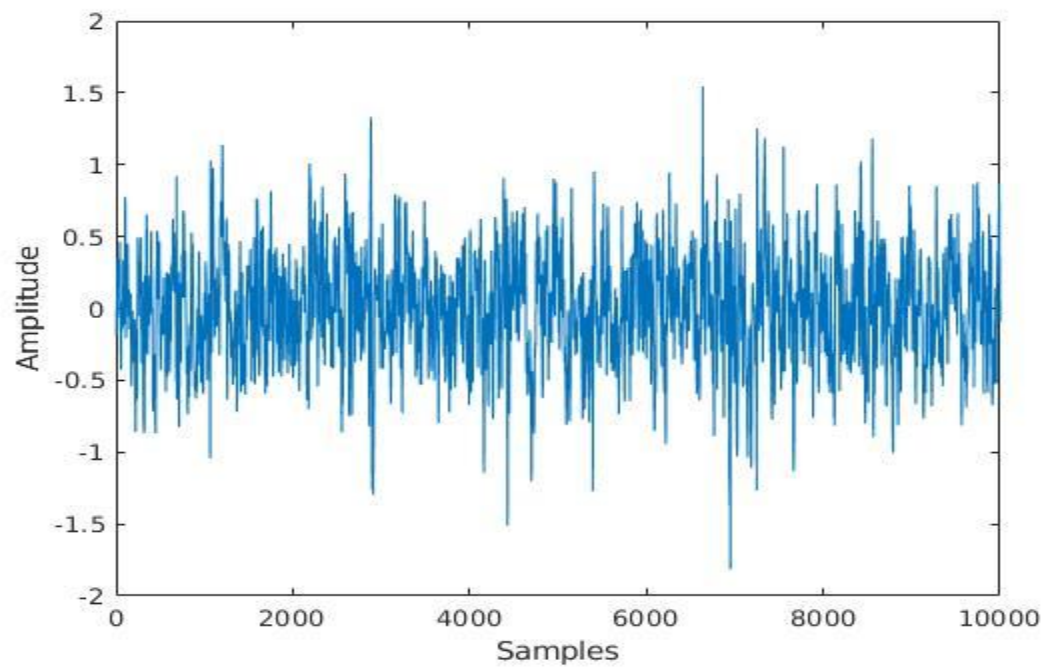


Figure 1.1.7 : Data_Test_7

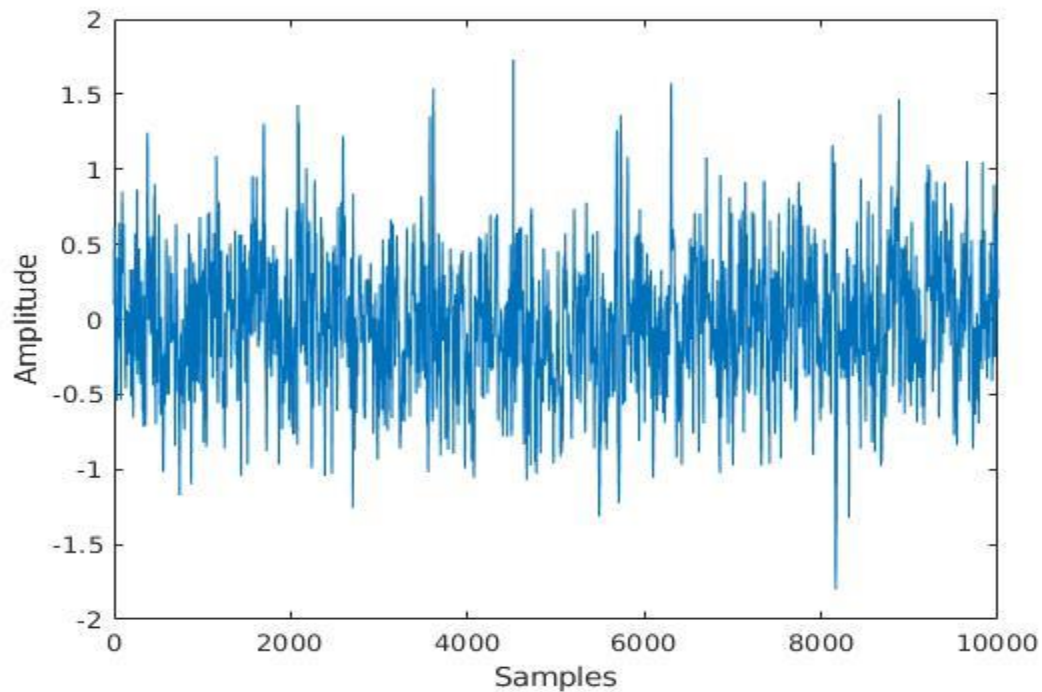


Figure 1.1.8 : Data_Test_8

Παρατηρείται όλο και περισσότερη δραστηριότητα στα γραφήματα όσο προχωράμε απο το ένα αρχείο στο επόμενο. Το γεγονός αυτό δηλώνει πως τα αρχικά σήματα δεν εμπεριέχουν τόσο μεγάλη επίδραση από θόρυβο ενώ όταν φτάνουμε στο τελευταίο η παρουσία του θορύβου είναι ιδιαίτερα αισθητή.

Ερώτημα 1.2 : Στην συνέχεια παρουσιάζονται τα γραφήματα των κορυφών που καταγράφονται κατά την εκτέλεση του κώδικα με την μέθοδο της κατωφλίωσης ($T = k \cdot \sigma_n$) για τιμές του k να κυμαίνονται από 0.1 έως 5 με βήμα 0.1 , για κάθε αρχείο Data_Test_#. Ουσιαστικά εφόσον υπολογιστεί η τιμή της σ_n για κάθε σήμα , ελέγχεται για όλες αυτές τις τιμές του k πόσα spikes θα καταμετρηθούν και αποθηκεύονται στον πίνακα με όνομα c1, οπότε κάνοντας plot του πίνακα αυτού και των τιμών του k προκύπτουν οι παρακάτω καμπύλες:

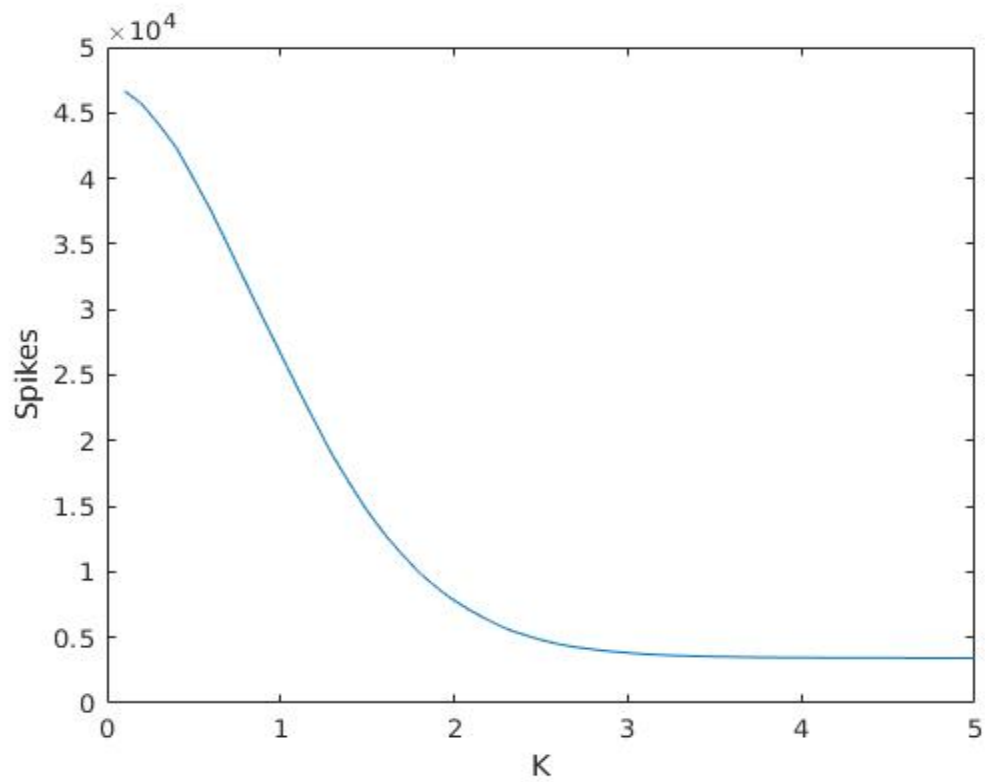


Figure 1.2.1 : Data_Test_1

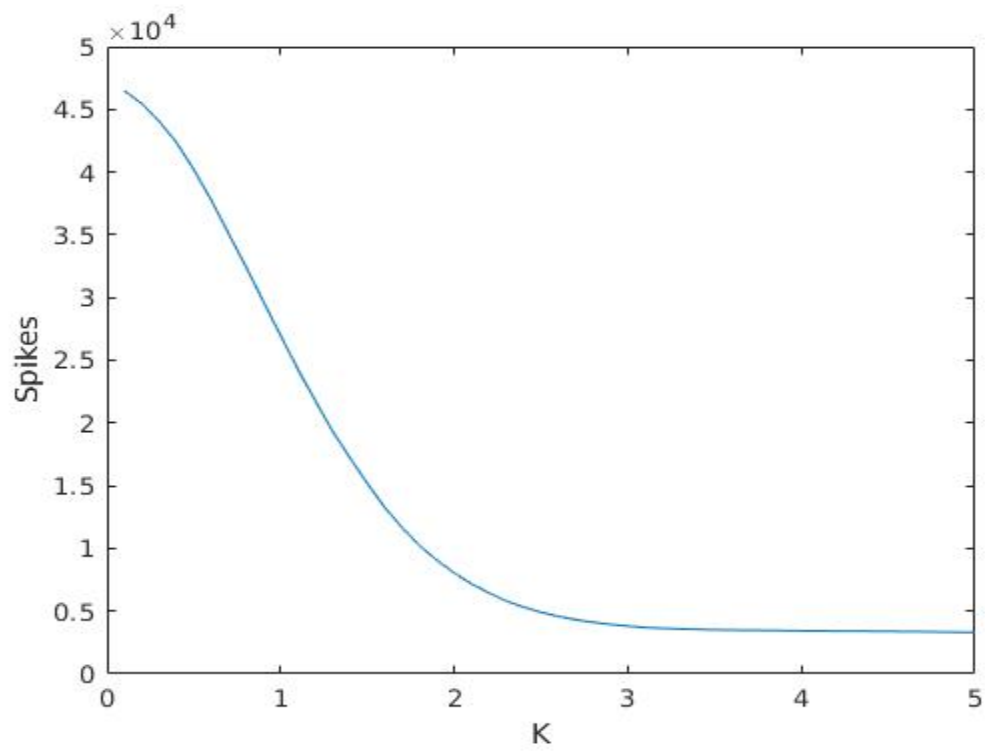


Figure 1.2.2 : Data_Test_2

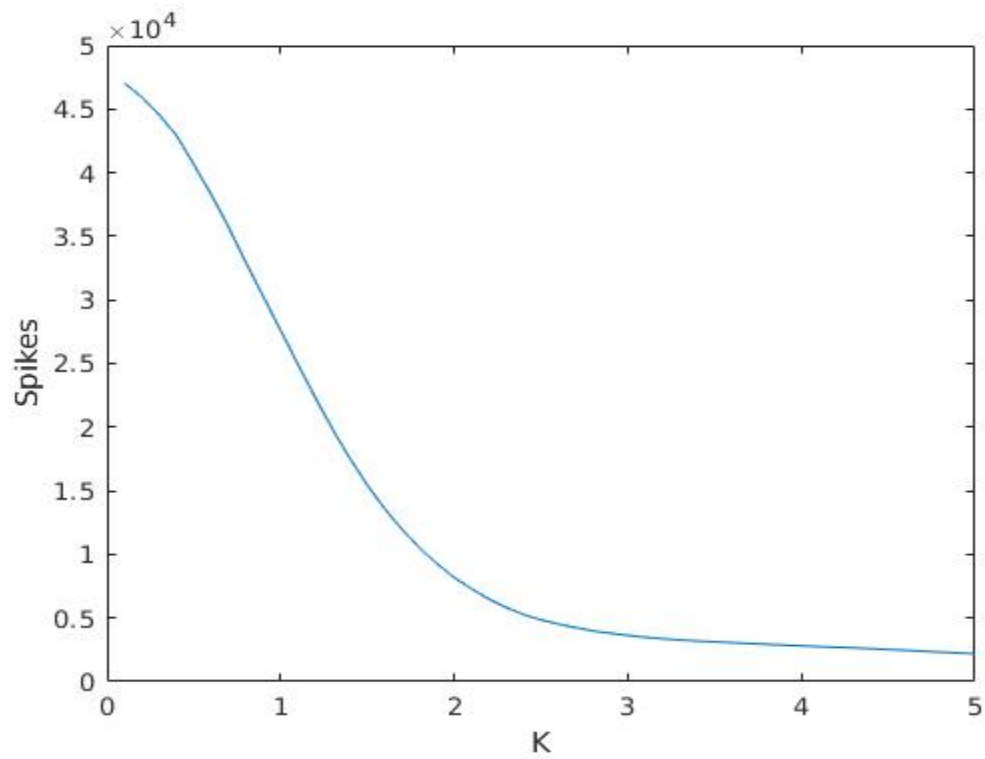


Figure 1.2.3 : Data_Test_3

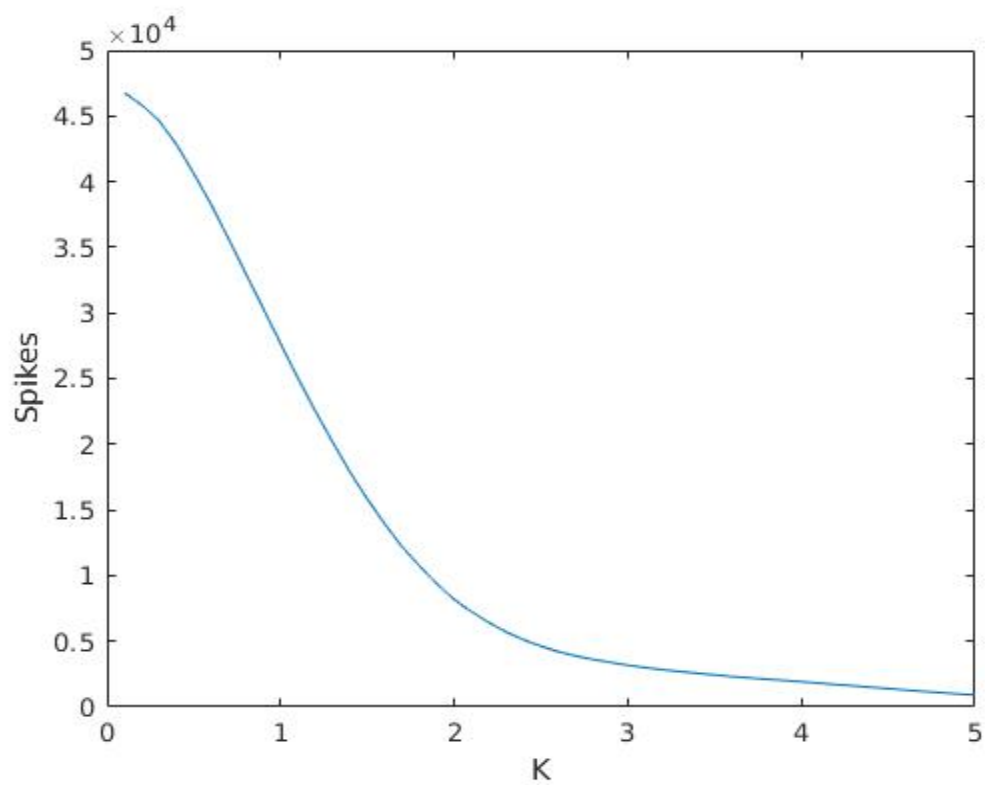


Figure 1.2.4 : Data_Test_4

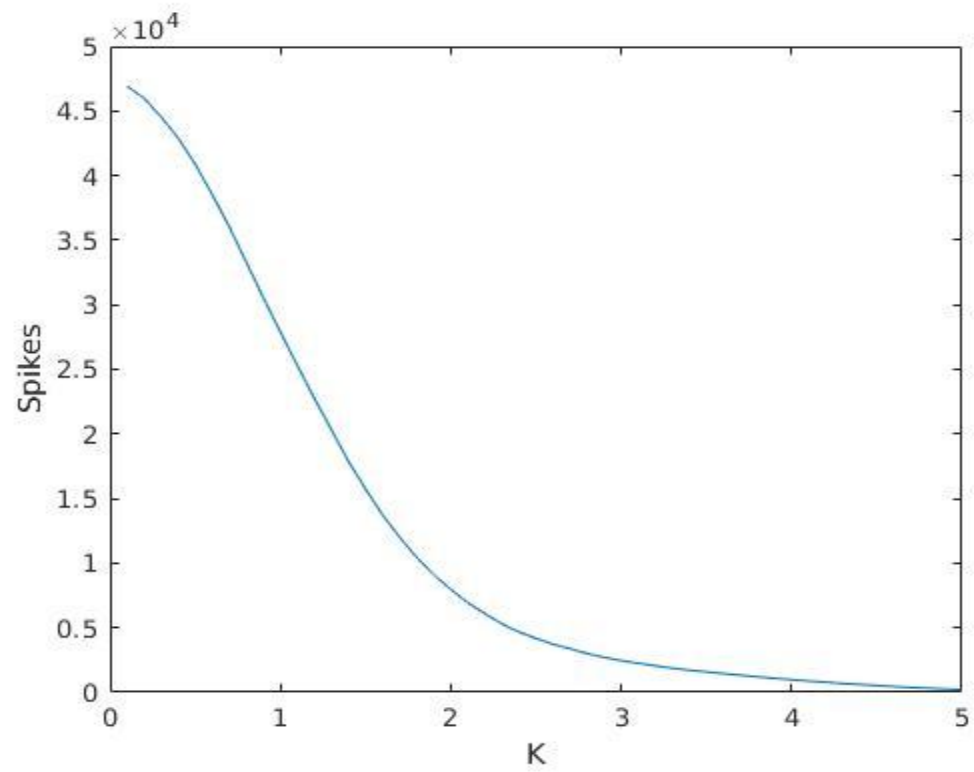


Figure 1.2.5 : Data_Test_5

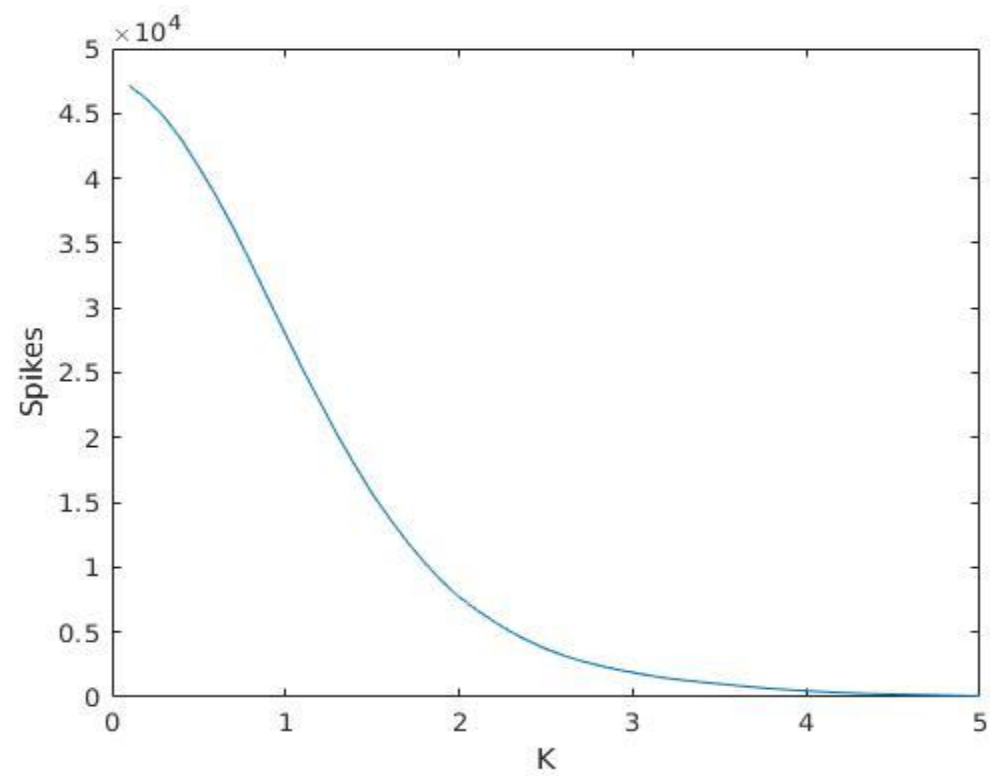


Figure 1.2.6 : Data_Test_6

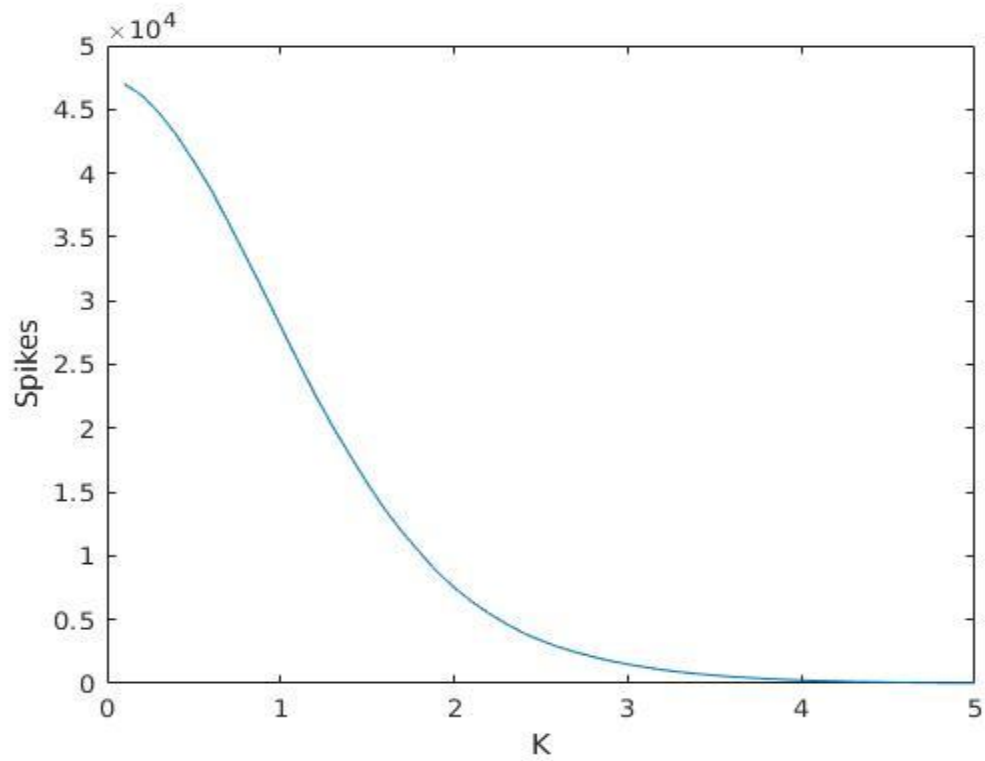


Figure 1.2.7 : Data_Test_7

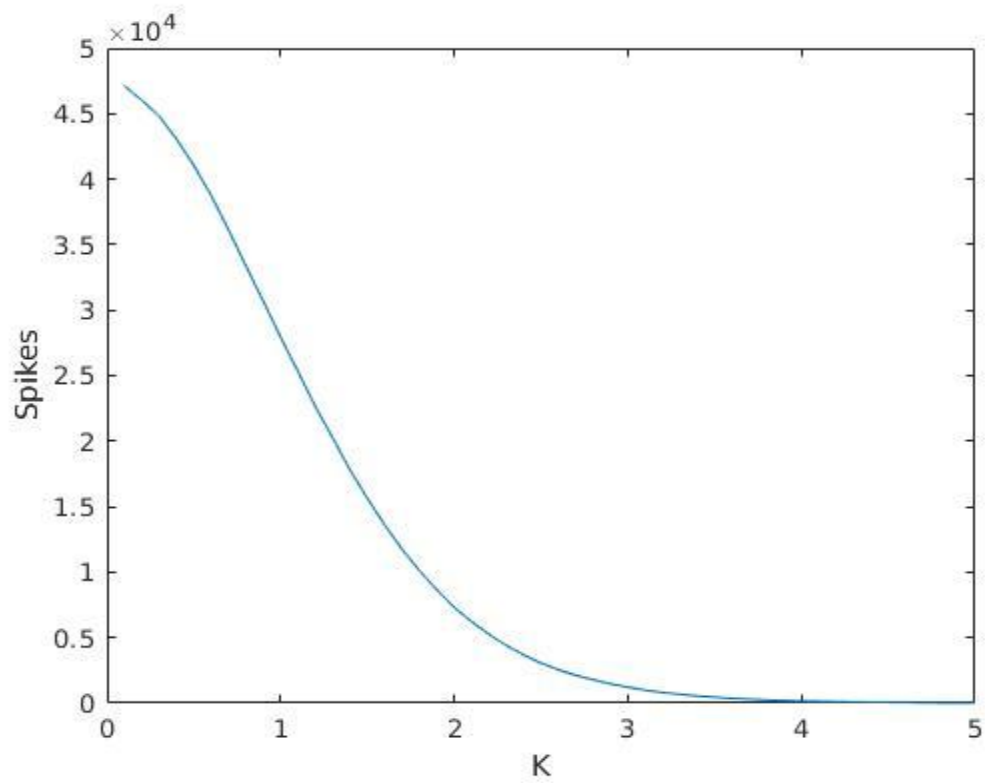


Figure 1.2.8 : Data_Test_8

Στα παραπάνω σχήματα φαίνεται πως όσο αυξάνεται ο συντελεστής k η τιμή του κατωφλίου αυξάνεται , οπότε λιγότερες τιμές ξεπερνούν την τιμή αυτή και με τον τρόπο αυτό μειώνεται ο αριθμός των υποτιθέμενων spikes. Επίσης στα αρχικά σήματα τα οποία δεν περιέχουν μεγάλη επίδραση θορύβου φαίνεται πως ενώ αυξάνεται το k παραμένει σχεδόν “σταθερός” ο αριθμός των spikes , ενώ στα τελευταία φαίνεται πως συνεχώς μειώνεται. Αυτό συμβαίνει διότι στα αρχικά εφόσον δεν υπάρχει η επίδραση του θορύβου μετά από κάποια τιμή του k πάνω από την τιμή του κατωφλίου θα υπάρχουν μόνο πραγματικά spikes.

Ερώτημα 1.3: Παρατηρώντας τα παραπάνω γραφήματα , φαίνεται πως το σημείο στο οποίο η τιμή του απαριθμητή $c1$ πλησιάζει την τιμή της μεταβλητής spikeNum, η οποία μας δίνεται , βρίσκεται περίπου στο σημείο από το οποίο διέρχεται μια ευθεία που περνά από την αρχή των αξόνων και έχει ένα συγκεκριμένο συντελεστή λ ($y = \lambda \cdot x$, όπου y είναι ο αριθμός των spikes και x ο συντελεστής k). Υπολογίζοντας λοιπόν , τα λ αυτά και για τα 8 σήματα και ελέγχοντας για ποιο από αυτά τα λ ξεχωριστά, έχουμε τις λιγότερες απώλειες μεταξύ των καταμετρημένων κορυφών και των πραγματικών , φαίνεται πως τελικά η τιμή του λ είναι 1213. Άρα αν φέρουμε μια ευθεία ($y = 1213 \cdot x$) από την αρχή των αξόνων θα τέμνει την καμπύλη του προηγούμενου ερωτήματος κοντά στο σημείο που αντιστοιχεί στον αριθμό των πραγματικών spikes.

➤ ΘΕΜΑ 2ο :

Ερώτημα 2.1 : Μέσω του εμπειρικού κανόνα του ερωτήματος 1.3 επιλέχθηκαν οι τιμές του συντελεστή k και ο αριθμός των spikes για κάθε σήμα που έχει δοθεί. Συγκεκριμένα για το αρχείο Data_Eval_E_1 ο αριθμός των spikes είναι 3653 , για το Data_Eval_E_2 είναι 3770 , για το Data_Eval_E_3 είναι 3730 και για το Data_Eval_E_4 είναι 3740. Στην συνέχεια δημιουργείται ο πίνακας spikeThresholdEst# για κάθε αρχείο στον οποίο αποθηκεύονται οι χρονικές στιγμές κατά τις οποίες το σήμα ξεπερνά την τιμή του κατωφλίου. Στην συνέχεια ελέγχοντας τις 40 γειτονικές τιμές του σήματος για κάθε spike δημιουργείται ο πίνακας wzspikeTimesEst# ο οποίος έχει αποθηκευμένες τις χρονικές στιγμές στις οποίες εμφανίζονται η κορυφές . (“ wz stands for with zeros “ .Για κάποιο λόγο στον πίνακα αυτόν εμφανίζονται μηδενικά , αλλά παρατηρώντας τις τιμές που προκύπτουν αν αφαιρεθούν τα μηδενικά δεν υπάρχει κάποιο σφάλμα , οπότε στην συνέχεια αφαιρούνται τα μηδενικά και ο νέος πίνακας που προκύπτει ονομάζεται spikeTimesEst)

Ερώτημα 2.2 : Σε αυτό το σημείο εφόσον στο προηγούμενο ερώτημα είχαν βρεθεί οι χρονικές στιγμές στις οποίες βρίσκονται τα spikes , δημιουργείται ο πίνακας wzspikesEst# (“wz” για τον ίδιο λόγο με παραπάνω) στον οποίο τοποθετούνται οι 40 γειτονικές τιμές , γύρω από την χρονική στιγμή στην οποία παρατηρείται κορυφή, από τον πίνακα data του κάθε αρχείου. Στην συνέχεια εμφανίζονται τα γραφήματα των

κυματομορφών των κορυφών και για τα 4 αρχεία.

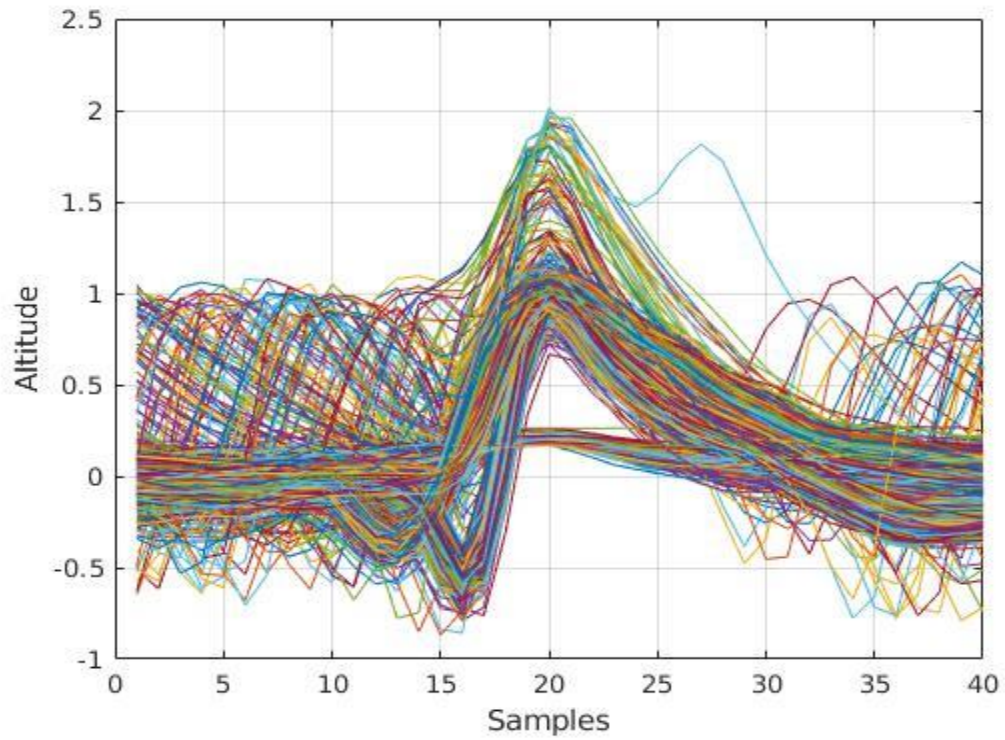


Figure 2.2.1 : Data_Eval_E_1

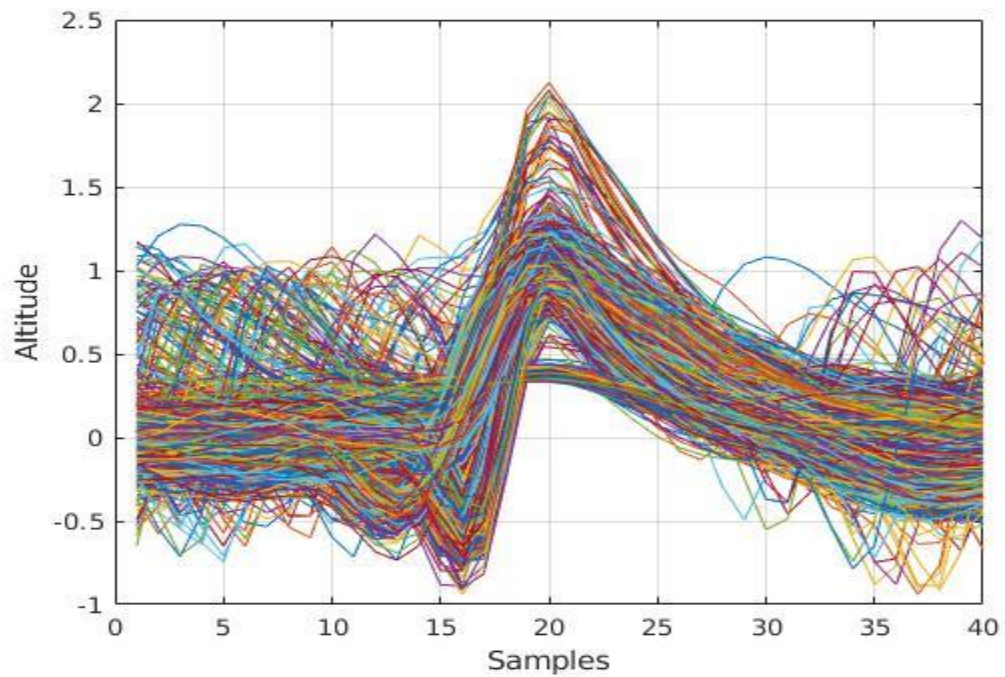


Figure 2.2.2 : Data_Eval_E_2

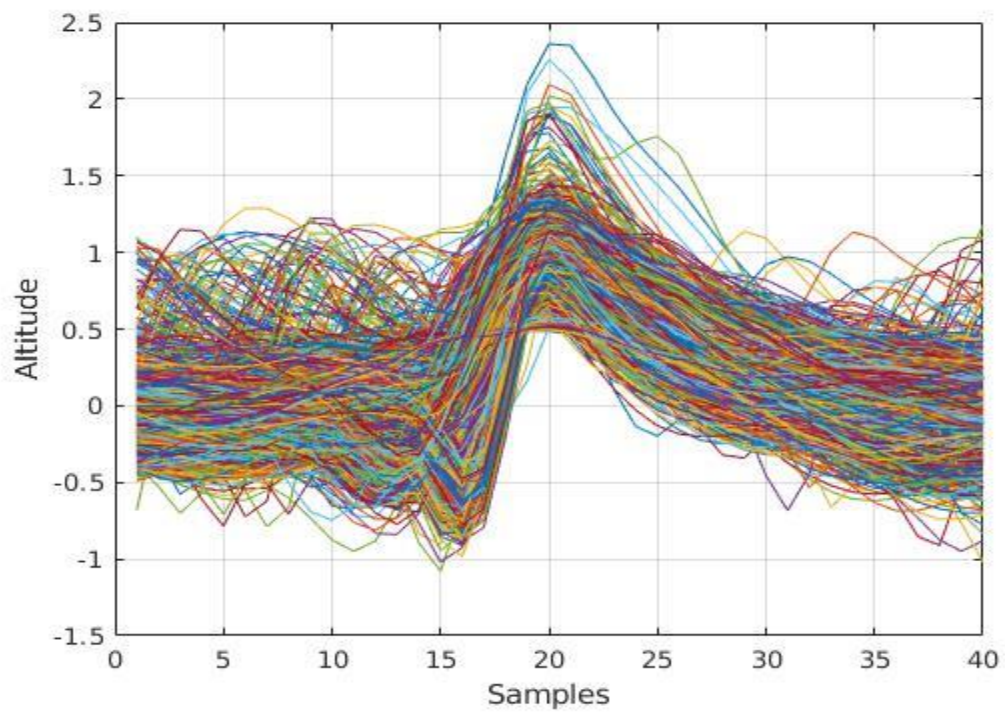


Figure 2.2.3 : Data_Eval_E_3

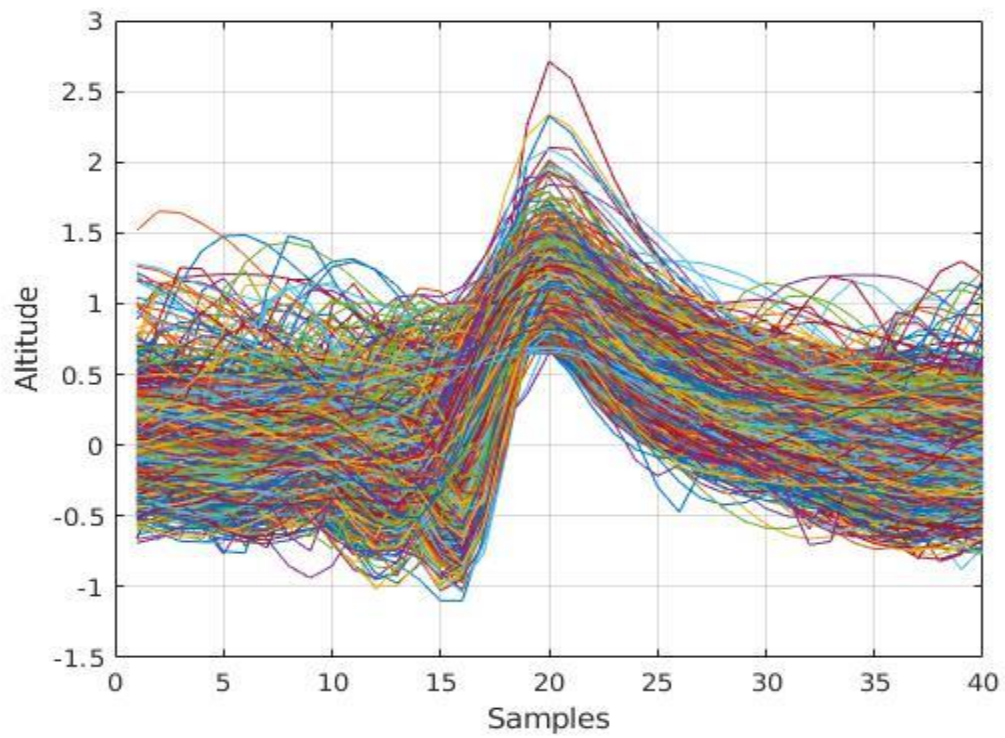


Figure 2.2.4 : Data_Eval_E_4

Στα παραπάνω γραφήματα παρατηρείται πως το σύνολο των κυματομορφών κυμαίνεται μεταξύ συγκεκριμένων τιμών και πέρα από αυτό δημιουργείται ένα spike το οποίο όπως φαίνεται έχει επηρεαστεί από θόρυβο. Η επίδραση αυτή είναι πιο έντονη στην αρχή του spike ,διότι σε αυτό το σημείο φαίνεται πώς υπάρχουν περισσότερες διαφορές και οι διαφορές αυτές είναι πιο έντονες στα δυο τελευταία σχήματα.

Ερώτημα 2.3 : Στο ερώτημα αυτό για να υπολογίσουμε τον αριθμό των spikes που αντιστοιχούν σε πραγματικά spikes και όχι σε θόρυβο ελέγχουμε την διαφορά των τιμών του πίνακα newSpikeTimesEst# και των spikeTimes που δίνονται να μην είναι μεγαλύτερη από 40, επίσης υπάρχει ένα ακόμα κριτήριο για να αφαιρεθεί κάποια τιμή από τον πίνακα newSpikeTimesEst# το οποίο περιγράφεται μέσα στον κώδικα . Ο αριθμός των πραγματικών spikes καταχωρείται στον πίνακα με όνομα errcounter#.

Ερώτημα 2.4: Επιλέγοντας σαν δύο χαρακτηριστικά την μέγιστη διαφορά μεταξύ δύο τιμών του κάθε spike και το μέγιστο μέτρο του μετασχηματισμού Fourier ως ζεύγη , ενώ έχει γίνει αντιστοίχιση του πίνακα spikeClass που δίνεται με τα spikes που έχουν βρεθεί απο μας στους πίνακες num#, παράγονται τα παρακάτω σχήματα :

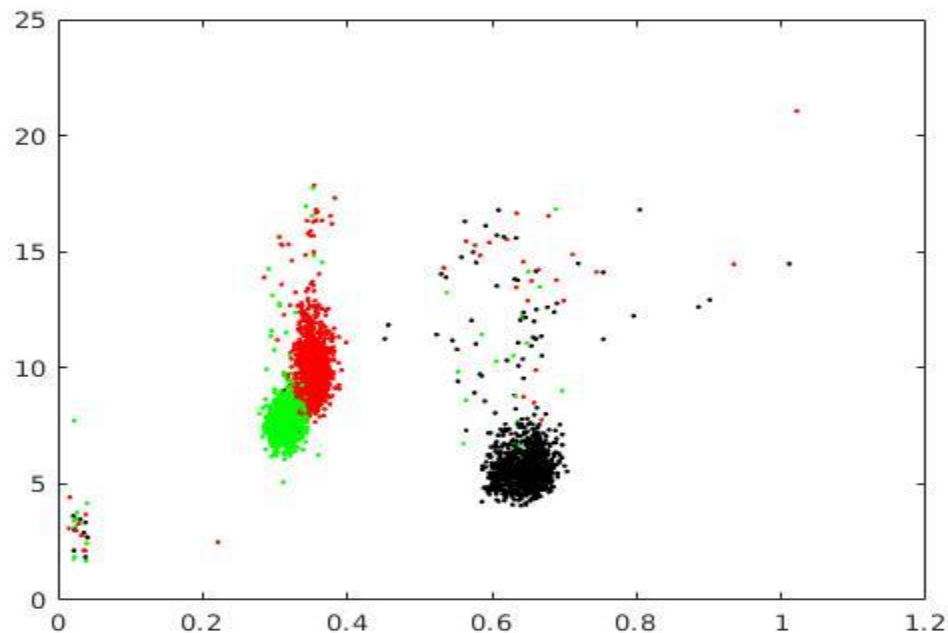


Figure 2.4.1 : Data_Eval_1

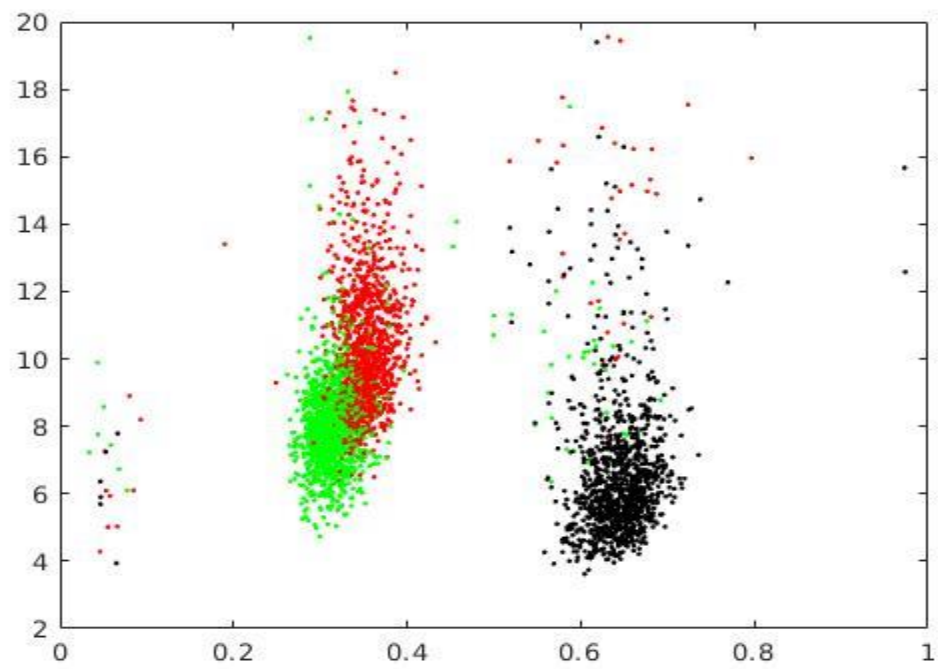


Figure 2.4.2 : Data_Eval_2

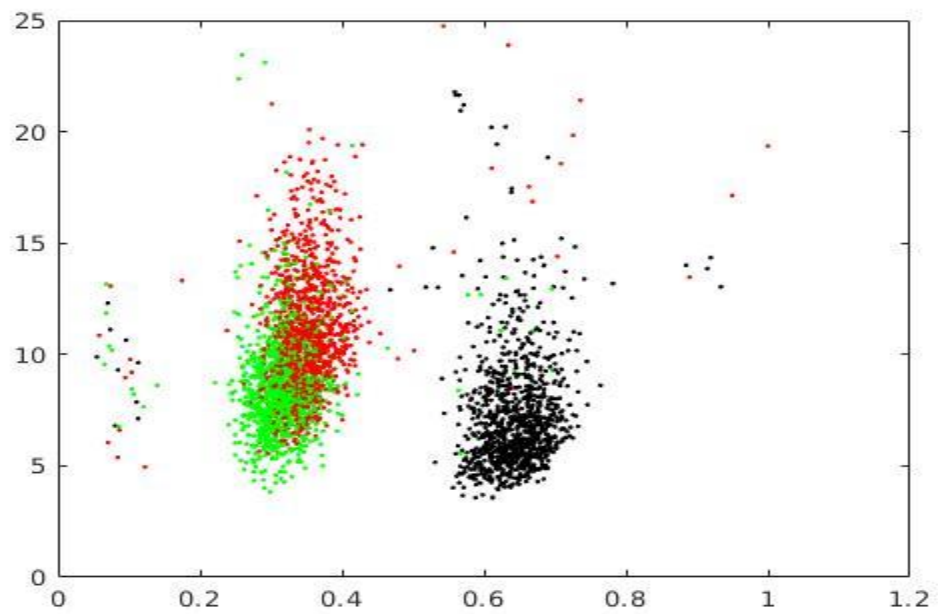


Figure 2.4.3 : Data_Eval_3

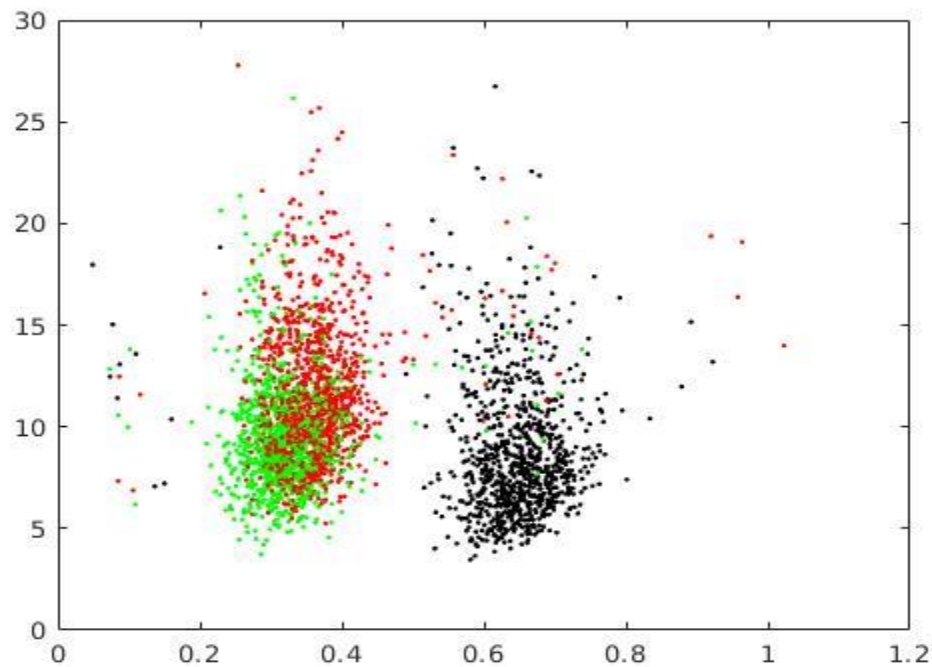


Figure 2.4.4 : Data_Eval_4

Στα γραφήματα αυτά φαίνεται πάλι πως, επειδή στα δύο τελευταία σήματα η επίδραση του θορύβου είναι μεγαλύτερη, η διαχώριση των spikes δεν είναι τόσο αποτελεσματική (Χρώματα πράσινο και κόκκινο).

Ερώτημα 2.5: Χρησιμοποιώντας τα συγκεκριμένα χαρακτηριστικά κάθε φορά που αυξάνεται ο αριθμός τους το ποσοστό επιτυχίας που επιστρέφει ο αλγόριθμος MyClassify μεγαλώνει. Βέβαια για κάθε χαρακτηριστικό το ποσοστό αυξάνεται περισσότερο ή λιγότερο. Δηλαδή όταν χρησιμοποιούνται χαρακτηριστικά όπως ο αριθμός των φορών που το spike διέρχεται από το μηδέν, ο οποίος είναι περίπου ίδιος για όλα τα spikes και επειδή οι τιμές είναι διακριτές, το ποσοστό δεν αυξάνεται πολύ. Ενώ από την άλλη η μέγιστη διαφορά μεταξύ δύο τιμών του κάθε spike βοηθάει περισσότερο στην διαχώριση των νευρώνων λόγω των διαφορετικών τιμών της μεταβλητής diff. Άρα κάθε χαρακτηριστικό μπορεί να προσφέρει διαφορετική αύξηση στο ποσοστό του classification. Τα τελικά ποσοστιαία αποτελέσματα κυμαίνονται από 89-96 τοις εκατό.