



Πανεπιστήμιο Δυτικής Μακεδονίας
Τμήμα Μηχανολόγων Μηχανικών
Πολυτεχνική Σχολή

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΣΥΓΚΡΙΤΙΚΗ ΔΙΕΡΕΥΝΗΣΗ ΜΕΘΟΔΩΝ ΕΝΗΜΕΡΩΣΗΣ ΜΟΝΤΕΛΩΝ
ΠΕΠΕΡΑΣΜΕΝΩΝ ΣΤΟΙΧΕΙΩΝ ΜΕ ΓΡΑΜΜΙΚΑ ΚΑΙ ΜΗ ΓΡΑΜΜΙΚΑ
ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ

COMPERATIVE INVESTIGATION OF OPTIMIZATION METHODS IN
UPDATING FINITE ELEMENT MODELS WITH LINEAR AND NON-
LINEAR CHARACTERISTICS

ΦΡΑΓΚΙΑΔΗΣ ΓΕΩΡΓΙΟΣ
ΑΕΜ 1717

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ : ΑΡΑΪΛΟΠΟΥΛΟΣ ΑΛΕΞΑΝΔΡΟΣ

ΚΟΖΑΝΗ 2024



Πανεπιστήμιο Δυτικής Μακεδονίας
Τμήμα Μηχανολόγων Μηχανικών
Πολυτεχνική Σχολή

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΣΥΓΚΡΙΤΙΚΗ ΔΙΕΡΕΥΝΗΣΗ ΜΕΘΟΔΩΝ ΕΝΗΜΕΡΩΣΗΣ ΜΟΝΤΕΛΩΝ
ΠΕΠΕΡΑΣΜΕΝΩΝ ΣΤΟΙΧΕΙΩΝ ΜΕ ΓΡΑΜΜΙΚΑ ΚΑΙ ΜΗ ΓΡΑΜΜΙΚΑ
ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ

COMPERATIVE INVESTIGATION OF OPTIMIZATION METHODS IN
UPDATING FINITE ELEMENT MODELS WITH LINEAR AND NON-
LINEAR CHARACTERISTICS

ΦΡΑΓΚΙΑΔΗΣ ΓΕΩΡΓΙΟΣ
ΑΕΜ 1717

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ : ΑΡΑΪΛΟΠΟΥΛΟΣ ΑΛΕΞΑΝΔΡΟΣ

ΚΟΖΑΝΗ 2024

Περίληψη

Ο σκοπός της παρούσας διπλωματικής είναι να μελετήσει την απόδοση μεθόδων βελτιστοποίησης για ενημέρωση μοντέλων πεπερασμένων στοιχείων. Η ενημέρωση μοντέλων πεπερασμένων στοιχείων έχει πολλές χρήσεις που ποικίλουν από ενημέρωση μοντέλων ώστε να αποκτήσουμε μοντέλα με μεγαλύτερη πιστότητα για να βελτιώσουμε διαδικασίες σχεδιασμού, ανάπτυξης, προσημείωσης δοκιμής δομών και προϊόντων μέχρι εφαρμογές παρακολούθησης δομικής ακεραιότητας. Συγκεκριμένα θα χρησιμοποιήσουμε τρία μοντέλα για να συγκρίνουμε την απόδοση του CMA-ES αλγορίθμου, δύο διαφορετικών υλοποιήσεων του αλγορίθμου PSO και την απόδοση της συνάρτησης `fmincon` του MATLAB. Τα μοντέλα που θα εξετάσουμε είναι ένα μοντέλο μισού οχήματος (half-car), ένα μοντέλο δικτύωματος και ένας μη-γραμμικός ταλαντωτής `duffing`.

Abstract

The purpose of this thesis is to investigate the performance of optimization methods for updating finite elements models. Finite Elements Model updating has many applications that span from acquiring a model with increased fidelity in order to facilitate improved developing, simulating and testing processes, to Structural Health Monitoring. In detail three models will be used to compare the CMA-ES algorithm with two variants of PSO algorithms and MATALAB'S fmincon function. These models represent a half-car, a truss and a non-linear duffing oscillator

Ευχαριστίες

Ευχαριστώ τον Θεό, την οικογένεια μου, το διδακτικό προσωπικό του Πανεπιστημίου Δυτικής Μακεδονίας για την μόρφωση που μου παρείχε και τον κ. Αραϊλόπουλο Αλέξανδρο για την αρωγή και την καθοδήγηση που μου παρείχε για την διεκπεραίωση της παρούσας διπλωματικής εργασίας.

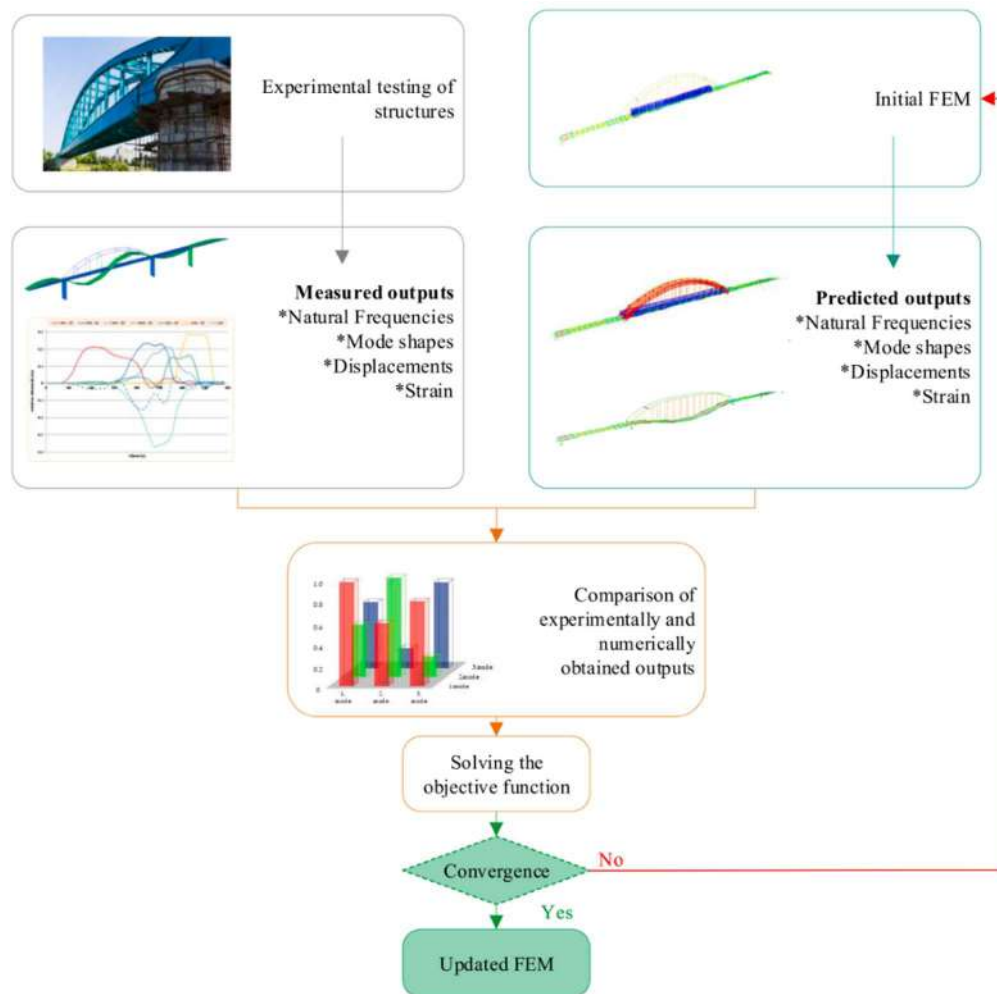
Περιεχόμενα

Περίληψη.....	3
Abstract	4
Ευχαριστίες.....	5
1. Ενημέρωση Μοντέλου και Πρόβλημα βελτιστοποίησης	8
2. Εξελικτικοί Αλγόριθμοι και Στρατηγικές Εξέλιξης.....	11
2.1. Εισαγωγή	11
2.2 Αρχικοποίηση του πληθυσμού.....	13
2.3. Αξιολόγηση και επιλογή.	14
2.4. Ανασυνδυασμός.....	16
2.5. Μετάλλαξη	17
2.6. Ανασυνδυασμός και Επιδιόρθωση.....	17
2.7. Το όφελος του ανασυνδυασμού – Γενετική επιδιόρθωση	18
2.8. Βαθμός Προόδου και Επιδιόρθωση	20
3. Αυτοπροσαρμογή των ενδογενών παραμέτρων.....	25
3.1. Κανόνας του $1/5^{0u}$ και Αυτοπροσαρμογή.	25
3.2. Αποτυχαιοποίηση.....	29
3.2.1. 1 ^ο Επίπεδο	29
3.2.2. 2 ^ο Επίπεδο.	32
3.2.3. Προσαρμογή του μητρώου Συνδιακύμανσης.....	35
4. Στρατηγική εξέλιξης CMA-ES.	39
5. Νοημοσύνη Σμήνους	43
5.1. Αλγόριθμος PSO.....	45
5.2. Παράμετροι Αλγόριθμου PSO	47
5.3. Αρχικοποίηση πληθυσμού PSO	48
5.4. Γειτονιές σωματιδίων.....	51
5.5. Είναι Στρατηγική εξέλιξης (ES) ο PSO ;	54
5.6. Εκδοχές Αλγορίθμων SPSO που χρησιμοποιήθηκαν	55
6. Επιλογή παραμέτρων των αλγορίθμων και μέτρα σύγκρισης.	57
7. Αποτελέσματα δοκιμών ενημέρωσης μοντέλου.....	59
7.1. Ενημέρωση μοντέλου Δικτυώματος.	59
7.2. Μοντέλο μισού αυτοκινήτου (Half-car)	61
7.2.1 Δοκιμή μοντέλου half-car με τυχαίο αρχικό σημείο.....	80

7.2.3. Σχολιασμός Αποτελεσμάτων CMA-ES	89
7.2.4 Δοκιμή μοντέλου half-car με διαφορετική τοποθέτηση.....	94
7.3. Ενημέρωση μοντέλου με μη γραμμικότητες.....	98
7.3.1 Δοκιμή του ταλαντωτή duffing σε διαφορετική διέγερση (Περίπτωση 2)	114
Συμπεράσματα	125
Βιβλιογραφικές Αναφορές	126

1. Ενημέρωση Μοντέλου και Πρόβλημα βελτιστοποίησης

Στην σημερινή εποχή τα αριθμητικά μοντέλα κατέχουν σημαντική θέση στην επιστήμη του μηχανικού. Μέθοδοι όπως αυτή των πεπερασμένων στοιχείων μπορούν να μας βοηθήσουν να προσεγγίσουμε λύσεις περίπλοκων συστημάτων διαφορικών εξισώσεων που περιγράφουν φυσικά συστήματα ώστε να σχεδιάσουμε και να δοκιμάσουμε δομές πριν τις κατασκευάσουμε ή να προβλέψουμε την συμπεριφορά υφιστάμενων δομών και να τις βελτιώσουμε. Ανάμεσα όμως στο αριθμητικό μοντέλο και στην πραγματική δομή μπορεί να προκύψει απόκλιση. Το σφάλμα αυτό μπορεί να προέρχεται από λανθασμένες παραδοχές γύρω από τις οριακές συνθήκες, λανθασμένες τιμές στις παραμέτρους των ιδιοτήτων των υλικών, κακή διακριτοποίηση, απλοποιήσεις λόγο παραδοχών κ.ά. Έτσι κρίνεται αναγκαίο να ενημερωθεί το αριθμητικό μοντέλο ώστε να περιγράψει με μεγαλύτερη πιστότητα την υπό εξέταση δομή. Η τελευταία μπορεί να θεωρηθεί ένα σύστημα το οποίο μπορούμε να τροφοδοτήσουμε με ένα σήμα εισόδου(διέγερση) και να μετρήσουμε το σήμα εξόδου(απόκριση). Τα δεδομένα από τέτοιες δοκιμές μπορούν να χρησιμοποιηθούν ώστε συγκρίνοντας τα αποτελέσματα του αριθμητικού μοντέλου με τα πειραματικά δεδομένα, να προσαρμόσουμε το αριθμητικό μας μοντέλο. Αυτή η διαδικασία ονομάζεται ενημέρωση μοντέλου και εντάσσεται στην ευρύτερη έννοια της αναγνώρισης συστήματος. Οι χρήσεις της μεθόδου αυτής ποικίλουν από μια απλή ενημέρωση παραμέτρων ώστε το μοντέλο να αποκτήσει μεγαλύτερη πιστότητα μέχρι παρακολούθηση δομικής ακεραιότητας κ.ά. Έχοντας τα πειραματικά δεδομένα και ενώ έχουμε καθορίσει τις παραμέτρους που θέλουμε να αναγνωρίσουμε, επόμενο βήμα είναι η ποσοτικοποίηση της διαφοράς ανάμεσα στα πειραματικά δεδομένα και την απόκριση του μοντέλου μας. Η συνάρτηση που περιγράφει την διαφορά αυτή καλείται αντικειμενική συνάρτηση και στόχος μας γίνεται η επιλογή παραμέτρων ώστε να ελαχιστοποιηθεί αυτή για να ελαχιστοποιηθεί ακολούθως η διαφορά του μοντέλου μας από το πραγματικό σύστημα. Οπότε στον πυρήνα της διαδικασίας ενημέρωσης μοντέλου έχουμε να κάνουμε με ένα πρόβλημα βελτιστοποίησης. Στο παρακάτω διάγραμμα παρουσιάζεται η εν γέννη ροή της ενημέρωσης μοντέλου.



Εικόνα 1. Διάγραμμα ροής Διαδικασίας Ενημέρωσης μοντέλου [1].

Γενικά ένα πρόβλημα βελτιστοποίησης έχει ως αντικείμενο να βελτιστοποιήσει κάποιο δοσμένο στόχο ή κάποια συνάρτηση (ή συναρτήσεις) F αναζητώντας τιμές γύρω από ένα σύνολο μεταβλητών \mathbf{y} (παράμετροι σχεδιασμού ή αντικειμενικές μεταβλητές), ώστε

$$F(\mathbf{y}) \rightarrow \text{opt για } \mathbf{y} \in \mathcal{Y}$$

Το \mathcal{Y} μπορεί να είναι οποιοδήποτε σύνολο όπως ο N – διάστατος χώρος αναζήτησης με πραγματικές τιμές στο \mathbb{R}^N , ή ο χώρος ακεραίων \mathbb{Z}^N , ο διάδικός \mathbb{B}^N κ.ά. [2]. Συνήθως έχουμε να κάνουμε με προβλήματα στον \mathbb{R}^N . Το πρόβλημα βελτιστοποίησης μπορεί να υπόκειται σε διάφορους περιορισμούς όπως :

$$g_i(\mathbf{y}) \leq 0, i \in I = \{1, \dots, m\}, \text{ ανισωτικοί περιορισμοί}$$

$$h_j(\mathbf{y}) = 0, i \in J = \{1, \dots, r\}, \text{ περιορισμοί ισότητας}$$

ενώ το σύνολο

$$\mathcal{B} = \{\mathbf{y} \in \mathbb{R}^N : g_i(\mathbf{y}) \leq 0, \forall i \in I, h_j(\mathbf{y}) = 0, \forall i \in J\}$$

είναι το σύνολο των εφικτών λύσεων που ορίζει τον χώρο αναζήτησης του προβλήματος [3]. Η αντικειμενική συνάρτηση στην ενημέρωση μοντέλου (Model Updating) μπορεί να λάβει διάφορες μορφές αναλόγως με τα πειραματικά δεδομένα που λαμβάνουμε και επιλέγουμε

να συγκρίνουμε. Για παράδειγμα αν χρησιμοποιούμε δεδομένα από την ιδιομορφική ανάλυση της κατασκευής με $D = \{\widehat{\omega}_r, \widehat{\boldsymbol{\varphi}}_r \in R^{N_o}, r = 1, \dots, m\}$, όπου $\widehat{\omega}_r$ ιδιοσυχνότητες και $\widehat{\boldsymbol{\varphi}}_r$ ιδιοδιανύσματα που προέρχονται από μετρήσεις στην δομή σε N_o βαθμούς ελευθερίας και m το πλήθος των ιδιομορφών, τότε αναζητούμε το σύνολο \mathbf{y} των παραμέτρων ώστε οι ιδιομορφές και οι ιδιοσυχνότητες που προκύπτουν από το μοντέλο ως $\{\omega_r(\mathbf{y}), \boldsymbol{\varphi}_r(\mathbf{y}) \in R^{N_o}, r = 1, \dots, m\}$, να παρουσιάζουν την μικρότερη διαφορά από τις αντίστοιχες πειραματικές [4]. Η διαφορά αυτή για τις ιδιοσυχνότητες ποσοτικοποιείται με την παρακάτω σχέση :

$$\varepsilon_{\omega_r}(\mathbf{y}) = \frac{\omega_r^2(\mathbf{y}) - \widehat{\omega}_r^2}{\widehat{\omega}_r^2} \quad (1)$$

Επομένως η αντικειμενική μας συνάρτηση λαμβάνει την μορφή $F(\mathbf{y}) = \sum_{r=1}^m \varepsilon_{\omega_r}(\mathbf{y})$ (2) και έχουμε να κάνουμε με ένα πρόβλημα ελαχιστοποίησης. Παρόμοια αν τα πειραματικά δεδομένα αποτελούνται από την απόκριση μερικών σημείων της κατασκευής στον χρόνο μετρώντας για παράδειγμα την επιτάχυνση x_{ij} , όπου i η θέση του αισθητήρα από τον οποίο προέρχεται η μέτρηση και j το χρονικό βήμα, τότε τοποθετώντας N αισθητήρες σε N σημεία η αντικειμενική συνάρτηση λαμβάνει την μορφή

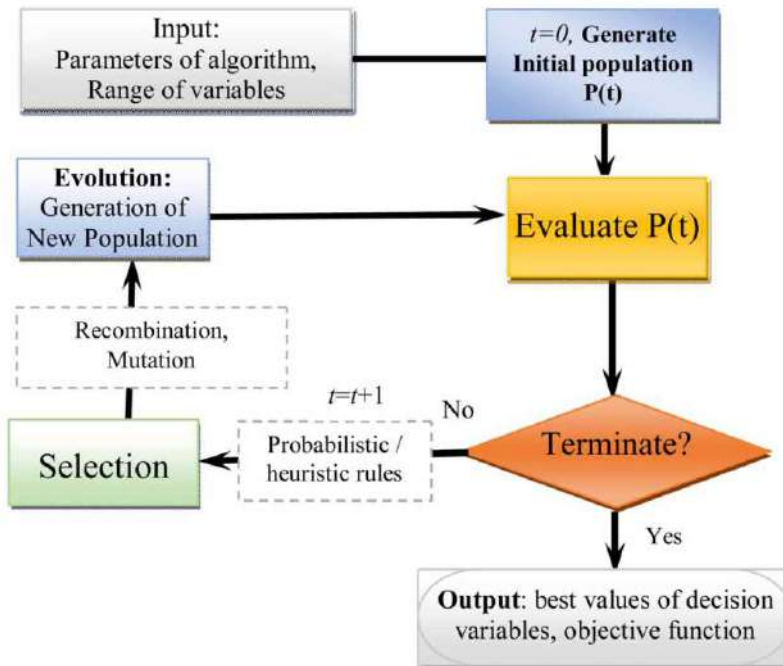
$$F(\mathbf{y}) = \frac{1}{N} \sum_{i=1}^N \frac{\sum_{j=1}^M (x(\mathbf{y})_{ij} - \hat{x}_{ij})^2}{\sum_{j=1}^M (\hat{x}_{ij})^2} \quad (3)$$

όπου M ο συνολικός αριθμός βημάτων στον χρόνο [4].

Οι μέθοδοι επίλυσης ενός προβλήματος βελτιστοποίησης ποικίλουν όμως μπορούν να χωριστούν σε δύο μεγάλες κατηγορίες, ντετερμινιστικές και στοχαστικές. Οι ντετερμινιστικές μέθοδοι συνήθως χρησιμοποιούν παραγώγους και κατηγοριοποιούνται περαιτέρω σε δεύτερης τάξης αν χρησιμοποιούν παραγώγους δεύτερης τάξης (Newton-like), πρώτης τάξης αν χρησιμοποιούν παραγώγους πρώτης τάξης (μεθόδοι κλιμακωτής μείωσης και απότομης κατάβασης) και σε μεθόδους άμεσης αναζήτησης όπου χρειάζεται μόνο η τιμή της αντικειμενικής συνάρτησης. Στις ντετερμινιστικές μεθόδους συνήθως επιλέγεται μια αρχική λύση και η πληροφορία που προέρχεται από τα μαθηματικά κριτήρια ακολουθείται με αιτιοκρατικό τρόπο ώστε να βελτιωθεί η λύση. Αν η λύση που επιλέχθηκε είναι μακριά από το ολικό βέλτιστο ή/και το πρόβλημα παρουσιάζει ασυνέχειες, τότε είναι πολύ πιθανό να παγιδευτεί ο αλγόριθμος σε κάποιο τοπικό ακρότατο. Επομένως η λύση σε αυτή την περίπτωση εξαρτάται σε μεγάλο βαθμό από την εκλογή της αρχικής λύσης. Αντίθετα στους στοχαστικούς αλγόριθμους βελτιστοποίησης η μετάβαση από μια λύση στην επόμενη περιλαμβάνουν στοχαστικές διαδικασίες που εισάγουν τυχαιότητα. Μια ιδιαίτερη κατηγορία στοχαστικών αλγορίθμων βελτιστοποίησης είναι αυτή των Εξελικτικών Αλγορίθμων. Στην συνέχεια της εργασίας θα χρησιμοποιήσουμε αλγορίθμους που ανήκουν σε αυτή την κατηγορία για να πετύχουμε ενημέρωση παραμέτρων τριών μοντέλων πεπερασμένων στοιχείων που αντιπροσωπεύουν ένα δικτύωμα, μοντέλο μισού οχήματος και έναν ταλαντωτή duffing. Στη μέθοδο ενημέρωσης μοντέλου πειραματικά δεδομένα συγκρίνονται με τα αποτελέσματα του μοντέλου ώστε να ενημερώσουμε ή να αναγνωρίσουμε μερικές επιλεγμένες παραμέτρους. Απουσία τέτοιων δεδομένων, στην παρούσα εργασία εκλέγουμε ένα ρεαλιστικό σύνολο παραμέτρων και θεωρούμε ως πειραματικά δεδομένα την απόκριση των μοντέλων για το σύνολο αυτό. Κρατώντας την απόκριση αυτή, στην συνέχεια αγνοούμε τις τιμές αυτές του συνόλου αυτού και αναζητούμε να τις αναγνωρίσουμε, προσομοιώνοντας έτσι την πραγματική διαδικασία που εφαρμόζεται.

2. Εξελικτικοί Αλγόριθμοι και Στρατηγικές Εξέλιξης

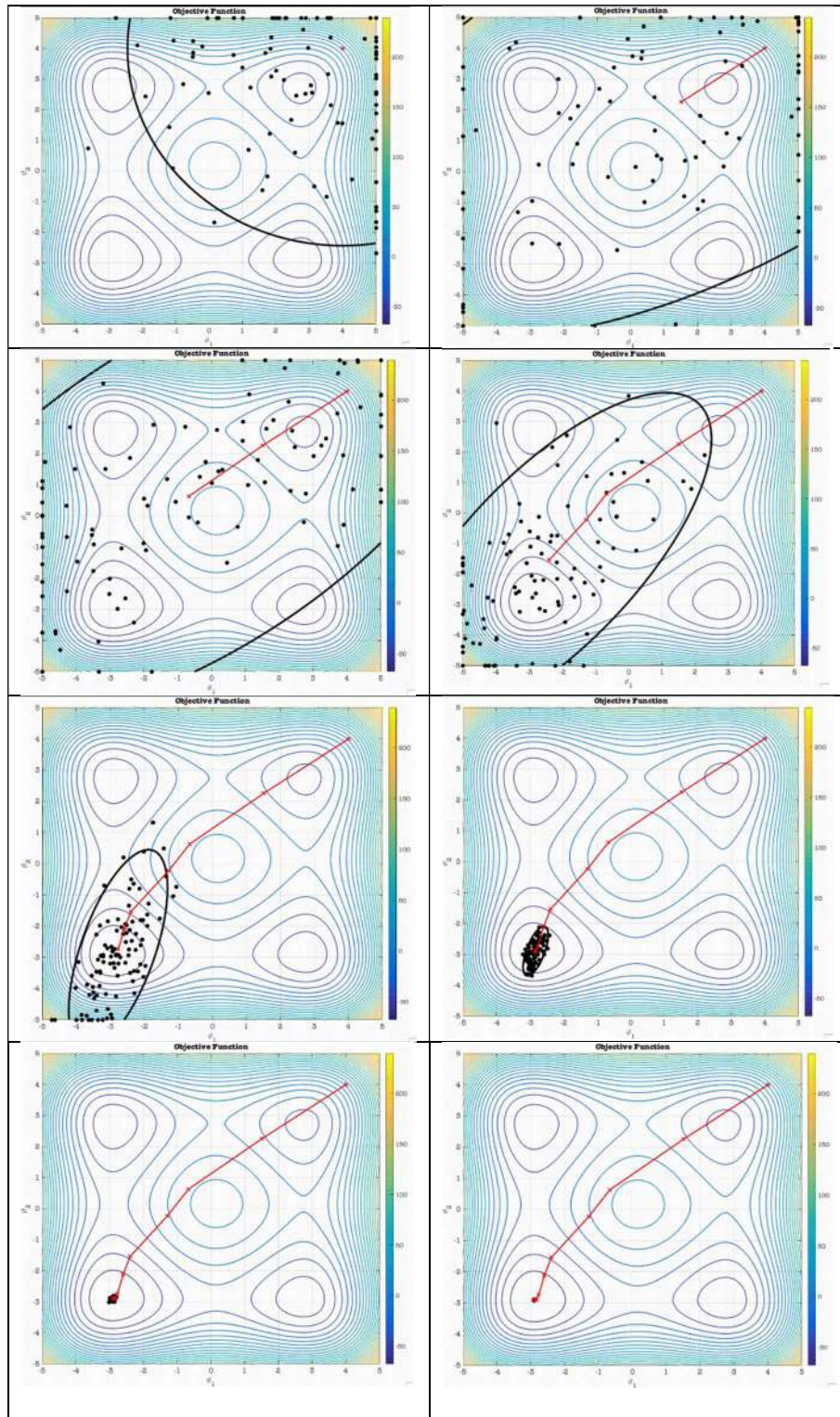
2.1. Εισαγωγή



Εικόνα 2 [5]. Διάγραμμα Ροής Εξελικτικών Αλγορίθμων.

Οι εξελικτικοί αλγόριθμοι (Evolution Algorithms, EA's) παρουσιάζουν τα τελευταία χρόνια εξαιρετικό ενδιαφέρον στον τομέα της βελτιστοποίησης [2]. Αυτοί οι αλγόριθμοι είναι επαναληπτικές στοχαστικές διαδικασίες εμπνευσμένες από την βιολογία. Βασικά τους στοιχεία είναι η δημιουργία ενός πληθυσμού, όπου κάθε άτομο(υποψήφιος) αντιπροσωπεύει ένα σημείο στον χώρο των δυνατών λύσεων, ένα κριτήριο για την αξιολόγηση της καταλληλότητας (fitness) του σημείου και μερικές διαδικασίες εμπνευσμένες από την βιολογία (ανασυνδυασμός , μετάλλαξη κ.ά.). Το κριτήριο αυτό που αξιολογεί την καταλληλότητα είναι η τιμή της αντικειμενικής συνάρτησης του προβλήματος βελτιστοποίησης για το σετ παραμέτρων του σημείου. Αν το πρόβλημα βελτιστοποίησης είναι διαμορφωμένο έτσι ώστε να είναι ένα πρόβλημα ελαχιστοποίησης τότε σκοπός των αλγορίθμων αυτών είναι να βρεθεί ένα σημείο για το οποίο η αντικειμενική συνάρτηση δίνει την ελάχιστη δυνατή τιμή. Στην παρούσα διπλωματική έχουμε να κάνουμε με προβλήματα ελαχιστοποίησης και οι έννοιες της συνάρτησης καταλληλότητας και της αντικειμενικής συνάρτησης στην χρησιμοποιούνται ταυτόσημα. Έτσι μια χαμηλή τιμή σε αυτήν την εξίσωση αντιπροσωπεύει ένα "καλό" σημείο το οποίο είναι αρκετά κατάλληλο. Τα καταλληλότερα άτομα του πληθυσμού επιλέγονται βάση του κριτηρίου και κατά αναλογία με τους όρους της βιολογίας, τα άτομα αυτά γίνονται οι "γονείς" και μέσω των στοχαστικών διαδικασιών δημιουργούνται νέες γενιές ατόμων ώστε με αυτή την επαναληπτική διαδικασία ο συνολικός πληθυσμός να μετακινείται προς καλύτερες λύσεις με την προσδοκία να συγκλίνουμε στην βέλτιστη (ή σχεδόν βέλτιστή λύση) λύση [5] [6]. Βασικές κατηγορίες των Εξελικτικών Αλγορίθμων είναι γενετικοί Αλγόριθμοι (Genetic Algorithms, GA's) και οι Στρατηγικές Εξέλιξης (Evolution Strategies ES's). Στην συνέχεια της παρούσας εργασίας θα μελετήσουμε με

περισσότερη λεπτομέρεια τις Στρατηγικές εξέλιξης και πιο συγκεκριμένα τον αλγόριθμο CMA-ES (Covariance Matrix Adaptation – Evolution Strategy) και τον αλγόριθμο βελτιστοποίησης σμήνους σωματιδίου PSO (Particle Swarm Optimization).

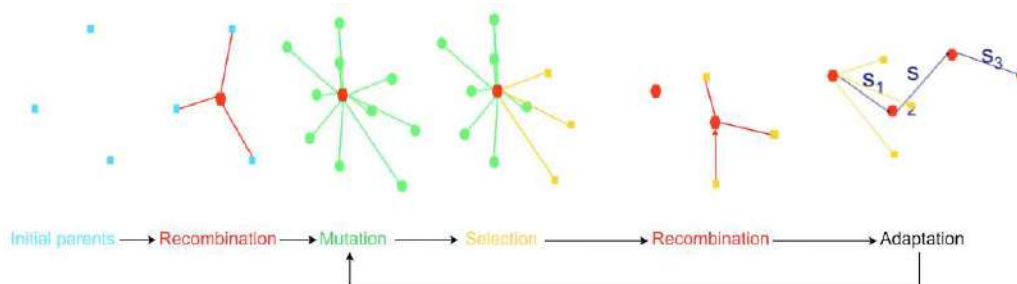


Εικόνα 3 [7]. Εξέλιξη των υποψήφιων λύσεων (μαύρες κουκίδες) στον χώρο αναζήτησης. Με κυανό χρώμα απεικονίζονται χαμηλές τιμές της αντικειμενικής συνάρτησης καταλληλότητας.

Στις ES, χρησιμοποιώντας τα σύμβολα που συναντάμε συνήθως για τον μνημονικό συμβολισμό των διαφόρων στρατηγικών, τα βήματα που ακολουθούνται συνήθως είναι τα εξής :

- Αρχικοποίηση ενός πληθυσμού λ απογόνων, συνήθως με τυχαίο τρόπο.
- Αξιολόγηση της καταλληλότητας των απογόνων.
- Επιλογή των " μ " καλύτερων γονέων, ανασυνδυασμός των γονέων σε έναν γονέα συνήθως και μετάλλαξη ώστε να δημιουργηθούν " λ " νέοι απόγονοι.
- Τέλος ελέγχεται το κριτήριο τερματισμού και σε περίπτωση που δεν ικανοποιείται η παραπάνω διαδικασία επαναλαμβάνεται .

Στην συνέχεια εξετάζουμε με μεγαλύτερη λεπτομέρεια τα βήματα αυτά.



Εικόνα 4 [7]. Συνοπτικά οι διαδικασίες στις ES, για μια επανάληψη.

2.2 Αρχικοποίηση του πληθυσμού

Η αρχικοποίηση συνήθως γίνεται με τυχαίο τρόπο αν και μερικές φορές αν υπάρχει εκ των προτέρων κάποια γνώση του προβλήματος και γνωρίζουμε σε έναν βαθμό ποια είναι μια "καλή" λύση, μπορεί αυτή η γνώση να ενσωματωθεί στην αρχικοποίηση με μερικούς τρόπους όπως για παράδειγμα θέτοντας το σημείο που αντιπροσωπεύει την λύση στον αρχικό πληθυσμό και τα υπόλοιπα σημεία τυχαία γύρω από αυτή την λύση [6]. Ακόμα, μπορούμε να θέσουμε πιο αυστηρά όρια στον χώρο αναζήτησης (search space) γύρω από μερικές διαστάσεις αξιοποιώντας την γνώση μας γύρω από το ποιος μπορεί να είναι "καλές" λύσεις. Αυτές οι τεχνικές χρησιμοποιούνται συνήθως σε εφαρμογές όπου είναι κρίσιμη η ταχύτητα με την οποία ο αλγόριθμος θα φτάσει σε κάποια λύση. Παρόλο αυτά έτσι μπορεί να εμποδίσουμε τον αλγόριθμο να φτάσει σε θεμελιώδεις διαφορετικές λύσεις εν συγκρίσει με αυτές που προτείνουμε βάση της εκ των προτέρων γνώσης μας [6]. Έτσι σε μερικούς αλγόριθμους προτιμάται η δημιουργία αρχικού πληθυσμού που θα παρέχει μια σχετικά ευρεία κάλυψη του χώρου αναζήτησης. Κάτι τέτοιο συνήθως συμβαίνει σε αλγόριθμους που εξαρτώνται σημαντικά από τον αρχικό πληθυσμό όπως ο αλγόριθμος Βελτιστοποίησης Σμήνους Σωματιδίων (Particle Swarm Optimization – PSO) [8]. Τέτοιοι αλγόριθμοι είναι πιθανό να παγιδευτούν γύρω από ένα τοπικό ελάχιστο ή μέγιστο (αν έχουμε ένα πρόβλημα ελαχιστοποίησης ή μεγιστοποίησης αντίστοιχα) κυρίως επειδή έπειτα από την αρχικοποίηση τους απουσιάζουν διαδικασίες τυχαιοποίησης και διότι ακολουθούν σε έναν βαθμό κάποιου είδους ντετερμινιστική προσέγγιση [9].

Τα παραπάνω είναι συνδεδεμένα με δύο έννοιες, την εξερεύνηση (exploration) και εκμετάλλευση (exploitation), που περιγράφουν μερικά από τα ποιοτικά χαρακτηριστικά αυτών των αλγορίθμων. Ένας σχετικά μεγάλος βαθμός εξερεύνησης θα καλύψει ένα μεγάλο

μέρος του χώρου αναζήτησης συνήθως μέσω της τυχαιότητας στην δημιουργία των ατόμων του πληθυσμού. Αν επιμείνει όμως για μεγάλο διάστημα, ο αλγόριθμος μπορεί να εξερευνήσει υπερβολικά πολλές περιοχές καθυστερώντας την σύγκλιση. Με την εκμετάλλευση εννοούμε τον βαθμό με τον οποίο χρησιμοποιούμε και εκμεταλλευόμαστε τα άτομα-σημεία που δίνουν ικανοποιητικές λύσεις για να φτάσουμε στην βέλτιστη λύση. Σημειώνουμε πως αυτές οι δύο έννοιες είναι ανταγωνιστικές, οπότε αν έχουμε υψηλό βαθμό εξερεύνησης αυτά τα άτομα που εμφανίζουν ικανοποιητικές λύσεις μπορεί να αντικατασταθούν δυσχεραίνοντας την εύρεση της βέλτιστης λύσης. Μια σχετικά μεγάλη εκμετάλλευση προσπαθεί να δώσει μια λύση όσο το δυνατόν κοντά στην βέλτιστη για τον διαθέσιμο πληθυσμό. Έτσι αν υπάρχει υψηλή εξάρτηση του αλγορίθμου μας από τον αρχικό πληθυσμό, υπάρχει κίνδυνος να παγιδευτεί ο αλγόριθμος σε ένα τοπικό ελάχιστο ή μέγιστο. Φαίνεται πως θα πρέπει να υπάρχει μια ισορροπία μεταξύ εξερεύνησης και εκμετάλλευσης ή ακόμα και μία μεταβολή αυτών των μεγεθών κατά την εξέλιξη των αλγορίθμων [9].

2.3. Αξιολόγηση και επιλογή.

Επιστρέφοντας στο πρόβλημα βελτιστοποίησης $F(\mathbf{y}) \rightarrow opt.$, μπορούμε να πούμε πως οι ES λειτουργούν πάνω σε πληθυσμούς \mathcal{B} που αποτελούνται από a άτομα όπως περιγράψαμε προηγουμένως. Ένα άτομο του πληθυσμού a_k που αντιπροσωπεύει μια υποψήφια λύση αποτελείται από ένα σύνολο που περιέχει το διάνυσμα των αντικειμενικών παραμέτρων \mathbf{y}_k (σημείο στον χώρο αναζήτησης) και την τιμή της αντικειμενικής συνάρτησης για το σημείο αυτό, $F(\mathbf{y}_k)$ [10]. Έτσι η αντικειμενική συνάρτηση αντιστοιχεί στα άτομα του πληθυσμού ή αλλιώς στις υποψήφιες λύσεις, μια τιμή καταλληλότητας ώστε να αξιολογήσουμε και να κατατάξουμε τις πιθανές λύσεις για να επιλέξουμε ανάλογα με την υλοποίηση του αλγορίθμου έναν αριθμό από “μ” (γονείς) καλύτερων λύσεων οι οποίοι θα παράξουν τους νέους απογόνους για να οδηγηθεί ο αλγόριθμος στην βέλτιστη λύση. Η αντικειμενική συνάρτηση ορίζει τον αντικειμενικό χώρο (objective space) ο οποίος είναι το σύνολο όλων των πιθανών τιμών καταλληλότητας της αντικειμενικής συνάρτησης. Αν έχουμε μια μόνο αντικειμενική συνάρτηση στο πρόβλημα της βελτιστοποίησης τότε συνήθως ο χώρος αυτός είναι ένα υποσύνολο του \mathbb{R} . Ενώ αν έχουμε προβλήματα πολλαπλών αντικειμενικών συναρτήσεων (πολλαπλών στόχων ή Multiobjective) τότε είναι υποσύνολο του \mathbb{R}^m όπου m ο αριθμός των στόχων ή των αντικειμενικών συναρτήσεων.

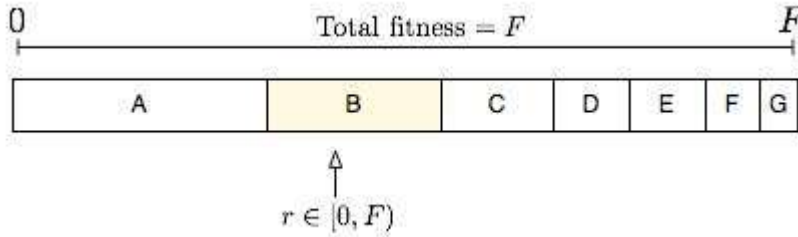
Συνήθως σε πολλές στρατηγικές εξέλιξης το σύνολο a_k κάθε ατόμου, περιέχει επιπλέον μερικές παραμέτρους που ονομάζονται ενδογενείς, έστω \mathbf{s}_k ώστε εν τέλη να έχουμε $a_k := (\mathbf{y}_k, \mathbf{s}_k, F(\mathbf{y}))$ [6]. Οι ενδογενείς παράμετροι είναι παράμετροι της στρατηγικής εξέλιξης που την ρυθμίζουν κατά κάποιο τρόπο και βαθμό και για αυτό καλούνται και στρατηγικές παράμετροι. Αυτές θα εξεταστούν παρακάτω με περισσότερη λεπτομέρεια. Πρέπει να σημειωθεί πως αυτές οι παράμετροι δεν επηρεάζουν το υπολογισμό της αντικειμενικής συνάρτησης και την καταλληλότητα του υποψηφίου, όμως μπορούν να κληροδοτηθούν στους απογόνους ανάλογα από τους με την καταλληλότητα των γονέων που επιλέγονται. Συμπληρωματικά αναφέρουμε πως εφόσον ορίζονται ενδογενείς παράμετροι, δημιουργείται ένας διαχωρισμός σε ενδογενείς και εξωγενείς παραμέτρους. Οι εξωγενείς δεν εντοπίζονται στο σύνολο με τις παραμέτρους του υποψηφίου και μένουν σταθερές κατά την εξέλιξη του αλγορίθμου. Τέτοιες είναι για παράδειγμα μπορεί να είναι ο αριθμός των γονέων και τον απογόνων.

Η διαδικασία της επιλογής είναι μείζονος σημασίας για να οδηγηθεί ένας αλγόριθμος σε καλύτερες λύσεις καθώς αφαιρεί τις υποψήφιες λύσεις με συγκριτικά “κακή” τιμή

συνάρτησης καταλληλότητας και επιλέγει μερικές καλές λύσεις για να γίνουν οι γονείς που θα παράξουν νέους υποψήφιους. Αυτή η διαδικασία καθορίζει τον τρόπο με τον οποίο θα ανανεώνεται ο πληθυσμός από γενιά σε γενιά. Υπάρχουν πολλές κατηγοριοποιήσεις βάση του τρόπου επιλογής των ατόμων που θα αποτελέσουν τους “γονείς” αλλά και κατηγορίες αναφορικά με την αναλογία επιλογής γονέων μέσα από “μ” υποψήφιους, ή ακόμα και σχετικά τον χειρισμό των ατόμων που δεν επιλέγονται ως γονείς.

Έτσι έχει υιοθετηθεί ο μνημονικός κανόνας $(\mu/\rho^+, \lambda)$ για την διαφοροποίηση του τρόπου πραγμάτωσης των ES. Ο αριθμός του πληθυσμού των γονέων συμβολίζεται με “μ”, από τους οποίους επιλέγονται “ρ” για να συμμετάσχουν στις διαδικασίες που θα περιγράψουμε (όπως ο ανασυνδυασμός) με $\rho \leq \lambda$. Με “λ” συμβολίζουμε τον αριθμό των απογόνων για κάθε επανάληψη. Τα σύμβολα “+” και “,” δεικνύουν τον χειρισμό που γίνεται στους γονείς μετά την επιλογή τους. Αν έχουμε το σύμβολο “+” σημαίνει πως η “ηλικία” των γονέων δεν λαμβάνεται υπόψιν, δηλαδή παραμένουν στον αλγόριθμο οι παλαιοί γονείς κάθε γενιάς. Οπότε για να επιλέξουμε τους νέους γονείς διαλέγουμε τους καλύτερους από τον πληθυσμό $\mu + \lambda$, που αποτελείται από τους παλιούς γονείς και τον τρέχων πληθυσμό, εξού και ο συμβολισμός [11]. Αυτή η προσέγγιση ονομάζεται ελιτιστική ή k-ελιτιστική αν k γονείς αποθηκεύονται ως η ελίτ [6]. Εν αντιθέσει σε έναν $(\mu/\rho, \lambda)$ αλγόριθμο όλοι οι γονείς αφαιρούνται (πεθαίνουν) μετά την επιλογή και επιβιώνει μόνο ένα σημείο που προκύπτει από τον ανασυνδυασμό τους. Αυτό το σημείο θα υποστεί μετάλλαξη και θα δημιουργήσει λ απογόνους. Έτσι οι νέοι γονείς επιλέγονται πάντα από τον τρέχων πληθυσμό λ. Συμπεραίνουμε πως στους (μ, λ) αλγόριθμους θα πρέπει $\lambda \geq \mu$ και πως στους $(\mu + \lambda)$ είναι δυνατή η τιμή $\lambda = 1$. Επιπρόσθετα μπορούμε να βρούμε επιπλέον δείκτες στην παράμετρο ρ που δηλώνουν το είδος του ανασυνδυασμού που θα εξετάσουμε στην συνέχεια. Για παράδειγμα ο αλγόριθμος CMA-ES που θα μελετήσουμε είναι ένας $(\mu/\mu_w, \lambda)$ όπου w δηλώνει πως ανασυνδυασμός είναι σταθμισμένος.

Οι συγκεκριμένες τεχνικές εφαρμογής της επιλογής είναι διάφορες. Αναφέρουμε ενδεικτικά μερικές. Στην επιλογή με διαγωνισμό κάθε ένας από τους υποψήφιους λαμβάνει μια πιθανότητα επιλογής όπου οι υποψήφιοι με καλύτερη τιμή καταλληλότητας λαμβάνουν μεγαλύτερη πιθανότητα. Αυτό γίνεται συνήθως με την σχέση $p(1 - p)^{j-1}$ όπου j η θέση στην κατάταξη και p η πιθανότητα επιλογής που την ορίζει ο ερευνητής. Ορίζεται το μέγεθος τους διαγωνισμού, έστω t, και στην συνέχεια t υποψήφιοι να επιλέγονται τυχαία από τον πληθυσμό βάση της πιθανότητας που έχουμε ορίσει, ώστε να διαγωνιστούν μεταξύ τους με τον καλύτερο να λαμβάνει μια από τις μ (ή ρ) διαθέσιμες θέσεις [6]. Αυτή η διαδικασία επαναλαμβάνεται μέχρι να συμπληρωθούν τα άτομα που θα αποτελέσουν τους γονείς. Στην διαδικασία ρουλέτας ή επιλογής κατ’ αναλογία, κάθε υποψήφιος λαμβάνει μια πιθανότητα επιβίωσης διαιρώντας την τιμή καταλληλότητας του με το άθροισμα των τιμών των άλλων υποψηφίων με την λογική πως η πιθανότητα επιβίωσης του εξαρτάται από το πόσο μεγαλύτερη ικανότητα έχει σχετικά με τους υπόλοιπους [6]. Επομένως η πιθανότητα επιβίωσης όλων των υποψηφίων αθροίζει στην μονάδα και στον κάθε υποψήφιο αντιστοιχείται ένα διάστημα μέσα στο [0,1] ίσο με την πιθανότητα επιβίωσης του. Ένας υποψήφιος επιλέγεται αν ένας τυχαίος αριθμός στο διάστημα [0,1] εντοπίζεται μέσα στο διάστημα που έχει αντιστοιχηθεί στον υποψήφιο [6].



Εικόνα 5 [12]. Επιλογή κατ' αναλογία.

Αυτές οι διαδικασίες περιέχουν κάποια στοχαστικότητα. Στους ES αλγόριθμους συναντάμε πιο συχνά ντετερμινιστικές τεχνικές. Ο πληθυσμός κατατάσσεται βάση των τιμών της συνάρτησης καταλληλότητας και επιλέγεται ένας αριθμός καλύτερων υποψηφίων που έχουμε ορίσει για να γίνουν οι γονείς. Για παράδειγμα σε έναν $(\mu/\mu, \lambda)$ αλγόριθμο, κατατάσσουμε τους λ υποψήφιους της τρέχουσας γενιάς βάση της καταλληλότητας και επιλέγουμε του μ καλύτερους.

Μια παράμετρος που λαμβάνεται υπόψιν είναι η πίεση της επιλογής (selection pressure) λ/μ που καθορίζει στην ουσία την πιθανότητα επιβίωσης ενός ατόμου σε μία γενιά-επανάληψη. Γενικά μια χαμηλή πίεση επιλογής (μικρός λόγος λ/μ) μπορεί να οδηγήσει σε μια πολύ αργή σύγκλιση ενώ μια μεγάλη πίεση (μεγάλος λόγος λ/μ) συνήθως οδηγεί τον αλγόριθμο σε σύγκλιση του αλγορίθμου σε μια μικρή περιοχή του χώρου αναζήτησης και πιθανή πρόωμη σε σύγκλιση σε λύση που δεν είναι η βέλτιστη [6]. Στις (μ, λ) στρατηγικές εξέλιξης είναι απαραίτητο να υπάρχει πίεση επιλογής και άρα $\mu < \lambda$. Αν δεν υπάρχει πίεση επιλογής, δηλαδή αν $\mu = \lambda$ με όλο τον πληθυσμό να επιλέγεται για γονείς τότε ο αλγόριθμος εκτελεί έναν τυχαίο περίπατο στον χώρο αναζήτησης [10].

2.4. Ανασυνδυασμός

Αφού έχουν επιλεγεί οι γονείς, επόμενη διαδικασία είναι ο ανασυνδυασμός και στην συνέχεια η μετάλλαξη. Ο ανασυνδυασμός συνδυάζει πληροφορίες από μ ή ρ γονείς που επιλέχθηκαν ώστε να δημιουργηθεί ένας απόγονος, έστω r , ο οποίος μετέπειτα θα υποστεί μετάλλαξη. Ξεχωρίζουμε τρεις κατηγορίες ανασυνδυασμών, τον διακριτό ή κυρίαρχο (discrete or dominant), τον ενδιάμεσο και τον σταθμισμένο, με τον τελευταίο να είναι γενίκευση του δεύτερου. Στον διακριτό αν έχουμε να κάνουμε με ένα πρόβλημα βελτιστοποίησης διάστασης N , και έχουμε ρ διανύσματα γονέων τα οποία έχουν διάσταση $k = N$, άρα και k συνιστώσες, τότε ο διακριτός ανασυνδυασμός δημιουργεί έναν απόγονο r όπου κάθε συνιστώσα k του απογόνου λαμβάνεται από την k συνιστώσα από κάποιον από τους ρ γονείς που επιλέγεται τυχαία για την ανάθεση της τιμής αυτής [11], [6]. Η παρακάτω σχέση περιγράφει ακριβώς αυτή την διαδικασία.

$$r = \sum_{i=1}^N (e_i^T y_{m_i}) e_i, \quad m_i = \text{Random}\{1, \dots, \rho\}$$

Ο ενδιάμεσος λαμβάνει όλους τους γονείς υπόψιν και υπολογίζει την μέση τιμή κάθε διάστασης (που αντιπροσωπεύει την μεταβλητή σχεδιασμού) ανάμεσα σε όλους τους γονείς και μεταθέτει την τιμή αυτή σε κάθε αντίστοιχη διάσταση του απογόνου.

$$r = \frac{1}{\rho} \sum_{i=1}^{\rho} y_i$$

Ο σταθμισμένος ανασυνδυασμός είναι γενίκευση του προηγούμενου ανασυνδυασμού όπου έχουμε τον σταθμισμένο μέσο όλων των ρ γονέων και συνήθως $\mu = \rho$. Οι τιμές των συντελεστών βαρύτητας εξαρτώνται από την κατάταξη των ατόμων βάση της καταλληλότητας [11].

$$\mathbf{r} = \frac{1}{\sum_{i=1}^{\rho} w_i} \sum_{i=1}^{\rho} w_i \mathbf{y}_i$$

2.5. Μετάλλαξη

Η μετάλλαξη εισάγει μικρές διαφοροποιήσεις, κάποια μεταβλητότητα, προσθέτοντας μια διαταραχή (perturbation) στο αποτέλεσμα του ανασυνδυασμού, τον απόγονο. Αυτή η διαταραχή επιλέγεται από μια πολυμεταβλητή κανονική κατανομή με μέσο το $\mathbf{0}$, που ανάλογα με τον αλγόριθμο μπορεί να είναι η $N(\mathbf{0}, \mathbf{I})$, η $N(\mathbf{0}, \text{diag}(\delta^2))$ ή τέλος η $N(\mathbf{0}, \mathbf{C})$ εκμεταλλευόμενοι το πλήρες μητρώο συνδιακύμανσης $\mathbf{C} \in \mathbb{R}^{n \times n}$. Ισχύει πως $\mathbf{x} + N(\mathbf{0}, \mathbf{C}) \sim N(\mathbf{x}, \mathbf{C})$, πράγμα που σημαίνει πως αναμενόμενη τιμή των απόγονων καθορίζεται από το \mathbf{x} [11], το αποτέλεσμα του ανασυνδυασμού των γονέων στην περίπτωση μας.

Παρακάτω παρουσιάζεται ένας ψευδοκώδικας για τις στρατηγικές εξέλιξης όπου N η διάσταση του προβλήματος, λ ο αριθμός απογόνων, μ ο αριθμός των γονέων, ρ ο αριθμός των γονέων που επιλέγονται για ανασυνδυασμό, P ο πληθυσμός που, \mathbf{y}_k το διάνυσμα σημείου k του πληθυσμού, s_k το σύνολο των ενδογενών παραμέτρων για το σημείο k και $F(\mathbf{y}_k)$ η τιμή της αντικειμενικής συνάρτησης για το σημείο αυτό.

Δεδομένα N, λ, ρ, μ

Αρχικοποίηση $P = \{\mathbf{y}_k, s_k, F(\mathbf{y}_k) \mid 1 \leq k \leq \mu\}$

Μέχρις ότου! (συνθήκη τερματισμού)

Για $k \in \{1, \dots, \lambda\}$

$(\mathbf{y}_k, s_k) = \text{ανασυνδυασμός}(\text{επιλογή γονέων}(\rho, P))$

$s_k \leftarrow \text{μετάλλαξη}_s(s_k)$

$\mathbf{y}_k \leftarrow \text{μετάλλαξη}_y(s_k, \mathbf{y}_k) \in \mathbb{R}^N$

$P \leftarrow P \cup \{(\mathbf{y}_k, s_k, F(\mathbf{y}_k)) \mid 1 \leq k \leq \mu\}$

$P \leftarrow \text{επιλογή βάση ηλικίας}(P) \% \text{ ταυτότητα αν έχουμε " + " στρατηγική}$

$P \leftarrow \text{επιλογή } \mu \text{ καλύτερων}(\mu, P)$

2.6. Ανασυνδυασμός και Επιδιόρθωση

Ένα ερώτημα που μπορεί να τεθεί είναι γιατί οι τελεστές ανασυνδυασμού δουλεύουν ευεργετικά; Ο ανασυνδυασμός από μόνος του συνήθως δεν έχει κάποια θετική επίδραση. Είναι χρήσιμος σε συνδυασμό με την μετάλλαξη με μια σχετικά μεγάλη δύναμη μετάλλαξης (mutation strength) και μια πίεση επιλογής [2]. Οπότε είναι ανάγκη να εξετάσουμε τον ανασυνδυασμό μαζί με την διαδικασία-τον τελεστή της μετάλλαξης.

Η μετάλλαξη όπως την περιγράψαμε προηγουμένως πραγματοποιεί μια δειγματοληψία. Έτσι πρέπει να επιλέξουμε την κατανομή δειγματοληψίας. Έχοντας τον χώρο αναζήτησης αν θεωρήσουμε ως την τυπική απόκλιση "σ" ως μόνη τη μόνη ενδογενής γνωστή παράμετρο τότε η κατανομή που δίνει την μέγιστη εντροπία είναι η πολυμεταβλητή κανονική κατανομή.

Σύμφωνα με την Αρχή της Μεγιστοποίησης της Εντροπίας η πιο αμερόληπτη κατανομή είναι αυτή που μεγιστοποιεί την εντροπία ικανοποιώντας παράλληλα τους δεδομένους περιορισμούς στους οποίους υπόκειται [13]. Έτσι δεν εισάγουμε μεροληψία, δεν υποθέτουμε δηλαδή κάποια γνώση που δεν έχουμε που θα μπορούσε να είναι λανθασμένη. Η αμεροληψία γενικά είναι μια από τις αρχές σχεδιασμού των Στρατηγικών εξέλιξης. Η μετάλλαξη έχει ως στόχο να εισάγει νέα αμερόληπτη πληροφορία. Ο τελεστής της επιλογής από την άλλη πλευρά εισάγει μια μεροληπτική τάση στην πληροφορία μετακινώντας τον πληθυσμό προς καλύτερες λύσεις με καλύτερες τιμές στην αντικειμενική συνάρτηση [11].

Οι γενετικοί αλγόριθμοι (GA) που ανήκουν στην ίδια οικογένειας με τους ES και αναπτύχθηκαν νωρίτερα από τους τελευταίους είχαν ως τελεστή ανασυνδυασμού την διασταύρωση (crossover). Η διασταύρωση στους GA έχει την ίδια χρήση με τον ανασυνδυασμός στις ES, απλά η δομή των δεδομένων στους GA αντιπροσωπεύεται με άλλον τρόπο από ότι στους ES. Για παράδειγμα, συνοπτικά, οι GA πρωτοεμφανίστηκαν δυαδικές ακολουθίες bit και συμβολοσειρές που αντιπροσωπεύουν το γονιδίωμα στο οποίο εφαρμόζονται οι τελεστές της διασταύρωσης και της μετάλλαξης. Για να ερμηνευθεί η ωφέλεια από τον τελεστή της διασταύρωσης υιοθετήθηκε η βασική θεώρηση των Building Blocks(BBH) στους GA η οποία εξηγεί πως ο αλγόριθμος αναγνωρίζει διαφορετικά “σχήματα” (Building Blocks) υψηλής καταλληλότητας τα οποία συνδυάζονται στο γονιδίωμα για να δημιουργηθεί “καλύτερος” απόγονος. Εν ολίγης βάση της θεώρησης αυτής, τα διαφορετικά “καλά χαρακτηριστικά” από διαφορετικούς γονείς μεταδίδονται στους απογόνους μέσω της διασταύρωσης [10]. Αν και διαισθητικά ήταν θελκτική η εξήγηση αυτή εντούτοις ήταν δύσκολο να υποστηριχθεί αυτή η υπόθεση. Με την ανάπτυξη των στρατηγικών εξέλιξης εμφανίστηκε μια άλλη υπόθεση, αυτή τη γενετικής επιδιόρθωσης (Genetic Repair - GR) με μία αντιδιαμετρική υπόθεση σε σύγκριση με αυτή των BBH, υποστηρίζοντας πως δεν είναι τα διαφορετικά επιθυμητά χαρακτηριστικά διαφορετικών γονέων που μεταδίδονται στους απογόνους, αλλά τα κοινά χαρακτηριστικά [10]. Δηλαδή ο ανασυνδυασμός εξάγει τις ομοιότητες των γονέων. Είναι εύλογο να υποθέσουμε πως εφόσον οι γονείς επιλέχθηκαν ως καταλληλότεροι τότε τα “κοινά χαρακτηριστικά” που είναι παρόμοια μεταξύ τους στις τιμές στις συνιστώσες των γονέων είναι πιθανότερο να μας οδηγήσουν σε καλύτερες λύσεις όντας ωφέλιμα σχετικά με την αντικειμενική συνάρτηση (καταλληλότητας). Έτσι ένας τελεστής που εισάγει μεταβλητότητα το καλύτερο που μπορεί να κάνει είναι να διατηρήσει τις “ωφέλιμες” αυτές συνιστώσες και τις υπόλοιπες συνιστώσες οι οποίες μπορεί να είναι επιβλαβείς προς την βελτίωση να τις μεταλλάξει [10].

2.7. Το όφελος του ανασυνδυασμού – Γενετική επιδιόρθωση

Εξετάζοντας την χρησιμότητα του ανασυνδυασμού υπό το πρίσμα της γενετικής επιδιόρθωσης ακολουθούμε την ανάλυση των Beyer και Schwefel [10] για έναν (μ/μ) ενδιάμεσο ανασυνδυασμό και ισότροπες μεταλλάξεις. Έτσι έχουμε το κεντροειδές του πληθυσμού :

$$\langle \mathbf{y} \rangle = \frac{1}{\mu} \sum_{m=1}^{\mu} \mathbf{y}_{m;\lambda}$$

Για να παράξουμε τα λ άτομα της νέας γενιάς εφαρμόζουμε λ μεταλλάξεις $\mathbf{z} = \sigma(\mathcal{N}_l(0,1), \dots, \mathcal{N}_l(0,1))$ σε κάθε συνιστώσα του κεντροειδούς. Το νέο κεντροειδές, $\langle \mathbf{y} \rangle^{g+1}$, λαμβάνεται προσθέτοντας στο παλιό κεντροειδές το μέσο διάνυσμα με τις μ καλύτερες μεταλλάξεις .

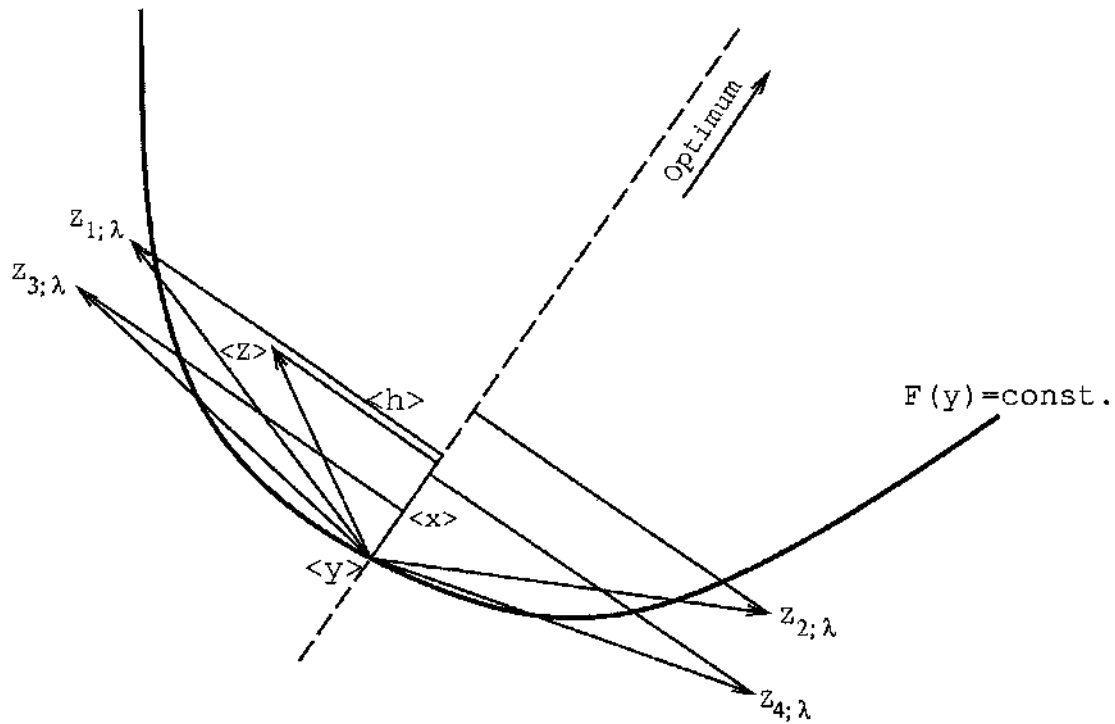
$$\langle \mathbf{y} \rangle^{g+1} = \langle \mathbf{y} \rangle + \frac{1}{\mu} \sum_{m=1}^{\mu} \mathbf{z}_{m;\lambda} \quad (4)$$

Αν αναλύσουμε κάθε μετάλλαξη \mathbf{z} σε μία συνιστώσα προς την επιθυμητή κατεύθυνση που βελτιώνει την καταλληλότητα (την συνάρτηση καταλληλότητας) $x\mathbf{e}_{opt}$ και σε κάθετη σε αυτό που αντιπροσωπεύει τις συνιστώσες \mathbf{h} που απομακρύνουν τους απογόνους από την βέλτιστη λύση τότε

$$\mathbf{z} = x\mathbf{e}_{opt} + \mathbf{h}, \mu\epsilon \mathbf{h} \cdot \mathbf{e}_{opt} = 0 \quad (5)$$

Αντικαθιστώντας της σχέση (5) στην (4) έχουμε :

$$\langle \mathbf{y} \rangle^{g+1} = \langle \mathbf{y} \rangle + \langle \mathbf{z} \rangle = \langle \mathbf{y} \rangle + \mathbf{e}_{opt} \frac{1}{\mu} \sum_{m=1}^{\mu} x_{m;\lambda} + \frac{1}{\mu} \sum_{m=1}^{\mu} \mathbf{h}_{m;\lambda} = \langle \mathbf{y} \rangle + \langle x \rangle \mathbf{e}_{opt} + \langle \mathbf{h} \rangle$$



Εικόνα 6 [10]. Ανάλυση των μεταλλάξεων σε μια διεύθυνση προς το βέλτιστο σημείο και μία κάθετη σε αυτή

Λόγο της διαδικασίας της επιλογής οι x συνιστώσες θα έχουν κάποια τάση προς την κατεύθυνση βελτίωσης της τιμής της αντικειμενικής συνάρτησης. Οι \mathbf{h} συνιστώσες όμως είναι σχεδόν ισότροπα κατανεμημένες ορθογώνια στο \mathbf{e}_{opt} . Αυτό μπορεί να φανεί στο σχήμα σχηματίζοντας τα τέσσερα καλύτερα \mathbf{z} διανύσματα (σχεδόν συμμετρικά) . Παίρνοντας τον μέσο για να καθορίσουμε το \mathbf{z} , μπορούμε να συμπεράνουμε ότι λίγο πολύ οι συνιστώσες που συνεισφέρουν στην βελτίωση διατηρούνται. Η τιμή τους x είναι στα όρια $\min \{x_{1;\lambda}, x_{2;\lambda}, \dots, x_{\mu;\lambda}\} \leq \langle x \rangle \leq \max \{x_{1;\lambda}, x_{2;\lambda}, \dots, x_{\mu;\lambda}\}$.Στις "επιβλαβής" όμως συνιστώσες του \mathbf{h} παίρνοντας την μέση τιμή μειώνεται το μήκος των βλαβερών συνιστωσών. Αυτό συμβαίνει λόγω της ισοτροπίας και της ανεξαρτησίας των \mathbf{h} διανυσμάτων. Με άλλα λόγια τα \mathbf{h} είναι σχεδόν ασυσχέτιστα.

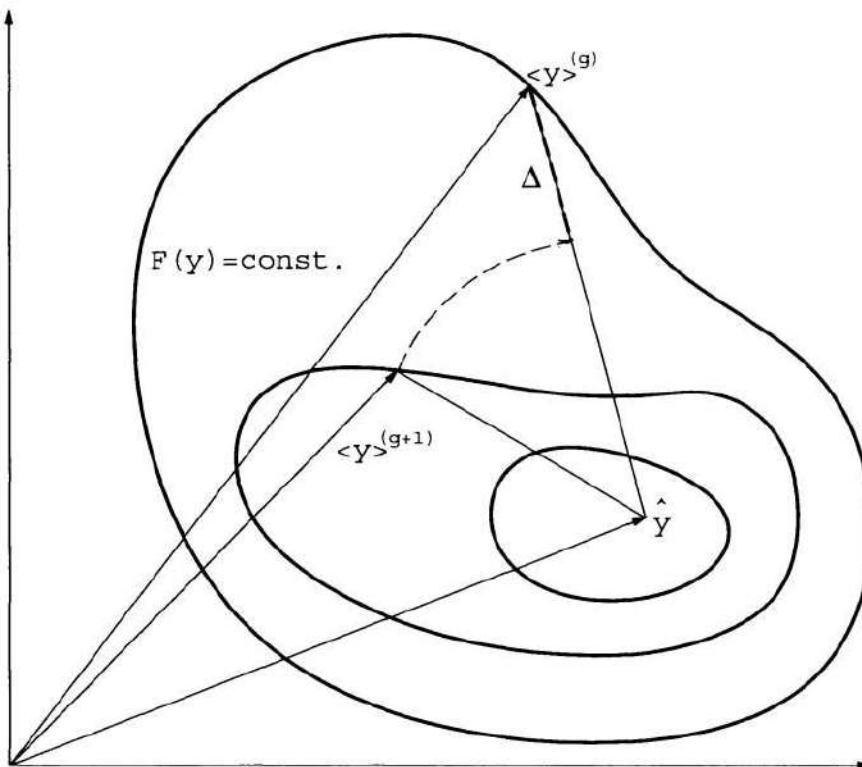
2.8. Βαθμός Προόδου και Επιδιόρθωση

Για να ποσοτικοποιήσουν την βελτίωση που προκαλεί ο ανασυνδυασμός ο H.G. Beyer [2], όρισε ένα βαθμό προόδου (progress rate) και σύγκρισε την στρατηγική (μ, λ) στην οποία δεν υπάρχει ανασυνδυασμός αλλά μόνο επιλογή και μετάλλαξη, με την $(\mu/\mu_l, \lambda)$ σε αντικειμενική συνάρτηση σφαίρας σε N διαστάσεις.

Συνοπτικά παρουσιάζουμε τα σημαντικότερα σημεία της ανάλυσης ώστε να είναι φανερή η προέλευση της τελικής σχέσης του βαθμού προόδου, η οποία εμφανίζεται σε πολλές δημοσιεύσεις γύρω από τις ES. Αρχικά ορίστηκε ένα μετρό (βαθμός προόδου) φ που μετρά την απόσταση στον χώρο αναζήτησης. Συγκεκριμένα μετρά την αναμενόμενη τιμή της αλλαγής της απόστασης των υποψήφιων λύσεων στον πληθυσμό από γενιά g σε $g + 1$ σε σχέση με ένα σταθερό σημείο αναφορά $\hat{\mathbf{y}}$. Ο στόχος της ανάλυσης είναι να συγκριθεί ο βαθμός προόδου $\varphi_{\mu, \lambda}$ για την (μ, λ) μέθοδο όπου δεν έχουμε ανασυνδυασμό με τον βαθμό προόδου $\varphi_{\mu/\mu_l, \lambda}$ όπου υπάρχει ανασυνδυασμός.

Αν $\langle \mathbf{y} \rangle_{\mu}^{(.)} = \frac{1}{\mu} \sum_{m=1}^{\mu} \mathbf{y}_m^{(.)}$ τελεστής για το κεντροειδής του πληθυσμού των γονέων για κάθε γενιά τότε

$$\varphi_{\mu/\rho, \lambda} = E \left\{ \left\| \hat{\mathbf{y}} - \langle \mathbf{y} \rangle^{(g)} \right\| - \left\| \hat{\mathbf{y}} - \langle \mathbf{y} \rangle^{(g+1)} \right\| \right\} \quad (6).$$



Εικόνα 7 [2]. Αποστάσεις $\hat{\mathbf{y}}$ από τα κεντροειδής των (g) και $(g+1)$

Στην $(\mu/\mu_l, \lambda)$ έχουμε ανασυνδυασμό στον οποίο συμμετέχουν όλοι οι μ γονείς που επιλέχθηκαν με ίση βαρύτητα. Αν $l = (1, \dots, \lambda)$, $m = (1, \dots, \mu)$ και $m; \lambda$ σημαίνει πως επιλέγεται ο m από λ , τότε οι απόγονοι που αποτελούν τον πληθυσμό της γενιάς g είναι οι

$\tilde{\mathbf{y}}_l^{(g)}$ και η διαδικασία της επιλογής αφήνει $\mathbf{y}_m^{(g)} := \tilde{\mathbf{y}}_{m;\lambda}^{(g)}$ γονείς για ανασυνδιασμό. Οπότε το κεντροειδής του πληθυσμού υπολογίζεται ως

$$\langle \mathbf{y} \rangle^{(g)} = \frac{1}{\mu} \sum_{m=1}^{\mu} \mathbf{y}_m^{(g)} \quad (7)$$

Οι απόγονοι $\tilde{\mathbf{y}}_l$ παράγονται προσθέτοντας στο κεντροειδής $\langle \mathbf{y} \rangle$ διανύσματα μετάλλαξης $\mathbf{z} = \sigma(\mathcal{N}(0,1), \dots, \mathcal{N}(0,1))^T$. Οπότε οι απόγονοι που θα αποτελέσουν την γενιά $g + 1$ υπολογίζονται από το κεντροειδή της προηγούμενης γενιάς προσθέτοντας το διάνυσμα μετάλλαξης.

$$\tilde{\mathbf{y}}_l^{(g+1)} = \langle \mathbf{y} \rangle^{(g)} + \mathbf{z}_l = \frac{1}{\mu} \sum_{m=1}^{\mu} \tilde{\mathbf{y}}_{m;\lambda}^{(g)} + \mathbf{z}_l \quad (8)$$

Θέτοντας το σημείο $\hat{\mathbf{y}}$ στο βέλτιστο σημείο, στο κέντρο του συστήματος που για την σφαίρα είναι το βέλτιστο, μπορούμε να πούμε πως το κεντροειδής $\langle \mathbf{y} \rangle$ των γονέων είναι ταυτόχρονα το διάνυσμα της ακτίνας της σφαίρας :

$$\langle \mathbf{y} \rangle^{(g)} = \frac{1}{\mu} \sum_{m=1}^{\mu} \mathbf{y}_m^{(g)} =: \mathbf{R} \quad (9)$$

άρα $\|\langle \mathbf{y} \rangle^{(g)}\| = \|\mathbf{R}\|$. Αντίστοιχα για την γενιά $g + 1$ έχουμε :

$$\langle \mathbf{y} \rangle^{(g+1)} = \frac{1}{\mu} \sum_{m=1}^{\mu} \mathbf{y}_m^{(g+1)} = \frac{1}{\mu} \sum_{m=1}^{\mu} \tilde{\mathbf{y}}_{m;\lambda}^{(g+1)} = \tilde{\mathbf{R}} \quad (10)$$

Όπου $m; \lambda$ συμβολίζει την επιλογή των m καλύτερων γονέων.

Έτσι εισάγοντας την (9) και την (10) στην (6), έχουμε :

$$\varphi_{\mu/\rho,\lambda} = E \left\{ \left\| \langle \mathbf{y} \rangle^{(g)} \right\| - \left\| \langle \mathbf{y} \rangle^{(g+1)} \right\| \right\} = E \left\{ R - \left\| \langle \mathbf{y} \rangle^{(g+1)} \right\| \right\} \quad (11)$$

Στην (10) έχουμε τους απογόνους που αποτελούν τον πληθυσμό της $(g + 1)$ να ισούνται με $\tilde{\mathbf{y}}_l^{(g+1)} = \langle \mathbf{y} \rangle^{(g)} + \mathbf{z}_l$. Εφαρμόζοντας την επιλογή σε αυτούς έχουμε

$$\tilde{\mathbf{y}}_{m;\lambda}^{(g+1)} = \left(\langle \mathbf{y} \rangle^{(g)} + \mathbf{z}_l \right)_{m;\lambda} = \mathbf{R} + \mathbf{z}_{m;\lambda} \quad (12)$$

Δηλαδή ο m απόγονος παράγεται από την “ m -καλύτερη” μετάλλαξη $\mathbf{z}_{m;\lambda}$

Στη συνέχεια αναλύονται οι μεταλλάξεις σε μια συνιστώσα x προς την βέλτιστη λύση και \mathbf{h} κάθετη

$$\tilde{\mathbf{y}}_{m;\lambda} = R \mathbf{e}_R - x_{m;\lambda} \mathbf{e}_R + \mathbf{h}_{m;\lambda}, \text{ με } \mathbf{e}_R = \mathbf{R}/\|\mathbf{R}\| \text{ και } \mathbf{e}_R \cdot \mathbf{h}_{m;\lambda} = 0$$

όπου $x_{m;\lambda}$ είναι η συνιστώσα $z_{m;\lambda}$ στην \mathbf{R} κατεύθυνση και $\mathbf{h}_{m;\lambda}$. Επομένως έχουμε

$$\langle \mathbf{y} \rangle^{(g+1)} = (R - \langle x \rangle) \mathbf{e}_R + \langle \mathbf{h} \rangle \quad (13)$$

με

$$\langle x \rangle = \frac{1}{\mu} \sum_{m=1}^{\mu} x_{m;\lambda} \text{ και } \langle \mathbf{h} \rangle = \frac{1}{\mu} \sum_{m=1}^{\mu} \mathbf{h}_{m;\lambda}$$

Αντικαθιστώντας της (13) στην (11) :

$$\varphi_{\mu/\rho,\lambda} = E\{\Delta R\} = E\{R - \tilde{R}\} = E\left\{R - \sqrt{(R - \langle x \rangle)^2 + \langle \mathbf{h} \rangle^2}\right\} \quad (14)$$

Θεωρούμε μικρή διακύμανση των $\langle x \rangle$ και $\langle \mathbf{h} \rangle^2$ γύρω από τους μέσους τους $\overline{\langle x \rangle}$ και $\overline{\langle \mathbf{h} \rangle^2}$ και από την (14) έχουμε $\Delta R = R - \sqrt{(R - \langle x \rangle)^2 + \langle \mathbf{h} \rangle^2}$. Αναπτύσσοντας το ανάπτυγμα Taylor για την προηγούμενη σχέση γύρω από $\langle x \rangle = \overline{\langle x \rangle}$ και $\langle \mathbf{h} \rangle^2 = \overline{\langle \mathbf{h} \rangle^2}$ έχουμε

$$\begin{aligned} \Delta R(\langle x \rangle, \langle \mathbf{h} \rangle^2) &= \\ &= R - \sqrt{(R - \langle x \rangle)^2 + \langle \mathbf{h} \rangle^2} + \frac{R - \overline{\langle x \rangle}}{\sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}}} (\langle x \rangle - \overline{\langle x \rangle}) - \frac{\frac{1}{2}(\langle \mathbf{h} \rangle^2 - \overline{\langle \mathbf{h} \rangle^2})}{\sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}}} \\ &- \frac{1}{2} \frac{\overline{\langle \mathbf{h} \rangle^2}}{\left(\sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}}\right)^3} (\langle x \rangle - \overline{\langle x \rangle})^2 + \frac{1}{8} \frac{\left(\langle \mathbf{h} \rangle^2 - \overline{\langle \mathbf{h} \rangle^2}\right)^2}{\left(\sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}}\right)^3} - \frac{2}{4} \frac{R - \overline{\langle x \rangle}}{\left(\sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}}\right)^3} (\langle x \rangle - \\ &\overline{\langle x \rangle})(\langle \mathbf{h} \rangle^2 - \overline{\langle \mathbf{h} \rangle^2}) + \dots \end{aligned} \quad (15)$$

Παίρνοντας της αναμενόμενη τιμή της παραπάνω σχέσης ο δεύτερος και ο τρίτος από τους γραμμικούς όρους είναι ίσοι με το μηδέν καθώς $E(\langle x \rangle - \overline{\langle x \rangle}) = E[\langle x \rangle] - E[\overline{\langle x \rangle}] = \overline{\langle x \rangle} - \overline{\langle x \rangle} = 0$ και $E[(\langle \mathbf{h} \rangle^2 - \overline{\langle \mathbf{h} \rangle^2})] = E[\langle \mathbf{h} \rangle^2] - E[\overline{\langle \mathbf{h} \rangle^2}] = Var(\langle \mathbf{h} \rangle) + E[\langle \mathbf{h} \rangle^2] - E[Var(\langle \mathbf{h} \rangle) + E[\langle \mathbf{h} \rangle^2]] = Var(\langle \mathbf{h} \rangle) + E[\langle \mathbf{h} \rangle^2] - Var(\langle \mathbf{h} \rangle) - E[E[\langle \mathbf{h} \rangle^2]] = Var(\langle \mathbf{h} \rangle) + E[\langle \mathbf{h} \rangle^2] - Var(\langle \mathbf{h} \rangle) - E[\langle \mathbf{h} \rangle^2] = 0$

Και η (15) γίνεται

$$\begin{aligned} \overline{\Delta R(\langle x \rangle, \langle \mathbf{h} \rangle^2)} &= R - \sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}} - \frac{1}{2} \frac{\overline{\langle \mathbf{h} \rangle^2}}{\left(\sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}}\right)^3} (\langle x \rangle - \overline{\langle x \rangle})^2 + \\ &\frac{1}{8} \frac{\left(\langle \mathbf{h} \rangle^2 - \overline{\langle \mathbf{h} \rangle^2}\right)^2}{\left(\sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}}\right)^3} - \frac{2}{4} \frac{R - \overline{\langle x \rangle}}{\left(\sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}}\right)^3} (\langle x \rangle - \overline{\langle x \rangle})(\langle \mathbf{h} \rangle^2 - \overline{\langle \mathbf{h} \rangle^2}) \end{aligned} \quad (16)$$

Στην συνέχεια εφαρμόζουμε κανονικοποίηση $\varphi^* = \varphi \frac{N}{R}$ και $\sigma^* = \sigma \frac{N}{R}$ και κρατώντας τον πρώτο όρο και λαμβάνοντας υπόψιν την τάξη μεγέθους των άλλων όρων έχουμε

$$\varphi_{\mu/\mu,\lambda}^* = N \left(1 - \frac{1}{R} \sqrt{(R - \overline{\langle x \rangle})^2 + \overline{\langle \mathbf{h} \rangle^2}} + O\left(\frac{1}{\sqrt{N}}\right) \right) \quad (17)$$

Για τον υπολογισμό του βαθμού για την μέθοδο $\mu/\mu_l, \lambda$ και $\mu/\mu, \lambda$ προόδου το ζήτημα είναι να υπολογιστούν τα $\overline{\langle x \rangle}$ και $\overline{\langle \mathbf{h} \rangle^2}$ για την υλοποίηση αυτή. Μετά από ανάλυση ο H.G. Beyer [2], για την περίπτωση $\mu/\mu_l, \lambda$ καταλήγει στην σχέση

$$\varphi_{\mu/\mu,\lambda}^*(\sigma^*) = c_{\mu/\mu,\lambda} \sigma^* \frac{1 + \frac{\sigma^{*2}}{2\mu N}}{\sqrt{1 + \frac{\sigma^{*2}}{2N}} \sqrt{1 + \frac{\sigma^{*2}}{\mu N}}} - N \left(\sqrt{1 + \frac{\sigma^{*2}}{\mu N}} - 1 \right) + O\left(\frac{1}{\sqrt{N}}\right) \quad (18)$$

οπου $c_{\mu/\mu,\lambda}$ συντελεστής προόδου ο οποίος παίρνει τιμές σύμφωνα με τον παρακάτω πίνακα ανάλογα με τον αριθμό γονέων και απογόνων.

$\mu \backslash \lambda$	10	20	30	40	50	100	150	200	300
1	1.539	1.867	2.043	2.161	2.249	2.508	2.649	2.746	2.878
2	1.270	1.638	1.829	1.957	2.052	2.328	2.478	2.580	2.718
3	1.065	1.469	1.674	1.810	1.911	2.201	2.357	2.463	2.607
4	0.893	1.332	1.550	1.694	1.799	2.101	2.263	2.372	2.520
5	0.739	1.214	1.446	1.596	1.705	2.018	2.185	2.297	2.449
10	0.000	0.768	1.061	1.242	1.372	1.730	1.916	2.040	2.206
20	-	0.000	0.530	0.782	0.950	1.386	1.601	1.742	1.928
30	-	-	0.000	0.414	0.634	1.149	1.390	1.545	1.746
40	-	-	-	0.000	0.343	0.958	1.225	1.393	1.608
50	-	-	-	-	0.000	0.792	1.085	1.265	1.494
100	-	-	-	-	-	0.000	0.542	0.795	1.088

Πίνακας 1. Συντελεστές Προόδου για την $\mu/\mu,\lambda$ μέθοδο [2].

Για διάσταση $N \rightarrow \infty$ έχουμε

$$\varphi_{\mu/\mu,\lambda}^*(\sigma^*) = c_{\mu/\mu,\lambda} \sigma^* - \frac{\sigma^{*2}}{2\mu} \quad (19)$$

και για την περίπτωση της μεθόδου χωρίς ανασυνδυασμό ο H.G. Beyer [2] αποδεικνύει πως

$$\varphi_{\mu,\lambda}^*(\sigma^*) = c_{\mu,\lambda} \sigma^* - \frac{\sigma^{*2}}{\mu} \quad (20)$$

με συντελεστές προόδου αυτούς που παρουσιάζονται παρακάτω.

$\mu \backslash \lambda$	5	10	20	30	40	50	100	150	200	250
1	1.16	1.54	1.87	2.04	2.16	2.25	2.51	2.65	2.75	2.82
2	0.92	1.36	1.72	1.91	2.04	2.13	2.40	2.55	2.65	2.72
3	0.68	1.20	1.60	1.80	1.93	2.03	2.32	2.47	2.57	2.65
4	0.41	1.05	1.49	1.70	1.84	1.95	2.24	2.40	2.51	2.59
5	0.00	0.91	1.39	1.62	1.77	1.87	2.18	2.34	2.45	2.53
10	-	0.00	0.99	1.28	1.46	1.59	1.94	2.12	2.24	2.33
20	-	-	0.00	0.76	1.03	1.20	1.63	1.84	1.97	2.07
30	-	-	-	0.00	0.65	0.89	1.41	1.64	1.79	1.90
40	-	-	-	-	0.00	0.57	1.22	1.49	1.65	1.77
50	-	-	-	-	-	0.00	1.06	1.35	1.53	1.65
100	-	-	-	-	-	-	0.00	0.81	1.07	1.24

Πίνακας 2 [2]. Συντελεστές Προόδου για την $\mu,\mu/\lambda$ μέθοδο

Στις πάνω σχέσεις η διάσταση μπορεί να τείνει στο άπειρο, $N \rightarrow \infty$, όμως σε πεπερασμένες αλλά μεγάλες διαστάσεις μπορεί να χρησιμοποιηθεί ως μέτρο εκτίμησης στην συνάρτηση σφαίρας [11]. Συγκρίνοντας τις εξισώσεις προόδου για τις (μ, λ) και τις $(\mu/\mu_I, \lambda)$ βλέπουμε πως παρουσιάζουν εξαιρετική ομοιότητα, με τους συντελεστές όμως $c_{\mu/\mu_I, \lambda}$ της $(\mu/\mu_I, \lambda)$ να είναι καλύτεροι από αυτούς της (μ, λ) αλλά και ο αρνητικός όρος, αυτός που αφαιρεί απόδοση, να είναι μικρότερος κατά $1/\mu$ πράγμα που σημαίνει ότι μπορούμε να χρησιμοποιήσουμε μεγαλύτερα βήματα μετάλλαξης όταν έχουμε ανασυνδυασμό. Αν ερευνήσουμε την ανάλυση στην εξαγωγή της σχέσης του βαθμού προόδου στο βιβλίο του H.G. Beyer βρίσκουμε πως ο αρνητικός όρος σχετίζεται με τις \mathbf{h} συνιστώσες των μεταλλάξεων \mathbf{z} . Έτσι όπως περιγράψαμε προηγουμένως ο ανασυνδυασμός μειώνει το μήκος του $\langle \mathbf{h} \rangle^2$ στο $1/\mu$ του μήκους.

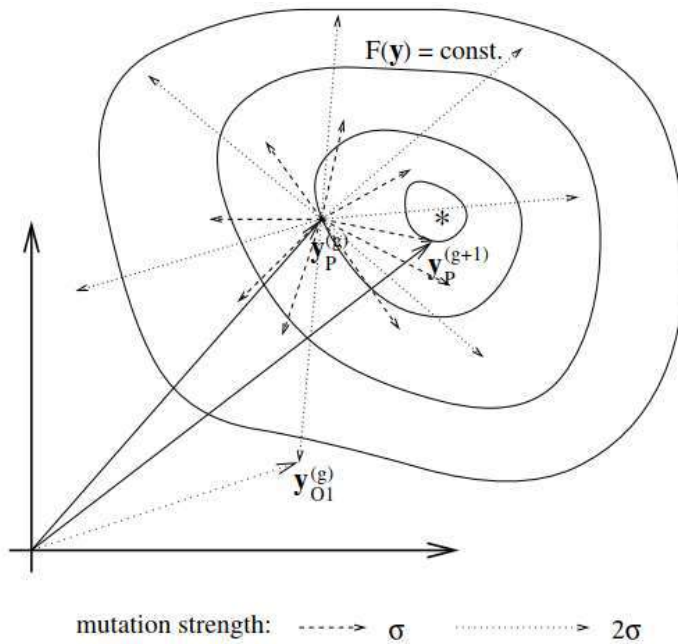
3. Αυτοπροσαρμογή των ενδογενών παραμέτρων

3.1. Κανόνας του $1/5^{th}$ και Αυτοπροσαρμογή.

Όπως έχουμε αναφέρει με την μετάλλαξη ο αλγόριθμος παράγει απογόνους μέσω μια πολυμεταβλητής κανονικής κατανομή με μέσο τον μοναδικό γονέα που έχει προκύψει από τον ανασυνδυασμό. Αν πρόκειται για την $N(\mathbf{0}, \text{diag}(\delta^2))$ κατανομή, μπορεί να επίσης να νοηθεί προσθέτοντας στον γονέα το διάνυσμα $\mathbf{z} = \sigma(\mathcal{N}_1(0,1), \dots, \mathcal{N}_1(0,1))$ όπως αναφέραμε προηγουμένως. Η διακύμανση σ^2 της κατανομής, στις στρατηγικές εξέλιξης ονομάζεται δύναμη μετάλλαξης (mutation strength) ή μέγεθος βήματος (step size) και κατέχει σημαντικό ρόλο. Η απόδοση των Στρατηγικών Εξέλιξης εξαρτάται σημαντικά από τη δύναμη μετάλλαξης. Η εξάρτηση αυτή από την δύναμη μετάλλαξης μπορεί να γίνει φανερή από όλα τα μέτρα που ορίζουν την πρόοδο του αλγορίθμου [2]. Ένα τέτοιο μέτρο εξετάσαμε προηγουμένως. Αυτή η παράμετρος είναι μια ενδογενής παράμετρος των αλγορίθμων Στρατηγικής Εξέλιξης. Αναφέραμε στην αρχή πως οι ενδογενείς παράμετροι είναι παράμετροι που ελέγχουν τον αλγόριθμο, όμως μπορούν να είναι και αυτές αντικείμενο της βελτιστοποίησης του αλγορίθμου. Επομένως στην εποχή όπου εξελίσσονταν οι ES, έγινε σύντομα φανερό στους ερευνητές πως θα πρέπει να ελέγξουν τις ενδογενείς μεταβλητές, όπως το μέγεθος βήματος, για να πετύχουν καλύτερη απόδοση.

Μια πρώτη προσπάθεια κατά την πρώιμη έρευνα γύρω από τις ES κατέληξε στον σχετικά απλό κανόνα του $1/5^{th}$ που αν και απλός είναι αρκετά διαφωτιστικός για την σημασία της ανάγκης προσαρμογής του μεγέθους του βήματος.

Θεωρούμε στρατηγική εξέλιξης (1+1) με μετάλλαξη μέσω ισότροπης κανονικής κατανομής, σταθερό μέγεθος βήματος σ , με στόχο την ελαχιστοποίηση αντικειμενικής συνάρτησης $F(\mathbf{y}) = F_{\text{sphere}}(\mathbf{y}) = \|\mathbf{y}\|^2$ που αντιπροσωπεύει σφαίρα σε N διαστάσεις. Έχουμε μόνο έναν γονέα σε κάθε γενιά (g) από τον οποίο παράγεται ο απόγονος από την μετάλλαξη. Η αναμενόμενη τιμή του μήκος της μετάλλαξης είναι ανάλογη του σ . Λόγο της συμμετρίας της μετάλλαξης καθώς $\sigma \rightarrow 0$, η πιθανότητα να βελτιωθεί το σημείο και να γίνει γονέας είναι $P_s \rightarrow 1/2$. Αν έχουμε $N = 2$, στον \mathbb{R}^2 απεικονίζοντας με ισοκαμπύλες τις τιμές $F(\mathbf{y}) = \text{σταθ}$ και το εσωτερικό της καμπύλης $F(\mathbf{y}_p)$ ως τομέας "επιτυχίας" με το βέλτιστο βρίσκεται στο σημείο με *, μπορούμε να καταλάβουμε διαισθητικά πως καθώς μικραίνει το σ περίπου οι μισές μεταλλάξεις της γενιά g προς $g + 1$ θα είναι εντός του τομέα επιτυχίας. Όμως ο βαθμός προόδου είναι ανάλογος του σ , έτσι καθώς $\sigma \rightarrow 0$ επίσης $\varphi \rightarrow 0$.



Εικόνα 8 [10]. Δύναμη μετάλλαξης σε στρατηγική (1+1)

Αυξάνοντας το σ όμως, με $\sigma \rightarrow \infty$ το να βρεθεί η μετάλλαξη εντός του τομέα που ορίζουμε, γίνεται εξαιρετικά απίθανο με $P_s \rightarrow 0$ άρα και $\varphi \rightarrow 0$. Μεταξύ αυτών των δύο ακραίων καταστάσεων θα υπάρχει ένα φάσμα τιμών για το σ που θα βελτιστοποιεί σχεδόν την απόδοση του βαθμού προόδου φ . Ο Rechenberg [10] σύγκρινε την απόδοση της στρατηγικής εξέλιξης για δύο τελείως διαφορετικές αντικειμενικές συναρτήσεις. Σε συνάρτηση σφαίρας και συνάρτηση διαδρόμου (corridor function) που είναι μια γραμμική συνάρτηση με $2(N - 1)$ ανισοτικούς περιορισμούς. Βρέθηκε πως οι ιδανικές τιμές για το σ που μεγιστοποιούν τον βαθμό απόδοσης αντιστοιχούν σε πιθανότητα επιτυχίας $P_s \approx 0.27 \approx 1/(3,7)$ και $P_s \approx 0.184 \approx 1/(5,39)$ αντίστοιχα. Αυτές οι δύο αντικειμενικές συναρτήσεις κατά έναν βαθμό αντιπροσωπεύουν μοντέλα από αντικειμενικές συναρτήσεις που συναντώνται σε προβλήματα βελτιστοποίησης. Έτσι ορίστηκε ο κανόνας του $1/5^{00}$ ο οποίος δηλώνει πως στις (1+1) ES για να πετύχουμε σχεδόν βέλτιστο βαθμό απόδοσης θα πρέπει να ρυθμίζουμε την δύναμη μετάλλαξης έτσι ώστε η πιθανότητα επιτυχία όπως πάνω να είναι περίπου ίση με $1/5$ [10].

Κινούμενοι στην κατεύθυνση προσαρμογής των ενδογενών παραμέτρων, στην συνέχεια οι ερευνητές υπέθεσαν πως καλά σημεία στον χώρο αναζήτησης προέρχονται από καλές παραμέτρους στις στρατηγικές εξέλιξης [14]. Έτσι οι στρατηγικές εξέλιξης οδηγήθηκαν στην αυτοπροσαρμογή.

Η συλλογιστική είναι πως οι ενδογενείς παράμετροι συζεύονται με τις παραμέτρους βελτιστοποίησής (τον πληθυσμό) ώστε κάθε υποψήφιος να έχει τις δικές του ενδογενείς παραμέτρους όπου κάθε υποψήφιος ορίζεται $a_k := (\mathbf{y}_k, \mathbf{s}_k, F(\mathbf{y}))$ με \mathbf{s}_k ενδογενείς παραμέτρους. Αυτές οι μεταβλητές πάλι μπορούν να υποστούν ανασυνδυασμό και μετάλλαξη. Με την διαδικασία της επιλογής τώρα, που γίνεται βάση της καταλληλότητας του υποψηφίου, επιπλέον επιλέγονται άρα και κληροδοτούνται οι ενδογενείς μεταβλητές μαζί με τις αντικειμενικές μεταβλητές του υποψηφίου. Οι ενδογενείς μεταβλητές ελέγχουν κατά κάποιο τρόπο τις διαδικασίες που παράγουν τους υποψήφιους, όπως την μετάλλαξη. Έτσι οι ενδογενείς μεταβλητές έχουν μεγαλύτερη πιθανότητα επιβίωσης όταν προκαλούν μεταβολές που οδηγούν σε καλύτερες τιμές των αντικειμενικών μεταβλητών και επομένως της

αντικειμενικής συνάρτησης. Άρα αναμένεται πως τα άτομα του πληθυσμού μπορούν να “μαθαίνουν” τις βέλτιστες ενδογενείς στρατηγικές παραμέτρους κατά της διάρκειας του αλγόριθμου [10].

Σε αυτό το σημείο θα πρέπει να γίνει ένας διαχωρισμός που έχει να κάνει με το μέγεθος βήματος ή αλλιώς την δύναμη μετάλλαξης που παρουσιάσαμε. Πλέον ο όρος καθολικό μέγεθος βήματος (global step size) αναφέρεται σε ένα βαθμωτό μέγεθος ενώ το ατομικό (individual) μέγεθος βήματος (individual step size) είναι διάνυσμα. Αν και έχουν να κάνουν με την ίδια έννοια, την διακύμανση σε μία κατανομή, η διαφοροποίηση έχει να κάνει με τον τρόπο εφαρμογής στις διάφορες στρατηγικές εξέλιξης και θα γίνει φανερή παρακάτω.

Η εφαρμογή της αυτοπροσαρμογής ποικίλει από υλοποίηση σε υλοποίηση. Αρχικά μπορεί να ενσωματώνει μια μόνο παράμετρο όπως μια κοινή δύναμη μετάλλαξης (ή μέγεθος βήματος) για όλες τις μεταβλητές σχεδιασμού, προσαρμόζοντας έτσι μόνο το καθολικό μέγεθος βήματος, με την γενική μορφή της υλοποίησης αυτήν να απεικονίζεται από την επόμενη σχέση.

$$\begin{aligned}\sigma' &= \sigma \exp(\tau_0 N(0,1)) \\ y'_i &= y_i + \sigma' N_i(0,1)\end{aligned}$$

Το βαθμωτό μέγεθος τ_0 στην πάνω σχέση ελέγχει τον ρυθμό αυτοπροσαρμογής. Πιο προηγμένη μέθοδος θεωρεί δύναμη μετάλλαξης σ_i για κάθε μεταβλητή y_i , το λεγόμενο ατομικό μέγεθος βήματος (individual step size).

$$\begin{aligned}\sigma'_i &= \sigma_i \exp(\tau' N(0,1) + \tau N_i(0,1)) \\ y'_i &= y_i + \sigma'_i N_i(0,1)\end{aligned}$$

Το τ' καθορίζει τον συνολικό ρυθμό αυτοπροσαρμογής και η παράμετρος τ ελέγχει τον ρυθμό προσαρμογής για κάθε μεταβλητή. Έτσι σε αυτή την εφαρμογή είναι δυνατόν να προσαρμοστεί και το ατομικό αλλά και το καθολικό μέγεθος βήματος.

Μια τελευταία υλοποίησης έχει στόχο να ελέγξει την συσχέτιση μεταξύ των αντικειμενικών μεταβλητών. Εκτός από την ενδογενή μεταβλητή σ_i , έχουμε επιπλέον ένα σύνολο γωνιών περιστροφής α_{ij} που περιγράφει την συσχέτιση μεταξύ μεταβλητών y_i και y_j .

$$\begin{aligned}\sigma'_i &= \sigma_i \exp(\tau' N(0,1) + \tau N_i(0,1)) \\ \alpha'_{ij} &= \alpha_{ij} + \beta N_{ij}(0,1) \\ y' &= y + \mathbf{N}(\mathbf{0}, \mathbf{C}(\sigma', \alpha'))\end{aligned}$$

Όπου β ελέγχει τον ρυθμό γωνιών περιστροφής και $\mathbf{C}(\sigma', \alpha')$ πίνακας συνδιακύμανσης [6].

Η τελευταία προσέγγιση οδήγησε στην ανάπτυξη της στρατηγικής εξέλιξης MSC-ES (Mutative Self-Adaptation of Covariances) ή CORR-ES από τον Schwefel το 1981. Παρουσιάζουμε συνοπτικά τον αλγόριθμο.

(μ,λ)-MSC-ES
Αρχικοποίηση $P^{(0)} \leftarrow \{(x_1, \sigma_1, \alpha_1), \dots, (x_\mu, \sigma_\mu, \alpha_\mu)\}$
$t \leftarrow 0$
Επανάλαβε
$t \leftarrow t + 1$
% Ανασυνδιασμός
$\bar{y} \leftarrow \frac{1}{\mu} \sum_{i=1}^{\mu} y_i$

```


$$\bar{\sigma} \leftarrow \frac{1}{\mu} \sum_{i=1}^{\mu} \sigma_i$$

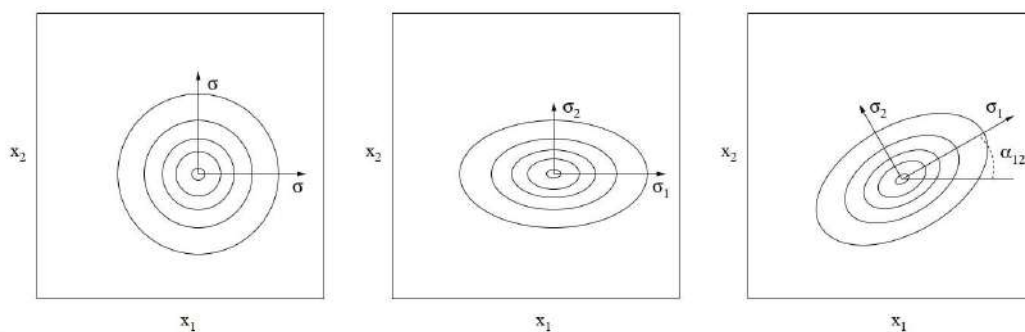

$$\bar{\alpha} \leftarrow \frac{1}{\mu} \sum_{i=1}^{\mu} \alpha_i$$

για  $i = 1$  έως  $\lambda$  επανέλαβε
   $\eta \leftarrow \tau' N(0,1)$ 
   $\sigma_i \leftarrow \bar{\sigma} \exp(\eta + \tau N(\mathbf{0}, \mathbf{I}))$ 
   $\alpha_i \leftarrow \bar{\alpha} + \beta N(\mathbf{0}, \mathbf{I})$ 
   $\mathbf{C} \leftarrow \prod_{i=1}^{n-1} \prod_{j=i+1}^n R_{ij}, R_{ij} = \cos(\alpha_{ij})$ 
   $\mathbf{y}_i \leftarrow \bar{\mathbf{y}} + \mathbf{C} \circ \sigma_i \circ N(\mathbf{0}, \mathbf{I})$ 
   $\varphi_i \leftarrow f(\mathbf{y}_i)$ 
για-τέλος
% Επιλογή
 $p^{(t)}$  τα  $\mu -$  καλύτερα  $(\mathbf{x}_i, \sigma_i, \alpha_i)$  από  $1 \leq i \leq \lambda$ 
Μέχρι (Κριτήριο Τερματισμού)

```

Πίνακας 3. Ψευδοκώδικας για την MSC-ES.

Από τον ψευδοκώδικα για την MSC φαίνεται ξεκάθαρα πως το ο υποψήφιος αποτελείται από το σύνολο $(\mathbf{x}_i, \sigma_i, \alpha_i)$ όπου ενσωματώνονται και οι ενδογενείς παράμετροι οι οποίοι επιλέγονται για ανασυνδυασμό βάση της καταλληλότητας του \mathbf{x}_i του συνόλου στο οποίο ανήκουν. Επίσης υπάρχουν βαθμωτά μεγέθη $(\tau', \eta, \tau, \beta)$ που καθορίζουν τον ρυθμό προσαρμογής των ενδογενών παραμέτρων, τα οποία μεγέθη μπορεί να έχουν και αυτά στοχαστικές διαδικασίες μέσω δειγματοληψίας από κανονική κατανομή. Παρακάτω φαίνονται, για δύο διαστάσεις, τα αποτελέσματα που έχουν οι τρεις προηγούμενες διαδικασίες αυτοπροσαρμογής στην δειγματοληπτική κατανομή που πραγματοποιεί την μετάλλαξη όπου οι ισοκαμπύλες συμβολίζουν ίση πυκνότητα πιθανότητας.



Εικόνα 9 [6]. Προσαρμογή της δειγματοληπτικής κατανομής.

Παρατηρούμε ότι στην πρώτη εικόνα η κατανομή παραμένει ισότροπη και αυτό που μπορούμε να ελέγξουμε μέσω του καθολικού βήματος είναι το μέγεθος των ομόκεντρων κύκλων. Στην δεύτερη περίπτωση μέσω των ατομικών μεγεθών βήματος μπορούμε να έχουμε ελλειψοειδής ισοκαμπύλες με τους κύριους άξονες των καμπυλών να είναι ευθυγραμμισμένοι με το σύστημα συντεταγμένων και τέλος με την τελευταία μέθοδο μπορούμε να παράξουμε οποιαδήποτε κατανομή αφού μπορεί να περιστραφεί. Μπορούμε να πούμε πως διαισθητικά το να εκμεταλλευόμαστε τον πίνακα συνδιακύμανσης όπως στις δύο τελευταίες περιπτώσεις μπορεί να βελτιώσει την απόδοση του αλγορίθμου εισάγοντας μια μεροληψία στην μετάλλαξη προς μια πιο υποσχόμενη κατεύθυνση, αναφορικά πάντα με τις τιμές της αντικειμενικής συνάρτησης.

Ακολουθώντας αυτή την λογική οι αλγόριθμοι αυτοπροσαρμογής μπορούν να χωριστούν με βάση την πολυπλοκότητα των ενδογενών στρατηγικών παραμέτρων που ελέγχουν. Στο μητρώο συνδιακύμανσης επειδή είναι θετικά ορισμένο μπορεί να γίνει μια ιδιοανάλυση (Ανάλυση πίνακα σε ιδιάζουσες τιμές – Singular Value Decomposition) εκφράζοντας το ως $\mathbf{C} = \mathbf{B}\mathbf{D}^2\mathbf{B}^T$ με τις στήλες του \mathbf{B} να περιέχουν ορθοκανονική βάση με των ιδιοδιανυσμάτων και \mathbf{D} διαγώνιος πίνακας με τις τετραγωνικές ρίζες των ιδιοτιμών.

Γενικά για μια πολυμεταβλητή κανονική κατανομή $N(\mathbf{m}, \mathbf{C})$, ισχύει

$$\begin{aligned} N(\mathbf{m}, \mathbf{C}) &\sim \mathbf{m} + N(\mathbf{0}, \mathbf{C}) \\ &\sim \mathbf{m} + \mathbf{C}^{\frac{1}{2}} N(\mathbf{0}, \mathbf{I}) \\ &\sim \mathbf{m} + \mathbf{B}\mathbf{D}\mathbf{B}^T N(\mathbf{0}, \mathbf{I}), \text{ με } \mathbf{B}^T N(\mathbf{0}, \mathbf{I}) \sim N(\mathbf{0}, \mathbf{I}) \\ &\sim \mathbf{m} + \mathbf{B}\mathbf{D} N(\mathbf{0}, \mathbf{I}) \end{aligned}$$

Ξεχωρίζουμε τρεις ειδικές περιπτώσεις γύρω από την κατανομή που χρησιμοποιείται στις στρατηγικές εξέλιξης άρα και κατά συνέπεια στις μεταβλητές που προσαρμόζουν οι αλγόριθμοι. Οι τρεις κατανομές αυτές είναι οι $N(\mathbf{0}, \mathbf{I})$, $N(\mathbf{0}, \text{diag}(\delta^2))$ και $N(\mathbf{0}, \mathbf{C})$.

Στην πρώτη περίπτωση χρησιμοποιώντας $\mathbf{B} = \mathbf{I}, \mathbf{D} = \sigma\mathbf{I}$ παράγονται μεταλλάξεις μέσω $\mathbf{y}' = \mathbf{y} + \delta N(\mathbf{0}, \mathbf{I})$ και η παράμετρος που ελέγχεται είναι το καθολικό βήμα δ πραγματοποιώντας δειγματοληψίες σαν την πρώτη του παραπάνω σχήματος. Στην δεύτερη περίπτωση με $\mathbf{B} = \mathbf{I}, \mathbf{D} = \text{diag}(\delta)$ έχουμε $\mathbf{y}' = \mathbf{y} + \mathbf{I}\text{diag}(\delta)N(\mathbf{0}, \mathbf{I}) = \mathbf{y} + N(\mathbf{0}, \text{diag}(\delta^2))$ όπου η παράμετρος που ελέγχεται είναι τα ατομικά μεγέθη βήματος δ παράγοντας δειγματοληψίες σαν την δεύτερη του σχήματος και στην τελευταία περίπτωση προσαρμόζεται επιπλέον και ο \mathbf{B} πίνακας ώστε να έχουμε $\mathbf{y}' = \mathbf{y} + \mathbf{B}\text{diag}(\delta)N(\mathbf{0}, \mathbf{I}) = \mathbf{y} + \mathbf{B}N(\mathbf{0}, \text{diag}(\delta^2)) = \mathbf{y} + N(\mathbf{0}, \mathbf{C})$ [15].

3.2. Αποτυχαιοποίηση

3.2.1. 1^ο Επίπεδο

Στις παραπάνω μεθόδους αυτοπροσαρμογής μείζονα ρόλο κατέχει ο ακριβής τρόπος εφαρμογής των παραμέτρων που ελέγχουν τον ρυθμό προσαρμογής και ο τρόπος με τον οποίο εφαρμόζεται η δειγματοληψία για τις διάφορες παραμέτρους. Εξετάζοντας τις πρώτες εφαρμογές της αυτοπροσαρμογής για μία (1,λ) στρατηγική έχουμε για την παραγωγή των νέων απογόνων $N(\text{Newer})$ από τον παλαιό γονέα $E(\text{Elder})$ ως εξής :

$$\mathbf{y}_{N_k}^g = \mathbf{y}_E^g + \xi^k \xi^k \delta^g \circ \mathbf{z}^k \quad (21),$$

όπου δ^g τα ατομικά βήματα στην γενιά g , το $\xi^k = \alpha$ ή $1/\alpha$ με ίση πιθανότητα και $\alpha = 1.4$ το οποίο ελέγχει την προσαρμογή του καθολικού βήματος του απογόνου k , $\xi^k = (\xi_1, \dots, \xi_n)$ ελέγχει το ατομικό βήμα του απογόνου k με το ξ_i να ακολουθεί την ίδια κατανομή με το ξ (ή κάποια πιο σύνθετη κατανομή) και $\mathbf{z}^k = (z_1, \dots, z_n)$ με το $z_i \sim (0,1)$.

Αξιολογούνται οι λ υποψήφιοι και επιλέγεται ο καλύτερος γονέας για την νέα γενιά $\mathbf{y}_E^{g+1} = \mathbf{y}_{N_{sel}}^g$, μαζί με το ατομικό μέγεθος βήματος που αντιστοιχεί σε αυτόν ως

$$\delta^{g+1} = \xi^{sel} \xi^{sel} \delta^g \quad (22).$$

Ένα θέμα που προκύπτει στην παραπάνω μεθοδολογία είναι πως σύμφωνα με τους Hansen και Ostermeier [16], υπάρχει πιθανότητα μια μεγάλη ή μικρή αλλαγή στο μέγεθος βήματος

να μην συμβαίνει λόγο μεγάλης ή μικρής τιμής των παραμέτρων ξ αντίστοιχα. Μπορεί να έχουμε μικρή τιμή στην μεταβλητή που ελέγχει το μέγεθος βήματος, ξ , αλλά όμως να προκληθεί μεγάλη μεταβολή λόγω της τυχαίας δειγματοληψίας \mathbf{z} [16]. Δηλαδή δεν μπορούμε να γνωρίζουμε αν καλές τιμές στις αντικειμενικές μεταβλητές προέρχονται σίγουρα από καλό βήμα ή λόγο ενός πλεονεκτικού σημείου που έτυχε να παραχθεί από την δειγματοληψία [15]. Σε δεύτερο λόγο επιδιώκουμε να διατηρήσουμε υψηλή διακύμανση της δύναμης μετάλλαξης (που παράγει τους απογόνους) μέσα σε μια γενιά αλλά να αποφύγουμε μεγάλες διακυμάνσεις μεγέθους βήματος μεταξύ διαδοχικών γενεών. Αυτές οι δύο απαιτήσεις με τον τρόπο που είναι δομημένες αυτές οι στρατηγικές είναι αντικρουόμενες [16], καθώς η παράμετρος ξ βρίσκεται και στην σχέση (21) που εισάγει διακύμανση στους υποψήφιους μια γενιάς αλλά και στην σχέση (22) που συνδέει και προσαρμόζει τα διαδοχικά βήματα ανάμεσα σε δύο γενιές.

Για να αντιμετωπιστούν αυτά τα προβλήματα του παραπάνω αλγόριθμου οι Hansen και Ostermeier [16] αρχικά, αντί να εισάγουν μια νέα ξεχωριστή διακύμανση ξ^k στο μέγεθος του βήματος και μια ξεχωριστή δειγματοληψία \mathbf{z}^k στην μετάλλαξη του υποψηφίου, χρησιμοποίησαν την ίδια δειγματοληψία \mathbf{z} της μετάλλαξης των υποψηφίων και για την διακύμανση στο μέγεθος του βήματος [11]. Συγκεκριμένα κάνανε χρήση του μήκους του πιο επιτυχημένου βήματος $\mathbf{z}_{sel} \sim N(\mathbf{0}, \mathbf{I})$, της δειγματοληψίας της μετάλλαξης (δηλαδή αυτής της μετάλλαξης που παρήγαγε τον καλύτερο υποψήφιο σε μια γενιά) για τον έλεγχο της προσαρμογής του ατομικού (individual) μεγέθους το βήματος. Έτσι αντιμετωπίστηκε το πρώτο ζήτημα όπου ήταν πιθανό διαφορετικές τιμές στην παράμετρο ξ^k μπορούσαν να παράξουν ίδιες μεταλλάξεις στην σχέση (20) επειδή το \mathbf{z} και το ξ μπορούν να μεταβάλλονται ανεξάρτητα.

Πιο συγκεκριμένα επιλέγεται για κάθε γενιά το διάνυσμα της μετάλλαξης που παρήγαγε τον καλύτερο απόγονο και υπολογίζεται το (ατομικό) μέγεθος βήματος μέσω της διαφοράς του διανύσματος αυτού από την αναμενόμενη τιμή του μεγέθους ενός διανύσματος που παράγεται από την $N(\mathbf{0}, \mathbf{I})$ κατανομή ως $\xi_{z_{sel}} = (\xi_1, \dots, \xi_n)$, $\xi_i = \exp \left\{ |z_i^{sel}| - \sqrt{\frac{2}{\pi}} \right\}$. Αυτές οι συνιστώσες δομούν το διάνυσμα ατομικών μεγέθους βημάτων δ^g . Η εκθετική συνάρτηση χρησιμοποιείται διότι ο μέσος των απόλυτων τιμών των μεγεθών των μεταλλάξεων που επιλέγονται, $|z_i^{sel}|$, είναι μικρότερο του ενός και αυτό οδηγεί σε σταδιακή μείωση των μεγεθών του ατομικού βήματος (οι σχέσεις είναι αναδρομικές) κάτι που διορθώνεται με αυτόν τον εκθετικό μετασχηματισμό.

Οπότε διατηρώντας την φιλοσοφία της αυτοπροσαρμογής οι ερευνητές στόχευσαν στην αυτοπροσαρμογή χωρίς ανεξάρτητη στοχαστική μεταβολή των ενδογενών παραμέτρων οι οποίοι τώρα υποβάλλονται σε μετάλλαξη μέσω των ίδιων στοχαστικών διαδικασιών που λαμβάνει χώρα για τις αντικειμενικές μεταβλητές. Αυτές οι διαδικασίες χαρακτηρίστηκαν ως αποτυχατοποιημένες (derandomized). Οποιοσδήποτε αλγόριθμος με απευθείας μετάλλαξη στις ενδογενείς παραμέτρους της στρατηγικής είναι υποκείμενος σε θόρυβο, καθώς η επιλογή κατεργάζεται την προσαρμογή των αντικειμενικών μεταβλητών ενώ οι ενδογενείς μεταβλητές αλλάζουν με έναν "χαλαρό" στοχαστικό τρόπο μαζί με τις αντικειμενικές μεταβλητές [16].

Όσον αφορά το δεύτερο πρόβλημα για την διατήρηση υψηλής διακύμανσης ανάμεσα στους υποψήφιους μιας γενιάς αλλά όχι ανάμεσα στα διαδοχικά βήματα, οι ερευνητές εισήγαγαν μια απόσβεση στις παραμέτρους όπως φαίνεται από την επόμενη σχέση

$$\delta^{g+1} = (\xi^{sel})^\beta (\xi^{sel})^{\beta_{scal}} \delta^g \quad (23)$$

με β και β_{scal} να παίρνουν τιμές στο (0,1) πραγματοποιώντας μια πιο αργή προσαρμογή με μεγαλύτερη ακρίβεια. Επομένως προσαρμογή του βήματος κατά ξ δεν γίνεται σε μια γενιά, αλλά σε $1/\beta > 1$ γενιές. Έτσι μειώνεται η στοχαστική διακύμανση στα διαδοχικά βήματα και η πληροφορία γύρω από το μέγεθος βήματος γίνεται πιο αξιόπιστη ώστε μεγάλα βήματα να δίνουν την πληροφορία πως το βήμα θα πρέπει να μεγαλώσει περισσότερο.

Αυτός ο αλγόριθμος ονομάστηκε DR1 (DeRandomization 1) και παρουσιάζονται τα βασικά του στοιχεία και ένας ψευδοκώδικας για αυτόν στους παρακάτω πίνακες.

Μετάλλαξη και αυτοπροσαρμογή στον Αλγόριθμο DR1
$\mathbf{y}_{N_k}^g = \mathbf{y}_E^g + \xi^k \delta^g \circ \mathbf{z}^k$
$\mathbf{y}_E^{g+1} = \mathbf{y}_{N_{sel}}^g$
$\delta^{g+1} = (\xi^{sel})^\beta (\xi_{z^{sel}})^{\beta_{scal}} \delta^g$
<p>Όπου,</p> <ul style="list-style-type: none"> • $\mathbf{y}_{E/N}^g$, διάνυσμα αντικειμενικής παραμέτρου, E=Elder (γονέας), N=Newer (απόγονος). • δ^g, διάνυσμα ατομικού μεγέθους βήματος για τη γενιά g. • $\xi^k = \alpha$ ή $1/\alpha$, με ίση πιθανότητα και $\alpha = 1.4$. Καθολικό μέγεθος βήματος απογόνου k. • $\xi_{z^{sel}} = (\xi_1, \dots, \xi_n)$, $\xi_i = \exp \left\{ z_i^{sel} - \sqrt{\frac{2}{\pi}} \right\}$, ατομικό μέγεθος βήματος γενιάς g με $\sqrt{\frac{2}{\pi}}$ να είναι η αναμενόμενη τιμή του z_i. • $\mathbf{z}^k = (z_1, \dots, z_n)$, με το $z_i \sim N(0,1)$. • $\beta = \sqrt{\frac{2}{n}}$ και $\beta_{scal} = 1/n$, παίρνουν τιμές στο (0,1).

Πίνακας 4. DR1 Αλγόριθμος.

DR1
<p>Αρχικοποίηση $\mathbf{x}, \delta \leftarrow (1, \dots, 1)^T$ $t \leftarrow 0$ Επανάλαβε $t \leftarrow t + 1$ για $i = 1$ έως λ επανάλαβε $\mathbf{z}_i \leftarrow N(\mathbf{0}, \mathbf{I})$ $\mathbf{x}_i \leftarrow \mathbf{x} + \xi \cdot \delta \circ \mathbf{z}_i$, όπου $P(\xi = \alpha) = P\left(\xi = \frac{1}{\alpha}\right) = \frac{1}{2}$, με $\alpha = 1.4$ $\varphi_i \leftarrow f(\mathbf{x}_i)$ για-τέλος $sel \leftarrow i$ η καλύτερη τιμή του φ_i $\mathbf{x} \leftarrow \mathbf{x}_{sel}$</p>

$$\xi_{z_{sel}} = \left(\exp\left(|z_{sel_1}| - \sqrt{\frac{2}{\pi}}\right), \dots, \exp\left(|z_{sel_i}| - \sqrt{\frac{2}{\pi}}\right) \right)$$

$$\delta \leftarrow (\xi_{sel})^\beta (\xi_{z_{sel}})^{\beta_{scal}} \circ \delta$$

Μέχρι (Κριτήριο Τερματισμού)

Πίνακας 5. Ψευδοκώδικας για τον DR1 αλγόριθμο.

3.2.2. 2^ο Επίπεδο.

Ο ρυθμός μεταβολής των στρατηγικών παραμέτρων που ορίζεται ως η διαφορά ανάμεσα στις στρατηγικές παραμέτρους δύο διαδοχικών γενιών, είναι πολύ σημαντικός παράγοντας και μας δίνει μια ένδειξη αναφορικά με την ταχύτητα προσαρμογής άρα και την απόδοση του αλγορίθμου. Αυτός ο ρυθμός μεταβολής εξαρτάται από το μέγεθος του βήματος των στρατηγικών παραμέτρων. Αν επιδιώξουμε να διατηρήσουμε έναν μέγιστο ρυθμό, συναντάμε ένα πάνω όριο που έχει να κάνει με το πόση πληροφορία μπορεί να μεταφερθεί μεταξύ των γενιών. Αν περάσουμε αυτό το άνω όριο οδηγούμαστε σε στοχαστική συμπεριφορά. Από την άλλη το μέγεθος βήματος που δίνει τον βέλτιστο ρυθμό μεταβολής των στρατηγικών παραμέτρων είναι αρκετά μικρό ώστε να μην υπάρχει μεγάλη διαφοροποίηση ανάμεσα στους απογόνους που παράγονται.

Έτσι όπως δομήθηκε ο DR1 για να αποφευχθεί στοχαστική συμπεριφορά στην προσαρμογή του ατομικού μεγέθους βήματος, η διακύμανση του (individual) ατομικού μεγέθους βήματος από γενιά σε γενιά είναι μικρή. Επιπλέον αν και ο MSC λειτουργούσε καλά όσον αφορά την προσαρμογή του καθολικού βήματος, ωστόσο η προσαρμογή του ατομικού μεγέθους βήματος σε μικρούς πληθυσμούς ήταν αδύνατος, (χρήση πληθυσμού 10*n οδηγούσε σε τυχαίο περίπατο αναφορικά με τα ατομικά μεγέθη βήματος). Αυτό ήταν ένα ακόμα πρόβλημα που έλυσε ο DR1. Ο DR1 αλγόριθμος ήταν δυνατόν να πραγματοποιηθεί και σε μια $(\mu/\mu_1, \lambda)$ εκδοχή όπου παράγονται όμως πολύ μικρά (individual) μεγέθη βήματος. Συνδυάζοντας όμως τις δύο παραπάνω ιδιαιτερότητες, δηλαδή τον μικρό πληθυσμό και τα μικρά ατομικά μεγέθη βήματος ο ρυθμός προόδου, $\varphi_{\mu/\mu_1, \lambda}^*$ (σ^*) που ορίσαμε προηγουμένως είναι μικρός. Έχοντας οπότε μικρό πληθυσμό και μικρό μέγεθος βήματος δυσχεραίνεται ο ρυθμός προόδου. Έτσι έγινε φανερό στους ερευνητές πως η πληροφορία για την αυτοπροσαρμογή του βήματος είναι καλύτερο να συλλεχθεί σε μια αλληλουχία γενεών με μικρό πληθυσμό [17].

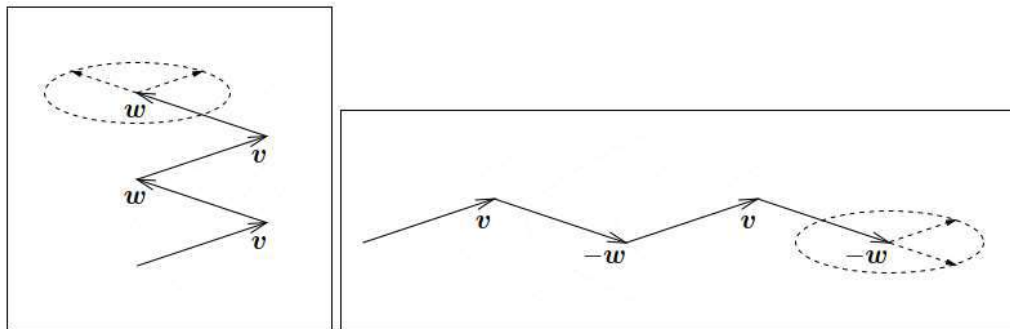
Ο DR1 μεταβάλλεται ως εξής, αντικαθίσταται το ξ^k και στην θέση του έρχεται το δ^g καθολικό ή γενικό μέγεθος βήματος, το οποίο βασίζεται στην νέα έννοια της συσσώρευσης που θα αναλύσουμε όπως και το ατομικό (ή local(τοπικό)) μέγεθος βήματος δ_{scal}^g .

Επομένως πλέον για την προσαρμογή του ατομικού μεγέθους του βήματος χρησιμοποιείται το άθροισμα διαδοχικών επιτυχημένων βημάτων μετάλλαξης, αθροίζοντας δηλαδή τις δειγματοληψίες που παράγουν τον καλύτερο υποψήφιο σε κάθε γενιά αν πρόκειται για $(1, \lambda)$ αλγόριθμο ή όλες αυτές που παρήγαγαν τους μ υποψήφιους που επιλέγονται για ανασυνδιασμό εάν πρόκειται για $(\mu/\mu_1, \lambda)$ αλγόριθμο [7], [17]. Αυτό το άθροισμα ονομάζεται μονοπάτι εξέλιξης (evolution path) και η διαδικασία υπολογισμού ονομάζεται συσσώρευση (cumulation), καθώς με την συσσώρευση επιπρόσθετα χρησιμοποιείται όχι μόνο το μέγεθος του πιο επιτυχημένου βήματος της προηγούμενης γενιάς αλλά χρησιμοποιείται πληροφορία και από τα προηγούμενα επιτυχημένα βήματα των προηγούμενων γενιών συσσωρεύοντας πληροφορία καθώς περνούν οι γενιές [15]. Το μονοπάτι εξέλιξης υπολογίζεται ως εξής :

$$\mathbf{p}^g = (1 - c)\mathbf{p}^{g-1} + c\mathbf{z}_{sel}$$

όπου c ένας παράγοντας απόσβεσης. Εκτός από μια επίδραση παρόμοια του ανασυνδιασμού, αν οι μεταλλάξεις που επιλέγονται (\mathbf{z}_{sel}) από γενιά σε γενιά, ήταν ασυσχέτιστες η μεθοδολογία δεν θα είχε νόημα. Αν τα διαδοχικά βήματα είναι παράλληλα συσχετιζόμενα τότε αυξάνεται το άθροισμα και αυτό σημαίνει πως τα διαδοχικά βήματα πηγαίνουν προς την ίδια κατεύθυνση. Στην αντίθετη περίπτωση αν το άθροισμα είναι σχετικά μικρό τα διαδοχικά βήματα είναι αντιπαράλληλα συσχετιζόμενα. Άρα για να μεγιστοποιήσουμε την απόδοση του αλγορίθμου θα πρέπει να κάνουμε μεγαλύτερα βήματα (αύξηση του μεγέθους του βήματος) στις κατευθύνσεις που το μονοπάτι εξέλιξης μεγάλο. Αντίθετα θα πρέπει να πραγματοποιούμε μικρότερα βήματα στην κατεύθυνση όπου το μονοπάτι είναι μικρό [18]. Θα μπορούσαμε να πούμε διαισθητικά πως το μονοπάτι εξέλιξης μας δίνει ένα μέτρο για να καταλάβουμε αν κινούμαστε προς την σωστή κατεύθυνση ή όχι, δίνοντας μας την δυνατότητα να αυξήσουμε την ταχύτητα με την οποία κινούμαστε σε μια κατεύθυνση που φαίνεται “σωστή”.

Μια πτυχή της χρησιμότητας του μονοπατιού εξέλιξης φαίνεται με το παρακάτω ιδεατό παράδειγμα δύο περιπτώσεων όπου θεωρούμε δύο ακολουθίες από τέσσερα διαδοχικά βήματα στις γενιές [18].



Εικόνα 10 [16]. Δύο περιπτώσεις με διαφορετικά μονοπάτια όπου σε εφαρμογή χωρίς συσσώρευση καταλήγουμε λανθασμένα στην ίδια κατανομή

Φαίνονται πως καθώς εξελίσσονται οι γενιές ακολουθείται τελείως διαφορετική κατεύθυνση στα δύο σχήματα, πράγμα που θα μπορούσε να σημαίνει πως έχουμε να κάνουμε με τελείως διαφορετικό πρόβλημα βελτιστοποίησης. Ωστόσο η μόνη διαφορά στα βήματα των δύο περιπτώσεων έγκειται μόνο στο πρόσημο του ενός διανύσματος, του w . Αν δεν χρησιμοποιήσουμε την συσσώρευση, αλλά ξεχωριστά τα βήματα σε κάθε γενιά τότε για τις δύο διαφορετικές περιπτώσεις προσαρμόζοντας το ατομικό μέγεθος βήματος για παράδειγμα με τον DR1 αλγόριθμο με το Hadamard (component-wise) \circ πολλαπλασιασμό στην σχέση $\delta^{g+1} = (\xi^{sel})^\beta (\xi_{z_{sel}})^{\beta_{scal}} \circ \delta^g$, σε ζυγά βήματα θα είχαμε το ίδιο μέγεθος ατομικού βήματος και θα παράγαμε την ίδια κατανομή (μέσω $\mathbf{y}' = \mathbf{y} + \mathbf{I}diag(\delta)\mathbf{N}(\mathbf{0}, \mathbf{I})$) όπως φαίνεται στο σχήμα με διακεκομμένη γραμμή. Κάτι τέτοιο προφανώς θα ήταν λανθασμένο καθώς στην πρώτη περίπτωση εισάγει μια μεροληψία προς την λανθασμένη κατεύθυνση.

Ο αλγόριθμος που χρησιμοποιεί την έννοια της συσσώρευσης ονομάστηκε DR2 και τα δ^g και δ_{scal}^g υπολογίζονται σύμφωνα με τις παρακάτω σχέσεις όπου στην πρώτη για το καθολικό βήμα υπολογίζουμε την απόλυτη τιμή του διανύσματος \mathbf{p} ενώ στις δεύτερη τις απόλυτες τιμές των συνιστωσών του διανύσματος \mathbf{p} .

$$\delta^{g+1} = \delta^g \left(\exp \left(\frac{\|\mathbf{p}^g\|}{\sqrt{n} \sqrt{\frac{c}{2-c}}} - 1 + \frac{1}{5n} \right) \right)^\beta$$

$$\delta_{scal}^{g+1} = \delta_{scal}^g \circ \left(\frac{|\mathbf{p}^g|}{\sqrt{n} \sqrt{\frac{c}{2-c}}} + 0.35 \right)^{\beta_{scal}}$$

Το μέγεθος του μονοπατιού συγκρίνεται με την αναμενόμενη τιμή $E(\|N(\mathbf{0}, I)\|) \approx \sqrt{n}$
 Στην εξίσωση προσαρμογής του βήματος (general και individual) δ^{g+1} και δ_{scal}^{g+1} , όρος $\sqrt{\frac{c}{2-c}}$
 είναι παράγοντας που κανονικοποιεί την διακύμανση .

Μετάλλαξη και αυτοπροσαρμογή στον Αλγόριθμο DR2
$\mathbf{y}_{N_k}^g = \mathbf{y}_E^g + \delta^g \delta_{scal}^g \circ \mathbf{z}^k$
$\mathbf{y}_E^{g+1} = \mathbf{y}_{Nsel}^g$
$\mathbf{p}^g = (1 - c)\mathbf{p}^{g-1} + c\mathbf{z}_{sel}$
$\delta^{g+1} = \delta^g \left(\exp \left(\frac{\ \mathbf{p}^g\ }{\sqrt{n} \sqrt{\frac{c}{2-c}}} - 1 + \frac{1}{5n} \right) \right)^\beta$ $\delta_{scal}^{g+1} = \delta_{scal}^g \circ \left(\frac{ \mathbf{p}^g }{\sqrt{n} \sqrt{\frac{c}{2-c}}} + 0.35 \right)^{\beta_{scal}}$
<p>Όπου,</p> <ul style="list-style-type: none"> • $\mathbf{y}_{E/N}^g$ διάνυσμα αντικειμενικής παράμετρου, E=Elder (γονέας), N=Newer (απόγονος) • δ^g ατομικό διάνυσμα μεγέθους βήματος για γενιά g • δ^g καθολικό μέγεθος βήματος γενιάς g • $\mathbf{z}^k = (z_1, \dots, z_n)$ με το $z_i \sim N(0,1)$ • $c = \sqrt{1/n}$ καθορίζει το πόσο γρήγορα αποσβέννεται η συνεισφορά των προηγούμενων γενεών.

- $\beta = \sqrt{1/n}$ και $\beta_{scal} = 1/n$, μικρές τιμές δίνουν πιο ακριβής προσαρμογή αλλά πιο αργή.

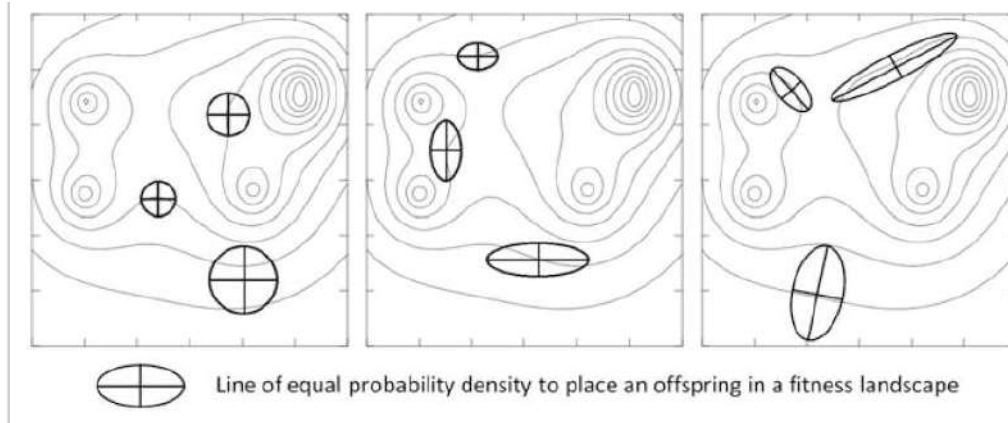
Πίνακας 6. Αλγόριθμος DR2

DR2 (1,λ)
<p>Αρχικοποίηση $\mathbf{x}, \boldsymbol{\delta}_{scal} \leftarrow (1, \dots, 1)^T, \delta \leftarrow 1, \mathbf{p} \leftarrow \mathbf{0}$ $t \leftarrow 0$ Επανάλαβε $t \leftarrow t + 1$ για $i = 1$ έως λ επανάλαβε $\mathbf{z}_i \leftarrow N(\mathbf{0}, \mathbf{I})$ $\mathbf{y}_i \leftarrow \mathbf{y} + \delta \cdot \boldsymbol{\delta}_{scal} \circ \mathbf{z}_i$ $\varphi_i \leftarrow f(\mathbf{x}_i)$ για-τέλος $sel \leftarrow i$ η καλύτερη τιμή του φ_i $\mathbf{p} \leftarrow (1 - c)\mathbf{p} + \mathbf{z}_{sel}$ $\delta' = \delta \left(\exp \left(\frac{\ \mathbf{p}\ }{\sqrt{n} \sqrt{\frac{c}{2-c}}} - 1 + \frac{1}{5n} \right) \right)^\beta$ $\boldsymbol{\delta}'_{scal} = \boldsymbol{\delta}_{scal} \circ \left(\frac{\ \mathbf{p}^g\ }{\sqrt{n} \sqrt{\frac{c}{2-c}}} + 0.35 \right)^{\beta_{scal}}$ $\mathbf{y} \leftarrow \mathbf{y}_{sel}$ $\mathbf{p} \leftarrow \mathbf{p}'$ $\delta \leftarrow \delta'$ $\boldsymbol{\delta}_{scal} \leftarrow \boldsymbol{\delta}'_{scal}$ Μέχρι (Κριτήριο Τερματισμού)</p>

Πίνακας 7. Ψευδοκώδικας αλγόριθμου DR2.

3.2.3. Προσαρμογή του μητρώου Συνδιακύμανσης

Οι αλγόριθμοι DR1 και DR2 που παρουσιάσαμε δεν εκμεταλλεύονται στο έπακρο την κατανομή της δειγματοληψίας. Εκμεταλλεύονται την κατανομή $N(\mathbf{0}, \text{diag}(\boldsymbol{\delta}^2))$ προσαρμόζοντας το μέγεθος βήματος ώστε να μπορούν να πραγματοποιήσουν δειγματοληψίες έχοντας κατανομές με ισοκαμπύλες πυκνότητας πιθανότητας που είτε είναι σφαιρικές είτε ελλειψοειδής παράλληλες με τους άξονες ώστε να είμαστε πάντα “ευθυγραμμισμένοι” με το σύστημα συντεταγμένων. Έτσι δεν μπορούμε να λύσουμε όλα τα προβλήματα με μεγάλη απόδοση. Αν για παράδειγμα έχουμε ένα πρόβλημα ελαχιστοποίησης με μακριά κοιλάδα μπορούμε να είμαστε αποδοτικοί μόνο αν αυτή είναι ευθυγραμμισμένη με έναν από τους άξονες μας [11].



Εικόνα 11 [3]. Ισοκαμπύλες πυκνότητας πιθανότητας στον χώρο αναζήτησης για συνάρτηση καταλληλότητας.

Από την άλλη ο αλγόριθμος MSC, που προσαρμόζει το μητρώο συνδιακύμανσης, εκτός από τα μειονεκτήματα που αναφέραμε και οδήγησαν στην αποτυχαιοποίηση των DR1 DR2, μειονεκτεί επίσης λόγω απουσίας της ιδιότητας της αμεταβλητότητας όπως θα δούμε.

Έτσι οι N. Hansen, A. Ostermeier και A. Gawelczyk [18] ανέπτυξαν την στρατηγική εξέλιξης (1,λ)-GSA-ES (generating set adaptation) (ή και AI) ώστε να παράγει μεταλλάξεις από δειγματοληψίες οποιασδήποτε κανονικής πολυμεταβλητής κατανομής εκμεταλλευόμενη το πλήρες μητρώο συνδιακύμανσης το οποίο στοχεύει ο αλγόριθμος να προσαρμόσει [3]. Ο στόχος του αλγόριθμου δεν είναι η απευθείας χρήση του μητρώου συνδιακύμανσης αλλά ο μετασχηματισμός ενός ισότροπου τυχαίου διανύσματος $\mathbf{z} = N(\mathbf{0}, \mathbf{I})$ σε ένα συσχετισμένο τυχαίο διάνυσμα μέσω πολλαπλασιασμού με έναν πίνακα $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_m)$ [3].

Η μετάλλαξη των αντικειμενικών μεταβλητών συνήθως γίνεται προσθέτοντας ένα κανονικά κατανομημένο διάνυσμα \mathbf{z}' στο διάνυσμα των αντικειμενικών μεταβλητών με $\mathbf{z}' = \delta \cdot (z_1 \mathbf{b}_1 + \dots + z_n \mathbf{b}_n)$, όπου n ο αριθμός των μεταβλητών, δ καθολικό μέγεθος βήματος, $z_i \sim N(0,1)$, $\mathbf{b}_i = \mathbf{e}_i$ τετριμμένη βάση στον \mathbb{R}^n και $\mathbf{z}' \sim N(\mathbf{0}, \delta^2 \mathbf{I})$. Για να πάρουμε ανισότροπη ελλειψοειδής μετάλλαξη παράλληλη με τους άξονες, πολλαπλασιάζουμε κάθε $z_i \mathbf{b}_i$ με διαφορετικό μέγεθος βήματος σ_i . Μπορούμε να μεταβάλουμε την κατανομή \mathbf{z}' αλλάζοντας τα διανύσματα της βάσης \mathbf{b}_i αποσπώντας το \mathbf{z}' από την δοσμένη βάση. Έτσι μπορούμε να παράξουμε οποιαδήποτε κατανομή, για παράδειγμα μπορούμε να θεωρήσουμε ένα σύνολο διανυσμάτων ορθογώνιο στο \mathbf{b}_i . Από την άλλη η κατανομή είναι πάντα κανονική επειδή προστίθενται κανονικές κατανομές.

Η προσαρμογή θα επιτευχθεί αντικαθιστώντας σταδιακά τα \mathbf{b}_i μια διαδικασία ανεξάρτητη του συστήματος συντεταγμένων. Κάθε φορά το \mathbf{b}_i θα αντικαθίστανται από το “καλύτερο” διάνυσμα \mathbf{z}_i κάθε γενιάς, το \mathbf{z}_{sel} . Σημειώνουμε πως κάθε φορά θα αντικαθίσταται το παλαιότερο \mathbf{b}_i διάνυσμα, διατηρώντας τις “ενημερωμένες” πληροφορίες των πρόσφατων διανυσμάτων. Αλλάζοντας το \mathbf{b}_i για το \mathbf{z}_{sel} έχουμε μεγαλύτερη πιθανότητα να παράξουμε μεταλλάξεις παρόμοιες του τελευταίου στο μέλλον. Αυτή η διαδικασία και η σταδιακή αντικατάσταση όλων των διανυσμάτων δεν εξαρτάται από το αρχικό δεδομένο σύστημα συντεταγμένων αλλά από την διαμόρφωση της αντικειμενικής συνάρτησης. Έτσι έχουμε για κάθε $k = (1, \dots, \lambda)$, με E_j να συμβολίζει τον j γονέα με $j = (1, \dots, \mu)$, με N_k ο k απόγονος όπου $k = (1, \dots, \lambda)$, με $\zeta_k = 1, \dots, \mu$ με ίση πιθανότητα ($\mu = 1$ τότε αγνοούμε το ζ_k) τις παρακάτω σχέσεις που περιγράφει την διαδικασία αυτή.

$$\mathbf{y}^{N_k} = \mathbf{y}^{E_{\zeta_k}} + \xi_k \delta^{E_{\zeta_k}} \mathbf{x}_k,$$

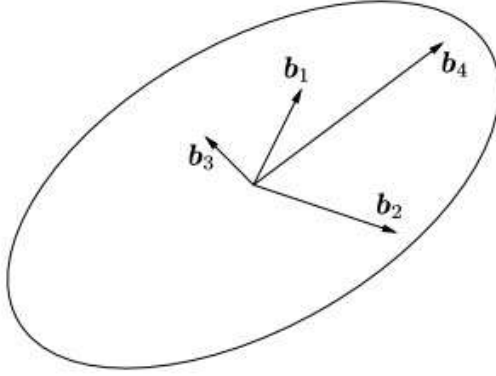
$$\mathbf{x}_k = c_m \cdot \mathbf{B}^{E_{\zeta_k}} \mathbf{z}_k = c_m \sum_{j=1}^m (z_k)_j \mathbf{b}_j^{E_{\zeta_k}}$$

$$\delta^{N_k} = \delta^{E_{\zeta_k}} (\xi_k)^\beta$$

$$\mathbf{b}_1^{N_k} = (1 - c) \cdot \mathbf{b}_1^{E_{\zeta_k}} + c \cdot (c_u \xi_{sel} \mathbf{x}_{sel})$$

$$\mathbf{b}_{i+1}^{N_k} = \mathbf{b}_i^{E_{\zeta_k}} \forall i \in \{1, \dots, m-1\}$$

με $\xi = 1.5$ ή $1/1.5$ με ίση πιθανότητα, $\mathbf{z} = (z_1, \dots, z_m) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ με $z_i \sim \mathcal{N}(0,1)$, δ καθολικό μέγεθος βήματος, $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_m) \in \mathbb{R}^n \times \mathbb{R}^m$ με $\mathbf{b}_i \in \mathbb{R}^n$ και το \mathbf{B} να μετασχηματίζει το \mathbf{z} . Αρχικοποιείται ο \mathbf{B} με $\mathbf{b}_1 = \mathbf{0}$ και $\mathbf{b}_2^{E_j}, \dots, \mathbf{b}_m^{E_j} \sim \mathcal{N}(\mathbf{0}, (1/n)\mathbf{I})$. $m \in \{n^2, \dots, 2n^2\}$ ο αριθμός των διανυσμάτων για τον \mathbf{B} (Generating Set). $c_m = (1/\sqrt{m})(1 + 1/m)$, $c = 1/\sqrt{n}$ και $\beta = 1/\sqrt{n}$. Σημειώνεται πως το νέο \mathbf{b}_i υπολογίζεται ως σταθμισμένος μέσος με απόσβεση όλων των προηγούμενων βημάτων, δηλαδή των καλύτερων βημάτων τα οποία συσσωρεύονται.



Εικόνα 12. Προσαρμογή του μητρώου συνδιακύμανσης με την AI(GSA).

Περαιτέρω εξέλιξη στο ίδιο μοτίβο είναι ο αλγόριθμος All που προέρχεται από έναν συνδυασμό του GSA και του DR2. Η κατανομή με την οποία γίνεται η δειγματοληψία είναι αποτέλεσμα του αθροίσματος μια κανονικής κατανομής με ελλειψοειδής παράλληλα στους άξονες (uncorrelated) και μια αυθαίρετης μονοδιαστατης κανονικής κατανομής με μέσο μηδέν. Αυτή η τελευταία κατανομή συμβάλει στην προσαρμογή της κατεύθυνσης προσθέτοντας τις επιλεγμένες μεταλλάξεις (mutation steps) κατά τις διαδοχικές γενιές. Δηλαδή η γραμμή μεταξύ των γονέων των γενιών γίνεται η βάση για την προσαρμογή της κατεύθυνσης. Λόγο της συνεισφοράς όμως της πρώτης κατανομής ο αλγόριθμος αυτός δεν είναι ανεξάρτητος του συστήματος συντεταγμένων, όπως ανεξάρτητος δεν είναι και ο MSC.

Η εξάρτηση από το σύστημα συντεταγμένων ή όχι, και εν γένη η αμεταβλητότητα ή αναλλοιώτητα (invariance) είναι σημαντική ιδιότητα και είναι απαραίτητη συνθήκη για επιτυχής και αξιόπιστη προσαρμογή, καθώς εξασφαλίζει προβλέψιμη συμπεριφορά, και λειτουργία του αλγορίθμου [19]. Ικανοποιώντας την συνθήκη αυτή αποφεύγεται η προτίμηση σε συγκεκριμένη αναπαράσταση της αντικειμενικής συνάρτησης [11]. Συγκεκριμένα αναφερόμαστε σε αμεταβλητότητα (invariance) γύρω από μεταφορά (translation), περιστροφή και αντικατοπτρισμό (reflection) του χώρου των αντικειμενικών μεταβλητών ώστε αν συμβεί κάποιος από τους μετασχηματισμούς αυτούς να μην επηρεάζεται η συμπεριφορά του αλγορίθμου [18]. Αν κάποια αντικειμενική συνάρτηση F και $q : \mathbb{R}^n \rightarrow \mathbb{R}$,

οι ιδιότητες της αναλλοιωότητας που θέλουμε να πετύχουμε είναι οι γύρω από : Μεταφορά $F: \mathbf{y} \rightarrow q(\mathbf{y} - \boldsymbol{\alpha}), \boldsymbol{\alpha} \in \mathbb{R}^n$, Περιστροφή $F: \mathbf{y} \rightarrow q(\mathbf{U}\mathbf{y}), \mathbf{U}$ ορθογώνιος και Order-preserving $F: \mathbf{y} \rightarrow g(q(\mathbf{y})), g: \mathbb{R} \rightarrow \mathbb{R}$ [19].

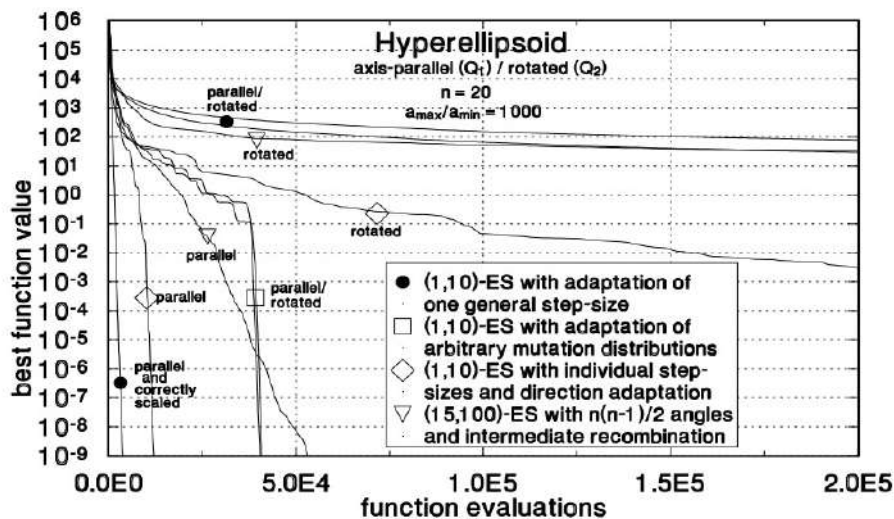
Χωρίς τις ιδιότητες αυτές θα ήταν δύσκολο να βγάλουμε συμπεράσματα για έναν αλγόριθμο και να προβλέψουμε την συμπεριφορά του σε διαφορετικά προβλήματα με διαφορετικές αντικειμενικές συναρτήσεις. Εισάγοντας διάφορες υλοποιήσεις των τελεστών που έχουμε παρουσιάσει αυτές οι ιδιότητες μπορούν να χαθούν ενώ αντίθετα προσεκτικές υλοποιήσεις μπορεί να εισάγουν νέες ιδιότητες αναλλοιωότητας [19].

Η MSC που παρουσιάσαμε δεν ικανοποιεί το κριτήριο της αναλλοιωότητας γύρω από περιστροφές καθώς εξαρτάται από το σύστημα συντεταγμένων. Θέτοντας για αντικειμενική συνάρτηση $Q_2(\mathbf{y})$ ένα αυθαίρετα τυχαία προσανατολισμένο υπερελλειψοειδής με λόγο μεταξύ του μεγαλύτερου και μικρότερου άξονα ως 1000 και $Q_1(\mathbf{y})$ υπερελλειψοειδής παράλληλο με τους άξονες του δοσμένου συστήματος οι N. Hansen, A. Ostermeier και A Gawelczyk [16] σύγκριναν τους AI, AII και τον MSC.

$$Q_1(\mathbf{y}) = \sum_{i=1}^n \left(1000^{\frac{i-1}{n-1}} \langle \mathbf{y}, \mathbf{e}_i \rangle \right)^2$$

$$Q_2(\mathbf{y}) = \sum_{i=1}^n \left(1000^{\frac{i-1}{n-1}} \langle \mathbf{y}, \mathbf{o}_i \rangle \right)^2$$

Όπου $\langle \cdot, \cdot \rangle$ βαθμωτό γινόμενο και τα διανύσματα $\mathbf{o}_1, \dots, \mathbf{o}_n \in \mathbb{R}^n$ και αποτελούν ορθοκανονική βάση με τυχαίο προσανατολισμό.



Εικόνα 13 Απόδοση στρατηγικών εξέλιξης $AI(\square), AII(\diamond), MSC(\nabla)$ και $(1,10)$ -ES (\bullet) με προσαρμογή καθολικού μεγέθους βήματος στα υπερελλειψοειδή Q_1 και Q_2

Τα αποτελέσματα έδειξαν πως οι AII και MSC στην αντικειμενική συνάρτηση Q_2 που ορίσαμε, με άξονες περιστραμένους συγκριτικά με το δοσμένο, έχουν υποδεέστερη απόδοση από τον AI επειδή οι δύο πρώτοι είναι εξαρτημένοι από το σύστημα συντεταγμένων ενώ ο τελευταίος είναι ανεξάρτητος. Παρατηρούμε στο διάγραμμα πως η απόδοση του αλγορίθμου AI που συμβολίζεται με (\square) , δεν μεταβάλλεται είτε αναζήτηση γίνεται σε συνάρτηση παράλληλη με τους άξονες είτε σε περιστραμένους.

4. Στρατηγική εξέλιξης CMA-ES.

Συνεχείς προσπάθειες ανάπτυξης των ES οδήγησαν στο λεγόμενο δεύτερο επίπεδο αποτυχαιοποίησης, την πλήρη αποτυχαιοποίηση και στην CMA-ES. Στόχος είναι ο απευθείας έλεγχος του ρυθμού μεταβολής των στρατηγικών παραμέτρων, η αύξηση της πιθανότητας αναπαραγωγής επιτυχημένων βημάτων και να διατηρούνται οι στρατηγικοί παράμετροι αμετάβλητοι όταν έχουμε τυχαία επιλογή (random selection) πράγμα που συμβαίνει όταν οι τιμές που επιστρέφει η αντικειμενική συνάρτηση είναι ανεξάρτητες των αντικειμενικών μεταβλητών \mathbf{y} . Ο CMA-ES αλγόριθμος προέρχεται ως καρπός και εξέλιξη όλων των προηγούμενων αλγορίθμων που παρουσιάσαμε και είναι ακόλουθος του AI (ή GSA-Generating Set Adaptation) όπου η προσαρμογή του μητρώου συνδυακύμανσης είναι λίγο πιο σύνθετη και προστίθεται έλεγχος του συνολικού μεγέθους βήματος μέσω συσσώρευσης και όχι μέσω αυτοπροσαρμογής και χωρίς τυχαία δειγματοληψία σε αυτό.

Ο CMA-ES είναι ένας $(\mu/\mu_w, \lambda)$ αλγόριθμος και χρησιμοποιεί όπως οι προηγούμενοι δειγματοληψία από πολυμεταβλητή κανονική κατανομή ως

$$\mathbf{x}_k = \mathbf{B}\mathbf{D}\mathbf{z}_k \sim N(\mathbf{0}, \mathbf{C})$$

Όπου ο πίνακας συνδιακύμανσης \mathbf{C} , $\mathbf{z}_k \sim N(\mathbf{0}, \mathbf{I})$ και απο ιδιοανάλυση (ιδιάζουσων Τιμών) του μητρώου \mathbf{C} έχουμε, $\mathbf{C} = \mathbf{B}\mathbf{D}^2\mathbf{B}^T$, με της στήλες του \mathbf{B} να περιέχουν ορθοκανονική βάση με των ιδιοδιανυσμάτων και \mathbf{D} διαγώνιος πίνακας με τις τετραγωνικές ρίζες των ιδιοτιμών.

Γενικά για μια κατανομή $N(\mathbf{m}, \mathbf{C})$, ισχύει

$$\begin{aligned} N(\mathbf{m}, \mathbf{C}) &\sim \mathbf{m} + N(\mathbf{0}, \mathbf{C}) \\ &\sim \mathbf{m} + \mathbf{C}^{\frac{1}{2}} N(\mathbf{0}, \mathbf{I}) \\ &\sim \mathbf{m} + \mathbf{B}\mathbf{D}\mathbf{B}^T N(\mathbf{0}, \mathbf{I}), \text{ με } \mathbf{B}^T N(\mathbf{0}, \mathbf{I}) \sim N(\mathbf{0}, \mathbf{I}) \\ &\sim \mathbf{m} + \mathbf{B}\mathbf{D} N(\mathbf{0}, \mathbf{I}) \end{aligned}$$

Έτσι οι λ απόγονοι παράγονται από τους μ γονείς με την παρακάτω σχέση

$$\mathbf{y}_k^{g+1} = \mathbf{m}^g + \sigma^g \mathbf{x}_k^g \sim N\left(\mathbf{m}^{(g)}, \left((\sigma^{(g)})^2 \mathbf{C}^{(g)}\right)\right)$$

Οι μ καλύτεροι γονείς επιλέγονται και ανασυνδυάζονται ώστε να προκύψει ο ένας γονέας, ο μέσος $\mathbf{m}^{(g+1)}$.

$$\mathbf{m}^{(g+1)} = \sum_{i=1}^{\mu} w_i \mathbf{y}_i^{(g+1)} \cong \mathbf{m}^{(g)} + c_m \sigma^{(g)} \sum_{i=1}^{\mu} w_i \mathbf{x}_i^{(g+1)}$$

Συνήθως εισάγεται μια παράμετρος απόσβεσης (learning rate) c_m στο δεξί μέλος της πάνω ισότητας ώστε να έχουμε $\mathbf{m}^{(g)} + c_m \sigma^{(g)} \sum_{i=1}^{\mu} w_i \mathbf{x}_i^{(g+1)}$, κάτι που κάνει την πάνω ισότητα να μην ισχύει παρά μόνο όταν $c_m = 1$. Αυτή η τιμή, $c_m = 1$ συναντάται συχνά, όμως η τιμή $c_m < 1$ μπορεί να πλεονεκτεί σε συναρτήσεις με μεγάλο θόρυβο [20]. Ορίζεται μια νόρμα (ή μέτρο) μ_{eff} ίση με $\mu_{eff} = \left(\frac{\|\mathbf{w}\|_1}{\|\mathbf{w}\|_2}\right)^2 = \frac{\|\mathbf{w}\|_1^2}{\|\mathbf{w}\|_2^2} = \frac{(\sum_{i=1}^{\mu} |w_i|)^2}{\sum_{i=1}^{\mu} w_i^2} = \frac{1}{\sum_{i=1}^{\mu} w_i^2}$.

Άρα $1 \leq \mu_{eff} \leq \mu$, παρέχοντας έτσι έναν οδηγό για την επιλογή των w_i . Στην ουσία το μ_{eff} γενικεύει την έννοια των μ πατέρων και είναι ένα μέτρο της ποσότητας της πληροφορίας που χρησιμοποιείται. Όπως είδαμε προηγουμένως για ίσους συντελεστές βαρύτητας στις $(\mu/\mu_l, \lambda)$, με τον ανασυνδυασμό μειώνεται η διακύμανση κατά έναν παράγοντα μ , ακολούθως εδώ στις $(\mu/\mu_w, \lambda)$ μειώνεται κατά μ_{eff} . Μια ικανοποιητική τιμή συνήθως για το

μ_{eff} λαμβάνεται ως $\mu_{eff} \approx \lambda/4$. Οι δημιουργοί του αλγορίθμου προτείνουν ως προκαθορισμένη τιμή $w_i = \frac{\ln \frac{\lambda+1}{2} - \ln i}{\sum_{j=1}^{\mu} (\ln \frac{\lambda+1}{2} - \ln j)}$, $i \in \{1, \dots, \mu\}$ [15].

Ο CMA-ES χρησιμοποιεί όπως ο DR2 την συσσώρευση , το μονοπάτι εξέλιξης, για την προσαρμογή του μεγέθους του βήματος (ή δύναμης μετάλλαξης) σ . Μια από τις διαφοροποιήσεις που εισάγει ο CMA-ES είναι πως προσαρμογή της πολυμεταβλητής κατανομής διαχωρίζεται σε μια ξεχωριστή προσαρμογή του μεγέθους του βήματος και ξεχωριστή προσαρμογή του πίνακα συνδιακύμανσης, ενώ η συσσώρευση του μονοπατιού εξέλιξης χρησιμοποιείται και για τον υπολογισμό του μητρώου συνδιακύμανσης. Ένας λόγος για τον διαχωρισμό αυτό είναι πως είναι απαραίτητο ο ρυθμός μεταβολής του καθολικού βήματος να είναι μεγαλύτερος από τον ρυθμό προσαρμογής του μητρώου συνδιακύμανσης [14], ώστε ο αλγόριθμος αρχικά να έχει αρκετή εξερεύνηση στον χώρο αναζήτησης .

Για τον υπολογισμό του μητρώου συνδιακύμανσης από δοσμένο πληθυσμό, θεωρώντας πως ο πληθυσμός είναι τυχαίες μεταβλητές (και όχι παράγωγο δειγματοληψίας) μπορούμε να χρησιμοποιήσουμε την παρακάτω εκτιμήτρια συνάρτηση όπου αντί για τον μέσο του παρόντος πληθυσμού ($\frac{1}{\lambda} \sum_{j=1}^{\lambda} (\mathbf{y}_j^{(g+1)})$), χρησιμοποιούμε τον πραγματικό μέσο ($\mathbf{m}^{(g)}$) της δειγματοληπτικής κατανομής εκφράζοντας έτσι μέσω του $\mathbf{C}_{\lambda}^{(g+1)}$ την διακύμανση των δειγματοληπτικών βημάτων ($\mathbf{y}_i^{(g+1)} - \mathbf{m}^{(g)}$).

$$\mathbf{C}_{\lambda}^{(g+1)} = \frac{1}{\lambda} \sum_{i=1}^{\lambda} (\mathbf{y}_i^{(g+1)} - \mathbf{m}^{(g)})(\mathbf{y}_i^{(g+1)} - \mathbf{m}^{(g)})^T$$

Υπολογίζοντας έναν “καλύτερο” πίνακα συνδιακύμανσης μετά την επιλογή

$$\mathbf{C}_{\mu}^{(g+1)} = \frac{1}{\lambda} \sum_{i=1}^{\mu} (\mathbf{y}_{i:\lambda}^{(g+1)} - \mathbf{m}^{(g)})(\mathbf{y}_{i:\lambda}^{(g+1)} - \mathbf{m}^{(g)})^T$$

Η συνάρτηση για την $\mathbf{C}_{\mu}^{(g+1)}$ είναι εκτιμήτρια για την κατανομή των επιλεγμένων βημάτων και χρησιμοποιώντας τον $\mathbf{C}_{\mu}^{(g+1)}$ για την δειγματοληψία τείνουν να αναπαράγονται τα επιτυχημένα επιλεγμένα βήματα [20]

Για να πετύχει όμως ο αλγόριθμος ταχεία απόδοση το μέγεθος του πληθυσμού θα πρέπει να είναι μικρό. Επίσης μετά την διαδικασία της επιλογής ελαττώνεται ακόμη περισσότερο το δείγμα της εκτιμήτριας. Έτσι η προηγούμενη εξίσωση δεν μπορεί να παράσχει αξιόπιστη εκτιμήτρια . Για να διορθωθεί το πρόβλημα αυτό μπορούν να χρησιμοποιηθεί πληροφορία από προηγούμενες γενιές . Μετά το πέρας μερικών γενιών ο μέσος του μητρώου συνδιακύμανσης γίνεται μία αξιόπιστη εκτιμήτρια συνάρτηση για το μητρώο συνδιακύμανσης

$$\mathbf{C}^{(g+1)} = \frac{1}{g+1} \sum_{i=0}^g \frac{1}{\sigma_i^2} \mathbf{C}_{\mu}^{i+1}$$

Όπου $1/\sigma_i^2$ κανονικοποίηση του μέγεθος βήματος για τις διάφορες γενιές ώστε να είναι συγκρίσιμοι οι πίνακες.

$$\text{Αν } \mathbf{C}^{(0)} = \mathbf{I} \text{ τότε } \mathbf{C}^{(g+1)} = (1 - c_{\mu})\mathbf{C}^{(g)} + c_{\mu} \frac{1}{(\sigma^{(g)})^2} \mathbf{C}_{\mu}^{g+1}$$

$$= (1 - c_\mu)C^{(g)} + c_\mu \sum_{i=1}^{\mu} w_i x_{i:\lambda}^{(g+1)} (x_{i:\lambda}^{(g+1)})^T$$

Το $c_\mu \in [0,1]$ ονομάζεται ρυθμός μάθησης (learning rate) και είναι μια παράμετρος χαλάρωσης. Οι καλές τιμές για τον ρυθμό μάθησης δεν εξαρτώνται από την αντικειμενική συνάρτηση και μια συντηρητική προσέγγιση για μια καλή τιμή είναι $c_\mu \approx \mu_{eff}/n^2$ [20].

Με την διαμόρφωση της παραπάνω εξίσωσης για τον υπολογισμό του μητρώου συνδιακύμανσης δεν λαμβάνουμε υπόψιν το πρόσημο των βημάτων, καθώς $xx^T = -x(-x)^T$. Για να επανεισάγουμε αυτή την πληροφορία χρησιμοποιούμε επιπλέον μια τεχνική το μητρώο συνδιακύμανσης παρόμοια με αυτή που εφαρμόστηκε στον αλγόριθμο *AI* χρησιμοποιώντας διάνυσμα συσσώρευσης. Θεωρούμε την ακολουθία διαδοχικών βημάτων που πραγματοποιεί η στρατηγική κατά τις γενιές και την ονομάζουμε ως μονοπάτι εξέλιξης. Το άθροισμα των βημάτων ονομάζεται συσσώρευση.

$$\mathbf{p}_c^{(g+1)} = (1 - c_c)\mathbf{p}_c^{(g)} + \sqrt{c_c(2 - c_c)\mu_{eff}} \frac{(\mathbf{m}^{(g+1)} - \mathbf{m}^{(g)})}{c_m \sigma^{(g)}}$$

Ο όρος $\sqrt{c_c(2 - c_c)\mu_{eff}}$ πραγματοποιεί κανονικοποίηση ώστε να διατηρείται σταθερή η διακύμανση της κατανομής των $\mathbf{p}^g, \mathbf{p}^{g-1}$ και στον γραμμικό συνδυασμό τους και η σταθερά c_c είναι πάλι ένας βαθμός μάθησης.

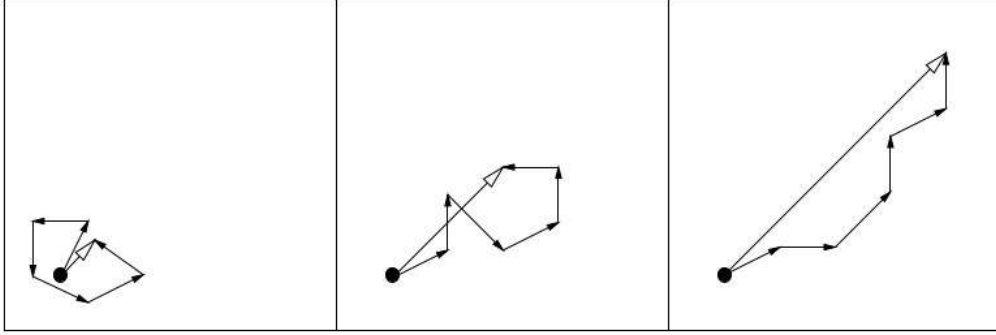
Έτσι η πρώτη τάξης (rank one) προσαρμογή του μητρώου συνδιακύμανσης είναι η ακόλουθη

$$C^{(g+1)} = (1 - c_1)C^{(g)} + c_1 \mathbf{p}_c^{(g+1)} \mathbf{p}_c^{(g+1)T}$$

Χρησιμοποιώντας το μονοπάτι εξέλιξης για την προσαρμογή του μητρώου συνδιακύμανσης έχουμε μεγάλη απόδοση ακόμα και όταν έχουμε μικρό πληθυσμό καθώς εκμεταλλευόμαστε τις συσχετίσεις μεταξύ των διαδοχικών βημάτων. Άρα ο πίνακας συνολικά προσαρμόζεται λόγω της κατανομής της προηγούμενης γενιάς, του πληθυσμού της γενιάς και των προηγούμενων επιτυχημένων βημάτων όπως φαίνεται στην σχέση :

$$C^{(g+1)} = \left(1 - c_1 - c_\mu \sum w_j\right) C^{(g)} + c_1 \mathbf{p}_c^{(g+1)} \mathbf{p}_c^{(g+1)T} + c_\mu \sum_{i=1}^{\mu} w_i x_{i:\lambda}^{(g+1)} (x_{i:\lambda}^{(g+1)})^T$$

Η συσσώρευση χρησιμοποιείται επίσης για την προσαρμογή του μεγέθους του βήματος. Τα επιτυχημένα διαδοχικά βήματα αθροίζονται και συγκρίνονται με την αναμενόμενη τιμή του μεγέθους ενός διανύσματος από την αμερόληπτη κατανομή $N(0, I)$. Αν το άθροισμα αυτό, που ονομάζεται μονοπάτι εξέλιξης, είναι συγκριτικά μικρό τότε σημαίνει πως τα διαδοχικά βήματα είναι συσχετισμένα αντιπαράλληλα και αλληλοαναιρούνται στο άθροισμα, οπότε το μέγεθος του βήματος θα πρέπει να μειωθεί. Αντίθετα αν το μονοπάτι εξέλιξης είναι μεγάλο τα διαδοχικά βήματα είναι προς την ίδια κατεύθυνση και η απόσταση προς την κατεύθυνση αυτή μπορεί να καλυφθεί με λιγότερα βήματα τα οποία είναι μεγαλύτερα και έτσι αυξάνεται το μέγεθος βήματος.



Εικόνα 14 [20]. Συσσώρευση και συσχέτιση διαδοχικών βημάτων.

Το μονοπάτι εξέλιξης για τον έλεγχο του βήματος \mathbf{p}_σ , διαφέρει από αυτό που χρησιμοποιείται για την προσαρμογή του μητρώου διακύμανσης \mathbf{p}_c . Θέλουμε να προσαρμόσουμε το μέγεθος βήματος ανεξάρτητα από το μητρώο συνδιακύμανσης και η αναμενόμενη τιμή του μήκους του \mathbf{p}_σ να είναι ανεξάρτητη την κατεύθυνσης που εισάγει το μητρώο συνδιακύμανσης ώστε $\mathbf{p}_\sigma \sim N(\mathbf{0}, \mathbf{I})$ αν $\mathbf{p}_\sigma^{(0)} \sim N(\mathbf{0}, \mathbf{I})$. Άρα

$$\mathbf{p}_\sigma^{(g+1)} = (1 - c_\sigma)\mathbf{p}_\sigma^{(g)} + \sqrt{c_\sigma(2 - c_\sigma)\mu_{eff}}\mathbf{C}^{(g)^{-\frac{1}{2}}}\frac{(\mathbf{m}^{(g+1)} - \mathbf{m}^{(g)})}{c_m\sigma^{(g)}}$$

Ο μετασχηματισμός $\mathbf{C}^{(g)^{-\frac{1}{2}}}$ ανάγει τα διανύσματα στην κατανομή $N(\mathbf{0}, \mathbf{I})$ αναιρώντας την μεροληψία που εισάγει η προσαρμογή του μητρώου συνδιακύμανσης από τα προηγούμενα βήματα. Δηλαδή έχουμε $\mathbf{C}^{(g)^{-\frac{1}{2}}} = \mathbf{B}^{(g)}\mathbf{D}^{(g)^{-1}}\mathbf{B}^{(g)T}$ και $N(\mathbf{0}, \mathbf{C}) \sim \mathbf{BD}N(\mathbf{0}, \mathbf{I})$ οπότε $\mathbf{D}^{-1}\mathbf{B}^{-1}N(\mathbf{0}, \mathbf{C}) \sim N(\mathbf{0}, \mathbf{I})$, με $\mathbf{B}^{-1} = \mathbf{B}^T$.

Για να προσαρμόσουμε το μέγεθος βήματος σ , συγκρίνουμε το μέγεθος του μονοπατιού εξέλιξης \mathbf{p}_σ με την αναμενόμενη τιμή $E\|N(\mathbf{0}, \mathbf{I})\|$

$$\ln \sigma^{(g+1)} = \ln \sigma^{(g)} + \frac{c_d}{d_d} \left(\frac{\|\mathbf{p}_\sigma^{(g+1)}\|}{E\|N(\mathbf{0}, \mathbf{I})\|} - 1 \right)$$

5. Νοημοσύνη Σμήνους

Κατά την εποχή που αναπτύσσονταν οι εξελικτικοί αλγόριθμοι όπως οι στρατηγικές εξέλιξης δύο Αμερικάνοι επιστήμονες ο J. Kennedy και ο R. C. Eberhart, ένας ψυχολόγος και ένας ηλεκτρολόγος μηχανικός αντίστοιχα, μελετούσαν την προσομοίωση της κοινωνικής συμπεριφοράς του ανθρώπου. Εφαλτήριο και βάση ήταν η συλλογική συμπεριφορά οργανισμών όπως των ψαριών ή των πτηνών τα οποία οργανώνονται σε κοπάδια, σε σμήνη, ή εντόμων όπως τα μυρμήγκια ή οι μέλισσες. Παρουσιάζει εξαιρετικό ενδιαφέρον το πως οργανισμοί οι οποίοι δεν είναι ιδιαίτερα ευφυής έχοντας περιορισμένες δυνατότητες επικοινωνίας, καταφέρνουν εντούτοις συναρπαστικές ευφυείς συλλογικές συμπεριφορές όταν οργανώνονται σε “σμήνη” πράγμα που μπορεί να οριστεί σαν ευφυΐα σμήνους (Swarm Intelligence) ή συλλογική ευφυΐα .

Για παράδειγμα οι τερμίτες χτίζουν μερικές από τις πιο περίπλοκες και μεγάλες δομές που συναντώνται στη φύση με χαρακτηριστικά που μπορούν να ρυθμίζουν τις κλιματολογικές συνθήκες εντός της αποικίας πράγμα που έχει εμπνεύσει σχεδιασμό κτηρίων με βιομιμητισμό βασιζόμενο σε αυτές τις αποικίες [21]. Η κίνηση των μυρμηγκιών ατομικά είναι σχεδόν τυχαία με τάση να κινούνται προς μονοπάτια από τα οποία έχουν περάσει άλλα μυρμήγκια, παρόλα αυτά συλλογικά παρουσιάζουν μοναδικά χαρακτηριστικά ένα από τα οποία είναι η ικανότητα επιλογής του συντομότερου μονοπατιού από κάποιον στόχο όπως το φαγητό ή ακόμα η προτίμηση προς της πιο πλούσια πηγή φαγητού όταν δύο πηγές είναι παρούσες. Όλα αυτά πραγματοποιούνται διαμέσου ενός μηχανισμού επικοινωνίας, τις φερομόνες που εναποθέτουν αυτά τα έντομα, και ακολουθώντας μερικούς απλούς κανόνες συμπεριφοράς [22]. Τα κοπάδια ψαριών και τα σμήνη πουλιών προσαρμόζουν την κίνηση τους για να αποφύγουν θηρευτές, να εντοπίσουν τροφή και να ρυθμίσουν τις κλιματολογικές συνθήκες στις οποίες βρίσκονται όπως η θερμοκρασία [23]. Συχνά τα σμήνη σχηματίζουν εντυπωσιακά και πολύπλοκα σχήματα και μοτίβα που συνεχώς μεταβάλλονται ενώ παραμένει η συνοχή του σμήνους.



Εικόνα 15 [24].Σμήνη πουλιών.

Αυτή η εντυπωσιακή συλλογική συμπεριφορά οδήγησε ερευνητές από διαφορετικά πεδία να μελετήσουν τα σμήνη και να προσπαθήσουν να αναπαράξουν σε προσημειώσεις τις κινήσεις τους είτε επειδή είχαν κάποια εφαρμογή στο πεδίο τους (Craig W. Reynolds – Computer Graphics) είτε επειδή παρόμοια μοτίβα εμφανίζονταν σε προβλήματα που προσπαθούσαν να λύσουν (Frank Heppner and Conway’s Game of Life) [22]. Έπειτα από παρατηρήσεις στην συμπεριφορά των σμηνών αυτών ώστε να μπορέσουν να αναπαράξουν την κίνηση τους, σημείωσαν πως οι τάσεις που κυριαρχούσαν ήταν η έλξη μεταξύ των μελών (παράλληλα όμως με την αποφυγή της σύγκρουσης με τα υπόλοιπα μέλη σε περίπτωση που πλησίαζαν αρκετά), η προσαρμογή της ταχύτητας σε σχέση με αυτή από τα γειτονικά μέλη και η τάση των πτηνών να κινούνται προς το κέντρο του σμήνους όπως το αντιλαμβάνονται. Το τελευταίο ίσως έχει να κάνει με την μειωμένη πιθανότητα να γίνει θήραμα ένα μέλος όταν βρίσκεται στο κέντρο του σμήνους [22]. Στην συνέχεια πρόσθεσαν έναν στόχο (όπως σπόροι φαγητού) και τυχαιότητα που μπορεί να προκύψει για παράδειγμα από ριπές ανέμου. Είναι χαρακτηριστικό πως δεν υπάρχει κάποιο αρχηγικό μέλος που καθορίζει την συμπεριφορά και ένα μέλος του σμήνους ήταν δυνατό να εκκινήσει έναν ελιγμό οποιαδήποτε στιγμή [22]. Ομοίως και στην συμπεριφορά του σμήνους των προηγούμενων οργανισμών που αναφέραμε είναι απύουσα μια συγκεντρωτική δομή ελέγχου. Ίσως το πιο σημαντικό χαρακτηριστικό είναι πως τα άτομα που αποτελούν τον πληθυσμό του κοπαδιού ή του σμήνους αλληλοεπιδρούν μοιράζοντας κατά κάποιο τρόπο με τα υπόλοιπα άτομα αυτό που ανακαλύπτουν γύρω από τον στόχο όπως το φαγητό ή τις προηγούμενες εμπειρίες τους.

Οι J.Kennedy και R. Eberhart [23] έστησαν μια προσομοίωση αρχικοποιώντας στο επίπεδο έναν πληθυσμό που αντιπροσωπεύει τα πουλιά ενός σμήνους (σωματίδια) και έδωσαν σε κάθε σωματίδιο μια αρχική ταχύτητα. Ο αλγόριθμος τους οδηγούσε τα σωματίδια να προσαρμόζουν την ταχύτητα τους σε αυτή του πιο κοντινού γείτονα και προσέθεσαν και έναν βαθμό τυχαιότητας με αποτέλεσμα μια χορογραφία πιστή στις κινήσεις ενός πραγματικού σμήνους. Προσθέτοντας στην συνέχεια μια έλξη στα σωματίδια προς κάποιο σημείο όπως κάτι που θα μπορούσε να αντιπροσωπεύσει έναν πόλο έλξης για τα σωματίδια του σμήνους (όπως μια φωλιά ή φαγητό) τα σωματίδια προσγειώνονται στο σημείο αυτό. Αυτό ώθησε τους ερευνητές να εξετάσουν αν ο αλγόριθμος αυτός έχει δυνατότητες χρήσης σε προβλήματα βελτιστοποίησης. Τα σωματίδια πλέον δεν γνωρίζουν την θέση του φαγητού αλλά εισάγεται μια συνάρτηση που εκφράζει την απόσταση του σωματιδίου από το φαγητό με σκοπό την ελαχιστοποίηση της απόστασης.

$$E_{val} = \sqrt{(presentx - target)^2} + \sqrt{(presenty - target)^2}$$

Επιπλέον κάθε σωματίδιο δύναται να θυμάται την καλύτερη τιμή E_{val} που έχει συναντήσει καθώς και την θέση στην οποία τη συνάντησε, $pbest$ (personal best). Καθώς το σωματίδιο κινείται στον χώρο προσαρμόζει την ταχύτητα του αναφορικά με το που βρίσκεται σε σχέση με το $pbest$. Αν στις δύο συνιστώσες ξεχωριστά έχει υπερκεράσει ή βρίσκεται όπισθεν του σημείο τότε αφαιρείται ή προστίθεται αντίστοιχα ένας σταθμισμένος τυχαίος παράγοντας. Επίσης η συνολικά καλύτερη θέση που έχει βρεθεί από κάποιο σωματίδιο, $gbest$ (global best), γνωστοποιείται σε κάθε σωματίδιο. Έτσι αντίστοιχα προσαρμόζεται η ταχύτητα πάλι βάση της θέσης σχετικά με τη $gbest$ θέση.

Τα αποτελέσματα του αλγορίθμου ήταν συναρπαστικά! Τα σωματίδια μετά από λίγες επαναλήψεις καταλήγουν ταχύτητα στον στόχο. Οπότε έγινε φανερό στους ερευνητές ότι ο αλγόριθμος μπορούσε εύκολα να βελτιστοποιήσει απλές γραμμικές συναρτήσεις και πως είχε

προοπτικές στον χώρο της βελτιστοποίησης. Αυτή η μέθοδος βελτιστοποίησης ονομάστηκε Particle Swarm Optimization (PSO).

5.1. Αλγόριθμος PSO

Σήμερα ο αλγόριθμος PSO θεωρείται πως ανήκει στην οικογένεια των εξελικτικών αλγορίθμων και συγκεκριμένα στην υποκατηγορία Νοημοσύνης Σμήνους (Swarm Intelligence) στην οποία κατέχει εξέχουσα θέση.

Όντας ένας αλγόριθμος βελτιστοποίησης έχει ως στόχο να βελτιστοποιήσει μια αντικειμενική συνάρτηση. Δηλαδή αν θεωρήσουμε πως έχουμε ένα πρόβλημα ελαχιστοποίησης με n -διάστατο χώρο αναζήτησης με πραγματικές τιμές το πρόβλημα είναι

$$\min F(\mathbf{y}), \mathbf{y} \in \mathcal{Y}$$

όπου $\mathcal{Y} \in \mathbb{R}^n$. Η μόνη παραδοχή που γίνεται για την αντικειμενική συνάρτηση είναι πως ορίζεται και έχουμε την τιμή $F(\mathbf{y})$ για κάθε $\mathbf{y} \in \mathcal{Y}$. Η συνέχεια της συνάρτησης και η διαφορισμότητα δεν είναι αναγκαίες.

Όπως και στις στρατηγικές εξέλιξης ο μηχανισμός αναζήτησης βασίζεται σε έναν πληθυσμό που αντιπροσωπεύουν σημεία στον χώρο αναζήτησης άρα και υποψήφιες λύσεις που εδώ καλούνται σωματίδια (particles). Αυτά τα σωματίδια διατηρούν στην μνήμη τους την καλύτερη θέση που έχουν επισκεφθεί στην ιστορία τους, την θέση δηλαδή που έχει δώσει για τα ίδια την καλύτερη τιμή της αντικειμενικής συνάρτησης. Αυτή η θέση σημειώνεται συνήθως ως ***pbest*** (personal best). Επίσης διατηρείται στην μνήμη τους η βέλτιστη θέση που έχει βρεθεί από το σμήνος συνολικά ***gbest*** (global best), η οποία μπορεί να έχει βρεθεί από οποιοδήποτε σωματίδιο του σμήνους. Σχετικά με το τελευταίο τονίζουμε πως με αυτό τον μηχανισμό πρωτοεμφανίστηκε ο PSO από τους δημιουργούς του ο οποίος συνήθως καλείται Κανονικός PSO (Canonical PSO). Επειδή ο PSO είναι ιδιαίτερα δημοφιλής και εξαιρετικά απλός έχουν προκύψει διάφορες παραλλαγές που σκοπεύουν να βελτιώσουν τα χαρακτηριστικά του. Μερικές από τις παραλλαγές επικεντρώνονται στην αλλαγή του τελευταίου μηχανισμού που αναφέραμε όπου σε αυτές δεν είναι διαθέσιμη στα σωματίδια η καλύτερη θέση που έχει βρεθεί από το σμήνος αλλά η καλύτερη θέση που έχει βρεθεί από την “γειτονιά” που ανήκει το σωματίδιο. Επιπρόσθετα σε μερικούς αλγόριθμους μπορεί η πληροφορία που αξιοποιείται να μην είναι απλά η καλύτερη θέση της γειτονιάς. Τα σωματίδια που ανήκουν στην ίδια γειτονιά μπορούν να ανταλλάξουν πληροφορία μεταξύ τους γνωστοποιώντας την καλύτερη θέση στην οποία έχουν βρεθεί. Η τοπολογική γειτονιά αυτή μπορεί να οριστεί με πολλούς τρόπους και θα εξεταστεί στα επόμενα κεφάλαια. Τα σωματίδια καθώς αλλάζουν θέση μεροληπτούν προς τις δύο αυτές θέσεις ***pbest*** και ***gbest***. Δηλαδή προς θέσεις που έχουν ανακαλύψουν τα ίδια τα σωματίδια καθώς και σε υποσχόμενες θέσεις που έχουν ανακαλύψει τα υπόλοιπα σωματίδια του σμήνους.

Αν n η διάσταση του χώρου αναζήτησης, N το πλήθος των σωματιδίων και A_i το i -οστό σωματίδιο τότε το σύνολο που αποτελεί το σμήνος είναι το

$$\mathcal{S} = \{A_1, A_2, \dots, A_N\}$$

και

$$A_i^{(t)} = \langle \mathbf{y}_i^{(t)}, \mathbf{v}_i^{(t)}, \mathbf{p}_i^{(t)}, NB_i^{(t)} \rangle$$

το κάθε σωματίδιο. Κάθε σωματίδιο αποτελείται από τα διανύσματα $\mathbf{y}_i^{(t)}$, $\mathbf{v}_i^{(t)}$ και $\mathbf{p}_i^{(t)}$. Το διάνυσμα $\mathbf{y}_i^{(t)} = (y_{i1}^{(t)}, y_{i2}^{(t)}, \dots, y_{in}^{(t)}) \in \mathcal{Y}$ ορίζει κάθε στιγμή, t , την θέση στον χώρο αναζήτησης και το διάνυσμα $\mathbf{v}_i^{(t)} = (v_{i1}^{(t)}, v_{i2}^{(t)}, \dots, v_{in}^{(t)})$ που καλείται ταχύτητα του σωματιδίου που μεταβάλλει την θέση του σωματιδίου. Το τρίτο διάνυσμα είναι η θέση **pbest** που αναφέραμε μέχρι την επανάληψη t , $\mathbf{p}_i^{(t)} = (p_{i1}^{(t)}, p_{i2}^{(t)}, \dots, p_{in}^{(t)}) \in \mathcal{Y}$ [25].

Οπότε $\mathbf{p}_i^{(t)} = \arg \min_{\{\mathbf{y}_i^{(k)}, k \leq t\}} f(\mathbf{y}_i^{(k)})$. Με $NB_i^{(t)}$ συμβολίζεται η γειτονιά του i -οστού σωματιδίου και $NB_i^{(t)}$ περιέχει τους δείκτες των σωματιδίων την γειτονιάς. Αν $I = \{1, 2, \dots, N\}$ το σύνολο των δεικτών του πληθυσμού, τότε $NB_i^{(t)} \subseteq I$ και $\mathbf{p}_{gi}^{(t)} = \arg \min_{\{\mathbf{p}_j^{(t)}, j \in NB_i^{(t)}\}} f(\mathbf{p}_j^{(t)})$ [4].

Σε κάθε επανάληψη ενημερώνονται οι συνιστώσες της ταχύτητας και οι θέσεις ως εξής

$$\begin{aligned} v_{ij}^{(t+1)} &= w v_{ij}^{(t)} + C_1 (p_{ij}^{(t)} - y_{ij}^{(t)}) + C_2 (p_{gij}^{(t)} - y_{ij}^{(t)}) \\ y_{ij}^{(t+1)} &= y_{ij}^{(t)} + v_{ij}^{(t+1)} \end{aligned}$$

Οι όροι C_1 και C_2 ακολουθούν μια ομοιόμορφη κατανομή, $C_1 \sim \mathcal{U}(0, c_1)$, $C_2 \sim \mathcal{U}(0, c_2)$, και οι σταθερές c_1, c_2 καλούνται σταθερές επιτάχυνσης και ελέγχουν το μέγεθος της έλξης προς τα $\mathbf{p}_i^{(t)}$ και $\mathbf{p}_{gi}^{(t)}$ αντίστοιχα [4]. Ο όρος $(p_{ij}^{(t)} - y_{ij}^{(t)})$ καλείται όρος προσωπικής γνώσης ή εμπειρίας (cognitive term) και επί της ουσίας ποσοτικοποιεί την απόδοση του σωματιδίου σε σύγκριση με προηγούμενες αποδόσεις, έχοντας ως αποτέλεσμα την εισαγωγή της τάσης να επανέρχεται σε προηγούμενες καλές θέσεις που ικανοποίησαν το σωματίδιο [5]. Ο όρος $(p_{gij}^{(t)} - y_{ij}^{(t)})$ καλείται όρος κοινωνικότητας (social term) και αντίστοιχα ποσοτικοποιεί την απόδοση του σωματιδίου σε σύγκριση με ένα σύνολο από σωματίδια. Με w συμβολίζεται ένας όρος που καλείται συνήθως αδράνεια και αποσβένει την συμβολή των προηγούμενων τιμών κατά τις επαναλήψεις. Αδράνεια λέγεται επειδή αν σκεφτούμε πως το σωματίδιο ταξιδεύει στον χώρο αναζήτησης, το βάρος της αδράνειας έχει ως αποτέλεσμα να μην αλλάζει ξαφνικά η κατεύθυνση του αλλά να μεταβάλλεται η τροχιά του σταδιακά.

Οι καλύτερες προσωπικές θέσεις ανανεώνονται ακολούθως με την σχέση :

$$\mathbf{p}_i^{(t)} = \begin{cases} \mathbf{y}_i^{(t+1)}, & \text{εάν } f(\mathbf{y}_i^{(t+1)}) \leq f(\mathbf{p}_i^{(t)}) \\ \mathbf{p}_i^{(t)}, & \text{διαφορετικά} \end{cases}$$

<p>PSO ψευδοκώδικας</p> <p>Αρχικοποίηση $\mathbf{y}_i^{(0)}, \mathbf{v}_i^{(0)}, \mathbf{p}_i^{(0)}, \mathbf{p}_{gi}^{(0)}$</p> <p>$t \leftarrow 0$</p> <p>Επανάλαβε</p> <p>$t \leftarrow t + 1$</p> <p>για $i = 1$ έως N επανάλαβε</p> <p>$\varphi_i \leftarrow f(\mathbf{y}_i)$</p> <p>εάν $\varphi_i \leq f(\mathbf{p}_i^{(t)})$</p> <p>$\mathbf{p}_i^{(t)} \leftarrow \mathbf{y}_i$</p> <p>αλλιώς</p> <p>$\mathbf{p}_i^{(t)} \leftarrow \mathbf{p}_i^{(t)}$</p>
--

<p>τέλος-εάν</p> $v_i^{(t+1)} = \chi v_i^{(t)} + c_1 (p_i^{(t)} - y_i^{(t)}) + c_2 (p_{g_i}^{(t)} - y_i^{(t)})$ $y_i^{(t)} = y_i^{(t)} + v_i^{(t+1)}$ <p>για-τέλος</p> <p>Μέχρι (Κριτήριο Τερματισμού)</p>

Πίνακας 8. Ψευδοκώδικας για τον Αλγόριθμο PSO.

5.2. Παράμετροι Αλγόριθμου PSO

Ίσως η πιο σημαντική παράμετρος είναι το βάρος αδράνειας που εφαρμόζεται στην ταχύτητα της προηγούμενης επανάληψης. Γενικά χρησιμοποιείται για να προσφέρει μια ισορροπία στη έκταση της αναζήτησης της λύσης στο χώρο αναζήτησης, άρα στην εξερεύνηση. Μεγάλη τιμή στην παράμετρο αυτή θα οδηγήσει σε μεγάλη περιοχή αναζήτησης, καθολική αναζήτηση (global search), ενώ μια μικρή τιμή οδηγεί σε πιο τοπικού χαρακτήρα αναζήτηση (local search) [26]. Προσαρμόζοντας τις τιμές κατάλληλα επιδιώκουμε να πετύχουμε μια ισορροπία γύρω από τις έννοιες την εξερεύνησης (exploration) και εκμετάλλευσης (exploitation) που αναφέραμε προηγουμένως. Οπότε μια μέθοδος ιδανικά θα ξεκινούσε με μεγάλο βάρος επιδιώκοντας καθολική αναζήτηση στον χώρο και στην πορεία με την πάροδο των επαναλήψεων θα θέλαμε να μειωθεί το βάρος για να έχουμε τοπική αναζήτηση και να οδηγηθούμε σε σύγκλιση.

Ένας από τους τρόπους που εφαρμόζεται η προσαρμογή αυτή είναι μέσω μιας γραμμικής μείωσης του βάρους με τις επαναλήψεις ανάμεσα από δύο ακραίες τιμές.

$$w^{(t)} = w_{max} - \frac{t}{t_{max}} (w_{max} - w_{min})$$

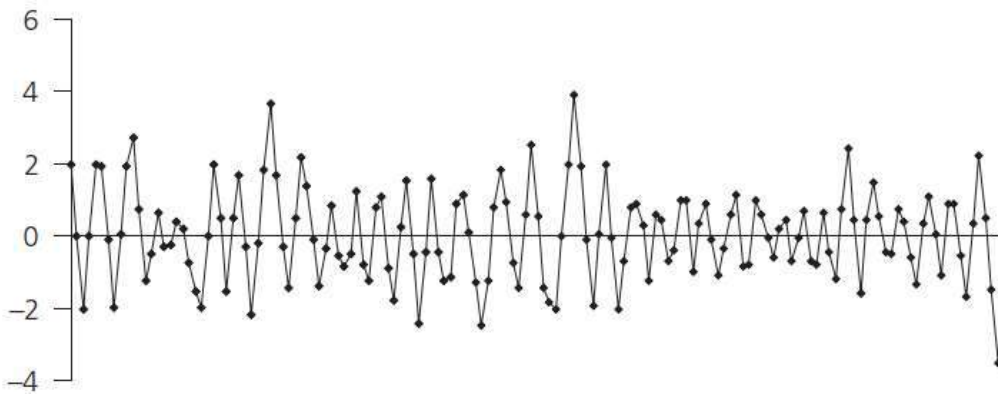
Εμπειρικές μελέτες έδειξαν πως καλές τιμές για τις ακραίες τιμές είναι $w_{max} = 0.9$ και $w_{min} = 0.4$. Παρά την τακτική αυτή οι ταχύτητες μπορεί να αυξηθούν πολύ ειδικά αν τα σωματίδια βρίσκονται μακριά από τα τοπικά και καθολικά βέλτιστα σημεία κάτι που μπορεί να οδηγήσει τα σωματίδια εκτός του χώρου αναζήτησης. Προκύπτει οπότε ανάγκη για περιορισμό της ταχύτητας γύρω από μία μέγιστη τιμή η οποία εξαρτάται από το πρόβλημα βελτιστοποίησης [27]. Συνήθως σε προβλήματα όπου ο χώρος αναζήτησης έχει πάνω και κάτω όρια, η μέγιστη ταχύτητα για κάθε διάσταση ορίζεται ως ένα ποσοστό του του μεγέθους του χώρου στη διάσταση αυτή, ήτοι $V_{maxj} = \delta(y_{maxj} - y_{mi_j})$, όπου y_{maxj}, y_{mi_j} τα άνω και κάτω όρια αντίστοιχα. Υπάρχουν πολλοί τρόποι ελέγχου της ταχύτητας όπως αντικατάσταση της ταχύτητας του σωματιδίου με την μέγιστη ταχύτητα που έχουμε ορίσει μόλις η ταχύτητα του σωματιδίου την ξεπερνά. Ο περιορισμός της ταχύτητας αν και δεν περιορίζει την θέση ενός σωματιδίου εντούτοις την μεταβάλλει.

Στο παρακάτω διάγραμμα εμφανίζεται η τροχιά ενός σωματιδίου αν δεν υπάρχει περιορισμός ταχύτητας και βάρος αδράνειας, θεωρώντας επίσης τα υπόλοιπα σωματίδια σταθερά. Η ταχύτητα μετά από 150 επαναλήψεις αρχίζει να αυξάνεται ανεξέλεγκτα



Εικόνα 16 [22]. “Εκρηξη” θέση σωματιδίου μετά από λίγες επαναλήψεις όταν δεν υπάρχει περιορισμός στην ταχύτητα.

Στο παρακάτω διάγραμμα βλέπουμε με τις ίδιες παραδοχές την τροχιά του σωματιδίου αυτή την φορά με περιορισμό στην μέγιστη ταχύτητα αλλά χωρίς βάρους αδράνειας



Εικόνα 17 [22]. Θέση σωματιδίου αυτή την φορά με περιορισμό της ταχύτητας χωρίς βάρους αδράνειας όμως .

Ο Frans van den Bergh [28] ερεύνησε την σύγκλιση του αλγορίθμου μελετώντας την τροχιά ενός σωματιδίου σε μία διάσταση θεωρώντας σταθερά τα υπόλοιπα σωματίδια άρα και με σταθερές τις $pbest$ και $gbest$ θέσεις. Ερευνώντας την σύγκλιση της σειράς της αναδρομικής σχέσης της θέσης του σωματιδίου συμπέραναν πως σύγκλιση επιτυγχάνεται εάν $w > \frac{1}{2}(c_1 + c_2) - 1$, όπου c_1 και c_2 τα επάνω όρια της $c_1 \sim \mathcal{U}(0, c_1)$ $c_2 \sim \mathcal{U}(0, c_2)$ [28]. Χρησιμοποιώντας κατάλληλες τιμές για τις παραμέτρους c_1, c_2 και w μπορεί να μην είναι αναγκαία η χρήση περιορισμού ταχύτητας, δημοφιλείς τιμές είναι οι $c_1 = c_2 = 1.49618$ και $w = 0.7298$ [28].

5.3. Αρχικοποίηση πληθυσμού PSO

Ένα ζήτημα είναι ο τρόπος με τον οποίο αρχικοποιούνται οι θέσεις του πληθυσμού των σωματιδίων. Σε περίπτωση που δεν έχουμε γνώση περί της αντικειμενικής συνάρτησης, εύκολα μπορούμε διαισθητικά να ισχυριστούμε πως μια αρχικοποίηση που παρέχει αρκετή κάλυψη του χώρου ή μια που δημιουργεί πυρήνες, ίσως να οδηγήσει σε σύγκλιση σε κάποιο τοπικό ακρότατο. Γενικά σε black-box προβλήματα βελτιστοποίησης προτιμάται τυχαία και ομοιόμορφη κατανομή των σωματιδίων στον χώρο αναζήτησης καθώς δεν υπάρχει γνώση που να μας οδηγήσει σε μεροληψία υπέρ κάποια τοποθεσίας [25]. Οπότε οι θέσεις και οι ταχύτητες αρχικοποιούνται συνήθως ως εξής :

$$y_{ij}^{(0)} \sim \mathcal{U}(y_j^{min}, y_j^{max})$$

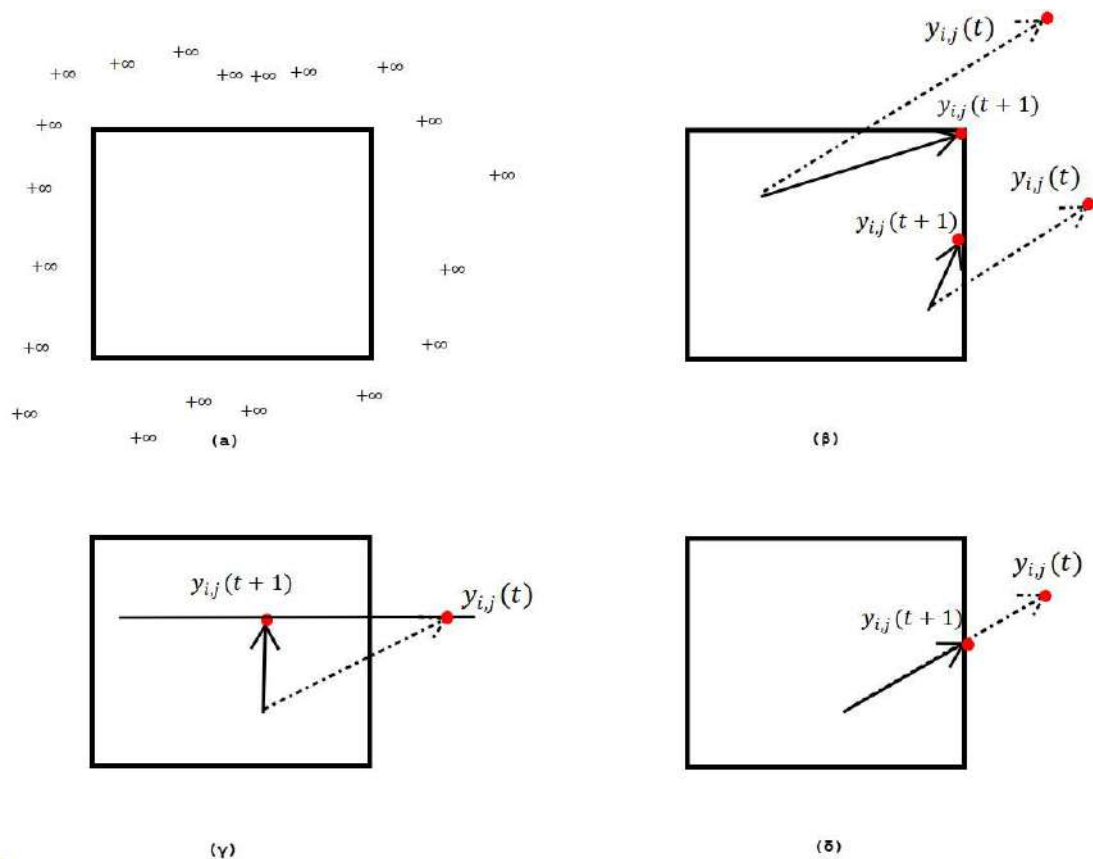
$$v_{ij}^{(0)} \sim \mathcal{U}(-v_j^{max}, v_j^{max})$$

Διάφοροι μέθοδοι έχουν αξιοποιηθεί για πιο ομοιόμορφη κατανομή των σωματιδίων κατά την αρχικοποίηση ώστε να επιτευχθεί καλύτερη εξερεύνηση όπως ο ορθογώνιος σχηματισμός που χρησιμοποιείται στον σχεδιασμό πειραμάτων ή η ψηφίδωση Voronoi. Γενικά όμως στους αλγόριθμους βελτιστοποιήσεις μια βασική συνθήκη σχεδιασμού είναι η ανεκτικότητα του αλγορίθμους σε μικρές διαταραχές των αρχικών συνθηκών. Οπότε σε μια ιδανική υλοποίηση του PSO, η απόδοση θα έπρεπε να είναι ανεξάρτητη της αρχικοποίησης, πράγμα όμως που δεν ισχύει στην πραγματικότητα.

Ένα πρόβλημα που προκύπτει με την ομοιόμορφη αρχικοποίηση εντοπίζεται στους χώρους με υψηλές διαστάσεις. Σε χώρους με υψηλές διαστάσεις ο όγκος είναι συγκεντρωμένος στα όρια του χώρου [29] [30]. Οπότε με την ομοιόμορφη αρχικοποίηση πολλά σωματίδια λαμβάνουν αρχικές θέσεις πολύ κοντά τουλάχιστον σε ένα όριο το χώρο αναζήτησης [30]. Αυτό έχει ως αποτέλεσμα πολλά σωματίδια να ξεφεύγουν από τα όρια του χώρου αναζήτησης [30]. Οπότε προκύπτει η ανάγκη να αντιμετωπιστεί το φαινόμενο αυτό μεταβάλλοντας την θέση των σωματιδίων που ξεπερνούν τα όρια. Αυτό το ζήτημα πρέπει να αντιμετωπιστεί ούτως ή άλλως αλλά στις συνθήκες που αναφέραμε το φαινόμενο είναι περισσότερο έντονο. Οι τεχνικές για τον έλεγχο της θέσης είναι πολλές και η βέλτιστη τεχνική εξαρτάται από το πρόβλημα βελτιστοποίησης και την αντικειμενική συνάρτηση που αντιμετωπίζουμε.

Σύμφωνα με τους Sabine Helwig and Rolf Wanka [30] μερικές από τις πιο κοινές τεχνικές που χρησιμοποιούνται στο σωματίδιο που φεύγει εκτός ορίων είναι :

- “θανατικής ποινής” (death penalty) η οποία θέτει την τιμή της αντικειμενικής συνάρτησης που αντιστοιχεί για το σωματίδιο ίση με $+\infty$ αν πρόκειται για πρόβλημα ελαχιστοποίηση και $-\infty$ για προβλήματα μεγιστοποίησης, “τιμωρώντας” έτσι την συμπεριφορά του σωματιδίου που βγαίνει εκτός ορίου.
- “Nearest”, τοποθέτηση του σωματιδίου στο πλησιέστερο όριο
- “Nearest + Turb” η τοποθέτηση πάλι στο πλησιέστερο όριο προσθέτοντας όμως μια τυχαιότητα στην ταχύτητα του
- “Random”, επαναφορά του σωματιδίου στον χώρο αναζήτησης μεταβάλλοντας με ομοιόμορφη κατανομή τη j – οστή συνιστώσα του σωματιδίου στα όρια του χώρου,
- “Shr” επαναφορά και συρρίκνωση της ταχύτητας του.

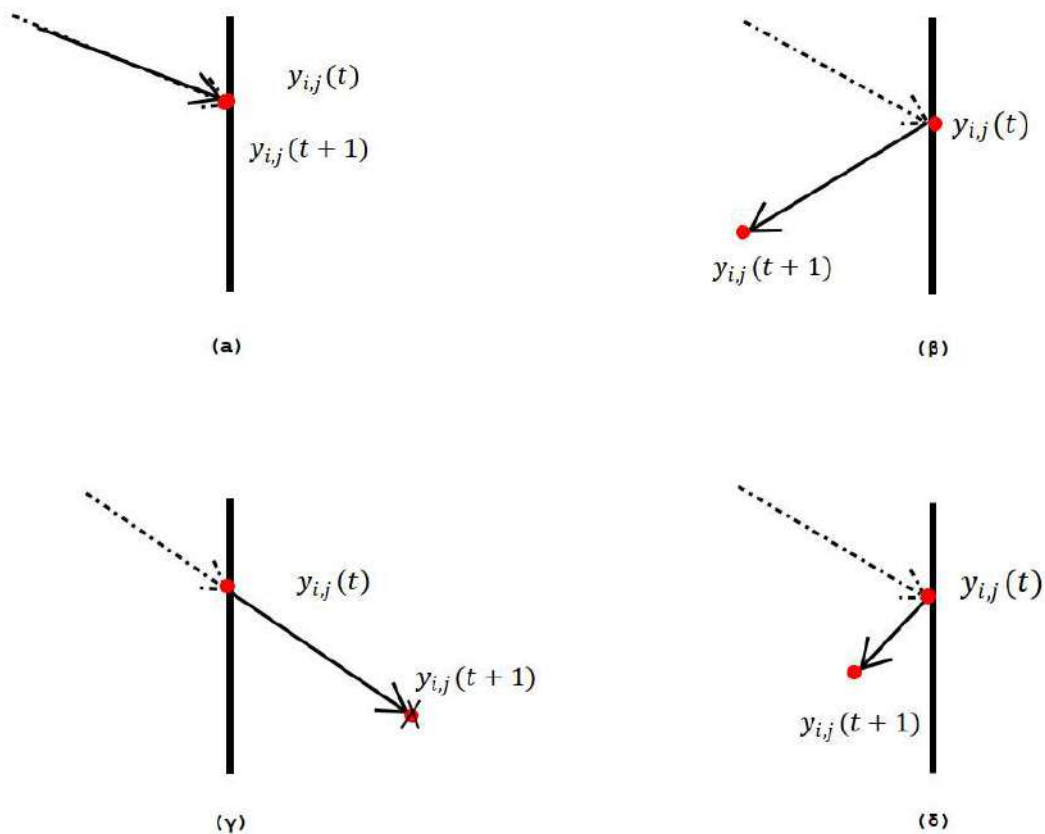


Εικόνα 18. Τεχνικές διαχείρισης σωματιδίων που φεύγουν από τον χώρο αναζήτησης. Θανατικής ποινής (α), Nearst+Turb (β), Random (γ), Nearest (δ).

Συγκρίνοντας τις τεχνικές αυτές οι Sabine Helwig and Rolf Wanka [30] στις συναρτήσεις σφαίρας, Rastrigin, Rosenbrock, Schwefel, Michalewicz και Griewank επιβεβαίωσαν πως η επιλογή της τεχνικής που θα εφαρμοστεί εξαρτάται από το πρόβλημα που έχουμε να αντιμετωπίσουμε. Τις χειρότερες επιδόσεις είχαν οι "Nearest" και "Nearest + Turb" κατά μέσο όρο (εκτός από την συνάρτηση Schwefel στην οποία είχαν εξαιρετικές επιδόσεις) και η Shr ήταν ελαφρώς χειρότερη των Inf και Random. Τρέχοντας δοκιμές όμως στις ίδιες συναρτήσεις αυτή τη φορά όμως σε χώρο αναζήτησης που δεν είναι ισορροπημένος (unbalanced) με το βέλτιστο δηλαδή να μην βρίσκεται στον κέντρο του χώρου αναζήτησης, τα αποτελέσματα είναι αντίστροφα. Οι τεχνικές που οδηγούσαν σε κακές επιδόσεις τώρα οδηγούν σε γρήγορη σύγκλιση [30].

Πολλοί ερευνητές εφαρμόζουν τεχνικές όπου τα όρια του χώρου αντιπροσωπεύουν έναν νοητό τοίχο και μόλις κάποιο σωματίδιο ξεπερνά τα όρια είναι σαν να προσκρούει σε αυτόν έτσι ξεχωρίζουν μερικές τεχνικές :

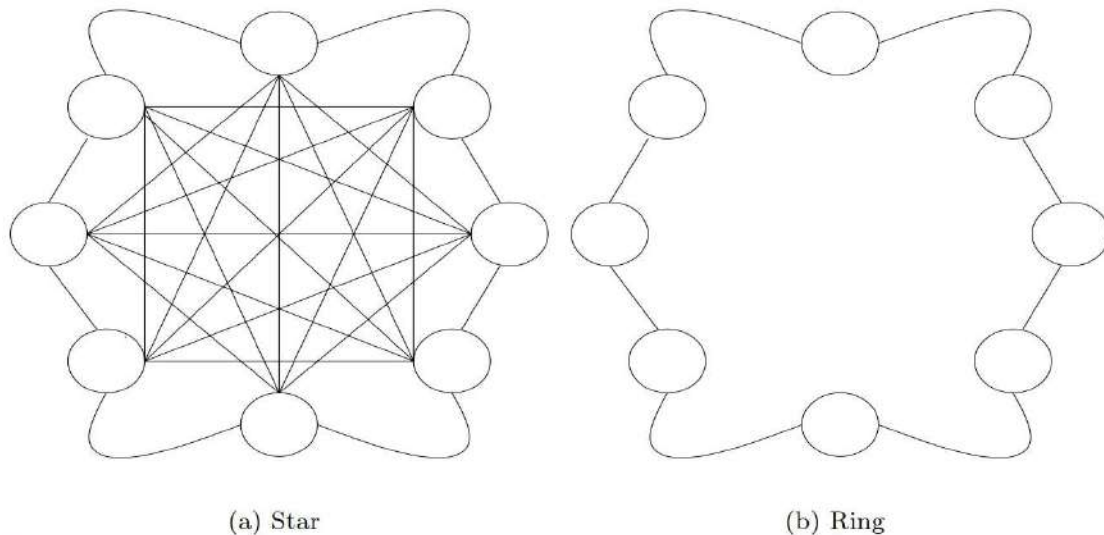
- Τοίχος απορρόφησης το σωματίδιο επανατοποθετείται στο όριο με εκμηδένιση της ταχύτητας του
- Τοίχος αντικατοπτρισμού όπου κάθε συνιστώσα της ταχύτητας αλλάζει πρόσημο ώστε να ανακλαστεί το σωματίδιο πίσω τον χώρο.
- Αόρατος τοίχος όπου επιτρέπεται να ξεφύγουν τα σωματίδια αλλά ο αλγόριθμος τα αγνοεί όσον αφορά τον υπολογισμό της αντικειμενικής συνάρτησης για την θέση του σωματιδίου έως ότου επιστρέψουν.
- Τοίχος απόσβεση όπου κάθε σωματίδιο που ανακλάται αλλά επιπλέον αποσβένεται επίσης η ταχύτητα του.



Εικόνα 19. Τοίχος απορρόφησης (α), Τοίχος αντικατοπτρισμού(β), Αόρατος Τοίχος (γ), Τοίχος αντικατοπτρισμού με απόσβεση (δ)

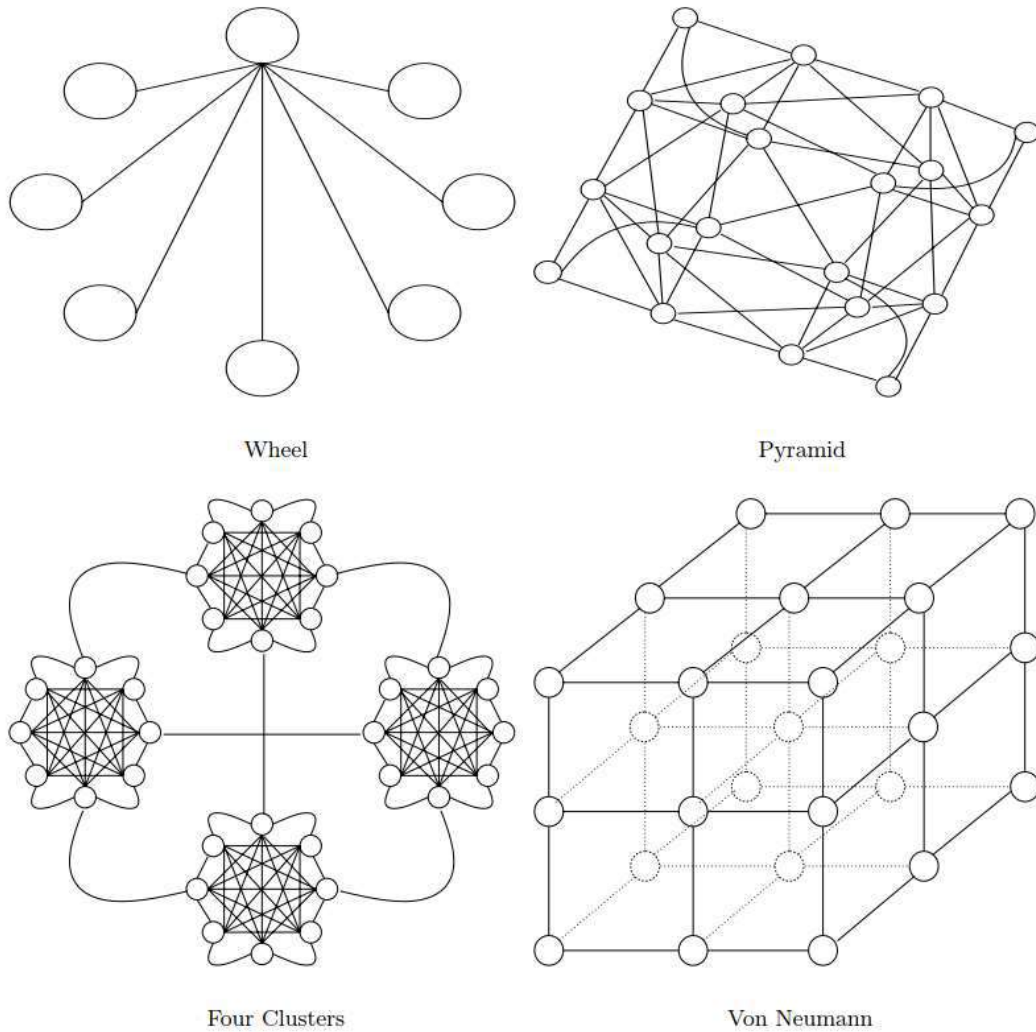
5.4. Γειτονιές σωματιδίων.

Το κεντρικό θέμα στην Νοημοσύνη Σμήνους είναι η έννοια της αλληλεπίδρασης και της ανταλλαγής πληροφορίας μέσα σε ένα σύνολο που καλείται γειτονιά. Οπότε επόμενο είναι να κατέχει σημαντικό ρόλο το πως ορίζεται μια γειτονιά. Αρχικά μπορούμε να χωρίσουμε τις γειτονιές σε δύο μεγάλες κατηγορίες την *gbest* και την *lbest*. Στην *gbest* η γειτονιά κάθε σωματιδίου είναι όλο το σμήνος και η τοπολογία του δικτύου είναι τύπου πλήρως κατανεμημένη (fully connected) [27]. Κάθε σωματίδιο συγκρίνει την πληροφορία που έχει αποκτηθεί από όλα τα σωματίδια. Στην *lbest* κατηγορία μπορούμε να έχουμε διάφορες τοπολογίες δικτύου όπου κάθε σωματίδιο επηρεάζεται από τους k άμεσους γείτονες που ανήκουν στην γειτονιά του. Μια από τις πιο κλασσικές *lbest* τοπολογίες είναι του δακτυλίου στην οποία κάθε σωματίδιο επικοινωνεί με τους $k = 2$ άμεσους γείτονες.



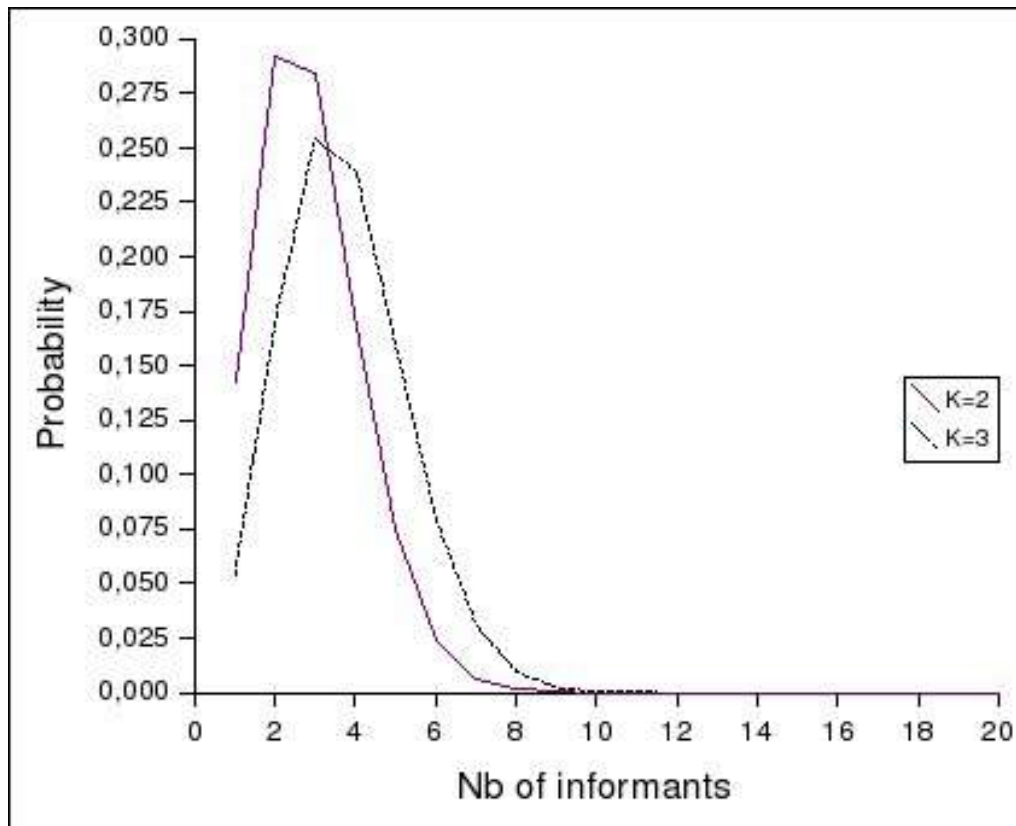
Εικόνα 20 [27]. Τοπολογίες Αστέρα και δακτυλίου

Η ροή της πληροφορίας σε ένα δίκτυο εξαρτάται από τον βαθμό συνδεσιμότητας μεταξύ των σωματιδίων, τον βαθμό ομαδοποίησης (clustering) και την απόσταση των σωματιδίων [27]. Σε υψηλά συνδεδεμένα δίκτυα τα πιο πολλά σωματίδια μπορούν να επικοινωνούν μεταξύ τους με αποτέλεσμα η πληροφορία γύρω από το καλύτερο σωματίδιο να μεταδίδεται γρήγορα στο δίκτυο πράγμα που σημαίνει ταχύτερη σύγκλιση [27]. Είναι όμως πιθανό όμως γίνεται φανερό ο κίνδυνος για πρόωρη σύγκλιση σε τοπικό ελάχιστο. Από την άλλη αραιά συνδεδεμένα δίκτυα με μεγάλο βαθμό ομαδοποίησης μπορεί να μην καλυφθεί ο χώρος αναζήτησης ικανοποιητικά και να μοιράζεται περιορισμένη πληροφορία με τον κίνδυνο πάλι να μην βρεθεί η βέλτιστη λύση [27]. Για την *gbest* τοπολογία διαισθητικά μπορεί κάποιος να εικάσει πως ο πληθυσμός απλά θα ακολουθήσει το καλύτερο σωματίδιο και θα παγιδευτεί με μεγάλη πιθανότητα, δοκιμές όμως δείχνουν ότι η *gbest* τεχνική συμπεριφέρεται πολύ καλά σε όλες τις δοκιμαστικές συναρτήσεις [22]. Στις τοπολογίες δακτυλίου κάθε σωματίδιο επικοινωνεί με τους άμεσους γείτονές του προσπαθώντας να μιμηθεί τους “καλύτερους” μετακινούμενο προς τις καλύτερες λύσεις που βρίσκονται την γειτονιά. Οι γειτονιές όμως αλληλεπικαλύπτονται ώστε να μεταφέρεται πληροφορία από την μία γειτονιά σε άλλη. Η πληροφορία μεταδίδεται με αργότερο ρυθμό άρα έχουμε μικρότερη ταχύτητα σύγκλισης, εξερευνάται όμως ο χώρος σε μεγαλύτερη έκταση [27].



Εικόνα 21 [27]. Άλλες τοπολογίες

Η επιλογή της βέλτιστης τοποθεσίας εξαρτάται από το πρόβλημα βελτιστοποίησης εν προκειμένω [22]. Για την μελέτη υιοθετούμε δύο τοπολογίες για δύο διαφορετικές εκδοχές του αλγορίθμου PSO. Στην πρώτη περίπτωση έχουμε την τοπολογία στοχαστικού αστέρα όπως περιγράφηκε από τον M. Clerc [31]. Στην τοπολογία αυτή κάθε σωματίδιο μοιράζει την πληροφορία που διαθέτει σε K τυχαία άλλα σωματίδια κατά μέσο όρο. Στην εφαρμογή αυτή επιλέγεται το ποιο σωματίδιο ενημερώνει ποιο. Συγκεκριμένα αν S ο αριθμός των σωματιδίων δημιουργούμε έναν ταυτοτικό πίνακα L μεγέθους $S \times S$. Για κάθε γραμμή i πραγματοποιούμε μια δειγματοληψία ομοιόμορφη κατανομής στο σύνολο $\{1, \dots, S\}$ εκλέγοντας με αντικατάσταση K αριθμούς k_1, \dots, k_K και θέτουμε $L(i, k_n) = 1$. Φυσικά εφόσον ο ίδιο αριθμός μπορεί να επιλεγεί πολλές φορές, τότε μπορεί να θέσουμε πολλές φορές μονάδα στην ίδια θέση. Κάθε στήλη και γραμμή του πίνακα L αντιπροσωπεύει ένα σωματίδιο. Αν για κάποια στήλη $L(i, j) = 1$ σημαίνει πως το σωματίδιο i μοιράζει πληροφορία προς το σωματίδιο j . Έτσι για κάθε στήλη j ο αριθμός των σειρών που έχουν την τιμή 1 αποτελούν την γειτονιά του σωματιδίου j . Με αυτό τον τρόπο για κάθε σωματίδιο, ο αριθμός των σωματιδίων από τα οποία λαμβάνει πληροφορία ανήκει στο σύνολο $\{1, \dots, S\}$. Κατά μέσο όρο σε κάθε γειτονιά ανήκουν K σωματίδια όμως αυτός ο αριθμός μπορεί να γίνει πολύ μεγαλύτερος με μικρότερη πιθανότητα όμως. Η κατανομή για το K είναι το άθροισμα $S - 1$ ανεξάρτητων δοκιμών Bernoulli [31].



Εικόνα 22 [31]. Η Συνάρτηση πυκνότητας πιθανότητας που αντιπροσωπεύει την πιθανότητα ένα σωματίδιο να έχει έναν συγκεκριμένο αριθμό σωματιδίων που το ενημερώνουν

5.5. Είναι Στρατηγική εξέλιξης (ES) ο PSO ;

Με μια πρώτη εξέταση των Swarm Intelligence (SI) μεθόδων και του PSO αλγορίθμου ένα φυσικό ερώτημα που προκύπτει είναι εάν και κατά πόσο ο PSO είναι εξελικτικός αλγόριθμος αφού φαίνεται πως απουσιάζουν οι τελεστές που συναντούμε σε αυτούς. Υπάρχουν πολλές οπτικές με τις οποίες μπορεί να εξετάσει κανείς την ομοιότητα του με τις ES και τα EA. Αρχικά ο PSO διατηρεί ένα πληθυσμό όπως οι εξελικτικοί αλγόριθμοι και οι στρατηγικές εξέλιξης. Ο πληθυσμός στις στρατηγικές εξέλιξης όπως είδαμε αποτελείται από γονείς που συνδυάζονται και παράγουν απογόνους που αντικαθιστούν τον προηγούμενο πληθυσμό. Θα μπορούσαμε να πούμε πως στον PSO όλα τα σωματίδια πεθαίνουν σε κάθε επανάληψη και αντικαθίστανται στην επόμενη επανάληψη [22]. Επίσης αν θεωρηθεί πως οι p_{best} θέσεις αποτελούν μέρος του πληθυσμού τότε υπάρχει και ένα είδους τελεστή επιλογής, η παρακάτω εξίσωση αντιπροσωπεύει αυτό σαν να ανταγωνίζονται οι απόγονοι αντικαθιστώντας του γονείς [28].

$$p_i^{(t)} = \begin{cases} y_i^{(t+1)}, & \text{εάν } f(y_i^{(t+1)}) \leq f(p_i^{(t)}) \\ p_i^{(t)}, & \text{διαφορετικά} \end{cases}$$

Επίσης η εξίσωση ανανέωσης της ταχύτητας

$$v_{ij}^{(t+1)} = w v_{ij}^{(t)} + c_1 (p_{ij}^{(t)} - y_{ij}^{(t)}) + c_2 (p_{gij}^{(t)} - y_{ij}^{(t)})$$

μπορεί να συγκριθεί με τον τελεστή μετάλλαξης όπου η δύναμη μετάλλαξης εξαρτάται από την απόσταση του σωματιδίου από τους δύο "γονείς" της περίπτωσης μας p_i και p_{g_i}

Ιδιαίτερα αν θέσουμε το βάρος αδράνειας ίσο με μηδέν και αντικαταστήσουμε την ταχύτητα με τις θέσεις τότε η ομοιότητα γίνεται μεγαλύτερη

$$y_{ij}^{(t+1)} = y_{ij}^{(t)} + C_1 (p_{ij}^{(t)} - y_{ij}^{(t)}) + C_2 (p_{gij}^{(t)} - y_{ij}^{(t)})$$

Είπαμε πως οι ES και EA αναπτύχθηκαν μιμούμενη την φύση την βιολογία την εξελικτική βιολογία. Πιο σύγχρονοι ερευνητές όπως ο S. Kauffman αμφισβητούν δύο βασικά σημεία του Νέο-Δαρβινισμού. Το πρώτο είναι πως η προέλευση την ζωής κατά τύχη ή λόγο μετάλλαξης είναι εξαιρετικά απίθανη στη χρονική στιγμή που εμφανίστηκε στη γη και κατά δεύτερο λόγο πως η εξέλιξη σύνθετης μορφής ζωής μόνο διαμέσου μετάλλαξης είναι εξαιρετικά απίθανη. Στον αντίποδα προτείνουν πως η πολυπλοκότητα των βιολογικών συστημάτων και των οργανισμών ίσως να προκύπτουν από Αυτο-οργάνωση, μια έννοια η οποία μπορεί να έχει πολλούς ορισμούς. Γενικά περιγράφει μια διαδικασία όπου τα συστήματα να εμφανίζουν μια αυθόρμητη τάξη μέσα από τοπικές αλληλεπιδράσεις από ένα αρχικά σύστημα με αταξία [32] ή την τάση να προσπαθεί να αυτό-οργανωθεί σε πιο περίπλοκες δομές προερχόμενο από απλούστερες [20]. Το σμήνος σωματιδίων είναι ένα αυτοοργανώμενο σύστημα που η συνολική δυναμική του προέρχεται από τοπικούς κανόνες [20]. Άρα εφόσον οι EA και οι ES ξεκίνησαν από την βιολογία εμπνεόμενες από την θεωρία της εξέλιξης, τώρα με μια νέα οπτική για την βιολογία επανατοποθετεί τον PSO ακριβώς σε αυτή την κατηγορία.

5.6. Εκδοχές Αλγορίθμων SPSO που χρησιμοποιήθηκαν

Για την μελέτη της παρούσας εργασίας χρησιμοποιήθηκαν δύο εκδοχές του αλγόριθμου PSO. Ο πρώτος ακολουθεί κυρίως τον κανονικό PSO και ονομάστηκε SPSO (Standard PSO). Η αρχικοποίηση του πληθυσμού έγινε μέσω ομοιόμορφης κατανομής. Για την διαχείριση των σωματιδίων που ξεφεύγουν από τον χώρο αναζήτησης εφαρμόστηκε η τεχνική του τοίχους αντικατοπτρισμού με απόσβεση. Η τοπολογία της γειτονιάς του SPSO είναι αυτή του στοχαστικού αστέρα που περιγράψαμε. Στην συνέχεια αυτός ο αλγόριθμος καλείται SPSO.

Η δεύτερη εκδοχή του PSO που χρησιμοποιήθηκε είναι ο Orthogonal Learning PSO όπως περιγράφεται από τους Z. Zhan κ.α. [33], [34]. Στους PSO αλγορίθμους το σωματίδιο μεταβάλλει την θέση του βάση της προσωπικής του καλύτερης θέση και της καλύτερης θέσης που έχει ανακαλύψει το καλύτερο σωματίδιο της γειτονιάς με το άθροισμα της σχέσης $v_{ij}^{(t+1)} = w v_{ij}^{(t)} + C_1 (p_{ij}^{(t)} - y_{ij}^{(t)}) + C_2 (p_{gij}^{(t)} - y_{ij}^{(t)})$. Σύμφωνα με τους δημιουργούς του OPSO αυτός ο τρόπος δεν είναι απαραίτητα ο καλύτερος τρόπος να συνδυαστεί η πληροφορία από τα δύο διανύσματα. Για παράδειγμα οι δύο θέσεις μπορεί να είναι αντιδιαμετρικές σε σχέση με το σωματίδιο που πρόκειται να ενημερωθεί. Επιπλέον τα δύο διανύσματα μπορεί να περιέχουν καλές τιμές για μερικές συνιστώσες βελτιώνοντας την λύση αλλά σε μερικές άλλες μπορεί να περιέχουν κακές συνιστώσες σε άλλες δυσχεραίνοντας της λύση.

Για να εκμεταλλευτούμε με καλύτερο τρόπο τα δύο αυτά διανύσματα δημιουργούμε έναν κλασματικό ορθογώνιο παραγοντικό σχεδιασμό με δύο επίπεδα, τα οποία είναι τα διανύσματα των θέσεων του $pbest$ και $gbest$, παράγοντες N όσες οι διαστάσεις του προβλήματος ενημέρωσης μοντέλου και σχηματίστηκαν M συνδιασμοί καταλήγοντας σε ένα $L_M(2^N)$ ορθογώνιο πίνακα. Σκοπός είναι ο υπολογισμός της επίδρασης κάθε επιπέδου σε κάθε παράγοντα βάση της συνάρτησης καταλληλότητας ώστε να καθορίσουμε το καλύτερο επίπεδο για κάθε παράγοντα. Με αυτό τον τρόπο μέσω των δύο διανυσμάτων δημιουργούμε έναν νέο διάνυσμα το οποίο θα χρησιμοποιήσει το σωματίδιο για να ανανεώσει τη θέση του. Αυτή η διαδικασία συμβαίνει για ένα σωματίδιο κάθε φορά που η προσωπική καλύτερη θέση $pbest$ του σωματιδίου αποτυγχάνει να βελτιωθεί μετά από έναν αριθμό επαναλήψεων. Με

περισσότερη λεπτομέρεια δημιουργείται ένας πίνακας όπου στήλες αντιπροσωπεύουν τους παράγοντες που μεταβάλλονται στο “πείραμα” και οι γραμμές το επίπεδο που επιλέγεται για κάθε παράγοντα. Ο πίνακας αυτός περιέχει τιμές με 0 ή 1, οι οποίες αντιπροσωπεύουν αν σε έναν συνδυασμό χρησιμοποιείται η *pbest* ή *gbest* θέση. Αν ένα κελί του πίνακα m, q περιέχει για παράδειγμα την τιμή 0, τότε αυτό σημαίνει πως στον συνδυασμό m για τον q παράγοντα, χρησιμοποιείται η *pbest* θέση. Για όλους αυτούς τους συνδυασμούς υπολογίζουμε την συνάρτηση καταλληλότητας. Αν f_m η τιμή της συνάρτησης καταλληλότητας για τον m συνδυασμό τότε η επίδραση S_{nq} , του q επιπέδου στον n – οστό παράγοντα υπολογίζεται ως εξής :

$$S_{nq} = \frac{\sum_{m=1}^M f_m \times z_{mnq}}{\sum_{m=1}^M z_{mnq}},$$

όπου για κάθε συνδυασμό m , το $z_{mnq} = 1$ αν ο n παράγοντας είναι από το q επίπεδο αλλιώς $z_{mnq} = 0$. Με αυτό τον τρόπο καθορίζεται το νέο διάνυσμα επιλέγοντας για κάθε συνιστώσα (παράγοντα), τιμή από το επίπεδο (διάνυσμα) που έχει την χαμηλότερη τιμή S_{nq} αν πρόκειται για πρόβλημα ελαχιστοποίησης.

Παρακάτω παρουσιάζεται ένας με ένα παράδειγμα για ένα χημικό πείραμα ώστε να γίνουν αυτές οι διαδικασίες πιο κατανοητές.

DECIDING THE BEST COMBINATION LEVELS OF THE CHEMICAL EXPERIMENTAL FACTORS USING AN OED METHOD

Combinations	A: Temperature (°C)	B: Time (Min)	C: Alkali (%)	Results
C_1	(1) 80	(1) 90	(1) 5	$F_1 = 31$
C_2	(1) 80	(2) 120	(2) 6	$F_2 = 54$
C_3	(1) 80	(3) 150	(3) 7	$F_3 = 38$
C_4	(2) 85	(1) 90	(2) 6	$F_4 = 53$
C_5	(2) 85	(2) 120	(3) 7	$F_5 = 49$
C_6	(2) 85	(3) 150	(1) 5	$F_6 = 42$
C_7	(3) 90	(1) 90	(3) 7	$F_7 = 57$
C_8	(3) 90	(2) 120	(1) 5	$F_8 = 62$
C_9	(3) 90	(3) 150	(2) 6	$F_9 = 64$
Levels	Factor Analysis			
L_1	$(F_1 + F_2 + F_3)/3 = 41$	$(F_1 + F_4 + F_7)/3 = 47$	$(F_1 + F_6 + F_8)/3 = 45$	
L_2	$(F_4 + F_5 + F_6)/3 = 48$	$(F_2 + F_5 + F_8)/3 = 55$	$(F_2 + F_4 + F_9)/3 = 57$	
L_3	$(F_7 + F_8 + F_9)/3 = 61$	$(F_3 + F_6 + F_9)/3 = 48$	$(F_3 + F_5 + F_7)/3 = 48$	
OED Results	A3	B2	C2	

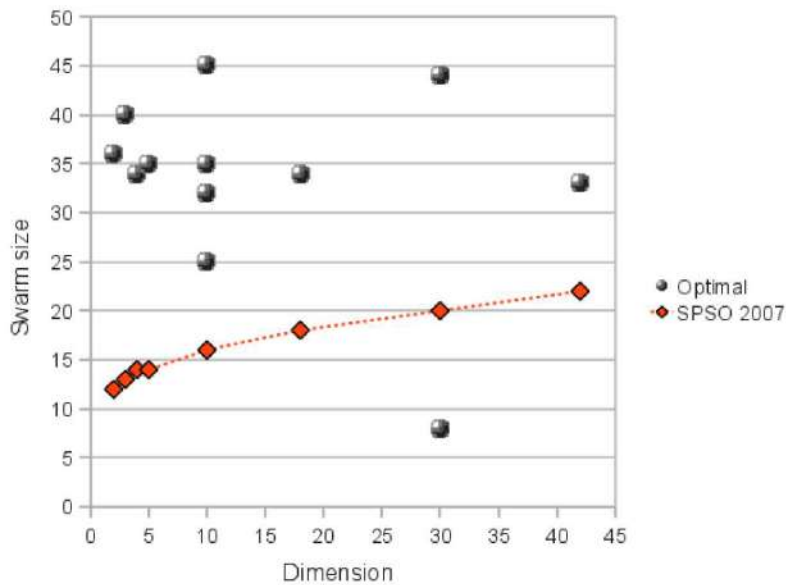
Εικόνα 23 [33]. Παράδειγμα Ορθογώνιου σχεδιασμού Πειράματος.

6. Επιλογή παραμέτρων των αλγορίθμων και μέτρα σύγκρισης.

Οι αλγόριθμοι που χρησιμοποιήθηκαν στο περιβάλλον του MATLAB είναι ο CMA-ES, στην έκδοση 3.61.beta όπως βρίσκεται στην ιστοσελίδα του CMAP εργαστηρίου του πολυτεχνικού Ινστιτούτου του Παρισιού [35] και οι PSO με τοπολογία στοχαστικού άστρου και OPSO με ορθογώνιο σχεδιασμό που περιγράφηκαν προηγουμένως.

Οι αλγόριθμοι που παρουσιάσαμε περιέχουν ένα μεγάλο εύρος παραμέτρων που επηρεάζουν τις επιμέρους διαδικασίες τους, επομένως και την απόδοσή τους. Όταν αλγόριθμοι βελτιστοποίησης πρόκειται να συγκριθούν μεταξύ τους ένα ερώτημα που τίθεται είναι αν είναι θεμιτή, αναγκαία ή λανθασμένη η ρύθμιση και η βελτιστοποίηση αυτών των παραμέτρων πριν την σύγκριση. Δηλαδή μια μετα-μοντελοποίηση. Αν και στις πρακτικές εφαρμογές των αλγορίθμων αυτό είναι αναγκαίο, εντούτοις στην σύγκριση θα πρέπει να χρησιμοποιείται το προεπιλεγμένο σύνολο παραμέτρων που έχει βρεθεί που δίνει σχετικά καλά αποτελέσματα σε ένα ευρύ φάσμα προβλημάτων ώστε η σύγκριση να είναι αμερόληπτη και για να δοκιμαστεί η "ανθεκτικότητα" (robustness) των αλγορίθμων. Οπότε με βάση αυτή την λογική χρησιμοποιούνται οι προεπιλεγμένες τιμές των παραμέτρων που προτείνουν οι δημιουργοί των αλγορίθμων και αυτές που εντοπίζονται στην βιβλιογραφία. Μία από τις παραμέτρους που δεν είναι τόσο αυστηρά ορισμένη και η επιλογή της επαφίεται στον χρήστη ανάλογα με το πρόβλημα που αντιμετωπίζει είναι η επιλογή του μεγέθους του πληθυσμού.

Όπως έχουμε αναφέρει στους PSO αλγορίθμους γενικότερα, απουσιάζει κάποιος μηχανισμός τυχαιοποίησης πέρα από την αρχικοποίηση του πληθυσμού. Άρα φανερό είναι πως χρειάζεται ένα σχετικά μεγάλο μέγεθος πληθυσμού για να αποφευχθεί πιθανός εγκλωβισμός σε τοπικό ακρότατο. Έχουν προταθεί πολλές σχέσεις για την επιλογή του μεγέθους του πληθυσμού σε συνάρτηση με την διάσταση του προβλήματος. Για παράδειγμα στον SPSO 2007 αλγόριθμο ο πληθυσμός καθορίζεται ανάλογα με την διάσταση N από την σχέση, $10 + 2 * \sqrt{N}$. Από δοκιμές φαίνεται όμως πως πολλές φορές αυτές οι σχέσεις υποεκτιμούν το μέγεθος του πληθυσμού. Τα αποτελέσματα από μερικές δοκιμές σε ένα σετ προβλημάτων προτείνουν μια τιμή που κυμαίνεται από 25 έως 45 σωματίδια με τους δημιουργούς της SPSO 2011 να προτείνουν 40 σωματίδια ως προεπιλεγμένη τιμή [36]. Στον CMA-ES η σχέση που προτείνεται είναι $4 + \lfloor 3 * \log(N) \rfloor$, με του δημιουργούς να τονίζουν πως αν και είναι δυνατή η αύξηση του πληθυσμού, ωστόσο δεν είναι απαραίτητη για την σωστή λειτουργία του αλγορίθμου διότι οι συντελεστές μάθησης αποτρέπουν τον υποβιβασμό του πληθυσμού σε ένα υποσύνολο του χώρου αναζήτησης, πράγμα που είναι η πιο συχνή αιτία της αποτυχίας μιας στρατηγικής εξέλιξης, καθώς οδηγεί σε εγκλωβισμό της λύσης.



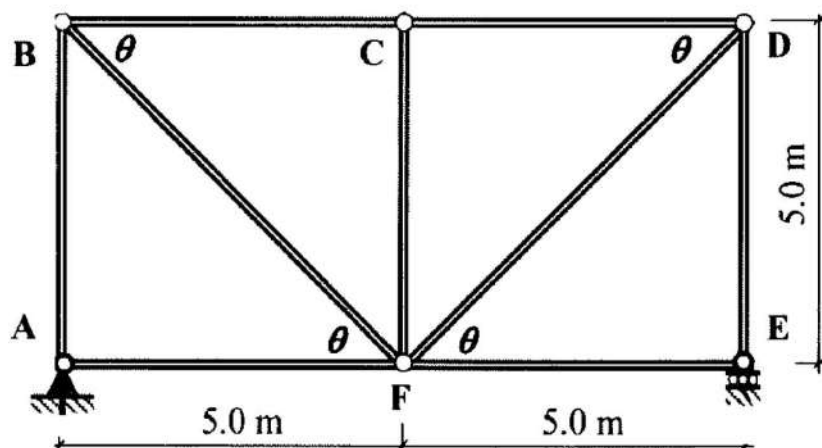
Εικόνα 24 [36]. Ιδανικό μέγεθος πληθυσμού σωματιδίων ανάλογα με την διάσταση του προβλήματος

Όσον αφορά τα μέτρα σύγκρισης αν και δεν έχει καθοριστεί ακόμη κάποιο αυστηρό πλαίσιο για την σύγκριση της απόδοσης των αλγορίθμων, γίνονται προσπάθειες προς αυτή την κατεύθυνση. Γενικά μπορούμε να πούμε πως τα μέτρα σύγκρισης έχουν ως στόχο να ποσοτικοποιήσουν την αξιοπιστία, την αποδοτικότητα και την αποτελεσματικότητα των αλγορίθμων. Συγκεκριμένα με την αποδοτικότητα κρίνουμε την ικανότητα του αλγορίθμου να βρίσκει ικανοποιητικές λύσεις συγκριτικά με το υπολογιστικό κόστος, την ταχύτητα με την οποία πλησιάζει την λύση ή ακόμα και τον χρόνο που χρειάζεται για να εκτελέσει μια επανάληψη μετρώντας την πολυπλοκότητα του αλγορίθμου. Ένας αξιόπιστος αλγόριθμος ικανοποιεί τους περιορισμούς που θέτουμε, έχει υψηλό λόγο επιτυχιών προς τις προσπάθειες, τα αποτελέσματα του διέπονται από επαναληψιμότητα και είναι ακριβής υπό την έννοια ότι οι λύσεις που παράγει έχουν συγκριτικά με άλλους αλγορίθμους μέση τιμή πιο κοντά στην βέλτιστη και μικρότερη τυπική απόκλιση. Για να περιγράψουμε αυτές τις ποσότητες στα μέτρα όπου χρειάζεται κάποιου είδους “διάσταση” χρόνου. Χρησιμοποιούμε τον αριθμό κλήσεων της αντικειμενικής συνάρτησης καθώς θεωρούμε πως είναι η πιο κοστοβόρα διαδικασία ενώ επίσης το υπολογιστικό κόστος της κλήσης είναι ανεξάρτητο από τον αλγόριθμο βελτιστοποίησης που εφαρμόζουμε. Έτσι δεν λαμβάνουμε υπόψιν την πολυπλοκότητα των αλγορίθμων που θα εφαρμόσουμε. Για να κριθεί αν μια δοκιμή του αλγορίθμου είναι επιτυχής, τέθηκαν άνω και κάτω όρια γύρω από κάθε μεταβλητή που πρόκειται να αναγνωριστεί. Αν το αποτέλεσμα μιας δοκιμής του αλγορίθμου ικανοποιήσει ταυτόχρονα όλους του περιορισμούς για τις μεταβλητές, τότε θεωρείται μια επιτυχημένη δοκιμή. Με βάση αυτά τα όρια, ένα μέτρο απόδοσης είναι η πρώτη στιγμή όπου το αποτέλεσμα του αλγορίθμου ικανοποιεί τα όρια αυτά για όλες τις παραμέτρους, με την επιπλέον απαίτηση να παραμένουν τα αποτελέσματα εντός των επιτρεπτών ορίων μέχρι τον τερματισμό του αλγορίθμου. Δηλαδή αφού εντοπίζεται το τελικό αποτέλεσμα της αναγνώρισης, αν αυτό είναι εντός των ορίων, αναζητείται προς τα πίσω ο αριθμός επαναλήψεων για τις οποίες βρισκόμαστε εντός των ορίων. Επόμενα μέτρα είναι η τιμή της αντικειμενικής συνάρτησης στο τέλος του αλγορίθμου και το ποσοστό των δοκιμών που πέτυχαν για κάθε αλγόριθμο.

7. Αποτελέσματα δοκιμών ενημέρωσης μοντέλου.

7.1. Ενημέρωση μοντέλου Δικτυώματος.

Το πρώτο μοντέλο που θα μελετήσουμε είναι αυτό ενός δικτυώματος.



Εικόνα 25. Μοντέλο δικτυώματος.

Το δικτύωμα με τις στηρίξεις που εικονίζονται, είναι κατασκευασμένο από αλουμίνιο με μέτρο ελαστικότητας $E = 69GPa$, πυκνότητα $\rho = 2700kg/m^3$, το μήκος των ράβδων είναι 5 μέτρα και η επιφάνεια της διατομής των ράβδων είναι $360m^2$. Αναζητούμε να αναγνωρίσουμε το μέτρο Ελαστικότητας και την επιφάνεια των διατομών των ράβδων. Οι ράβδοι προσομοιώνονται με αξονικά πεπερασμένα στοιχεία. Για την ενημέρωση του μοντέλου συγκρίνουμε τις ιδιοσυχνότητες της κατασκευής επιλύοντας το ιδιοπρόβλημα. Οι πειραματικές (πραγματικές) ιδιοσυχνότητες συγκρίνονται με τις ενημερωμένες μέσω της σχέσης (1) που παρουσιάσαμε.

$$\varepsilon_{\omega_r}(\mathbf{y}) = \frac{\omega_r^2(\mathbf{y}) - \hat{\omega}_r^2}{\hat{\omega}_r^2} \quad (1)$$

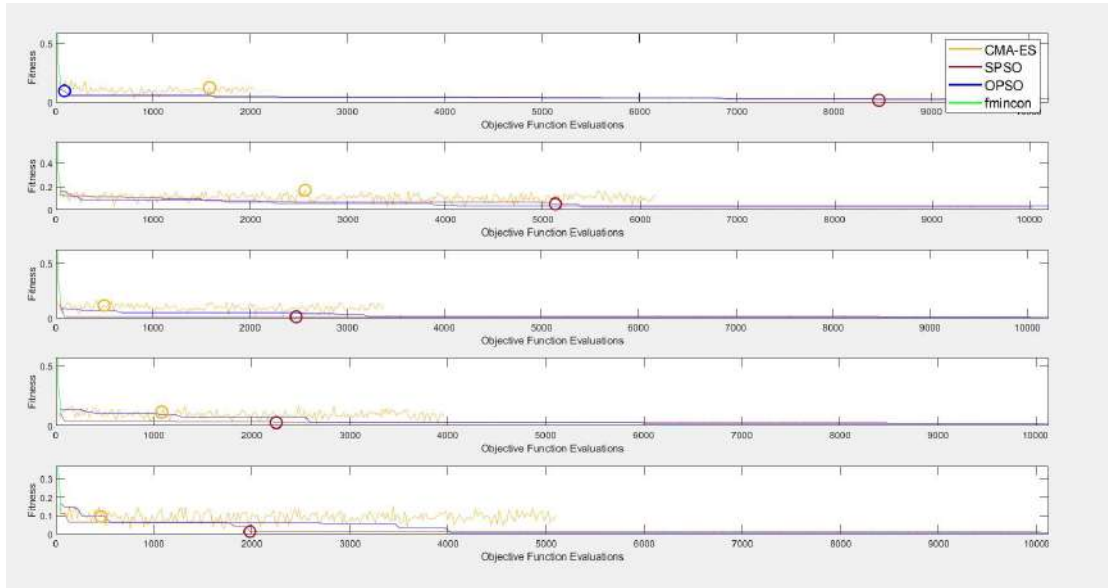
Ακολούθως η αντικειμενική συνάρτηση διαμορφώνεται ως εξής $F(\mathbf{y}) = \sum_{r=1}^m \varepsilon_{\omega_r}(\mathbf{y})$.

Η CMA-ES πάλι έχει την καλύτερη επίδοση με την SPSO να έχει την δεύτερη καλύτερη. Ενώ η fmincon παγιδεύεται πάλι σε τοπικό ελάχιστο χωρίς να συγκλίνει προς την βέλτιστη λύση.

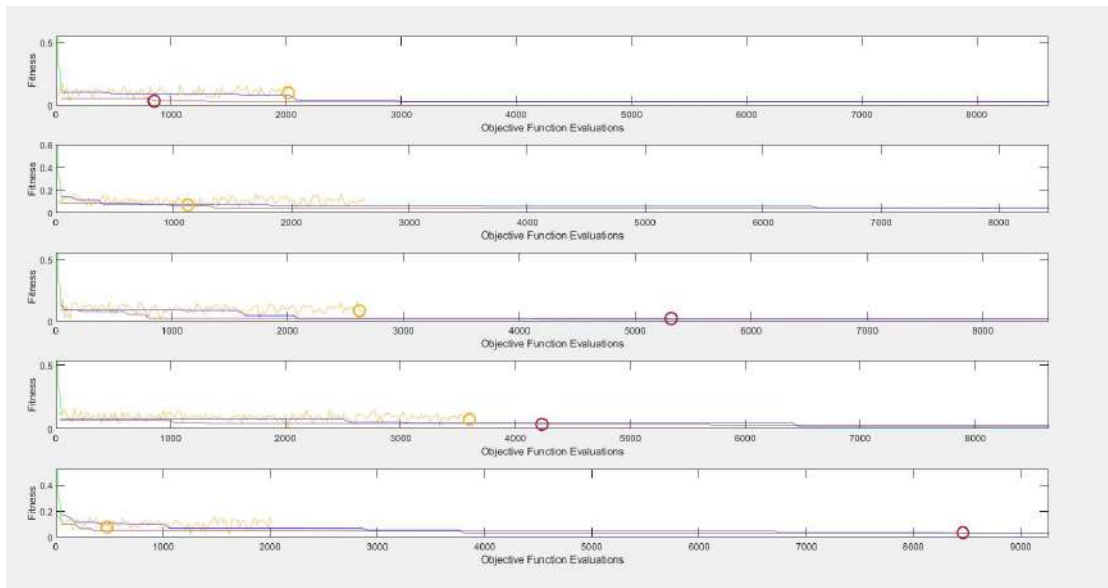
Δοκιμή/Αριθμός Υπολογισμών	CMA-ES	SPSO	OPSO	fmincon
1	1580	8460	90	-
2	2560	5130	0	-
3	500	2475	0	-
4	1080	2250	910	-
5	460	1980	943	-
6	2020	855	2097	-

7	1120	-	0	-
8	2620	5310	1632	-
9	3600	4230	0	-
10	480	8460	1061	-
Μέσος όρος Υπολογισμών	1602	4350	1122	-

Εικόνα 26. Σύγκριση μεθόδων σε ενημέρωση μοντέλου δικτυώματος.



Εικόνα 27. Απόδοση μεθόδων σε ενημέρωση μοντέλου δικτυώματος



Εικόνα 28. Απόδοση μεθόδων σε ενημέρωση μοντέλου δικτυώματος

7.2. Μοντέλο μισού αυτοκινήτου (Half-car)

Το επόμενο μοντέλο του οποίου θα δοκιμάσουμε να αναγνωρίσουμε τις παραμέτρους είναι ενός μισού οχήματος. Το πρώτο μέρος της κατασκευής αντιπροσωπεύει το όχημα και το δεύτερο, μετά το στρεπτικό ελατήριο μια καρότσα και το φορτίο της m_5 . Η m_4 αντιπροσωπεύει τον κινητήρα του οχήματος, το ελατήριο με σταθερά ελαστικότητας k_4 προσομοιώνει την έδραση του κινητήρα και η δύναμη που εφαρμόζεται την ταλάντωση του κινητήρα. Το σκάφος του οχήματος προσομοιώνεται με ομογενή δοκό και χωρίζεται σε 8 ισομήκη καμπτικά στοιχεία. Οι μάζες m_1, m_2, m_3 αντιπροσωπεύουν τους τροχούς. Τα συστήματα των αναρτήσεων αντιπροσωπεύονται από τα διακριτά τους στοιχεία k_1, c_1 για την πρώτη, k_2, c_2 για την δεύτερη και k_3, c_3 για την τρίτη. Το ελαστικό προσομοιώνεται από τα διακριτά στοιχεία στιβαρότητας και απόσβεσης με τους δείκτες w . Οι παράμετροι της κατασκευής που θα προσπαθήσουμε να αναγνωρίσουμε είναι τα διακριτά στοιχεία των αναρτήσεων, η έδραση της καρότσας και του κινητήρα και το μέτρο ελαστικότητας του υλικού. Οι πραγματικές τιμές των παραμέτρων αυτών είναι :

$$k_1 = 1000 \frac{N}{m}, k_2 = 24950.639 \frac{N}{m}, k_3 = 28216.326 \frac{N}{m}, k_4 = 250 \frac{kN}{m}, k_5 = 300 \frac{kN}{m}$$

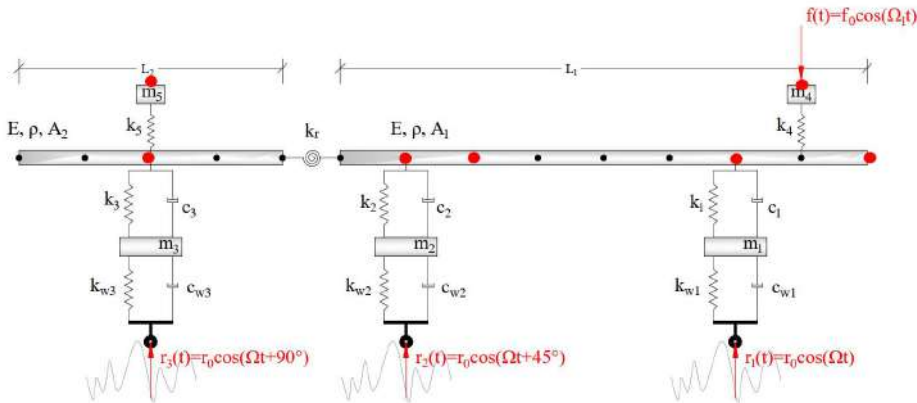
$$c_1 = 1284.6199 \frac{Ns}{m}, c_2 = 3713.1995 \frac{Ns}{m}, c_3 = 2786.4977 \frac{Ns}{m}, E = 210 \text{ GPa}.$$

Τέλος η πυκνότητα του υλικού του σκάφους είναι $\rho = 7580 \text{ kg/m}^3$, οι δοκοί του οχήματος και της καρότσας έχουν μήκος τέσσερα και δύο μέτρα αντίστοιχα, ενώ η δοκός του οχήματος έχει ύψος και πλάτος 0.17 και 0.3 μέτρα αντίστοιχα και η δοκός της καρότσας έχει ύψος και πλάτος 0.1 και 0.2 μέτρα αντίστοιχα.

Για την ενημέρωση του μοντέλου επιλέχθηκε να χρησιμοποιηθεί η επιτάχυνση των βαθμών ελευθερίας των κόμβων που εμφανίζονται στο παρακάτω σχήμα με κόκκινη κουκίδα. Η αντικειμενική συνάρτηση έχει την μορφή που παρουσιάσαμε στο πρώτο κεφάλαιο

$$F(\mathbf{y}) = \frac{1}{N} \sum_{i=1}^N \frac{\sum_{j=1}^M (x(\mathbf{y})_{ij} - \hat{x}_{ij})^2}{\sum_{j=1}^M (\hat{x}_{ij})^2} \quad (3)$$

Η απόκριση του μοντέλου στον χρόνο υπολογίζεται κατασκευάζοντας ένα μοντέλο χώρου κατάστασης από τα μητρώα του μοντέλου πεπερασμένων στοιχείων. Για το μοντέλο χώρου κατάστασης χρησιμοποιείται η συνάρτηση *ss* του MATLAB και για την απόκριση η *lsim*.



Εικόνα 29. Μοντέλο μισού οχήματος

Στην συνέχεια τέθηκε ένα όριο σφάλματος για την κάθε μεταβλητή για κρίνουμε την επιτυχής αναγνώριση της παραμέτρου. Το αρχικό σημείο εκτίμησης (Nominal Point) θεωρήθηκε το σημείο με τιμές παραμέτρων $k_1 = 12000 \frac{N}{m}$, $k_2 = 12000 \frac{N}{m}$, $k_3 = 12000 \frac{N}{m}$, $k_4 = 150 \frac{kN}{m}$, $k_5 = 150 \frac{kN}{m}$, $c_1 = 1000 \frac{Ns}{m}$, $c_2 = 1000 \frac{Ns}{m}$, $c_3 = 1000 \frac{Ns}{m}$,

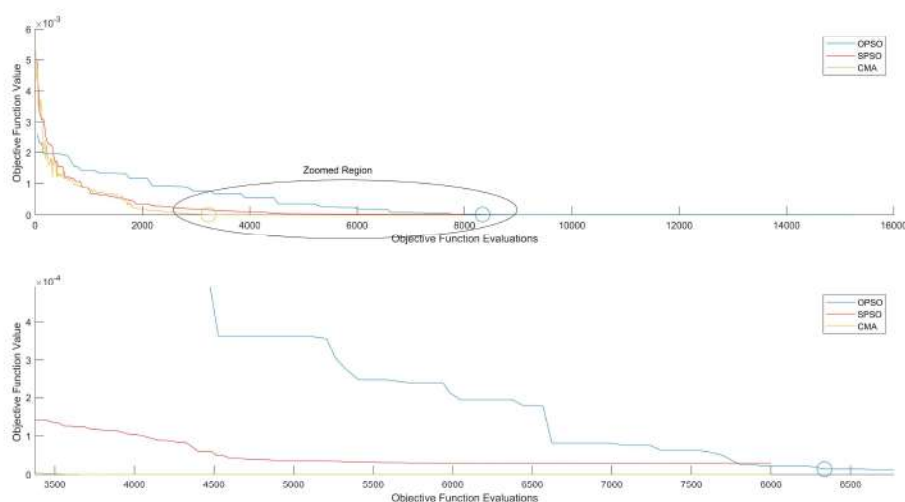
$E = 200 \text{ GPa}$. Οι απόκλιση των τιμών των μεταβλητών αυτών σε ποσοστό επί τοις εκατό είναι κατά σειρά, 20%, 51.9%, 57.5%, 22.15%, 73.1%, 64.1%, 40%, 50% και 5%.

Το σημείο τέθηκε αυτό ως interior point στην fmincon συνάρτηση του MATLAB και ως αρχικός γονέας στην CMA-ES. Στους PSO αλγόριθμους συνήθως συνηθίζεται η αρχικοποίηση του σμήνους μέσω ομοιόμορφης κατανομής ανάμεσα στα άνω και κάτω όρια του χώρου αναζήτησης. Αρχικά επιλέχθηκε να "προσομοιωθεί" κατά κάποιο τρόπο το πρώτο βήμα της CMA-ES στους SPSO και OPSO, αρχικοποιώντας τον πληθυσμό του σμήνους γύρω από το Nominal σημείο μέσω κανονικής κατανομής με μέσο το σημείο αυτό και τυπική απόκλιση την ίδια τυπική απόκλιση που εισάγουμε στην CMA-ES.

Τα κάτω όρια του χώρου αναζήτησης για τις μεταβλητές k_1, k_2, k_3 είναι $500 \frac{N}{m}$ ενώ τα άνω όρια είναι $250000 \frac{N}{m}$. Για τις c_1, c_2, c_3 τα κάτω όρια είναι $100 \frac{Ns}{m}$ και άνω $100000 \frac{Ns}{m}$. Για τις k_4, k_5 τα κάτω όρια είναι $150 \frac{kN}{m}$ και άνω $500 \frac{kN}{m}$. Τέλος για το μέτρο ελαστικότητας τα όρια είναι 220 GPa και 180 GPa άνω και κάτω αντίστοιχα.

Οι δοκιμές έγιναν με μέγιστο όριο τους 8000 υπολογισμούς της αντικειμενικής συνάρτησης, πραγματοποιώντας τον αντίστοιχο αριθμό επαναλήψεων στους αλγόριθμους ανάλογα με τον πληθυσμό τους σε κάθε γενιά, ώστε να φτάσουν τους 8000 υπολογισμούς της αντικειμενικής συνάρτησης. Για τον OPSO ακολουθήθηκε μια ελαφρώς διαφορετική προσέγγιση αναφορικά με την ποσότητα των υπολογισμών της αντικειμενικής συνάρτησης. Αυτό διότι στον OPSO ύστερα από κάποιον αριθμό επαναλήψεων όπου η προσωπική καλύτερη τιμή ενός σωματιδίου μένει στάσιμος χωρίς βελτίωση (προτείνεται η τιμή 4 ή 5), καλείται η συνάρτηση ορθογώνιου σχεδιασμού πειραμάτων η οποία με τη σειρά της θα πραγματοποιήσει υπολογισμούς της αντικειμενικής συνάρτησης για έναν σταθερό αριθμό ο οποίος για παράδειγμα για σχεδιασμό πειράματος εννέα μεταβλητών και δύο επιπέδων είναι 16 φορές. Αυτή η διαδικασία είναι μια επιπλέον διαδικασία η οποία δεν είναι μια από αυτές που αποτελούν τον κορμό των PSO. Αντίθετα σε παρόμοια περίπτωση στασιμότητας ο SPSO δημιουργεί τυχαία νέες συνδέσεις στον πληθυσμό μέσω της τοπολογίας του στοχαστικού άστρου χωρίς να υπολογίζει επιπλέον φορές την αντικειμενική συνάρτηση. Έτσι χωρίς να γνωρίζουμε εκ των προτέρων πόσες φορές επιπλέον ο OPSO θα υπολογίσει την αντικειμενική συνάρτηση για δοσμένο αριθμό επαναλήψεων, το άνω όριο του δεν σταμάτησε στις 8000 επαναλήψεις. Τα αποτελέσματα έδειξαν πως αν συνυπολογίσουμε το επιπλέον υπολογιστικό κόστος της OPSO υπάρχει μια καθυστέρηση στην σύγκλιση του OPSO στις πιο πολλές περιπτώσεις. Ο αλγόριθμος OPSO όμως κατάφερε να συγκλίνει και τις 30 φορές σε αποδεκτή λύση με τους επιπλέον υπολογισμούς της αντικειμενικής συνάρτησης να μην είναι πολλές σχετικά. Στις 30 δοκιμές κατά μέσο όρο ο OPSO βρήκε αποδεκτές λύση έπειτα από 9064 υπολογισμούς της συνάρτησης καταλληλότητας ξεπερνώντας το όριο κατά 1000 επαναλήψεις. Επίσης παρατηρείται πως σε πολλές περιπτώσεις που δεν κατάφερε ο SPSO να εντοπίσει αποδεκτή λύση, η τιμή της συνάρτησης καταλληλότητας για τον SPSO γίνεται σχεδόν παράλληλη με τον οριζόντιο άξονα σε κάποια επίπεδο, πράγμα που δείχνει πως σε αυτές τις περιπτώσεις ακόμη και αν ο SPSO δέχεται τις επιπλέον επαναλήψεις του OPSO, πιθανότατα δεν θα μπορούσε να εντοπίσει αποδεκτή λύση. Παρακάτω παρουσιάζουμε μια

τέτοια περιπτώσεις και τα αποτελέσματα συγκεντρωτικά σε πίνακα. Συνολικά ο CMA αλγόριθμος είναι ο πιο γρήγορος στην σύγκλιση και παρουσιάζει καλύτερες τιμές στην συνάρτηση καταλληλότητας. Η fmincon σταμάτησε πρόωρα τις επαναλήψεις και το MATLAB εμφάνισε ένα μήνυμα που περιγράφει τον λόγο της διακοπής το οποίο εμφανίζεται όταν σταματά λόγω παγίδευσης σε τοπικό ελάχιστο ή λόγω παρουσίας ασυνέχειας.



Εικόνα 30. Σύγκριση απόδοσης των μεθόδων όπου η καμπύλη της τιμή καταλληλότητας όπου ο SPSO αποτυγχάνει.

Για την κάθε μέθοδο έτρεξαν 30 δοκιμές και στον πίνακα 9. παρουσιάζεται για κάθε μέθοδο ο αριθμός κλήσης της αντικειμενικής συνάρτησης έπειτα από τον οποίο η μέθοδος συνέχισε να παραμένει μέχρι το τέλος εντός των ορίων επιτυχίας που έχουμε ορίσει. Στον πίνακα 10. συγκρίνεται η τιμή της συνάρτησης καταλληλότητας στο τέλος κάθε δοκιμής.

Δοκιμή	CMA-ES	SPSO	OPSO	fmincon
1	4240	-	10648	-
2	3760	6960	10856	-
3	4400	6520	9392	-
4	4170	-	8536	-
5	4170	-	8552	-
6	3920	-	8664	-
7	-	-	8744	-
8	3870	3320	9560	-
9	4410	-	9704	-
10	4930	-	8328	-
11	4080	-	10088	-
12	4640	-	7056	-
13	4100	6080	9840	-

14	4700	-	9872	-
15	5310	4600	8752	-
16	4460	7320	8544	-
17	4080	7320	8544	-
18	3540	-	8904	-
19	3970	4040	9736	-
20	4840	-	7624	-
21	4050	5200	7720	-
22	4370	-	8216	-
23	4180	-	10304	-
24	4680	4960	8856	-
25	3690	-	10768	-
26	3720	7920	9200	-
27	4040	-	8128	-
28	4210	-	10008	-
29	4140	-	7768	-
30	4420	6320	9024	-
Μέσος όρος Υπολογισμών	4244	5880	9064	-

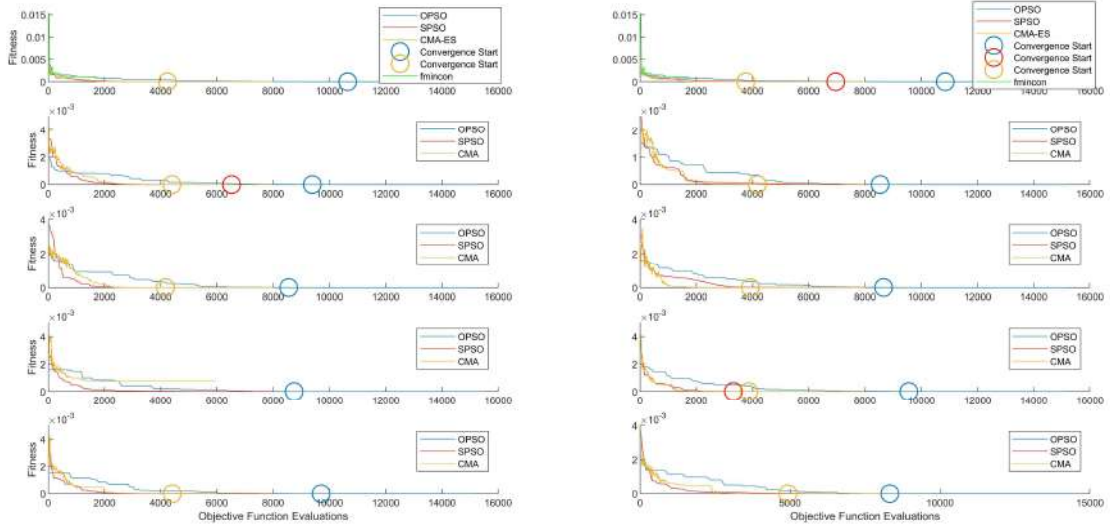
Πίνακας 9. Αποτελέσματα για το half car βάση ταχύτητας σύγκλισης.

Δοκιμή	CMA-ES	SPSO	OPSO	fmincon
1	5.74e-12	4.5655e-06	6.302e-09	0.0011189
2	6.39e-12	3.318e-07	6.183e-09	0.0011189
3	7.18e-12	6.382e-07	6.522e-09	0.0011189
4	7.82e-12	3.5463e-06	2.219e-09	0.0011189
5	6.82e-12	1.4427e-06	1.043e-09	0.0011189
6	6.01e-12	9.9935e-06	4.831e-09	0.0011189
7	0.00076330507161	9.7037e-06	3.695e-09	0.0011189
8	4.65e-12	1.495e-07	6.053e-09	0.0011189
9	6.83e-12	1.60544e-05	4.471e-09	0.0011189
10	5.64e-12	9.28e-07	1.973e-09	0.0011189
11	8.52e-12	1.7203e-06	3.022e-09	0.0011189
12	8.64e-12	3.6897e-06	3.426e-09	0.0011189
13	5.89e-12	4.006e-07	3.076e-09	0.0011189

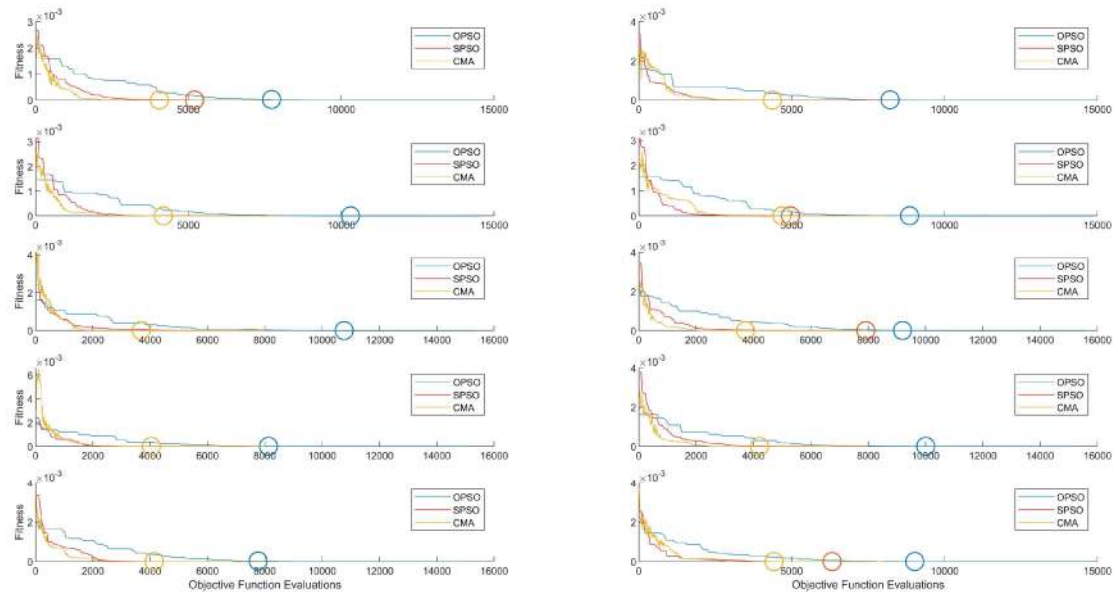
14	5.42e-12	4.8006e-06	2.223e-09	0.0011189
15	5.31e-12	3.21e-07	3.71e-09	0.0011189
16	7.36e-12	4.06e-07	3.036e-09	0.0011189
17	8.85e-12	4.06e-07	3.036e-09	0.0011189
18	5.18e-12	9.473e-07	1.0291e-08	0.0011189
19	6.39e-12	5.282e-07	9.239e-09	0.0011189
20	6.41e-12	1.70548e-05	6.921e-09	0.0011189
21	5.03e-12	5.026e-07	1.074e-09	0.0011189
22	7.82e-12	9.9372e-06	3.288e-09	0.0011189
23	6.58e-12	9.388e-07	8.108e-09	0.0011189
24	7.64e-12	1.549e-07	6.932e-09	0.0011189
25	1.418e-11	1.07014e-05	1.3802e-08	0.0011189
26	4.85e-12	8.444e-07	5.546e-09	0.0011189
27	8.9e-12	7.8969e-06	4.016e-09	0.0011189
28	4.28e-12	8.4202e-06	4.98e-09	0.0011189
29	5.37e-12	1.6793e-06	2.898e-09	0.0011189
30	7.78e-12	2.297e-07	6.097e-09	0.0011189

Πίνακας 10. Αποτελέσματα για το μοντέλο half car βάση τιμής συνάρτησης καταλληλότητας

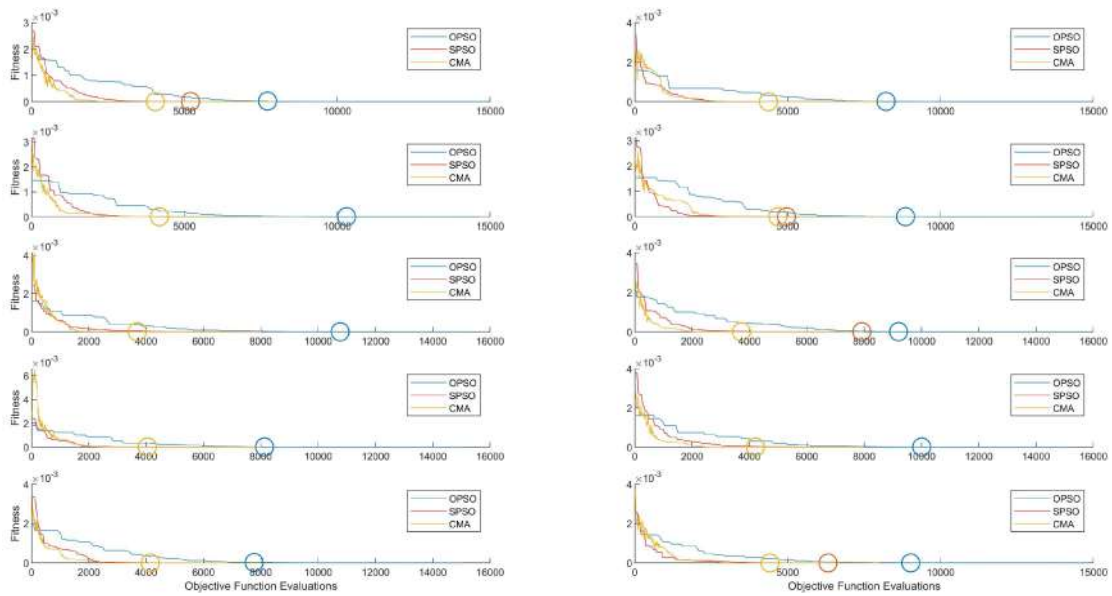
Παρακάτω παρουσιάζεται η απόδοση των αλγορίθμων συγκριτικά για τις 30 δοκιμές όπου με το σύμβολο του κύκλου συμβολίζεται το πρώτο σημείο όπου ο αλγόριθμος συνέκλινε εντός των σωστών ορίων για όλες τις παραμέτρους και στην συνέχεια μέχρι την τελευταία επανάληψη παρέμεινε εντός αυτών των ορίων. Η απόδοση της fmincon παρουσιάζεται μόνο για τις δύο πρώτες δοκιμές διότι η απόδοση της είναι σταθερή, όντας αιτιοκρατική και επειδή η συμπερίληψη της μεταβάλλει την κλίμακα των αξόνων ώστε να μην διακρίνονται τα χαρακτηριστικά των υπόλοιπων αλγορίθμων.



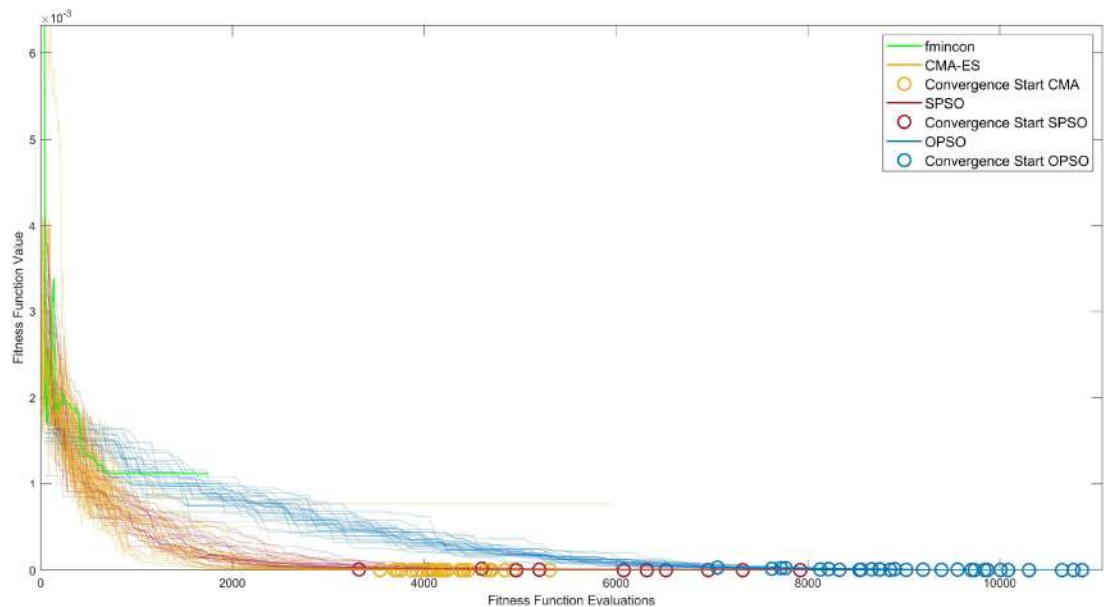
Εικόνα 31. Αποτελέσματα μεθόδων βάση τιμής αντικειμενικής συνάρτησης για αριθμό κλήσεων αντικειμενικής συνάρτησης.



Εικόνα 32. Αποτελέσματα μεθόδων βάση τιμής αντικειμενικής συνάρτησης για αριθμό κλήσεων αντικειμενικής συνάρτησης.

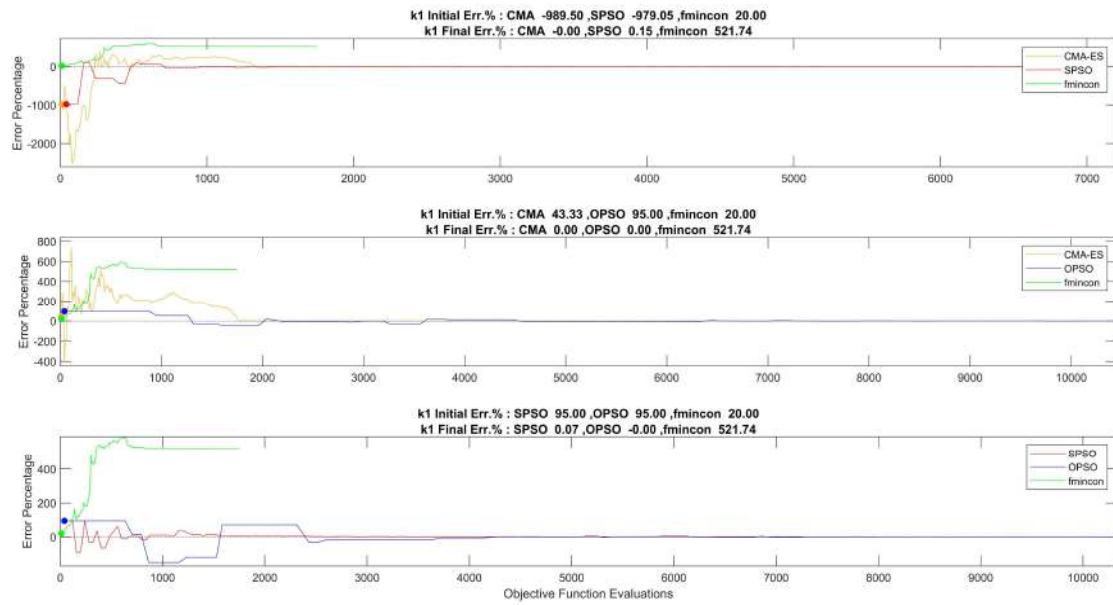


Εικόνα 33. Αποτελέσματα μεθόδων βάσει τιμής αντικειμενικής συνάρτησης για αριθμό κλήσεων αντικειμενικής συνάρτησης

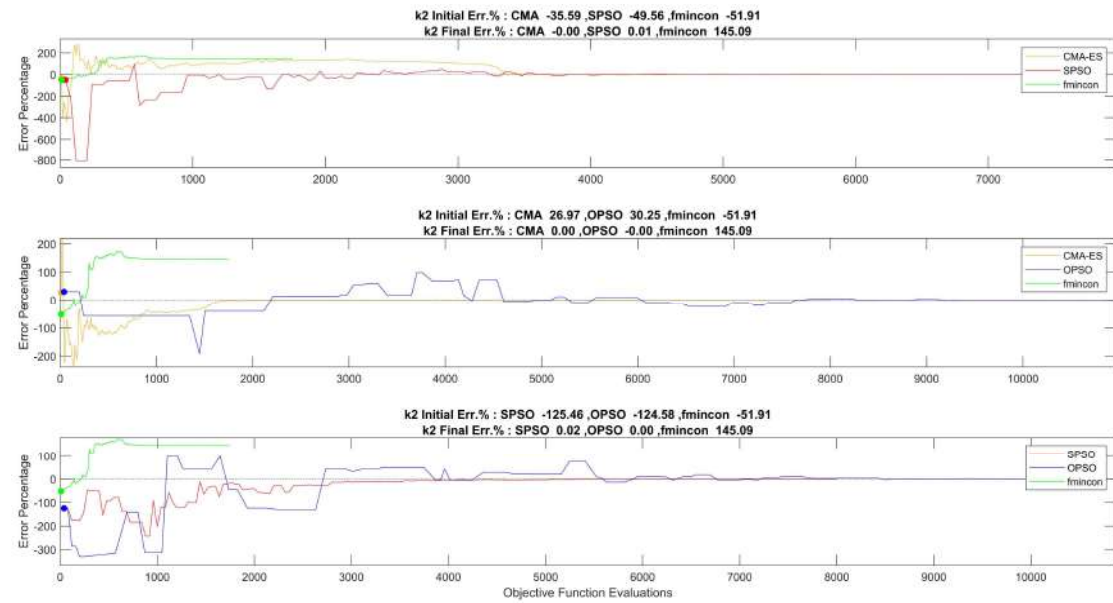


Εικόνα 344. Αποτελέσματα μεθόδων για όλες τις δοκιμές μαζί, βάσει τιμής αντικειμενικής συνάρτησης για αριθμό κλήσεων αντικειμενικής συνάρτησης.

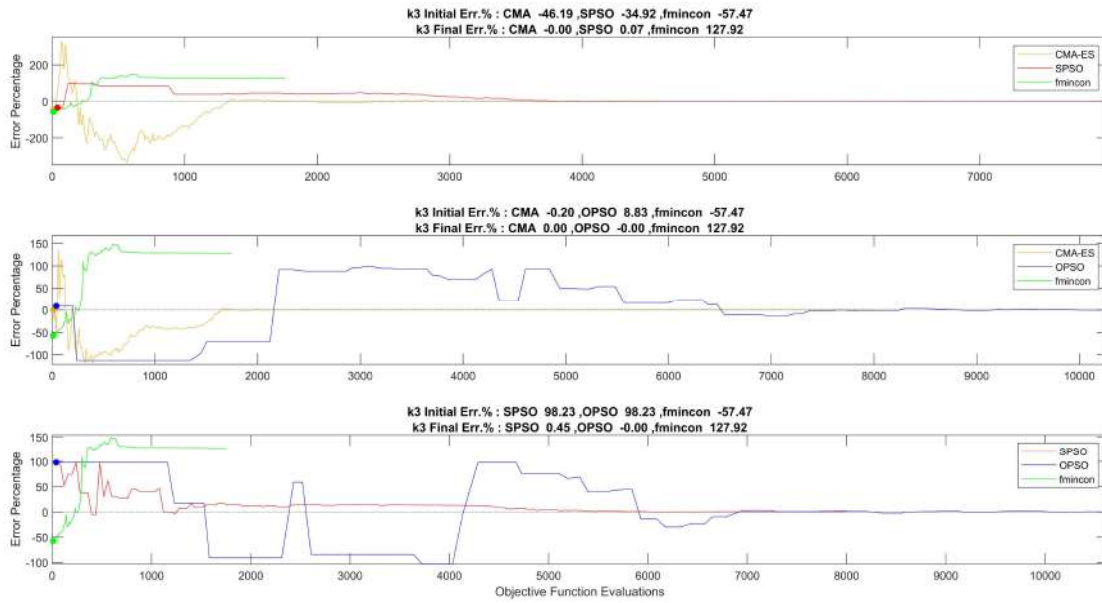
Στην συνέχεια παρουσιάζουμε την μεταβολή των παραμέτρων ξεχωριστά. Επιλέγουμε για κάθε μέθοδο τις αντίστοιχες δοκιμές ώστε τα σημεία εκκίνησης για κάθε μέθοδο να είναι όσο το δυνατόν κοντά. Σημειώνουμε πως μπορεί να συμπεριλαμβάνονται για τα παρακάτω διαγράμματα περιπτώσεις όπου δεν αναγνωρίστηκε σωστά η παράμετρος για κάποια μέθοδο.



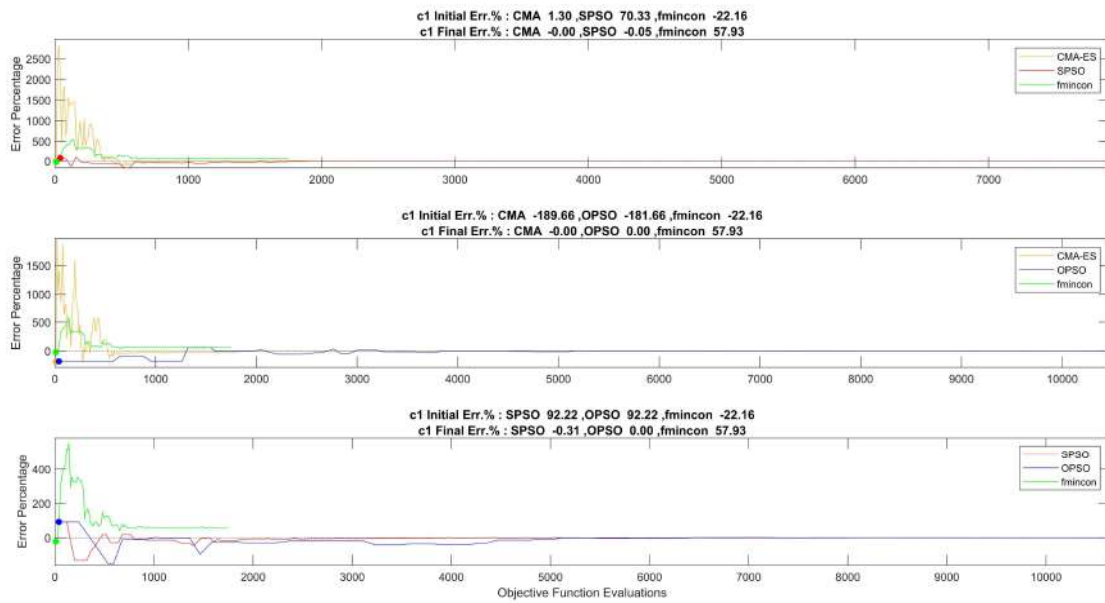
Εικόνα 35. Αναζήτηση Παραμέτρου k_1 για όλες τις μεθόδους



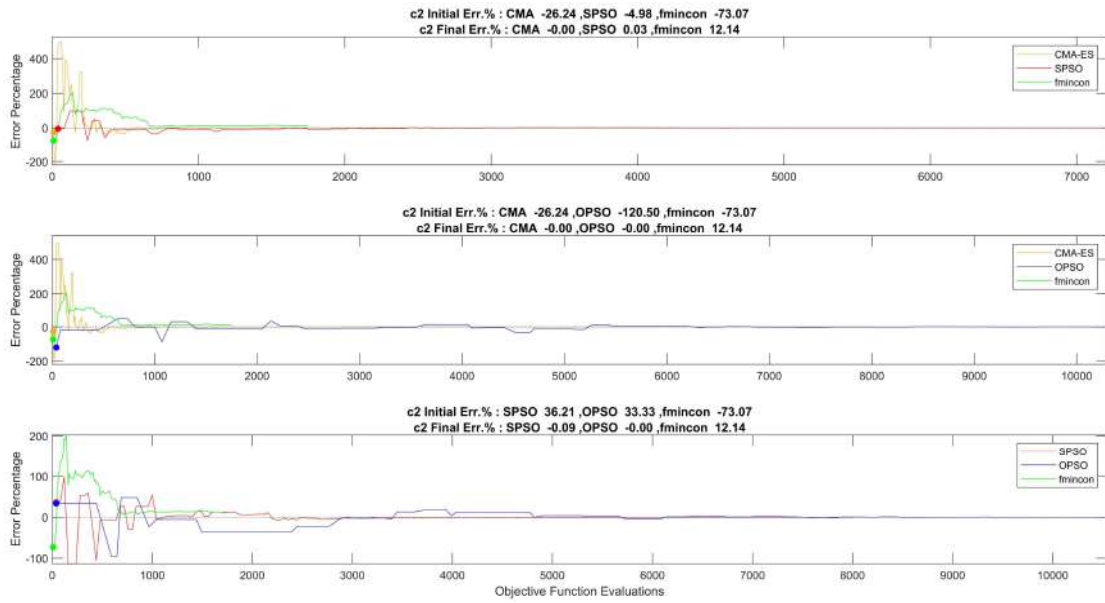
Εικόνα 36. Αναζήτηση Παραμέτρου k_2 για όλες τις μεθόδους



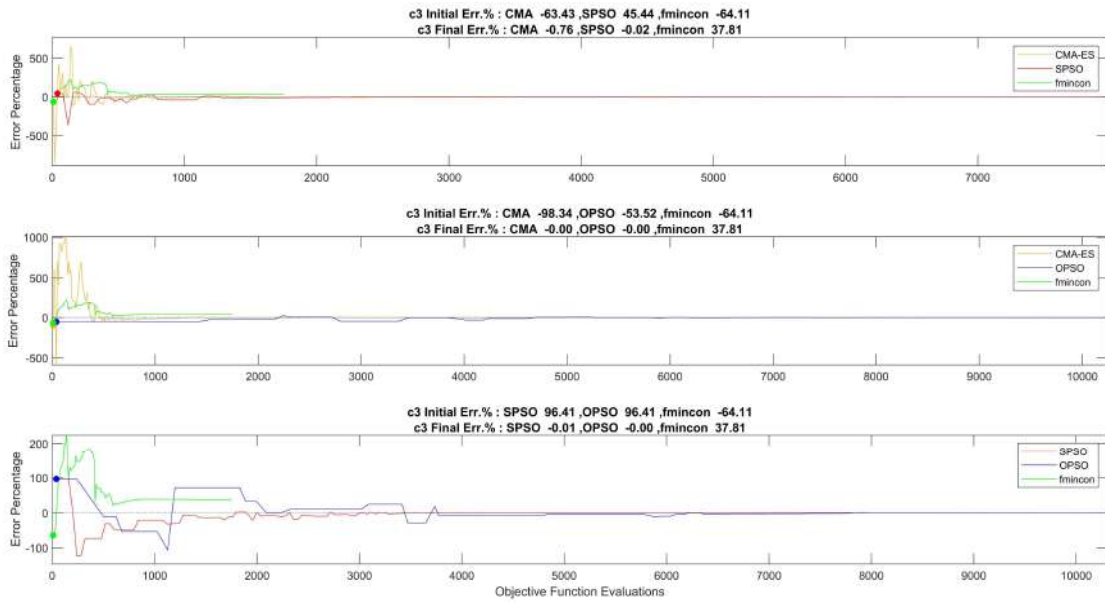
Εικόνα 37. Αναζήτηση Παραμέτρου k_3 για όλες τις μεθόδους



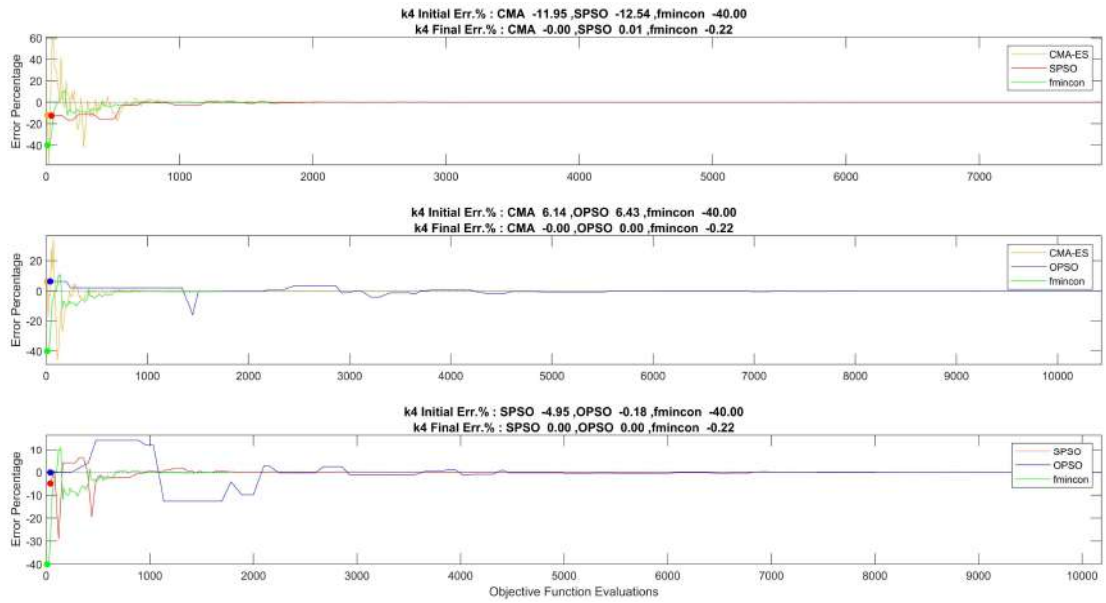
Εικόνα 38. Αναζήτηση Παραμέτρου c_1 για όλες τις μεθόδους



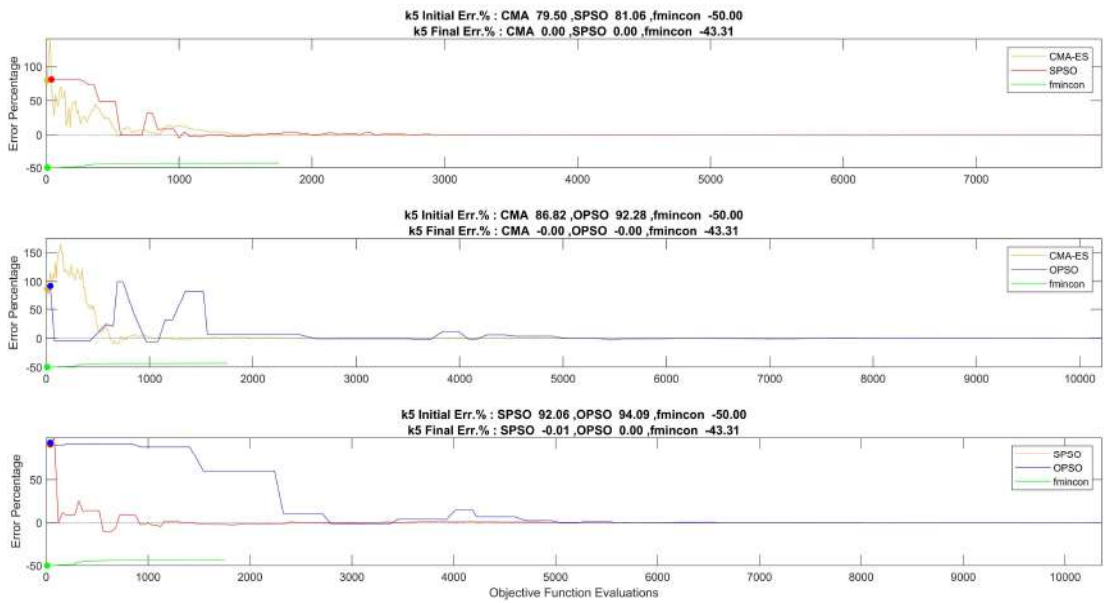
Εικόνα 39. Αναζήτηση Παραμέτρου c2 για όλες τις μεθόδους



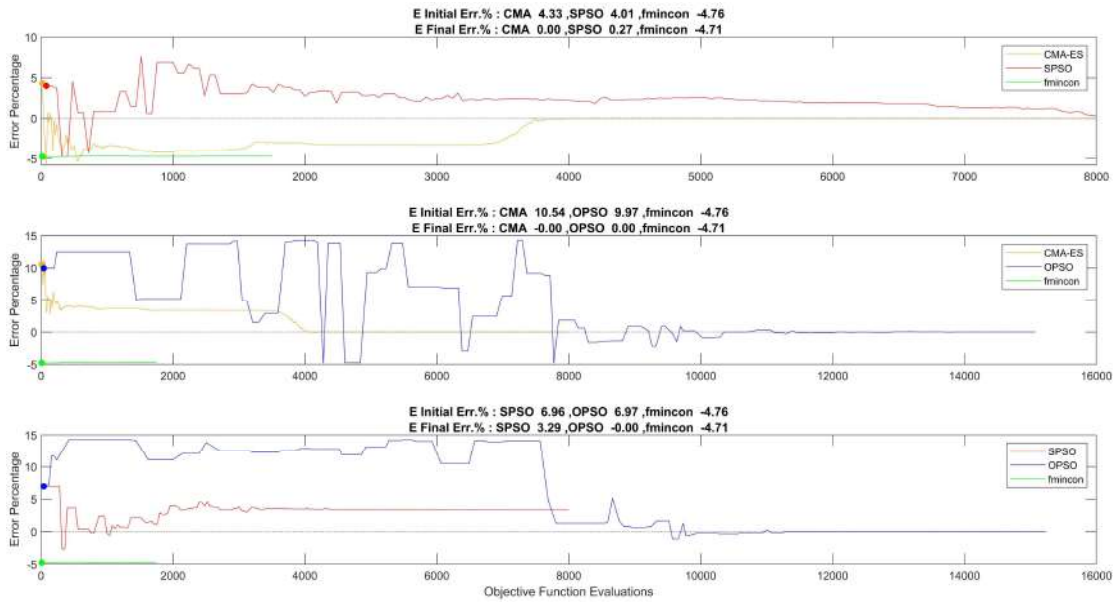
Εικόνα 40. Αναζήτηση Παραμέτρου c3 για όλες τις μεθόδους



Εικόνα 41. Αναζήτηση Παραμέτρου k4 για όλες τις μεθόδους

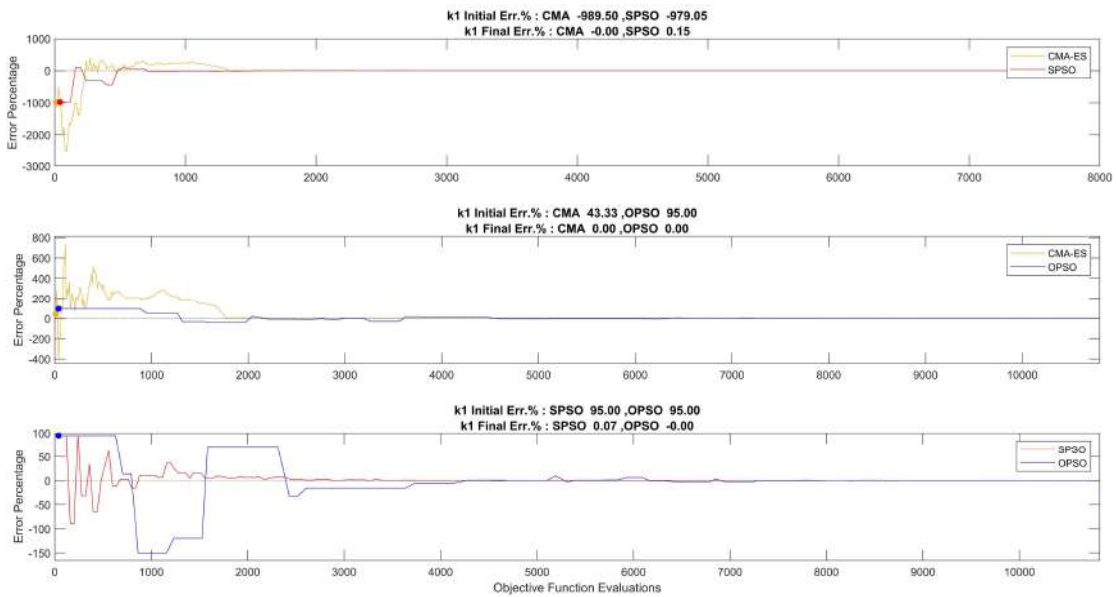


Εικόνα 42. Αναζήτηση Παραμέτρου k5 για όλες τις μεθόδους

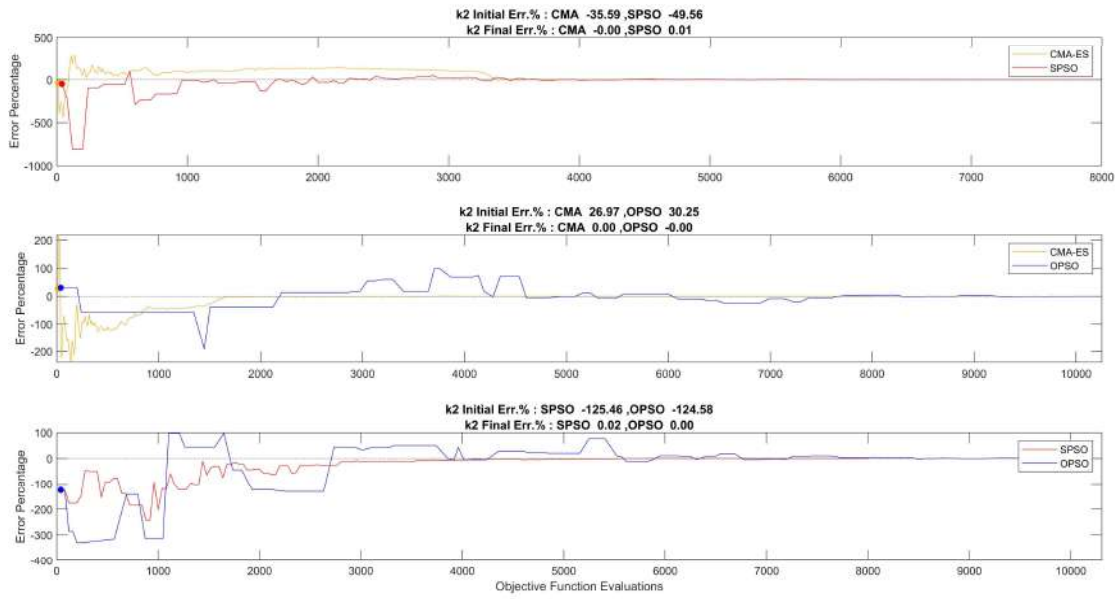


Εικόνα 43. Αναζήτηση Παραμέτρου E για όλες τις μεθόδους

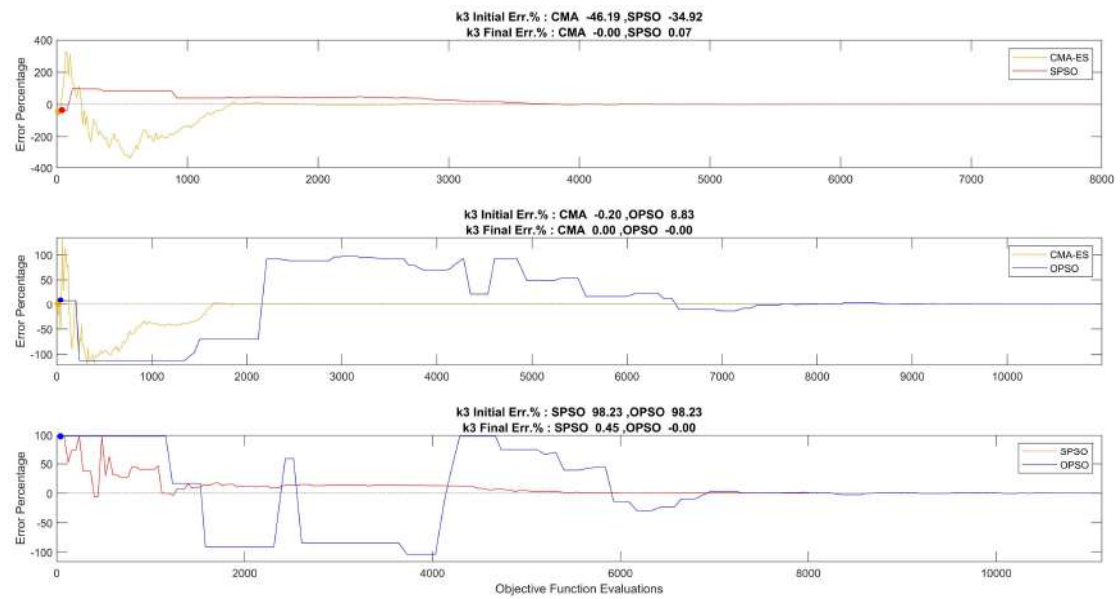
Παρακάτω παρουσιάζουμε διαγράμματα σαν τα προηγούμενα χωρίς την fmincon ώστε να συγκρίνουμε καλύτερα τις υπόλοιπες μεθόδους.



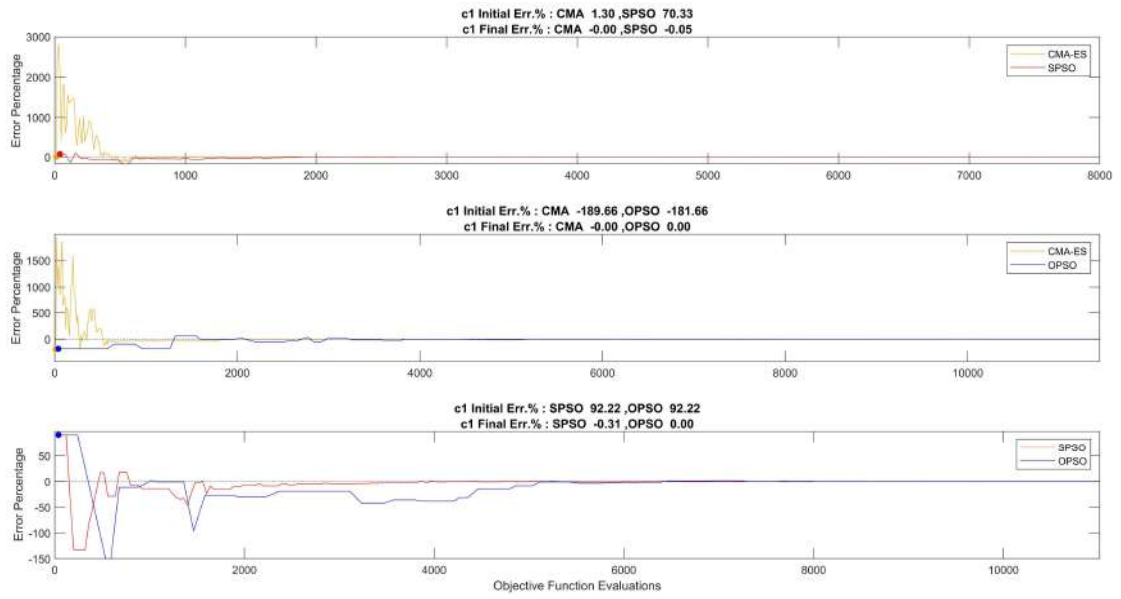
Εικόνα 44. Αναζήτηση Παραμέτρου k1



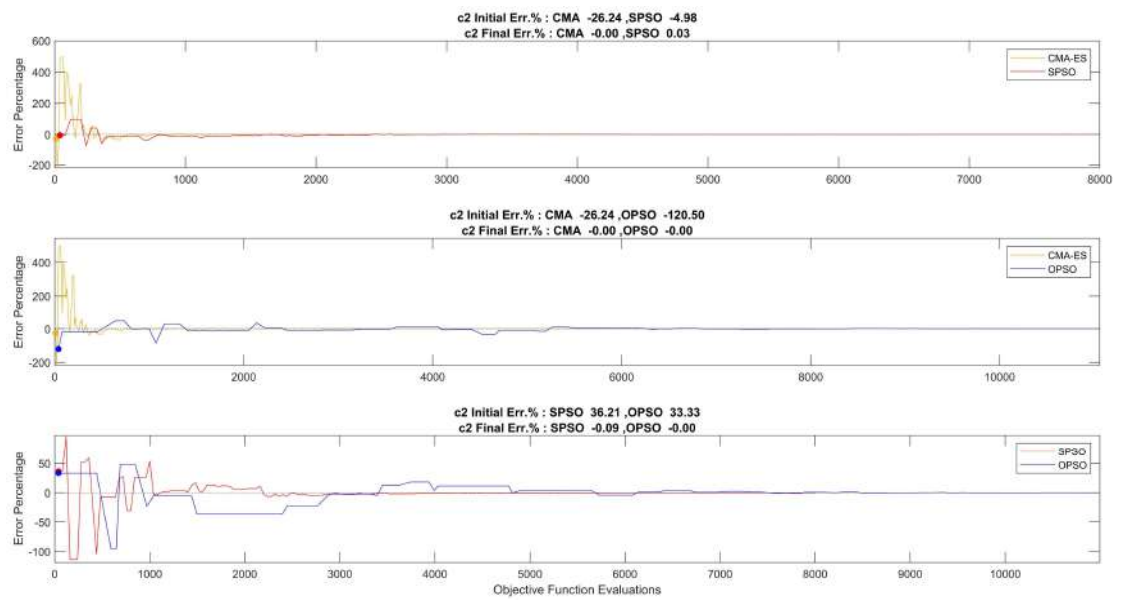
Εικόνα 45. Αναζήτηση Παραμέτρου k2



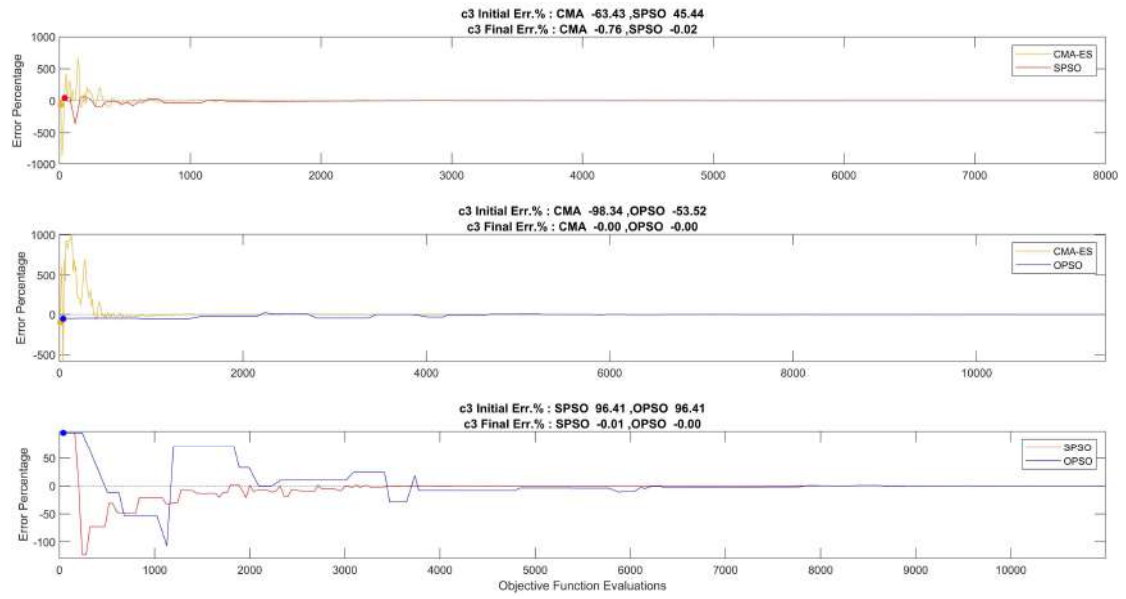
Εικόνα 46. Αναζήτηση Παραμέτρου k3



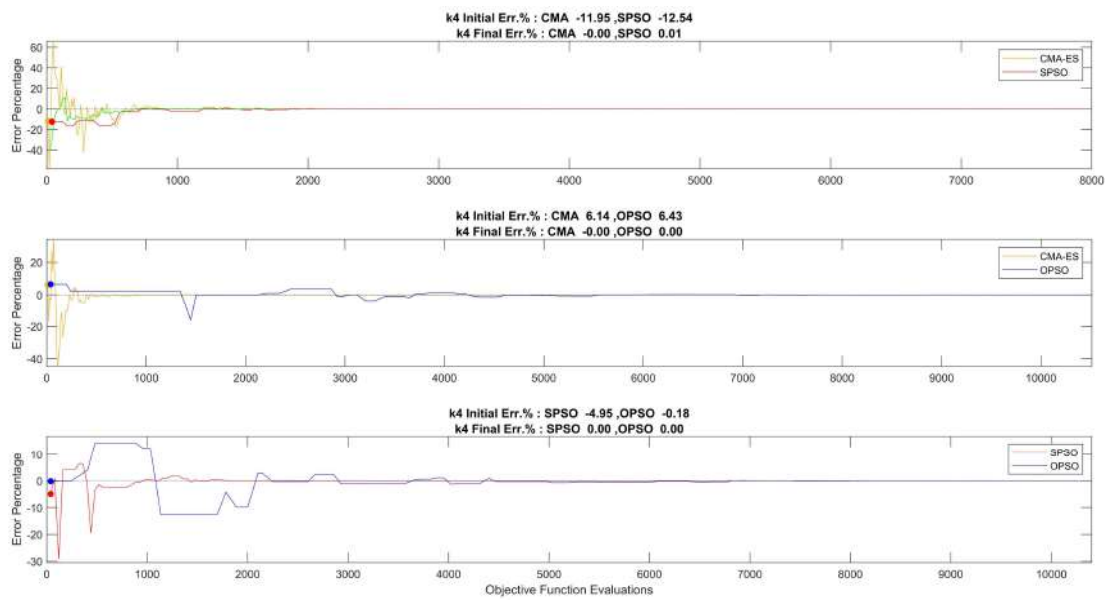
Εικόνα 47. Αναζήτηση Παραμέτρου c1



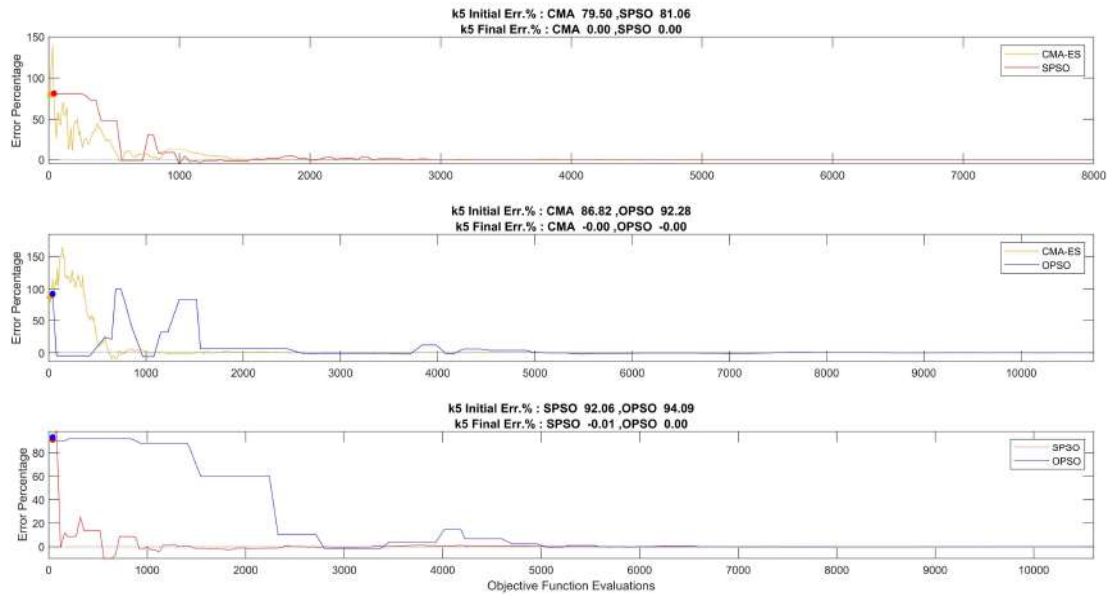
Εικόνα 48. Αναζήτηση Παραμέτρου c2



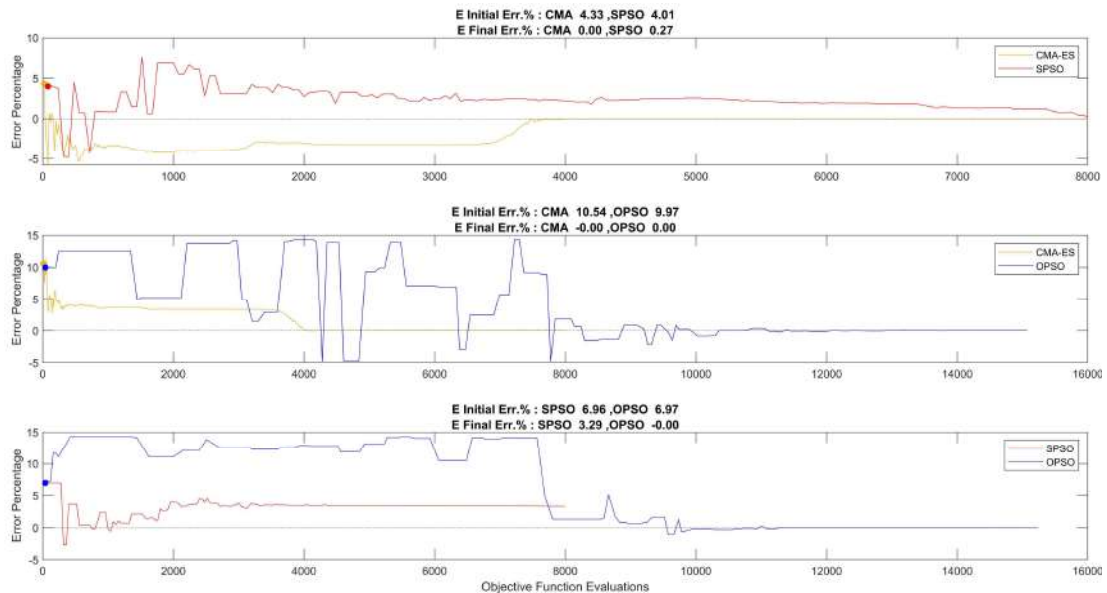
Εικόνα 49. Αναζήτηση Παραμέτρου c3



Εικόνα 50. Αναζήτηση Παραμέτρου k4



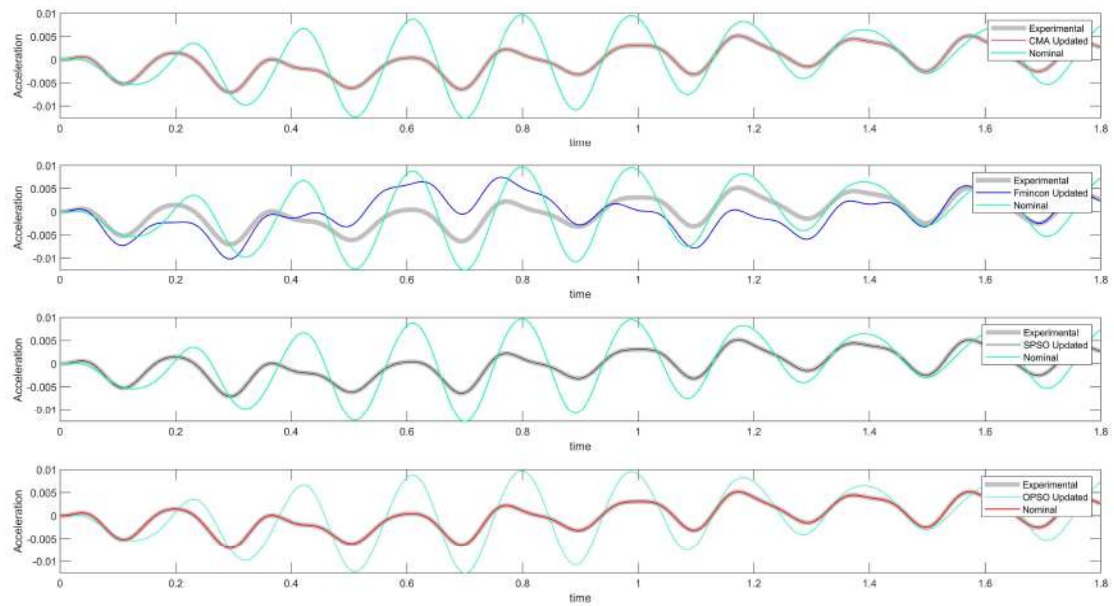
Εικόνα 51. Αναζήτηση Παραμέτρου $k5$



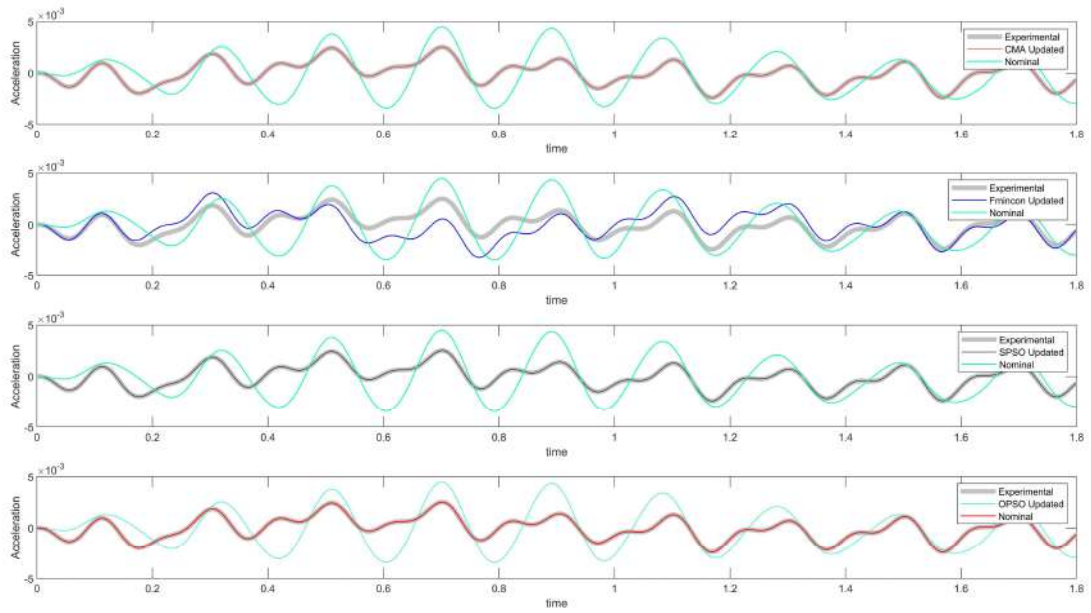
Εικόνα 52. Αναζήτηση Παραμέτρου E

Εύκολα παρατηρούμε πως στις πρώτες επαναλήψεις η CMA-ES εφαρμόζει έναν μεγάλο βαθμό εξερεύνησης ξεφεύγοντας από το εγγύς διάστημα γύρω από το σημείο εκκίνησης. Παρ' όλα καθώς εξελίσσονται οι γενιές του πληθυσμού οι παράμετροι της προσαρμόζονται και επιστρέφει σύντομα κοντά στην περιοχή βέλτιστης τιμής για τις παραμέτρους. Καθώς πλησιάζει την βέλτιστη τιμή η εξερεύνηση της διατηρείται αλλά σε μικρότερο διάστημα, πράγμα που απεικονίζεται με τις μικρές αλλαγές στην καμπύλη που μοιάζει σαν να ταλαντώνεται με πολύ μικρό εύρος γύρω από μια τιμή. Η OPSO πραγματοποιεί πιο απότομες αλλαγές διότι αφού είναι μια *gbest* PSO, οι τιμές για διαγράμματα προέρχονται από το καθολικά βέλτιστο σωματίδιο. Αυτές οι απότομες αλλαγές ενισχύονται από την αλλαγή του διανύσματος μέσω του ορθογώνιο σχεδιασμού πειράματος όταν οι *pbest* θέσεις φτάνουν σε τέλος.

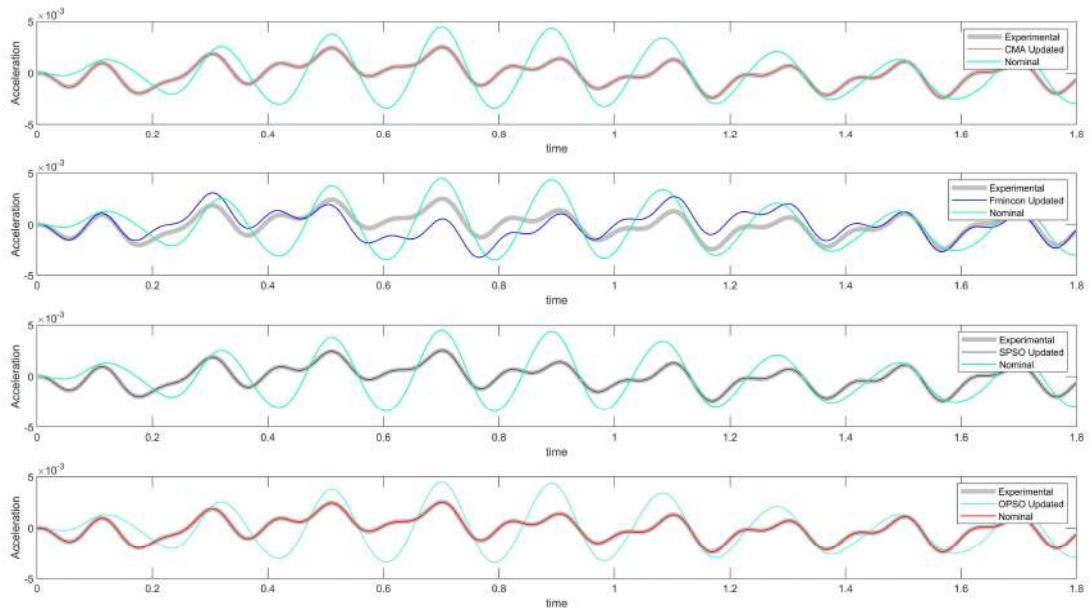
Στα παρακάτω διαγράμματα παρουσιάζεται η απόκριση της επιτάχυνσης στον χρόνο μερικών επιλεγμένων κόμβων, χρησιμοποιώντας για τιμές των παραμέτρων αυτές που αναγνώρισαν οι μέθοδοι στην πιο αποτελεσματική δοκιμή τους. Με γκρι χρώμα είναι η πραγματική απόκριση (ή αλλιώς η πειραματική) η οποία συγκρίνεται για κάθε μέθοδο μέθοδο ξεχωριστά.



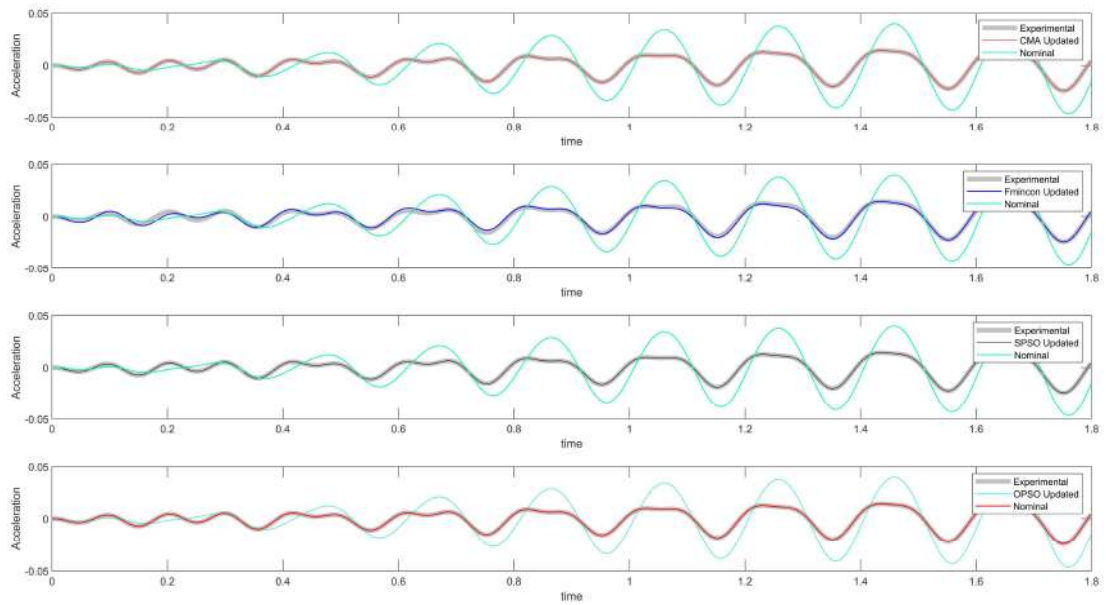
Εικόνα 53. Απόκριση βαθμού ελευθερίας 1



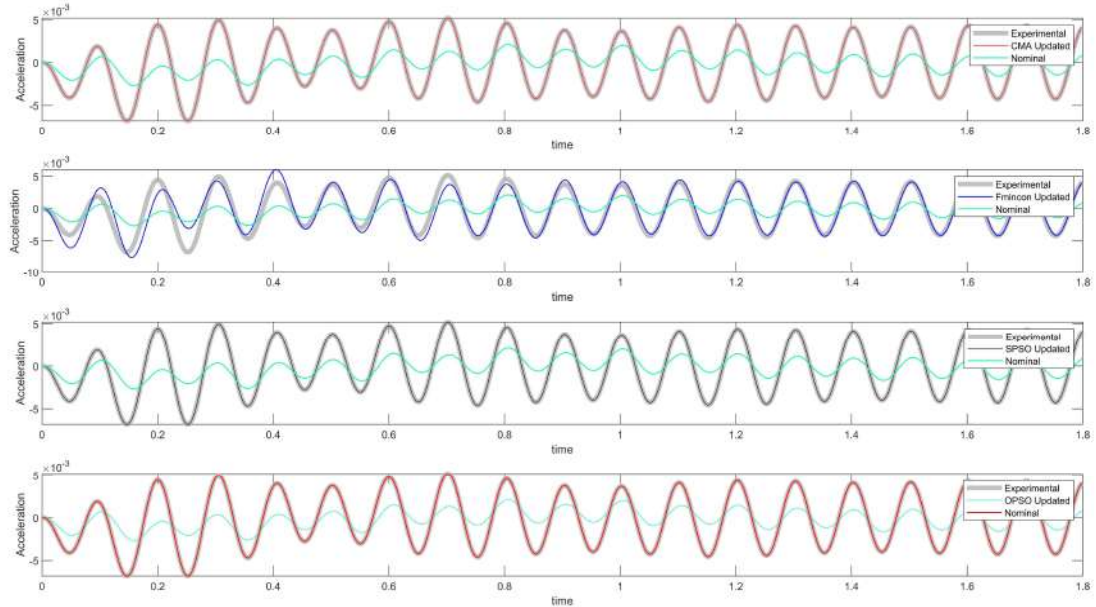
Εικόνα 54. Απόκριση βαθμού ελευθερίας 2



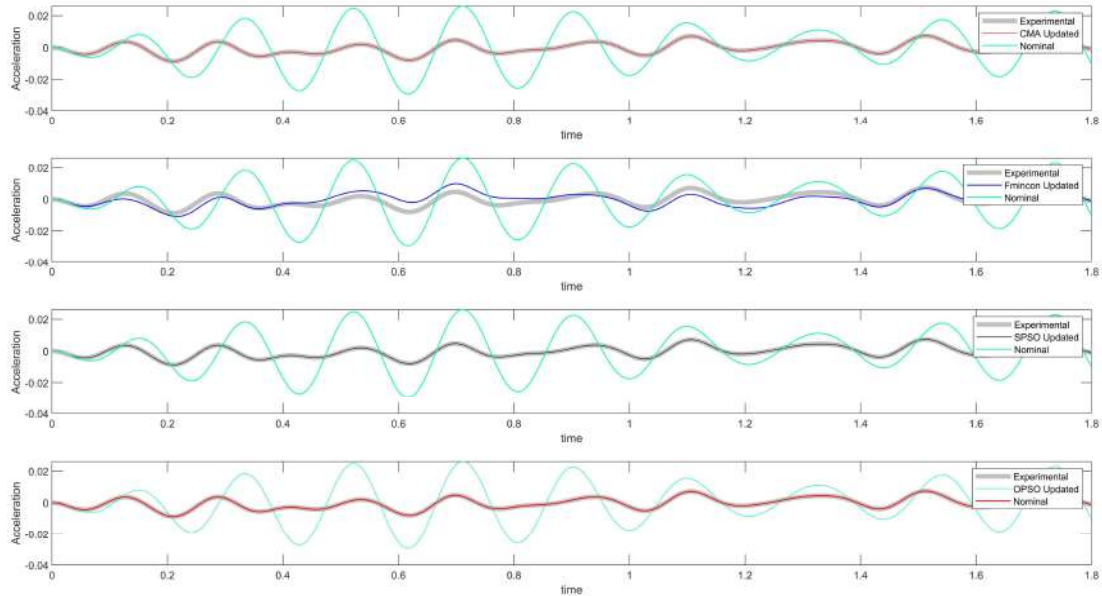
Εικόνα 55. Απόκριση βαθμού ελευθερίας 3



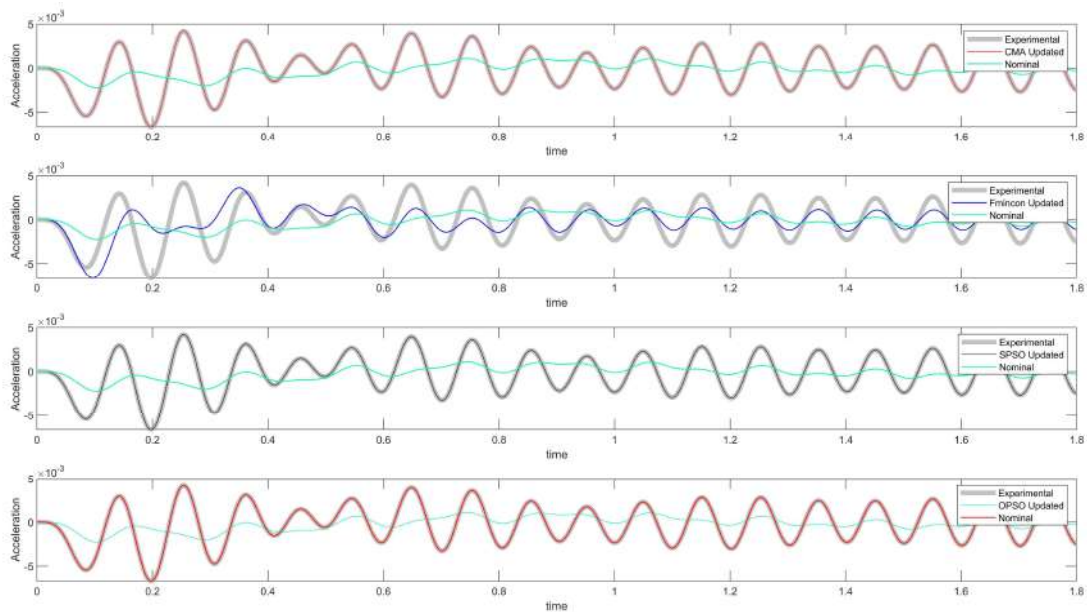
Εικόνα 56. Απόκριση βαθμού ελευθερίας 19



Εικόνα 57. Απόκριση βαθμού ελευθερίας 23



Εικόνα 58. Απόκριση βαθμού ελευθερίας 32



Εικόνα 59 Απόκριση βαθμού ελευθερίας 33

7.2.1 Δοκιμή μοντέλου half-car με τυχαίο αρχικό σημείο.

Στους PSO αλγόριθμους έπειτα από την αρχικοποίηση απουσιάζει κάποια εκτεταμένη διαδικασία τυχαιοποίησης. Οπότε η απόδοση των αλγορίθμων αυτών εξαρτάται σε μεγάλο βαθμό από την αρχικοποίηση του πληθυσμού. Για να είναι η σύγκριση πιο δίκαια δοκιμάσαμε ξανά τις μεθόδους χρησιμοποιώντας για τις PSO αρχικοποίηση πληθυσμού με σωματίδια που εκλέγουμε από ομοιόμορφη κατανομή ανάμεσα στα όρια του χώρου αναζήτησης. Ομοίως για τις CMA-ES και fmincon το αρχικό σημείο εκλέγεται από δειγματοληψία από ομοιόμορφη κατανομή εντός των ορίων του χώρου. Με αυτή την ομοιόμορφη αρχικοποίηση, η απόδοση των PSO βελτιώθηκε και ο αλγόριθμος SPSO πετυχαίνει αυτή τη φορά να έχει 25 επιτυχής δοκιμές έναντι των 12 που είχε προηγουμένως, συγκλίνοντας μάλιστα πιο γρήγορα. Ο

αλγόριθμος OPSO βελτίωσε και αυτός την απόδοση του φτάνοντας σε σύγκλιση μετά περίπου από 1000 επαναλήψεις λιγότερες κατά μέσο όρο όπως και η CMA-ES.

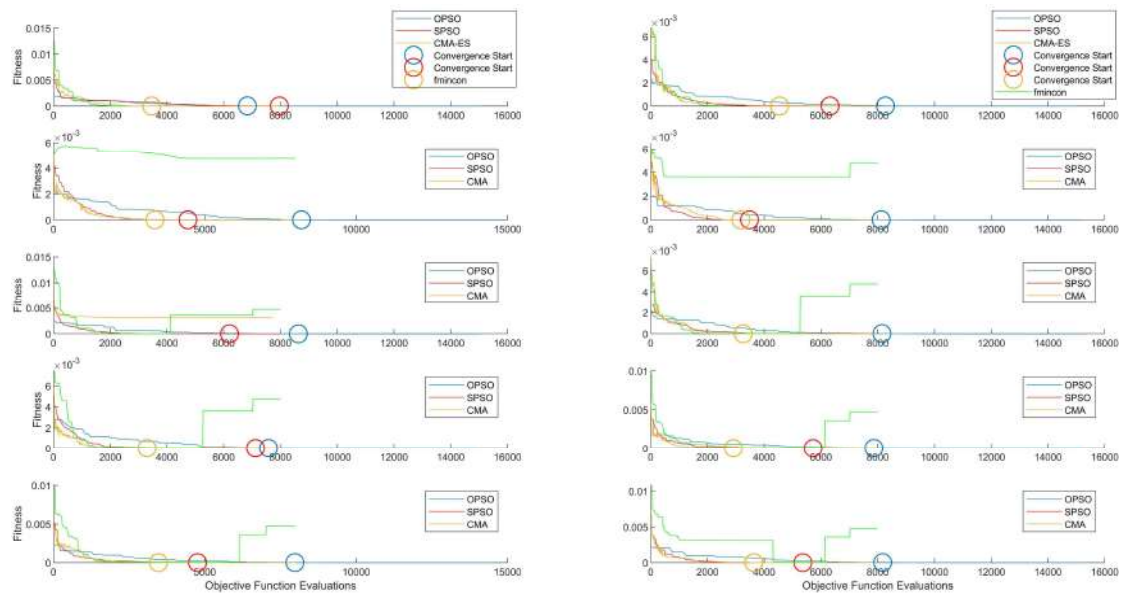
Δοκιμή	CMA-ES	SPSO	OPSO	fmincon
1	3460	7960	6832	-
2	4540	6320	8272	-
3	3350	4440	8192	-
4	3180	3480	8120	-
5	-	6200	8632	-
6	3260	-	8152	-
7	3300	7120	7576	-
8	2910	5720	7864	-
9	3470	4760	7976	-
10	3630	5360	8176	-
11	3900	-	7136	-
12	3290	4360	8688	-
13	3230	-	8024	-
14	2540	4640	8056	-
15	4700	-	7456	-
16	2900	4960	7472	-
17	-	4240	8104	-
18	2870	7760	7848	-
19	2960	4400	8576	-
20	3730	7080	8704	-
21	3190	5080	8472	-
22	4080	6320	8272	-
23	3340	-	8440	-
24	4160	6800	7056	-
25	3560	5520	8944	-
26	3180	5720	8080	-
27	3280	6320	8272	-
28	2840	6320	8272	-
29	3600	4440	8192	-
30	3740	6320	8272	-
Μέσος όρος Υπολογισμών	3435,3	5665,6	8070	-

Εικόνα 60. Σύγκριση ταχύτητας σύγκλισης για το μοντέλο half car με τυχαίο αρχικό σημείο για CMA και fmincon και πληθυσμός ομοιόμορφης κατανομής για τους PSO.

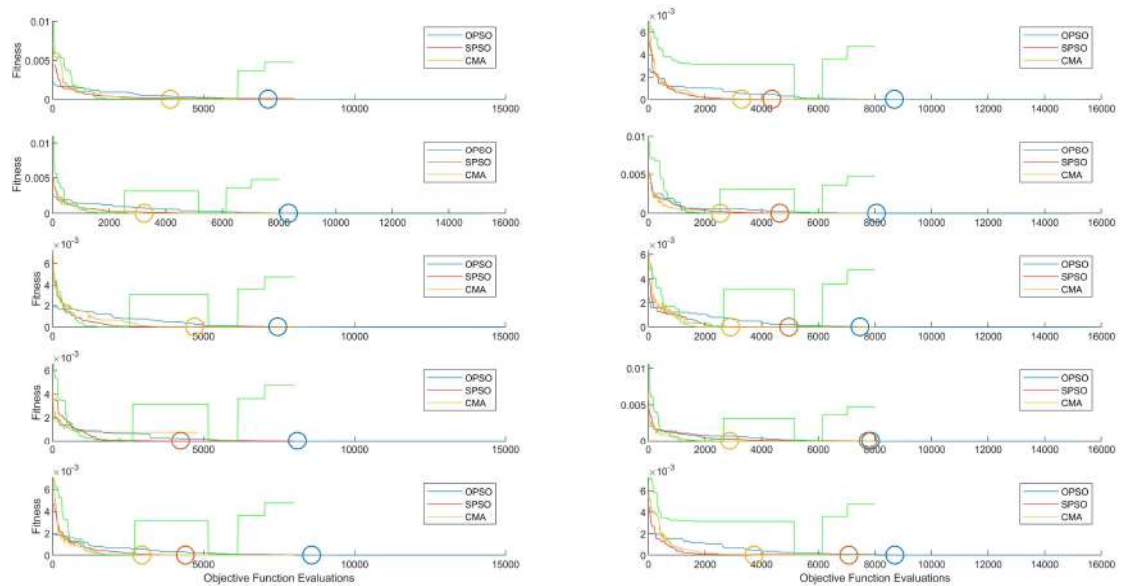
Δοκιμή	CMA-ES	SPSO	OPSO	fmincon
1	6.51e-12	2.14e-5	2.24e-05	0,00026
2	5.79e-12	5.30e-6	2.48e-05	0,00026
3	5.48e-12	6.82e-7	1.34e-05	0,00026
4	3.38e-12	1.63e-8	1.868e-05	0,00026
5	0.003128	1.26e-6	6.96e-06	0,00026
6	9.24e-12	2.49e-5	2.75e-05	0,00026
7	6.92e-12	1.26e-6	1.67e-05	0,00026
8	6.82e-12	1.47e-6	3.32e-05	0,00026
9	6.38e-12	5.43e-6	2.10e-05	0,00026
10	4.78e-12	3.26e-6	3.20e-05	0,00026
11	3.94e-12	9.25e-5	1.41e-05	0,00026
12	5.20e-12	7.3e-7	2.51e-05	0,00026
13	4.67e-12	2.84e-5	2.16e-05	0,00026
14	6.63e-12	1.83e-6	1.27e-05	0,00026
15	5.305e-12	2.43e-5	2.54e-05	0,00026
16	5.82e-12	1.25e-6	1.36e-05	0,00026
17	0.0007633	7.92e-7	1.094e-05	0,00026
18	7.63e-12	1.32e-5	2.03e-05	0,00026
19	6.28e-12	2.84e-7	2.12e-05	0,00026
20	7.62e-12	4.5e-6	4.70e-05	0,00026
21	8.19e-12	6.40e-7	2.02e-05	0,00026
22	7.20e-12	5.3e-6	2.48e-05	0,00026
23	1.19e-11	5.88e-5	1.50-05	0,00026
24	4.60e-12	1.48e-5	2.30e-05	0,000162
25	6.06e-12	9.09e-6	5.48e-05	0,000162
26	1.35e-11	3.22e-6	1.69e-05	0,000162
27	5.73e-12	5.30e-6	2.48e-05	0,000162
28	5.80e-12	5.30e-6	2.48e-05	0,000162
29	6.58e-12	6.82e-7	1.34e-05	0,000324
30	7.75e-12	5.30e-6	2.48e-05	0,000324

Εικόνα 61. Τιμή καταλληλότητας για το μοντέλο half car με τυχαίο αρχικό σημείο για CMA και fmincon και πληθυσμός ομοιόμορφης κατανομής για τους PSO

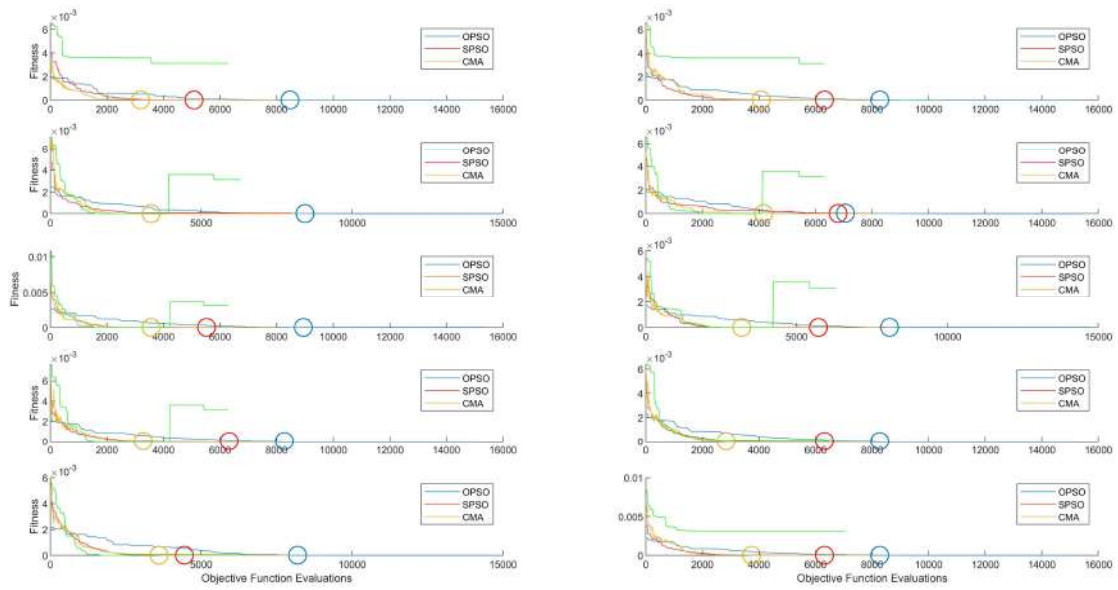
Όπως και πριν παρουσιάζεται η απόδοση των αλγορίθμων συγκριτικά για κάθε δοκιμή, στην συνέχεια παρουσιάζουμε την μεταβολή των παραμέτρων ξεχωριστά διατηρώντας αυτή φορά και την απόδοση της fitness καθώς η απόδοση της πλέον αλλάζει με την αλλαγή του interior point.



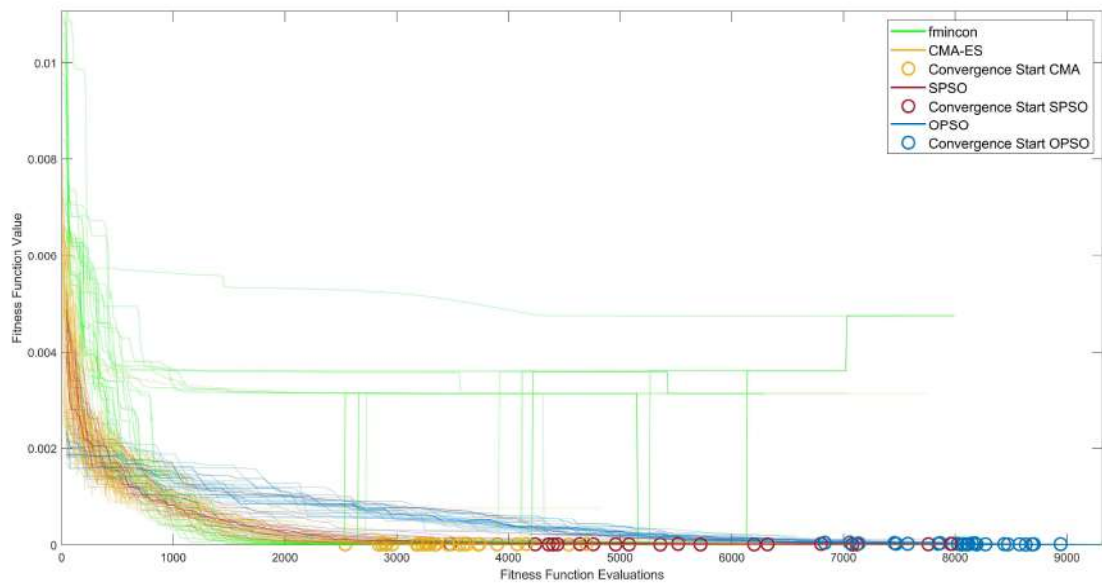
Εικόνα 62. Αποτελέσματα μεθόδων βάση τιμής αντικειμενικής συνάρτησης για αριθμό κλήσεων αντικειμενικής συνάρτησης.



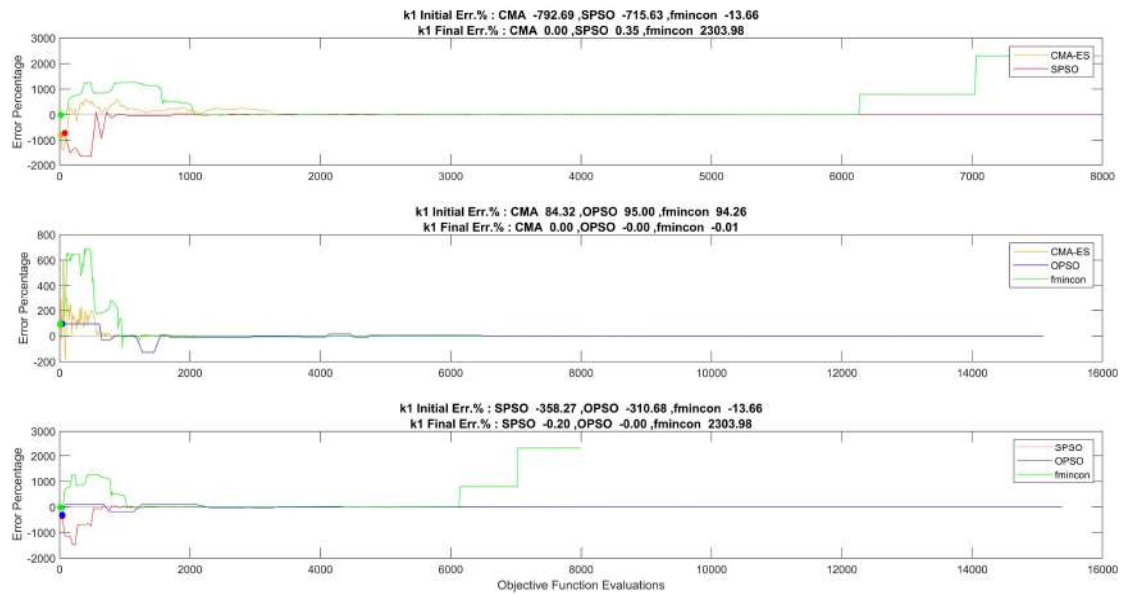
Εικόνα 63. Αποτελέσματα μεθόδων βάση τιμής αντικειμενικής συνάρτησης για αριθμό κλήσεων αντικειμενικής συνάρτησης.



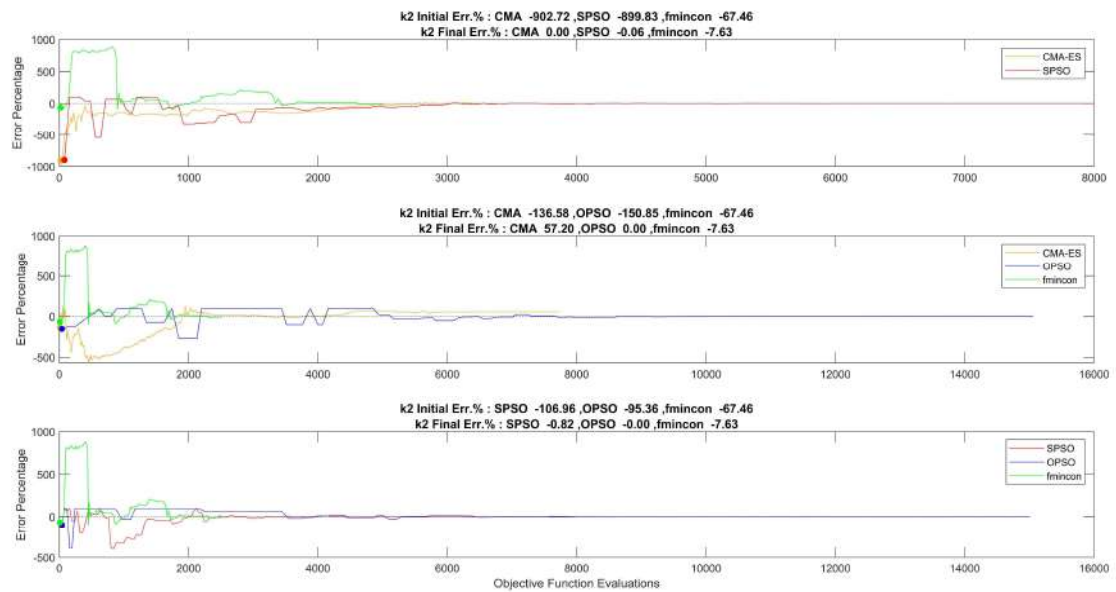
Εικόνα 64. Αποτελέσματα μεθόδων βάσει τιμής αντικειμενικής συνάρτησης για αριθμό κλήσεων αντικειμενικής συνάρτησης.



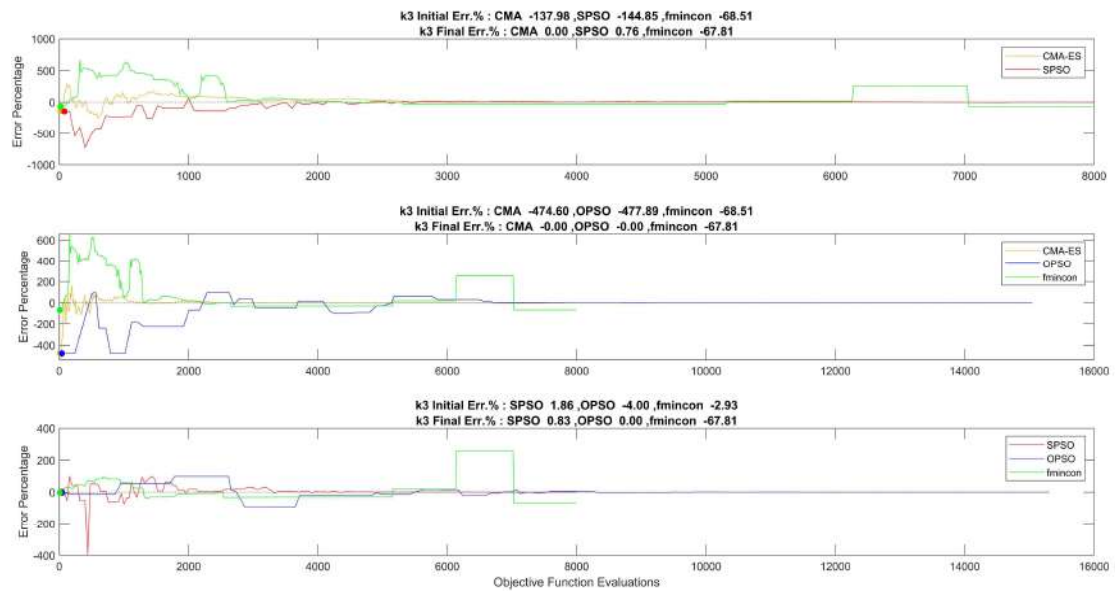
Εικόνα 65. Αποτελέσματα μεθόδων για όλες τις δοκιμές μαζί, βάσει τιμής αντικειμενικής συνάρτησης για αριθμό κλήσεων αντικειμενικής συνάρτησης.



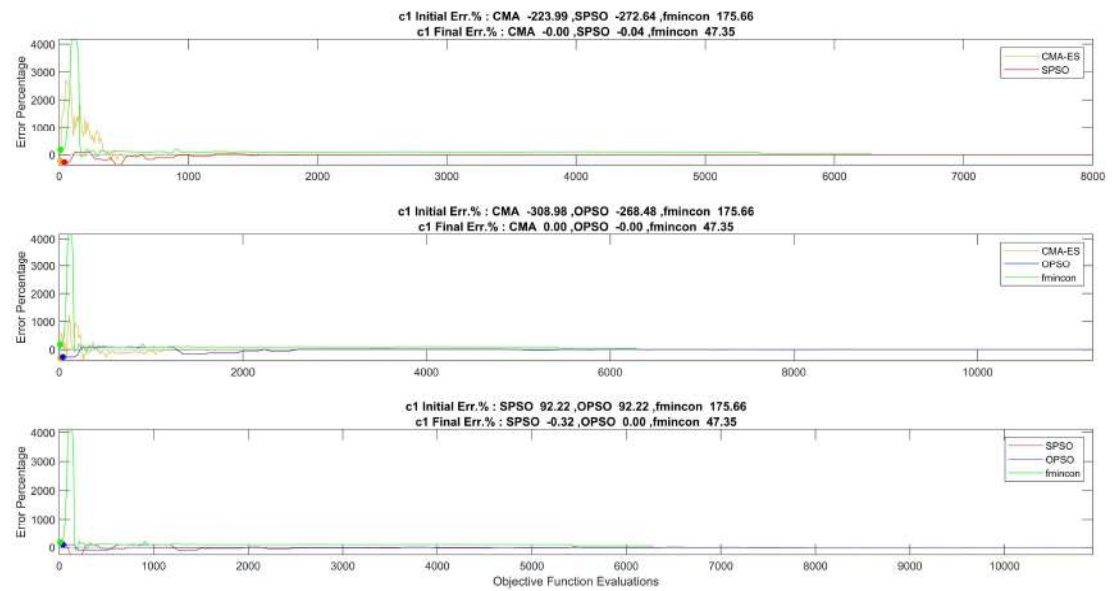
Εικόνα 66 Αναζήτηση Παραμέτρου k1



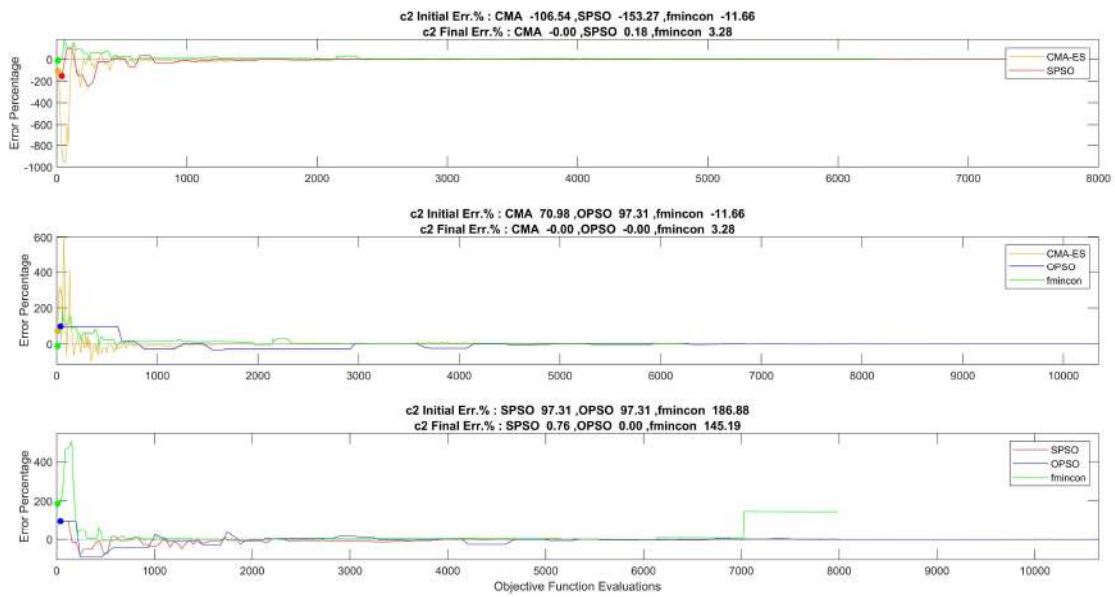
Εικόνα 67. Αναζήτηση Παραμέτρου k2



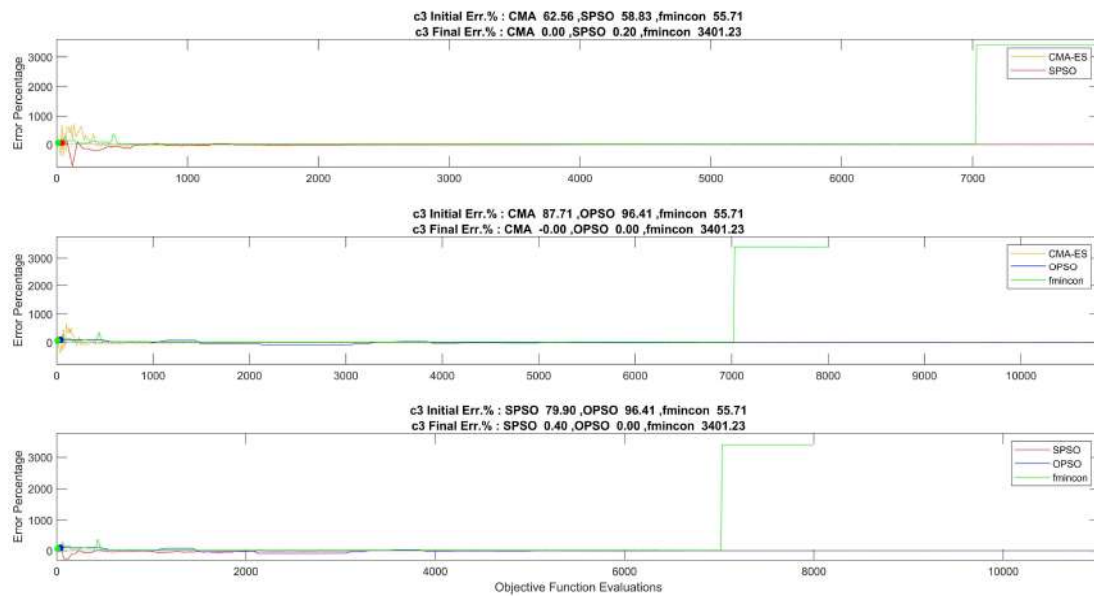
Εικόνα 68. Αναζήτηση Παραμέτρου k_3



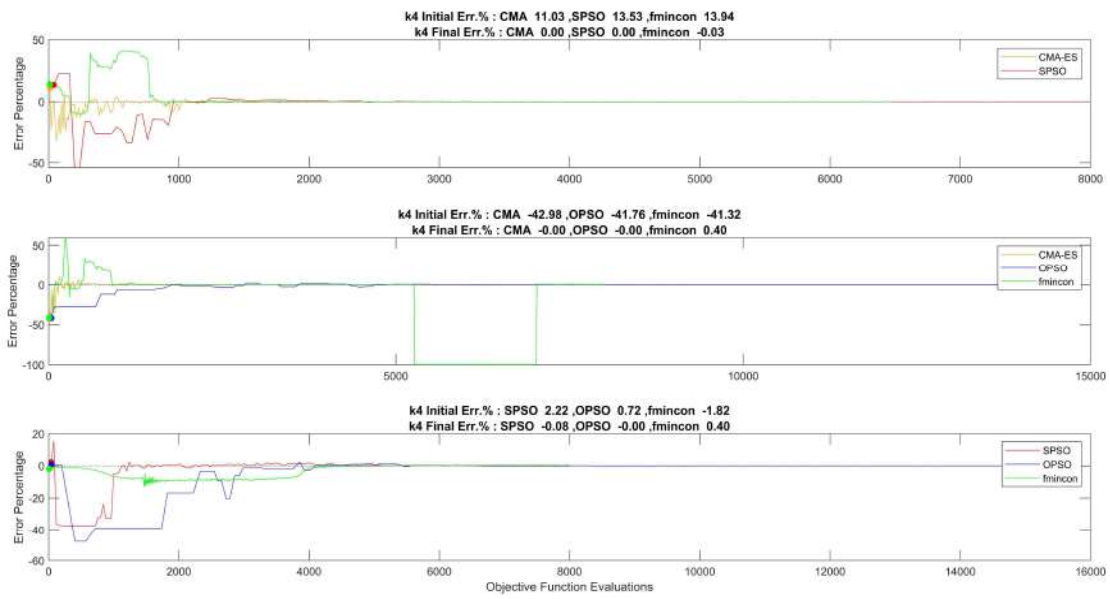
Εικόνα 69. Αναζήτηση Παραμέτρου c_1



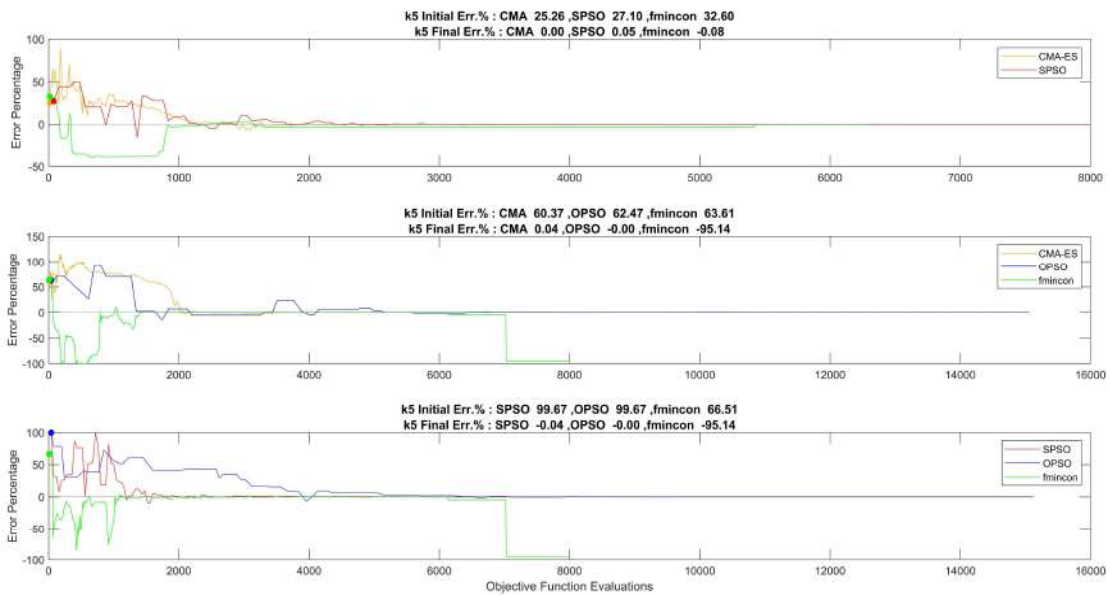
Εικόνα 70. Αναζήτηση Παραμέτρου c2



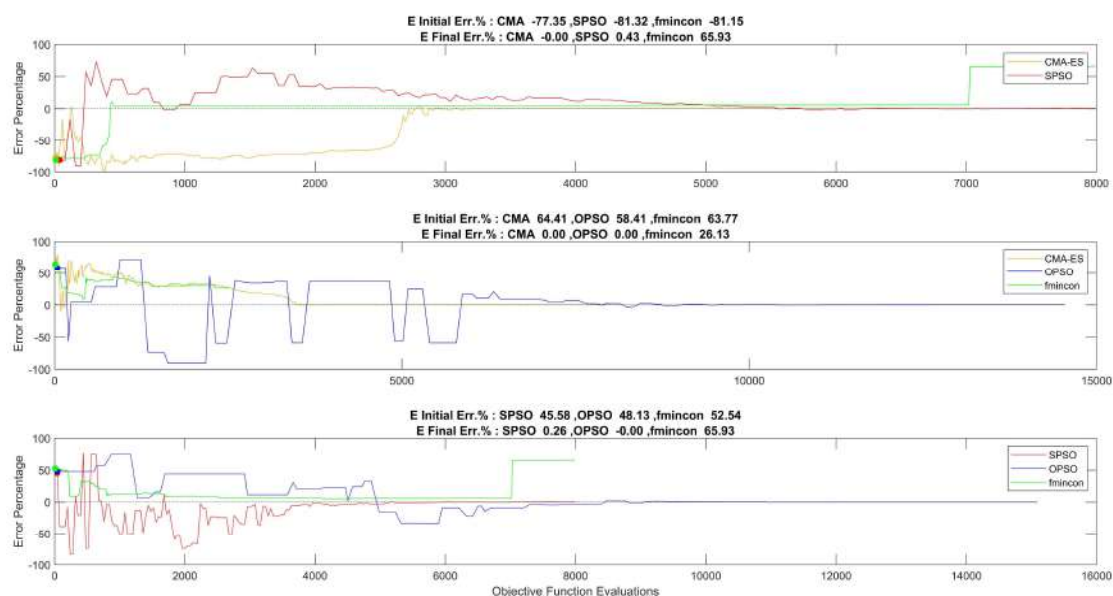
Εικόνα 71. Αναζήτηση Παραμέτρου c3



Εικόνα 72. Αναζήτηση Παραμέτρου k_4



Εικόνα 73. Αναζήτηση Παραμέτρου k_5



Εικόνα 74. Αναζήτηση Παραμέτρου E

7.2.3. Σχολιασμός Αποτελεσμάτων CMA-ES

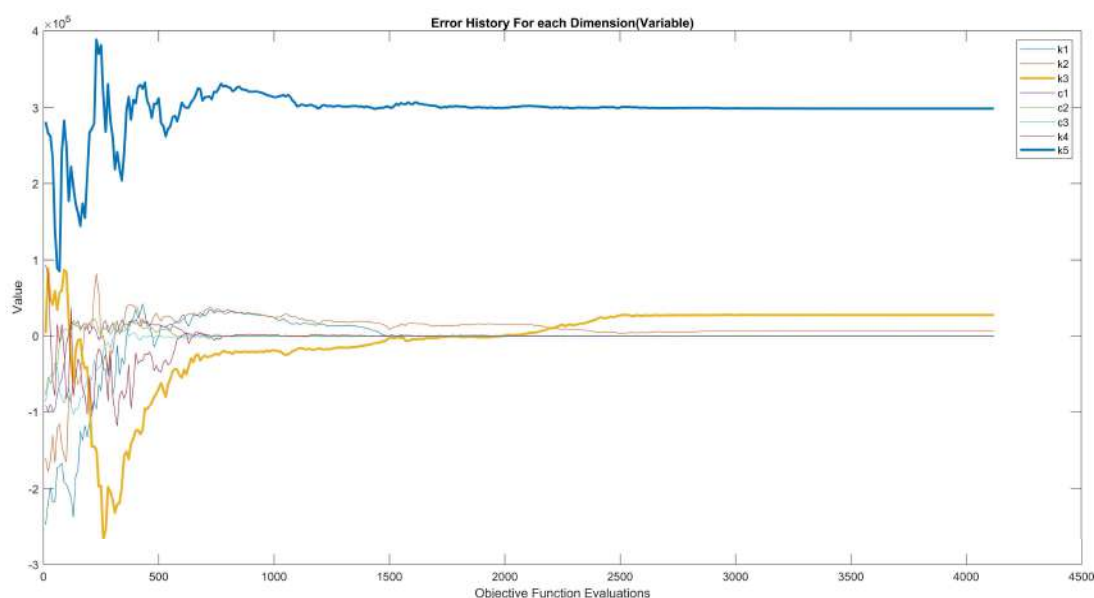
Ενδιαφέρον έχει να εξετάσουμε τις περιπτώσεις στις οποίες απέτυχε ο CMA-ES. Γενικά στους εξελικτικούς αλγόριθμους μια ένδειξη γύρω από την αιτία αποτυχίας είναι ο εγκλωβισμός του πληθυσμού. Στην ουσία ο πληθυσμός υποβιβάζεται σε έναν υποχώρου του αρχικού χώρου αναζήτησης [7]. Στους PSO αλγόριθμους τα σωματίδια μπορεί να εγκλωβιστούν σε τοπικά ελάχιστα περιορίζοντας τον χώρο αναζήτησης σε ένα υποσύνολο. Στην CMA όμως στρατηγική τότε περιορίζεται ο χώρος; Η μετάλλαξη που εφαρμόζουμε στην CMA-ES είναι η

$$\mathbf{y} + \mathcal{N}(\mathbf{0}, \mathbf{C})$$

Στην ουσία καθώς προσαρμόζουμε τον μητρώ συνδιακύμανσης εφαρμόζουμε έναν γραμμικό μετασχηματισμό $\mathbf{C}^{1/2}$ μέσω του μητρώου συνδιακύμανσης. Παρόλου που ο μετασχηματισμός μπορεί να είναι πλήρης τάξης (full rank) παράγοντας τον χώρο κατά την διάρκεια της βελτιστοποίησης, εντούτοις μπορεί να προκύψει το παρακάτω φαινόμενο. Η προσαρμογή του μητρώου συνδιακύμανσης έχει ως στόχο να παράξει μεταλλάξεις προς υποσχόμενες κατευθύνσεις κάνοντας ουσιαστικά μια Ανάλυση Κύριων Συνιστωσών (PCA). Αν ένας από τους αξονες των συνιστωσών (ή αλλιώς ένας ένα από τα διανύσματα της βάσης του μετασχηματισμού) μικρίνει αρκετά, το γεγονός αυτό αυτομάτως σημαίνει πως δεν θα μπορούμε να έχουμε μεγάλες δειγματοληψίες προς εκείνη την κατεύθυνση. Μπορεί ο χώρος να παράγεται αλλά ουσιαστικά για την βελτιστοποίηση έχει πέσει η τάξη και έχουμε υποβιβαστεί σε χαμηλότερη διάσταση. Διαισθητικά αν το πρόβλημα μας ήταν δύο διαστάσεων θα μπορούσαμε να παράξουμε δειγματοληψίες σχεδόν γύρω από μία γραμμή ή αν έχουμε τρεις διαστάσεις το πρόβλημα μας έχει περιοριστεί σε κοντά επίπεδο. Αυτή η συμπεριφορά θα μπορούσε να είναι θεμιτή στον CMA-ES, όμως αν αυτό γίνει όταν ο αλγόριθμος δεν έχει συγκλίνει σε αποδεκτή λύση ή αν η κατεύθυνση που “κατέρρευσε” συμφωνεί με αυτή μιας μεταβλητής του αρχικού συστήματος συντεταγμένος η οποία απέχει από μια “καλή” τιμή, υπάρχει περίπτωση να μην βρεί ο αλγόριθμος αποδεκτή λύση ειδικά αν έχουμε περιορισμένο αριθμό επαναλήψεων. Οι Nikolaus Hansen, Raymond Ros κα. [37] αναφέρουν πως αυτή η τάση “κατάρρευσης” παρουσιάζεται σε όλους τους αλγόριθμους

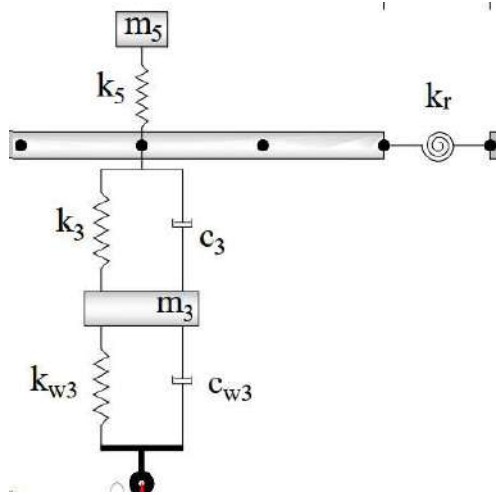
αναζήτησης με γενικές ιδιότητες αναλλοιώτητας. Στους PSO αλγορίθμους που παρουσιάσαμε οι οποίες δεν κατέχουν τέτοιες ιδιότητες, δεν παρατηρείται το φαινόμενο αυτό [37]. Στον CMA-ES για να αποτραπεί αυτή η περίπτωση χρησιμοποιούνται οι ρυθμοί ή βαθμοί μάθησης (learning rates) που παρουσιάσαμε. Μεγάλες τιμές στις σταθερές αυτές οδηγούν σε μια ταχεία προσαρμογή του μητρώου συνδιακύμανση και κατάρρευση του χώρου. Μικρές τιμές αντίθετα αποτρέπουν την κατάρρευση αλλά καθυστερούν την σύγκλιση του αλγορίθμου. Οι δημιουργεί του CMA-ES προτείνουν μια τιμή σαν προεπιλεγμένη για να αποφευχθεί η κατάρρευση καθιστώντας παράλληλα τον αλγόριθμο αποδοτικό.

Παρόλα αυτά το φαινόμενο αυτό μπορεί να συμβεί. Ας εξετάσουμε μια δοκιμή της CMA όπου απέτυχε να αναγνωρίσει τις παραμέτρους k_1 k_2 c_3 και k_5 . Γενικά στις πρώτες επαναλήψεις υπάρχει αρκετός “θόρυβος” καθώς ο αλγόριθμος αρχικά επιτρέπει μια πιο μεγάλη εξερεύνηση. Αυτό που παρατηρούμε εύκολα ότι η παράμετρος k_5 ξεκινά μακριά από την βέλτιστη τιμή της και η k_3 σχετικά κοντά. Καθώς ξεκινούν οι επαναλήψεις η παράμετρος k_5 πολύ γρήγορα απομακρύνεται από την περιοχή καλών τιμών και κινείται σχεδόν αντιδιαμετρικά της k_3 .



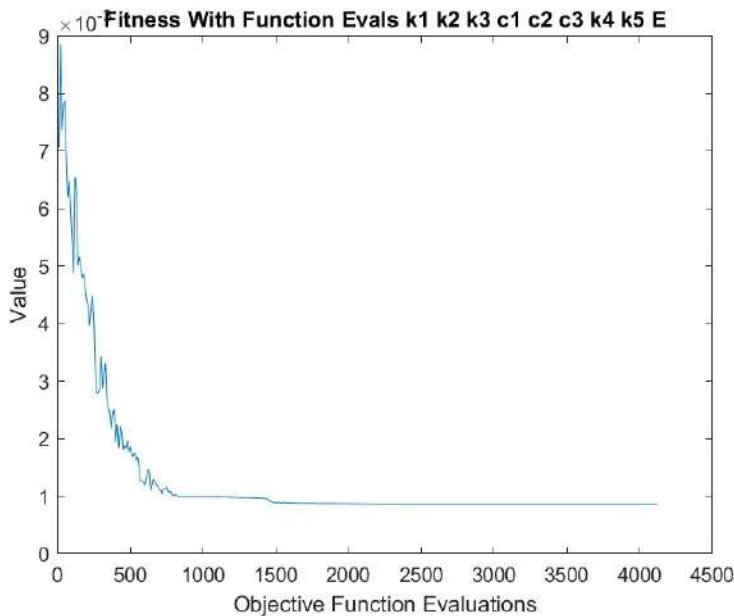
Εικόνα 75. Μια αποτυχημένη δοκιμή για την CMA-ES.

Αν εξετάσουμε το μοντέλο θα δούμε πως το ελατήριο με στιβαρότητα k_3 και αυτό με k_5 είναι αντιδιαμετρικά τοποθετημένα . Οπότε ίσως η κίνηση της μεταβλητής k_5 στον αλγόριθμο γίνεται για να αντισταθμίσει την αλλαγή της σταθεράς του ελατηρίου της k_3 μεταβλητής ώστε οι κόμβοι να πλησιάζουν στην πραγματική μετατόπιση.



Εικόνα 76. Αντιδιαμετρική τοποθέτηση των ελατηρίων k_5 και k_3 .

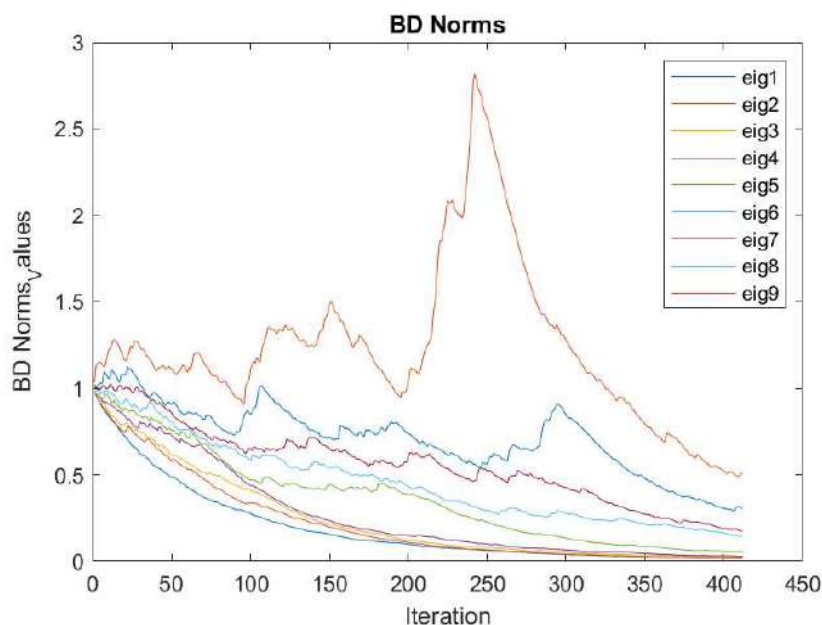
Καθώς συμβαίνουν αυτές οι αλλαγές στις μεταβλητές, το συνολικό σφάλμα μειώνεται και οι υπόλοιπες μεταβλητές συνεχίζουν να κινούνται προς τις σωστές κατευθύνσεις. Έτσι η απόκριση των επιλεγμένων κόμβων που παράγει το μοντέλο συνεχίζει να πλησιάζει την πειραματική απόκριση παρόλο που μερικές μεταβλητές απομακρύνονται από την πειραματική τιμή τους. Μετά από μερικές επαναλήψεις ακόμη και η μεταβλητή k_5 που έκανε αρχικά μια υπερεκτίμηση, σταδιακά δείχνει να βελτιώνεται. Αντίθετα η μεταβλητή k_3 φαίνεται να έχει εγκλωβιστεί σε μια λανθασμένη τιμή. Μια πιθανή εξήγηση είναι ο υποβιβασμός της τάξης που αναφέραμε.



Εικόνα 77. Εξέλιξη τιμής συνάρτησης καταλληλότητας για την αποτυχημένη δοκιμή της CMA-ES.

Έχοντας κάνει ιδιοανάλυση στο μητρώο συνδιακύμανσης υπολογίζουμε τις νόρμες από των ιδιοδιανυσμάτων διαμέσου των επαναλήψεων του αλγορίθμου. Εξετάζοντας το διάγραμμα παρατηρούμε πως σταδιακά η ευκλείδεια νόρμα των ιδιοδιανυσμάτων μειώνεται με πολλές από τις διαστάσεις να πλησιάζουν στο μηδέν και μία κατεύθυνση να μεγαλώνει στην μέση της εκτέλεσης του αλγορίθμου δείχνοντας πως ενισχύεται η αναζήτηση προς μια κατεύθυνση. Με αυτή την πληροφορία μόνο, μπορούμε να αποφανθούμε για το μέγεθος των

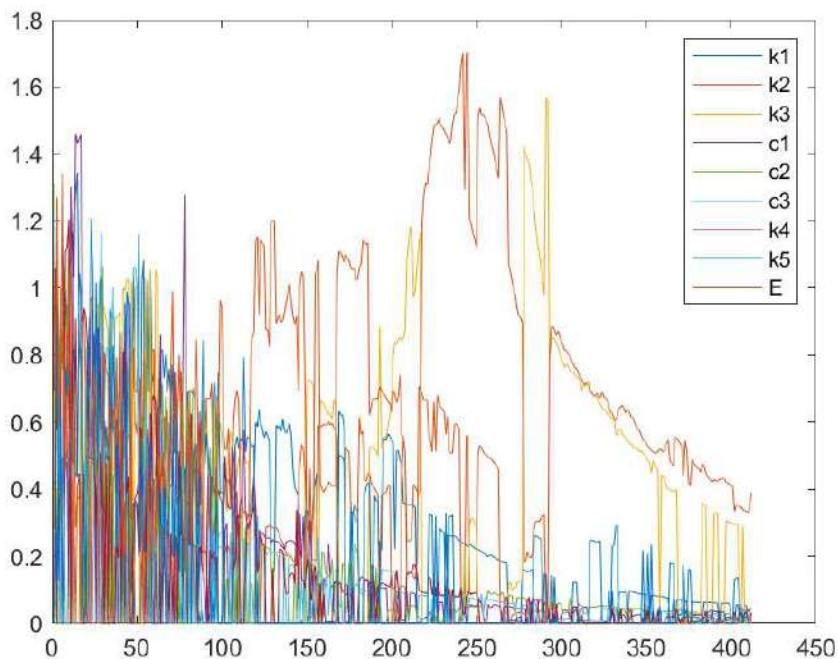
διανυσμάτων των κύριων συνιστωσών άρα και το μέγεθος βήματος στην δειγματοληψία στις διάφορες διαστάσεις. Καθώς όμως περιστρέφεται η κατανομή στον χώρο δεν γνωρίζουμε κατά πόσο συμπίπτουν οι συνιστώσες της βάσης της περιστραμμένης κατανομής με τις αρχικές κατευθύνσεις του ορθοκανονικού συστήματος που αντιπροσωπεύουν τις μεταβλητές που αναζητούμε.



Εικόνα 78. Νόρμες για τα ιδιοδιανύσματα του μητρώου συνδιακύμανσης.

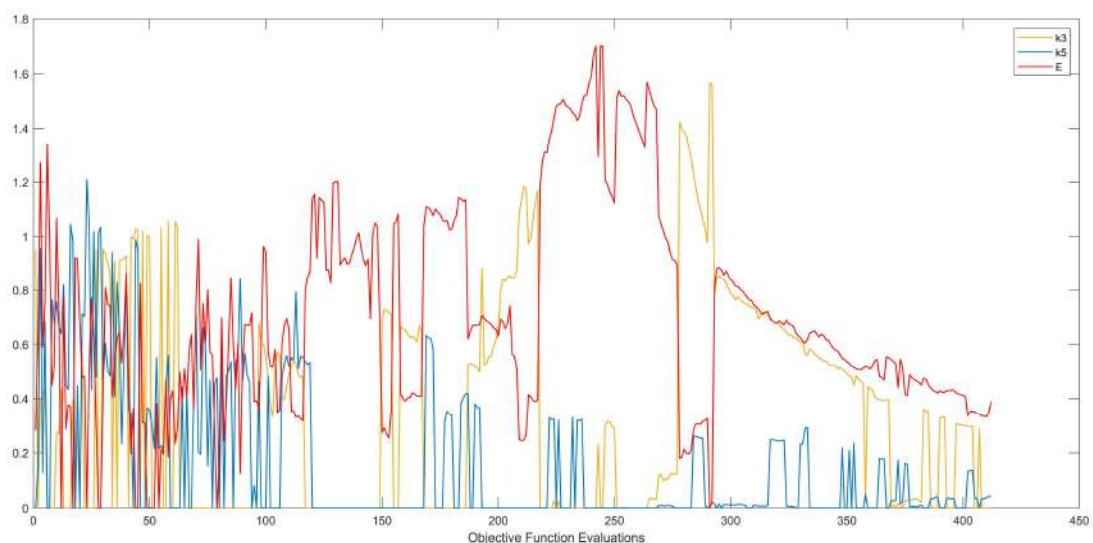
Έτσι δημιουργήθηκε ένα μέτρο για να κρίνουμε προς τα ποια κατεύθυνση η δειγματοληψία παράγει τα πιο ισχυρά βήματα σχετικά με τις μεταβλητές που αναζητούμε. Για κάθε γενιά υπολογίζουμε το εσωτερικό γινόμενο των ιδιοδιανυσμάτων του πίνακα συνδιακύμανσης με το κάθε διάνυσμα από την τετριμμένη βάση πολλαπλασιαζόμενα με τις ιδιοτιμές. Στόχος είναι να συμπεράνουμε προς τα ποια κατεύθυνση, σχετικά με τις μεταβλητές, παράγονται οι μεγαλύτερες δειγματοληψίες. Για παράδειγμα για ένα ιδιοδιάνυσμα πολλαπλασιασμένο με τις ιδιοτιμές, εντοπίζουμε το διάνυσμα της τετριμμένης βάσης που δίνει το μεγαλύτερο εσωτερικό γινόμενο. Έτσι θεωρούμε πως συγκεκριμένο ιδιοδιάνυσμα εισάγει μια μεροληψία προς την κατεύθυνση του τετριμμένου το οποίο συμφωνεί με μια κατεύθυνση που αντιπροσωπεύει μια συγκεκριμένη μεταβλητή και λαμβάνουμε την τιμή του υπόψιν. Αυτή η διαδικασία επαναλαμβάνεται για κάθε ιδιοδιάνυσμα του πίνακα συνδιακύμανσης εκτιμώντας με πιο διάνυσμα της τετριμμένης βάσης ταυτίζεται περισσότερο αθροίζοντας το αποτέλεσμα αν πρόκειται για το ίδιο διάνυσμα της τετριμμένης βάσης. Έτσι καταλήγουμε να έχουμε για κάθε επανάληψη του αλγορίθμου ένα είδους μέτρο για την δύναμη μετάλλαξης που αντιστοιχεί σε κάθε μεταβλητή.

Στο παρακάτω διάγραμμα παρουσιάζονται τα αποτελέσματα αυτής της διαδικασίας που αντικατοπτρίζουν κατά πόσο ο αλγόριθμος αναζητεί προς τις κατευθύνσεις των μεταβλητών πράγμα που σχετίζεται με την κατάρρευση και τον υποβιβασμό της διάστασης που αναφέραμε.



Εικόνα 79. Μέτρο δύναμης μετάλλαξης σχετικά με τις κατευθύνσεις των παραμέτρων προς αναγνώριση.

Κρατώντας μόνο τις k_3 , k_5 και E μεταβλητές βλέπουμε πως στην αρχή η δύναμη μετάλλαξης προς τις k_3 και είναι συγκρίσιμη με αυτή των υπολοίπων, όμως για ένα διάστημα περίπου μετά από 100 επαναλήψεις ο αλγόριθμος σταματά αρκετές επαναλήψεις να αναζητά λύσεις προς τις κατευθύνσεις των k_3 και k_5 . Μετά από μερικές επαναλήψεις η δύναμη μετάλλαξης επανέρχεται για την k_5 , κάτι που ταιριάζει με την επιστροφή της προς ένα φάσμα καλύτερων λύσεων. Κάτι τέτοιο όμως δεν γίνεται για την παράμετρο k_3 , προς την κατεύθυνση της οποίας μειώνεται όλο και περισσότερο η δύναμη μετάλλαξης, πράγμα που σημαίνει ότι η παράμετρος παγιδεύεται στην τιμή που έτυχε να έχει νωρίτερα η οποία είναι μακριά της πειραματικής.



Εικόνα 80. . Μέτρο δύναμης μετάλλαξης σχετικά με τις κατευθύνσεις των παραμέτρων προς αναγνώριση. Μεταβλητές k_3, k_5, E .

Το φαινόμενο αυτό μπορεί να αντιμετωπιστεί είτε αυξάνοντας τον πληθυσμό της CMA-ES είτε μειώνοντας τον ρυθμό μάθησης. Το πρώτο μπορεί να εφαρμοστεί χωρίς κάποια αμφιβολία για τον αντίκτυπο που θα υπάρξει στην απόδοση του αλγορίθμου ενώ ο δεύτερος τρόπος αντιμετώπισης χρίζει μια πιο προσεκτική προσέγγιση στην αλλαγή αυτή.

Τέλος παρατηρήσαμε πως στις πιο πολλές περιπτώσεις που οι αλγόριθμοι πετυχαίνουν, το μέτρο ελαστικότητας συγκλίνει σε αποδεκτή λύση σχετικά νωρίς. Αντίθετα στις περιπτώσεις που ο αλγόριθμος αποτυγχάνει, αν το μέτρο ελαστικότητας καταφέρει να πλησιάσει προς την περιοχή αποδεκτής λύσης αυτό το κάνει πολύ πιο μετά. Είναι λογικό πως το μέτρο ελαστικότητας θα πρέπει να κατέχει κάποιον σχετικά σημαντικό ρόλο στην απόκριση του μοντέλου και ίσως η δυσκολία στην σύγκλιση αυτής της μεταβλητής δυσχεραίνει συνολικά την απόδοση.

7.2.4 Δοκιμή μοντέλου half-car με διαφορετική τοποθέτηση αισθητήρων

Με βάση την παρατήρηση αλλάζουμε του κόμβους από τους οποίους παίρνουμε τις μετρήσεις ώστε να γίνει το μοντέλο πιο ευαίσθητο στην αλλαγή του μέτρου ελαστικότητας ώστε να εντοπίζεται πιο γρήγορα. Συγκεκριμένα με βάση μια εκτίμηση “καλών” παραμέτρων αναζητούμε τους κόμβους στο σασί του οχήματος που έχουν μεγαλύτερο μέγεθος επιτάχυνσης. Επιλέγουμε δύο κόμβους και αντικαθιστούμε με τους προηγούμενους δύο που ανήκουν στο σασί του οχήματος. Η αρχικοποίηση που χρησιμοποιήθηκε είναι η ίδια που εφαρμόστηκε στην πρώτη περίπτωση με το ίδιο αρχικό σημείο.

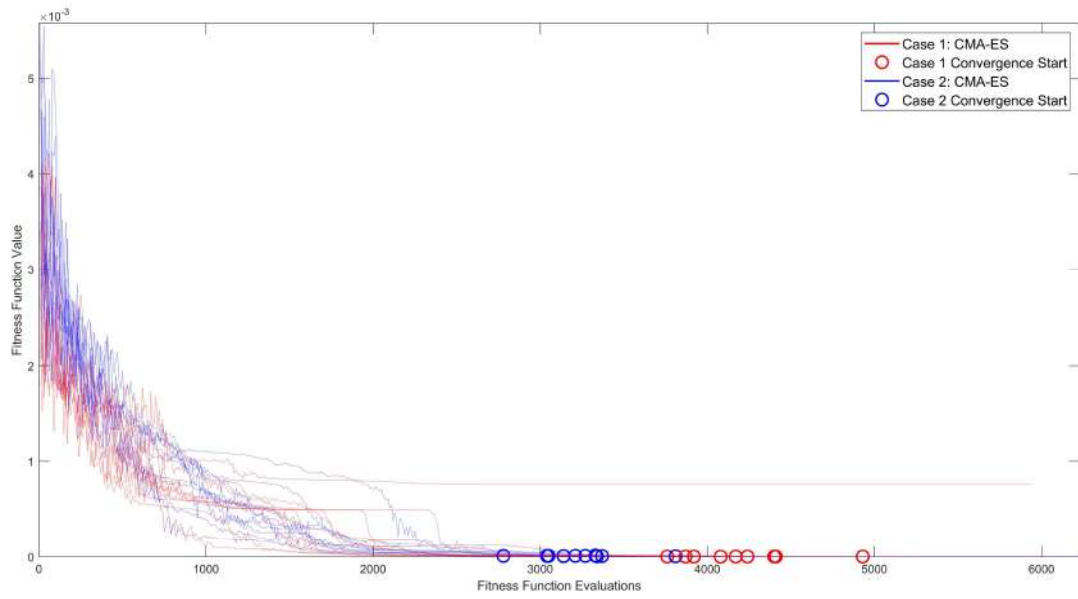
Παρακάτω παρουσιάζονται τα αποτελέσματα.

Δοκιμή	CMA-ES	SPSO	OPSO	fmincon
1	2780	-	8216	-
2	3140	4280	9960	-
3	3270	5480	10632	-
4	3810	-	9104	-
5	3330	-	9160	-
6	3140	6240	8872	-
7	3210	4760	9264	-
8	3050	7600	8080	-
9	3040	7520	7488	-
10	3340	6840	8600	-
11	3370	-	8288	-
12	3580	-	8808	-
13	3210	4640	9336	-
14	2900	5440	9008	-
15	3150	4760	8800	-
16	3410	6440	9040	-
17	2910	6600	8848	-

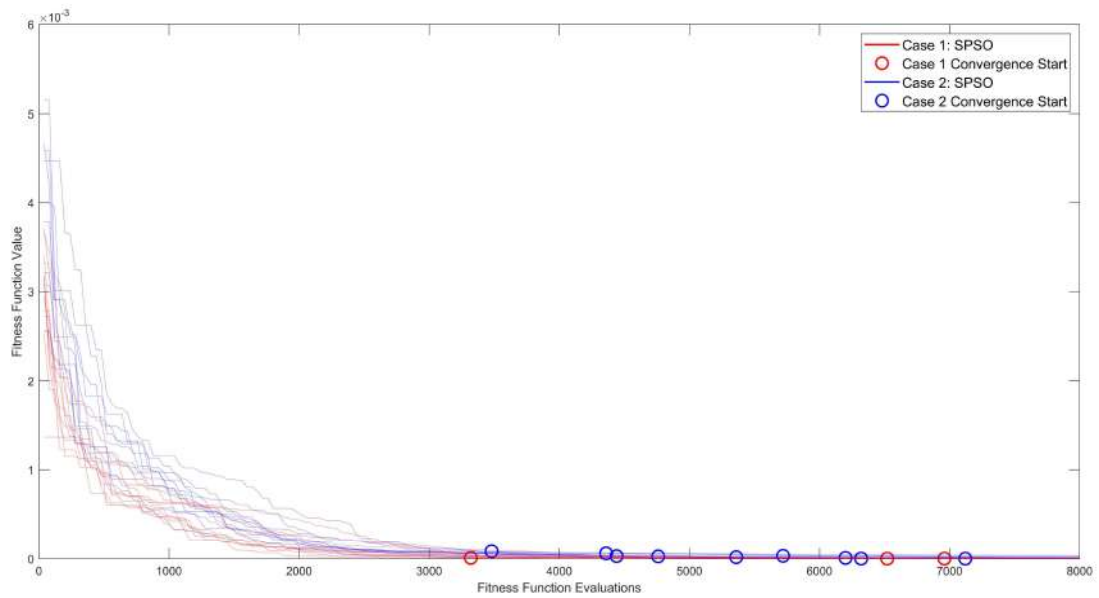
18	3270	6600	8848	-
19	3110	-	9240	-
20	3130	4560	8336	-
21	3340	-	10056	-
22	2410	5920	7848	-
23	3060	-	8208	-
24	3050	4040	8832	-
25	5250	4920	9320	-
26	3450	-	9056	-
27	3040	2880	10200	-
28	3410	6360	8352	-
29	3020	-	9704	-
30	3390	-	8776	-
<i>Μέσος όρος Υπολογισμών</i>	2780	5572	8942	-

Με αυτή την εκλογή κόμβων για τοποθέτηση αισθητήρων, αυτή την φορά ο αλγόριθμος CMA-ES σε 30 δοκιμές δεν αποτυγχάνει καμία φορά και μάλιστα βελτιώνεται η απόδοση του με τον αλγόριθμο να συγκλίνει κατά μέσο όρο μετά από 2780 επαναλήψεις σε αντίθεση με την πρώτη περίπτωση που συγκλίνει κατά μέσο όρο μετά από 4244 επαναλήψεις. Ο SPSO συγκλίνει 19 φορές σωστά, σε αντίθεση με τις 12 φορές της πρώτης περίπτωσης. Η απόδοση της OPSSO μεθόδου δεν μεταβάλλεται αρκετά.

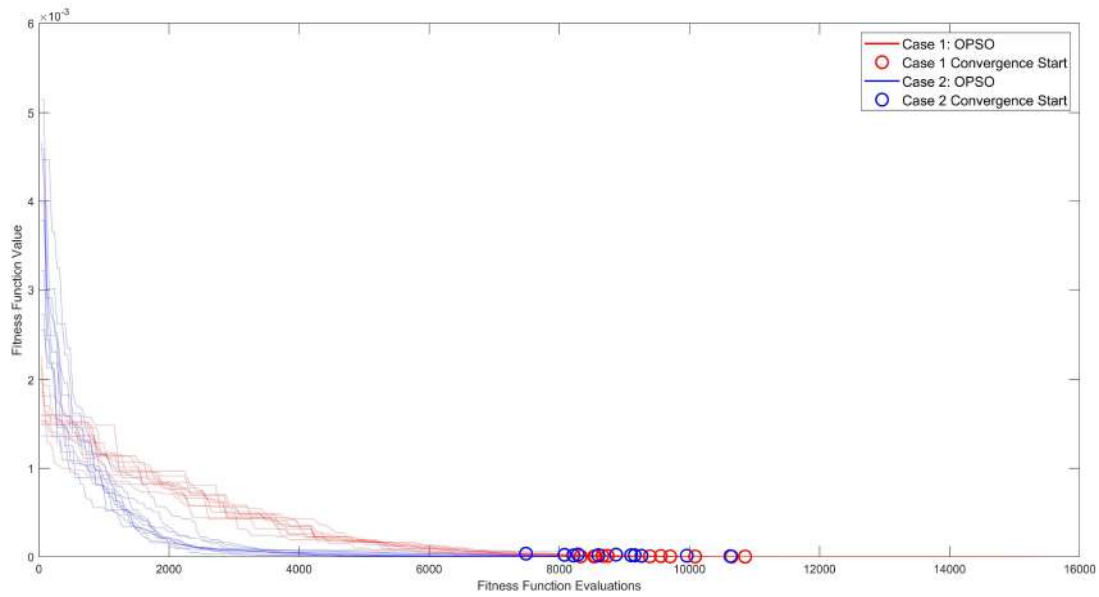
Παρακάτω συγκρίνεται για κάθε μέθοδο η απόδοση της πρώτης περίπτωσης (CASE 1) με την απόδοση της δεύτερης περίπτωσης (CASE 2) στην οποία επιλέξαμε να αντικαταστήσουμε τους αισθητήρες δύο κόμβων με δύο άλλους στους οποίους η απόκριση της επιτάχυνσης έχει μεγαλύτερο μέγεθος .



Εικόνα 81. Απόδοση της CMA-ES για CASE 1 και CASE 2

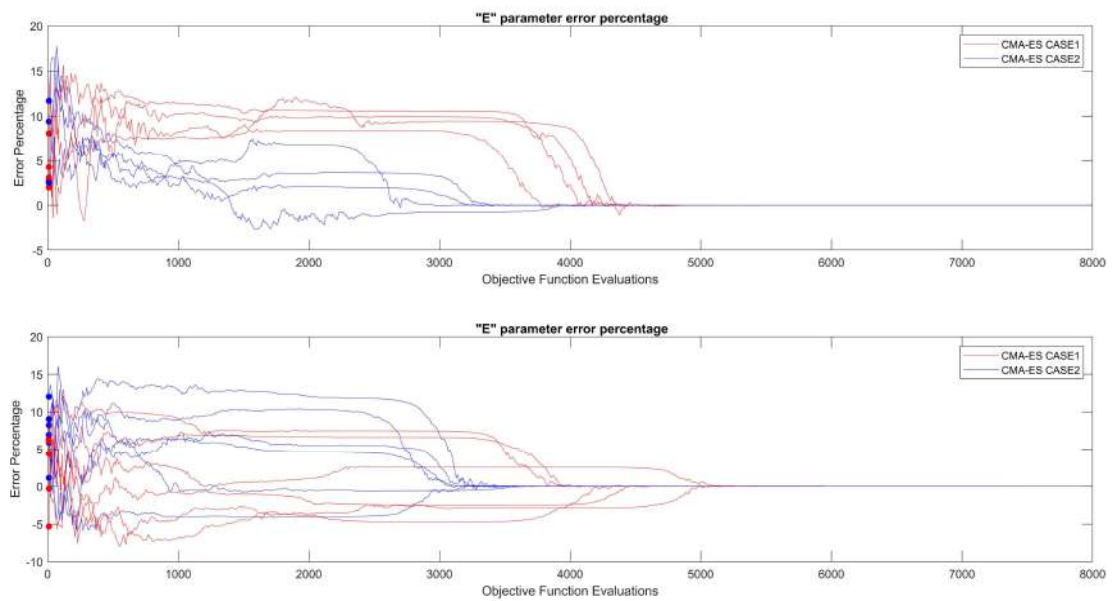


Εικόνα 82. Απόδοση της SPSO για CASE 1 και CASE 2

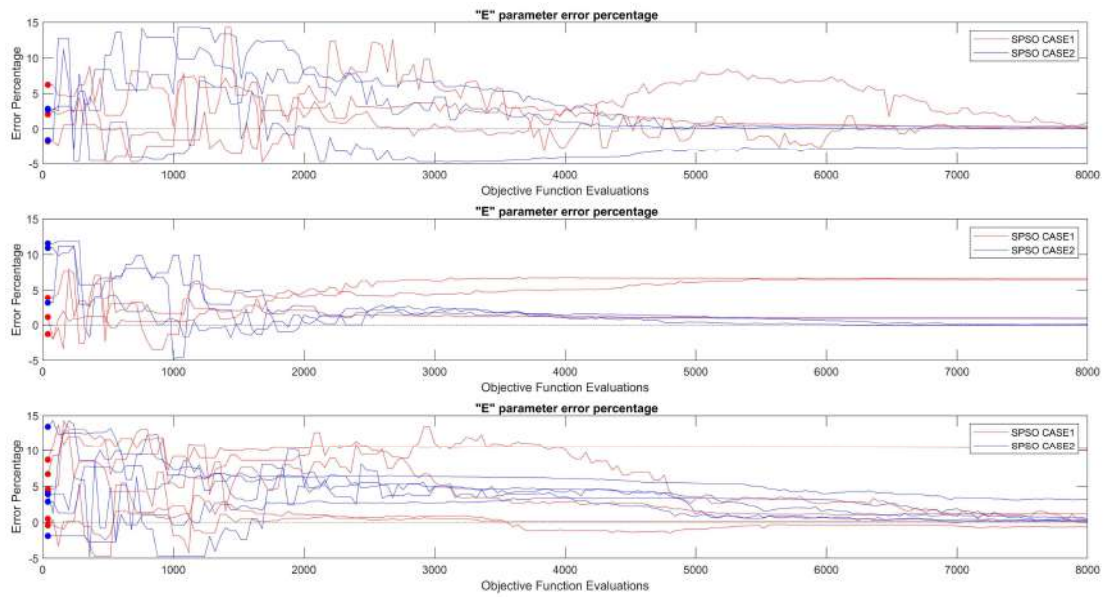


Εικόνα 83. Απόδοση της OP SO για CASE 1 και CASE 2.

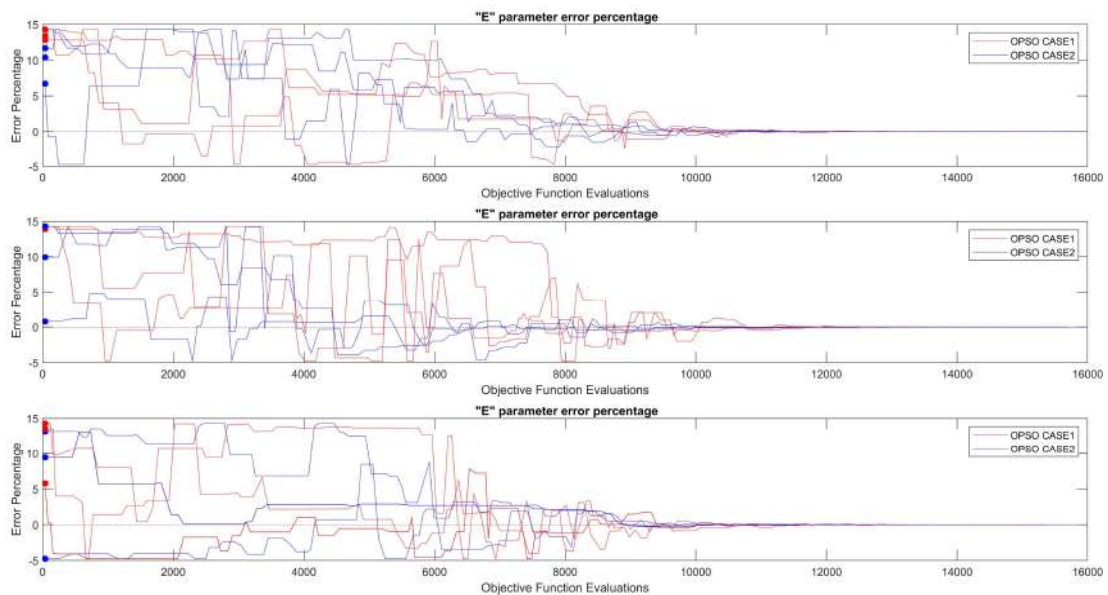
Για να ερευνήσουμε την εικασία όπου η αναγνώριση της παραμέτρου “E” πιθανώς επηρεάζει τα παραπάνω αποτελέσματα, συγκρίνουμε την αναζήτηση της μεταβλητής “E” για 9 δοκιμές για την περίπτωση CASE 1 και CASE 2 για κάθε μέθοδο.



Εικόνα 84. Αναζήτηση της παραμέτρου “E” μέσω της CMA-ES για CASE1 και CASE2



Εικόνα 85. Αναζήτηση της παραμέτρου "E" μέσω της SPSO για CASE1 και CASE2



Εικόνα 86. Αναζήτηση της παραμέτρου "E" μέσω της OPSO για CASE1 και CASE2

7.3. Ενημέρωση μοντέλου με μη γραμμικότητες.

Γενικά σε όλες τις πρακτικές εφαρμογές μηχανικών κατασκευών παρουσιάζονται φαινόμενα μη γραμμικότητας τα οποία μπορεί να προκαλούνται από διάφορες αιτίες. Γενικά μπορούν να χωριστούν σε τρεις μεγάλες κατηγορίες, σε γεωμετρικές, υλικού και μη γραμμικότητες επαφής. Στην κατηγορία των γεωμετρικών η μη γραμμικότητα προκαλείται λόγω μεγάλων παραμορφώσεων που επηρεάζουν σημαντικά την γεωμετρία της υπό μελέτη κατασκευής. Σε αυτή του υλικού προκαλείται λόγω της σχέσης τάσης-παραμόρφωσης η οποία μπορεί να αποχωριστεί από τη γραμμική συμπεριφορά. Το τελευταίο εμφανίζεται υλικά που έχουν εν γένει μη-γραμμική συμπεριφορά όπως σε πλαστικά πολυμερή υλικά που εμφανίζουν

ιξωδοελαστική ελαστική συμπεριφορά αλλά ακόμα και σε μέταλλα όταν η τάση ξεπερνά το όριο διαρροής. Μη-γραμμικότητα μπορεί να παρουσιαστεί σε στηρίξεις κατασκευών όπου περιστροφές και μετατοπίσεις υποστηρίζονται διαφορετικά από την στήριξη σε διαφορετικές κατευθύνσεις ή στις περιπτώσεις όπου υπάρχει κάποιο κενό στην στήριξη λόγω του οποίου δεν υπάρχει αντίδραση στην στήριξη μέχρι να κλείσει το κενό αυτό. Επίσης μη γραμμικότητα εμφανίζεται σε περιπτώσεις όπου υπάρχει τριβή οι οποίες εμπίπτουν στην κατηγορία επαφής.

Η μη γραμμικότητα είναι σχετικά δύσκολη να αναλυθεί επειδή παραβιάζει μερικές αρχές που αποτελούν την βάση της ανάλυσης που γίνεται στις γραμμικές περιπτώσεις. Η πιο βασική αρχή που παραβιάζεται είναι η αρχή της υπέρθεσης.

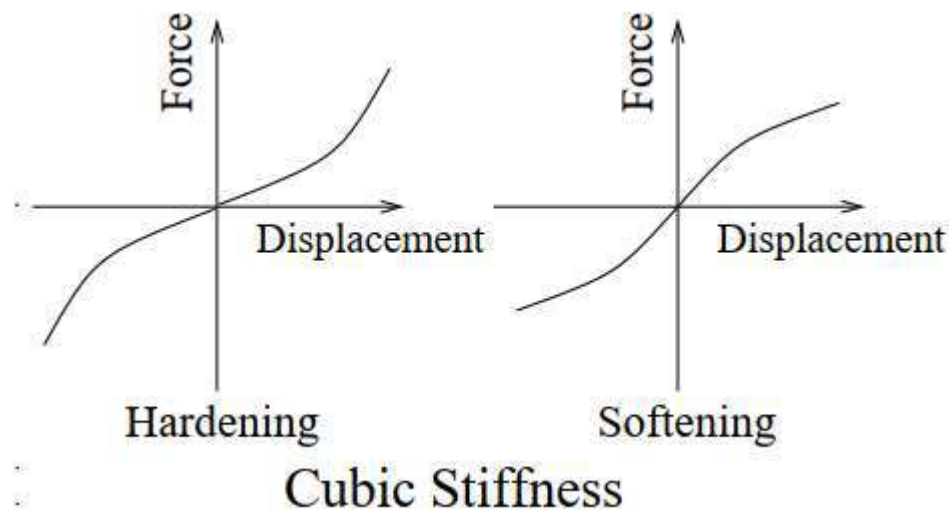
Η αρχή της υπέρθεσης δηλώνει πως όταν η απόκριση σε ένα σύστημα λόγος μιας δύναμης $x_1(t)$ είναι η $y_1(t)$ και μία άλλη δύναμη $x_2(t)$ προκαλεί την απόκριση $y_2(t)$ τότε η απόκριση στην δύναμη $\alpha x_1(t) + \beta x_2(t)$ είναι η $\alpha y_1(t) + \beta y_2(t)$. Μια απόρροια της συνθήκης αυτής βρίσκεται και στο πεδίο των συχνοτήτων όπου γραμμικά συστήματα που υποβάλλονται σε αρμονική διέγερση με σταθερή συχνότητα έχουν απόκριση στην ίδια συχνότητα. Μία πιο αδύναμη περίπτωση της αρχής της υπέρθεσης είναι αυτή της ομοιογένειας. Βάση αυτής αν η διέγερση $x_1(t)$ προκαλεί $y_1(t)$ απόκριση τότε $\alpha x_1(t) \rightarrow \alpha y_1(t)$. Έτσι αν $\alpha x_1(t) \rightarrow \alpha y_1(t)$, τότε στο πεδίο των συχνοτήτων $\alpha X(\omega) \rightarrow \alpha Y(\omega)$ και $H(\omega) = \frac{Y(\omega)}{X(\omega)} = \frac{\alpha Y(\omega)}{\alpha X(\omega)} = H(\omega)$. Οπότε η συνάρτηση μεταφοράς είναι ανεξάρτητη του μεγέθους της διέγερσης. Στα μη-γραμμικά συστήματα τα παραπάνω δεν ισχύουν με την συνάρτηση μεταφοράς, άρα και την απόκριση συχνότητας, να εξαρτάται από το μέγεθος της διέγερσης και την απόκριση να παρουσιάζει ανώτερες αρμονικές της βασικής συχνότητας [38]. Λόγο αυτών των ιδιοτήτων παρουσιάζεται πρόβλημα στην αξιοποίηση της απόκρισης συχνότητας στην αναγνώριση παραμέτρων των μη-γραμμικών συστημάτων.

Μια διαφορική εξίσωση δευτέρας τάξης που περιγράφει μη-γραμμικούς ταλαντωτές με απόσβεση είναι η εξίσωση Duffing.

$$m\ddot{y} + c\dot{y} + k_1y + k_3y^3 = x(t)$$

Αυτή η εξίσωση περιέχει κυβικό όρο στην στιβαρότητα και αντιπροσωπεύει μια προσέγγιση χαμηλής τάξης αναπτύγματος Taylor για μοντέλα με γενικότερη μη-γραμμικότητα στην στιβαρότητα με περιττή συνάρτηση δύναμης επαναφοράς, $m\ddot{y} + c\dot{y} + k_1y + f_s(y) = x(t)$, όπου $f_s(y)$ η δύναμη επαναφοράς η οποία είναι περιττή συνάρτηση με $f_s(y) = -f_s(-y)$, πράγμα που σημαίνει πως η δύναμη επαναφοράς έχει κατεύθυνση πάντα προς το κέντρο με μέγεθος ανεξάρτητο του πρόσημου της μετατόπισης [38].

Εξετάζοντας την δύναμη επαναφοράς $f_s(y) = k_1y + k_3y^3$ ξεχωρίζουμε μερικές περιπτώσεις. Συγκεκριμένα η σταθερά στιβαρότητας k_3 μπορεί να λάβει θετικές ή αρνητικές τιμές. Αν το μέτρο k_3 είναι θετικό τότε έχουμε χαρακτηριστικά σκλήρυνσης ενώ αν η σταθερά k_3 έχει αρνητικές τιμές τότε η στιβαρότητα έχει χαρακτηριστικά χαλάρωσης. Συστήματα με $k_3 < 0$ είναι αφύσικα για μεγάλες τιμές μετατόπισης καθώς μετά από κάποια τιμή μετατόπισης η δύναμη επαναφοράς αλλάζει πρόσημο και οδηγεί την απόκριση του συστήματος στο άπειρο [38]. Τέτοια συστήματα στην πραγματικότητα έχουν πάντα πολυωνυμικούς όρους μεγαλύτερης τάξης με θετικού συντελεστές που κυριαρχούν στις μεγάλες μετατοπίσεις επαναφέροντας της ισορροπία [38].



Εικόνα 87 [38]. Συμπεριφορά δύναμης επαναφοράς για στιβαρότητα με κυβικούς όρους σκλήρυνσης και χαλάρωσης

Εξετάζοντας την απόκριση αρμονικής ισορροπίας ενός ταλαντωτή duffing ακολουθώντας την ανάλυση των K. Worden και G.R. Tomlinson [38], θεωρούμε πως η απόκριση σε μια ημιτονοειδής διέγερση είναι και αυτή ημιτονοειδής ίδια συχνότητας. Εισάγοντας για λύση μια δοκιμή $y = Y\sin(\omega t)$ και μεταφέροντας τον όρο της φάσης στην διέγερση ώστε να απλοποιηθούν οι πράξεις καθιστώντας την είσοδο Y ως πραγματικό αριθμό έχουμε :

$$\begin{aligned}
 m\ddot{y} + c\dot{y} + k_1y + k_3y^3 &= X \sin(\omega t - \varphi) \\
 -m\omega^2 Y \sin(\omega t) + c\omega Y \cos(\omega t) + k_1 Y \sin(\omega t) + k_3 Y^3 \sin^3(\omega t) &= X \sin(\omega t - \varphi) \\
 -m\omega^2 Y \sin(\omega t) + c\omega Y \cos(\omega t) + k_1 Y \sin(\omega t) + k_3 Y^3 \left\{ \frac{3}{4} \sin(\omega t) - \frac{1}{4} \sin(3\omega t) \right\} &= X \sin(\omega t - \varphi) \\
 &= X \sin(\omega t) \cos \varphi - X \cos(\omega t) \sin \varphi
 \end{aligned}$$

Εξισώνοντας του συντελεστές των $\sin(\omega t)$ και $\cos(\omega t)$ έχουμε,

$$-m\omega^2 Y + k_1 Y + \frac{3}{4} k_3 Y^3 = X \cos \varphi \quad (25)$$

$$c\omega Y = -X \sin \varphi \quad (26)$$

Τετραγωνίζοντας και αθροίζοντας τις σχέσεις (25) και (26) έχουμε :

$$X^2 = Y^2 \left[\left\{ -m\omega^2 + k_1 + \frac{3}{4} k_3 Y^2 \right\}^2 + c^2 \omega^2 \right] \quad (27)$$

Οπότε συνδιάζοντας τη σχέση (27) έχουμε την εξίσωση απολαβής (gain) :

$$\left| \frac{Y}{X} \right| = \frac{1}{\left[\left\{ -m\omega^2 + k_1 + \frac{3}{4} k_3 Y^2 \right\}^2 + c^2 \omega^2 \right]^{1/2}} \quad (28)$$

και την φάση

$$\varphi = \tan^{-1} = \frac{-c\omega}{-m\omega^2 + k_1Y + \frac{3}{4}k_3Y^2} \quad (29)$$

καταλήγουμε στην συνάρτηση απόκρισης συχνότητας :

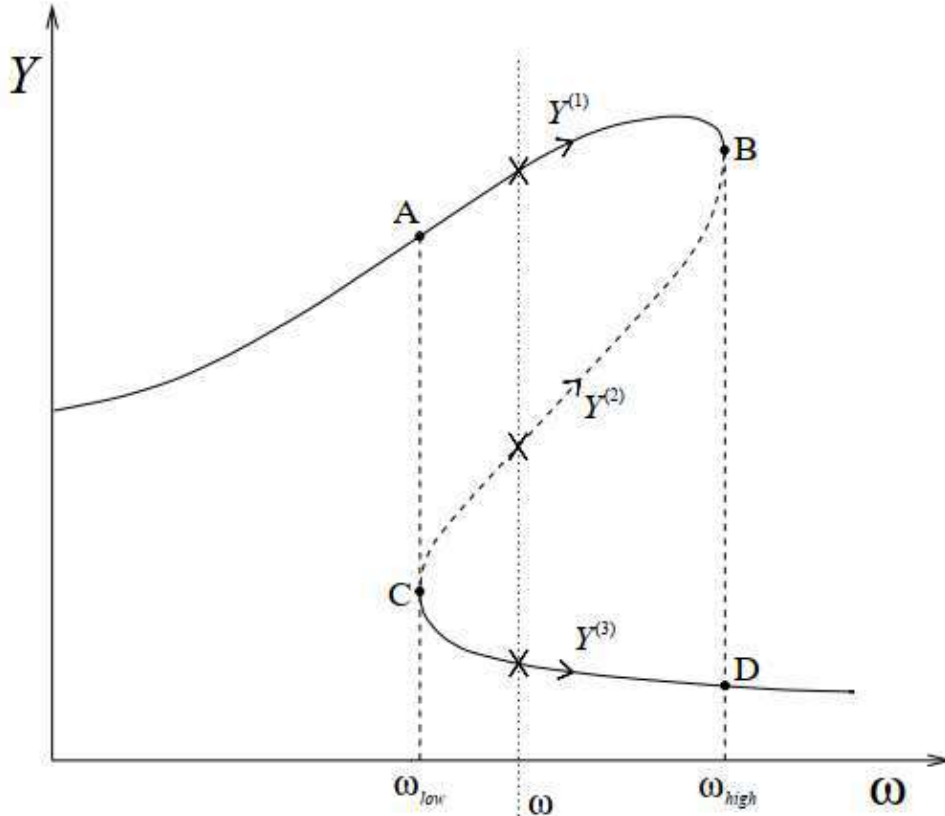
$$A(\omega) = \frac{1}{k + \frac{3}{4}k_3Y^2 - m\omega^2 + ic\omega} \quad (30)$$

Αυτή μπορεί να θεωρηθεί ως η συνάρτηση απόκρισης συχνότητας του γραμμικοποιημένου συστήματος $m\ddot{y} + c\dot{y} + k_{eq}y + k_3y^3 = X\sin(\omega t - \varphi)$ όπου η ισοδύναμη στιβαρότητα εξαρτάται από το μέγεθος της διέγερσης και ισούται με $k_{eq} = k + \frac{3}{4}k_3Y^2$. Έτσι σε σταθερή διέγερση η ιδιοσυχνότητα της συνάρτησης απόκρισης συχνότητας είναι

$$\omega_n = \sqrt{\frac{k + \frac{3}{4}k_3Y^2}{m}}$$

και εξαρτάται από το Y άρα και έμμεσα από το πλάτος της διέγερσης X . Αν $k_3 > 0$ τότε η ιδιοσυχνότητα αυξάνεται με το πλάτος X ενώ αν $k_3 < 0$ τότε αντίθετα μειώνεται αυξανόμενου του πλάτους X .

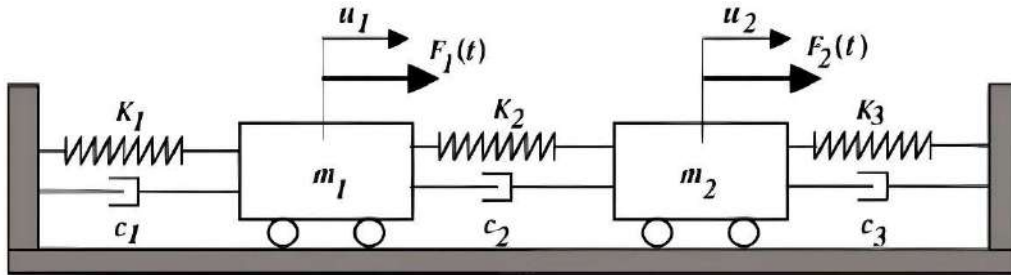
Για δεδομένο X και ω η απόκριση Y λύνεται από την εξίσωση (27). Για μικρή διέγερση η συνάρτηση απόκρισης συχνότητας είναι σχεδόν όμοια με αυτή του γραμμικού συστήματος και λαμβάνουμε μοναδική λύση Y για κάθε διέγερση ω . Καθώς όμως το πλάτος X αυξάνεται απομακρυνόμαστε από την γραμμική συμπεριφορά και σε μια κρίσιμη τιμή X_{crit} μετά την οποία υπάρχουν τρεις λύσεις για την απόκριση σε ένα φάσμα $[\omega_{low}, \omega_{high}]$ όπως παρουσιάζεται στο παρακάτω σχήμα.



Εικόνα 88 [38]. Μεταβολή της συνάρτησης απόκρισης συχνότητας και σημείο διαχωρισμού.

Αυτό είναι ένα κλασικό παράδειγμα σημείου διαχωρισμού (bifurcation point) για την παράμετρο X . Ενώ η παράμετρος X είναι συνεχής, καθώς μεταβάλλεται, ο αριθμός των λύσεων και τα σημεία ισορροπίας μεταβάλλονται ξαφνικά. Αν ακολουθήσουμε την καμπύλη, καθώς αυξάνεται η συχνότητα μόλις φτάσουμε στη συχνότητα ω_{low} δύο νέες αποκρίσεις γίνονται πιθανές οι οποίες παραμένουν μέχρι να φτάσουμε την συχνότητα ω_{high} . Ονομάζουμε τον κλάδο από το A έως το B ως $Y^{(1)}$, τον κλάδο B έως το C ως $Y^{(2)}$ και τον κλάδο από το C έως το D ως $Y^{(3)}$. Μόλις φτάσου στο σημείο B του πρώτου κλάδου, αν αυξήσουμε περαιτέρω την συχνότητα διέγερσης τότε η απόκριση μεταπίπτει απευθείας στο σημείο D του κλάδου $Y^{(3)}$ παρουσιάζοντας μια ασυνέχεια η οποία ονομάζεται άλμα. Παρόμοιο φαινόμενο συμβαίνει ξεκινώντας από τον κλάδο $Y^{(3)}$ μόλις φτάσουμε στο σημείο C μειώνοντας την συχνότητα. Περαιτέρω μείωση τώρα μας οδηγεί στο σημείο A. Η λύση του κλάδου $Y^{(2)}$ είναι ασταθής και δεν παρατηρείται ποτέ στην πραγματικότητα. Σε περίπτωση που η σταθερά k_3 είναι θετική τότε η συχνότητα συντονισμού αυξάνεται και τα άλματα που περιγράψαμε συμβαίνουν στη δεξιά πλευρά της κορυφής όπως περιγράψαμε. Αντίθετα με $k_3 < 0$ τα άλματα συμβαίνουν στην αριστερή πλευρά της κορυφής και η συχνότητα συντονισμού μεταβαίνει σε χαμηλότερες τιμές. Παρόμοια φαινόμενα ασυνέχειας παρατηρούνται και στην φάση φ .

Στην περίπτωση μας έγινε προσπάθεια αναγνώρισης παραμέτρων ενός ταλαντωτή Duffing δύο βαθμών ελευθερίας. Το μοντέλο παρουσιάζεται στο παρακάτω σχήμα



Εικόνα 89 Ταλαντωτής Duffing δύο βαθμών ελευθερίας

Κυβική στιβαρότητα εφαρμόστηκε τα ελατήρια με σταθερές,

$$K_i(u) = k_i u + a_i u^3, i = 1,2,3$$

Οι εξισώσεις κίνησης που προκύπτουν είναι :

$$m_1 \ddot{u}_1 + (c_1 + c_2) \dot{u}_1 - c_2 \dot{u}_2 + (k_1 + k_2) u_1 - k_2 u_2 + a_1 u_1^3 - a_2 (u_2 - u_1)^3 = \delta_1 \sin(\Omega_1 t)$$

$$m_2 \ddot{u}_2 + (c_3 + c_2) \dot{u}_2 - c_2 \dot{u}_1 + (k_3 + k_2) u_2 - k_2 u_1 + a_3 u_2^3 + a_2 (u_2 - u_1)^3 = \delta_2 \sin(\Omega_2 t)$$

Για τα ελατήρια επιλέχθηκε στιβαρότητα με θετικούς συντελεστές a_i , ώστε τα ελατήρια να έχουν σκληρυνόμενη συμπεριφορά. Συνολικά οι συντελεστές και οι διεγέρσεις επιλέχθηκαν ώστε να μην παρουσιάζει χαοτική συμπεριφορά το σύστημα. Οι πραγματικές τιμές των παραμέτρων προς αναγνώριση είναι :

$k_1 = 5.4 \frac{N}{m}, k_2 = 6.9 \frac{N}{m}, k_3 = 7.2 \frac{N}{m}, c_1 = 0.01 \frac{Ns}{m}, c_2 = 0.04 \frac{Ns}{m}, c_3 = 0.1 \frac{Ns}{m}, a_1 = 1 \frac{N}{m^3}, a_2 = 2 \frac{N}{m^3}, a_3 = 1 \frac{N}{m^3}$. Τα πλάτη των ημιτονοειδών διεγέρσεων στην δοκιμή F_1 και F_2 τέθηκαν ως $f_1 = 7.5$ και $f_2 = 4$ αντίστοιχα. Για την επίλυση του συστήματος επιλέχθηκε η ode45 στο MATLAB η οποία βασίζεται στην μέθοδο Runge-Kutta 5^{ης} τάξης.

Για την ενημέρωση του μοντέλου και την αναγνώριση των παραμέτρων αρχικά επιλέχθηκε σύγκριση χρονοιστοριών απόκρισης για διεγερση σε δύο περιπτώσεις με δύο συχνότητες ίδιες και για τις δύο μάζες. Στην πρώτη περίπτωση σε 0.1592Hz αρχικά και για τις δύο μάζες και στην δεύτερη σε 0.7962Hz ομοίως. Στην συνέχεια τέθηκε ένα όριο σφάλματος στο 10% για κρίνουμε την επιτυχής αναγνώριση της παραμέτρου. Το αρχικό σημείο εκτίμησης Nominal Point θεωρήθηκε το σημείο με τιμές παραμέτρων $k_1 = 10 \frac{N}{m}, k_2 = 10 \frac{N}{m}, k_3 = 10 \frac{N}{m}, c_1 = 0.05 \frac{Ns}{m}, c_2 = 0.05 \frac{Ns}{m}, c_3 = 0.05 \frac{Ns}{m}, a_1 = 5 \frac{N}{m^3}, a_2 = 5 \frac{N}{m^3}, a_3 = 5 \frac{N}{m^3}$. Η διαφορά από τις πραγματικές τιμές εκφρασμένη σε ποσοστό επί τοις εκατό για την κάθε μεταβλητή κατά σειρά είναι, 85.2%, 45%, 39%, 400%, 25%, 50%, 20%, 50% και 20%.

Το σημείο τέθηκε αυτό ως interior point στην fmincon συνάρτηση του MATLAB και ως αρχικός γονέας στην CMA-ES. Στους PSO αλγόριθμους συνήθως συνηθίζεται η αρχικοποίηση του σμήνους μέσω ομοιόμορφης κατανομής ανάμεσα στα άνω και κάτω όρια του χώρου αναζήτησης. Αρχικά επιλέχθηκε να "προσομοιωθεί" κατά κάποιο τρόπο το πρώτο βήμα της CMA-ES στους SPSO και OPSO, αρχικοποιώντας τον πληθυσμό του σμήνους γύρω από το

Nominal σημείο μέσω κανονικής κατανομής με μέσο το σημείο αυτό και τυπική απόκλιση την ίδια τυπική απόκλιση που εισάγουμε στην CMA-ES.

Βάση του ορίου του 10% για τις παραμέτρους η CMA-ES κατάφερε να αναγνωρίσει σωστά όλες τις παραμέτρους 13 φορές στις 20. Συνολικά αναγνώρισε τις k_1 16 φορές, k_2 16 φορές, k_3 16 φορές, c_1 16 φορές, c_2 16 φορές, c_3 16 φορές, a_1 16 φορές, a_2 13 φορές και την a_3 17 φορές.

Δοκιμές Αριθμός	Ποσοστά Σφάλματος Κάθε Μεταβλητής για CMA-ES								
	k_1	k_2	k_3	c_1	c_2	c_3	α_1	α_2	α_3
1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	22	0
6	0	0	0	0	0	0	0	0	0
7	53	47	67	122	21	12	30	150	39
8	53	47	67	122	21	11	33	150	22
9	0	0	0	0	0	0	0	10	0
10	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0
13	53	47	67	123	21	11	33	150	23
14	0	0	0	0	0	0	0	15	0
15	0	0	0	0	0	0	0	0	0
16	53	47	67	120	20	12	35	150	8
17	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	20	0
20	0	0	0	0	0	0	0	0	0

Πίνακας 11. Αποτελέσματα για την CMA-ES σε μοντέλο ταλαντωτή *duffing*.

Η OPPO δεν κατάφερε καμία φορά να αναγνωρίσει όλες τις παραμέτρους σωστά κυρίως λόγω αδυναμίας αναγνώρισης μίας παραμέτρου, της a_2 .

Συνολικά κατάφερε να αναγνωρίσει τις k_1 19 φορές, k_2 20 φορές, k_3 19 φορές, c_1 1 φορά, c_2 15 φορές, c_3 46 φορές, a_1 6 φορές, a_2 0 φορές και την a_3 5 φορές.

Δοκιμές Αριθμός	Ποσοστά Σφάλματος Κάθε Μεταβλητής για OPPO								
	k_1	k_2	k_3	c_1	c_2	c_3	α_1	α_2	α_3
1	1	1	1	4	0	1	6	75	10
2	11	5	12	22	41	14	3	75	40
3	2	2	2	28	6	4	50	129	77
4	4	2	3	88	10	14	27	74	49
5	1	2	1	11	5	1	36	150	56
6	1	0	0	27	5	3	38	20	63
7	1	1	1	43	5	5	11	73	4
8	0	0	0	13	2	1	11	35	18
9	1	1	1	18	2	3	12	75	25

10	0	1	1	24	5	2	47	71	77
11	2	2	3	67	4	12	34	75	32
12	3	3	3	73	11	6	8	149	9
13	2	2	2	37	4	3	7	149	13
14	6	3	5	49	2	10	50	150	111
15	1	1	1	37	5	5	31	74	50
16	6	2	7	90	16	19	17	72	9
17	3	3	3	73	11	6	8	149	9
18	2	2	2	37	4	3	7	149	13
19	6	3	5	49	2	10	50	150	111
20	1	1	1	37	5	5	31	74	50

Πίνακας 12. Αποτελέσματα για την OPPO σε μοντέλο ταλαντωτή *duffing*.

Ο SPSO αλγόριθμος αναγνώρισε σωστά όλες τις παραμέτρους 5 φορές στις 20 δοκιμές. Συνολικά κατάφερε να αναγνωρίσει τις k_1 19 φορές, k_2 19 φορές, k_3 19 φορές, c_1 12 φορές, c_2 20 φορές, c_3 19 φορές, a_1 5 φορές, a_2 8 φορές και την a_3 5 φορές. Οι μεταβλητές που δεν μπόρεσαν να αναγνωριστούν τις περισσότερες φορές ήταν αυτές που συμμετέχουν στην μη γραμμικότητα της στιβαρότητας.

Δοκιμές	Ποσοστά Σφάλματος Κάθε Μεταβλητής για SPSO								
Αριθμός	k_1	k_2	k_3	c_1	c_2	c_3	α_1	α_2	α_3
1	2	2	1	14	5	2	50	117	77
2	1	2	1	3	0	1	11	114	16
3	1	0	0	8	3	1	30	26	47
4	1	1	1	0	1	0	12	61	18
5	1	0	0	12	3	2	22	22	39
6	1	0	0	11	0	1	35	9	58
7	53	47	68	82	4	20	22	150	50
8	1	1	0	18	5	2	47	32	76
9	1	1	0	5	2	1	20	51	31
10	1	1	0	4	1	1	15	35	23
11	0	0	0	2	0	0	3	10	5
12	0	0	0	1	0	0	1	1	2
13	1	1	0	14	3	2	24	38	39
14	0	0	0	0	0	0	0	3	1
15	0	0	0	4	1	0	13	1	22
16	1	1	1	14	2	2	12	69	18
17	0	0	0	1	0	0	1	1	2
18	1	1	0	14	3	2	24	38	39
19	0	0	0	0	0	0	0	3	1
20	0	0	0	4	1	0	13	1	22

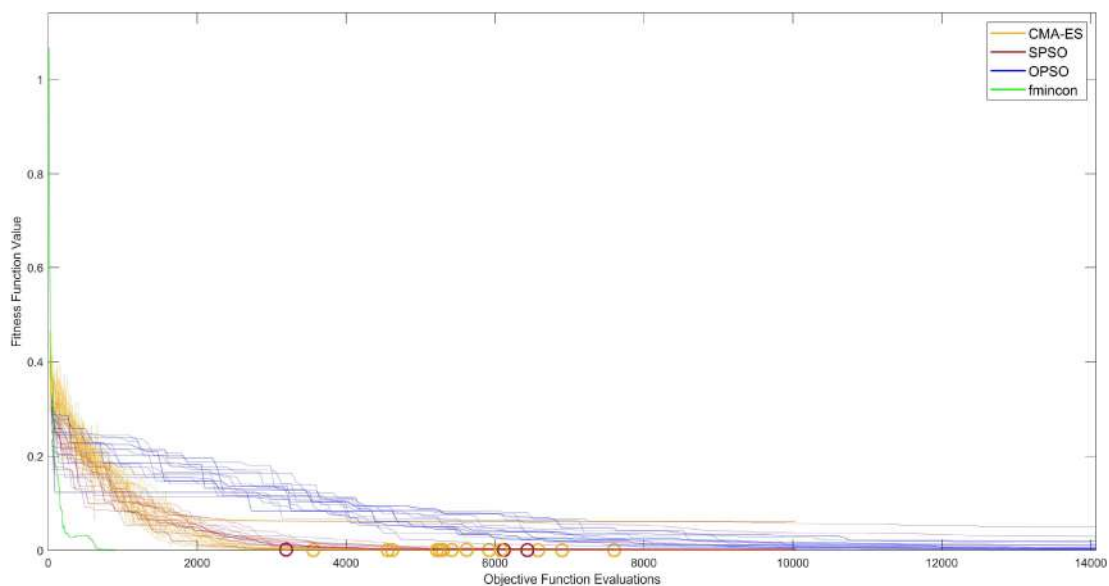
Πίνακας 13. Αποτελέσματα για την SPSO σε μοντέλο ταλαντωτή *duffing*.

Η *fmfcn* σε όλες τις δοκιμές εγκλωβίστηκε σε τοπικό ελάχιστο χωρίς να καταφέρει καμία επιτυχημένη δοκιμή και έτσι η μέθοδος σταμάτησε πρόωρα. Κατάφερε όμως να αναγνωρίσει τις k_1 20 φορές, k_2 20 φορές, k_3 20 φορές, c_1 0 φορές, c_2 20 φορές, c_3 20 φορές, a_1 0 φορές, a_2 20 φορές και την a_3 0 φορές.

Δοκιμές	Ποσοστά Σφάλματος Κάθε Μεταβλητής για fmincon								
Αριθμός	k_1	k_2	k_3	c_1	c_2	c_3	α_1	α_2	α_3
1	0	0	0	20	2	2	11	4	22
2	0	0	0	20	2	2	11	4	22
3	0	0	0	20	2	2	11	4	22
4	0	0	0	20	2	2	11	4	22
5	0	0	0	20	2	2	11	4	22
6	0	0	0	20	2	2	11	4	22
7	0	0	0	20	2	2	11	4	22
8	0	0	0	20	2	2	11	4	22
9	0	0	0	20	2	2	11	4	22
10	0	0	0	20	2	2	11	4	22
11	0	0	0	20	2	2	11	4	22
12	0	0	0	20	2	2	11	4	22
13	0	0	0	20	2	2	11	4	22
14	0	0	0	20	2	2	11	4	22
15	0	0	0	20	2	2	11	4	22
16	0	0	0	20	2	2	11	4	22
17	0	0	0	20	2	2	11	4	22
18	0	0	0	20	2	2	11	4	22
19	0	0	0	20	2	2	11	4	22
20	0	0	0	20	2	2	11	4	22

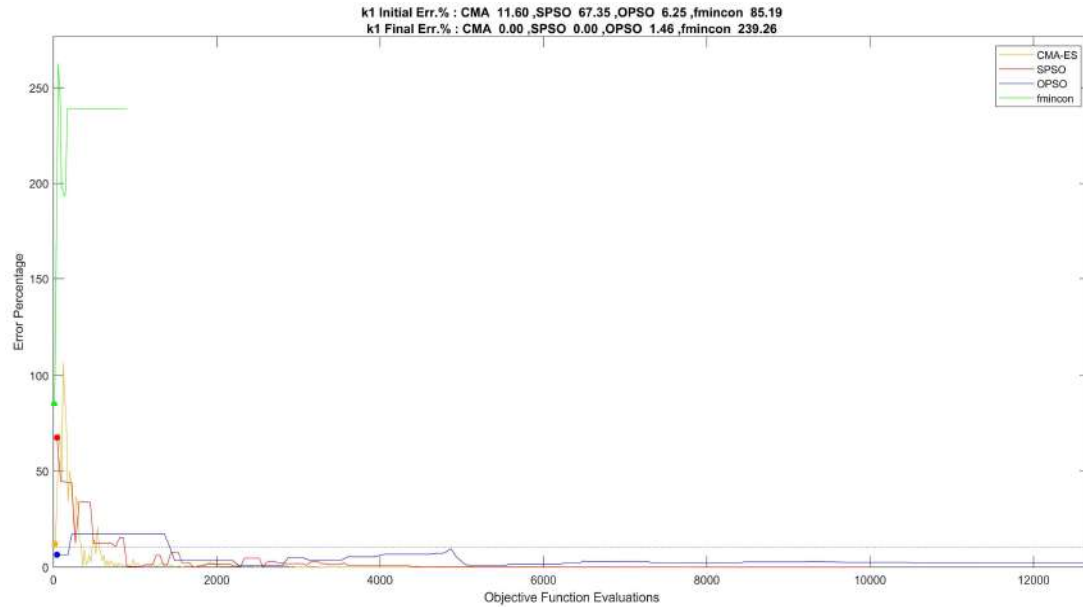
Πίνακας 14. Αποτελέσματα για την *fmincon* σε μοντέλο ταλαντωτή *duffing*.

Παρακάτω παρουσιάζεται συνολικά η μεταβολή της συνάρτησης καταλληλότητας κατά τις επαναλήψεις και για τους τέσσερις αλγόριθμους. Η πρώτη επανάληψη έπειτα από την οποία η δοκιμή ενός αλγόριθμου συνέκλινε αναγνωρίζοντας σωστά όλες τις μεταβλητές συμβολίζεται με έναν κύκλο.

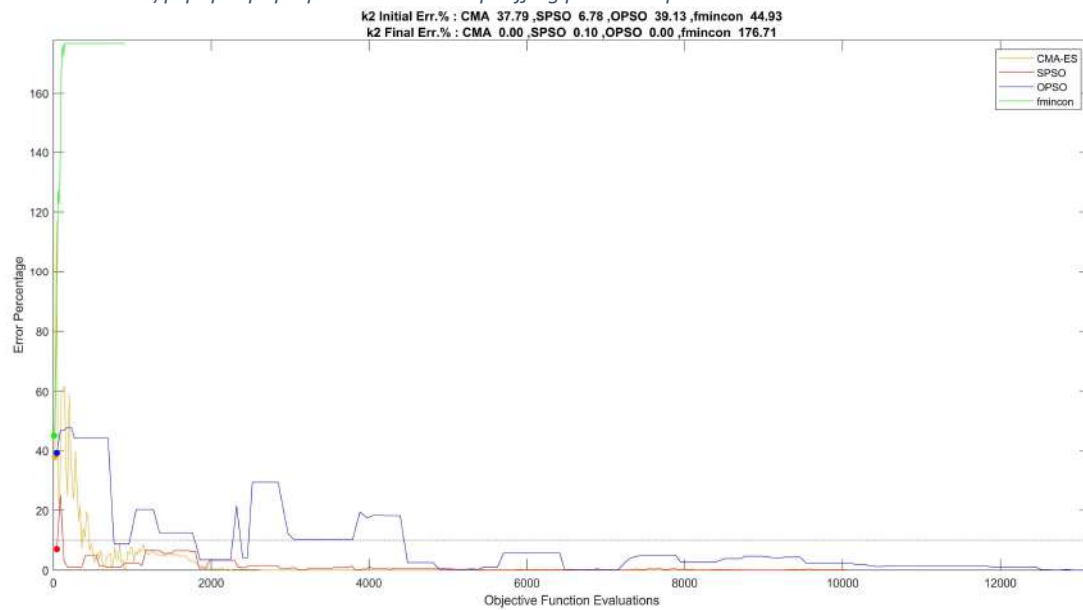


Εικόνα 90. Τιμή αντικειμενικής συνάρτησης με τον αριθμό υπολογισμών της αντικειμενικής συνάρτησης για όλες τις μεθόδους και τις δοκιμές.

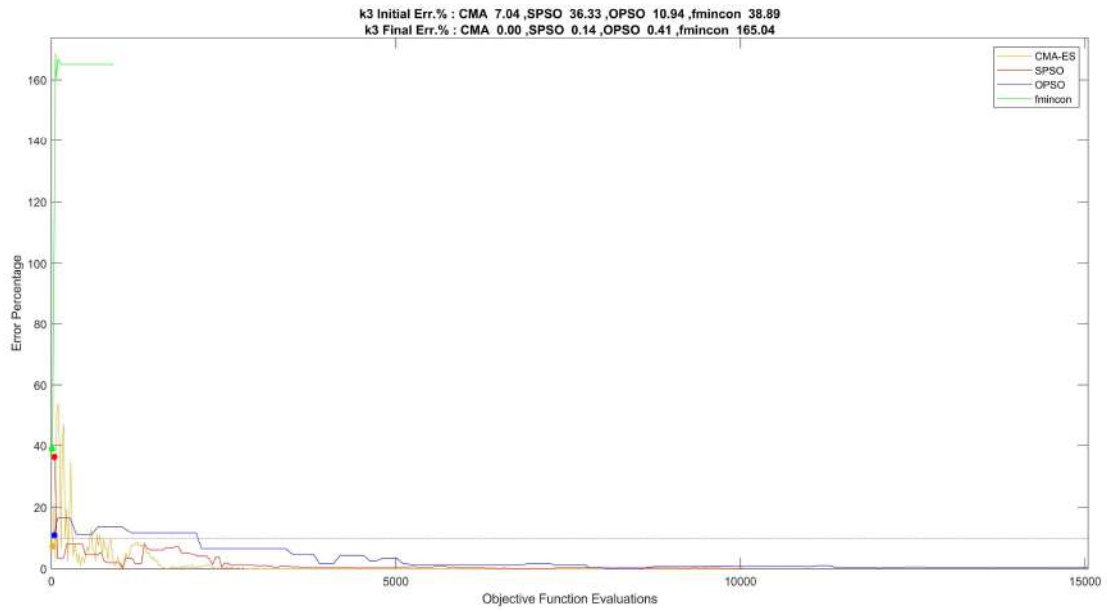
Στην συνέχεια παρουσιάζουμε την μεταβολή των παραμέτρων ξεχωριστά. Επιλέγουμε για κάθε μέθοδο τις αντίστοιχες δοκιμές ώστε τα σημεία εκκίνησης για κάθε μέθοδο να είναι όσο το δυνατόν κοντά. Σημειώνουμε πως μπορεί να συμπεριλαμβάνονται για τα παρακάτω διαγράμματα περιπτώσεις όπου δεν αναγνωρίστηκε σωστά η παράμετρος για κάποια μέθοδο.



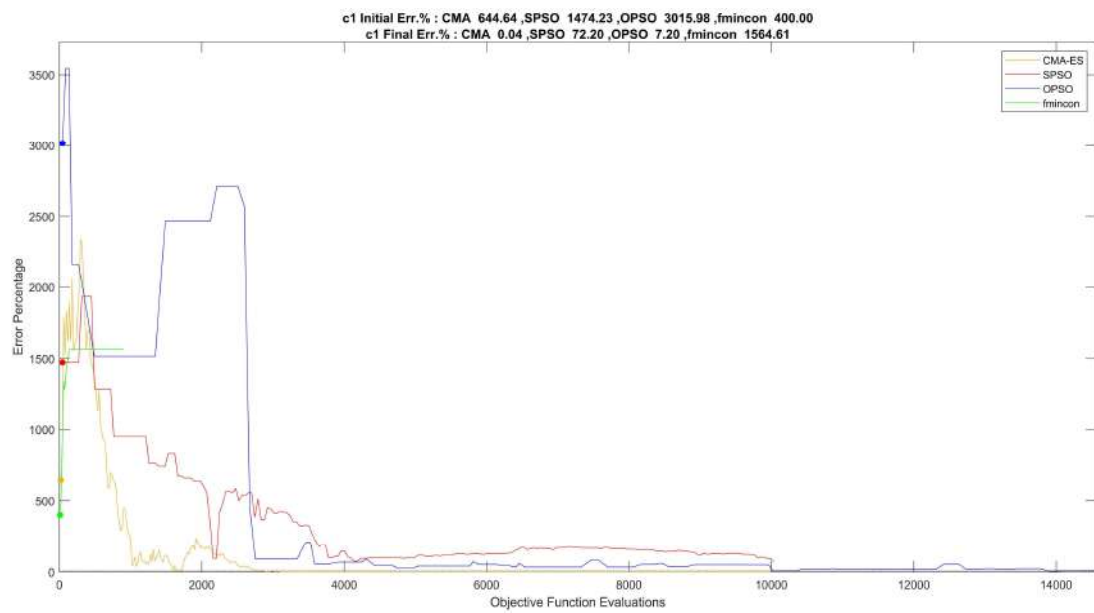
Εικόνα 91. Αναζήτηση παραμέτρου k_1 ταλαντωτή *duffing* μέσω των μεθόδων.



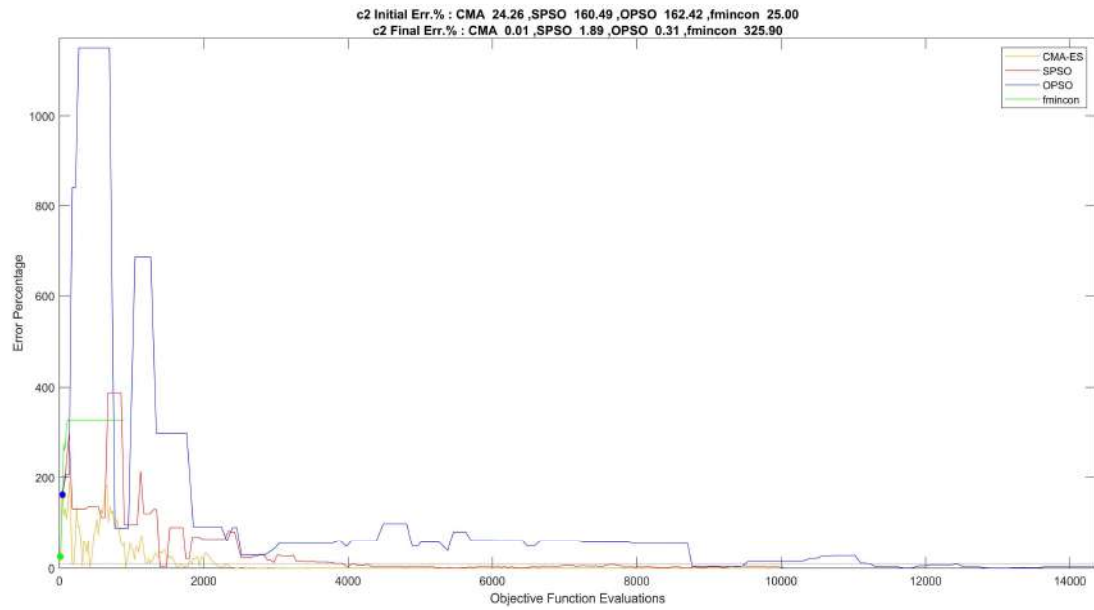
Εικόνα 92. Αναζήτηση παραμέτρου k_2 ταλαντωτή *duffing* μέσω των μεθόδων.



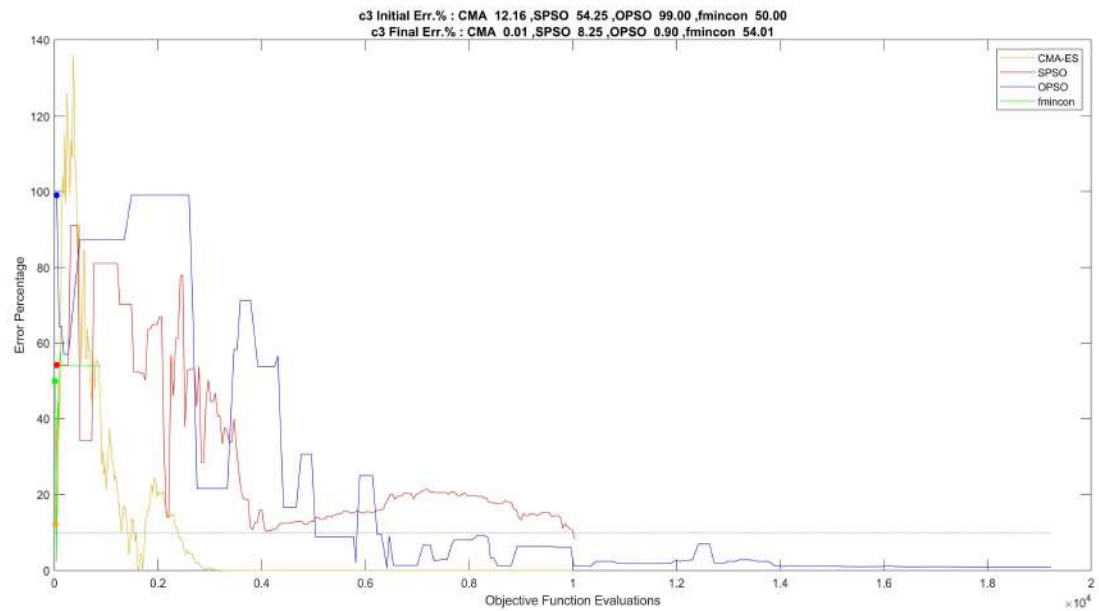
Εικόνα 93. Αναζήτηση παραμέτρου k3 ταλαντωτή duffing μέσω των μεθόδων.



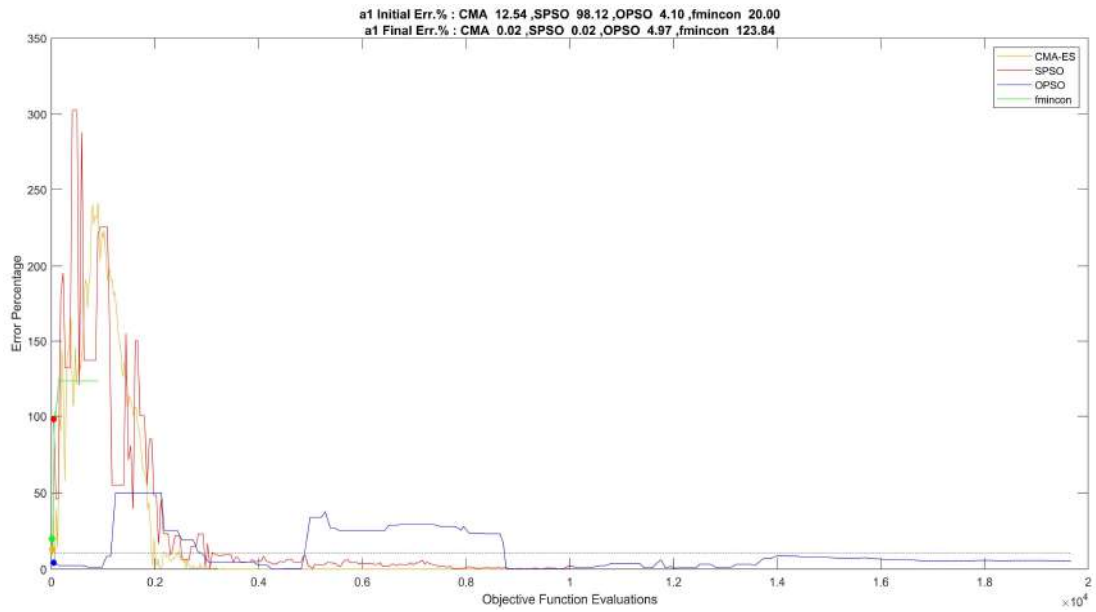
Εικόνα 94. Αναζήτηση παραμέτρου c1 ταλαντωτή duffing μέσω των μεθόδων.



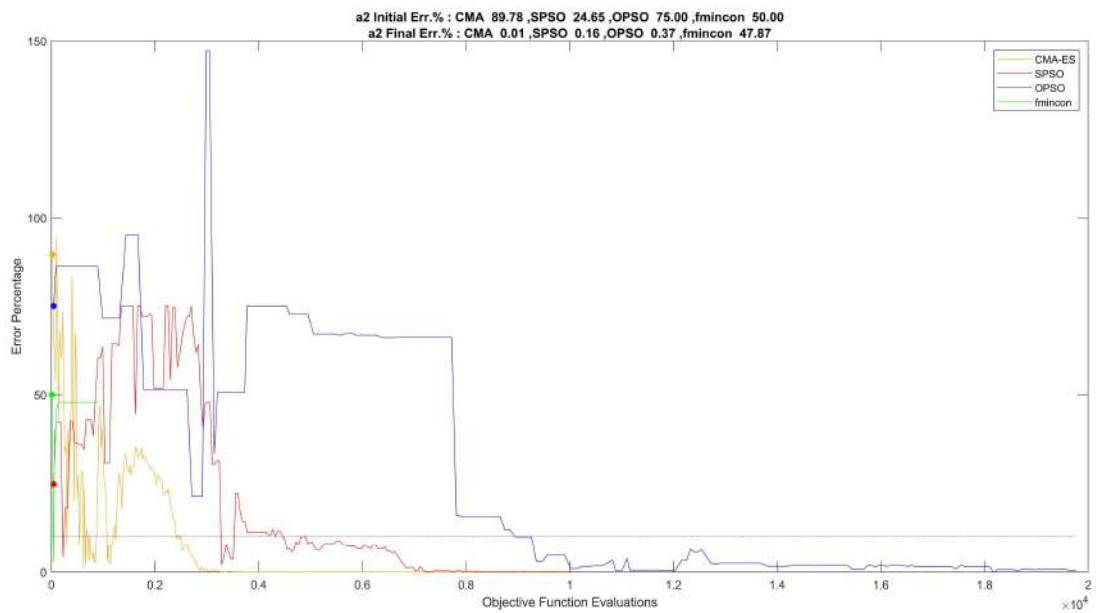
Εικόνα 95. Αναζήτηση παραμέτρου c2 ταλαντωτή duffing μέσω των μεθόδων



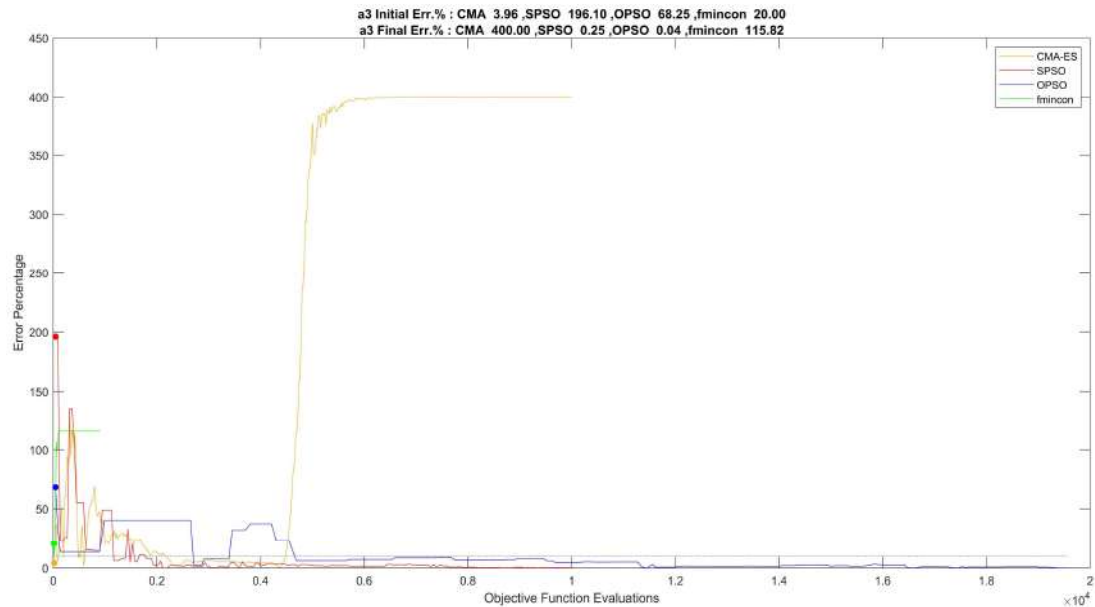
Εικόνα 96. Αναζήτηση παραμέτρου c3 ταλαντωτή duffing μέσω των μεθόδων



Εικόνα 97. Αναζήτηση παραμέτρου a_1 ταλαντωτή *duffing* μέσω των μεθόδων

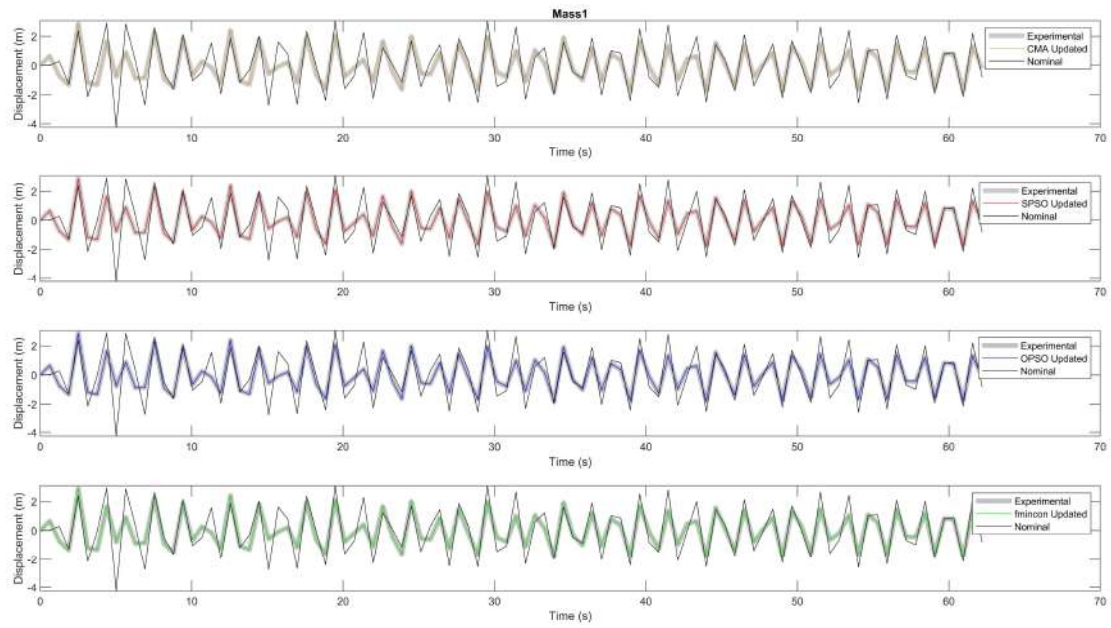


Εικόνα 98. Αναζήτηση παραμέτρου a_2 ταλαντωτή *duffing* μέσω των μεθόδων

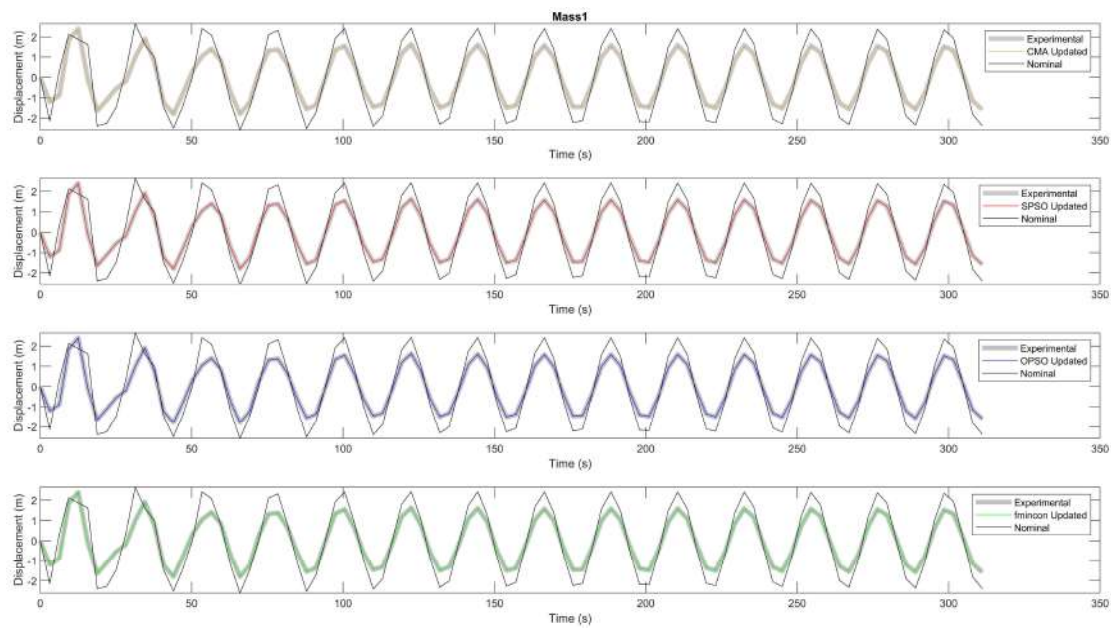


Εικόνα 99. Αναζήτηση παραμέτρου a_3 ταλαντωτή *duffing* μέσω των μεθόδων

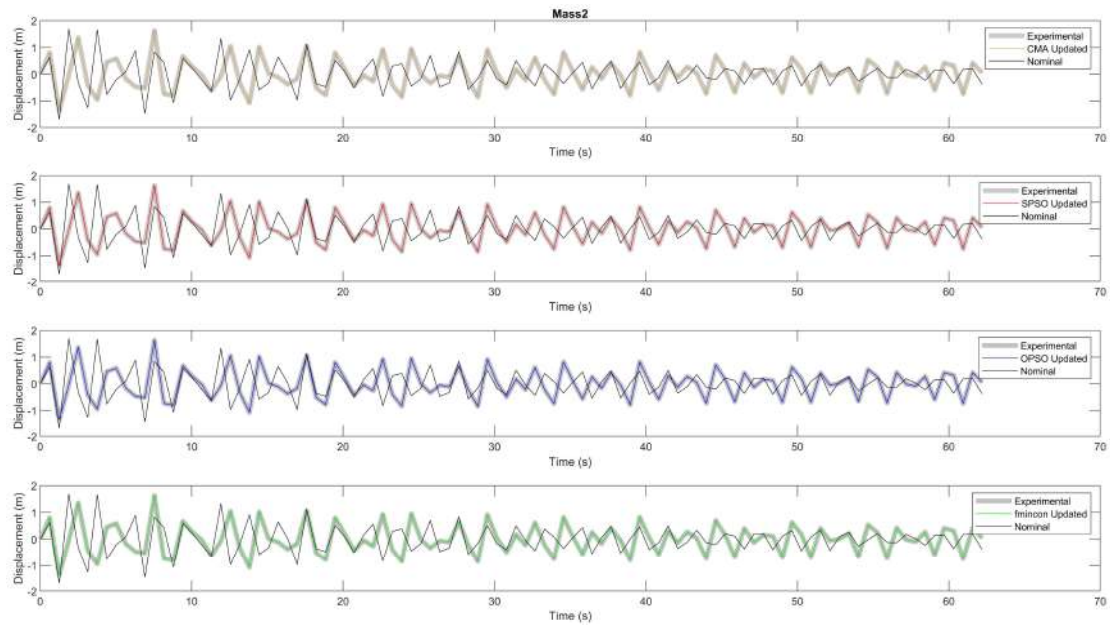
Παρακάτω παρουσιάζεται η απόκριση των μαζών για την διέγερση που περιγράψαμε. Συγκεκριμένα για κάθε μέθοδο επιλέξαμε τις παραμέτρους που προήλθαν από την καλύτερη δοκιμή της κάθε μεθόδου. Οι παράμετροι για την CMA-ES είχαν πλήρη ταύτιση με τις πειραματικές, όπως και για της SPSO εκτός από τις παραμέτρους a_2 και a_3 οι οποίες είχαν μια απόκλιση της τάξεως 3% και 1% αντίστοιχα. Για την OPSPSO μέθοδο οι μεταβλητές k_1 , k_2 και k_3 είχαν απόκλιση 1% , ενώ οι c_1 , c_2 , c_3 , a_1 και a_3 παρουσίασαν απόκλιση 4%, 0%, 1% ,6 % και 10 % αντίστοιχα. Για την OPSPSO όμως μεγάλη απόκλιση είχε η παράμετρος a_2 με τιμή αυτή να βρίσκεται στο 75%. Για την απόκριση που προέρχεται από τις παραμέτρους που αναγνώρισε η fmincon με την c_1 και την a_3 να έχουν το μεγαλύτερο σφάλμα 11% και 22% αντίστοιχα.



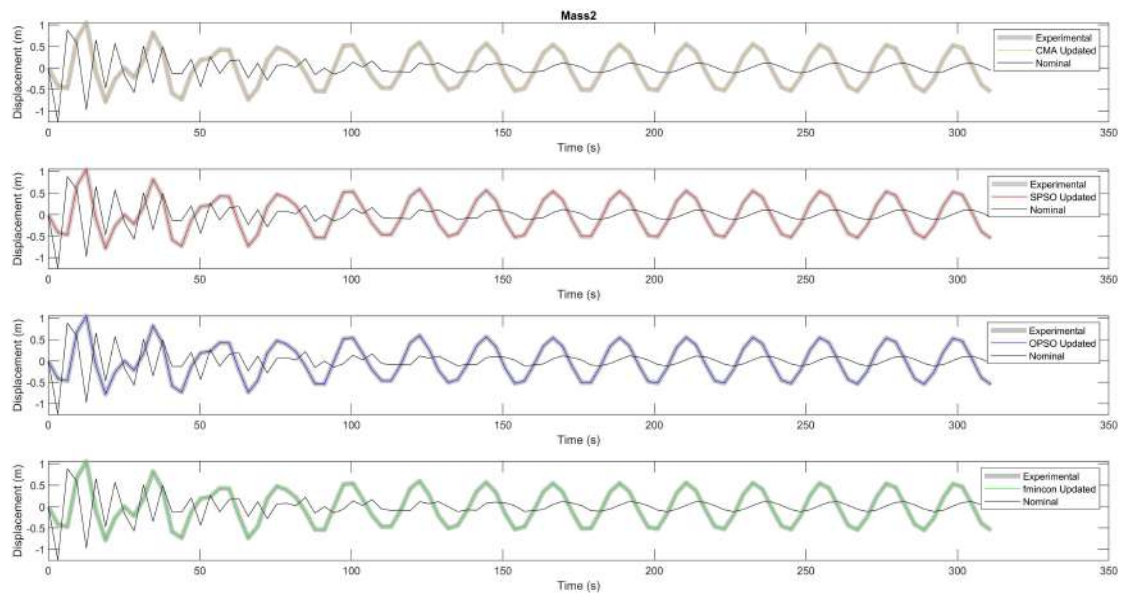
Εικόνα 100. Απόκριση μάζας 1, για την πρώτη διέγερση.



Εικόνα 101. Απόκριση μάζας 1, για την δεύτερη διέγερση.

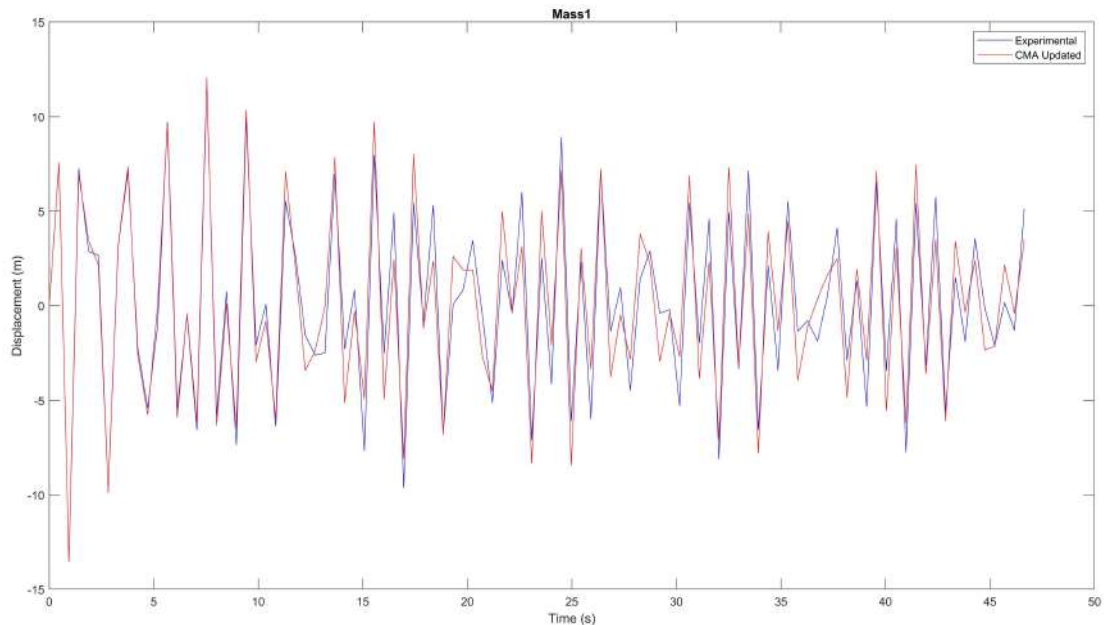


Εικόνα 102. Απόκριση μάζας 2, για την πρώτη διέγερση.



Εικόνα 103. Απόκριση μάζας 2, για την δεύτερη διέγερση.

Παρά την μεγάλη απόκλιση της παραμέτρων a_2 για την OPSP και της c_1 για την fmincon, παρατηρούμε πως η απόκριση των μαζών στην διέγερση που εφαρμόσαμε ταυτίζεται με την πειραματική και για αυτές τις μεθόδους. Αυτό παρατηρείται λόγω της διέγερσης που επιλέξαμε στην οποία οι παράμετροι αυτοί δεν κατέχουν μεγάλο ρόλο. Αυτό επιβεβαιώνεται χρησιμοποιώντας το ίδιο σετ των παραμέτρων της OPSP τρέχοντας το μοντέλο σε μια άλλη διέγερση με μεγαλύτερο πλάτος διέγερσης. Στην απόκριση που παρουσιάζεται παρακάτω φαίνεται η απόκλιση από τα πειραματικά δεδομένα. Παρόμοια αποτελέσματα αναμένουμε και για τις παραμέτρους που προέρχονται από την fmincon. Αυτό το φαινόμενο γίνεται ξεκάθαρο παρακάτω όπου εξετάζεται η ευαισθησία των μεταβλητών.



Εικόνα 104. Απόκριση του συστήματος *duffing* χρησιμοποιώντας το καλύτερο σεντ παραμέτρων που προέρχεται από την OPPO, σε διαφορετική διέγερση όμως.

Παρατηρούμε πως στους SPSO και OPPO οι μεταβλητές που δεν μπόρεσαν να αναγνωριστούν ήταν κυρίως οι τρεις παράμετροι της στιβαρότητας του κυβικού όρου. Συγκεκριμένα η μεγαλύτερη δυσκολία ήταν στην αναγνώριση της α_2 που αντιπροσωπεύει τον κυβικό όρο του ελατηρίου που συνδέει τις μάζες. Σε δεύτερο λόγο επίσης η c_1 παράμετρος ήταν επίσης δύσκολο να αναγνωριστεί.

7.3.1 Δοκιμή του ταλαντωτή *duffing* σε διαφορετική διέγερση (Περίπτωση 2)

Για να βελτιώσουμε τα αποτελέσματα δοκιμάσαμε το σύστημα σε δύο διαφορετικά σεντ διεγέρσεων. Συγκεκριμένα επιλέχθηκε ένα μεγαλύτερο πλάτος διέγερσης ώστε οι α_1, α_2 και α_3 να έχουν μεγαλύτερη σημασία στην απόκριση, $f_1 = f_2 = 10 \text{ N}$ και δύο διαφορετικές συχνότητες διέγερσης για τις δύο μάζες με 0.1592 Hz για την πρώτη και 0.3183 Hz για τη δεύτερη. Λόγο της διαφοράς φάσης στις δύο διεγέρσεις αναμένεται το μοντέλο να έχει μεγαλύτερη ευαισθησία στις τιμές του ελατηρίου που συνδέει τις δύο μάζες. Στην δεύτερη περίπτωση επιλέχθηκε ένα χαμηλότερο πλάτος με $f_1 = f_2 = 3 \text{ N}$ ώστε να μην υπάρχει μεγάλη διαφορά στην συνεισφορά των μη-γραμμικών και των γραμμικών όρων που συμμετέχουν στην στιβαρότητα. Σε αυτή την περίπτωση οι συχνότητες διέγερσης είναι 0.0796 Hz και 0.0398 Hz .

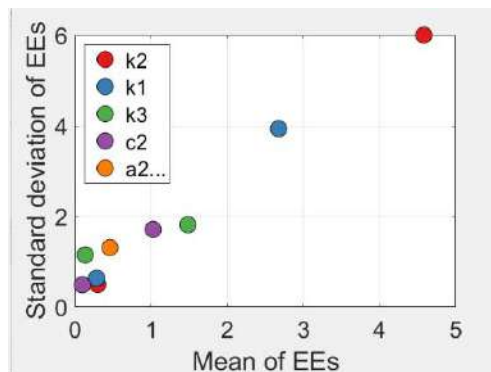
Μεταβάλλοντας την διέγερση και το πλάτος των δυνάμεων αναμένουμε οι αλλαγές των παραμέτρων αυτών να έχουν μεγαλύτερη "βαρύτητα" στην αντικειμενική συνάρτηση. Ονομάζουμε την πρώτη περίπτωση που δοκιμάσαμε Περίπτωση 1 και αυτή την δεύτερη περίπτωση ως Περίπτωση 2. Για να εκτιμήσουμε την ευαισθησία της αντικειμενικής συνάρτησης σε αλλαγές από τις παραμέτρους που θέλουμε να αναγνωρίσουμε εφαρμόζουμε την μέθοδο Morris. Στην μέθοδο αυτή κάθε φορά μεταβάλλεται μια παράμετρος και οι άλλοι μένουν σταθερές, οπότε από άποψη δειγματοληψίας είναι μια OAT (One at a Time) μέθοδος. Δύο μέτρα ευαισθησίας χρησιμοποιούνται στην μέθοδο αυτή για να τον κάθε παράγοντα, ένα μέτρο μ που υπολογίζει την συνολική επίδραση που έχει ο παράγοντας στο αποτέλεσμα και το μέτρο σ που αντιπροσωπεύει συνολικά τις επιδράσεις δεύτερης και ανώτερης τάξης την ευαισθησία.

Αν ένα μοντέλο έχει μεταβλητές εισόδου $X_i, i = 1, \dots, k$ και διακριτοποιηθεί το πεδίο ορισμού σε ένα πλέγμα p επιπέδων, Ω , τότε για κάθε τιμή του X οι στοιχειώσεις επιπτώσεις ή επιδράσεις του i – οστού παράγοντα υπολογίζονται ως εξής

$$d_i(\mathbf{X}) = \frac{[y(X_1, \dots, X_{i-1}, X_i + \Delta, X_{i+1}, \dots, X_k) - y(\mathbf{X})]}{\Delta},$$

όπου $\Delta \in \{1/(p-1), \dots, 1 - 1/(p-1)\}$ με $\Delta + X$ να ανήκει ακόμα στο πεδίο ορισμού. Η κατανομή F_i των στοιχειωδών επιδράσεων υπολογίζεται πραγματοποιώντας μια δειγματοληψία από r σημεία $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(r)}$ από τον Ω και υπολογίζοντας r στοιχειώδης επιδράσεις για κάθε $d_i(\mathbf{X}^{(1)}), d_i(\mathbf{X}^{(2)}), \dots, d_i(\mathbf{X}^{(r)})$ για κάθε παράμετρο εισαγωγής i . Η μέση τιμή της κατανομής και η τυπική απόκλιση είναι τα μέτρα της συνολικής επίδρασης και των επιδράσεων ανώτερης τάξης μ και σ που αναφέραμε αντίστοιχα. Για την δειγματοληψία χρησιμοποιείται σχεδιασμός Λατινικού Υπερκύβου.

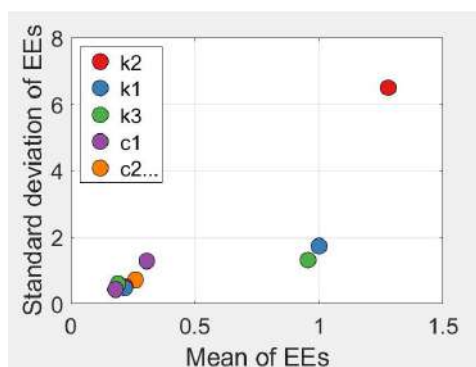
Παρακάτω παρουσιάζονται οι επιδράσεις για την πρώτη δοκιμή όπου παρατηρούνται σχετικά μεγάλες διαφορές. Για τον υπολογισμό των επιδράσεων χρησιμοποιήθηκε το “SAFE” toolbox στο MATLAB που αναπτύχθηκε από την F. Pianosi [39].



Παράμετρος	μ	σ
k_1	2.674	3.932
k_2	4.582	5.999
k_3	1.482	1.813
c_1	0.284	0.633
c_2	1.025	1.711
c_3	0.297	0.496
a_1	0.134	1.153
a_2	0.457	1.308
a_3	0.098	0.489

Εικόνα 105. Επιδράσεις των παραμέτρων για την Περίπτωση 1.

Έπειτα από τις αλλαγές στην διαμόρφωση της αντικειμενικής συνάρτησης πίνακα οι επιδράσεις των μεταβλητών είναι πιο κοντά μεταξύ τους.



Παράμετρος	μ	σ
k_1	1.000	1.738
k_2	1.279	6.497
k_3	0.956	1.315
c_1	0.306	1.288
c_2	0.261	0.723
c_3	0.190	0.601
a_1	0.219	0.528
a_2	0.217	0.478
a_3	0.180	0.433

Εικόνα 106. 107 Επιδράσεις των παραμέτρων για την Περίπτωση 2

Στον επόμενη πίνακα έχουμε τα αποτελέσματα για την CMA-ES στην Περίπτωση 2, η οποία πετυχαίνει 19 φορές 20 φορές σε αντίθεση με την Περίπτωση 1 όπου συνέκλινε στα σωστά όρια μόλις 13 φορές στις 20.

Δοκιμές	Ποσοστά Σφάλματος Κάθε Μεταβλητής για CMA-ES								
Αριθμός	k_1	k_2	k_3	c_1	c_2	c_3	α_1	α_2	α_3
1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0
3	104	44	24	1751	841	115	156	7	50
4	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0

Ο αλγόριθμος SPSO έχει 10 επιτυχημένες δοκιμές στις 20 διπλασιάζοντας την απόδοση του συγκριτικά με την Περίπτωση 1. Συνολικά κατάφερε να αναγνωρίσει τις k_1 20 φορές, k_2 20 φορές, k_3 20 φορές, c_1 10 φορές, c_2 20 φορές, c_3 16 φορές, α_1 20 φορές, α_2 19 φορές και την α_3 20 φορές.

Δοκιμές	Ποσοστά Σφάλματος Κάθε Μεταβλητής για SPSO								
Αριθμός	k_1	k_2	k_3	c_1	c_2	c_3	α_1	α_2	α_3
1	0	0	0	57	1	7	0	0	0
2	0	0	0	72	1	8	1	3	1
3	0	0	0	17	3	2	0	1	0
4	0	0	0	426	6	54	2	1	3
5	1	1	0	163	1	21	6	11	1
6	0	0	0	1	2	0	3	0	0
7	0	0	0	54	2	6	0	0	0
8	0	0	0	6	0	1	0	1	0
9	0	0	0	0	0	0	0	0	0
10	0	0	0	2	1	0	0	0	0
11	0	0	0	18	0	2	0	0	0
12	0	0	0	143	1	18	0	3	0
13	0	0	0	8	0	1	0	0	0

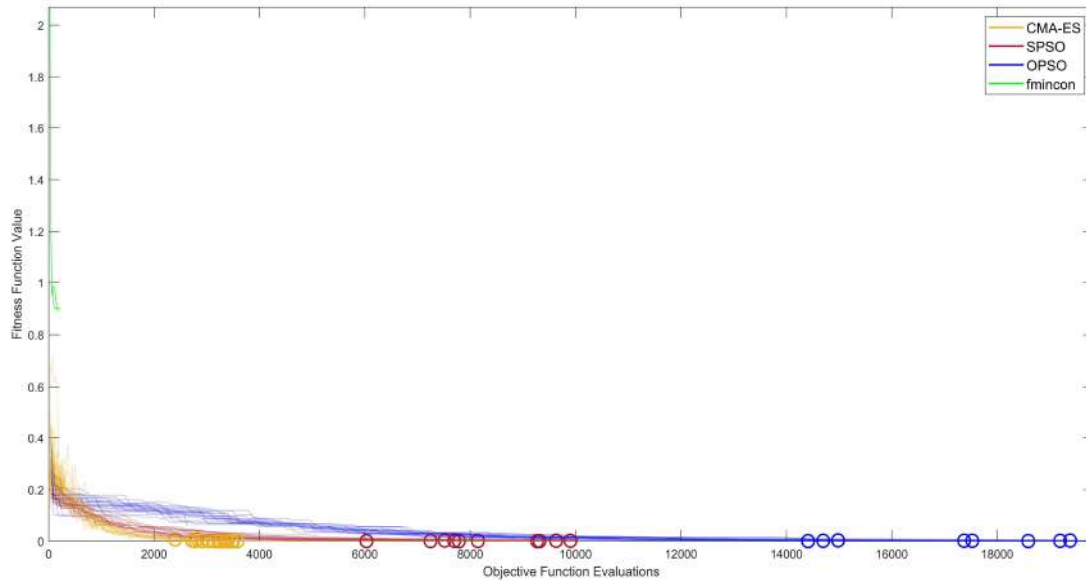
14	0	0	0	8	0	1	0	0	0
15	1	0	0	7	2	2	6	1	0
16	0	0	0	41	0	5	1	1	1
17	0	0	0	5	1	0	0	0	0
18	0	0	0	5	0	1	3	4	0
19	0	0	0	0	0	0	0	0	0
20	1	1	1	99	4	11	5	3	2

Η απόδοση του OPSO επίσης βελτιώθηκε, αναγνωρίζοντας αυτή τη φορά σωστά 8 φορές στις 20. Υπενθυμίζουμε πως στην Περίπτωση 1 δεν κατάφερε καμία φορά να αναγνωρίσει σωστά όλες τις παραμέτρους. Συνολικά κατάφερε να αναγνωρίσει τις k_1 20 φορές, k_2 20 φορές, k_3 20 φορές, c_1 8 φορές, c_2 20 φορές, c_3 19 φορές, a_1 20 φορές, a_2 20 φορές και την a_3 20 φορές.

Δοκιμές	Ποσοστά Σφάλματος Κάθε Μεταβλητής για OPSO								
	Αριθμός	k_1	k_2	k_3	c_1	c_2	c_3	a_1	a_2
1	1	0	0	16	1	2	7	2	0
2	0	0	0	7	0	1	0	0	0
3	0	1	0	1	1	0	5	7	0
4	0	0	0	5	0	0	0	0	0
5	1	0	0	0	1	1	3	1	0
6	0	0	0	4	0	0	0	0	0
7	0	0	0	19	3	2	2	1	0
8	1	0	0	35	2	4	5	1	1
9	0	0	0	15	0	2	2	1	0
10	0	0	0	45	0	6	1	1	0
11	1	0	0	10	1	2	4	2	0
12	1	0	1	76	1	8	2	1	1
13	0	0	0	21	0	3	5	2	0
14	0	0	0	35	1	4	1	0	0
15	0	0	0	11	0	1	0	1	0
16	0	0	0	24	1	3	2	0	0
17	1	0	0	2	0	0	4	1	0
18	1	0	0	2	1	1	3	1	0
19	1	0	1	14	0	1	10	3	0
20	0	0	0	42	1	5	3	0	0

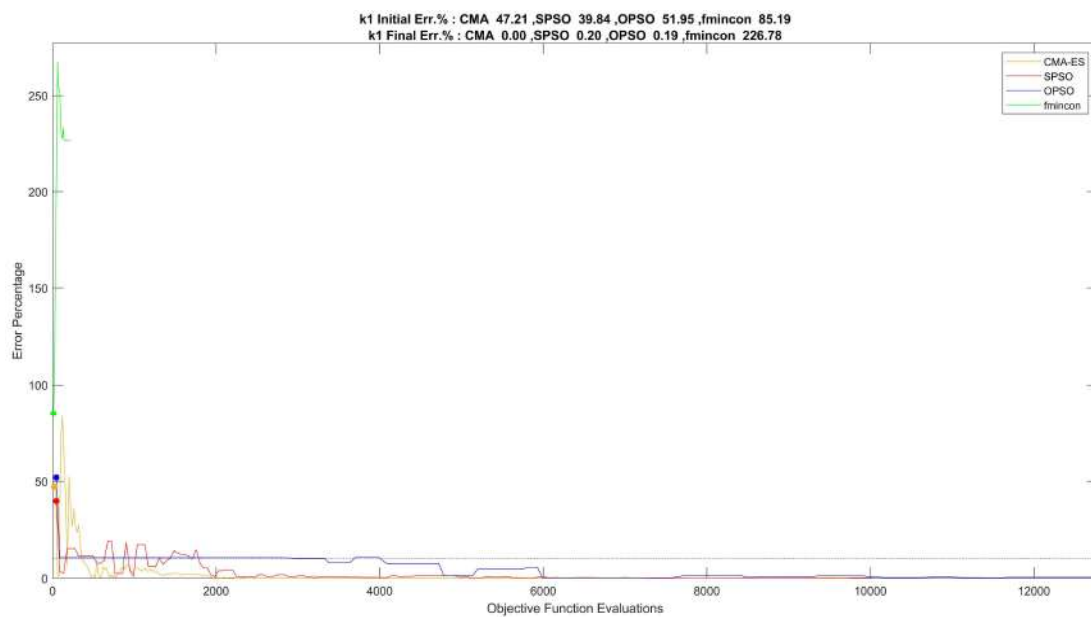
Η fmincon και πάλι δεν κατάφερε να αναγνωρίσει καμία φορά σωστά τις παραμέτρους

Παρακάτω παρουσιάζεται συνολικά για την Περίπτωση 2 η μεταβολή της συνάρτησης καταλληλότητας κατά τις επαναλήψεις και για τους τέσσερις αλγόριθμους. Η πρώτη επανάληψη έπειτα από την οποία η δοκιμή ενός αλγόριθμος συνέκλινε αναγνωρίζοντας σωστά όλες τις μεταβλητές συμβολίζεται με έναν κύκλο.

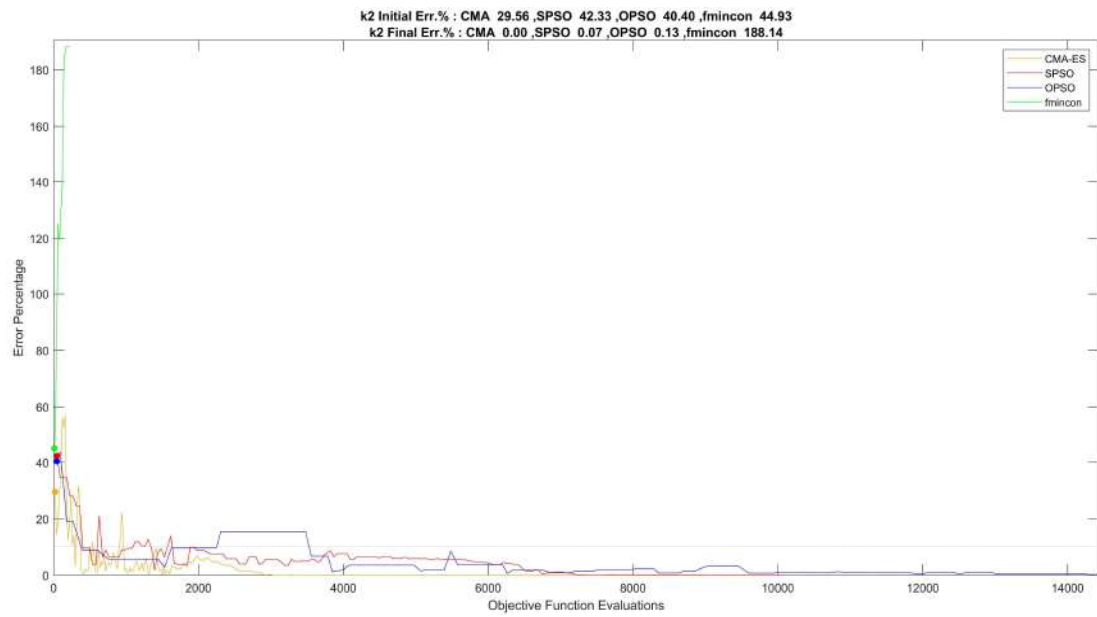


Εικόνα 108. Απόδοση των μεθόδων στην Περίπτωση 2 του ταλαντωτή *duffing*.

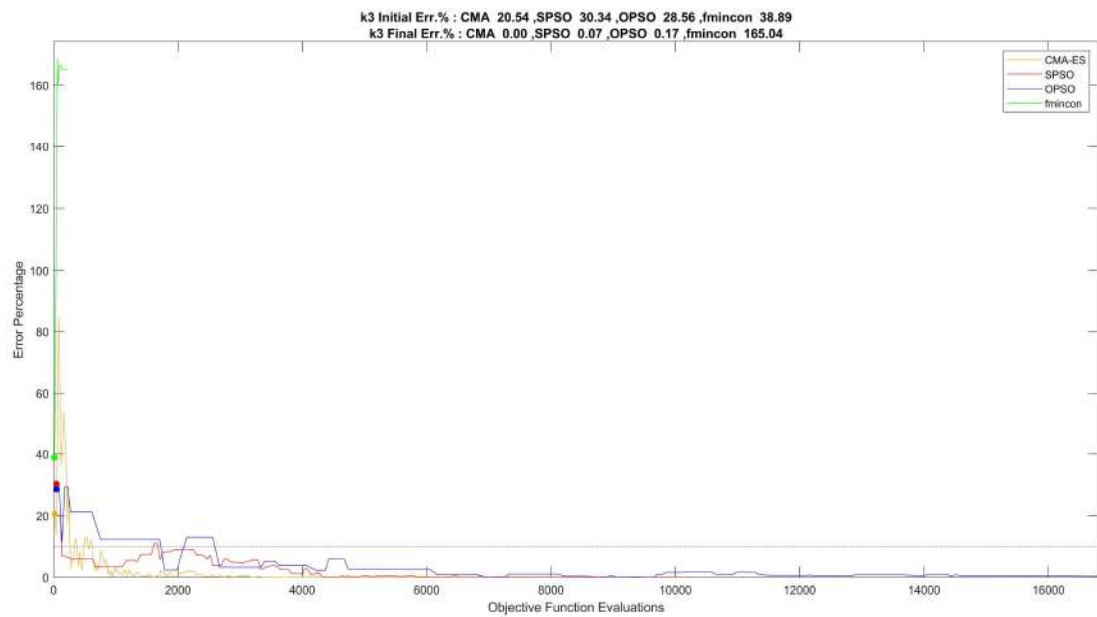
Στην συνέχεια παρουσιάζουμε την μεταβολή των παραμέτρων ξεχωριστά. Επιλέγουμε για κάθε μέθοδο τις αντίστοιχες δοκιμές ώστε τα σημεία εκκίνησης για κάθε μέθοδο να είναι όσο το δυνατόν κοντά. Σημειώνουμε πως μπορεί να συμπεριλαμβάνονται για τα παρακάτω διαγράμματα περιπτώσεις όπου δεν αναγνωρίστηκε σωστά η παράμετρος για κάποια μέθοδο.



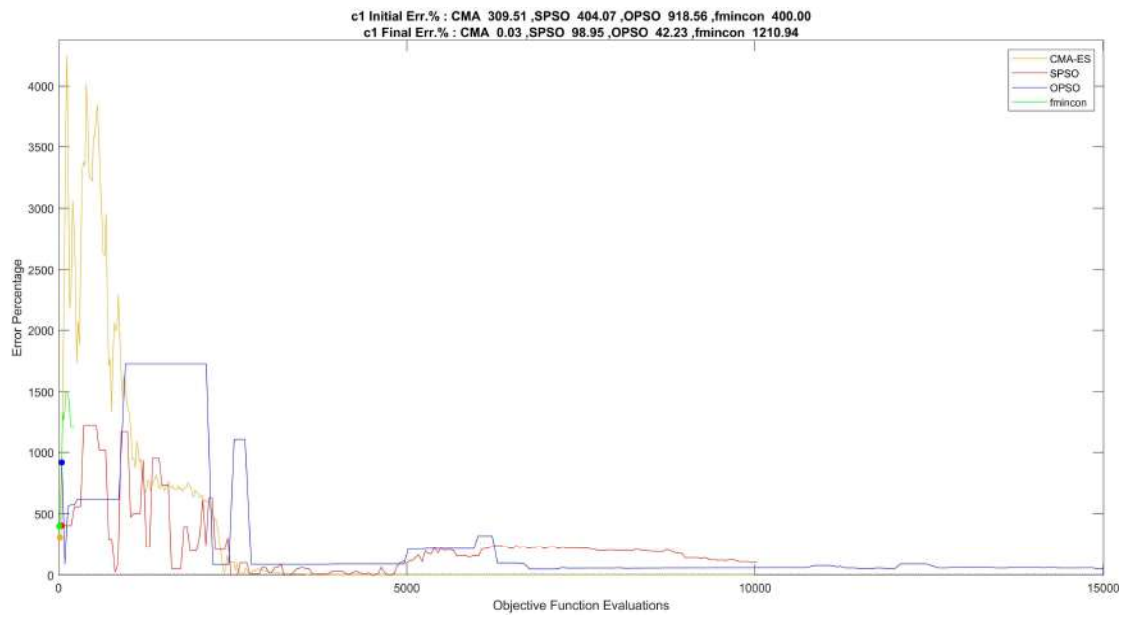
Εικόνα 109. Αναζήτηση παραμέτρου k_1 του ταλαντωτή *duffing*.



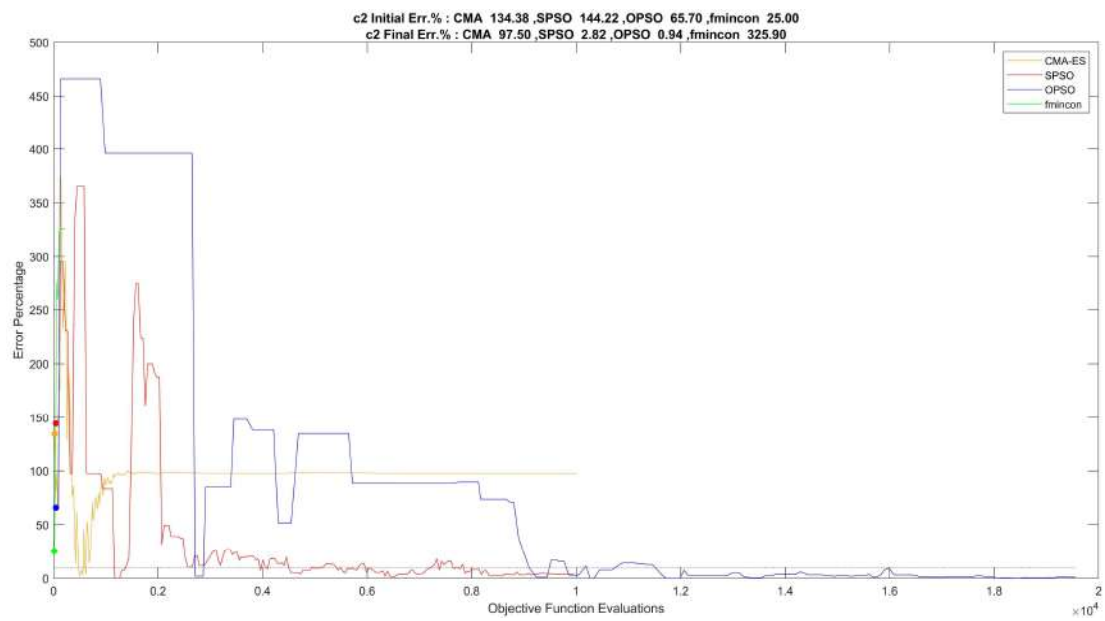
Εικόνα 110. Αναζήτηση παραμέτρου k_2 του ταλαντωτή *duffing*.



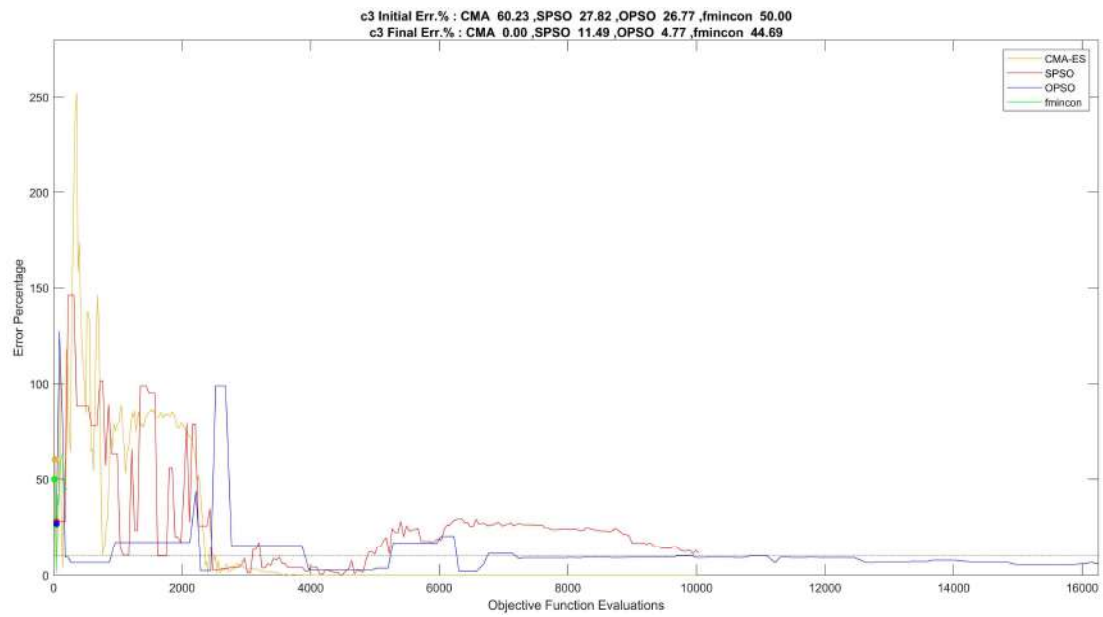
Εικόνα 111. Αναζήτηση παραμέτρου k_3 του ταλαντωτή *duffing*.



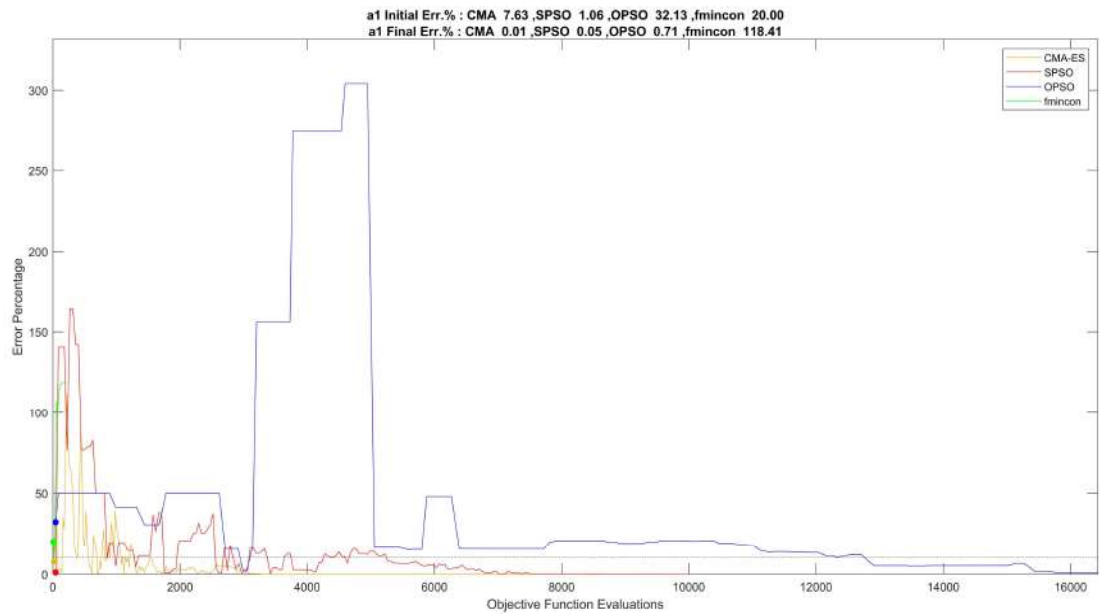
Εικόνα 112. Αναζήτηση παραμέτρου c1 του ταλαντωτή duffing.



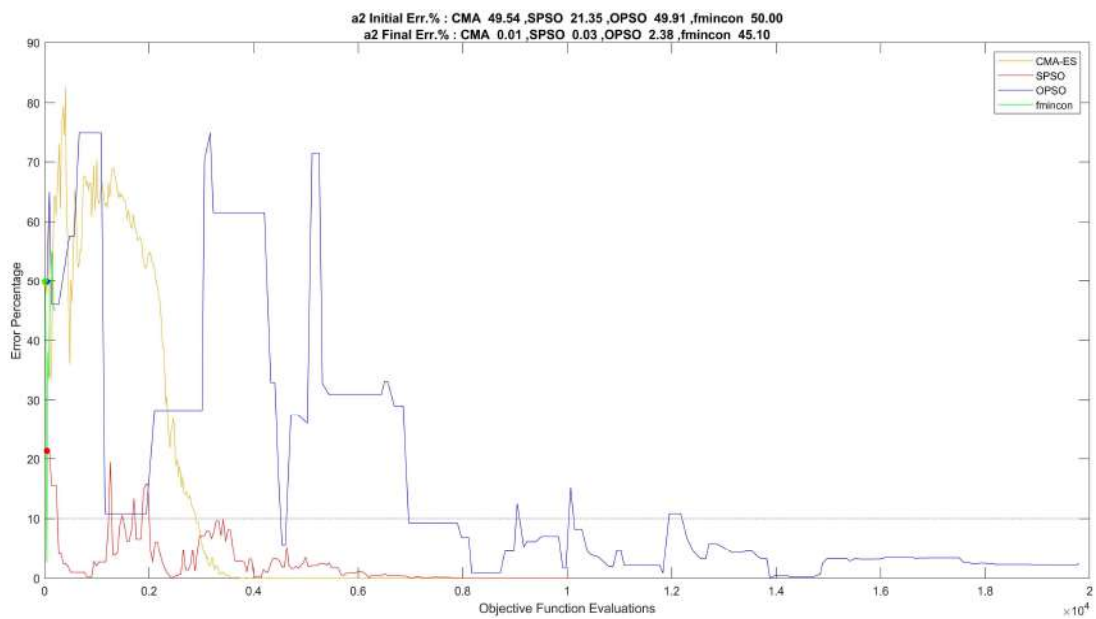
Εικόνα 113. Αναζήτηση παραμέτρου c2 του ταλαντωτή duffing.



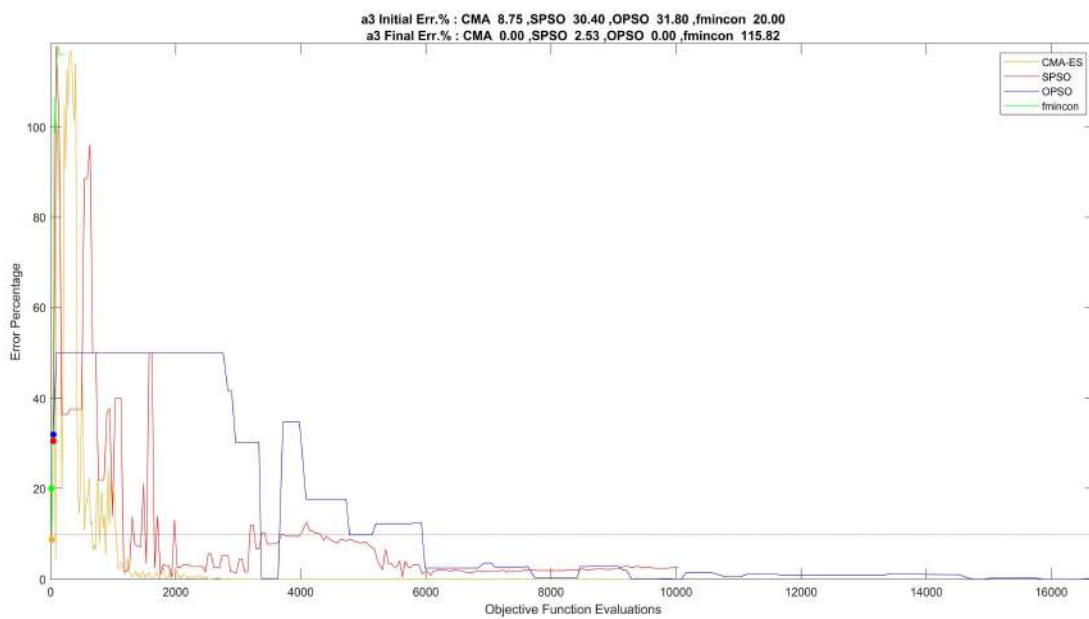
Εικόνα 114. Αναζήτηση παραμέτρου c3 του ταλαντωτή duffing.



Εικόνα 115. Αναζήτηση παραμέτρου a1 του ταλαντωτή duffing.

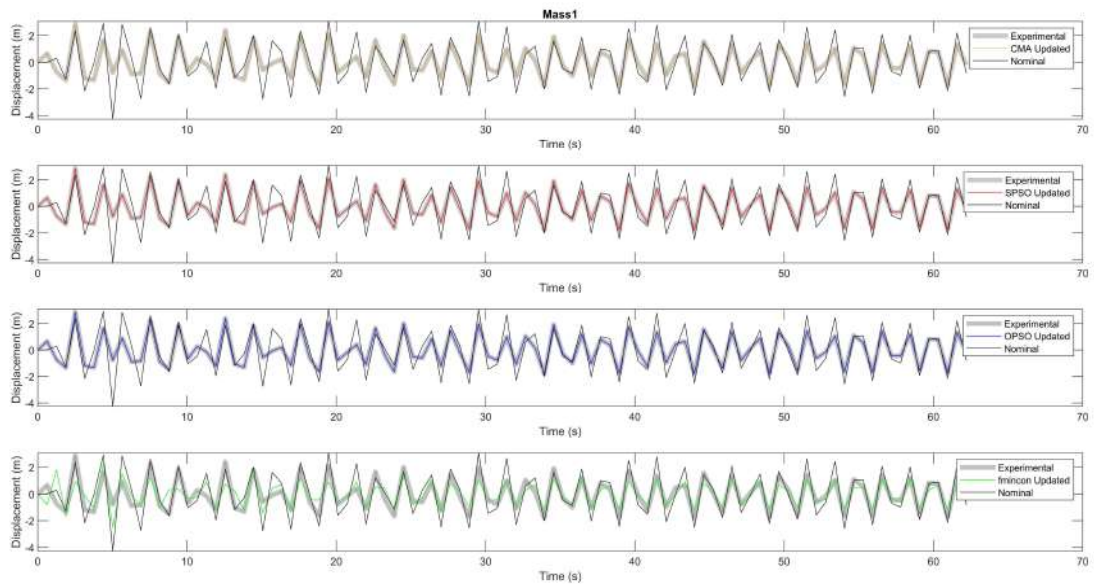


Εικόνα 116. Αναζήτηση παραμέτρου a_2 του ταλαντωτή *duffing*.

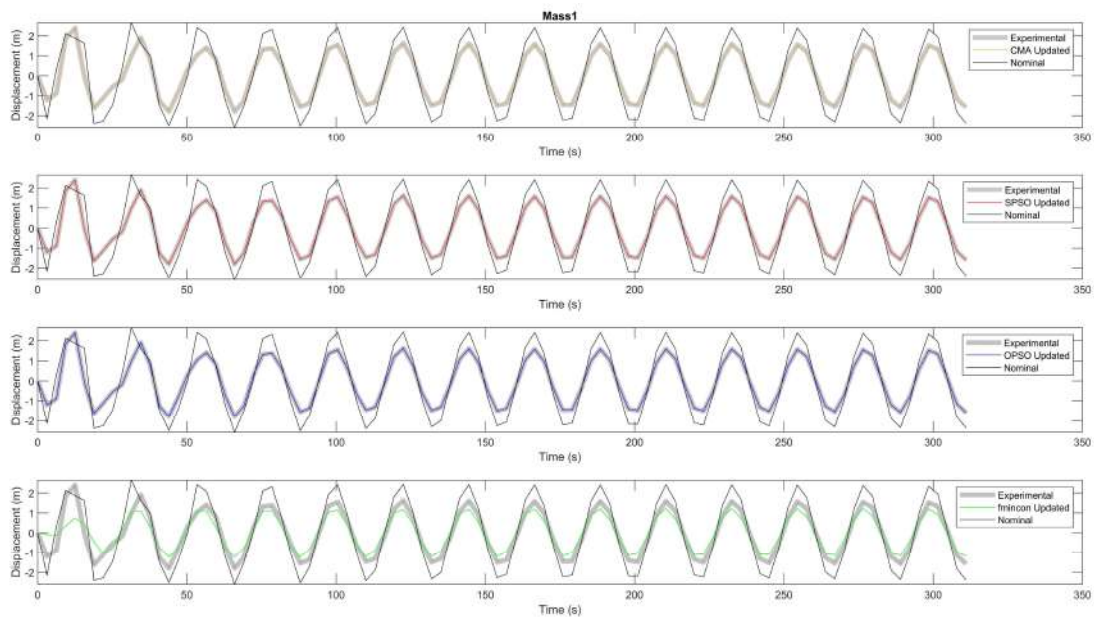


Εικόνα 117. Αναζήτηση παραμέτρου a_3 του ταλαντωτή *duffing*.

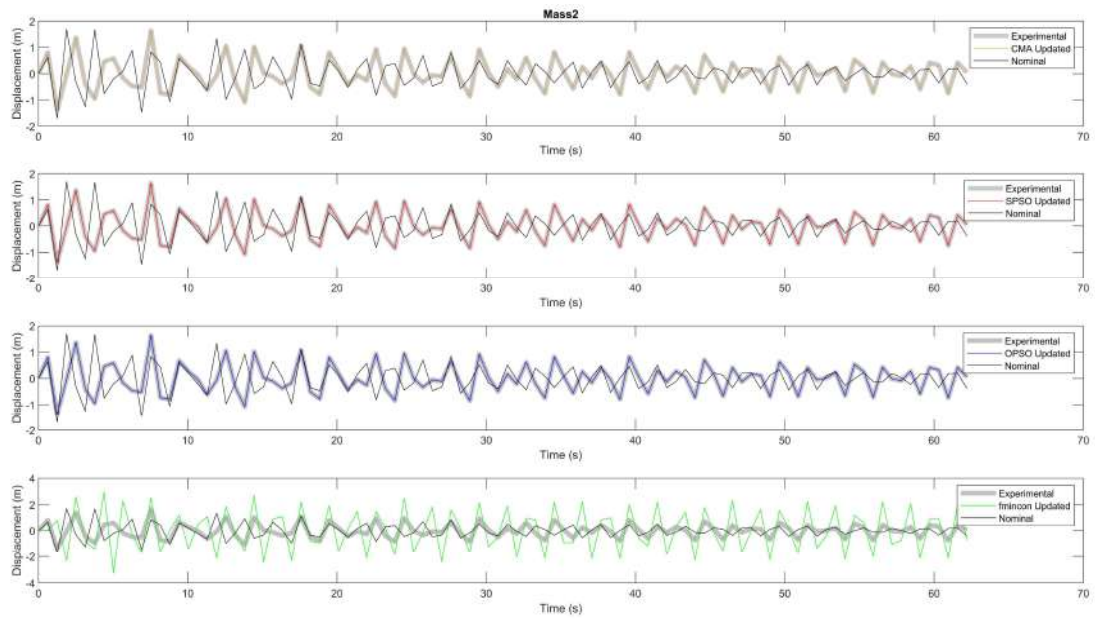
Παρακάτω παρουσιάζεται η απόκριση των μαζών για την διέγερση που περιγράψαμε.



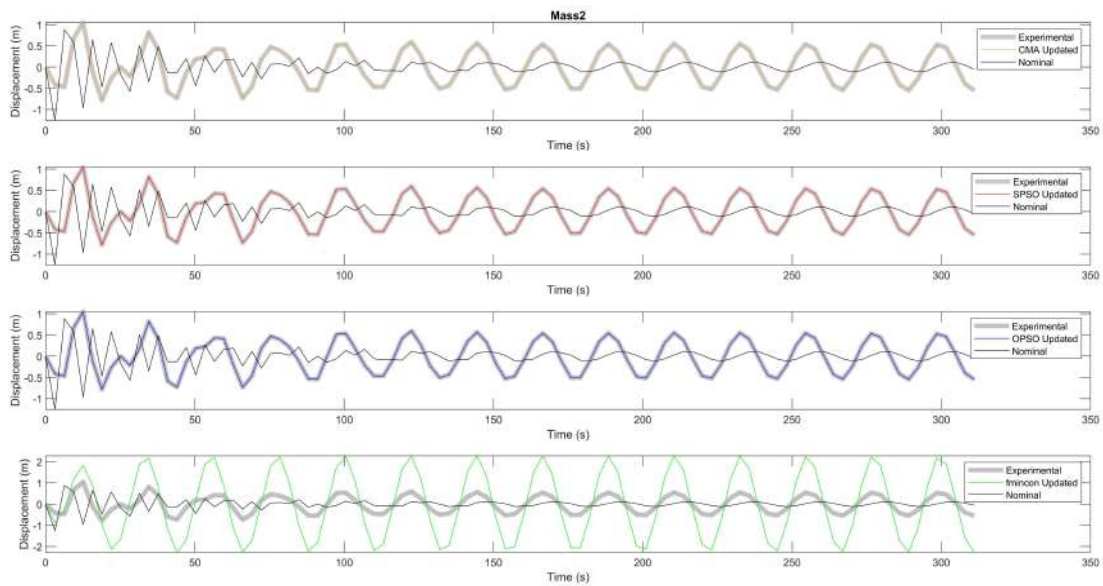
Εικόνα 118. Απόκριση μάζας 1, για την πρώτη διέγερση της Περίπτωσης 2



Εικόνα 119. Απόκριση μάζας 1, για την δεύτερη διέγερση της Περίπτωσης 2



Εικόνα 120. Απόκριση μάζας 2, για την πρώτη διέγερση της Περίπτωσης 2



Εικόνα 121. Απόκριση μάζας 2, για την δεύτερη διέγερση της Περίπτωσης 2

Συμπεράσματα

Η μέθοδος ενημέρωσης μοντέλου είναι μια αρκετά ισχυρή μέθοδος για την αναγνώριση παραμέτρων και πολλές είναι οι τεχνικές και οι αλγόριθμοι που χρησιμοποιούνται για την υλοποίηση της μεθόδου. Γενικά δεν υπάρχει κάποιος αλγόριθμος ο οποίος μπορεί να χαρακτηριστεί ως καθολικά βέλτιστος για την μέθοδο ενημέρωσης μοντέλου. Η απόδοση των αλγορίθμων εξαρτάται από την συγκεκριμένη μορφή του προβλήματος. Ακόμα και διαφορετική εκδοχή της ίδιας μεθόδου μπορεί να επιφέρει διαφορετικά αποτελέσματα. Η CMA-ES δικαίως θεωρείται μια state-of-the-art μέθοδος όπως επιβεβαιώθηκε και από τα αποτελέσματα της παρούσας εργασίας. Οι PSO αλγόριθμοι παρέχουν καλά αποτελέσματα και υλοποίηση τους είναι σχετικά εύκολη και οι διαδικασίες τους είναι σχετικά απλές. Συμπεραίνουμε επίσης πως στην ενημέρωση μοντέλου μεγάλη βαρύτητα έχει η ακριβής διαμόρφωση της αντικειμενικής συνάρτησης. Επιπλέον κρίσιμο σημείο είναι η εκλογή των κόμβων όπου πρόκειται να τοποθετηθούν οι αισθητήρες. Τέλος η εξέταση της ευαισθησίας της αντικειμενικής συνάρτησης στις παραμέτρους που πρόκειται να αναγνωριστούν είναι μια πρακτική που μπορεί να βοηθήσει στην καθοδήγηση της υλοποίησης της μεθόδου ενημέρωσης μοντέλου.

Βιβλιογραφικές Αναφορές

- [1] S. Ereiz, I. Duvnjak, J. Fernando and Jiménez-Alonso, “Review of finite element model updating methods for structural applications,” *Structures*, vol. 41, pp. 684-723, 2022,.
- [2] B. Hans-Georg, *The Theory of Evolution Strategies*, Berlin: Springer Berlin, Heidelberg, 2001.
- [3] T. Bäck, C. Foussette and P. Krause, *Natural Computing Series, Contemporary Evolution Strategies*, 2013: Springer Berlin, Heidelberg, 2013.
- [4] A. Arailopoulos, «Optimal Finite Element Modeling of Mechanical Systems with Linear and Nonlinear Characteristics using Numerical and Experimental Methods,» *Ph.D. dissertation*. Department of Mechanical Engineering University of Western Macedonia, 2019.
- [5] M. J. Reddy and D. N. Kumar, “Invited Review. Evolutionary algorithms, swarm intelligence methods, and their applications in water resources engineering: a state-of-the-art review,” *H2Open Journal*, vol. 3, no. 1, 2020.
- [6] R. K. Ursem, «Models for Evolutionary Algorithms and Their Applications in System Identification and Control Optimization» *Ph.D. dissertation, Department of Computer Science, University of Aarhus, Denmark, 2003.*
- [7] Harvard University, CSElab (Computational Science & Engineering Laboratory), «Derandomized Evolution Strategies and Local Learning Algorithms» <https://cse-lab.seas.harvard.edu/research/numerics/derandomized-evolution-strategies-and-local-learning-algorithms>.
- [8] L. YiFei, C. MaoSen, T. N. Hoa et. al., «Metamodel-assisted hybrid optimization strategy for model updating using vibration response data,» *Advances in Engineering Software*, vol. 185, 2023.
- [9] T. M. Vala et. al., «Revisiting the performance of evolutionary algorithms”, *Expert Systems with Applications*, vol. 1, 2021.
- [10] H. Beyer and Schwefel, «Evolution strategies – A comprehensive introduction,» *Natural Computing*, vol. 1, pp. 3-52, 2002.
- [11] N. Hansen, D. V. Arnold και A. Auger, *Evolution Strategies*, J. Kacprzyk and W. Pedrycz, edit., Berlin, Germany: Springer Handbook of Computational Intelligence, 2015.
- [12] «Fitness proportionate selection,» Wikipedia, [Ηλεκτρονικό]. Available: https://en.wikipedia.org/wiki/Fitness_proportionate_selection. [Πρόσβαση 6 Mar. 2024].

- [13] H. Kesavan και J. Kapur, «The generalized maximum entropy principle,» *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, pp. 1042-1052, 1989.
- [14] T. H. W. Bäck, A. V. Kononova, B. v. Stein, H. Wang, K. A. Antonov, R. T. Kalkreuth, J. d. Nobel, D. Vermetten, R. d. Winter and F. Ye, «Evolutionary Algorithms for Parameter Optimization-Thirty Years Later,» *Evol Comput*, vol. 31, pp. 81-122, 2023.
- [15] T. Bäck, C. Foussette and P. Krause, *Contemporary Evolution Strategies*, Berlin, Germany: Springer Berlin, Heidelberg, 2013.
- [16] A. Ostermeier, A. Gawelczyk and N. Hansen, «A Derandomized Approach to Self-Adaptation of Evolution Strategies,» *Evolutionary Computation*, vol. 2, no. 4, pp. 369-380, 1994.
- [17] A. Ostermeier, A. Gawelczyk and N. Hansen, «Step-Size Adaption Based on Non-Local Use of Selection Information.» in *Parallel Problem Solving from Nature - PPSN III, International Conference on Evolutionary Computation. The Third Conference on Parallel Problem Solving from Nature*, Jerusalem, Israel., 1994.
- [18] N. Hansen and A. Ostermeier, «Completely Derandomized Self-Adaptation in Evolution Strategies,» *Evolutionary Computation*, vol. 9, no. 2, pp. 159-195, 2001.
- [19] N. Hansen, «Invariance, Self-Adaptation and Correlated Mutations in Evolution Strategies,» in *Parallel Problem Solving from Nature-PPSN VI, 6th International Conference*, Paris, France, 2000.
- [20] N. Hansen, «The CMA Evolution Strategy: A Tutorial» 2005.
- [21] J. French και B.M. Ahmed, «The challenge of biomimetic design for carbon-neutral buildings using termite engineering,» *Insect Science*, vol. 17, 2010.
- [22] James Kennedy, R. Eberhart and Y. Shi, *Swarm Intelligence*, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc, 2001.
- [23] J. Kennedy και R. Eberhart, «Particle swarm optimization,» in *Proceedings of ICNN'95 - International Conference on Neural Networks, Perth, WA, Australia, 1995, .*
- [24] «How bird flocks are like liquid helium,» science.org, [Ηλεκτρονικό]. Available: <https://www.science.org/content/article/how-bird-flocks-are-liquid-helium>.
- [25] K. Parsopoulos, «Particle Swarm Method,» in *Handbook of Heuristics*, Springer, Cham., 2018.
- [26] D. Wang, D. Tan and L. Liu, «Particle swarm optimization algorithm: an overview,» *Soft Comput*, vol. 22, (2018).
- [27] A. P. Engelbrecht, *Computational Intelligence: An Introduction*, Wiley Publishing, 2007.
- [28] F. v. d. Bergh, «An Analysis of Particle Swarm Optimizers,» *Ph.D. dissertation Faculty of Natural and Agricultural Science University of Pretoria, Pretoria, South Africa, 2001*.

- [29] T. Strohmer, «Surprises in high dimensions,» [*Class Notes & Handouts*]. *Department of Mathematics, University of California, 180BigData*.
- [30] S. Helwig and R. Wanka, «Particle Swarm Optimization in High-Dimensional Bounded Search Spaces» in *2007 IEEE Swarm Intelligence Symposium*, Honolulu, HI, USA, 2007.
- [31] M. Clerc, «Back to random topology» 2007.
- [32] Y. Zhang, S. Wang and G. Ji «A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications» *Mathematical Problems in Engineering*, 2015.
- [33] Z. Z.-H. Zhang, Y. J Li και Y.-H. Shi, «Orthogonal learning particle swarm optimization,» *IEEE Transactions on Evolutionary Computation*, vol. 99, 2010.
- [34] G. Zhang και Y. Li, «Orthogonal Experimental Design Method Used in Particle Swarm Optimization for Multimodal Problems,» σε *Sixth International Conference on Advanced Computational Intelligence* , Hangzhou, China , 2013.
- [35] N. Hansen, «CMA-ES 3.61.beta. »
http://www.cmap.polytechnique.fr/~nikolaus.hansen/cmaes_inmatlab.html
- [36] M. Clerc, «Standard Particle Swarm Optimisation» 2012.
- [37] N. Hansen, R. Ros, N. Mauny, M. Schoenauer και A. Auger, «Impacts of invariance in search: When CMA-ES and PSO face ill-conditioned and non-separable problems» *Applied Soft Computing*, vol. 11, no. 8, 2011.
- [38] K. Worden και G. R. Tomlinson, *NONLINEARITY IN STRUCTURAL DYNAMICS : Detection, Identification and Modelling*, Institute of Physics Publishing Bristol and Philadelphia, 2001.
- [39] F. Pianosi, F. Sarrazin και T. Wagener, «A Matlab toolbox for Global Sensitivity Analysis,» *Environmental Modelling & Software*, τόμ. 70, pp. 80-85, 2015.
- [40] R. Martí, P. Pardalos και M. Resende, *Handbook of Heuristics*, Springer, Cham., 2018.