



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**  
**ΣΧΟΛΗ ΧΗΜΙΚΩΝ ΜΗΧΑΝΙΚΩΝ**  
ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ  
ΣΠΟΥΔΩΝ «ΥΠΟΛΟΓΙΣΤΙΚΗ ΜΗΧΑΝΙΚΗ»

**ΕΞΕΛΙΚΤΙΚΟΣ ΑΛΓΟΡΙΘΜΟΣ ΜΕ ΠΡΟΣΑΡΜΟΓΗ ΤΟΥ  
ΜΗΤΡΩΟΥ ΣΥΝΔΙΑΚΥΜΑΝΣΗΣ (EA-CMA) ΚΑΙ ΧΡΗΣΗ  
ΣΕ ΠΡΟΒΛΗΜΑΤΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ**

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**ΔΗΜΗΤΡΙΟΣ ΛΑΚΑΦΩΣΗΣ**

**ΕΠΙΒΛΕΠΩΝ:**

ΚΥΡΙΑΚΟΣ Χ. ΓΙΑΝΝΑΚΟΓΛΟΥ, ΚΑΘΗΓΗΤΗΣ ΕΜΠ  
ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΤΟΜΕΑΣ ΡΕΥΣΤΩΝ  
ΜΟΝΑΔΑ ΠΑΡΑΛΛΗΛΗΣ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΡΕΥΣΤΟΜΗΧΑΝΙΚΗΣ &  
ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

ΑΘΗΝΑ, ΣΕΠΤΕΜΒΡΙΟΣ 2021

## ΠΡΟΛΟΓΟΣ

Η μεταπτυχιακή αυτή εργασία συντάχθηκε κατά το ακαδημαϊκό έτος 2020 – 2021 στο πλαίσιο του Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών «Υπολογιστική Μηχανική» του Εθνικού Μετσόβιου Πολυτεχνείου.

Εκπονήθηκε υπό την επίβλεψη του Καθηγητή Κυριάκου Χ. Γιαννάκογλου, τον οποίο θα ήθελα να ευχαριστήσω θερμά για ποικίλους λόγους. Θα αναφέρω μόνο τον κατά την άποψή μου σημαντικότερο: είναι υπεύθυνος για τη γνωριμία μου με τη βελτιστοποίηση, ένα αντικείμενο που μου ήταν παντελώς άγνωστο πριν το μάθημά του «Μέθοδοι Αιτιοκρατικής & Στοχαστικής Βελτιστοποίησης και Εφαρμογές» που διδάσκεται κατά το 2<sup>ο</sup> εξάμηνο του μεταπτυχιακού προγράμματος σπουδών. Παρά τις δυσκολίες που επέφερε ο ιός COVID-19 με την τηλεκπαίδευση στην παράδοση των μαθημάτων, το ενδιαφέρον και η μεταδοτικότητα του κύριου Γιαννάκογλου με έκαναν να αγαπήσω το αντικείμενο, με τελικό αποτέλεσμα τη συνεργασία μας για την εκπόνηση της μεταπτυχιακής μου εργασίας. Ευελπιστώ μέσω αυτής να μπορέσω να ανταποδώσω μερικώς για όλες τις γνώσεις που έλαβα, κάνοντας μια χρήσιμη, τελική συμβολή προτού αποχωρήσω από το μεταπτυχιακό.

Εκφράζω ένα εξίσου μεγάλο ευχαριστώ στη Δρ. Βαρβάρα Γ. Ασούτη για την παροχή υλικού που χρησιμοποιήθηκε τόσο στην τελική μορφή του προγράμματος H/Y με τη μέθοδο CMA, όσο και για την επίλυση του προβλήματος CFD που πραγματοποιήθηκε στο πλαίσιο της εργασίας. Τέλος, θα ήθελα να συγχαρώ τα προαναφερθέντα πρόσωπα, αλλά και την υπόλοιπη ερευνητική ομάδα της Μονάδας Παράλληλης Υπολογιστικής Ρευστομηχανικής & Βελτιστοποίησης που δεν γνώρισα και ασχολείται με την ανάπτυξη του λογισμικού EASY για την εξαιρετική δουλειά που έχουν αφιερώσει στο προϊόν αυτό. Θαρρώ ότι η έλξη μου στο αντικείμενο και ο λόγος που αποφάσισα να ασχοληθώ με αυτό οφείλεται σε μεγάλο βαθμό στη λειτουργικότητα και ευκολία χρήσης του EASY, το οποίο υπήρξε και η πρώτη μου επαφή με τους εξελικτικούς αλγόριθμους.



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**  
**ΣΧΟΛΗ ΧΗΜΙΚΩΝ ΜΗΧΑΝΙΚΩΝ**  
ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ  
ΣΠΟΥΔΩΝ «ΥΠΟΛΟΓΙΣΤΙΚΗ ΜΗΧΑΝΙΚΗ»

## **ΕΞΕΛΙΚΤΙΚΟΣ ΑΛΓΟΡΙΘΜΟΣ ΜΕ ΠΡΟΣΑΡΜΟΓΗ ΤΟΥ ΜΗΤΡΩΟΥ ΣΥΝΔΙΑΚΥΜΑΝΣΗΣ (EA-CMA) ΚΑΙ ΧΡΗΣΗ ΣΕ ΠΡΟΒΛΗΜΑΤΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ**

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**ΔΗΜΗΤΡΙΟΣ ΛΑΚΑΦΩΣΗΣ**

**ΕΠΙΒΛΕΠΩΝ:**

**ΚΥΡΙΑΚΟΣ Χ. ΓΙΑΝΝΑΚΟΓΛΟΥ, ΚΑΘΗΓΗΤΗΣ ΕΜΠ**

ΑΘΗΝΑ, ΣΕΠΤΕΜΒΡΙΟΣ 2021

Η μεταπτυχιακή αυτή εργασία έχει ως σκοπό τη δημιουργία ενός προγράμματος υπολογιστή ικανού να εφαρμόσει και να πιστοποιήσει έναν τροποποιημένο Εξελικτικό Αλγόριθμο (EA) CMA. Ο απώτερος στόχος της εργασίας είναι να επιλυθούν δοκιμαστικά προβλήματα βελτιστοποίησης με τον νέο αλγόριθμο και το λογισμικό EASY της Μονάδας Παράλληλης Υπολογιστικής Ρευστομηχανικής & Βελτιστοποίησης του Εθνικού Μετσόβιου Πολυτεχνείου ώστε να φανεί εάν είναι σκόπιμο το τελευταίο να ενσωματώσει στοιχεία της μεθόδου CMA. Τα πρώτα δύο κεφάλαια συνοψίζουν το αντικείμενο της βελτιστοποίησης με ιδιαίτερη έμφαση στην κατηγορία των Εξελικτικών Αλγορίθμων όπου ανήκει και ο CMA. Θίγονται έννοιες απαραίτητες προκειμένου το υπόλοιπο κείμενο να μπορεί να κατανοηθεί από κάθε αναγνώστη ανεξαρτήτως υποβάθρου. Στο τρίτο κεφάλαιο γίνεται μια εξίσου σύντομη επισκόπηση των βασικών ιδιοτήτων της κανονικής κατανομής που αποτελεί και τη βάση των διεργασιών του CMA. Στο τέταρτο κεφάλαιο αναλύεται ο αλγόριθμος CMA και επισημαίνονται οι αλλαγές που έχουν γίνει επί αυτού. Στη συνέχεια, επιλύονται τρία κλασικά προβλήματα ψευδομηχανικής, καθώς και μια πραγματική εφαρμογή Υπολογιστικής Ρευστοδυναμικής (CFD), η οποία αφορά το σχεδιασμό καμπής 90 μοιρών σε κλειστό αγωγό. Όλα τα προβλήματα λύνονται με τον CMA και με το λογισμικό EASY (γενικευμένος EA) και παρουσιάζονται τα διαγράμματα σύγκλισης του καθενός. Τέλος, στο έβδομο κεφάλαιο γίνεται μια γενική ανακεφαλαίωση και διατυπώνονται κάποια πρώτα συμπεράσματα σχετικά με την αποτελεσματικότητα του τροποποιημένου CMA έναντι του γενικευμένου EA.



**NATIONAL TECHNICAL UNIVERSITY OF ATHENS**  
**SCHOOL OF CHEMICAL ENGINEERING**  
JOINT POSTGRADUATE COURSE  
«COMPUTATIONAL MECHANICS»

**EVOLUTIONARY ALGORITHM WITH COVARIANCE  
MATRIX ADAPTATION (EA-CMA) AND USE ON  
OPTIMIZATION PROBLEMS**

MASTER THESIS

**DIMITRIOS LAKAFOSIS**

**SUPERVISOR:**

**KYRIAKOS C. GIANNAKOGLU, PROFESSOR NTUA**

**ATHENS, SEPTEMBER 2021**

This master thesis has the goal of creating a computer program capable of implementing a modified CMA Evolutionary Algorithm (EA). The ultimate goal of the thesis is to solve test optimization problems with both the new algorithm and the EASY software of the Parallel CFD & Optimization Unit of the National Technical University of Athens in order to show whether the latter should employ some elements of the former method. The first two chapters summarize the subject of optimization with special focus on the category of Evolutionary Algorithms where CMA belongs. Necessary concepts are touched upon so that the rest of the text can be comprehended by readers. In the third chapter there's a similarly short review of the basic properties of the normal distribution which is the basis of all CMA's processes. In the fourth chapter, the CMA algorithm is analyzed and the modifications that have been implemented are highlighted. Afterwards three benchmark pseudo-engineering optimization design problems are solved, as well as a real CFD application, which involves the design of a 90 degree curved pipe. All problems are solved by both CMA and EASY (generalized EA) and the corresponding convergence graphs are presented. Finally, in the seventh chapter a general recap is done and the conclusions regarding the effectiveness of the modified CMA versus the generalized EA are formulated.

## ΑΚΡΩΝΥΜΙΑ

CSA	:	Cumulative Step-Size Adaptation
DMAEA	:	Distributed Metamodel-Assisted Evolutionary Algorithm
EASY	:	Evolutionary Algorithm SYstem
MAEA	:	Metamodel-Assisted Evolutionary Algorithm
MACMA	:	Metamodel-Assisted Covariance Matrix Adaptation
SPEA	:	Strength Pareto Evolutionary Algorithm
CFD	:	Computational Fluid Dynamics Υπολογιστική Ρευστοδυναμική
CMA	:	Covariance Matrix Adaptation Προσαρμογή του Μητρώου Συνδιακύμανσης
EA	:	Evolutionary Algorithm Εξελικτικός Αλγόριθμος
ES	:	Evolution Strategy Στρατηγική Εξέλιξης
GA	:	Genetic Algorithm Γενετικός Αλγόριθμος
LCPE	:	Low-Cost Pre-Evaluation Προ-αξιολόγηση Χαμηλού Κόστους
MOO	:	Multi-Objective Optimization Βελτιστοποίηση Πολλών Στόχων
SOO	:	Single Objective Optimization Βελτιστοποίηση Ενός Στόχου
PDF	:	Probability Density Function Συνάρτηση Πυκνότητας Πιθανότητας
RBFN	:	Radial Basis Function Networks Δίκτυα Ακτινικών Συναρτήσεων Βάσης

# ΠΕΡΙΕΧΟΜΕΝΑ

	<u>Σελίδα</u>
<b>1. ΕΙΣΑΓΩΓΗ</b>	<b>1</b>
<b>2. ΕΞΕΛΙΚΤΙΚΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ</b>	<b>3</b>
2.1 ΓΕΝΙΚΑ	3
2.2 ΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ	5
<b>2.2.1 Επιλογή Γονέων</b>	<b>6</b>
<b>2.2.2 Δημιουργία Απόγονων</b>	<b>9</b>
2.3 ΣΥΝΘΕΤΑ ΠΡΟΒΛΗΜΑΤΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ	13
<b>2.3.1 Προβλήματα με Περιορισμούς</b>	<b>13</b>
<b>2.3.2 Προβλήματα με Πολλούς Στόχους</b>	<b>16</b>
<b>2.3.3 Προβλήματα με Υψηλό Κόστος</b>	<b>20</b>
<b>3. ΔΕΙΓΜΑΤΟΛΗΨΙΑ ΜΕΣΩ ΚΑΝΟΝΙΚΗΣ ΚΑΤΑΝΟΜΗΣ</b>	<b>22</b>
3.1 ΤΥΠΟΠΟΙΗΜΕΝΗ ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ	22
3.2 ΠΟΛΥΔΙΑΣΤΑΤΗ ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ	25
<b>4. Ο ΤΡΟΠΟΠΟΙΗΜΕΝΟΣ ΑΛΓΟΡΙΘΜΟΣ CMA</b>	<b>28</b>
4.1 ΔΕΙΓΜΑΤΟΛΗΨΙΑ & ΑΞΙΟΛΟΓΗΣΗ	28
4.2 ΥΠΟΛΟΓΙΣΜΟΣ ΤΗΣ ΜΕΣΗΣ ΤΙΜΗΣ	30
4.3 ΠΡΟΣΑΡΜΟΓΗ ΤΟΥ ΜΗΤΡΩΟΥ ΣΥΝΔΙΑΚΥΜΑΝΣΗΣ	33
<b>4.3.1 Μονοπάτι Εξέλιξης</b>	<b>33</b>
<b>4.3.2 Υπολογισμός του Μητρώου Συνδιακύμανσης</b>	<b>36</b>
<b>4.3.3 Υπολογισμός του Μεγέθους Βήματος</b>	<b>40</b>

<b>5. ΣΧΕΔΙΑΣΤΙΚΑ ΠΡΟΒΛΗΜΑΤΑ ΨΕΥΔΟΜΗΧΑΝΙΚΗΣ</b>	<b>43</b>
5.1 ΣΧΕΔΙΑΣΜΟΣ ΔΙΚΤΥΩΜΑΤΟΣ ΤΡΙΩΝ ΜΕΛΩΝ	44
5.2 ΣΧΕΔΙΑΣΜΟΣ ΣΥΓΚΟΛΛΗΤΗΣ ΔΟΚΟΥ	47
5.3 ΣΧΕΔΙΑΣΜΟΣ ΜΕΙΩΤΗΡΑ ΣΤΡΟΦΩΝ	53
<b>6. ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΚΛΕΙΣΤΟΥ ΑΓΩΓΟΥ ΜΕ ΚΑΜΠΗ 90°</b>	<b>60</b>
6.1 ΕΠΙΛΥΣΗ ΡΟΗΣ ΣΕ ΚΛΕΙΣΤΟ ΑΓΩΓΟ ΜΕ ΚΑΜΠΗ 90°	60
6.2 ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΤΟΥ ΣΧΕΔΙΑΣΜΟΥ ΚΑΜΠΗΣ 90°	63
<b>7. ΣΥΜΠΕΡΑΣΜΑΤΑ</b>	<b>69</b>
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ</b>	<b>74</b>

# ΚΕΦΑΛΑΙΟ 1

## ΕΙΣΑΓΩΓΗ

Με τον όρο βελτιστοποίηση νοείται η διαδικασία αναζήτησης της καταλληλότερης τιμής για μία ή περισσότερες δεδομένες συναρτήσεις, οι οποίες στο σύνολό τους ονομάζονται συναρτήσεις-στόχοι ή συναρτήσεις σχεδιασμού <sup>[1]</sup>. Στην απλή περίπτωση της μίας συνάρτησης-στόχου (Single Objective Optimization: SOO) <sup>[2]</sup>, η βέλτιστη τιμή εξορισμού βρίσκεται στο ολικό ακρότατο που αντιστοιχεί στις απαιτήσεις του σχεδιασμού: στο ολικό μέγιστο για ένα πρόβλημα μεγιστοποίησης και στο ολικό ελάχιστο για ένα πρόβλημα ελαχιστοποίησης. Η φύση του εκάστοτε προβλήματος συνήθως καθιστά αυτονόητο το είδος της αναζήτησης. Μια συνάρτηση κόστους πάντα πρέπει να ελαχιστοποιείται, ενώ αντιθέτως μια συνάρτηση καταλληλότητας να μεγιστοποιείται.

Στην πραγματικότητα όμως κανένα πρόβλημα βελτιστοποίησης με πρακτική εφαρμογή δεν είναι τόσο απλό στην επίλυσή του, αφού ακόμα και προβλήματα SOO μπορεί να παρουσιάζουν τρομερές περιπλοκές. Η αναζήτηση ενός ακρότατου στα μαθηματικά συνδέεται άρρηκτα με την τιμή της παραγώγου. Όταν η συνάρτηση-στόχος αποτελείται από δεκάδες ή εκατοντάδες μεταβλητές σχεδιασμού, είναι εύλογο ότι ο υπολογισμός της παραγώγου αποτελεί μια μη-τετριμμένη, αν όχι ρεαλιστικά αδύνατη διαδικασία. Επιπλέον, μπορεί η συνάρτηση να παρουσιάζει πολλαπλά τοπικά ακρότατα και ως επακόλουθο να δυσκολεύει δραματικά η εύρεση του ολικού ακρότατου. Κατ' επέκταση, σε προβλήματα με πολλές συναρτήσεις-στόχους (Multi-Objective Optimization: MOO) <sup>[3],[4]</sup> μπορεί ένα σετ των μεταβλητών σχεδιασμού να ελαχιστοποιεί μια συνάρτηση αλλά ταυτόχρονα να μεγιστοποιεί μια άλλη, οπότε χρειάζεται μια γκάμα προτεινόμενων λύσεων που να καλύπτουν ολόκληρο το εύρος του χωρίου σχεδιασμού, ώστε να μπορεί ο χρήστης να επιλέξει την πλέον καταλληλότερη για τον ίδιο. Τέλος, σε πραγματικά προβλήματα βελτιστοποίησης υπάρχουν επιπλέον και συναρτήσεις περιορισμών, συνήθως ανισότητες που πρέπει να πληρούνται από τη λύση προκειμένου να θεωρηθεί αποδεκτή.

Οι παραπάνω δυσκολίες επόμενο είναι να οδηγήσουν στην ανάπτυξη διάφορων μεθόδων βελτιστοποίησης, οι οποίες μπορούν να χωριστούν σε δύο ευρείς κατηγορίες: τις στοχαστικές (stochastic) και τις αιτιοκρατικές (deterministic) <sup>[2]</sup>. Οι στοχαστικές μέθοδοι αναζητούν τη βέλτιστη λύση ενός προβλήματος με ένα κράμα κυρίως τυχηματικών (randomized) αλλά και τυχαίων (random) διαδικασιών, ενώ οι αιτιοκρατικές αξιοποιούν



την ακριβή ή προσεγγιστική τιμή της παραγώγου στην τρέχουσα λύση ώστε να οδηγηθούν τελικά σε κάποιο ακρότατο. Η βασικότερη διαφορά μεταξύ των μεθόδων είναι ήδη εμφανής: οι στοχαστικές μέθοδοι δεν απαιτούν τον υπολογισμό ή την προσέγγιση της τιμής της παραγώγου προκειμένου να λειτουργήσουν. Αυτό τις κάνει πιο εύχρηστες, με την έννοια ότι ο χρήστης αρκεί να ρυθμίσει τον αλγόριθμο, να τον τροφοδοτήσει με τη συνάρτηση-στόχο και η βελτιστοποίηση γίνεται αυτόματα και χωρίς να χρειάζεται καμία απολύτως επέμβαση στον κώδικα που υλοποιεί τη μέθοδο. Συνάμα, ο αλγόριθμος δεν χρειάζεται να γνωρίζει ποια είναι η συνάρτηση-στόχος, αρκεί να μπορεί να τη χρησιμοποιήσει για την αξιολόγηση των υποψήφιων λύσεων που αυτόματα παράγει. Η βελτιστοποίηση αυτής της μορφής θυμίζει τη λογική του μαύρου κουτιού και για αυτό ονομάζεται black-box optimization <sup>[5]</sup>.

Πέραν αυτών των χαρακτηριστικών τους, οι στοχαστικές μέθοδοι μπορούν να εντοπίσουν πάντα το ολικό ακρότατο άμα ρυθμιστούν κατάλληλα, ενώ οι αιτιοκρατικές εγκυμονούν τον κίνδυνο να εγκλωβιστούν σε κάποιο τοπικό ακρότατο ανάλογα τη δοθείσα αρχικοποίηση. Όμως το μεγάλο πλεονέκτημα που έχει μια αιτιοκρατική μέθοδος έναντι μιας στοχαστικής είναι η ταχύτητά της, πράγμα που οφείλεται στ' ότι η έρευνα νέων λύσεων καθοδηγείται από την παράγωγο. Όσο πιο πολυδιάστατο είναι μάλιστα το πρόβλημα, τόσο περισσότερο ενισχύεται η διαφορά αυτή.

Στο πλαίσιο της παρούσας εργασίας γίνεται λόγος αποκλειστικά για στοχαστικές μεθόδους βελτιστοποίησης, συγκεκριμένα για την ευρύτετη κατηγορία των Εξελικτικών Αλγορίθμων (Evolutionary Algorithms: EA) <sup>[2], [6]</sup> στις οποίες ανήκει και η μέθοδος της προσαρμογής του μητρώου συνδιακύμανσης (Covariance Matrix Adaptation: CMA) <sup>[5], [7]</sup>. Είναι σημαντικό προτού παρουσιαστεί η μέθοδος ο αναγνώστης να έχει το απαραίτητο υπόβαθρο προκειμένου να την κατανοήσει, πράγμα που εκτός από την υποβόσκουσα μαθηματική θεωρία συμπεριλαμβάνει το σκεπτικό και τη φιλοσοφία που διέπει τους αλγόριθμους αυτούς, οι οποίοι ουσιαστικά προσομοιώνουν φυσικούς μηχανισμούς που προέρχονται από το επιστημονικό πεδίο της βιολογίας. Αφού καλυφθούν οι βασικές αυτές έννοιες, θα ακολουθήσουν εφαρμογές ενός τροποποιημένου CMA ώστε να φανεί αν ο νέος αλγόριθμος μπορεί να ενισχύσει με νέα στοιχεία τον γενικευμένο EA του λογισμικού EASY <sup>[8]</sup>.

## ΚΕΦΑΛΑΙΟ 2

### ΕΞΕΛΙΚΤΙΚΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ

Ξεκινώντας κρίνεται αναγκαίο να αναλυθεί ένα βασικό πρόβλημα βελτιστοποίησης προκειμένου ο αναγνώστης να εξοικειωθεί με τα κύρια χαρακτηριστικά του και τις δυσκολίες που εμπεριέχονται σε αυτό. Στη συνέχεια, ορίζεται ένας γενικευμένος ΕΑ και παρουσιάζεται ο τρόπος λειτουργίας του, ενώ όπου χρειάζεται γίνεται αναφορά και σε συγκεκριμένα είδη ΕΑ. Τέλος, τίγονται πιο σύνθετα προβλήματα βελτιστοποίησης και το πώς αυτά μπορούν να αντιμετωπιστούν αποτελεσματικά σε ένα περιβάλλον ΕΑ. Επισημαίνεται ότι σκοπός εδώ είναι μια σύντομη και κατά το δυνατόν περιεκτική περίληψη των παραπάνω αντικειμένων, καθώς θα μπορούσαν από μόνα τους να καλύψουν ολόκληρη την έκταση της εργασίας χωρίς να έχουν καλυφθεί στην ολότητά τους.

#### 2.1 ΓΕΝΙΚΑ

Οι κύριες διακρίσεις των προβλημάτων βελτιστοποίησης γίνονται ανάλογα με τον αριθμό των συναρτήσεων-στόχων, καθώς και με το αν υπάρχουν ή όχι περιορισμοί. Οι συναρτήσεις-στόχοι συμβολίζονται ως  $f_i$ , οι περιορισμοί που υπεισέρχονται και αυτοί συνήθως με τη μορφή συναρτήσεων συμβολίζονται ως  $g_i$ , ενώ το σύνολο των μεταβλητών από τις οποίες εξαρτώνται όλες οι συναρτήσεις-στόχοι και οι περιορισμοί ονομάζονται μεταβλητές σχεδιασμού και συμβολίζονται ως  $x_i$ . Έτσι, ένα πρόβλημα βελτιστοποίησης με  $n$  μεταβλητές σχεδιασμού έχει ως σκοπό την εύρεση ενός διανύσματος της μορφής:

$$\vec{X}^* = \left\{ \begin{array}{c} X_1 \\ X_2 \\ \dots \\ X_{n-1} \\ X_n \end{array} \right\}^* \quad (2.1)$$

όπου ο αστερίσκος υποδηλώνει το μοναδικό διάνυσμα  $\vec{X}$  που αποδίδει τη βέλτιστη τιμή σε όλες τις συναρτήσεις-στόχους ταυτόχρονα <sup>[1]</sup>.

Η έρευνα για κάθε μεταβλητή  $x_i$  γίνεται στο χωρίο που ορίζεται από τις τιμές  $x_i^{\min}$  και  $x_i^{\max}$ , δηλαδή είναι απαραίτητο να οριστεί μια ελάχιστη και μια μέγιστη τιμή για κάθε μεταβλητή σχεδιασμού. Το σύνολο των  $x_i^{\min}$  και  $x_i^{\max}$  για  $i = 1, 2, \dots, n$  ορίζει το χωρίο σχεδιασμού (design space) του προβλήματος. Η οριοθέτηση κάθε μεταβλητής σχεδιασμού

θα μπορούσε να θεωρηθεί ως επιβολή δύο περιορισμών της μορφής  $g_{2i-1}(x_i): x_i \geq x_i^{\min}$  και  $g_{2i}(x_i): x_i \leq x_i^{\max}$ , οπότε και θα χρειαζόντουσαν εκ των πραγμάτων τουλάχιστον  $2n$  περιορισμοί για κάθε πρόβλημα βελτιστοποίησης. Στην πράξη όμως αυτό δεν εφαρμόζεται ποτέ. Κανένας πρακτικός αλγόριθμος βελτιστοποίησης δεν θα πραγματοποιούσε την οριοθέτηση μέσω περιορισμών διότι μπορεί να επιβληθεί πολύ πιο εύκολα και αποτελεσματικά από το χρήστη ως είσοδος στο πρόγραμμα (user input).

Έστω λοιπόν ότι δίνεται ένα πρόβλημα ελαχιστοποίησης SOO με  $n$  οριοθετημένες μεταβλητές σχεδιασμού. Η συνάρτηση σχεδιασμού συμβολίζεται ως  $f(\vec{X})$ , οπότε αναζητείται το διάνυσμα  $\vec{X}^*$  που εξασφαλίζει την ελάχιστη τιμή  $f^* = f(\vec{X}^*)$ . Επισημαίνεται ότι στο πλαίσιο της black-box optimization υπάρχει ένα λογισμικό αξιολόγησης, το οποίο μπορεί να τροφοδοτείται συνεχώς με διανύσματα  $\vec{X}$  και να επιστρέφει την τιμή  $f(\vec{X})$ . Αυτό μπορεί να επαναληφθεί όσες φορές ο χρήστης επιθυμεί χωρίς όμως να έχει πρόσβαση στο λογισμικό, δηλαδή χωρίς να ξέρει απολύτως τίποτα για τη συνάρτηση  $f$  πέραν του συνολικού αριθμού και των ορίων των μεταβλητών σχεδιασμού της [5].

Αρχίζουν έτσι ήδη να διακρίνονται οι λόγοι που είναι σημαντικές οι μέθοδοι βελτιστοποίησης. Παρόλο που το πρόβλημα είναι απλό στην κατανόηση του δεν είναι καθόλου εύκολο να βρεθεί το βέλτιστο  $\vec{X}^*$ , αφού θεωρητικά υπάρχουν άπειρα υποψήφια διανύσματα  $\vec{X}$  στο χωρίο σχεδιασμού που πρέπει να αξιολογηθούν. Επίσης στην περίπτωση που η  $f$  είναι μία ιδιαίτερα πολύπλοκη συνάρτηση και χρειάζεται μεγάλο υπολογιστικό κόστος ή και χρονικό διάστημα για να υπολογιστεί, τότε ο χρήστης δεν έχει την ευχέρεια να δοκιμάζει αδιάκοπα πιθανές λύσεις.

Υπάρχουν δύο κύριες συλλογιστικές που μπορούν να ακολουθηθούν για να λυθεί το πρόβλημα. Πρώτον, είναι δυνατόν να οριστεί και να αξιολογηθεί ένα αυθαίρετο  $\vec{X}$  και εφαρμόζοντας σε αυτό αλληπάλληλες μετατροπές να βελτιώνεται συνεχώς η  $f(\vec{X})$  έως ότου να μην παρατηρείται σημαντική αλλαγή στην τιμή της. Εναλλακτικά μπορεί να γίνει μία δειγματοληψία (sampling) τυχαίων διανυσμάτων  $\vec{X}_i$  από ολόκληρο το χωρίο σχεδιασμού. Αυτά αξιολογούνται και βάσει των προκυπτουσών τιμών  $f(\vec{X}_i)$  αλλάζουν κατάλληλα τα κριτήρια της επόμενης δειγματοληψίας, ενώ παράλληλα αποθηκεύεται το βέλτιστο  $\vec{X}^*$  από αυτά. Λαμβάνονται όσα δείγματα είναι επιθυμητό και όταν τα νέα δείγματα δεν εμπεριέχουν καλύτερες λύσεις τότε το τελευταίο αποθηκευμένο  $\vec{X}^*$  αποτελεί τη βέλτιστη λύση. Βάσει των παραπάνω μεθοδολογιών, οι μέθοδοι βελτιστοποίησης μπορούν να

χωριστούν σε ατομικές (individual-based) και πληθυσμιακές (population-based) αντίστοιχα <sup>[2]</sup>. Τονίζεται ότι αυτή η διάκριση είναι παντελώς ξεχωριστή από τη διάκριση σε αιτιοκρατικές και στοχαστικές μεθόδους που έγινε στο Κεφάλαιο 1 καθώς υπάρχουν εξίσου πολλές ατομικές και πληθυσμιακές στοχαστικές μέθοδοι βελτιστοποίησης, ενώ οι αιτιοκρατικές είναι κατά κύριο λόγο ατομικές.

## 2.2 ΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ

Οι ΕΑ αποτελούν ένα σύνολο αλγορίθμων που εδραιώθηκαν τη δεκαετία του 1960 και έχουν ως κοινό χαρακτηριστικό ότι είναι εμπνευσμένοι, όπως υποδηλώνει άλλωστε και το όνομά τους, από τη θεωρία της εξέλιξης <sup>[2]</sup>, <sup>[6]</sup>. Εξ ορισμού, η εξέλιξη αφορά ένα σύνολο ατόμων ή πληθυσμό και πραγματοποιείται με το πέρασ γενεών, επομένως εύλογα οι ΕΑ κατά πλειοψηφία αποτελούν πληθυσμιακές μεθόδους στοχαστικής βελτιστοποίησης, αφού επεξεργάζονται πληθυσμούς υποψήφιων λύσεων που συνεχώς εξελίσσονται με έναν «οργανωμένα τυχαίο» τρόπο. Δύο από τις κυριότερες κατηγορίες ΕΑ είναι οι Γενετικοί Αλγόριθμοι (Genetic Algorithms: GA) και οι Στρατηγικές Εξέλιξης (Evolution Strategies: ES). Όταν πρωτοεμφανίστηκαν οι αλγόριθμοι αυτοί είχαν την ειδοποιό διαφορά ότι η κωδικοποίηση των υποψήφιων λύσεων γινόταν σε δυαδική μορφή για GA και σε πραγματική για ES (binary και real coding αντίστοιχα) <sup>[2]</sup>. Ο διαχωρισμός αυτός πλέον είναι άτοπος διότι έχουν αναπτυχθεί πολλές υβριδικές μέθοδοι που λειτουργούν με οποιονδήποτε από τους δύο παραπάνω τρόπους κωδικοποίησης και είναι επιλογή του προγραμματιστή ή του χρήστη ποια εν τέλει θα χρησιμοποιηθεί. Παρόλα αυτά, οι διαφορές που παρουσιάζονται ακόμα και σήμερα σε ορισμένα σημεία του τρόπου λειτουργίας τους μπορούν να αποδοθούν στο παλιό αυτό σχίσμα.

Κατά τη γενική ορολογία των ΕΑ, οι  $\mu$  σε αριθμό τρέχουσες λύσεις που επιλέγονται για να δημιουργηθεί μια νέα γενιά ονομάζονται γονείς (parents). Με κατάλληλες διεργασίες από τους  $\mu$  γονείς κάθε γενιάς παράγονται  $\lambda$  νέες υποψήφιες λύσεις που αποτελούν την επόμενη προς αξιολόγηση γενιά και ονομάζονται απόγονοι (offspring). Αναφέρεται ότι σχεδόν πάντα ισχύει  $\mu \leq \lambda$ . Βάσει των παραπάνω, η λειτουργία ενός γενικευμένου ΕΑ μπορεί να συνοψιστεί στα εξής επαναλαμβανόμενα βήματα <sup>[2]</sup>:

### 1. Αρχικοποίηση (Initialization)

Δημιουργία ενός αρχικού πληθυσμού  $\lambda$  ατόμων, συνήθως με τυχαίο τρόπο.

## 2. Αξιολόγηση (Evaluation)

Αξιολόγηση κάθε ατόμου του πληθυσμού με χρήση του λογισμικού αξιολόγησης και αποθήκευση του τρέχοντος βέλτιστου.

## 3. Επιλογή Γονέων (Parent Selection)

Επιλογή των  $\mu$  γονέων σύμφωνα με τα αποτελέσματα της αξιολόγησης, αλλά και με στοιχεία τυχαιότητας.

## 4. Δημιουργία Απογόνων (Offspring Creation)

Οι  $\mu$  γονείς παράγουν  $\lambda$  απογόνους που αποτελούν τη νέα γενιά μέσω των εξελικτικών τελεστών διασταύρωσης (crossover) ή ανασυνδυασμού (recombination) και μετάλλαξης (mutation).

## 5. Κριτήριο Τερματισμού (Termination Criterion)

Ελέγχεται κάποιο κριτήριο τερματισμού. Αν αυτό πληρείται, τότε ο αλγόριθμος σταματάει και επιστρέφει τη βέλτιστη λύση. Αν όχι, επιστροφή στο βήμα 2.

Τονίζεται ότι τα παραπάνω δεν αποτελούν παρά μια απλή πρόγερση των βασικών αρχών του εξελικτικού λογισμού. Το αντικείμενο είναι πολύ ευέλικτο και δίνει την ελευθερία να γίνουν αναρίθμητες τροποποιήσεις και επεμβάσεις στην παραπάνω στοιχειώδη διαδικασία. Τα βήματα 1, 2 και 5 είναι κατά κύριο λόγο τυποποιημένα και εμπεριέχονται σε κάθε μέθοδο βελτιστοποίησης. Η πεμπτουσία των ΕΑ, αυτό που τους διαφοροποιεί από τις υπόλοιπες μεθόδους αλλά και τους διακρίνει μεταξύ τους, εντοπίζεται στα βήματα 3 και 4 τα οποία και θα αναλυθούν στις δύο επόμενες υποενότητες.

### 2.2.1 Επιλογή Γονέων

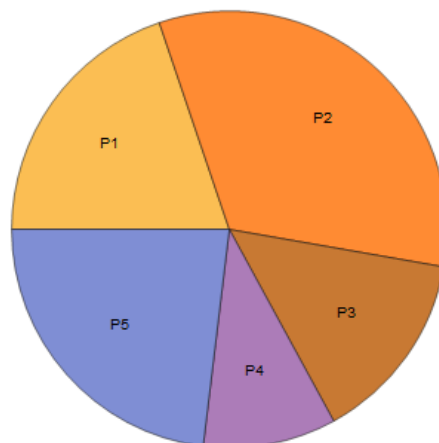
Σύμφωνα με τη θεωρία του Δαρβινισμού, τα καταλληλότερα άτομα ενός πληθυσμού έχουν την υψηλότερη πιθανότητα να αναπαραχθούν. Η αρχή αυτή ονομάζεται φυσική επιλογή ή επιβίωση του ισχυρότερου και οφείλει να προσομοιωθεί από μία εξελικτική μέθοδο. Γίνεται επομένως κατανοητό ότι η επιλογή γονέων αποτελεί θεμελιώδες χαρακτηριστικό των ΕΑ και έχει οριστεί ο συμβολισμός  $(\mu^+ \lambda)$ -ΕΑ προκειμένου να γίνεται απευθείας γνωστή στον αναγνώστη η τεχνική που εφαρμόζεται αλλά και οι αριθμοί γονέων και απογόνων <sup>[9]</sup>. Σε έναν  $(\mu, \lambda)$ -ΕΑ οι γονείς πεθαίνουν αμέσως μετά την αναπαραγωγή και δεν μπορούν να επιλεγούν ξανά για τη δημιουργία μελλοντικών γενιών, οπότε οι νέοι  $\mu$  γονείς επιλέγονται πάντα από τον τρέχοντα πληθυσμό μεγέθους  $\lambda$ . Αντιθέτως, ένας  $(\mu + \lambda)$ -ΕΑ αποθηκεύει ξεχωριστά τους τελευταίους γονείς και επιλέγει τους νέους από το σύνολο των  $\mu$  παλιών γονέων και του τρέχοντος πληθυσμού  $\lambda$ . Ο  $(\mu +$

λ)-EA λέγεται ότι έχει ελιτιστικό χαρακτήρα, πράγμα που του επιτρέπει να συγκλίνει σημαντικά πιο γρήγορα εν συγκρίσει με έναν (μ, λ)-EA, αλλά ταυτόχρονα караδοκεί πολύ υψηλότερος κίνδυνος εγκλωβισμού σε τοπικό ακρότατο λόγω της έλλειψης ποικιλίας χαρακτηριστικών στον πληθυσμό. Επισημαίνεται ότι ο ελιτισμός μπορεί να εφαρμοσθεί με ποικίλες μορφές και σε πολλά διαφορετικά σημεία στον αλγόριθμο αλλά, χάριν απλότητας κατά την παρουσίαση των μεθόδων επιλογής γονέων, ας γίνει η υπόθεση χρήσης ενός συχνότερου (μ, λ)-EA.

Ο πιο εύκολος και προφανής τρόπος για να πραγματοποιηθεί η επιλογή γονέων σε ένα πρόβλημα βελτιστοποίησης είναι η κατάταξη του τρέχοντος πληθυσμού σύμφωνα με την τιμή της συνάρτησης-στόχου που αντιστοιχεί σε κάθε υποψήφιο έτσι ώστε το πρώτο άτομο να είναι και το βέλτιστο. Τα μ πρώτα άτομα στην κατάταξη αποτελούν την ελίτ του πληθυσμού με τα καλύτερα χαρακτηριστικά, τα οποία και καλούνται να μεταλαμπαδεύσουν στην επόμενη γενιά. Ο τρόπος επιλογής αυτός ονομάζεται επιλογή με αποκοπή (truncation selection) και είναι αιτιοκρατικός καθώς δεν εμπεριέχει απολύτως κανένα στοχαστικό χαρακτηριστικό <sup>[9]</sup>. Οι ES χρησιμοποιούν συνήθως αυτήν τη μέθοδο επιλογής και επεξεργάζονται μικρούς πληθυσμούς, ενώ μάλιστα είναι πολύ σύνηθες να ισχύει  $\mu = 1$ . Ενδεικτικά αναφέρεται ότι ο πρώτος ES που δημοσιεύτηκε από τον Ingo Rechenberg το 1965 ήταν, σύμφωνα με το σημερινό συμβολισμό, ο  $(1 + 1)$ -ES <sup>[6]</sup>.

Οι GA αντιθέτως χρησιμοποιούν πιο περίπλοκες διαδικασίες επιλογής γονέων, καθώς και πολύ μεγαλύτερους πληθυσμούς που απαρτίζονται από δεκάδες άτομα. Το βασικότερο χαρακτηριστικό ενός GA είναι ο ίσος αριθμός γονέων και απογόνων, δηλαδή ισχύει  $\mu = \lambda$ . Η κλασικότερη μέθοδος είναι η επιλογή με τροχό ρουλέτας (roulette wheel selection) όπως περιγράφεται από τον David E. Goldberg <sup>[10]</sup>. Δημιουργείται μία ρουλέτα λ θέσεων ώστε να αντιστοιχεί μία θέση σε κάθε απόγονο, έχοντας μέγεθος ανάλογο της καταλληλότητάς του. Το αποτέλεσμα κάθε περιστροφής ορίζει ένα γονέα, οπότε η ρουλέτα θα χρειαστεί να γυρίσει μ συνολικά φορές ώστε να συμπληρωθούν όλοι οι γονείς. Αυτό σημαίνει ότι σαφώς τα καλύτερα άτομα θα λάβουν μεγαλύτερες πιθανότητες να επιλεγούν κατά τα γυρίσματα, αλλά και πως όλα τα άτομα του πληθυσμού έχουν κάποια πιθανότητα να αναπαραχθούν. Με αυτόν τον τρόπο μεταφέρονται χαρακτηριστικά στην επόμενη γενιά τα οποία, αν και σε αυτό το στάδιο της βελτιστοποίησης δεν θεωρούνται χρήσιμα, ενδεχομένως στο μέλλον να είναι. Ένα παράδειγμα εφαρμογής της παραπάνω μεθόδου δίνεται στο Σχήμα 2.1 για ένα πρόβλημα μεγιστοποίησης με  $\lambda = 5$ .

Απόγονος $\vec{X}_i$	Τιμή $f(\vec{X}_i)$	Πιθανότητα $P_i$
1	1.8	0.20
2	3.0	0.33
3	1.3	0.14
4	0.9	0.10
5	2.1	0.23
Άθροισμα	9.1	1.00



**Σχήμα 2.1:** Πίνακας με στοιχεία απογόνων ενός προβλήματος μεγιστοποίησης (αριστερά), Αναλογικός τροχός ρουλέτας που θα χρησιμοποιηθεί για την επιλογή γονέων (δεξιά).

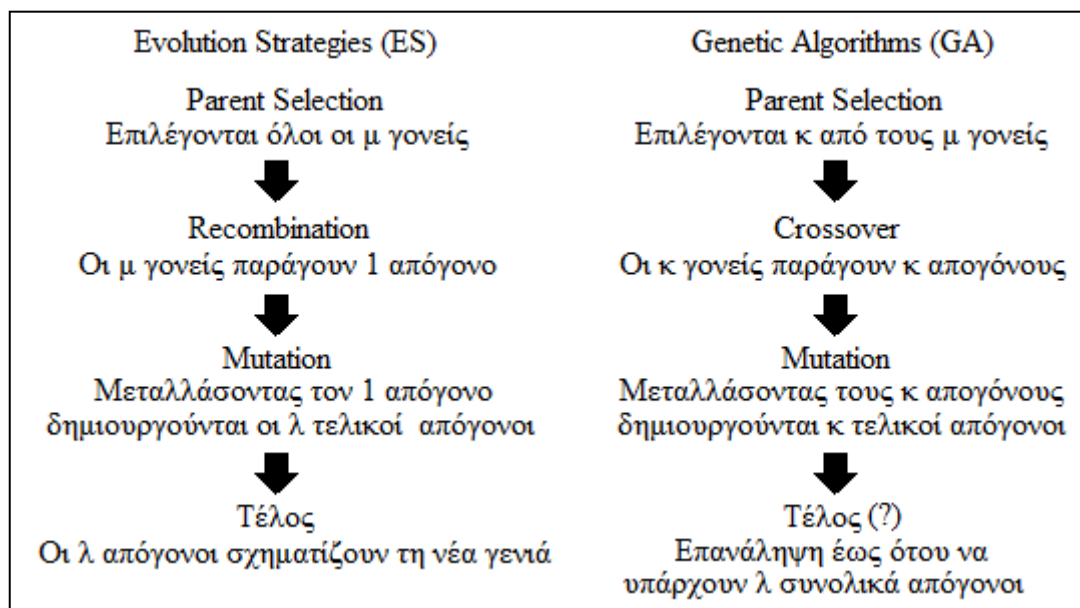
Τέλος, ένας άλλος τρόπος επιλογής που χρησιμοποιείται πιο συχνά από σύγχρονους GA είναι η επιλογή με διαγωνισμό (tournament selection)<sup>[2]</sup>. Προκειμένου να λειτουργήσει η μέθοδος πρέπει να οριστούν πρώτα δύο μεγέθη από το χρήστη: το μέγεθος του διαγωνισμού  $t$  και η πιθανότητα επιλογής  $p$ . Επιλέγονται αρχικά  $t$  τυχαία άτομα από ολόκληρο τον πληθυσμό, τα οποία κατατάσσονται βάσει της αξιολόγησής τους και διεξάγεται μεταξύ τους διαγωνισμός για το ποιο θα καταφέρει να γίνει γονέας. Η πιθανότητα να επιλεγεί κάθε διαγωνιζόμενος ισούται με:

$$\text{Sel\_Prob} = p(1 - p)^{(r-1)} \quad (2.2)$$

όπου  $r$  είναι η θέση του στην κατάταξη. Έτσι ο καλύτερος υποψήφιος έχει πιθανότητα  $p$  να γίνει γονέας, ο δεύτερος καλύτερος  $p(1 - p)$ , ο τρίτος καλύτερος  $p(1 - p)^2$  κ.ο.κ. έως τον τελευταίο υποψήφιο που αναλαμβάνει ότι απομένει ώστε να συμπληρωθούν όλες οι πιθανότητες. Όταν η διαδικασία τελειώσει οι υποψήφιοι που δεν επιλέχτηκαν επιστρέφουν στη δεξαμενή επιλογής και πραγματοποιείται νέος διαγωνισμός έως ότου βρεθούν όλοι οι  $\mu$  γονείς. Η μέθοδος αυτή έχει παρόμοια στοχαστικά στοιχεία με την προηγούμενη αλλά είναι πιο ευέλικτη διότι τα χαρακτηριστικά της επιλογής εξαρτώνται από τις τιμές των  $t$  και  $p$  που θα οριστούν. Μεγάλες τιμές  $t$  ή και  $p$  προσδίδουν ελιτιστικά χαρακτηριστικά στην επιλογή διότι μειώνονται οι πιθανότητες να επιλεχτούν οι πιο αδύναμοι υποψήφιοι, ενώ παράλληλα υπάρχει και η δυνατότητα για εντελώς τυχαία επιλογή γονέων άμα δοθεί  $t = 1$ . Αναφέρεται ακόμη ότι μπορεί να προγραμματιστεί και με ρύθμιση που να επιτρέπει την επανατοποθέτηση του επιλεγμένου γονέα στη δεξαμενή επιλογής, αυξάνοντας περαιτέρω τον ελιτισμό του αλγορίθμου.

## 2.2.2 Δημιουργία Απογόνων

Έχοντας πλέον επιλέξει τους  $\mu$  γονείς με κάποιον τρόπο, απομένει να δημιουργηθούν οι  $\lambda$  νέοι απόγονοι που θα αποτελέσουν τη νέα γενιά. Στα συγγράμματα της επιστημονικής κοινότητας των EA παρατηρείται μεγάλη σύγχυση αναφορικά με τη χρήση όρων που περιγράφουν τη διαδικασία δημιουργίας απογόνων, καθώς οι χρήστες των GA και των ES μπορεί να χρησιμοποιούν την ίδια ορολογία για να περιγράψουν μια διαφορετική έννοια ή και αντίστροφα. Ορίζονται τρεις όροι οι οποίοι, στο πλαίσιο της εργασίας, επισημαίνεται ότι χρησιμοποιούνται ως εξής: η διασταύρωση (crossover) για την περιγραφή διαδικασιών GA, ο ανασυνδυασμός (recombination) για την περιγραφή διαδικασιών ES και η μετάλλαξη (mutation) που υφίσταται και στις δύο αυτές κατηγορίες EA αλλά με διαφορετικό τρόπο και αποτέλεσμα. Στο Σχήμα 2.2 φαίνεται ένα ενδεικτικό διάγραμμα ροής της δημιουργίας απογόνων για ES και GA, όπου εμφανίζεται και ένα νέο μέγεθος  $\kappa$ . Το  $\kappa$  καθορίζει πόσοι γονείς ζευγαρώνουν για τη δημιουργία νέων απογόνων και όντας εμπνευσμένη από το φυσικό περιβάλλον, η κλασική θεωρία των GA θέτει  $\kappa = 2$ .



**Σχήμα 2.2:** Διαγράμματα ροής για τη δημιουργία νέας γενιάς από μια τυπική Στρατηγική Εξέλιξης (ES, αριστερά) και από έναν τυπικό Γενετικό Αλγόριθμο (GA, δεξιά).

Αρχικά θα αναλυθεί η δημιουργία απογόνων των ES. Έχοντας επιλέξει τους γονείς πιθανότατα με αποκοπή, στη βιβλιογραφία υπάρχουν δύο τρόποι ανασυνδυασμού με σκοπό τη δημιουργία ενός και μοναδικού απογόνου: ο διακριτός ή κυρίαρχος (discrete or dominant) και ο ενδιάμεσος (intermediate) <sup>[11]</sup>. Σύμφωνα με το διακριτό, η τιμή κάθε



μεταβλητής σχεδιασμού του απογόνου λαμβάνεται από έναν τυχαίο γονέα (κυριαρχία). Ο ενδιάμεσος ανασυνδυασμός, αντιθέτως, λαμβάνει υπόψη τα χαρακτηριστικά όλων των γονέων, υπολογίζοντας τη μέση τιμή κάθε μεταβλητής σχεδιασμού και αναθέτοντάς τη στον απόγονο. Οι δύο μέθοδοι ανασυνδυασμού εφαρμόζονται στο Σχήμα 2.3 σε ένα πρόβλημα με  $n = 5$  μεταβλητές σχεδιασμού για  $\mu = 3$ . Συμβολίζοντας τον απόγονο που προκύπτει με χρήση οποιασδήποτε εκ των δύο μεθόδων ως  $\vec{X}_0$ , για να ληφθούν οι  $\lambda$  τελικοί απόγονοι  $\vec{X}_i$  εφαρμόζονται  $\lambda$  συνολικά μεταλλάξεις της γενικής μορφής:

$$\vec{X}_i = \vec{X}_0 + \sigma \vec{z}_i \quad (2.3)$$

όπου το  $\sigma$  ονομάζεται δύναμη μετάλλαξης (mutation power) ή μέγεθος βήματος (step size) και:

$$\vec{z}_i = \begin{Bmatrix} N_1(0,1) \\ N_2(0,1) \\ \dots \\ N_{n-1}(0,1) \\ N_n(0,1) \end{Bmatrix} \quad (2.4)$$

Τα διανύσματα  $\vec{z}_i$  ονομάζονται διανύσματα μετάλλαξης (mutation vectors) ή βήματα (steps) και κάθε στοιχείο τους  $N_i(0,1)$  αποτελεί ένα τυχαίο δείγμα (sample) από την τυποποιημένη κανονική κατανομή  $N(0,1)$ . Φαίνεται έτσι ότι κάθε απόγονος  $\vec{X}_i$  διαφέρει από τους υπόλοιπους μόνο λόγω του μοναδικού διανύσματος  $\vec{z}_i$  το οποίο τον γέννησε.

Γονείς	$\vec{X}_1 = \begin{Bmatrix} 1.5 \\ 2.0 \\ 0.6 \\ 3.6 \\ 2.4 \end{Bmatrix}, \vec{X}_2 = \begin{Bmatrix} 0.5 \\ 1.7 \\ 1.9 \\ 5.0 \\ 1.1 \end{Bmatrix}, \vec{X}_3 = \begin{Bmatrix} 1.0 \\ 3.5 \\ 2.0 \\ 3.4 \\ 0.1 \end{Bmatrix}$
<u>Discrete Recombination</u> Πολλοί Πιθανοί Απόγονοι, π.χ.	$\vec{X}_0' = \begin{Bmatrix} 0.5 \\ 1.7 \\ 1.9 \\ 3.4 \\ 2.4 \end{Bmatrix}, \vec{X}_0'' = \begin{Bmatrix} 1.0 \\ 2.0 \\ 2.0 \\ 3.4 \\ 0.1 \end{Bmatrix}, \vec{X}_0''' = \begin{Bmatrix} 1.5 \\ 3.5 \\ 0.6 \\ 5.0 \\ 1.1 \end{Bmatrix}, \dots$
<u>Intermediate Recombination</u> Μοναδικός Πιθανός Απόγονος	$\vec{X}_0 = \begin{Bmatrix} 1.0 \\ 2.4 \\ 1.5 \\ 4.0 \\ 1.2 \end{Bmatrix}$

**Σχήμα 2.3:** Εφαρμογή διακριτού (πάνω) και ενδιάμεσου (κάτω) ανασυνδυασμού τριών γονέων σε πρόβλημα βελτιστοποίησης με πέντε μεταβλητές σχεδιασμού.

Στους GA, οι υποψήφιες λύσεις αντιμετωπίζονται ως δυαδικές συστοιχίες (binary strings) που ονομάζονται χρωμοσώματα σύμφωνα με την ορολογία της βιολογίας. Κάθε θέση ενός χρωμοσώματος ονομάζεται bit και μπορεί να λάβει τιμή 0 ή 1. Οι τελεστές διασταύρωσης και μετάλλαξης εφαρμόζονται απευθείας στα χρωμοσώματα των γονέων και παράγουν τους νέους απογόνους. Για την παρουσίαση των τελεστών αυτών γίνεται η υπόθεση ενός κλασικού GA όπου χρησιμοποιείται η παραπάνω μέθοδος δυαδικής κωδικοποίησης και σύμφωνα με το συμβολισμό που ορίστηκε στο Σχήμα 2.2 ισχύει  $\kappa = 2$ , άρα γίνεται χρήση 2 γονέων για την παραγωγή 2 απογόνων.

Η διασταύρωση ενεργοποιείται με μία μεγάλη πιθανότητα, κοντά στη μονάδα, προκειμένου να υπάρχει παράλληλα και η δυνατότητα μεταφοράς αναλλοίωτων γονικών χαρακτηριστικών στην επόμενη γενιά<sup>[2]</sup>. Ως κλασικοί τρόποι διασταύρωσης αναφέρονται ενδεικτικά η διασταύρωση ενός σημείου (1-point crossover), γενίκευση της οποίας αποτελεί η διασταύρωση  $\kappa$  σημείων ( $\kappa$ -point crossover), η διακριτή (discrete) και η ομοιόμορφη (uniform) διασταύρωση<sup>[12]</sup>. Κατά τις διασταυρώσεις με χρήση  $\kappa$  σημείων επιλέγονται  $\kappa$  τυχαίες θέσεις στα χρωμοσώματα των γονέων και πραγματοποιούνται τομές, χωρίζοντάς τα σε  $\kappa + 1$  επιμέρους συστοιχίες (substrings). Οι απόγονοι προκύπτουν από τυχαία επανασύνδεση των απομονωμένων αυτών τμημάτων, ενώ υπάρχει και ο εναλλακτικός τρόπος οι τομές να γίνονται ανά μεταβλητή σχεδιασμού (multivariate crossover). Η διακριτή διασταύρωση δημιουργεί έναν απόγονο θέτοντας την τιμή κάθε bit ίση με κάποιου τυχαίου γονέα με τρόπο παρόμοιο με το διακριτό ανασυνδυασμό των ES. Η διαδικασία επαναλαμβάνεται και για τον δεύτερο απόγονο. Η ομοιόμορφη διασταύρωση τελεί την ακριβώς ίδια πρώτη ενέργεια με τη διακριτή, αλλά ο δεύτερος απόγονος προκύπτει από τα bit που δεν χρησιμοποιήθηκαν για την παραγωγή του πρώτου. Όλες οι παραπάνω μέθοδοι διασταύρωσης παρουσιάζονται στο Σχήμα 2.4.

Όταν τελειώσει η διαδικασία της διασταύρωσης σαρώνεται κάθε bit των χρωμοσωμάτων των απογόνων που προέκυψαν και ενεργοποιείται ο τελεστής μετάλλαξης με κάποια πολύ μικρή πιθανότητα της τάξης του 2%. Σε περίπτωση που ενεργοποιηθεί, η τιμή του τρέχοντος bit αλλάζει από 0 σε 1 ή αντίστροφα. Οι τελικοί απόγονοι προστίθενται στη νέα γενιά και η διαδικασία του Σχήματος 2.2 επαναλαμβάνεται μέχρι να συγκεντρωθούν όλοι οι απαιτούμενοι απόγονοι για να σχηματιστεί η νέα γενιά.

Γονείς	<table border="1"> <tbody> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td></tr> </tbody> </table>	0	1	1	0	1	1	0	1	0	0	0	0	1	1	0	1	1	0	1	1
0	1	1	0	1	1	0	1	0	0												
0	0	1	1	0	1	1	0	1	1												
<u>1-Point Crossover</u>	<table border="1"> <tbody> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> </tbody> </table>	0	1	1	0	1	1	1	0	1	1	0	0	1	1	0	1	0	1	0	0
0	1	1	0	1	1	1	0	1	1												
0	0	1	1	0	1	0	1	0	0												
Δύο Απόγονοι																					
<u>2-Point Crossover</u>	<table border="1"> <tbody> <tr><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td></tr> </tbody> </table>	0	1	1	1	0	1	0	1	0	0	0	0	1	0	1	1	1	0	1	1
0	1	1	1	0	1	0	1	0	0												
0	0	1	0	1	1	1	0	1	1												
Δύο Απόγονοι																					
<u>Discrete Crossover</u>	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td></tr> </tbody> </table>	0	0	1	1	0	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0
0	0	1	1	0	1	1	1	0	1												
0	0	1	1	1	1	1	1	1	0												
Ένας Απόγονος																					
<u>Uniform Crossover</u>	<table border="1"> <tbody> <tr><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td></tr> </tbody> </table>	0	0	1	1	0	1	1	1	0	1	0	1	1	0	1	1	0	0	1	0
0	0	1	1	0	1	1	1	0	1												
0	1	1	0	1	1	0	0	1	0												
Δύο Απόγονοι																					

**Σχήμα 2.4:** Παραδείγματα εφαρμογής των βασικών τελεστών διασταύρωσης των GA. Στη διακριτή διασταύρωση (discrete crossover) πραγματοποιήθηκαν δύο εφαρμογές ώστε να φανεί ξεκάθαρα ο διαφορετικός τρόπος λειτουργίας της εν συγκρίσει με την ομοιόμορφη (uniform).

Κλείνοντας, αναφέρεται ότι ανάλογα τη θέση των bit τα οποία τροποποιούνται κατά τις παραπάνω διεργασίες μπορεί να γίνουν ουσιαστικά αμελητέες ή και παντελώς τρομακτικές αλλαγές στην πραγματική τιμή των μεταβλητών σχεδιασμού ενός απογόνου. Υπάρχουν όμως πολλοί τρόποι που ένας GA μπορεί να επιβάλει ελιτισμό και να αποφύγει τη χειροτέρευση των χαρακτηριστικών του πληθυσμού, π.χ. με μια ξεχωριστή πιθανότητα η βέλτιστη λύση να επιλεγεί αυτομάτως ως γονέας χωρίς να ενεργοποιηθεί η επιλογή γονέων ή με αντικατάσταση του χειρίστου απογόνου μιας κακής γενιάς με τον βέλτιστο της προηγούμενης. Καλό είναι, πάντως, τέτοιες ενέργειες να περιορίζονται ή να εφαρμόζονται με πολύ μικρή πιθανότητα, διαφορετικά ο πληθυσμός κινδυνεύει να ομογενοποιηθεί ταχύτατα κατ' εικόνα των αρχικών βέλτιστων λύσεων που ανιχνεύονται στις πρώτες γενιές. Χωρίς ποικιλία χαρακτηριστικών, ο τελεστής διασταύρωσης θα δημιουργεί σταδιακά όλο και πιο παρόμοια χρωμοσώματα για τους νέους απογόνους και κατά πάσα πιθανότητα ο GA θα συγκλίνει πρόωρα σε κάποιο τοπικό ακρότατο, αφού η μετάλλαξη από μόνη της δεν αρκεί για να εξασφαλίσει επαρκείς αλλαγές τιμής των bit στον πληθυσμό.

## 2.3 ΣΥΝΘΕΤΑ ΠΡΟΒΛΗΜΑΤΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

Χρησιμοποιώντας τις πληροφορίες της Ενότητας 2.2 μπορεί να δημιουργηθεί ένας ΕΑ ικανός να λύσει ένα βασικό πρόβλημα SOO όπως αυτό παρουσιάστηκε στην Ενότητα 2.1. Όμως ένα πραγματικό πρόβλημα βελτιστοποίησης σπάνια είναι τόσο απλό και συνήθως ανήκει σε μία ή περισσότερες από τις παρακάτω κατηγορίες, επομένως κρίνεται αναγκαίο να γίνουν ορισμένες τροποποιήσεις στο γενικό ΕΑ προκειμένου το πρόβλημα να μπορεί να αντιμετωπιστεί. Τονίζεται ότι οι τρόποι αντιμετώπισης που θα αναφερθούν ισχύουν μόνο για επίλυση με χρήση ΕΑ και διαφέρουν σημαντικά αν επιστρατευτούν αιτιοκρατικές ή λοιπές στοχαστικές μέθοδοι βελτιστοποίησης.

### 2.3.1 Προβλήματα με Περιορισμούς

Τα προβλήματα βελτιστοποίησης με φυσική σημασία είναι ως επί το πλείστον προβλήματα με περιορισμούς. Για παράδειγμα, αν πρέπει να σχεδιαστεί μια κατασκευή με χαμηλό χρηματικό κόστος, προφανώς η συνάρτηση-στόχος θα είναι το κόστος της σε ευρώ, αλλά ταυτόχρονα πρέπει να κατασκευαστεί βάσει των κανονιστικών διατάξεων ώστε να μην αστοχήσει. Τα όρια αντοχής επομένως πρέπει να ληφθούν ως περιορισμοί και να μην ξεπεραστούν σε καμία απολύτως περίπτωση. Επισημαίνεται ότι δεν υπάρχει λόγος να επιβάλλονται περιορισμοί στις συναρτήσεις-στόχους, στο παράδειγμα αυτό ένα μέγιστο επιτρεπτό κόστος, διότι οι λύσεις βελτιστοποιούνται σύμφωνα με αυτές και κρίνεται περιττό. Ένας περιορισμός μπορεί να συμπεριλαμβάνει ένα υποσύνολο των μεταβλητών σχεδιασμού ή και όλες μαζί ταυτόχρονα. Σε λογική black-box, μάλιστα, η ακριβής μορφή των περιορισμών είναι παντελώς άγνωστη και το λογισμικό αξιολόγησης οφείλει να μπορεί να τροφοδοτήσει τον ΕΑ όχι μόνο με τις τιμές των συναρτήσεων στόχων  $f_i(\vec{X})$ , αλλά και με τις τιμές των περιορισμών  $g_i(\vec{X})$  για κάθε υποψήφιο  $\vec{X}$ .

Τα ακατάλληλα άτομα του πληθυσμού που δεν τηρούν έναν ή παραπάνω από τους περιορισμούς σε καμία περίπτωση δεν πρέπει να αγνοηθούν από τον ΕΑ. Αν τους επιτραπεί να αναπαράγονται ελεύθερα, τότε ο πληθυσμός θα αποκτήσει σε βάθος χρόνου ακατάλληλα χαρακτηριστικά και η βελτιστοποίηση θα αποτύχει. Ένας χονδροκομμένος τρόπος που μπορεί να δουλέψει αλλά δεν αντιμετωπίζει το πρόβλημα αποτελεσματικά θα ήταν να αξιολογούνται οι απόγονοι την ώρα που γεννιούνται και, σε περίπτωση που κριθούν ακατάλληλοι, να διαγράφονται, με τη διαδικασία αυτή να επαναλαμβάνεται έως ότου βρεθεί κάποιος που τηρεί όλους τους περιορισμούς <sup>[13]</sup>. Το προκύπτον εδώ θέμα

είναι ότι σε περίπτωση που υπάρχουν αυστηροί περιορισμοί μπορεί να αποδειχτεί εξαιρετικά δύσκολο να βρεθεί ένας κατάλληλος απόγονος με αποτέλεσμα να απαιτούνται πάρα πολλές αξιολογήσεις ανά γενιά. Όντας το πιο ακριβό τμήμα λειτουργίας ενός EA, είναι πολύ πιο θεμιτό να αξιοποιούνται τα αποτελέσματα κάθε αξιολόγησης αντί να πραγματοποιούνται καταχρηστικά από τον αλγόριθμο.

Ένας ορθός και κομψότερος τρόπος χειρισμού των περιορισμών είναι να περιοριστεί η δυνατότητα των ακατάλληλων υποψηφίων να γίνονται γονείς κατά την επιλογή γονέων. Αυτή η τεχνική αντικατοπτρίζει καλύτερα τους μηχανισμούς της φύσης και μπορεί να υλοποιηθεί απλά και αποτελεσματικά επιβάλλοντας κάποια ποινή (penalty) στην τιμή της συνάρτησης-στόχου που αντιστοιχεί στα άτομα που δεν πληρούν όλες τις σχεδιαστικές προϋποθέσεις <sup>[8]</sup>. Ως επακόλουθο, οι ανεπιθύμητες λύσεις θα βρίσκονται χαμηλότερα στη γραμμική κατάταξη των ES και θα έχουν χαμηλότερες πιθανότητες να επιλεγούν από τις στοχαστικές διαδικασίες των GA. Άρα για κάθε απόγονο  $\vec{X}$  θα είναι:

$$f(\vec{X})_{ps} = f(\vec{X}) \quad (2.5)$$

αν τηρούνται όλοι οι περιορισμοί, αλλιώς:

$$f(\vec{X})_{ps} = f(\vec{X}) \pm \text{Penalty} \quad (2.6)$$

όπου  $f(\vec{X})_{ps}$  είναι η τιμή της συνάρτησης στόχου που θα χρησιμοποιηθεί για τον τελεστή της επιλογής γονέων. Στην Εξίσωση (2.6) η ποινή προστίθεται σε περίπτωση που λύνεται πρόβλημα ελαχιστοποίησης και αφαιρείται σε ένα πρόβλημα μεγιστοποίησης προκειμένου ο υποψήφιος να τιμωρηθεί κατάλληλα.

Το μόνο που απομένει να αποφασισθεί είναι το μέτρο της επιβαλλόμενης ποινής. Ως θανατική ποινή (death penalty) ορίζεται μια ποινή «άπειρου» μέτρου π.χ.  $10^{20}$  που πρακτικά εκμηδενίζει την πιθανότητα επιλογής ενός ατόμου ως γονέα. Ένας EA που λειτουργεί αποκλειστικά με χρήση θανατικής ποινής δεν θα είναι αποτελεσματικός διότι οι αξιολογήσεις των ακατάλληλων ατόμων ουσιαστικά παραμερίζονται, με την έννοια ότι δεν θα συνδράμουν καθόλου στην αναζήτηση του βέλτιστου. Επίσης, αναλογίζοντας τη φυσική έννοια της τιμωρίας, είναι πραγματικά άδικο να τιμωρούνται εξίσου ένας υποψήφιος που δεν τηρεί οριακά έναν περιορισμό με έναν άλλο υποψήφιο που δεν τηρεί κανέναν και μάλιστα παρουσιάζοντας μεγάλες αποκλίσεις από τους δοθέντες φραγμαούς. Βάσει των παραπάνω παρατηρήσεων λαμβάνονται οι εξής αποφάσεις:

1. Ένας υποψήφιος τιμωρείται ξεχωριστά για κάθε περιορισμό που δεν τηρεί.
2. Η τιμωρία αυτή εξαρτάται από τον βαθμό στον οποίο ο εν λόγω περιορισμός δεν τηρείται.
3. Η θανατική ποινή επιβάλλεται μόνο σε περιπτώσεις όπου διακρίνεται ακραία καταπάτηση οποιουδήποτε περιορισμού.

Έτσι, για κάθε περιορισμό της μορφής  $g_i \leq a_i$  ορίζεται ένα επιπλέον όριο  $b_i$  ώστε  $a_i \leq b_i$ , όπου  $a_i$  είναι το πραγματικό άνω όριο του περιορισμού  $g_i$  και  $b_i$  ένα προσαυξημένο άνω όριο που ορίζεται από το χρήστη. Αν ο συνολικός αριθμός των περιορισμών συμβολιστεί ως  $c$ , οι Εξισώσεις (2.5) και (2.6) ενώνονται σε μια ενιαία έκφραση:

$$f(\vec{X})_{ps} = f(\vec{X}) \pm \sum_i^c \text{Penalty}_i \quad (2.7)$$

όπου:

$$\bullet \text{ αν } g_i(\vec{X}) \leq a_i(\vec{X}) \quad \rightarrow \text{Penalty}_i = 0 \quad (2.8)$$

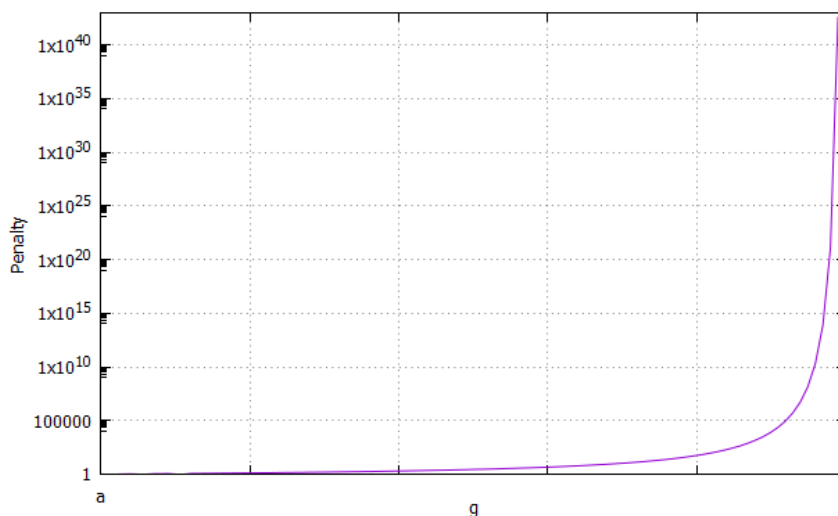
$$\bullet \text{ αν } a_i(\vec{X}) < g_i(\vec{X}) < b_i(\vec{X}) \rightarrow \text{Penalty}_i = \exp \left[ \frac{g_i(\vec{X}) - a_i(\vec{X})}{b_i(\vec{X}) - g_i(\vec{X})} \right] \quad (2.9)$$

$$\bullet \text{ αν } g_i(\vec{X}) \geq b_i(\vec{X}) \quad \rightarrow \text{Penalty}_i = \text{Death\_Penalty} \quad (2.10)$$

Αντίστοιχες σχέσεις μπορούν να κατασκευαστούν με την ίδια λογική για περιορισμούς της μορφής  $g_i \geq a_i$ , όπου κατά αναλογία πρέπει  $a_i \geq b_i$ . Η Εξίσωση (2.9) είναι μια πιο απλή έκδοση της αντίστοιχης που δίνεται στο εγχειρίδιο χρήσης του λογισμικού EASY <sup>[8]</sup> και επιβάλλει εκθετικά αυξανόμενη ποινή σε άτομα που δεν αξίζουν τη θανατική όπως φαίνεται στο Σχήμα 2.5. Διακρίνεται ότι η θανατική ποινή έχει ουσιαστικά αρχίσει να επιβάλλεται ήδη αρκετά πριν η  $g_i$  φτάσει την τιμή  $b_i$ , πράγμα που πρέπει να ληφθεί υπόψη όταν δίνεται η τιμή από το χρήστη. Προτείνεται κατά τις πρώτες δοκιμές βελτιστοποίησης ενός προβλήματος με περιορισμούς να δίνονται πολύ μεγάλες τιμές  $b_i$  ώστε να είναι σίγουρο ότι ο αλγόριθμος θα λειτουργήσει χωρίς επιπλοκές και, σε περίπτωση που διαπιστωθεί αργή σύγκλιση, να μειώνονται σταδιακά.

Τονίζεται ότι οποιαδήποτε λύση  $\vec{X}$  βρίσκεται στην περιοχή  $a_i(\vec{X}) < g_i(\vec{X}) < b_i(\vec{X})$  είναι ακατάλληλη, δηλαδή δεν μπορεί να θεωρηθεί ως πιθανή λύση του προβλήματος βελτιστοποίησης. Μπορεί, εν γένει, λόγω της παραπάνω μεθοδολογίας, να φαίνεται ότι έχει βρεθεί ένα νέο βέλτιστο άτομο  $\vec{X}^*$  λόγω της μικρής ποινής που του επιβλήθηκε κατά τον έλεγχο των περιορισμών, θα ήταν όμως σοβαρό λάθος να αποθηκευτεί και να αντικαταστήσει το τρέχον βέλτιστο κατάλληλο άτομο αφού δεν είναι βέβαιο ότι θα βρεθεί

νέος βέλτιστος απόγονος που τηρεί όλους τους περιορισμούς προτού τερματιστεί ο αλγόριθμος.



**Σχήμα 2.5:** Διάγραμμα του μέτρου της ποινής σε λογαριθμική κλίμακα που επιβάλλεται στην τιμή  $f$  ενός απογόνου στην περιοχή  $a \leq g \leq b$  υπό περιορισμό της μορφής  $g \leq a$ .

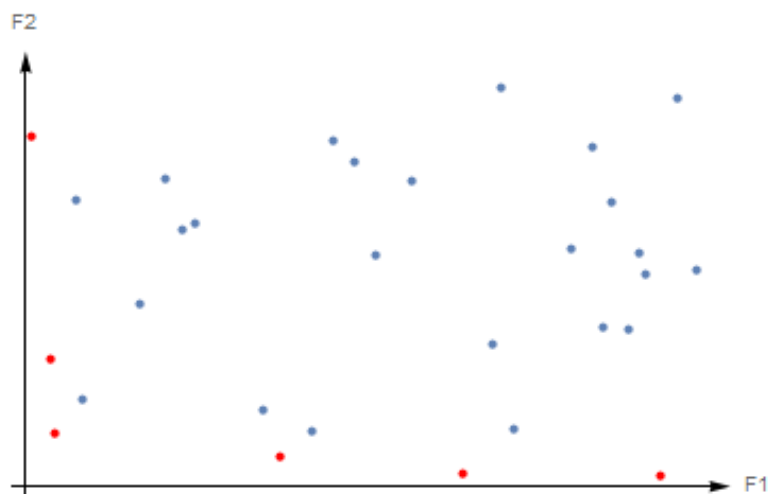
### 2.3.2 Προβλήματα με Πολλούς Στόχους

Έστω ότι δίνεται ένα πρόβλημα ελαχιστοποίησης  $m$  στόχων με  $n$  μεταβλητές σχεδιασμού, δηλαδή υπάρχουν  $m$  συναρτήσεις-στόχοι  $f_1(\vec{X})$ ,  $f_2(\vec{X})$ , ...,  $f_{m-1}(\vec{X})$ ,  $f_m(\vec{X})$  που εξαρτώνται από τις ίδιες  $n$  συνολικά κοινές μεταβλητές. Ορίζεται έτσι η διανυσματική συνάρτηση-στόχος προς ελαχιστοποίηση  $\vec{F}(\vec{X})$  ως <sup>[4]</sup>:

$$\vec{F}(\vec{X}) = \begin{pmatrix} f_1(\vec{X}) \\ f_2(\vec{X}) \\ \dots \\ f_{m-1}(\vec{X}) \\ f_m(\vec{X}) \end{pmatrix} \quad (2.11)$$

Το κύριο πρόβλημα που προκύπτει βάσει της θεωρίας των ΕΑ που παρουσιάστηκε στην Ενότητα 2.2 εντοπίζεται στον τομέα της επιλογής γονέων. Όλες οι μέθοδοι που προτάθηκαν στην Υποενότητα 2.2.2 για SOO πραγματοποιούσαν άμεσα (ES) ή έμμεσα (GA) κατάταξη των λύσεων από την καλύτερη στη χειρότερη και χρησιμοποιούσαν την τιμή της  $f(\vec{X})$  ως κριτήριο. Τώρα υπάρχουν όμως πολλές, εξίσου σημαντικές και σχεδόν πάντα αντικρουόμενες  $f_i(\vec{X})$  με αποτέλεσμα να μην είναι καθόλου προφανές ποιοι απόγονοι πρέπει να γίνουν γονείς.

Για την καλύτερη κατανόηση του προβλήματος και των τρόπων αντιμετώπισής του, σε ένα πρόβλημα ελαχιστοποίησης το διάνυσμα  $\vec{X}_P$  ονομάζεται κατά Pareto βέλτιστη λύση άμα δεν υπάρχει κανένα άλλο διάνυσμα  $\vec{X}_R$  στο χωρίο σχεδιασμού ώστε  $f_i(\vec{X}_R) \leq f_i(\vec{X}_P)$  για κάθε  $i = 1, 2, \dots, m$  και ταυτόχρονα ισχύει  $f_j(\vec{X}_R) < f_j(\vec{X}_P)$  για έναν τουλάχιστον δείκτη στόχου  $j$ . Αντίστροφα, το διάνυσμα  $\vec{X}_D$  είναι κυριαρχούμενο (dominated) από το  $\vec{X}_R$  άμα ισχύει  $f_i(\vec{X}_R) \leq f_i(\vec{X}_D)$  για κάθε  $i = 1, 2, \dots, m$  και  $f_j(\vec{X}_R) < f_j(\vec{X}_D)$  για έναν τουλάχιστον δείκτη στόχου  $j$  [2], [3], [4]. Σύμφωνα με τους παραπάνω ορισμούς προκύπτει ότι οι κατά Pareto βέλτιστες λύσεις είναι οι μόνες μη-κυριαρχούμενες λύσεις στο χωρίο σχεδιασμού και ακόμα πιο σημαντικά, ότι ένα κατά Pareto βέλτιστο διάνυσμα σχεδιασμού  $\vec{X}_P$  στο χωρίο σχεδιασμού αντιστοιχεί σε μία κατά Pareto βέλτιστη διανυσματική συνάρτηση-στόχο  $\vec{F}(\vec{X}_P)$  στο χωρίο των συναρτήσεων-στόχων. Ως παράδειγμα, δίνεται το Σχήμα 2.6, στο οποίο γίνεται ο διαχωρισμός τριάντα διανυσματικών συναρτήσεων στόχου ενός προβλήματος ελαχιστοποίησης δύο στόχων σε κυριαρχούμενες και μη.



**Σχήμα 2.6:** Εποπτική κατηγοριοποίηση των διανυσματικών συναρτήσεων-στόχων σε πρόβλημα ελαχιστοποίησης δύο στόχων, όπου με κόκκινο συμβολίζονται οι μη-κυριαρχούμενες και με μπλε οι κυριαρχούμενες.

Το σύνολο των κατά Pareto βέλτιστων λύσεων ονομάζεται μέτωπο Pareto και αποτελεί την τελική λύση ενός προβλήματος MOO. Πρέπει να γίνει αντιληπτό ότι οι διανυσματικές συναρτήσεις  $\vec{F}(\vec{X}_P)$  που ανήκουν στο μέτωπο και κατ' επέκταση, οι λύσεις  $\vec{X}_P$  που τους αντιστοιχούν δεν μπορούν να συγκριθούν μεταξύ τους με την κλασική έννοια. Καθεμία πρεσβεύει ένα διαφορετικό βέλτιστο αποτέλεσμα του σχεδιασμού και ο χρήστης καλείται να επιλέξει σύμφωνα με άλλα κριτήρια ποια λύση θα υιοθετήσει.



Υπάρχουν πολλές μέθοδοι επίλυσης προβλημάτων MOO, η πλειοψηφία εκ των οποίων βασίζονται στον ορισμό μιας νέας συνάρτησης-στόχου  $\Phi$  που αντικαθιστά τη διανυσματική συνάρτηση  $\vec{F}$  προκειμένου το πρόβλημα να μπορεί να επιλυθεί ως SOO [4]. Στο πλαίσιο της εξελικτικής βελτιστοποίησης πολλών στόχων η  $\Phi$  ορίζεται κυρίως βάσει του μετώπου Pareto, αφού το μεγάλο προτέρημα ενός EA είναι ότι σε κάθε γενιά υπάρχει ένα ενεργό μέτωπο μη-κυριαρχούμενων λύσεων. Έτσι είναι δυνατόν μέσω κατάλληλα επιλεγμένων  $\Phi$  που αντικατοπτρίζουν τα χαρακτηριστικά του τρέχοντος πληθυσμού να προσεγγίζεται όλο και περισσότερο το πραγματικό μέτωπο Pareto του προβλήματος. Αναφέρεται ότι η  $\Phi$  μπορεί εναλλακτικά να οριστεί ως μια σταθμισμένη ποσότητα των διαφορών  $f_i$  του προβλήματος, αλλά η σωστή ρύθμισή της απαιτεί εκ των προτέρων (a priori) εξαιρετική γνώση ή και κατανόηση του προβλήματος, οπότε μια τέτοια προσέγγιση δεν μπορεί να εφαρμοστεί υπό συνθήκες black-box.

Ο αλγόριθμος SPEA (Strength Pareto Evolutionary Algorithm) [14] των Eckart Zitzler και Lothar Thiele αποτελεί ένα χαρακτηριστικό παράδειγμα των αρχών MOO που χρησιμοποιεί για τον ορισμό της  $\Phi$  την κυριαρχία των υποψήφιων λύσεων. Για κάθε μη-κυριαρχούμενο άτομο  $i$  υπολογίζεται η ισχύς (strength)  $s_i$  ως:

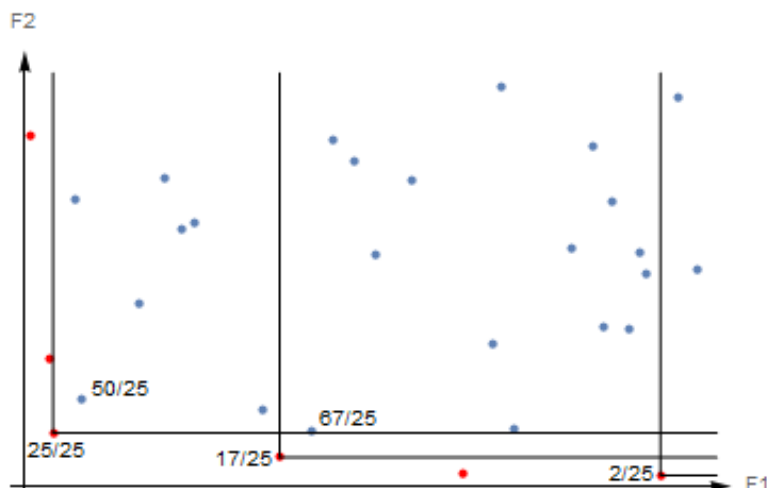
$$s_i = \frac{v_i}{N+1} \quad (2.12)$$

όπου  $v_i$  είναι ο αριθμός των ατόμων στα οποία το  $i$  κυριαρχεί και  $N$  ο συνολικός αριθμός των κυριαρχούμενων ατόμων του πληθυσμού. Ο παρονομαστής στην Εξίσωση (2.12) κατασκευάζεται έτσι ώστε σε κάθε περίπτωση να ισχύει  $0 \leq s_i < 1$ . Αφού υπολογιστούν όλα τα  $s_i$ , η τιμή  $\Phi$  που αντιστοιχεί σε κάθε άτομο και θα χρησιμοποιηθεί κατά τη διαδικασία της επιλογής γονέων προκύπτει ως εξής:

- για τα μη-κυριαρχούμενα άτομα  $i \rightarrow \Phi_i = s_i$  (2.13)

- για τα κυριαρχούμενα άτομα  $j \rightarrow \Phi_j = 1 + \sum_{i,i \geq j} s_i$  (2.14)

Ο όρος  $\sum_{i,i \geq j} s_i$  συμβολίζει το άθροισμα των ισχύων  $s_i$  των κυριαρχούντων στο  $j$  ατόμων, στο οποίο προστίθεται και η μονάδα ώστε να εξασφαλιστεί ότι  $\Phi_i < \Phi_j$ . Στο Σχήμα 2.7 δίνονται ενδεικτικά οι τιμές  $\Phi$  υποψήφιων λύσεων από το παράδειγμα του Σχήματος 2.6, όπου στον παρονομαστή της Εξίσωσης (2.12) προκύπτει  $N + 1 = 24 + 1 = 25$ .



**Σχήμα 2.7:** Τιμές συνάρτησης  $\Phi$  από τον αλγόριθμο SPEA σε πρόβλημα ελαχιστοποίησης δύο στόχων, με κόκκινο τα μη-κυριαρχούμενα άτομα και με μπλε τα κυριαρχούμενα.

Στο παραπάνω σχήμα οι ορθές γωνίες οριοθετούν το χωρίο στο οποίο κάθε βέλτιστη λύση κυριαρχεί. Φαίνεται ξεκάθαρα πως τα μη-κυριαρχούμενα άτομα που βρίσκονται στα άκρα του μετώπου έχουν μικρότερη τιμή  $\Phi_i$  από τα κεντρικά αφού κυριαρχούν σε λιγότερα άτομα, επομένως έχουν περισσότερες πιθανότητες να επιλεγούν κατά την επιλογή γονέων. Αν και αυτό δεν συμβαίνει πάντοτε, αποτελεί το συνήθη και επιθυμητό τρόπο λειτουργίας του SPEA ώστε να αποφευχθεί η συρρίκνωση του μετώπου κατά την επίλυση.

Για λόγους πληρότητας θα γίνει μια σύντομη αναφορά και στη βελτιωμένη εκδοχή του SPEA, τον αλγόριθμο SPEA2 <sup>[15]</sup>. Ο SPEA2 βασίζεται, όπως και ο προκάτοχός του, στην κυριαρχία των λύσεων, με μια σημαντική όμως αλλαγή στον τρόπο υπολογισμού της  $\Phi$ . Σύμφωνα με τους παραπάνω συμβολισμούς και ορολογίες, πλέον υπολογίζεται ελαφρώς διαφορετικά η ισχύς για όλα τα άτομα του αρχείου και του πληθυσμού ως:

$$s_i = \sum_{i,i \geq j} v_i \quad (2.15)$$

Φαίνεται έτσι ότι για τα μη-κυριαρχούμενα άτομα αποθηκευμένα στο αρχείο ισχύει  $s_i = 0$ . Η τελική τιμή  $\Phi_i$  λαμβάνεται ενιαία ως:

$$\Phi_i = s_i + d_i \quad (2.16)$$

όπου  $d_i$  είναι ένας όρος μικρότερος της μονάδας που ονομάζεται πυκνότητα (density) και αντιπροσωπεύει πόσο πυκνοκατοικημένη είναι η περιοχή γύρω από το άτομο  $i$ . Άτομα που είναι απομονωμένα λαμβάνουν χαμηλότερη τιμή πυκνότητας, άρα ευνοούνται κατά την επιλογή γονέων ώστε να εξερευνηθεί καλύτερα το εγγύς τους χωρίο, αλλά το  $s_i$  λαμβάνει γενικά μεγαλύτερες τιμές από το  $d_i$  οπότε και διαδραματίζει πολύ πιο καθοριστικό ρόλο.

### 2.3.3 Προβλήματα με Υψηλό Κόστος

Το κύριο αρνητικό χαρακτηριστικό των ΕΑ είναι ο μεγάλος αριθμός αξιολογήσεων που απαιτούν. Το υπολογιστικό κόστος ανά γενιά μπορεί να θεωρηθεί πρακτικά ίσο με  $l$  κλήσεις του λογισμικού αξιολόγησης, μία για κάθε απόγονο, αφού οι υπόλοιπες διαδικασίες ενός ΕΑ δεν είναι ιδιαίτερα απαιτητικές και δεν χρειάζεται να προσμετρηθούν. Ακόμα και αν θεωρηθεί εξασφαλισμένο ότι εν τέλει θα βρεθεί το ολικό ακρότατο της συνάρτησης-στόχου μετά από  $g$  γενιές, το συνολικό κόστος  $lg$  αξιολογήσεων μπορεί να είναι απαγορευτικό αναλόγως των απαιτήσεων του λογισμικού αξιολόγησης σε χρόνο και CPU. Θεωρώντας ότι η επέμβαση στο λογισμικό δεν είναι δυνατή λόγω συνθηκών black-box, η μόνη εφικτή λύση είναι η μείωση των απαιτούμενων κλήσεων του ανά γενιά, χωρίς όμως να δημιουργούνται άλλα θέματα ή επιπλοκές κατά τη λειτουργία του αλγορίθμου.

Μια σύγχρονη κλάση ΕΑ που δημιουργήθηκε για να αντιμετωπίσει το παραπάνω πρόβλημα είναι οι MAEA (Metamodel-Assisted Evolutionary Algorithms) [16], [17], [18]. Πρόκειται για ΕΑ που λειτουργούν με τις ίδιες ακριβώς τεχνικές επιλογής και αναπαραγωγής που έχουν ήδη θιχτεί, αλλά επιπλέον κάνουν χρήση μεταμοντέλων ικανών να προβλέψουν την τιμή της συνάρτησης-στόχου των υποψήφιων λύσεων. Ένα μεταμοντέλο μπορεί, χάριν απλότητας, να θεωρηθεί ως μία συνάρτηση με διάφορα βάρη τα οποία ρυθμίζονται ώστε να οδηγούν σε τελική τιμή παραπλήσια της συνάρτησης-στόχου. Προκειμένου να είναι δυνατή μια καλή προσέγγιση χρειάζεται μια βάση δεδομένων που να εμπεριέχει ζεύγη ήδη αξιολογημένων ατόμων και του αποτελέσματος της ακριβούς αξιολόγησής τους (exact evaluation) από το λογισμικό. Το μεταμοντέλο χρησιμοποιεί τα δεδομένα της βάσης για να εκπαιδευτεί στο να μπορεί να πραγματοποιήσει προσεγγιστικές αξιολογήσεις (inexact evaluations) των μελλοντικών υποψήφιων λύσεων, οι οποίες είναι πολύ πιο φτηνές υπολογιστικά από τις ακριβείς, μειώνοντας έτσι δραματικά το συνολικό κόστος της βελτιστοποίησης.

Η λειτουργία ενός MAEA μπορεί να χωριστεί σε δύο στάδια [16], [17]. Το πρώτο στάδιο αφορά την πλήρωση της βάσης δεδομένων με αρκετές εισόδους ώστε τα μεταμοντέλα που θα δημιουργηθούν κατά την εκτέλεση να είναι αποτελεσματικά. Σε περίπτωση που έχουν γίνει ήδη ακριβείς αξιολογήσεις στο παρελθόν με το ίδιο λογισμικό, τότε μπορούν να προστεθούν απευθείας στη βάση πριν την εκτέλεση του MAEA και αν κριθούν αρκετές για τη ρύθμιση των μεταμοντέλων, το στάδιο αυτό μπορεί να παραλειφθεί εντελώς. Διαφορετικά ο MAEA συμπεριφέρεται κατά τις πρώτες γενιές

ακριβώς όπως ένας συνηθισμένος ΕΑ, αξιολογώντας κάθε νέο υποψήφιο με το λογισμικό αξιολόγησης και τοποθετώντας τον στη βάση δεδομένων μαζί με την τιμή της συνάρτησης-στόχου που του αντιστοιχεί. Με το πέρας της πρώτης φάσης, ξεκινά το δεύτερο στάδιο. Για κάθε απόγονο που προκύπτει κατά το στάδιο της αναπαραγωγής δημιουργείται και ένα τοπικό μεταμοντέλο που εκπαιδεύεται βάσει παλιών, κοντινών υποψήφιας της βάσης δεδομένων. Με αυτόν τον τρόπο δημιουργούνται λ εξατομικευμένα μεταμοντέλα, ικανά να δώσουν μια προσεγγιστική αξιολόγηση των λ απογόνων. Στη συνέχεια, γίνεται μια κατάταξη του πληθυσμού σύμφωνα με τα αποτελέσματα και τα βέλτιστα άτομα τροφοδοτούνται στο λογισμικό αξιολόγησης ώστε να επαληθευθούν τα χαρακτηριστικά τους. Ο χρήστης μπορεί να καθορίσει πόσα άτομα θα αξιολογηθούν με ακρίβεια, έστω  $\lambda_e$  στον αριθμό, αλλά πρέπει να ισχύει  $\lambda_e \ll \lambda$  ώστε η χρήση μεταμοντέλων να μειώνει σημαντικά το κόστος. Τα ακριβή αποτελέσματα τοποθετούνται στη βάση δεδομένων ώστε να εμπλουτίζεται συνεχώς για μελλοντική χρήση και τελικά ενεργοποιείται ο τελεστής επιλογής γονέων του MAEA με τις ανανεωμένες τιμές των  $\lambda_e$  απογόνων. Η μεθοδολογία αυτή ονομάζεται προ-αξιολόγηση χαμηλού κόστους (Low-Cost Pre-Evaluation, LCPE) και είναι πολύ αποτελεσματική με βάση αποτελέσματα από δίκτυα ακτινικών συναρτήσεων βάσης (Radial Basis Function Networks, RBFN) ως μεταμοντέλα [19], [20], [21], ενώ μπορεί να λειτουργήσει πολύ αποδοτικά και σε προβλήματα MOO [22]. Υπάρχουν άλλες μέθοδοι που χρησιμοποιούν ένα καθολικό μεταμοντέλο για να αξιολογήσουν απογόνους από ολόκληρο το χωρίο σχεδιασμού, αλλά απαιτούν πολύ μεγαλύτερη προσοχή κατά την εκπαίδευσή τους.

## **ΚΕΦΑΛΑΙΟ 3**

### **ΔΕΙΓΜΑΤΟΛΗΨΙΑ ΜΕΣΩ ΚΑΝΟΝΙΚΗΣ ΚΑΤΑΝΟΜΗΣ**

Έχοντας πλέον καλύψει τις γενικές αρχές βελτιστοποίησης με χρήση ΕΑ, πρέπει να γίνει ιδιαίτερος λόγος για τις διεργασίες των ΕΣ όπου ανήκει και ο αλγόριθμος CMA. Στην Υποενότητα 2.2.2 αναφέρθηκε ότι οι ΕΣ λαμβάνουν δείγματα από την κανονική κατανομή για την αναζήτηση νέων υποψήφιων λύσεων και υπάρχει πληθώρα λόγων για τους οποίους γίνεται αυτό. Από φιλοσοφικής άποψης, είναι σίγουρα πρέπον αλγόριθμοι που βασίζονται σε φυσικές λειτουργίες να χρησιμοποιούν την κανονική κατανομή για τη μετάλλαξη των μεταβλητών σχεδιασμού, αφού ως γνωστό είναι η κατανομή χαρακτηριστικών που παρατηρείται ευρύτερα από όλες στη φύση. Επιπρόσθετα, είναι πολύ εύκολο προγραμματιστικά να ληφθεί δείγμα από μία κανονική κατανομή, το οποίο βοηθά στο πρακτικό τμήμα της υλοποίησης των ΕΣ. Το κύριο χαρακτηριστικό της όμως που αφορά άμεσα έναν αλγόριθμο αναζήτησης, συμπεριλαμβανομένων και των ΕΑ, είναι η ισοτροπία της. Η συμμετρία που επιδεικνύει δεν ευνοεί καμία κατεύθυνση για την έρευνα και έτσι η δειγματοληψία μέσω κανονικής κατανομής στερείται πάντα μεροληψίας (bias) <sup>[23]</sup>, η οποία μπορεί και προστίθεται ξεχωριστά σε άλλα σημεία του αλγορίθμου.

Στο σύντομο αυτό κεφάλαιο θα γίνει μια επισκόπηση των βασικών στατιστικών μεγεθών που απαιτούνται για να οριστεί μία κανονική κατανομή καθώς και των γενικών μαθηματικών της ιδιοτήτων. Οι γνώσεις αυτές είναι θεμελιώδεις για την κατανόηση και υλοποίηση οποιασδήποτε ΕΣ και πρέπει να παρουσιαστούν προκειμένου ο αναγνώστης να μπορεί να παρακολουθήσει, στη συνέχεια, τον τρόπο λειτουργίας του τροποποιημένου αλγορίθμου CMA-ES.

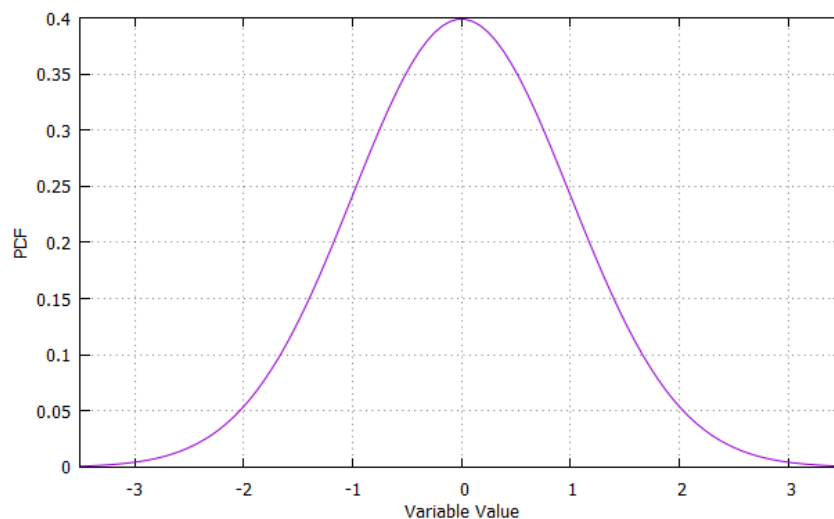
#### **3.1 ΤΥΠΟΠΟΙΗΜΕΝΗ ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ**

Μια μονοδιάστατη κανονική κατανομή συμβολίζεται στη βιβλιογραφία ως  $N(\mu, \sigma^2)$ . Οι δύο ποσότητες από τις οποίες εξαρτάται είναι βαθμωτές και ονομάζονται μέση τιμή  $\mu$  και διακύμανση  $\sigma^2$ . Η μέση τιμή εκφράζει την καλύτερη εκτίμηση ή την πιθανότερη τιμή της μεταβλητής που εξετάζεται, ενώ η διακύμανση εκφράζει τη μέση απόκλιση ενός δείγματος από τη μέση τιμή <sup>[24]</sup>.

Για να γίνει ξεκάθαρο το νόημα των μεγεθών αυτών, ας εξεταστεί αρχικά η μονοδιάστατη κανονική κατανομή με μηδενική μέση τιμή και μοναδιαία διακύμανση, ευρέως γνωστή ως τυποποιημένη κανονική κατανομή  $N(0,1)$ . Η μορφή της συνάρτησης πυκνότητας πιθανότητας (Probability Density Function, PDF) για οποιαδήποτε  $N(\mu, \sigma^2)$  βασίζεται στη μέση τιμή της  $\mu$  και στην τετραγωνική ρίζα της διακύμανσης  $\sigma$ , η οποία ονομάζεται τυπική απόκλιση. Ο γενικός τύπος είναι <sup>[24]</sup>:

$$\text{PDF}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5\left(\frac{x-\mu}{\sigma}\right)^2} \quad (3.1)$$

Στην περίπτωση της  $N(0,1)$  προφανώς ισχύει  $\mu = 0$  και  $\sigma = \sigma^2 = 1$ . Η απεικόνιση της PDF δίνεται στο Σχήμα 3.1.



**Σχήμα 3.1:** Συνάρτηση πυκνότητας πιθανότητας της τυποποιημένης κανονικής κατανομής  $N(0,1)$ .

Στο Σχήμα 3.1 διακρίνεται ξεκάθαρα ότι η πιθανότερη τιμή ενός τυχαίου δείγματος από οποιαδήποτε κανονική κατανομή ταυτίζεται με τη μέση τιμή  $\mu$ , στην προκειμένη περίπτωση με το μηδέν. Οι PDF έχουν επίσης μια πολύ χρήσιμη για τη στατιστική ανάλυση ιδιότητα: το ορισμένο ολοκλήρωμα μεταξύ δύο τιμών ισούται με την πιθανότητα ένα τυχαίο δείγμα να έχει τιμή μεταξύ τους. Οι βασικότερες πιθανότητες της μονοδιάστατης κανονικής κατανομής είναι <sup>[25]</sup>:

- Πιθανότητα τιμής δείγματος μεταξύ  $\mu - \sigma$  και  $\mu + \sigma \cong 68.27\%$
- Πιθανότητα τιμής δείγματος μεταξύ  $\mu - 2\sigma$  και  $\mu + 2\sigma \cong 95.45\%$
- Πιθανότητα τιμής δείγματος μεταξύ  $\mu - 3\sigma$  και  $\mu + 3\sigma \cong 99.73\%$

Θεωρητικά όλες οι κανονικές κατανομές μπορούν να δώσουν τιμή δείγματος από  $-\infty$  έως  $+\infty$ , αλλά οι πιθανότητες η τιμή αυτή να ξεφύγει από το  $\mu \pm 3\sigma$  είναι λιγότερες του 1% και μειώνονται εκθετικά όσο απομακρύνονται από τη μέση τιμή, οπότε μπορούν να θεωρηθούν ως ασήμαντες.

Η  $N(0,1)$  είναι αναμφισβήτητα η βασικότερη κανονική κατανομή από υπολογιστική αλλά και στατιστική άποψη. Υπολογιστικά διότι στη βιβλιογραφία υπάρχουν δεκάδες αλγόριθμοι που εξασφαλίζουν αμερόληπτη δειγματοληψία από την  $N(0,1)$ , ενδεικτικά η μέθοδος Box-Muller Transform <sup>[26]</sup>, ενώ στατιστικά επειδή ένα τυχαίο δείγμα της μπορεί να αναχθεί σε ένα τυχαίο δείγμα οποιασδήποτε άλλης  $N(\mu, \sigma^2)$ . Συμβολίζοντας την ισότητα στην κατανομή ως  $\sim$  ισχύει:

$$N(\mu, \sigma^2) \sim \mu + \sigma N(0,1) \quad (3.2)$$

Ας συγκριθεί η Εξίσωση (3.2) με την Εξίσωση (2.3) της Υποενότητας 2.2.2 η οποία και χρησιμοποιείται για τη δημιουργία απογόνων μίας γενικευμένης ES. Η Εξίσωση (2.3) αποτελεί εφαρμογή της (3.2) για ένα μονοδιάστατο πρόβλημα βελτιστοποίησης και γίνεται επομένως αντιληπτό ότι οι ES πραγματοποιούν δειγματοληψία κανονικών κατανομών με μέση τιμή τον απόγονο που δημιουργήθηκε από τον ανασυνδυασμό και τυπική απόκλιση ίση με τη δύναμη μετάλλαξης. Μάλιστα, ο τρόπος υπολογισμού και ανανέωσης των δύο αυτών παραμέτρων από γενιά σε γενιά είναι το στοιχείο που διαφοροποιεί κατά κύριο λόγο τις ES μεταξύ τους.

Τέλος, πρέπει να γίνει λόγος για μια ιδιότητα που θα φανεί ιδιαίτερα χρήσιμη κατά την υλοποίηση του αλγορίθμου CMA στη συνέχεια. Έστω ότι λαμβάνονται δύο ανεξάρτητα δείγματα  $z_1$  και  $z_2$  από δύο διαφορετικές, μονοδιάστατες κανονικές κατανομές ώστε  $z_1 \sim N(\mu_1, \sigma_1^2)$  και  $z_2 \sim N(\mu_2, \sigma_2^2)$ . Για το γραμμικό συνδυασμό τους ισχύει:

$$c_1 z_1 + c_2 z_2 \sim N(c_1 \mu_1 + c_2 \mu_2, c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2) \quad (3.3)$$

Εφαρμόζοντας την Εξίσωση (3.3) για την περίπτωση δύο ανεξάρτητων δειγμάτων από την ίδια μονοδιάστατη κανονική κατανομή  $N(\mu, \sigma^2)$ , το τελικό άθροισμά τους θα ανήκει στη νέα κανονική κατανομή  $N([c_1 + c_2]\mu, [c_1^2 + c_2^2]\sigma^2)$ , π.χ. για απλή άθροιση ( $c_1 = c_2 = 1$ ) προκύπτει τελικά η κανονική κατανομή  $N(2\mu, 2\sigma^2)$ . Η αλλαγή των χαρακτηριστικών μεγεθών της κανονικής κατανομής κατά την άθροιση δειγμάτων πρέπει να λαμβάνεται πάντα υπόψη, αλλιώς είναι αναμενόμενο να οδηγηθεί κανείς σε πολύ σοβαρά σφάλματα κατά την ανάπτυξη θεωριών στοχαστικών μεθόδων όπως είναι ο CMA.

### 3.2 ΠΟΛΥΔΙΑΣΤΑΤΗ ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ

Για ένα πρόβλημα βελτιστοποίησης που εξαρτάται από  $n$  μεταβλητές σχεδιασμού, η δειγματοληψία πρέπει να γίνει από  $n$ -διάστατες κανονικές κατανομές. Μια πολυδιάστατη κανονική κατανομή συμβολίζεται ως  $N(\vec{m}, \mathbf{C})$  και για να οριστεί πλήρως απαιτείται η μέση διανυσματική της τιμή  $\vec{m}$  διαστάσεων  $[n \times 1]$ , καθώς και ένα μητρώο συνδιακύμανσης  $\mathbf{C}$  διαστάσεων  $[n \times n]$ . Το  $\mathbf{C}$  εμπεριέχει στη διαγώνιό του όρους διακύμανσης  $\sigma_{x_i}^2$  για κάθε μεταβλητή σχεδιασμού, η τετραγωνική ρίζα των οποίων δίνει τις τυπικές αποκλίσεις  $\sigma_{x_i}$ . Οι νέοι όροι που βρίσκονται εκτός της διαγώνιου ονομάζονται συνδιακυμάνσεις, συμβολίζονται ως  $\sigma_{x_i x_j}$  και εκφράζουν πόσο η αλλαγή της μεταβλητής  $x_i$  επηρεάζει τη μεταβλητή  $x_j$  [24]. Εξ ορισμού ισχύει  $\sigma_{x_i x_j} = \sigma_{x_j x_i}$ , οπότε ένα μητρώο συνδιακύμανσης  $\mathbf{C}$  είναι πάντα συμμετρικό και έχει τη μορφή:

$$\mathbf{C} = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & \cdots & \sigma_{x_1 x_{n-1}} & \sigma_{x_1 x_n} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 & \cdots & \sigma_{x_2 x_{n-1}} & \sigma_{x_2 x_n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{x_{n-1} x_1} & \sigma_{x_{n-1} x_2} & \cdots & \sigma_{x_{n-1}}^2 & \sigma_{x_{n-1} x_n} \\ \sigma_{x_n x_1} & \sigma_{x_n x_2} & \cdots & \sigma_{x_n x_{n-1}} & \sigma_{x_n}^2 \end{bmatrix} \quad (3.4)$$

Επιπλέον κάθε μητρώο  $\mathbf{C}$  είναι θετικά ορισμένο, δηλαδή μπορεί να γραφεί στη μορφή [5]:

$$\mathbf{C} = \mathbf{B}\mathbf{D}^2\mathbf{B}^T \quad (3.5)$$

όπου  $\mathbf{B}$  είναι ένα μητρώο διαστάσεων  $[n \times n]$  που έχει ως στήλες τα ιδιοδιανύσματα του  $\mathbf{C}$  και  $\mathbf{D}$  είναι ένα διαγώνιο μητρώο διαστάσεων  $[n \times n]$  με στοιχεία τις ρίζες των ιδιοτιμών του  $\mathbf{C}$ . Βάσει της Εξίσωσης (3.5) προκύπτει ότι:

$$\mathbf{C}^{-1} = \mathbf{B}\mathbf{D}^{-2}\mathbf{B}^T \quad (3.6)$$

Τέλος, οι ισοϋψείς καμπύλες της PDF μιας πολυδιάστατης κανονικής κατανομής έχουν, στη γενική περίπτωση, μορφή ομόκεντρων ελλείψεων με κέντρο τη μέση τιμή  $\vec{m}$  που μπορούν να υπολογιστούν από τη σχέση:

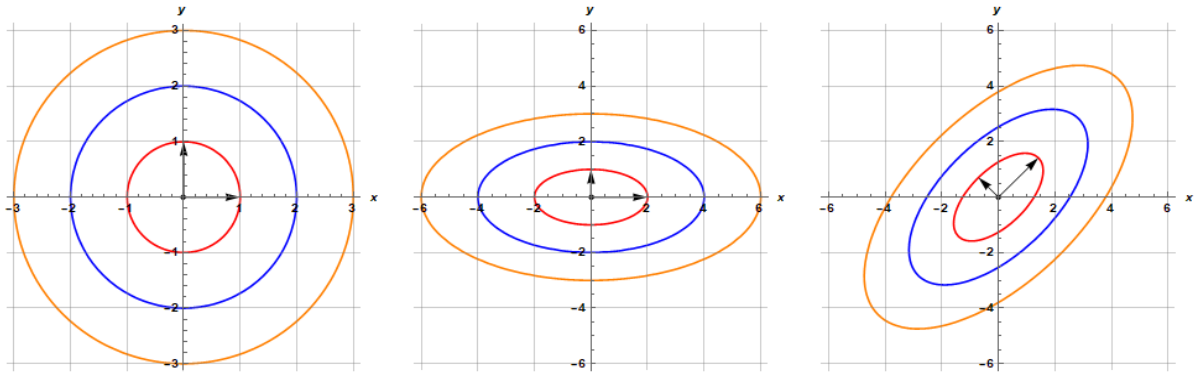
$$\vec{x}^T \mathbf{C}^{-1} \vec{x} = \sigma^2 \quad (3.7)$$

όπου το  $\sigma$  αντιπροσωπεύει πόσες τυπικές αποκλίσεις μακριά από το  $\vec{m}$  θα οριοθετηθεί το δείγμα.

Προκειμένου να κατανοηθεί πλήρως η έννοια του μητρώου συνδιακύμανσης  $\mathbf{C}$  και πώς αυτό επηρεάζει μορφολογικά μία πολυδιάστατη κανονική κατανομή  $N(\vec{m}, \mathbf{C})$ , στο



Σχήμα 3.2 παρουσιάζονται τρία παραδείγματα διδιάστατων κανονικών κατανομών  $N(\vec{0}, \mathbf{C})$  με την ίδια μέση διανυσματική τιμή ( $\vec{\mu} = \vec{0}$ ), αλλά με διαφορετικά μητρώα συνδιακύμανσης  $\mathbf{C}_i$ , εφαρμόζοντας την Εξίσωση (3.7) για  $\sigma = 1, 2, 3$ .



**Σχήμα 3.2:** Ισοϋψείς της συνάρτησης πυκνότητας πιθανότητας διδιάστατων κανονικών κατανομών  $N(\vec{0}, \mathbf{C}_i)$ , όπου αριστερά τίθεται  $\mathbf{C}_1 = \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ , στο κέντρο  $\mathbf{C}_2 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$  και δεξιά  $\mathbf{C}_3 = \begin{bmatrix} 2.5 & 1.5 \\ 1.5 & 2.5 \end{bmatrix}$ . Τα διανύσματα απεικονίζουν τους κύριους άξονες των ελλείψεων και ισούνται με τα ιδιοδιανύσματα κάθε μητρώου  $\mathbf{C}_i$ , πολλαπλασιασμένα με τη ρίζα της αντίστοιχης ιδιοτιμής. Το μητρώο  $\mathbf{C}_3$  προέκυψε περιστρέφοντας το  $\mathbf{C}_2$  αριστερόστροφα κατά  $45^\circ$ , οπότε έχουν ίσες ιδιοτιμές (μέτρο) αλλά διαφορετικά ιδιοδιανύσματα (διεύθυνση).

Η  $N(\vec{0}, \mathbf{I})$  (αριστερά στο Σχήμα 3.2) είναι η αντίστοιχη της  $N(0,1)$  σε δύο διαστάσεις. Οι ισοϋψείς της είναι ομόκεντροι κύκλοι διότι οι τυπικές αποκλίσεις  $\sigma_x$  και  $\sigma_y$  είναι μεταξύ τους ίσες. Ένα δείγμα προερχόμενο από οποιαδήποτε  $N(\vec{0}, \mathbf{I})$  διαστάσεων  $n$  είναι προφανώς ένα διάνυσμα με  $n$  στοιχεία και μπορεί να ληφθεί πραγματοποιώντας  $n$  ξεχωριστές δειγματοληψίες από τη μονοδιάστατη  $N(0,1)$ , μία για κάθε στοιχείο. Η  $N(\vec{0}, \mathbf{C}_2)$  (κεντρικά στο Σχήμα 3.2) διαφέρει συγκριτικά με την  $N(\vec{0}, \mathbf{I})$  μόνο στην τυπική διακύμανση  $\sigma_x$ , η οποία είναι διπλάσια και για αυτό οι ισοϋψείς της έχουν ελλειπτική μορφή. Έχοντας επομένως ένα δείγμα από την  $N(\vec{0}, \mathbf{I})$ , αρκεί να πολλαπλασιαστεί με το μητρώο  $\mathbf{D}_2$  ώστε το πρώτο στοιχείο του που αντιστοιχεί στη μεταβλητή  $x$  να διπλασιαστεί και το δείγμα έχει πλέον αναχθεί στην  $N(\vec{0}, \mathbf{C}_2)$ .

Η  $N(\vec{0}, \mathbf{C}_3)$  (δεξιά στο Σχήμα 3.2) αποτελεί τη γενική περίπτωση μιας διδιάστατης κανονικής κατανομής με πλήρες μητρώο συνδιακύμανσης. Αν και τα μητρώα  $\mathbf{C}_2$  και  $\mathbf{C}_3$

διαφέρουν πολύ φαινομενικά, οι ελλείψεις τους έχουν την ίδια ακριβώς μορφή αλλά διαφορετική διεύθυνση. Αυτό οφείλεται στους θετικούς όρους συνδιακύμανσης  $\sigma_{xy} = \sigma_{yx} > 0$  που πρεσβεύουν ότι όταν η μεταβλητή  $x$  αυξάνεται, την ίδια τάση έχει και η  $y$  και αντίστροφα. Αρνητικοί όροι συνδιακύμανσης σημαίνουν ότι η αύξηση της μίας μεταβλητής προκαλεί μείωση της άλλης, ενώ μηδενικοί όροι αντιπροσωπεύουν πλήρη ανεξαρτησία των μεταβλητών. Ένα δείγμα από την  $N(\vec{0}, \mathbf{C}_2)$  μπορεί να αναχθεί σε δείγμα της  $N(\vec{0}, \mathbf{C}_3)$  άμα περιστραφεί, το οποίο επιτυγχάνεται πολλαπλασιάζοντας με το μητρώο  $\mathbf{B}_3$  της  $\mathbf{C}_3$ .

Συνοψίζοντας όλους τους μετασχηματισμούς που αναφέρθηκαν παραπάνω σε μία σχέση, προκύπτει ότι για οποιαδήποτε πολυδιάστατη κανονική κατανομή  $N(\vec{m}, \mathbf{C})$  ισχύει:

$$N(\vec{m}, \mathbf{C}) \sim \vec{m} + \mathbf{BD}N(\vec{0}, \mathbf{I}) \quad (3.8)$$

Τονίζεται, τέλος, ότι καθώς οι διαστάσεις αυξάνονται, οι πιθανότητες ένα δείγμα να βρίσκεται κοντά στη μέση τιμή μειώνονται σημαντικά. Ενδεικτικά αναφέρεται ότι, ενώ η πιθανότητα το δείγμα να βρίσκεται μεταξύ  $\mu \pm \sigma$  στη μία διάσταση είναι περίπου 68.27%, στις δύο διαστάσεις η πιθανότητα το διάνυσμα ενός δείγματος να βρίσκεται εντός της έλλειψης που προκύπτει για  $\sigma = 1$  στην Εξίσωση (3.7) είναι 39.35%, για τρεις διαστάσεις 19.87%, για πέντε διαστάσεις 3.74% και για δέκα διαστάσεις οι πιθανότητες είναι μόλις 0.02% [25].

## ΚΕΦΑΛΑΙΟ 4

### Ο ΤΡΟΠΟΠΟΙΗΜΕΝΟΣ ΑΛΓΟΡΙΘΜΟΣ CMA

Η μέθοδος CMA δημιουργήθηκε από τους Nikolaus Hansen και Andreas Ostermeier το 1996 <sup>[7]</sup>. Έως τότε οι ES συνήθιζαν να χρησιμοποιούν ισοτροπικές κανονικές κατανομές της μορφής  $N(\vec{m}, \sigma^2 \mathbf{I})$  για τη δειγματοληψία, δηλαδή γινόταν η υπόθεση ότι οι μεταβλητές σχεδιασμού ήταν πλήρως ανεξάρτητες μεταξύ τους. Η CMA-ES αποτελεί μία από τις πρώτες μεθόδους που, όπως μαρτυρά άλλωστε και το όνομά της, έχει ως στόχο τον υπολογισμό του πλήρους μητρώου συνδιακύμανσης, λαμβάνοντας δείγματα από κανονικές κατανομές της μορφής  $N(\vec{m}, \sigma^2 \mathbf{C})$ . Πέραν αυτού, χαρακτηρίζεται από δύο καινοτόμες για την εποχή της έννοιες: την αποτυχαιοποίηση (derandomization) και τη συσσώρευση (cumulation) <sup>[13]</sup>. Στη συνέχεια, παρουσιάζεται βήμα προς βήμα ο τρόπος λειτουργίας ενός τροποποιημένου αλγόριθμου CMA με πραγματική κωδικοποίηση που χρησιμοποιεί στοιχεία εμπνευσμένα από το σύνολο ολόκληρης της θεωρίας των ΕΑ, ενώ συζητείται παράλληλα και η βασική ιδέα πίσω από τις διεργασίες που πραγματοποιούνται.

#### 4.1 ΔΕΙΓΜΑΤΟΛΗΨΙΑ & ΑΞΙΟΛΟΓΗΣΗ

Επεκτείνοντας την Εξίσωση (3.8) του προηγούμενου κεφαλαίου που αφορά κανονικές κατανομές της μορφής  $N(\vec{m}, \mathbf{C})$  σε μία κανονική κατανομή  $N(\vec{m}, \sigma^2 \mathbf{C})$  είναι:

$$N(\vec{m}, \sigma^2 \mathbf{C}) \sim \vec{m} + \sigma \mathbf{B} \mathbf{D} N(\vec{0}, \mathbf{I}) \quad (4.1)$$

Παρατηρείται ότι για τη δημιουργία νέων απογόνων κατά τη γενιά  $g + 1$  χρειάζεται να είναι ήδη γνωστή η μέση τιμή  $\vec{m}^{(g)}$ , το μητρώο συνδιακύμανσης  $\mathbf{C}^{(g)}$  και το μέγεθος του βήματος  $\sigma^{(g)}$  της προηγούμενης γενιάς  $g$ . Το βαθμωτό μέγεθος  $\sigma$  λειτουργεί ως η «γενική κλίμακα» της κατανομής και προστίθεται στη μέθοδο διότι το  $\mathbf{C}$  από μόνο του δεν μπορεί να προσαρμόσει την έρευνα στο χωρίο σχεδιασμού με βέλτιστο τρόπο <sup>[5]</sup>. Θεωρώντας λοιπόν ότι όλα τα παραπάνω είναι ήδη αποθηκευμένα, εφαρμόζεται αρχικά ιδιοανάλυση στο  $\mathbf{C}^{(g)}$  ώστε να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματά του, εκ των οποίων δημιουργούνται τα μητρώα  $\mathbf{D}^{(g)}$  και  $\mathbf{B}^{(g)}$  αντίστοιχα. Οι  $\lambda$  νέοι απόγονοι λαμβάνονται ως:

$$\vec{X}_i^{(g+1)} = \vec{m}^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{D}^{(g)} \vec{z}_i \quad (4.2)$$

όπου το διάνυσμα  $\vec{z}_i \sim N(\vec{0}, \mathbf{I})$  έχει ως στοιχεία τυχαία δείγματα από την τυποποιημένη κανονική κατανομή, όπως ορίστηκε στην Εξίσωση (2.4) της Υποενότητας 2.2.2.

Κατά τη δημιουργία ενός διανύσματος  $\vec{z}_i$ , οι πιθανότητες ένα στοιχείο του να πάρει τιμή εκτός του διαστήματος  $[-2, +2]$  είναι μικρότερες του 5%. Αυτό σημαίνει ότι η έρευνα είναι σχετικά περιορισμένη στην εγγύς περιοχή της μέσης τιμής, ανάλογα βέβαια τις διαστάσεις του προβλήματος και την τρέχουσα τιμή του βήματος. Σε περίπτωση που λύνεται ένα δύσκολο πρόβλημα βελτιστοποίησης όπου δεν βρίσκονται νέες καλύτερες λύσεις κατά τις πρώτες γενιές, υπάρχει υψηλός κίνδυνος εγκλωβισμού σε τοπικό ακρότατο διότι η τιμή του βήματος θα μειωθεί ραγδαία και οι ελλείψεις της  $N(\vec{m}, \sigma^2 \mathbf{C})$  θα περιορίσουν την έρευνα σε πολύ μικρά χωρία.

Για να αποφευχθεί αυτό εφαρμόζεται ένας νέος τελεστής μετάλλαξης στους απογόνους μέσω των  $\vec{z}_i$ , ο οποίος βασίζεται μεν στις αρχές των GA για μεγάλες και σημαντικές μεταλλάξεις στα χαρακτηριστικά του πληθυσμού αλλά σέβεται παράλληλα το καθεστώς χρήσης κανονικών κατανομών των ES. Ο χρήστης ορίζει μία μικρή πιθανότητα για την ενεργοποίηση του τελεστή, η οποία ελέγχεται προτού οριστεί κάθε στοιχείο διανύσματος  $\vec{z}_i$ . Άμα ενεργοποιηθεί, η τιμή του στοιχείου τίθεται απευθείας ίση με  $\pm 3$  χωρίς να ληφθεί δείγμα από την  $N(0,1)$ . Η επιλογή πρόσημου εξαρτάται από την τιμή του αντίστοιχου στοιχείου του διανύσματος  $\vec{p}_c^{(g)}$  που αντιπροσωπεύει την κατεύθυνση της μετατόπισης της μέσης τιμής από γενιά σε γενιά και θα αναλυθεί εκτενέστερα παρακάτω. Όταν το στοιχείο  $i$  του  $\vec{p}_c^{(g)}$  είναι θετικό σημαίνει ότι η μεταβλητή σχεδιασμού  $x_i$  έχει την τάση να αυξάνεται κατά την εξέλιξη, οπότε μπορεί να γίνει η υπόθεση ότι η βέλτιστη τιμή της βρίσκεται προς τα θετικά και το στοιχείο του  $\vec{z}_i$  λαμβάνει τιμή  $+3$ . Το αντίστροφο ισχύει για την αρνητική κατεύθυνση, ενώ στην σπάνια περίπτωση που η μέση τιμή  $\bar{x}_i$  έχει παραμείνει σταθερή για μεγάλο χρονικό διάστημα και το στοιχείο  $i$  του  $\vec{p}_c^{(g)}$  είναι μηδενικό, τότε το πρόσημο τίθεται τυχαία. Με αυτόν τον τρόπο, όχι μόνο αποφεύγεται ο εγκλωβισμός σε τοπικό ακρότατο, αλλά επιταχύνεται σημαντικά και ο ρυθμός σύγκλισης του αλγόριθμου.

Αφού έχει ολοκληρωθεί η δημιουργία τους, οι  $\lambda$  απόγονοι αξιολογούνται από το λογισμικό αξιολόγησης. Στον τροποποιημένο CMA προστίθεται επιπλέον η δυνατότητα προ-αξιολόγησης χαμηλού κόστους των απογόνων με χρήση μεταμοντέλων RBFN, όπως περιγράφεται στην Υποενότητα 2.3.3. Επιπλέον, για προβλήματα βελτιστοποίησης με

περιορισμούς επιβάλλονται εκθετικά αυξανόμενες ποινές σύμφωνα με τις Εξισώσεις (2.7) έως (2.10) της Υποενότητας 2.3.1, ενώ για προβλήματα ΜΟΟ χρησιμοποιείται η μέθοδος SPEA2 που παρουσιάστηκε στην Υποενότητα 2.3.2.

Έχοντας διαθέσιμες τις τελικές τιμές της συνάρτησης-στόχου που αντιστοιχούν σε κάθε υποψήφια λύση, ενεργοποιείται ο επιθυμητός τελεστής επιλογής γονέων από την Υποενότητα 2.2.1 για να βρεθούν οι  $\mu$  γονείς. Η προεπιλεγμένη μέθοδος είναι η κλασική επιλογή με αποκοπή, όπως συνηθίζεται στις ES, αλλά δίνεται και η δυνατότητα να πραγματοποιηθεί επιλογή μέσω διαγωνισμού, αν αυτό είναι επιθυμητό. Ο χρήστης ορίζει τις απαιτούμενες παραμέτρους των διαγωνισμών και η τελική πιθανότητα επιλογής κάθε υποψηφίου προκύπτει από την Εξίσωση (2.2).

#### 4.2 ΥΠΟΛΟΓΙΣΜΟΣ ΤΗΣ ΜΕΣΗΣ ΤΙΜΗΣ

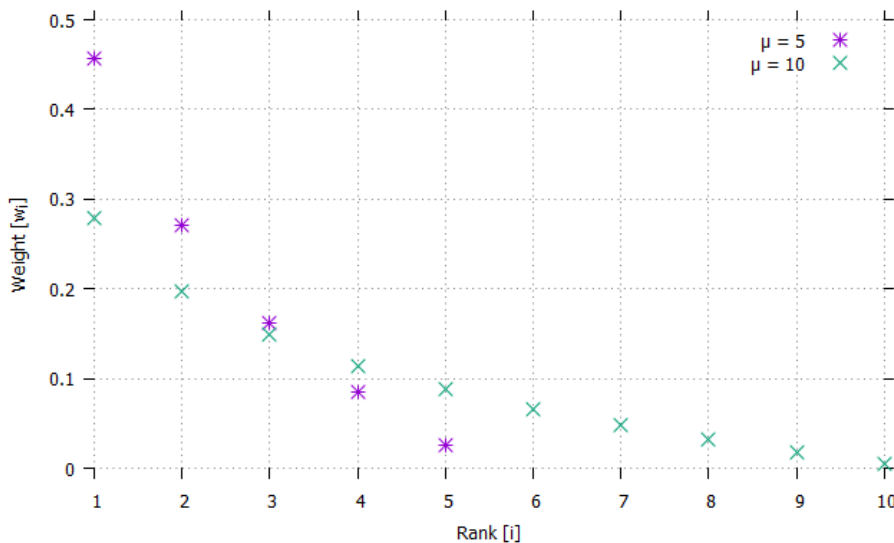
Η μέση τιμή  $\vec{m}^{(g+1)}$  της νέας γενιάς προκύπτει ενεργοποιώντας έναν νέο τελεστή εξέλιξης που αποτελεί γενίκευση του συνεχούς ανασυνδυασμού και ονομάζεται σταθμισμένος πολλαπλός ανασυνδυασμός (weighted multi-recombination) [9]. Για να λειτουργήσει, πρέπει αρχικά να οριστούν από το χρήστη  $\mu$  συνολικά βάρη  $w_i$  που συνήθως είναι κανονικοποιημένα, δηλαδή ισχύει  $\sum_{i=1}^{\mu} w_i = 1$ . Στην περίπτωση που τα βάρη οριστούν ως  $w_i = 1/\mu \forall i$ , ο σταθμισμένος ανασυνδυασμός ταυτίζεται με τον ενδιάμεσο της Υποενότητας 2.2.2 και υπολογίζει την απλή μέση τιμή των γονέων. Η βέλτιστη τιμή των  $w_i$  εξαρτάται από το πρόβλημα που εξετάζεται και ο Hansen προτείνει τον τύπο [5]:

$$w_i = \frac{\ln(\mu+0.5) - \ln(i)}{\sum_{j=1}^{\mu} [\ln(\mu+0.5) - \ln(j)]} \quad (4.3)$$

ως μια καλή, αρχική πρόβλεψη για πραγματικά προβλήματα βελτιστοποίησης. Θεωρώντας ότι οι γονείς  $\vec{X}_i$  κατατάσσονται από τον καλύτερο στο χειρότερο καθώς ο δείκτης  $i$  αυξάνεται, η Εξίσωση (4.3) παράγει φθίνοντα βάρη όπως φαίνεται στο Σχήμα 4.1. Έτσι, οι καλύτεροι γονείς έχουν και σαφώς μεγαλύτερη επιρροή στον καθορισμό του νέου κέντρου της κανονικής κατανομής. Η νέα μέση τιμή προκύπτει τελικά ως:

$$\vec{m}^{(g+1)} = \sum_{i=1}^{\mu} w_i \vec{X}_i^{(g+1)} \quad (4.4)$$

Λόγω του τρόπου γένεσής τους, υπενθυμίζεται ότι όλα τα διανύσματα  $\vec{X}_i^{(g+1)}$  προκύπτουν από την παλιά μέση τιμή  $\vec{m}^{(g)}$  ασκώντας μια μετάλλαξη όπως φαίνεται από την Εξίσωση (4.2), οπότε η Εξίσωση (4.4) μπορεί να ξαναγραφτεί ως:



**Σχήμα 4.1:** Κανονικοποιημένα βάρη  $w_i$  για  $\mu = 5$  και  $\mu = 10$ . Και στις δύο περιπτώσεις ισχύει  $\sum_{i=1}^{\mu} w_i = 1$ , αλλά διακρίνεται ότι μικρότεροι αριθμοί γονέων έχουν ως αποτέλεσμα μεγαλύτερο ελιτισμό κατά την τέλεση του ανασυνδυασμού λόγω του αυξημένου βάρους που λαμβάνουν οι καλύτεροι γονείς.

$$\bar{\mathbf{m}}^{(g+1)} = \bar{\mathbf{m}}^{(g)} + \sigma^{(g)} \sum_{i=1}^{\mu} w_i \vec{y}_i^{(g+1)} \quad (4.5)$$

όπου:

$$\vec{y}_i^{(g+1)} = \mathbf{B}^{(g)} \mathbf{D}^{(g)} \vec{z}_i \quad (4.6)$$

και είναι εμφανές ότι  $\vec{y}_i^{(g+1)} \sim N(\vec{0}, \mathbf{C}^{(g)})$ . Τα διανύσματα  $\vec{y}_i$  είναι εξαιρετικά μεγάλης σημασίας για τη μέθοδο διότι εμπεριέχουν όλες τις πληροφορίες που αφορούν το μητρώο συνδιακύμανσης και όπως θα φανεί στη συνέχεια χρησιμοποιούνται για την προσαρμογή του στο πρόβλημα που λύνεται. Πολλαπλασιάζοντας με το μέγεθος του βήματος  $\sigma^{(g)}$  επηρεάζεται προφανώς μόνο το μέτρο της μετάλλαξης και όχι η κατεύθυνσή της, οπότε το  $\sigma^{(g)}$  μπορεί να προσαρμοστεί ξεχωριστά και με άλλα κριτήρια όπως θα φανεί παρακάτω.

Ο κλασικός αλγόριθμος CMA σταματά εδώ και χρησιμοποιεί την Εξίσωση (4.5) για τον υπολογισμό της  $\bar{\mathbf{m}}^{(g+1)}$ . Αν και ο εν λόγω τύπος εξασφαλίζει συνήθως ταχύτατη σύγκλιση, αφού η μέση τιμή τείνει να μεταφέρεται στην περιοχή των βέλτιστων λύσεων λόγω επιλογής με αποκοπή, υπάρχει η περίπτωση πρόωρης σύγκλισης σε τοπικό ακρότατο άμα η συνάρτηση-στόχος εμπεριέχει κοιλάδες. Όταν η μέση τιμή εγκλωβιστεί σε μια κοιλάδα, η δειγματοληψία θα εντοπίζει πολλές καλές λύσεις στην εγγύς περιοχή με αποτέλεσμα να μειωθεί το εύρος της αναζήτησης και να τερματιστεί ο αλγόριθμος χωρίς

να εξερευνηθεί το υπόλοιπο χωρίο σχεδιασμού. Με σκοπό να μειωθεί όσο το δυνατόν περισσότερο η πιθανότητα αυτού του φαινομένου, ορίζεται ένας εναλλακτικός τρόπος υπολογισμού <sup>[27]</sup>:

$$\vec{m}^{(g+1)} = \vec{m}^{(g)} + c_m^{(g)} \sigma^{(g)} \sum_{i=1}^{\mu} w_i \vec{y}_i^{(g+1)} \quad (4.7)$$

όπου  $c_m^{(g)} \in [0,1]$  είναι μια παράμετρος χαλάρωσης της μετατόπισης της μέσης τιμής από γενιά σε γενιά. Η Εξίσωση (4.7) έχει ως στόχο τη μείωση της μετακίνησης της  $\vec{m}$  κατά τις πρώτες, αναγνωριστικές με το πρόβλημα γενιές, ώστε να μην μπορεί να οδηγηθεί εύκολα σε περιοχές με τοπικά ακρότατα όπου πιθανότατα θα εγκλωβιζόταν. Καθώς όμως η βελτιστοποίηση προχωράει και οδεύει στο τέλος της, η μέση τιμή οφείλει να μετακινείται με μεγαλύτερη ελευθερία, ειδικά περιορίζεται η αποτελεσματικότητα του αλγορίθμου αναζήτησης. Έτσι, η παράμετρος χαλάρωσης  $c_m$  θα πρέπει να τείνει στη μονάδα με το πέρας γενεών και τελικά υπολογίζεται, ανάλογα με το τρέχον στάδιο της βελτιστοποίησης, ως:

$$c_m^{(g)} = c_m^{(0)} \exp \left[ \ln \left( \frac{1}{c_m^{(0)}} \right) \frac{\text{evals}^{(g)}}{\text{max\_evals}} \right] \quad (4.8)$$

όπου  $\text{evals}^{(g)}$  είναι ο αριθμός των αξιολογήσεων που έχουν πραγματοποιηθεί από το λογισμικό αξιολόγησης έως τη γενιά  $g$  και  $\text{max\_evals}$  είναι ο μέγιστος αριθμός επιτρεπτών αξιολογήσεων, δηλαδή το κριτήριο τερματισμού του αλγόριθμου. Ο χρήστης ορίζει μια αρχική τιμή χαλάρωσης  $c_m^{(0)}$  και η τιμή της παραμέτρου  $c_m^{(g)}$  αυξάνεται σταδιακά μέχρι να φτάσει τη μονάδα στο τέλος της βελτιστοποίησης. Ο ρυθμός αύξησης της παραμέτρου είναι πλήρως τροποποιήσιμος μέσω του αρχείου εισόδου ανάλογα τις απαιτήσεις του προβλήματος. Μετά από δοκιμές <sup>[27]</sup> προέκυψε ότι η εφαρμογή της παραμέτρου χαλάρωσης  $c_m$  οδηγεί σε σημαντική βελτίωση της τελικής, καλύτερης λύσης σε προβλήματα με δύσκολες συναρτήσεις στόχου όπως είναι οι συναρτήσεις Ackley και Rastrigin, αλλά στη γενική περίπτωση η παράμετρος μπορεί να καθυστερήσει αχρείαστα τη σύγκλιση του αλγόριθμου. Στην περίπτωση αυτή μπορεί να τεθεί εξ αρχής  $c_m^{(0)} = 1 =$  σταθερό, οπότε και η Εξίσωση (4.8) ταυτίζεται με την (4.5). Κλείνοντας, αναφέρεται ότι στη γενική περίπτωση η αρχική μέση τιμή  $\vec{m}^{(0)}$  λαμβάνεται κατά την αρχικοποίηση τυχαία εντός του χωρίου σχεδιασμού, ενώ υπάρχει και η δυνατότητα να τεθεί απευθείας από το χρήστη άμα υπάρχει a priori γνώση του προβλήματος.

### 4.3 ΠΡΟΣΑΡΜΟΓΗ ΤΟΥ ΜΗΤΡΩΟΥ ΣΥΝΔΙΑΚΥΜΑΝΣΗΣ

Μέχρι αυτό το σημείο της λειτουργίας του αλγορίθμου, έχει υπολογιστεί το πρώτο απαιτούμενο στοιχείο για τη γένεση της επόμενης γενιάς, το κέντρο της νέας κανονικής κατανομής  $\bar{\mathbf{m}}^{(g+1)}$ . Εκκρεμούν το ανανεωμένο μητρώο συνδιακύμανσης  $\mathbf{C}^{(g+1)}$ , καθώς και το νέο μέγεθος του βήματος  $\sigma^{(g+1)}$  ώστε να μπορεί να οριστεί πλήρως το χωρίο αναζήτησης των νέων λύσεων. Ο τρόπος υπολογισμού των μεγεθών αυτών αποτελεί και το κυριότερο σημείο της μεθόδου, όντας το χαρακτηριστικό που τη διαφοροποιεί από τις υπόλοιπες ES.

#### 4.3.1 Μονοπάτι Εξέλιξης

Η μέθοδος CMA αποτελεί μία στοχαστική μέθοδο βελτιστοποίησης, των οποίων βασικό γνώρισμα είναι ότι δεν είναι αυτοσκοπός η εύρεση καλύτερης λύσης από γενιά σε γενιά, αλλά η βελτίωση σε βάθος χρόνου των χαρακτηριστικών των ατόμων του πληθυσμού. Αυτή η βελτίωση μπορεί να βοηθηθεί σε μεγάλο βαθμό από πληροφορίες προηγούμενων γενιών, οι οποίες σχηματίζουν στο σύνολό τους ένα μονοπάτι εξέλιξης (evolution path). Η ονομασία «μονοπάτι» προκύπτει από την ορολογία των ES, όπου κάθε δείγμα που λαμβάνεται θεωρείται ως ένα πιθανό βήμα της μεθόδου. Κάθε βήμα για να οριστεί πλήρως απαιτεί μια κατεύθυνση και ένα μέτρο. Αντικαθιστώντας την Εξίσωση (4.6) στην (4.2) προκύπτει ότι όλα τα δείγματα έχουν εν γένει τη μορφή:

$$\vec{X}_i^{(g+1)} = \bar{\mathbf{m}}^{(g)} + \sigma^{(g)} \vec{y}_i^{(g+1)} \quad (4.9)$$

και γίνεται ξεκάθαρο ότι τα διανύσματα  $\vec{y}_i^{(g+1)}$  που προκύπτουν από την ιδιοάναλυση του  $\mathbf{C}^{(g)}$  καθορίζουν την κατεύθυνση κάθε μεμονωμένου, πιθανού βήματος  $i$  της γενιάς  $g + 1$ , ενώ το  $\sigma^{(g)}$  καθορίζει το μέτρο όλων των  $\lambda$  πιθανών βημάτων. Το άθροισμα των  $\mu$  πετυχημένων βημάτων, δηλαδή αυτών που γεννούν γονείς μετά την αξιολόγησή τους, είναι αυτό που σχηματίζει τελικά το μονοπάτι της εξέλιξης.

Η διαδικασία υπολογισμού ενός μονοπατιού εξέλιξης ονομάζεται συσσώρευση και αποτελεί το πρώτο χαρακτηριστικό γνώρισμα της μεθόδου CMA [28]. Το μονοπάτι της εξέλιξης  $\vec{p}$  προκύπτει από το άθροισμα των διαδοχικών, πετυχημένων βημάτων της μεθόδου  $\vec{S}$ . Προκειμένου όμως η συσσώρευση να είναι αποτελεσματική, το  $\vec{p}$  και το  $\vec{S}$  πρέπει να ανήκουν στην ίδια κανονική κατανομή, ενώ παράλληλα τα πιο πρόσφατα



βήματα πρέπει να συνεισφέρουν περισσότερο στο σχηματισμό του μονοπατιού εξέλιξης. Λαμβάνοντας υπόψη τα παραπάνω, το τρέχον μονοπάτι στη γενιά  $g + 1$  υπολογίζεται ως:

$$\vec{p}^{(g+1)} = (1 - c)\vec{p}^{(g)} + \sqrt{c(2 - c)}\vec{S}^{(g+1)} \quad (4.10)$$

όπου  $c \in [0,1]$  είναι μία παράμετρος που καθορίζει τη διάρκεια ζωής πληροφοριών από παλιότερες γενιές. Ο όρος  $\sqrt{c(2 - c)}$  αποτελεί έναν παράγοντα κανονικοποίησης και είναι υπολογισμένος ώστε να διατηρεί σταθερή τη διακύμανση της κατανομής σύμφωνα με την Εξίσωση (3.3). Δεδομένου ότι τα διανύσματα  $\vec{p}^{(g)}$  και  $\vec{S}^{(g+1)}$  έχουν την ίδια διακύμανση, ο γραμμικός τους συνδυασμός  $\vec{p}^{(g+1)} = c_1\vec{p}^{(g)} + c_2\vec{S}^{(g+1)}$  με  $c_1 = (1 - c)$  και  $c_2 = \sqrt{c(2 - c)}$  έχει συντελεστή διακύμανσης  $c_1^2 + c_2^2 = (1 - c)^2 + (\sqrt{c(2 - c)})^2 = 1$ . Με αυτόν τον τρόπο εξασφαλίζεται σε όλες τις γενιές ότι  $\vec{p}^{(g+1)} \sim \vec{p}^{(g)}$ , δεδομένου ότι  $\vec{p}^{(g)} \sim \vec{S}^{(g+1)}$  [5], [13]. Η Εξίσωση (4.10) θεωρεί ένα βήμα  $\vec{S}^{(g)}$  ανά γενιά, το οποίο για τη μέθοδο CMA δεν είναι άλλο από το σταθμισμένο άθροισμα των μεταλλάξεων  $\vec{y}_i^{(g+1)}$  που χρησιμοποιήθηκε στην Εξίσωση (4.7) για τον υπολογισμό της νέας μέσης τιμής. Συμβολίζοντας το άθροισμα αυτό ως  $\vec{y}_w$ , το τελικό βήμα του αλγόριθμου CMA σε κάθε γενιά μπορεί να εκφραστεί συναρτήσει της μεταβολής της μέσης τιμής ως:

$$\vec{y}_w^{(g+1)} = \sum_{i=1}^{\mu} w_i \vec{y}_i^{(g+1)} = \frac{\vec{m}^{(g+1)} - \vec{m}^{(g)}}{c_m^{(g)} \sigma^{(g)}} \quad (4.11)$$

Με βάση τα παραπάνω θα δημιουργηθούν δύο μονοπάτια εξέλιξης  $\vec{p}_c$  και  $\vec{p}_\sigma$  για τον υπολογισμό του  $\mathbf{C}$  και  $\sigma$  αντίστοιχα. Προκειμένου να μπορούν τα μονοπάτια να χρησιμοποιηθούν αποτελεσματικά πρέπει να ισχύει  $\vec{p}_c \sim N(\vec{0}, \mathbf{C})$  και  $\vec{p}_\sigma \sim N(\vec{0}, \mathbf{I})$ , αφού με τον όρο  $\sqrt{c(2 - c)}$  έχει εξασφαλιστεί η ισότητα στη διασπορά και όχι στη μέση τιμή. Αν και εξ ορισμού είναι  $\vec{y}_i^{(g+1)} \sim N(\vec{0}, \mathbf{C}^{(g)})$ , το ίδιο δεν ισχύει και για το τελικό βήμα  $\vec{y}_w^{(g+1)}$  λόγω της χρήσης των βαρών  $w_i$ , αφού εφαρμόζοντας την Εξίσωση (3.3) για το γραμμικό συνδυασμό  $\sum_{i=1}^{\mu} w_i \vec{y}_i^{(g+1)}$  προκύπτει τελικά ο συντελεστής διακύμανσης  $w_1^2 + w_2^2 + \dots + w_{\mu-1}^2 + w_\mu^2 = \sum_{i=1}^{\mu} w_i^2$ , άρα είναι:

$$\vec{y}_w^{(g+1)} \sim N(\vec{0}, \sum_{i=1}^{\mu} w_i^2 \mathbf{C}^{(g)}) \sim \sqrt{\sum_{i=1}^{\mu} w_i^2} N(\vec{0}, \mathbf{C}^{(g)}) \quad (4.12)$$

Αυτό μπορεί να διορθωθεί μέσω του συντελεστή  $\mu_{\text{eff}}$  που δίνεται από τη σχέση [5]:

$$\mu_{\text{eff}} = \frac{\sum_{i=1}^{\mu} w_i}{\sum_{i=1}^{\mu} w_i^2} = \frac{1}{\sum_{i=1}^{\mu} w_i^2} \quad (4.13)$$

επειδή πολλαπλασιάζοντας την Εξίσωση (4.12) με  $\sqrt{\mu_{\text{eff}}}$  γίνεται:

$$\sqrt{\mu_{\text{eff}}}\vec{y}_w^{(g+1)} \sim \sqrt{\frac{1}{\sum_{i=1}^{\mu} w_i^2}} \sqrt{\sum_{i=1}^{\mu} w_i^2} N(\vec{0}, \mathbf{C}^{(g)}) \sim N(\vec{0}, \mathbf{C}^{(g)}) \quad (4.14)$$

Επομένως, πρέπει να χρησιμοποιείται παντού το γινόμενο  $\sqrt{\mu_{\text{eff}}}\vec{y}_w$  προκειμένου το μονοπάτι εξέλιξης  $\vec{p}_c$  να έχει την ίδια διακύμανση καθ' όλο το πέρασ της βελτιστοποίησης. Αντικαθιστώντας το θεωρητικό βήμα  $\vec{S}^{(g+1)}$  της Εξίσωσης (4.10) με το επεξεργασμένο, πραγματικό βήμα  $\sqrt{\mu_{\text{eff}}}\vec{y}_w^{(g+1)}$  λαμβάνεται η τελική σχέση:

$$\vec{p}_c^{(g+1)} = (1 - c_c)\vec{p}_c^{(g)} + \sqrt{c_c(2 - c_c)\mu_{\text{eff}}}\vec{y}_w^{(g+1)} \quad (4.15)$$

Απομένει να υπολογιστεί το μονοπάτι εξέλιξης για το μέγεθος του βήματος  $\vec{p}_\sigma$ , για το οποίο το  $\vec{y}_w$  πρέπει να αναχθεί στην κανονική κατανομή  $N(\vec{0}, \mathbf{I})$ . Από την Εξίσωση (3.8) είναι ήδη γνωστό ότι:

$$N(\vec{0}, \mathbf{C}) \sim \mathbf{B}DN(\vec{0}, \mathbf{I}) \rightarrow N(\vec{0}, \mathbf{I}) \sim \mathbf{D}^{-1}\mathbf{B}^{-1}N(\vec{0}, \mathbf{C}) \sim \mathbf{D}^{-1}\mathbf{B}^T N(\vec{0}, \mathbf{C}) \quad (4.16)$$

Στην παραπάνω εξίσωση φαίνεται ότι  $\mathbf{B}^{-1} = \mathbf{B}^T$ . Αυτή είναι μια βασική ιδιότητα των ορθογώνιων μητρώων στα οποία ανήκουν όλα τα μητρώα στροφής, συμπεριλαμβανομένου και του μητρώου  $\mathbf{B}$ . Πρακτικά, η Εξίσωση (4.16) τελεί την αντίστροφη διαδικασία από αυτήν που περιγράφεται στις σελίδες 26 και 27 της εργασίας. Λαμβάνει ένα διάνυσμα που ανήκει στην  $N(\vec{0}, \mathbf{C})$ , του ασκεί μία περιστροφή μέσω του  $\mathbf{B}^T$  ώστε να το επαναφέρει στο κύριο σύστημα συντεταγμένων και τέλος του αφαιρεί τις κλίμακες των τυπικών αποκλίσεων για κάθε μεταβλητή μέσω του  $\mathbf{D}^{-1}$ . Το  $\mathbf{D}$  είναι ένα διαγώνιο μητρώο, οπότε και το  $\mathbf{D}^{-1}$  είναι διαγώνιο και τα στοιχεία του υπολογίζονται πολύ εύκολα με απλή διαίρεση. Παρόλο που το προκύπτον διάνυσμα ανήκει πράγματι στην  $N(\vec{0}, \mathbf{I})$ , αν ο μετασχηματισμός σταματήσει εδώ θα έχουν χαθεί παντελώς οι πληροφορίες που αφορούν την κατεύθυνσή του αρχικού διανύσματος και η συσσώρευση θα αποτύχει πλήρως αφού τα διαδοχικά  $\vec{y}_w$  δεν θα είναι συγκρίσιμα. Επομένως γίνεται ένας επιπλέον πολλαπλασιασμός με το μητρώο στροφής  $\mathbf{B}$  ώστε το τελικό διάνυσμα να έχει την ίδια κατεύθυνση με το  $\vec{y}_w$ , αλλά διαφορετικό μέτρο. Εφαρμόζοντας τα παραπάνω, η Εξίσωση (4.10) διαμορφώνεται για τον υπολογισμό του  $\vec{p}_\sigma$  ως:

$$\vec{p}_\sigma^{(g+1)} = (1 - c_\sigma)\vec{p}_\sigma^{(g)} + \sqrt{c_\sigma(2 - c_\sigma)\mu_{\text{eff}}}\mathbf{B}\mathbf{D}^{-1}\mathbf{B}^T\vec{y}_w^{(g+1)} \quad (4.17)$$

Όσον αφορά την αρχικοποίηση των μονοπατιών εξέλιξης, είναι  $\vec{p}_c^{(0)} = \vec{p}_\sigma^{(0)} = 0$ . Για τις παραμέτρους  $c_c$  και  $c_\sigma$  προτείνεται  $c_c \approx c_\sigma$  και να βρίσκονται αμφότερες εντός του διαστήματος  $[1/n, \sqrt{2/n}]$  όπου  $n$  είναι ο αριθμός των μεταβλητών σχεδιασμού <sup>[13]</sup>. Η βέλτιστη επιλογή εξαρτάται φυσικά από το εκάστοτε πρόβλημα αλλά γενικά συνιστάται η άνω τιμή  $\sqrt{2/n}$  για  $n \leq 5$ , ενώ καθώς το  $n$  αυξάνεται η τιμή αυτή πρέπει να μειώνεται προς το κάτω όριο  $1/n$ .

### 4.3.2 Υπολογισμός του Μητρώου Συνδιακύμανσης

Δεδομένου ότι η γένεση απογόνων γίνεται μέσω μετάλλαξης της μέσης τιμής  $\vec{m}$  ή με άλλα λόγια, πραγματοποιώντας βήματα στην εγγύς περιοχή της όπως φαίνεται στην Εξίσωση (4.9), οι δημιουργοί της μεθόδου CMA επέλεξαν το μητρώο συνδιακύμανσης  $\mathbf{C}$  να κατασκευάζεται με σημείο αναφοράς την πραγματική αυτή μέση τιμή και όχι τη μέση τιμή που προκύπτει από το λαμβανόμενο ανά γενιά δείγμα. Η διαφορά μπορεί να φαίνεται εκ πρώτης όψεως μικρή, αλλά είναι μεγάλης σημασίας διότι επιτρέπει την αξιοποίηση των μονοπατιών εξέλιξης που υπολογίστηκαν προηγουμένως. Έτσι, το  $\mathbf{C}$  εκτιμά τη διακύμανση των υλοποιημένων βημάτων και όχι των τελικών σημείων ή απογόνων που προκύπτουν από αυτά <sup>[5]</sup>. Έχοντας υπόψη τα παραπάνω, μία πρώτη εκτίμηση του νέου μητρώου  $\mathbf{C}^{(g+1)}$  είναι η:

$$\mathbf{C}^{(g+1)} = \sum_{i=1}^{\mu} w_i \vec{y}_i^{(g+1)} \vec{y}_i^{(g+1)T} \quad (4.18)$$

Το παραπάνω μητρώο  $\mathbf{C}^{(g+1)}$  χρησιμοποιεί τα βάρη  $w_i$  για να σταθμίσει τη σημαντικότητα των επιλεγμένων βημάτων και τείνει να αναπαράγει τα επιτυχημένα βήματα από τα οποία προέκυψαν γονείς με μεγαλύτερη συχνότητα.

Ασχέτως άμα γίνεται αναφορά σε βήματα ή σε άτομα, για να υπολογιστεί μια καλή προσέγγιση οποιουδήποτε μητρώου συνδιακύμανσης απαιτείται ένα αρκετά μεγάλο πλήθος παρατηρήσεων ώστε το δείγμα να μπορεί να θεωρηθεί αντιπροσωπευτικό. Σε έναν ΕΑ, αυτό μπορεί να επιτευχθεί με αύξηση των απογόνων  $\lambda$  ανά γενιά ώστε να δοκιμάζονται πολλές υποψήφιες λύσεις και να διαμορφωθεί μια ορθή εικόνα της κατανομής των χαρακτηριστικών εντός του πληθυσμού. Από την Εξίσωση (4.18) μπορεί να φαίνεται ότι απαιτείται αύξηση του αριθμού γονέων  $\mu$ , αλλά αυτό δεν ισχύει διότι οι γονείς προκύπτουν από τους απογόνους. Ένας μεγάλος αριθμός  $\lambda$ , όμως, μπορεί να καθυστερήσει σημαντικά τη σύγκλιση της μεθόδου και πάει ενάντια στην κυριαρχούσα

τάση των ES που χρησιμοποιούν πληθυσμούς μικρού μεγέθους. Η λύση είναι να αξιοποιηθούν για τη δημιουργία του  $\mathbf{C}^{(g+1)}$  πληροφορίες όχι μόνο της γενιάς  $g + 1$ , αλλά και των προηγούμενων γενιών ώστε να υπάρχει ένα συνολικά ικανοποιητικό δείγμα. Αυτό επιτυγχάνεται διατηρώντας ένα γενικό μητρώο συνδιακύμανσης που ανανεώνεται από γενιά σε γενιά ως:

$$\mathbf{C}^{(g+1)} = (1 - c_\mu) \mathbf{C}^{(g)} + c_\mu^{(g)} \sum_{i=1}^{\mu} w_i \vec{y}_i^{(g+1)} \vec{y}_i^{(g+1)T} \quad (4.19)$$

όπου  $c_\mu^{(g)} \in [0,1]$  είναι μια παράμετρος που ελέγχει το ρυθμό μάθησης του μητρώου από τις νέες γενιές σε λογική ολότητα με τους συντελεστές  $c$  των μονοπατιών εξέλιξης. Για  $c_\mu^{(g)} = 1$  οι Εξισώσεις (4.18) και (4.19) προφανώς ταυτίζονται και το μητρώο  $\mathbf{C}^{(g+1)}$  ανανεώνεται πλήρως ανά γενιά. Ο νέος τρόπος υπολογισμού είναι ιδιαίτερα συμβατός με μικρούς πληθυσμούς διότι το μητρώο συνδιακύμανσης θα έχει προσεγγιστεί ικανοποιητικά με λίγες απαιτούμενες γενιές, οπότε και η έρευνα θα ευνοηθεί από τη χρήση του μητρώου πιο γρήγορα και για περισσότερες γενιές προτού τερματιστεί η μέθοδος.

Η ιδέα να χρησιμοποιηθεί το μονοπάτι εξέλιξης  $\vec{p}_c$  για την ανανέωση του  $\mathbf{C}$  προκύπτει από την παρατήρηση ότι η Εξίσωση (4.19), ως έχει, δεν λαμβάνει υπόψη το πρόσημο των βημάτων  $\vec{y}_i$  αφού  $\vec{y}_i \vec{y}_i^T = (-\vec{y}_i)(-\vec{y}_i^T)$  [5]. Αυτό έχει ως αποτέλεσμα να χάνονται πληροφορίες σχετικές με την κατεύθυνση της εξέλιξης, οι οποίες όμως υπάρχουν στο  $\vec{p}_c$  και μπορούν να αξιοποιηθούν μέσω αυτού. Η τελική εξίσωση για τον υπολογισμό του  $\mathbf{C}^{(g+1)}$  διαμορφώνεται ως:

$$\mathbf{C}^{(g+1)} = (1 - c_1 - c_\mu) \mathbf{C}^{(g)} + c_1^{(g)} \vec{p}_c^{(g)} \vec{p}_c^{(g)T} + c_\mu^{(g)} \sum_{i=1}^{\mu} w_i \vec{y}_i^{(g+1)} \vec{y}_i^{(g+1)T} \quad (4.20)$$

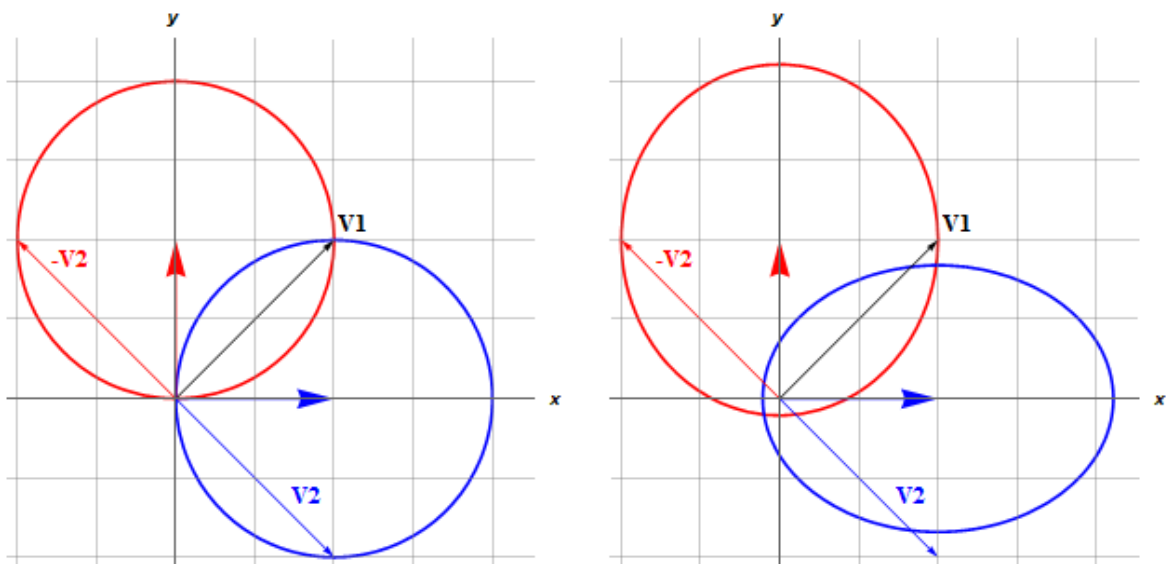
όπου  $c_1^{(g)} \in [0,1]$  είναι μία νέα παράμετρος εκμάθησης που αφορά τη συνεισφορά του μονοπατιού εξέλιξης. Πρέπει να ισχύει  $c_1^{(g)} + c_\mu^{(g)} \leq 1$ , ενώ η βέλτιστη επιλογή τιμής για τα  $c_1^{(g)}$  και  $c_\mu^{(g)}$  εξαρτάται από το πλήθος των μεταβλητών σχεδιασμού  $n$ . Σε παρόμοια λογική με αυτά που εφαρμόστηκαν για την παράμετρο χαλάρωσης της μέσης τιμής  $c_m$  πιο πριν, επιλέγεται  $c_\mu^{(0)} = 1/n$  στην αρχή της βελτιστοποίησης και η τιμή ανανεώνεται ανά γενιά ως:

$$c_\mu^{(g)} = c_\mu^{(0)} \exp \left[ \ln \left( \frac{1}{c_\mu^{(0)}} \right) \frac{\text{evals}^{(g)}}{\text{max\_evals}} \right] \quad (4.21)$$

Στη συνέχεια, ο αλγόριθμος θέτει  $c_1^{(g)} = 0.1 c_\mu^{(g)}$  και υπολογίζει το μητρώο  $\mathbf{C}^{(g+1)}$  από την Εξίσωση (4.20). Η απόφαση που πάρθηκε, δηλαδή ότι η επιρροή του μονοπατιού

εξέλιξης οφείλει να είναι σημαντικά μικρότερη απ' ό τι των βημάτων της γενιάς είναι εμπειρική και θα επεξηγηθεί με ένα παράδειγμα.

Στο Σχήμα 4.2 φαίνονται οι προκύπτουσες ελλείψεις της κανονικής κατανομής  $N(\bar{\mathbf{m}}^{(g+1)}, \mathbf{C}^{(g+1)})$  δύο διαφορετικών ζευγών βημάτων με ίσα βάρη  $w_1 = w_2 = 0.5$ . Το διάνυσμα  $\vec{V}_1$  με μαύρο παραμένει σταθερό και στις δύο περιπτώσεις, ενώ το δεύτερο διάνυσμα  $\vec{V}_2$  αλλάζει πρόσημο ανάλογα με το χρώμα του, δημιουργώντας έτσι δύο διαφορετικές κανονικές κατανομές όταν συνδυάζεται με το  $\vec{V}_1$ . Στο αριστερό σχήμα φαίνεται ότι, λόγω της συμμετρίας των διανυσμάτων περί των αξόνων, οι ισοϋψείς της κανονικής κατανομής έχουν μορφή κύκλων με ίδια ακτίνα ασχέτως άμα χρησιμοποιείται το  $\vec{V}_2$  ή το  $-\vec{V}_2$  ως γονέας. Δεξιά, όμως, όπου λαμβάνεται υπόψη και το μονοπάτι της εξέλιξης, οι κύκλοι έχουν επεκταθεί κατά τη διεύθυνση της μετατόπισης της μέσης τιμής διότι ο υπολογισμός του μονοπατιού βασίζεται σε αυτή. Προκύπτουν έτσι ελλείψεις που ευνοούν την εξερεύνηση του χωρίου προς τη διεύθυνση όπου φαίνεται να ενισχύει σε μεγαλύτερο βαθμό την έρευνα. Πιο αναλυτικά, αφού π.χ. τα διανύσματα  $\vec{V}_1$  και  $\vec{V}_2$  έχουν αντίθετους όρους κατά  $y$  και επιλέχθηκαν ως ισοβαρείς γονείς, μπορεί να γίνει η υπόθεση ότι η έρευνα κατά τον άξονα αυτόν δεν είναι κρίσιμη για τη βελτίωση της δειγματοληψίας.



**Σχήμα 4.2:** Ισοϋψείς καμπύλες κανονικών κατανομών που προκύπτουν προσαρμόζοντας το μητρώο συνδιακύμανσης με  $c_\mu = 1.0$  και  $c_1 = 0.0$  (αριστερά) και με  $c_\mu = c_1 = 0.5$  (δεξιά). Τα διανύσματα με λεπτή κεφαλή αποτελούν τα βήματα που επιλέχθηκαν ως γονείς, ενώ τα διανύσματα με χοντρή κεφαλή δείχνουν τη μεταβολή της μέσης τιμής.

Αντιθέτως, το γεγονός ότι οι δύο γονείς που επιλέχθηκαν έχουν ομόσημες τιμές κατά  $x$  αποτελεί βάσιμη ένδειξη ότι ο εν λόγω άξονας είναι πιο σημαντικός για τη βελτίωση της έρευνας, οπότε και αυτή ενισχύεται κατά  $x$ . Το σκεπτικό μπορεί να είναι απολύτως λογικό, αλλά φαίνεται ότι η νέα κατανομή ως αντίτιμο μειώνει τις πιθανότητες των γονέων να επιλεγούν ξανά εν συγκρίσει με το αριστερό σχήμα. Μάλιστα, για  $c_1 = 1.0$  η έλλειψη χάνει όλο το εύρος της κατά  $y$  και από δισδιάστατη γίνεται μονοδιάστατη όπως φαίνεται στο Σχήμα 3.1. Επομένως, η επιρροή του μονοπατιού εξέλιξης αν και ευεργετική πρέπει να είναι περιορισμένη, αφού εγκυμονεί ο κίνδυνος να ατροφήσουν τα στοιχεία του μητρώου συνδιακύμανσης. Επιπλέον, αναφέρεται ότι παρόλο που το  $\vec{p}_c$  εμπεριέχει πληροφορίες και για το πρόσημο της εξέλιξης (στο παράδειγμα του σχήματος είναι προς τα θετικά και στις δύο περιπτώσεις) και άρα υποδεικνύει κατεύθυνση, οι σχετικές με το πρόσημο πληροφορίες χάνονται αφού στην κύρια διαγώνιο του μητρώου ο πολλαπλασιασμός  $\vec{p}_c \vec{p}_c^T$  τετραγωνίζει τους όρους για να δημιουργηθούν οι τυπικές αποκλίσεις  $\sigma^2$ . Στον παρόντα όμως CMA το πρόσημο του  $\vec{p}_c$  αξιοποιείται από τον νέο τελεστή μετάλλαξης που παρουσιάστηκε στην Ενότητα 4.1 και ωθεί την έρευνα προς τη φορά του όταν ενεργοποιείται, οπότε οι πληροφορίες αυτές δεν χάνονται.

Τέλος, όσον αφορά την αρχικοποίηση του μητρώου συνδιακύμανσης, συμβολισμένη ως  $\mathbf{C}^{(0)}$ , συνήθως τίθεται  $\mathbf{C}^{(0)} = \mathbf{I}$  όπου  $\mathbf{I}$  είναι το μοναδιαίο μητρώο διαστάσεων  $[n \times n]$  και η κλίμακα της κανονικής κατανομής προσαρμόζεται στο πρόβλημα μέσω του μεγέθους του βήματος  $\sigma^{(0)}$ . Θεωρώντας έτσι ότι όλες οι μεταβλητές σχεδιασμού έχουν τα ίδια ή παρόμοια ελάχιστα και μέγιστα όρια, έστω  $x^{\min}$  και  $x^{\max}$  αντίστοιχα, ο Hansen προτείνει την αρχικοποίηση  $\sigma^{(0)} = 0.3(x^{\max} - x^{\min})$  [5]. Αν τα εύρη των ορίων διαφέρουν όμως σημαντικά, όπως π.χ. σε ένα πρόβλημα δύο μεταβλητών σχεδιασμού με  $x_1 \in [0, 1]$  και  $x_2 \in [0, 100]$ , τότε αυτός ο τρόπος δεν είναι ο βέλτιστος. Υπάρχουν ES που για να αποφύγουν αυτό το πρόβλημα χρησιμοποιούν διαφορετικό μέγεθος βήματος για κάθε μεταβλητή που αποθηκεύονται σε ένα διάνυσμα  $\vec{\sigma}$ , αλλά στη CMA-ES είναι πολύ απλούστερο η κλίμακα να προσαρμοστεί αυτόματα μέσω της αρχικοποίησης:

$$\mathbf{C}^{(0)} = \begin{bmatrix} (x_1^{\max} - x_1^{\min})^2 & 0 & \dots & 0 \\ 0 & (x_2^{\max} - x_2^{\min})^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & (x_n^{\max} - x_n^{\min})^2 \end{bmatrix} \quad (4.22)$$

Το πόσο μεγάλο θα είναι το τμήμα του χωρίου σχεδιασμού που καλύπτει η έλλειψη της αρχικής κανονικής κατανομής  $N(\bar{\mathbf{m}}^{(0)}, \sigma^{(0)2} \mathbf{C}^{(0)})$  εξαρτάται προφανώς και από την αρχικοποίηση του  $\sigma^{(0)}$ . Για τυχαία αρχική μέση τιμή  $\bar{\mathbf{m}}^{(0)}$  προτείνεται  $\sigma^{(0)} = 1.0$ , αλλά γενικότερα άμα ισχύει  $\sigma^{(0)} \in [0.3, 1.0]$  η μέθοδος θα μπορέσει να εξερευνήσει ολόκληρο το χωρίο. Άμα γίνει αρχικοποίηση της  $\bar{\mathbf{m}}^{(0)}$  από το χρήστη και είναι γνωστό ότι το ολικό ακρότατο βρίσκεται στην εγγύς περιοχή της, τότε φυσικά μπορεί να δοθεί και πολύ μικρότερη τιμή στο μέγεθος του βήματος, αναλόγως το πόσο κοντά εικάζεται η τοποθεσία του ακρότατου. Αντίθετα, σε δύσκολα προβλήματα βελτιστοποίησης, μπορεί να δοθεί αρχικό βήμα μεγαλύτερο της μονάδας ώστε να αποφευχθεί ο εγκλωβισμός της μεθόδου σε τοπικό ακρότατο.

### 4.3.3 Υπολογισμός του Μεγέθους Βήματος

Έχοντας πλέον ολοκληρώσει τους υπολογισμούς των μεγεθών  $\bar{\mathbf{m}}^{(g+1)}$  και  $\mathbf{C}^{(g+1)}$ , απομένει μονάχα η εύρεση της τιμής του νέου βήματος  $\sigma^{(g+1)}$  ώστε να έχουν τελειώσει όλες οι απαραίτητες εξελικτικές διεργασίες της γενιάς  $g$ . Η προσαρμογή του μεγέθους βήματος είναι ένα πρόβλημα που απασχολεί ανέκαθεν τους χρήστες των ES, διότι επηρεάζει σε τεράστιο βαθμό την απόδοση των αλγορίθμων αυτών. Για να γίνει κατανοητός ο προβληματισμός, ας εξεταστεί η περίπτωση σταθερού βήματος καθ' όλη τη διάρκεια ενός τρεξίματος. Θέτοντας μία σταθερή, μικρή τιμή βήματος η έρευνα θα περιορίζεται σε μια περιοχή κοντά στη μέση τιμή, το οποίο αν και εξασφαλίζει ότι η εξελικτική διαδικασία θα προχωρά με σταθερό ρυθμό, ο ρυθμός αυτός δεν θα είναι καθόλου ικανοποιητικός αν η αρχικοποίηση γίνει πολύ μακριά από τη βέλτιστη τιμή. Επιπρόσθετα, ο αλγόριθμος θα είναι πολύ επιρρεπής σε εγκλωβισμό σε τοπικά ακρότατα διότι δεν υπάρχει η δυνατότητα να υπερπηδήσει ενδεχόμενες κοιλάδες που εμπεριέχουν καλές, αλλά όχι βέλτιστες τιμές σχεδιασμού. Αντιθέτως, μεγάλες τιμές βήματος δεν μπορούν να ανιχνεύσουν την εγγύς περιοχή της μέσης τιμής, οπότε η εξελικτικότητα του αλγορίθμου περιορίζεται δραματικά και κατ' επέκταση η πορεία σύγκλισης θα είναι εξίσου μη ικανοποιητική με πριν <sup>[11]</sup>. Γίνεται επομένως κατανοητό ότι μία ES, προκειμένου να είναι αποδοτική, πρέπει να διαθέτει τη δυνατότητα προσαρμογής του μεγέθους του βήματος στο βέλτιστο δυνατό και ενώ υπάρχουν μεν πολλοί τρόποι για να επιτευχθεί αυτό στη βιβλιογραφία των ES, η μέθοδος που χρησιμοποιείται παραδοσιακά

από τον αλγόριθμο CMA ονομάζεται Cumulative Step-Size Adaptation (CSA) και θεσπίστηκε μάλιστα από τους ίδιους δημιουργούς [28].

Η μέθοδος CSA οφείλει τη μεγάλη της επιτυχία στα αιτιοκρατικά σε μεγάλο βαθμό χαρακτηριστικά της. Οι παλιότερες μέθοδοι για την προσαρμογή του βήματος στηρίζονται σε τοπικές πληροφορίες που προκύπτουν από κάθε γενιά ξεχωριστά, αλλά ο υπολογισμός του μονοπατιού εξέλιξης  $\vec{p}_\sigma$  στον CMA επιτρέπει μια πολύ πιο εξεζητημένη μεθοδολογία που λαμβάνει υπόψη ολόκληρη την πορεία της εξέλιξης. Η θεμελιώδης αρχή πίσω από τη CSA είναι ότι, δεδομένου ότι η δειγματοληψία είναι αμερόληπτη, το  $\vec{p}_\sigma$  πρέπει ανά πάσα στιγμή να έχει το ίδιο μήκος με ένα τυχαία κατανομημένο διάνυσμα  $N(\vec{0}, \mathbf{I})$  ιδίων διαστάσεων  $n$ , δηλαδή να ισχύει [13]:

$$\|\vec{p}_\sigma\| = E\|N(\vec{0}, \mathbf{I})\| \approx \sqrt{n} \left(1 - \frac{1}{4n} + \frac{1}{21n^2}\right) \quad (4.23)$$

Στην περίπτωση που η παραπάνω εξίσωση πράγματι ισχύει, τότε τα διαδοχικά βήματα  $\vec{y}_w$  ανά γενιά είναι μεταξύ τους περίπου κάθετα και ο αλγόριθμος εξερευνεί το χωρίο σχεδιασμού με βέλτιστο τρόπο. Αντιθέτως, εάν  $\|\vec{p}_\sigma\| \neq E\|N(\vec{0}, \mathbf{I})\|$  τότε τα βήματα είναι συσχετιζόμενα, δηλαδή παράλληλα μεταξύ τους, οπότε ο αλγόριθμος πρέπει να επιβάλει με κάποιον τρόπο την επιθυμητή συμπεριφορά που περιγράφηκε προηγουμένως. Σύμφωνα με τον Hansen [29], για  $\|\vec{p}_\sigma\| > E\|N(\vec{0}, \mathbf{I})\|$  τα βήματα έχουν την ίδια κατεύθυνση και αυξάνοντας επομένως το μέγεθός τους, δηλαδή το  $\sigma$ , η ίδια απόσταση μπορεί να καλυφθεί με πολύ μικρότερο αριθμό βημάτων. Στην αντίθετη περίπτωση όπου  $\|\vec{p}_\sigma\| < E\|N(\vec{0}, \mathbf{I})\|$ , τότε τα βήματα έχουν αντίθετη κατεύθυνση με τάση να αλληλοακυρώνονται και για να σταματήσει το εν λόγω φαινόμενο το  $\sigma$  οφείλει να μειωθεί. Ο τελικός τύπος προσαρμογής του μεγέθους του βήματος είναι:

$$\sigma^{(g+1)} = \sigma^{(g)} \exp \left( \frac{c_\sigma}{d_\sigma} \left( \frac{\|\vec{p}_\sigma^{(g+1)}\|}{E\|N(\vec{0}, \mathbf{I})\|} - 1 \right) \right) \quad (4.24)$$

Η μοναδική νέα παράμετρος που εμφανίζεται είναι ο συντελεστής απόσβεσης (damping)  $d_\sigma \approx 1$  που καθορίζει πόσο μεταβάλλεται το  $\sigma^{(g)}$ . Έτσι, τελειώνουν όλες οι απαραίτητες διεργασίες του CMA στην τρέχουσα γενιά  $g$  και ο αλγόριθμος μπορεί πλέον να προχωρήσει στη δημιουργία της επόμενης με δείκτη  $g + 1$ .

Κλείνοντας, αναφέρεται ότι η μέθοδος CSA είναι αυτή που προσδίδει στον CMA τα χαρακτηριστικά αποτυχαιοποίησης που προαναφέρθηκαν στην αρχή του κεφαλαίου.



Παρατηρώντας όλες τις διεργασίες που παρουσιάστηκαν φαίνεται ότι η στοχαστικότητα του αλγορίθμου προέρχεται εξ ολοκλήρου από τα διανύσματα μετάλλαξης  $\bar{z}_i$ . Αυτά είναι που καθορίζουν εν τέλει την προσαρμογή τόσο του μητρώου  $\mathbf{C}$  στην Εξίσωση (4.20) όσο και του μέτρου του βήματος  $\sigma$  στην Εξίσωση (4.24). Από τη στιγμή, όμως, που και οι δύο εξισώσεις χρησιμοποιούν άμεσα ή έμμεσα (μέσω των μονοπατιών εξέλιξης) τα  $\bar{z}_i$  για την κατάλληλη προσαρμογή των παραμέτρων του αλγορίθμου, οι στοχαστικές διακυμάνσεις περιορίζονται σημαντικά και έτσι ο CMA γίνεται πολύ πιο στιβαρός <sup>[13]</sup>. Σε αυτό το σημείο της εργασίας, το θεωρητικό τμήμα έχει πλέον καλυφθεί, ο τρόπος λειτουργίας του τροποποιημένου αλγορίθμου CMA έχει περιγραφεί πλήρως και απομένει μόνο να δοκιμαστεί η απόδοσή του σε προβλήματα βελτιστοποίησης.

## ΚΕΦΑΛΑΙΟ 5

### ΣΧΕΔΙΑΣΤΙΚΑ ΠΡΟΒΛΗΜΑΤΑ ΨΕΥΔΟΜΗΧΑΝΙΚΗΣ

Οι περισσότερες μέθοδοι βελτιστοποίησης δοκιμάζονται σε πρώτο στάδιο είτε με μαθηματικές συναρτήσεις σχεδιασμού (test functions) όπως π.χ. η συνάρτηση σφαίρας, είτε σε απλουστευμένα προβλήματα σχεδιασμού με εφαρμογή στη μηχανική (engineering design optimization problems). Οι κύριες διαφορές μεταξύ των δύο τρόπων δοκιμής είναι ότι οι μαθηματικές συναρτήσεις συνήθως δεν έχουν περιορισμούς και η τοποθεσία του ολικού ακρότατου είναι εξ αρχής γνωστή, ενώ αντιθέτως στα προβλήματα ψευδομηχανικής υπάρχουν πάντα περιορισμοί λόγω της φύσης τους και η βέλτιστη τιμή σχεδιασμού δεν είναι μαθηματικά προσβάσιμη a priori. Επιπρόσθετα, οι μαθηματικές συναρτήσεις έχουν το θετικό ότι ο χρήστης μπορεί να καθορίσει ο ίδιος τον αριθμό των μεταβλητών σχεδιασμού  $n$ , άρα και τη δυσκολία του προβλήματος, αλλά τα προβλήματα ψευδομηχανικής βασίζονται σε πραγματικές εφαρμογές μηχανικής, οπότε και οι συναρτήσεις τους είναι πιο αντιπροσωπευτικές προβλημάτων με πρακτική εφαρμογή. Στη βιβλιογραφία υπάρχει πληθώρα κειμένων που πραγματοποιούν διερεύνηση για τον κλασικό CMA ή διάφορες παραλλαγές του σε test functions <sup>[7], [13], [27], [29]</sup>, συνηθέστερα χωρίς την προσθήκη περιορισμών, οπότε στο πλαίσιο της παρούσας εργασίας επιλέγεται να γίνουν εφαρμογές ψευδομηχανικής. Επιλέχθηκαν τρία από τα κλασικότερα προβλήματα, τα οποία θα αναλυθούν από φυσικής άποψης και, στη συνέχεια, θα επιλυθούν από τον τροποποιημένο αλγόριθμο CMA χωρίς και με μεταμοντέλα για να διαπιστωθεί η αποδοτικότητά τους, καθώς και από έναν DMAEA (Distributed Metamodel-Assisted Evolutionary Algorithm) <sup>[17], [18]</sup> του λογισμικού EASY με δυαδική κωδικοποίηση.

Για λόγους πληρότητας αναφέρεται ότι ένας DMAEA πρόκειται ουσιαστικά για πολλαπλούς MAEA όπως περιγράφηκαν στην Υποενότητα 2.3.3, οι οποίοι ναί μεν χειρίζονται διαφορετικούς πληθυσμούς που εξελίσσονται ανεξάρτητα μεταξύ τους αλλά με τη δυνατότητα να ανταλλάσσουν άτομα ανά προκαθορισμένο αριθμό γενιών. Αυτό έχει ως αποτέλεσμα την περαιτέρω βελτίωση της σύγκλισης, αφού π.χ. ένας (30,60)-MAEA θα δώσει καλύτερα αποτελέσματα άμα εκτελεστεί ως τρεις ξεχωριστούς (10,20)-MAEA που συνεργάζονται και ανταλλάσσουν πληροφορίες μεταξύ τους. Ο εν λόγω αλγόριθμος θα συμβολιζόταν ως 3x(10,20)-MAEA και αποτελεί, σύμφωνα με την ορολογία, έναν DMAEA με τρεις δήμους (demes).

### 5.1 ΣΧΕΔΙΑΣΜΟΣ ΔΙΚΤΥΩΜΑΤΟΣ ΤΡΙΩΝ ΜΕΛΩΝ

Το πρώτο πρόβλημα που θα επιλυθεί είναι ίσως και το ευκολότερο που συναντάται στη βιβλιογραφία, όπου και αναφέρεται ως Three Bar Truss Design <sup>[30]</sup>. Αφορά το σχεδιασμό ενός δικτύωματος τριών μελών υπό δύο πιθανές φορτίσεις  $P$  όπως φαίνεται στο Σχήμα 5.1. Η φόρτιση θα έχει ως αποτέλεσμα να αναπτυχθούν ορθές τάσεις κατά μήκος των μελών. Λόγω συμμετρίας και ανάλογα την κατεύθυνση της φόρτισης, ένα από τα ακριανά μέλη βρίσκεται πάντα υπό θλίψη, ενώ το άλλο ακριανό και το μεσαίο μέλος εφελκύνονται. Θεωρώντας ότι οι διατομές των ακριανών μελών έχουν εμβαδό  $x_1$  και της μεσαίας  $x_2$ , ζητείται να ελαχιστοποιηθεί το βάρος της κατασκευής. Το βάρος ανάγεται στον όγκο και τελικά η συνάρτηση προς ελαχιστοποίηση είναι η:

$$f(\vec{X}) = (2\sqrt{2}x_1 + x_2)L \quad (5.1)$$

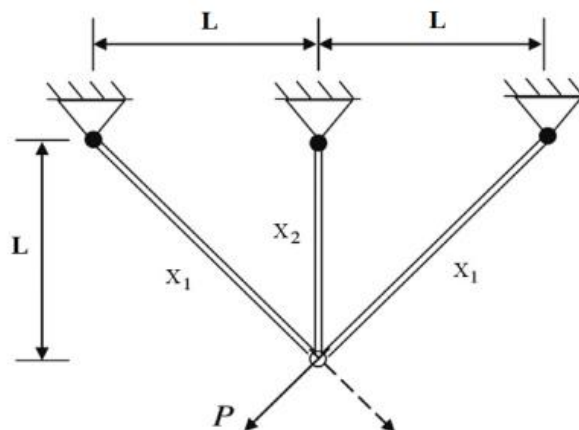
όπου  $x_1 \in [0.0, 1.0]$ ,  $x_2 \in [0.0, 1.0]$  και  $L = 100$  cm. Στο πρόβλημα υπεισέρχονται τρεις επιπλέον περιορισμοί που αφορούν τις μέγιστες επιτρεπτές τάσεις που θα αναπτυχθούν σε κάθε μέλος:

$$g_1(\vec{X}) = \frac{\sqrt{2}x_1 + x_2}{\sqrt{2}x_1^2 + 2x_1x_2} P - \sigma \leq 0.0 \quad (5.2)$$

$$g_2(\vec{X}) = \frac{x_2}{\sqrt{2}x_1^2 + 2x_1x_2} P - \sigma \leq 0.0 \quad (5.3)$$

$$g_3(\vec{X}) = \frac{1}{x_1 + \sqrt{2}x_2} P - \sigma \leq 0.0 \quad (5.4)$$

όπου  $P = 2$  kN/cm<sup>2</sup> και  $\sigma = 2$  kN/cm<sup>2</sup> είναι η μέγιστη επιτρεπτή τάση. Ο περιορισμός  $g_1$  αναφέρεται στον εφελκυσμό της ράβδου που βρίσκεται στην ίδια ευθεία με την  $P$ , ο  $g_2$  στον εφελκυσμό της μεσαίας ράβδου και ο  $g_3$  στη θλίψη της απομένουσας.



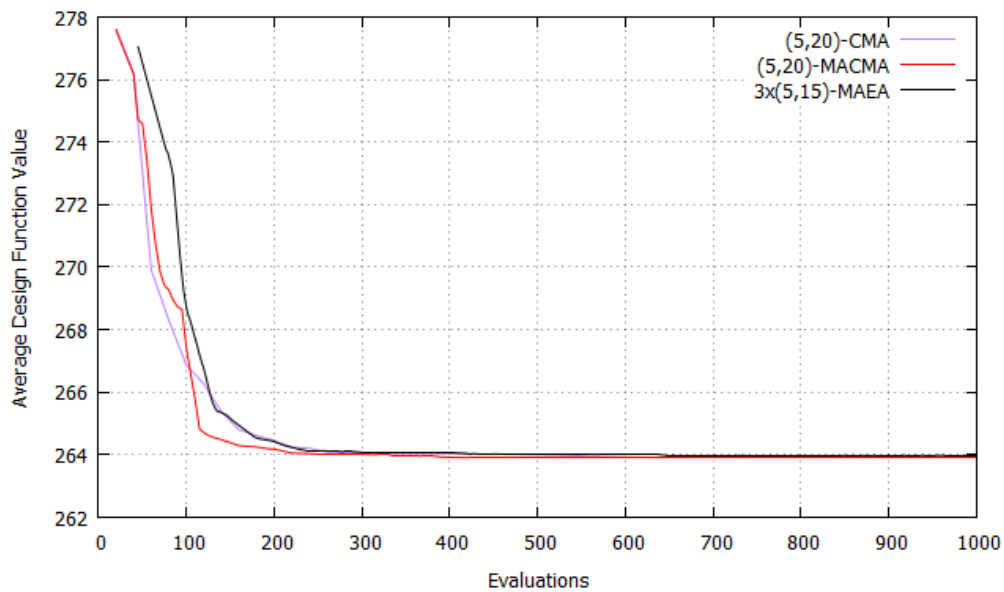
Σχήμα 5.1: Συμμετρικό δικτύωμα τριών μελών υπό φόρτιση  $P$  <sup>[30]</sup>

Πραγματοποιήθηκαν 10 τρεξίματα με κάθε αλγόριθμο και υπολογίστηκε το μέσο διάγραμμα σύγκλισης του καθένα. Οι βασικότερες ρυθμίσεις των CMA και του DMAEA δίνονται στον Πίνακα 5.1 και ο μέγιστος αριθμός αξιολογήσεων επιλέγεται ώστε ο αλγόριθμος CMA χωρίς μεταμοντέλα να έχει φτάσει σε μια ικανοποιητική τιμή της συνάρτησης σχεδιασμού. Σύμφωνα με την παραπάνω μεθοδολογία, για το πρόβλημα του δικτύματος τριών μελών τίθενται ως μέγιστο οι 1500 αξιολογήσεις. Τα διαγράμματα σύγκλισης σχεδιάστηκαν στο Σχήμα 5.2 έως τις 1000 αξιολογήσεις διότι στη συνέχεια φαίνεται να σταθεροποιούνται. Παρακάτω ακολουθούν πίνακες με τα στοιχεία κάθε επίλυσης που έγινε, όπου και συγκρίνονται τα στατιστικά και οι τελικές τιμές της συνάρτησης-στόχου που προέκυψαν (Πίνακας 5.2). Η σύγκριση επεκτείνεται και στα τελικά αποτελέσματα άλλων αλγορίθμων της βιβλιογραφίας (Πίνακας 5.3), ενώ τέλος παρουσιάζονται και οι ακριβείς τιμές των μεταβλητών σχεδιασμού για να φανεί άμα υπάρχει ουσιαστική διαφορά στην τελική λύση (Πίνακας 5.4).

**Πίνακας 5.1:** Ρυθμίσεις των αλγορίθμων CMA και DMAEA για το πρόβλημα σχεδιασμού δικτύματος τριών μελών.

CMA & MACMA			
Γονείς $\mu$	5	Χαλάρωση $c_m^{(0)}$	1.0
Απόγονοι $\lambda$	20	Αρχικό βήμα $\sigma^{(0)}$	1.0
Τελεστής επιλογής γονέων		Truncation selection	
Τελεστής δημιουργίας απογόνων		Weighted multi-recombination	
Πιθανότητα μετάλλαξης		5%	

DMAEA			
Γονείς $\mu$	5	Δήμοι	3
Απόγονοι $\lambda$	15	Ρυθμός ανταλλαγής	4 γενιές
Τελεστής επιλογής γονέων		Tournament selection	
Μέγεθος διαγωνισμού $t$		2	
Πιθανότητα επιλογής $p$		80%	
Τελεστής δημιουργίας απογόνων		2-point per variable crossover	
Πιθανότητα διασταύρωσης		80%	
Πιθανότητα μετάλλαξης		5%	



**Σχήμα 5.2:** Μέσο διάγραμμα σύγκλισης 10 τρεξιμάτων για το πρόβλημα σχεδιασμού του δικτύματος τριών μελών.

**Πίνακας 5.2:** Στατιστική ανάλυση των αποτελεσμάτων για το πρόβλημα σχεδιασμού του δικτύματος τριών μελών από την εκτέλεση 10 τρεξιμάτων με 1500 επιτρεπόμενες αξιολογήσεις.

Αλγόριθμος	CMA	MACMA	DMAEA
Βέλτιστη Τιμή	263.895843376	263.895843376	263.90149
Χείριστη Τιμή	263.895843387	263.895878678	264.18611
Μέση Τιμή	263.895843378	263.895848301	263.98060
Διάμεσος	263.895843376	263.895843391	263.96600
Τυπική Απόκλιση	4.00E-09	1.09E-05	8.50E-02

**Πίνακας 5.3:** Στατιστικά αποτελεσμάτων άλλων αλγορίθμων για το πρόβλημα σχεδιασμού του δικτύματος τριών μελών.

Αλγόριθμος	SC <sup>[31]</sup>	MBA <sup>[32]</sup>	ITLBO <sup>[33]</sup>	HBBO <sup>[34]</sup>	DAPSO <sup>[35]</sup>
Αξιολογήσεις	17610	13280	5000	7550	7131
Βέλτιστη Τιμή	263.8958	263.895852	263.8958434	263.8958433	263.8958436
Χείριστη Τιμή	263.9697	263.915983	263.8958434	263.8958433	263.9476331
Μέση Τιμή	263.9033	263.897996	263.8958434	263.8958433	263.9029260
Διάμεσος	263.8989	-	-	263.8958433	-
Τυπ. Απόκλιση	1.26E-02	3.93E-03	5.78E-14	7.08E-14	8.30E-03

**Πίνακας 5.4:** Βέλτιστες τιμές της συνάρτησης-στόχου και αντίστοιχες τιμές των μεταβλητών σχεδιασμού για το πρόβλημα σχεδιασμού του δικτύωματος τριών μελών με 1500 επιτρεπόμενες αξιολογήσεις.

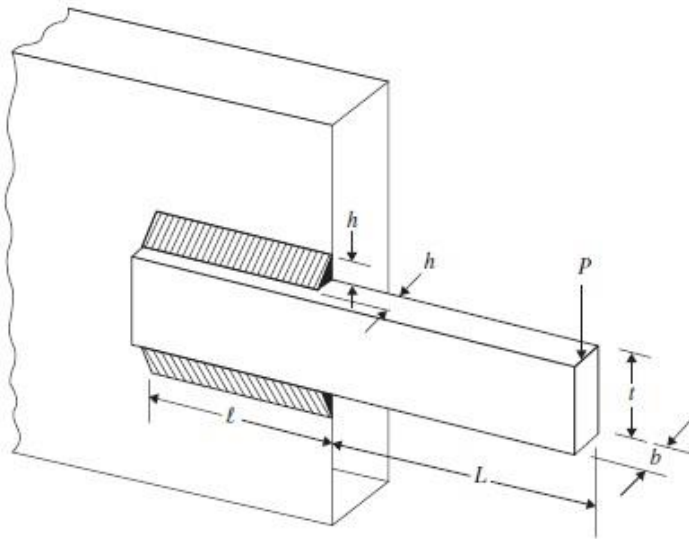
Αλγόριθμος	CMA	MACMA	DMAEA
$f(\vec{X})$	263.895843376	263.895843376	263.90149
$x_1$	0.78867513662	0.78867575375	0.78787878787
$x_2$	0.40824828473	0.40824653922	0.41055718475

Συγκρίνοντας τους Πίνακες 5.2 και 5.3, φαίνεται ότι ο τροποποιημένος αλγόριθμος CMA δίνει ίδια ή καλύτερη ποιότητα λύσεων σε σύγκριση με τους υπόλοιπους στοχαστικούς αλγορίθμους, οι οποίοι μάλιστα τερματίζονται σε τουλάχιστον τριπλάσιο αριθμό αξιολογήσεων. Επιπλέον, ο CMA έχει πολύ μικρή τυπική απόκλιση οπότε τα αποτελέσματά του είναι συνεπέστατα. Η χρήση μεταμοντέλων φαίνεται να αυξάνει αισθητά την τυπική απόκλιση της τελικής τιμής, αλλά εξακολουθεί να διατηρείται σε ικανοποιητικά επίπεδα. Στο Σχήμα 5.2 φαίνεται ότι τα μεταμοντέλα βοηθούν τον CMA στο να ανιχνεύσει καλύτερη λύση μετά από 100 αξιολογήσεις, αλλά εν τέλει η βέλτιστη λύση των 10 τρεξιμάτων δεν επηρεάζεται με ακρίβεια 9 δεκαδικών ψηφίων. Η βέλτιστη λύση του DMAEA είναι ελαφρώς χειρότερη συγκριτικά, με την τυπική απόκλιση να είναι σαφώς μεγαλύτερη. Είναι προφανές όμως ότι οι 1500 αξιολογήσεις αρκούν για να βρεθεί ίδια ποιότητα λύσης με τον CMA, αφού η διαφορά στο 2<sup>ο</sup> δεκαδικό ψηφίο είναι πρακτικά αμελητέα. Στον Πίνακα 5.4 φαίνεται άλλωστε και η μικρή διαφορά στις τιμές των μεταβλητών σχεδιασμού περίπου στο 3<sup>ο</sup> δεκαδικό. Γενικά, στο Σχήμα 5.2 δεν διακρίνεται σημαντική διαφοροποίηση των τριών αλγορίθμων στην περιοχή μετά τις 250 επαναλήψεις, οπότε θα πρέπει να εξεταστούν πιο δύσκολα προβλήματα για να ληφθούν σαφή συμπεράσματα σχετικά με το νέο αλγόριθμο.

## 5.2 ΣΧΕΔΙΑΣΜΟΣ ΣΥΓΚΟΛΛΗΤΗΣ ΔΟΚΟΥ

Το επόμενο πρόβλημα είναι γνωστό ως Welded Beam Design και αφορά το σχεδιασμό συγκολλητής δοκού με στόχο την ελαχιστοποίηση του κόστους κατασκευής της σε δολάρια <sup>[36]</sup>. Η δοκός βρίσκεται υπό φόρτιση με μια σημειακή δύναμη  $P$  [lb] στο ελεύθερο άκρο της όπως φαίνεται στο Σχήμα 5.3. Η φόρτιση αναμένεται να προκαλέσει κατακόρυφη μετατόπιση του ελεύθερου άκρου κατά  $\delta$  [in] με αποτέλεσμα να

παρουσιάζονται ορθές τάσεις λόγω κάμψης  $\sigma$  [psi] κατά μήκος της δοκού και διατμητικές τάσεις  $\tau$  [psi] στη ραφή.



**Σχήμα 5.3:** Συγκολλητή δοκός υπό φόρτιση P <sup>[1]</sup>

Όλες οι μεταβλητές σχεδιασμού διακρίνονται στο παραπάνω σχήμα και είναι με αύξουσα αρίθμηση: το πάχος της ραφής  $h = x_1$ , το μήκος της ραφής  $l = x_2$ , το ύψος της δοκού  $t = x_3$  και το πάχος της δοκού  $b = x_4$ , όλα σε [in]. Έτσι το διάνυσμα σχεδιασμού είναι το:

$$\vec{X} = \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{Bmatrix} = \begin{Bmatrix} h \\ l \\ t \\ b \end{Bmatrix} \quad (5.5)$$

με  $x_1 \in [0.1, 2.0]$ ,  $x_2 \in [0.1, 10.0]$ ,  $x_3 \in [0.1, 10.0]$  και  $x_4 \in [0.1, 2.0]$ .

Προτού γίνει η διατύπωση της συνάρτησης-στόχου και των περιορισμών πρέπει να αναφερθεί κάτι σημαντικό. Το συγκεκριμένο πρόβλημα έχει πολλές διαφορετικές εκδοχές και δεν είναι σπάνιο ερευνητές να συγκρίνουν τα αποτελέσματα της μεθόδου τους με αντίστοιχα αποτελέσματα άλλων εκδοχών, καταλήγοντας έτσι συχνά σε λανθασμένα συμπεράσματα για την υπεροχή του αλγορίθμου τους <sup>[37]</sup>. Επισημαίνεται ότι η εκδοχή που χρησιμοποιήθηκε για τη δοκιμή του CMA μπορεί να βρεθεί στο [36] ως Version II και δόθηκε ιδιαίτερη προσοχή ώστε να συγκριθούν αποτελέσματα αλγορίθμων που έχουν λύσει με σιγουριά τη συγκεκριμένη εκδοχή, προκειμένου να αποφευχθεί η σύγχυση που παρατηρείται πολλές φορές στη βιβλιογραφία και τα τελικά συμπεράσματα να είναι δίκαια και αντιπροσωπευτικά.

Σύμφωνα λοιπόν με τη δεύτερη εκδοχή του προβλήματος συγκολλητής δοκού, η συνάρτηση-στόχος είναι το κόστος σε δολάρια:

$$f(\vec{X}) = 1.10471x_1^2x_2 + 0.04811x_3x_4(14.0 + x_2) \quad (5.6)$$

και συμπεριλαμβάνει τα κόστη τοποθέτησης, συγκόλλησης και υλικών κατασκευής [38]. Επιπρόσθετα, στο πρόβλημα επιβάλλονται οι επτά εξής περιορισμοί:

$$g_1(\vec{X}) = \tau(\vec{X}) - \tau_{\max} \leq 0.0 \quad (5.7)$$

$$g_2(\vec{X}) = \sigma(\vec{X}) - \sigma_{\max} \leq 0.0 \quad (5.8)$$

$$g_3(\vec{X}) = x_1 - x_4 \leq 0.0 \quad (5.9)$$

$$g_4(\vec{X}) = 0.125 - x_1 \leq 0.0 \quad (5.10)$$

$$g_5(\vec{X}) = \delta(\vec{X}) - \delta_{\max} \leq 0.0 \quad (5.11)$$

$$g_6(\vec{X}) = P - P_c(\vec{X}) \leq 0.0 \quad (5.12)$$

$$g_7(\vec{X}) = 0.10471x_1^2 + 0.04811x_3x_4(14.0 + x_2) - 5.0 \leq 0.0 \quad (5.13)$$

όπου  $\tau_{\max} = 13600$  psi είναι η μέγιστη επιτρεπτή διατμητική τάση στη ραφή,  $\sigma_{\max} = 30000$  psi είναι η μέγιστη επιτρεπτή ορθή τάση στη δοκό και  $\delta_{\max} = 0.25$  in είναι η μέγιστη επιτρεπτή μετατόπιση του ελεύθερου άκρου. Αυτά τα μεγέθη οριοθετούνται πάντα από κανονισμούς και ελέγχονται από τους περιορισμούς  $g_1$ ,  $g_2$  και  $g_5$  αντίστοιχα. Οι περιορισμοί  $g_3$  και  $g_4$  εξασφαλίζουν ότι  $0.125 \leq x_1 \leq x_4$ , δηλαδή ότι το πάχος της ραφής είναι μικρότερο από το πάχος της δοκού, αλλά ταυτόχρονα μεγαλύτερο από ένα ελάχιστο επιτρεπτό όριο. Ο περιορισμός  $g_6$  τελεί τον έλεγχο αντοχής της κατασκευής έναντι λυγισμού, ενώ τέλος ο  $g_7$  δεν επιτρέπει στο συνολικό κόστος των υλικών να ξεπεράσει τα 5 δολάρια. Η ασκούμενη δύναμη στο άκρο τίθεται ίση με  $P = 6000$  lb. Τα ακριβή μεγέθη των ορθών και διατμητικών τάσεων  $\tau(\vec{X})$  και  $\sigma(\vec{X})$  αντίστοιχα, της μετατόπισης  $\delta(\vec{X})$  στο ελεύθερο άκρο και του φορτίου λυγισμού  $P_c(\vec{X})$  της δοκού εξαρτώνται από τις εκάστοτε τιμές σχεδιασμού και υπολογίζονται από τους ακόλουθους τύπους:

$$\tau(\vec{X}) = \sqrt{(\tau')^2 + 2\tau'\tau'' \frac{x_2}{2R} + (\tau'')^2} \quad (5.14)$$

$$\tau' = \frac{P}{\sqrt{2}x_1x_2} \quad (5.15)$$



$$\tau'' = \frac{MR}{J} \quad (5.16)$$

$$M = P \left( L + \frac{x_2}{2} \right) \quad (5.17)$$

$$R = \sqrt{\frac{x_2^2}{4} + \left( \frac{x_1 + x_3}{2} \right)^2} \quad (5.18)$$

$$J = 2 \left\{ \sqrt{2} x_1 x_2 \left[ \frac{x_2^2}{12} + \left( \frac{x_1 + x_3}{2} \right)^2 \right] \right\} \quad (5.19)$$

$$\sigma(\vec{X}) = \frac{6PL}{x_4 x_3^2} \quad (5.20)$$

$$\delta(\vec{X}) = \frac{4PL^3}{E x_3^3 x_4} \quad (5.21)$$

$$P_c(\vec{X}) = \frac{4.013E\sqrt{(x_3^2 x_4^6 / 36)}}{L^2} \left( 1 - \frac{x_3}{2L} \sqrt{\frac{E}{4G}} \right) \quad (5.22)$$

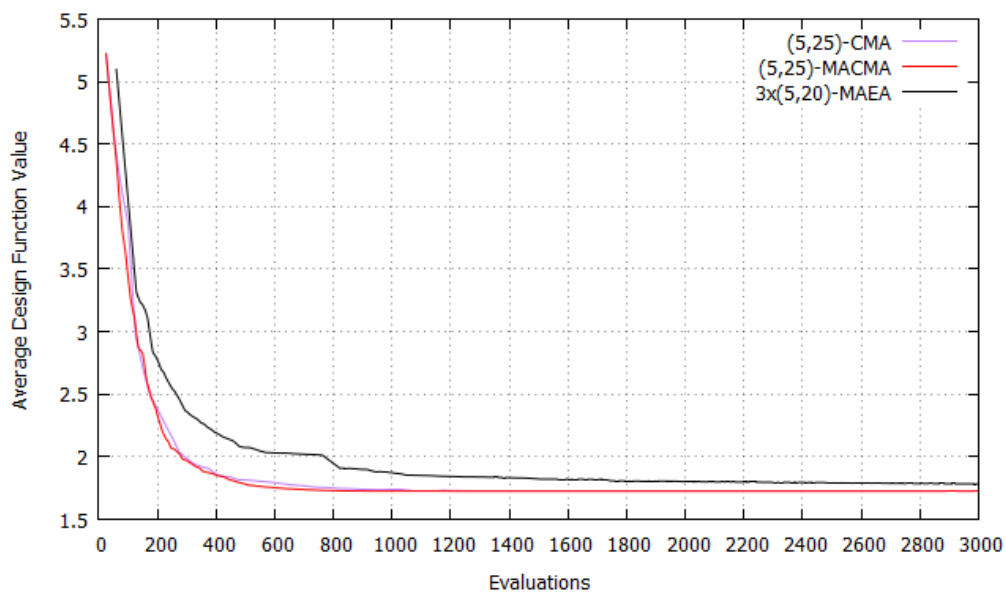
όπου  $L = 14$  in είναι το μήκος της δοκού άνευ συγκόλλησης,  $G = 12 \times 10^6$  psi είναι το μέτρο διάτμησης του υλικού κατασκευής της δοκού και  $E = 30 \times 10^6$  psi είναι το μέτρο ελαστικότητας του ίδιου υλικού. Το λογισμικό αξιολόγησης οφείλει να υπολογίζει όλες τις τιμές των Εξισώσεων (5.14) έως (5.22) για κάθε απόγονο, ώστε κατά τη βελτιστοποίηση να μπορεί να διαπιστωθεί εάν τηρούνται ή όχι οι απαιτούμενοι περιορισμοί.

Η αύξηση των μεταβλητών σχεδιασμού και των περιορισμών του προβλήματος έχει ως λογικό επακόλουθο την ανάγκη για αύξηση του μέγιστου αριθμού αξιολογήσεων σε 3000 προκειμένου να εντοπιστεί λύση ικανοποιητικής ποιότητας. Τα νέα διαγράμματα σύγκλισης δίνονται στο Σχήμα 5.4 και ακολουθούν οι αντίστοιχοι με πριν πίνακες με τα στατιστικά στοιχεία του τροποποιημένου CMA, του MACMA, του DMAEA και των λοιπών αλγορίθμων βελτιστοποίησης που έχουν επιλύσει το πρόβλημα Welded Beam II. Σχετικά με κάποιους αλγορίθμους της βιβλιογραφίας που έχουν ήδη αναφερθεί, ο SC <sup>[31]</sup> φαίνεται να επιλύει το Welded Beam I κατά το [36], ενώ οι βέλτιστες λύσεις των DAPSO <sup>[35]</sup>, του ABC2 <sup>[36]</sup> και του NM2 <sup>[37]</sup> είναι μη αποδεκτές σύμφωνα με το λογισμικό αξιολόγησης που χρησιμοποιήθηκε στην εργασία, οπότε και παραλείπονται.

**Πίνακας 5.5:** Ρυθμίσεις των αλγορίθμων CMA και DMAEA για το πρόβλημα σχεδιασμού της συγκολλητής δοκού.

CMA & MACMA			
Γονείς $\mu$	5	Χαλάρωση $c_m^{(0)}$	1.0
Απόγονοι $\lambda$	25	Αρχικό βήμα $\sigma^{(0)}$	1.0
Τελεστής επιλογής γονέων		Truncation selection	
Τελεστής δημιουργίας απογόνων		Weighted multi-recombination	
Πιθανότητα μετάλλαξης		5%	

DMAEA			
Γονείς $\mu$	5	Δήμοι	3
Απόγονοι $\lambda$	20	Ρυθμός ανταλλαγής	3 γενιές
Τελεστής επιλογής γονέων		Tournament selection	
Μέγεθος διαγωνισμού $t$		2	
Πιθανότητα επιλογής $p$		80%	
Τελεστής δημιουργίας απογόνων		2-point per variable crossover	
Πιθανότητα διασταύρωσης		85%	
Πιθανότητα μετάλλαξης		5%	



**Σχήμα 5.4:** Μέσο διάγραμμα σύγκλισης 10 τρεξιμάτων για το πρόβλημα σχεδιασμού της συγκολλητής δοκού.

**Πίνακας 5.6:** Στατιστική ανάλυση των αποτελεσμάτων για το πρόβλημα σχεδιασμού της συγκολλητής δοκού από την εκτέλεση 10 τρεξιμάτων με 3000 επιτρεπόμενες αξιολογήσεις.

Αλγόριθμος	CMA	MACMA	DMAEA
Βέλτιστη Τιμή	1.724852338	1.724852417	1.746258
Χείριστη Τιμή	1.724867809	1.724927054	1.831478
Μέση Τιμή	1.724853901	1.724861230	1.783269
Διάμεσος	1.724852343	1.724853760	1.784855
Τυπική Απόκλιση	4.63E-06	2.20E-05	2.23E-02

**Πίνακας 5.7:** Στατιστικά αποτελεσμάτων άλλων αλγορίθμων για το πρόβλημα σχεδιασμού της συγκολλητής δοκού.

Αλγόριθμος	MBA <sup>[32]</sup>	ITLBO <sup>[33]</sup>	HBBO <sup>[34]</sup>
Αξιολογήσεις	47340	10000	25050
Βέλτιστη Τιμή	1.724853	1.7248523	1.724852309
Χείριστη Τιμή	1.724853	1.7248523	1.724852309
Μέση Τιμή	1.724853	1.7248523	1.724852309
Διάμεσος	1.724853	1.7248523	1.724852309
Τυπική Απόκλιση	6.94E-19	6.77E-16	1.14E-15

**Πίνακας 5.8:** Βέλτιστες τιμές της συνάρτησης-στόχου και αντίστοιχες τιμές των μεταβλητών σχεδιασμού για το πρόβλημα σχεδιασμού της συγκολλητής δοκού με 3000 επιτρεπόμενες αξιολογήσεις.

Αλγόριθμος	CMA	MACMA	DMAEA
$f(\vec{X})$	1.724852338	1.724852417	1.746258844
$x_1$	0.205729639	0.205729642	0.201172447
$x_2$	3.470488716	3.470488814	3.600659340
$x_3$	9.036623923	9.036623647	8.974945054
$x_4$	0.205729639	0.205729653	0.208597948

Στο Σχήμα 5.4 διακρίνεται ότι ο τροποποιημένος CMA παρουσιάζει ελαφρώς καλύτερη πορεία σύγκλισης με τη χρήση μεταμοντέλων. Επιπρόσθετα, στον Πίνακα 5.6 φαίνεται ότι ο απλός CMA έχει ελαφρώς καλύτερα στατιστικά, αν και υπάρχει πλήρης ταύτιση με τον MACMA μέχρι το 3<sup>ο</sup> δεκαδικό ψηφίο. Η καλύτερη λύση που βρέθηκε από τον αλγόριθμο HBBO στον Πίνακα 5.7, δεδομένης και της ακρίβειάς της, έχει διαφορά μόλις 2.90E-08 συγκριτικά με του CMA, ενώ χρειάστηκε περίπου οκταπλάσιος αριθμός αξιολογήσεων για να βρεθεί. Στις 3000 αξιολογήσεις ο DMAEA έχει προσεγγίσει τη βέλτιστη λύση με ακρίβεια ενός δεκαδικού ψηφίου, αλλά είναι εμφανές από το Σχήμα 5.4 ότι δεν έχει προλάβει να συγκλίνει πλήρως αφού η πτωτική πορεία της μέσης συνάρτησης στόχου συνεχίζεται έως το τέλος του διαγράμματος σύγκλισης. Επίσης παρατηρείται ότι σε κανένα σημείο της βελτιστοποίησης το διάγραμμα του DMAEA δεν «ακούμπησε» τα αντίστοιχα των CMA, δηλαδή η σύγκλιση του ήταν αυστηρά χειρότερη. Συμπεραίνεται έτσι ότι ο DMAEA χρειάζεται έναν ελαφρώς μεγαλύτερο αριθμό αξιολογήσεων συγκριτικά με τον CMA προκειμένου να βρει παρόμοια λύση. Τέλος, τα βέλτιστα διανύσματα σχεδιασμού του Πίνακα 5.8 ταυτίζονται στο 6<sup>ο</sup> δεκαδικό για τους CMA, ενώ το αντίστοιχο διάνυσμα του DMAEA εν συγκρίσει παρουσιάζει σημαντική διαφορά κυρίως στις τιμές των μεταβλητών σχεδιασμού  $x_2$  και  $x_3$ , ενδεχομένως λόγω του μεγαλύτερου εύρους τιμών που μπορούν να λάβουν συγκριτικά με τις μεταβλητές  $x_1$  και  $x_4$ . Λαμβάνοντας υπόψη όλα τα παραπάνω, ο τροποποιημένος αλγόριθμος CMA έχει αισθητά καλύτερη απόδοση στο πρόβλημα της συγκολλητής δοκού απ' ότι ο DMAEA και γενικότερα φαίνεται να απαιτεί σημαντικά μικρότερο αριθμό απαιτούμενων αξιολογήσεων για να βρει λύση άριστης ποιότητας, συγκρίσιμη με των υπολοίπων αλγορίθμων της βιβλιογραφίας.

### 5.3 ΣΧΕΔΙΑΣΜΟΣ ΜΕΙΩΤΗΡΑ ΣΤΡΟΦΩΝ

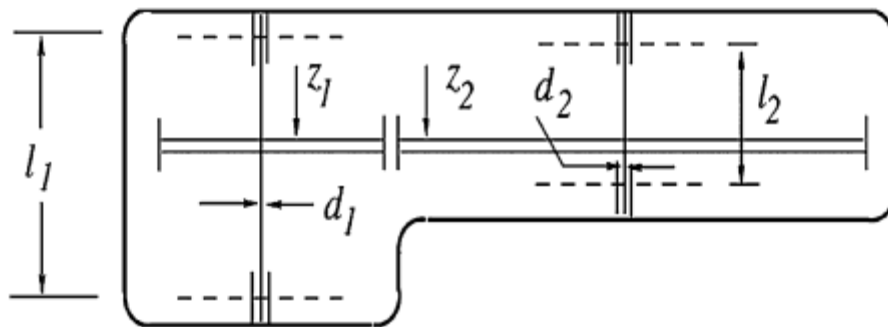
Το τελευταίο και μακράν πιο δύσκολο πρόβλημα ψευδομηχανικής από αυτά που θα εξεταστούν είναι το Speed Reducer Design <sup>[32]</sup>. Αφορά το σχεδιασμό ενός μειωτήρα στροφών, το εσωτερικό του οποίου φαίνεται στο Σχήμα 5.5, με σκοπό την ελαχιστοποίηση του βάρους του. Το πρόβλημα βελτιστοποίησης έχει συνολικά επτά μεταβλητές σχεδιασμού: το πάχος των οδοντωτών τροχών  $b$ , το μέτρο των δοντιών  $m$ , τον αριθμό των δοντιών στον κινητήριο τροχό  $z$ , τα μήκη του πρώτου και δεύτερου άξονα μεταξύ των στηρίξεων  $l_1$  και  $l_2$  και τις αντίστοιχες διαμέτρους των ιδίων αξόνων  $d_1$  και  $d_2$ . Η συνάρτηση-στόχος, δηλαδή το βάρος του μηχανήματος, προκύπτει ως:

$$f(\vec{X}) = 0.7854x_1x_2^2(3.3333x_3^2 + 14.9334x_3 - 43.0934) - 1.508x_1(x_6^2 + x_7^2) + 7.4777(x_6^3 + x_7^3) + 0.7854(x_4x_6^2 + x_5x_7^2) \quad (5.23)$$

όπου:

$$\vec{X} = \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{Bmatrix} = \begin{Bmatrix} b \\ m \\ z \\ l_1 \\ l_2 \\ d_1 \\ d_2 \end{Bmatrix} \quad (5.24)$$

με  $x_1 \in [2.6, 3.6]$ ,  $x_2 \in [0.7, 0.8]$ ,  $x_3 \in [17, 28]$ ,  $x_4 \in [7.3, 8.3]$ ,  $x_5 \in [7.3, 8.3]$ ,  $x_6 \in [2.9, 3.9]$  και  $x_7 \in [5.0, 5.5]$ . Επισημαίνεται ότι η μεταβλητή  $x_3$  που αφορά τον αριθμό των δοντιών μπορεί να πάρει μόνο ακέραιες τιμές λόγω της φυσικής της έννοιας, ενώ οι υπόλοιπες μεταβλητές είναι συνεχείς. Τα προβλήματα αυτού του είδους ονομάζονται mixed integer programming problems και είναι σημαντικά πιο δύσκολα στην επίλυση [32].



Σχήμα 5.5: Κάτοψη του μειωτήρα στροφών [32]

Οι περιορισμοί του προβλήματος είναι 11 στον αριθμό και αφορούν κυρίως τις τάσεις που αναπτύσσονται στα δόντια των τροχών ( $g_1$  και  $g_2$ ), τις παραμορφώσεις ( $g_3$  και  $g_4$ ) των δύο αξόνων, καθώς και τις τάσεις ( $g_5$  και  $g_6$ ) που αναπτύσσονται σε αυτούς. Οι υπόλοιποι περιορισμοί από  $g_7$  έως  $g_{11}$  αφορούν τη διαστασιολόγηση των εξαρτημάτων του μειωτήρα και αποτελούν μια μίξη εργονομικών και εμπειρικών κανονισμών [39]. Οι συναρτήσεις των περιορισμών είναι οι:

$$g_1(\vec{X}) = \frac{27.0}{x_1x_2^2x_3} - 1.0 \leq 0.0 \quad (5.25)$$

$$g_2(\vec{X}) = \frac{397.5}{x_1x_2^2x_3^2} - 1.0 \leq 0.0 \quad (5.26)$$

$$g_3(\vec{X}) = \frac{1.93x_4^3}{x_2x_6^4x_3} - 1.0 \leq 0.0 \quad (5.27)$$

$$g_4(\vec{X}) = \frac{1.93x_5^3}{x_2x_7^4x_3} - 1.0 \leq 0.0 \quad (5.28)$$

$$g_5(\vec{X}) = \frac{\sqrt{745\left[\frac{x_4}{x_2x_3}\right]^2 + 169 \times 10^5}}{110x_6^3} - 1.0 \leq 0.0 \quad (5.29)$$

$$g_6(\vec{X}) = \frac{\sqrt{745\left[\frac{x_5}{x_2x_3}\right]^2 + 1575 \times 10^5}}{85x_7^3} - 1.0 \leq 0.0 \quad (5.30)$$

$$g_7(\vec{X}) = \frac{x_2x_3}{40.0} - 1.0 \leq 0.0 \quad (5.31)$$

$$g_8(\vec{X}) = \frac{5x_2}{x_1} - 1.0 \leq 0.0 \quad (5.32)$$

$$g_9(\vec{X}) = \frac{x_1}{12x_2} - 1.0 \leq 0.0 \quad (5.33)$$

$$g_{10}(\vec{X}) = \frac{1.5x_6 + 1.9}{x_4} - 1.0 \leq 0.0 \quad (5.34)$$

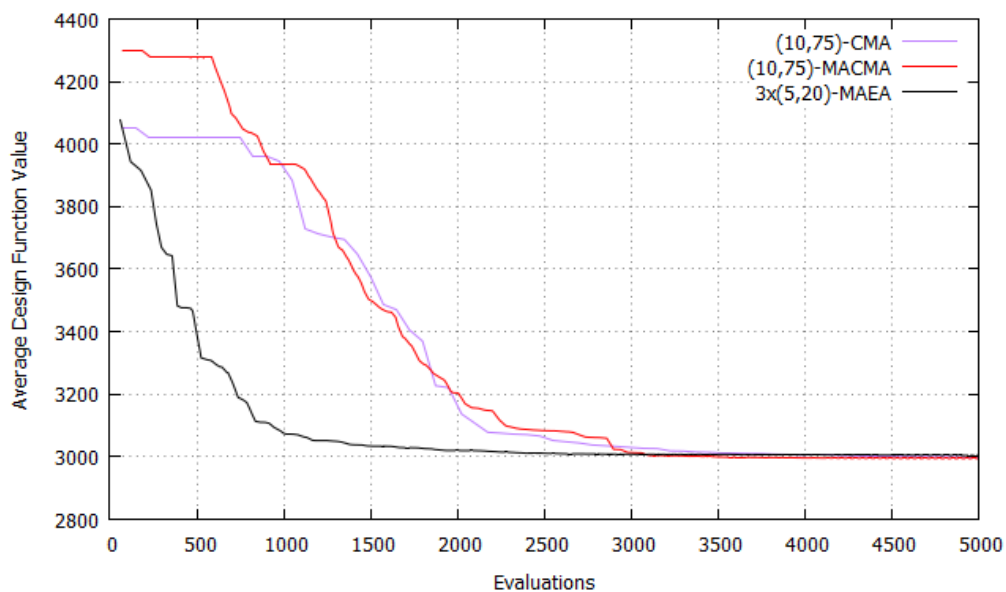
$$g_{11}(\vec{X}) = \frac{1.1x_7 + 1.9}{x_5} - 1.0 \leq 0.0 \quad (5.35)$$

Λόγω της περαιτέρω αύξησης στη δυσκολία του προβλήματος, οι επιτρεπόμενες αξιολογήσεις για κάθε τρέξιμο αυξάνονται σημαντικά στις 7000. Για πρώτη φορά, οι αλγόριθμοι CMA δυσκολεύονται να βρουν αποδεκτές λύσεις κατά τα αρχικά στάδια της βελτιστοποίησης και, συνεπώς, συγκλίνουν πρόωρα. Αποφασίζεται έτσι να αυξηθούν σημαντικά ο αριθμός απογόνων και το μέγεθος του αρχικού βήματος, ενώ παράλληλα ο τελεστής μετάλλαξης αποφασίζεται να μην ενεργοποιηθεί λόγω της έντονης ασυνέχειας του χωρίου σχεδιασμού που προκαλούν οι αυστηροί περιορισμοί. Λαμβάνοντας τα μέτρα αυτά, η αποτελεσματικότητα των CMA αναμένεται να μειωθεί σημαντικά κατά τις πρώτες γενιές, αλλά εν τέλει κρίνονται απαραίτητα για τη μακροπρόθεσμη επιτυχία και συνέπεια της βελτιστοποίησης. Αντιθέτως, ο DMAEA του EASY λειτουργεί ικανοποιητικά με τις ίδιες ρυθμίσεις που είχε προηγουμένως και έτσι παραμένει ίδιος. Παρατίθενται τα πλέον γνώριμα διαγράμματα σύγκλισης και οι συνοδευόμενοι πίνακες με τα στατιστικά στοιχεία των αλγορίθμων, ενώ ακολουθεί εκτενής σχολιασμός των αποτελεσμάτων.

**Πίνακας 5.9:** Ρυθμίσεις των αλγορίθμων CMA και DMAEA για το πρόβλημα σχεδιασμού του μειωτήρα στροφών.

CMA & MACMA			
Γονείς $\mu$	10	Χαλάρωση $c_m^{(0)}$	1.0
Απόγονοι $\lambda$	75	Αρχικό βήμα $\sigma^{(0)}$	3.0
Τελεστής επιλογής γονέων		Truncation selection	
Τελεστής δημιουργίας απογόνων		Weighted multi-recombination	
Πιθανότητα μετάλλαξης		0%	

DMAEA			
Γονείς $\mu$	5	Δήμοι	3
Απόγονοι $\lambda$	15	Ρυθμός ανταλλαγής	4 γενιές
Τελεστής επιλογής γονέων		Tournament selection	
Μέγεθος διαγωνισμού $t$		2	
Πιθανότητα επιλογής $p$		80%	
Τελεστής δημιουργίας απογόνων		2-point per variable crossover	
Πιθανότητα διασταύρωσης		80%	
Πιθανότητα μετάλλαξης		5%	



**Σχήμα 5.6:** Μέσο διάγραμμα σύγκλισης 10 τρεξιμάτων για το πρόβλημα σχεδιασμού του μειωτήρα στροφών.

**Πίνακας 5.10:** Στατιστική ανάλυση των αποτελεσμάτων για το πρόβλημα σχεδιασμού του μειωτήρα στροφών από την εκτέλεση 10 τρεξιμάτων με 7000 επιτρεπόμενες αξιολογήσεις.

Αλγόριθμος	CMA	MACMA	DMAEA
Βέλτιστη Τιμή	2994.486300186	2994.471290249	2997.831
Χείριστη Τιμή	3001.217924513	2994.472717096	3019.410
Μέση Τιμή	2995.770993295	2994.471757905	3002.998
Διάμεσος	2994.600219657	2994.471641999	3001.845
Τυπική Απόκλιση	2.38	4.51E-04	5.94

**Πίνακας 5.11:** Στατιστικά αποτελεσμάτων άλλων αλγορίθμων για το πρόβλημα σχεδιασμού του μειωτήρα στροφών.

Αλγόριθμος	SC <sup>[31]</sup>	MBA <sup>[32]</sup>	ITLBO <sup>[33]</sup>	HBBO <sup>[34]</sup>	DAPSO <sup>[35]</sup>
Αξιολογήσεις	54456	6300	10000	25100	7320
Βέλτιστη Τιμή	2994.744	2994.482453	2994.471066	2996.348165	2994.471066
Χείριστη Τιμή	3009.964	2999.652444	2994.471066	2996.348165	2994.471366
Μέση Τιμή	3001.758	2996.769019	2994.471066	2996.348165	2994.471072
Διάμεσος	3001.758	-	-	2996.348165	-
Τυπ. Απόκλιση	4.01	1.56	4.62E-13	1.14E-15	1.61E-05

**Πίνακας 5.12:** Βέλτιστες τιμές της συνάρτησης-στόχου και αντίστοιχες τιμές των μεταβλητών σχεδιασμού για το πρόβλημα σχεδιασμού του μειωτήρα στροφών με 7000 επιτρεπόμενες αξιολογήσεις.

Αλγόριθμος	CMA	MACMA	DMAEA
$f(\vec{X})$	2994.486300186	2994.471290249	2997.831792511
$x_1$	3.500001595	3.500000092	3.500293255
$x_2$	0.700000104	0.700000000	0.7
$x_3$	17	17	17
$x_4$	7.300380487	7.300000617	7.306842619
$x_5$	7.715392154	7.715322558	7.737927663
$x_6$	3.350245292	3.350214763	3.351612903
$x_7$	5.286656611	5.286654572	5.290322580



Στο Σχήμα 5.6 διακρίνεται ένας τελείως διαφορετικός τρόπος σύγκλισης μεταξύ των αλγορίθμων, πράγμα που παρατηρείται για πρώτη φορά στις δοκιμές. Αρχικά ο DMAEA, ο οποίος όπως προαναφέρθηκε δεν χρειάζεται καμία αύξηση στον πληθυσμό για να ανταπεξέλθει στο νέο πρόβλημα, παρουσιάζει ταχύτερη σύγκλιση έως τις 1000 αξιολογήσεις με μέση τιμή σχεδιασμού κάτω από 3100. Αντιθέτως, σε αυτό το στάδιο, οι δύο CMA μόλις έχουν αρχίσει να ανιχνεύουν αποδεκτές λύσεις που τηρούν όλους τους περιορισμούς και μπορούν επιτέλους να προσαρμόσουν ορθά το μητρώο συνδιακύμανσης. Στο διάστημα από τις 1000 έως τις 3000 αξιολογήσεις ο DMAEA συνεχίζει να βελτιώνει την τιμή σχεδιασμού με πιο χαμηλό ρυθμό εν συγκρίσει με πριν, ενώ οι CMA πλέον λειτουργούν επιθυμητά και τον έχουν σχεδόν φτάσει. Από εδώ και έως το τέλος της βελτιστοποίησης στις 7000 αξιολογήσεις, ο DMAEA αρχίζει να καθυστερεί σημαντικά και καταφέρνει μια μικρή βελτίωση της τιμής σχεδιασμού από 3008 σε 3003. Ο τροποποιημένος CMA, με και χωρίς χρήση μεταμοντέλων, τον ξεπερνάει τελικά στις 3975 και 3085 αξιολογήσεις αντίστοιχα.

Το ενδιαφέρον σχετικά με τον CMA είναι ότι, ενώ στα προηγούμενα προβλήματα τα μεταμοντέλα καταφέρνουν να βελτιώσουν μόνο την πορεία σύγκλισης, εδώ υπερτερούν κατά μεγάλο βαθμό και βάσει των στατιστικών αποτελεσμάτων του Πίνακα 5.10. Η διαφορά στη μέση τιμή και στην τυπική απόκλιση της συνάρτησης-στόχου δείχνει ότι ο MACMA επιδεικνύει πολύ μεγαλύτερη συνέπεια στην τελική ποιότητα της λύσης. Αυτό οφείλεται σε μεγάλο βαθμό στ' ότι οι 75 απόγονοι είναι πραγματικά μεγάλος αριθμός, ειδικά για μία ES. Έτσι, ο απλός CMA τελεί περιττές αξιολογήσεις ανά γενιά ενώ θα μπορούσε να προσαρμόσει το μητρώο συνδιακύμανσης κάλλιστα και με λιγότερες, αλλά για μικρότερους πληθυσμούς δεν θα εντόπιζε αποδεκτές λύσεις λόγω των αυστηρών περιορισμών στην αρχή της βελτιστοποίησης και θα αποτύγχανε. Αντιθέτως, ο MACMA κάνει 75 αξιολογήσεις κατά τις πρώτες γενιές ώσπου να πληρωθεί η βάση δεδομένων για την εκπαίδευση των μεταμοντέλων και μετά λειτουργεί με πολύ λιγότερες. Αυτό έχει ως αποτέλεσμα όχι μόνο να βρει την καλύτερη τιμή συνάρτησης σχεδιασμού αλλά και να έχει πολύ μικρότερη τυπική απόκλιση στα 10 τρεξίματα.

Σύμφωνα με τον Πίνακα 5.11, η βέλτιστη τιμή σχεδιασμού είναι 2994.471066 από τους αλγόριθμους ITLBO και DAPSO. Συγκρίνοντάς τους, ο DAPSO χρειάστηκε περίπου 3000 λιγότερες αξιολογήσεις οπότε και μπορεί να θεωρηθεί ως η καλύτερη μέθοδος από αυτές που παρουσιάζονται στον πίνακα. Ο MACMA έχει ως βέλτιστη τιμή 2994.471290,

δηλαδή παρουσιάζει απόκλιση στο 4<sup>ο</sup> δεκαδικό με 320 λιγότερες αξιολογήσεις. Η στατιστική ανάλυση του DAPSO είναι επίσης ελαφρώς καλύτερη απ' ό τι του MACMA. Αν και δεν μπορεί να προκύψει σαφές συμπέρασμα για το ποια είναι η βέλτιστη μεταξύ των δύο μεθόδων για το πρόβλημα του μειωτήρα στροφών, το σίγουρο είναι ότι ο MACMA είναι ο μόνος από τους αλγορίθμους που εξετάστηκαν που μπορεί να ανταγωνιστεί την επίδοση του DAPSO.

## ΚΕΦΑΛΑΙΟ 6

### ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΚΛΕΙΣΤΟΥ ΑΓΩΓΟΥ ΜΕ ΚΑΜΠΗ 90°

Η παρατηρούμενη αποτελεσματικότητα του τροποποιημένου αλγόριθμου CMA κατά την επίλυση των προβλημάτων ψευδομηχανικής αποτελεί ένα θετικό, πρώτο βήμα προς την πιστοποίησή του. Παρόλα αυτά, προτού ληφθούν τα τελικά συμπεράσματα περί της αποδοτικότητάς του, οφείλει να δοκιμαστεί και σε πραγματικά προβλήματα σχεδιασμού της επιστήμης του μηχανικού, λ.χ. στη μηχανική των ρευστών, η οποία αποτελεί την περιοχή επιστημονικού ενδιαφέροντος της ερευνητικής ομάδας στην οποία εκπονήθηκε η μεταπτυχιακή αυτή εργασία. Έτσι, ως τελική δοκιμή του τροποποιημένου CMA επιλέγεται να εξεταστεί μια εφαρμογή ρευστοδυναμικής που αφορά τη βελτιστοποίηση μορφής ενός κλειστού αγωγού με καμπή 90°. Αρχικά, παρουσιάζονται οι εξισώσεις του φυσικού προβλήματος που επιλύεται με το ανοιχτό λογισμικό Υπολογιστικής Ρευστοδυναμικής (Computational Fluid Dynamics: CFD) OpenFOAM. Κατόπιν ορίζεται το πρόβλημα βελτιστοποίησης, διαδικασία που εμπεριέχει τον ορισμό της συνάρτησης-στόχου και των μεταβλητών σχεδιασμού. Το πρόβλημα αυτό επιλύεται από τον CMA, με και χωρίς μεταμοντέλα, καθώς και από τον γενικευμένο EA του λογισμικού EASY, ώστε να διαπιστωθεί ποιος από τους αλγόριθμους έχει την καλύτερη επίδοση.

#### 6.1 ΕΠΙΛΥΣΗ ΡΟΗΣ ΣΕ ΚΛΕΙΣΤΟ ΑΓΩΓΟ ΜΕ ΚΑΜΠΗ 90°

Σε προβλήματα ρευστοδυναμικής σε κλειστούς αγωγούς, το πρόβλημα της ροής καταλήγει στον αριθμητικό υπολογισμό των πεδίων στατικής πίεσης  $p$  [Pa] και του μέτρου της ταχύτητας  $U$  [m/s] του ρευστού. Η ροή θεωρείται μόνιμη, δηλαδή οι συνιστώσες της ταχύτητας του ρευστού  $u_i$  σε κάθε σημείο του πεδίου ροής παραμένουν σταθερές προς το χρόνο, και ασυμπίεστη, δηλαδή η πυκνότητα του ρευστού  $\rho$  [kg/m<sup>3</sup>] παραμένει επίσης σταθερή ως προς το χρόνο και τον χώρο. Το εν λόγω φυσικό πρόβλημα διέπεται από τις διδιάστατες, χρονικά μόνιμες εξισώσεις Navier-Stokes, οι οποίες για ασυμπίεστη, στρωτή ροή προκύπτουν σε τανυστική γραφή ως:

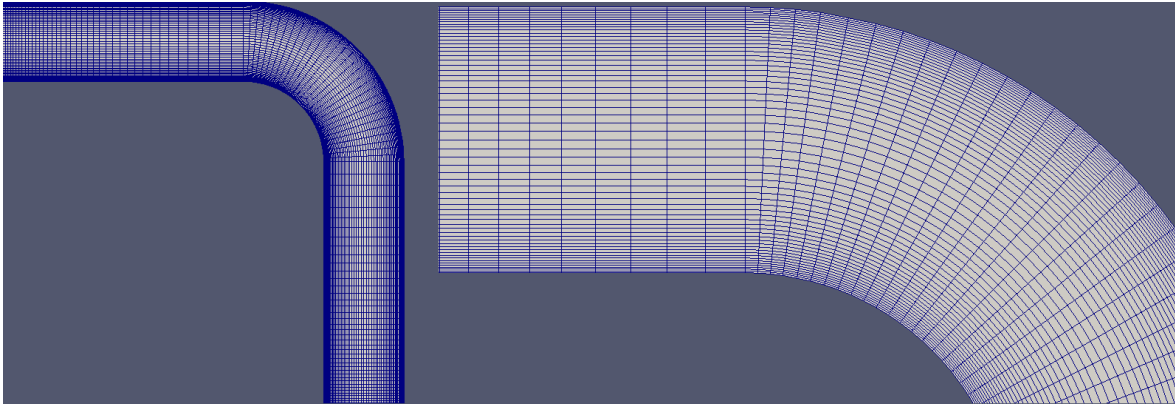
$$\frac{\partial u_i}{\partial x_i} = 0 \quad (6.1)$$

$$u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} \quad (6.2)$$

όπου  $\nu$  είναι η κινηματική συνεκτικότητα του ρευστού [ $\text{m}^2/\text{s}$ ] και οι δείκτες  $i$  και  $j$  λαμβάνουν τιμή 1 όταν αναφέρονται στον άξονα  $x_1 = x$  και τιμή 2 όταν αναφέρονται στον άξονα  $x_2 = y$ . Ως επακόλουθο, οι συνιστώσες της ταχύτητας  $u_1$  και  $u_2$  αντιστοιχούν στις κατά  $x$  και  $y$  συνιστώσες της ταχύτητας.

Η αριθμητική επίλυση των εξισώσεων Navier-Stokes υλοποιείται με τη χρήση ενός λογισμικού CFD. Ένα τέτοιο λογισμικό διακριτοποιεί τις διαφορικές εξισώσεις της ροής, χρησιμοποιεί επαναληπτικές μεθόδους για να τις λύσει αριθμητικά και τερματίζεται όταν η λύση σταθεροποιηθεί με κάποια ακρίβεια σφάλματος υπολοίπου που καθορίζει ο χρήστης. Τα στάδια χρήσης ενός λογισμικού CFD είναι τρία. Πρώτον, η σύνθεση του προβλήματος προς επίλυση, διαδικασία που περιλαμβάνει τόσο τη δημιουργία της γεωμετρίας του όσο και τη ρύθμιση των δεδομένων του προβλήματος, όπως είναι π.χ. οι οριακές συνθήκες, η επιλογή του επιλύτη, η γένεση του υπολογιστικού πλέγματος κτλ. Το επόμενο στάδιο εμπεριέχει την αριθμητική επίλυση του προβλήματος από το λογισμικό, η οποία ανάλογα την υπολογιστική ισχύ αλλά και το μέγεθος και την πολυπλοκότητα του προβλήματος, μπορεί να διαρκέσει από μερικά μόνο λεπτά έως και πολλές μέρες σε έναν ή περισσότερους επεξεργαστές. Τέλος, τα αποτελέσματα επεξεργάζονται από το χρήστη με την κατασκευή διαγραμμάτων και σχημάτων σε λογισμικά σχεδίασης και γραφικών.

Ξεκινώντας με τις ρυθμίσεις, δημιουργείται στο λογισμικό OpenFoam ένας κλειστός 2D αγωγός που αποτελείται από δύο ευθύγραμμα τμήματα μήκους 0.6 m και διαμέτρου  $D = 0.2$  m (αποτελούν την είσοδο και την έξοδο), τα οποία ενώνονται με μια καμπή  $90^\circ$  χωρίς καθοδηγητικά πτερύγια. Η καμπή έχει μέση ακτίνα  $R = 0.3$  m, οπότε το συνολικό μήκος του αγωγού είναι  $L = 2 * 0.6 + 2\pi R/4 = 1.67$  m. Επιλέγεται να επιλυθεί μία στρωτή ροή προκειμένου να μειωθούν οι χρονικές απαιτήσεις της βελτιστοποίησης, οπότε η μέση ταχύτητα του ρευστού τίθεται ίση με  $U = 1$  m/s και η κινηματική συνεκτικότητα ορίζεται ως  $\nu = 0.01$   $\text{m}^2/\text{s}$ . Προκύπτει έτσι ο αριθμός Reynolds  $Re = UD/\nu = 20$ , ο οποίος εξασφαλίζει δημιουργία στρωτής ροής <sup>[40]</sup>. Τα παραπάνω μεγέθη παραμένουν αναλλοίωτα για το υπόλοιπο της εργασίας. Στο Σχήμα 6.1 δίνεται η τελική γεωμετρία του αγωγού, μαζί με το υπολογιστικό πλέγμα που θα χρησιμοποιηθεί για την επίλυση. Το δομημένο πλέγμα, διαστάσεων  $110 \times 30$ , δημιουργείται αυτόματα από το λογισμικό OpenFOAM με τη λειτουργία blockMesh, ενώ παράλληλα ρυθμίζεται ώστε να είναι πυκνότερο κοντά στα στερεά τοιχώματα του αγωγού, όπως πρέπει να γίνεται κατά την επίλυση προβλημάτων CFD.



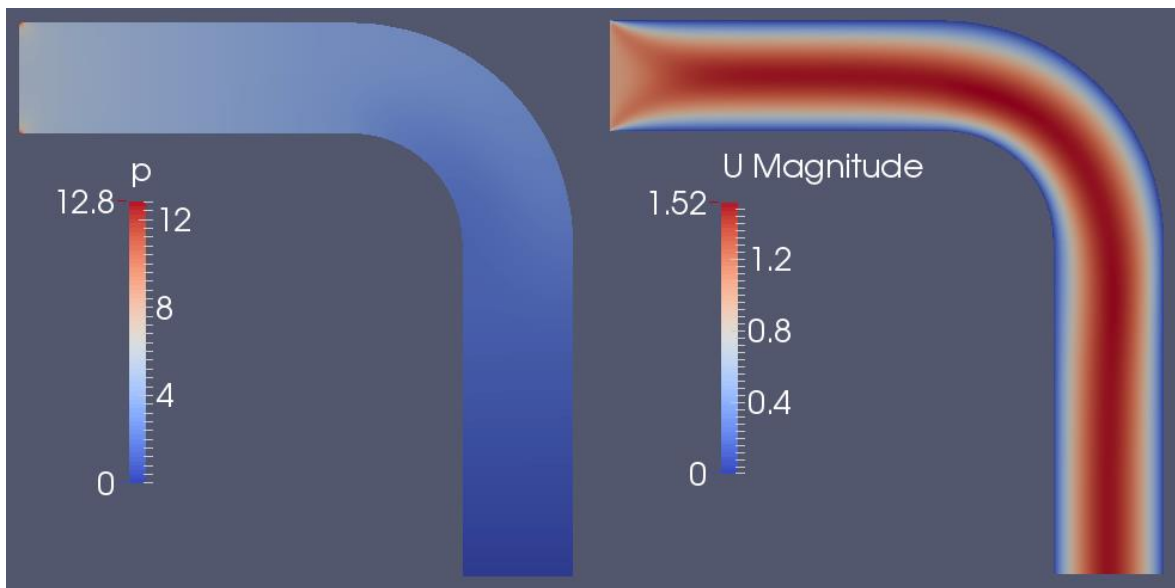
**Σχήμα 6.1:** Κλειστός αγωγός με καμπή  $90^\circ$  και δημιουργηθέν υπολογιστικό πλέγμα  $110 \times 30$  με πύκνωση στα σύνορα (αριστερά). Λεπτομέρειες του πλέγματος στην ευρύτερη περιοχή της καμπής (δεξιά).

Έχοντας τη γεωμετρία έτοιμη, στη συνέχεια πραγματοποιείται η ρύθμιση του επιλύτη. Αρχικά, δηλώνεται στο λογισμικό ότι το πρόβλημα αφορά μόνιμη ροή ενός ασυμπίεστου και νευτώνειου ρευστού, ώστε να γνωρίζει ότι πρέπει να λυθούν οι Εξισώσεις (6.1) και (6.2). Η επιτρεπόμενη μέγιστη τιμή υπολοίπου τίθεται σε  $10^{-7}$  και απομένει μόνο να οριστούν οι οριακές συνθήκες του προβλήματος. Για την επίλυση των εξισώσεων Navier-Stokes χρησιμοποιούνται δύο είδη οριακών συνθηκών: συνθήκες τύπου Dirichlet, οι οποίες επιβάλλονται σε κόμβους όπου είναι γνωστή η τιμή της μεταβλητής, και συνθήκες τύπου Neumann, όπου είναι γνωστή η τιμή της παραγώγου της μεταβλητής. Οι οριακές συνθήκες επιβάλλονται στις διατομές εισόδου, εξόδου και κατά μήκος των στερεών τοιχωμάτων του αγωγού ως εξής:

- Για τη στατική πίεση  $p$ :
  1. Μηδενική συνθήκη Neumann στην είσοδο:  $\frac{\partial p}{\partial x} = 0$
  2. Μηδενική συνθήκη Neumann στα τοιχώματα:  $\frac{\partial p}{\partial n} = 0$
  3. Μηδενική συνθήκη Dirichlet στην έξοδο:  $p = 0$
- Για τη  $x$  συνιστώσα της ταχύτητας  $u_1$ :
  1. Μοναδιαία συνθήκη Dirichlet στην είσοδο:  $u_1 = 1$
  2. Μηδενική συνθήκη Dirichlet στα τοιχώματα:  $u_1 = 0$
  3. Μηδενική συνθήκη Neumann στην έξοδο:  $\frac{\partial u_1}{\partial x} = 0$
- Για τη  $y$  συνιστώσα της ταχύτητας  $u_2$ :
  1. Μηδενική συνθήκη Dirichlet στην είσοδο:  $u_2 = 0$
  2. Μηδενική συνθήκη Dirichlet στα τοιχώματα:  $u_2 = 0$

### 3. Μηδενική συνθήκη Neumann στην έξοδο: $\frac{\partial u_2}{\partial y} = 0$

όπου  $n$  είναι η κάθετη στο στερεό τοίχωμα διεύθυνση. Οι ρυθμίσεις έχουν πλέον ολοκληρωθεί, οπότε το πρόβλημα λύνεται από το OpenFoam και λαμβάνονται τα αποτελέσματα του Σχήματος 6.2. Στο εν λόγω σχήμα παρατηρείται ότι η κατανομή της ταχύτητας  $U$  είναι ομοιόμορφη στην είσοδο λόγω της οριακής συνθήκης Dirichlet που έχει επιβληθεί και αρχίζει αμέσως να μεταβάλλεται ως ένα μήκος. Από εκεί και ύστερα η ροή φαίνεται ότι είναι πλήρως διαμορφωμένη, αφού διατηρεί την παραβολική, συμμετρική ως προς τον άξονα του αγωγού, κατανομή ταχύτητας που αναμένεται λόγω των συνθηκών μη ολίσθησης στα στερεά τοιχώματα του αγωγού. Βάσει των παραπάνω παρατηρήσεων το OpenFoam θεωρείται ότι, με τις παρούσες ρυθμίσεις, μπορεί να λειτουργήσει αξιόπιστα ως λογισμικό αξιολόγησης για το πρόβλημα βελτιστοποίησης (με CMA ή και EA) του σχεδιασμού της καμπής  $90^\circ$ , το οποίο ορίζεται στην επόμενη ενότητα.



**Σχήμα 6.2:** Πεδία της στατικής πίεσης  $p$  και του μέτρου της ταχύτητας  $U$  για το πρόβλημα μόνιμης και ασυμπίεστης ροής σε κλειστό αγωγό με καμπή  $90^\circ$ .

## 6.2 ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΤΟΥ ΣΧΕΔΙΑΣΜΟΥ ΚΑΜΠΗΣ $90^\circ$

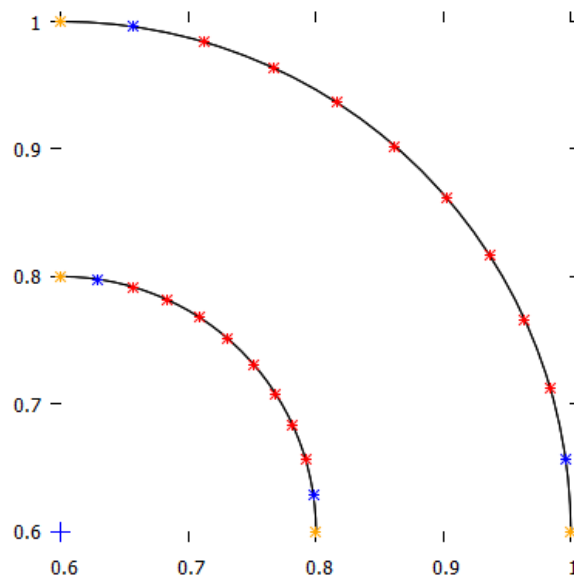
Το πρόβλημα βελτιστοποίησης αφορά την ελαχιστοποίηση των απωλειών ολικής πίεσης που παρατηρούνται κατά τη ροή στον κλειστό αγωγό του Σχήματος 6.1, κυρίως λόγω της ύπαρξης της καμπής  $90^\circ$ . Κατά τα πλέον γνωστά, για να οριστεί πλήρως ένα πρόβλημα βελτιστοποίησης απαιτείται μία συνάρτηση-στόχος, εξαρτώμενη από τις

λεγόμενες μεταβλητές σχεδιασμού. Οι συνολικές απώλειες ολικής πίεσης στο δίκτυο υπολογίζονται ως <sup>[41]</sup>:

$$f = - \int_{S_{I,O}} \left( \frac{p}{\rho} + \frac{U^2}{2} \right) \vec{U}_i \cdot \vec{n}_i dS \quad (6.3)$$

όπου  $S_I$  και  $S_O$  είναι οι διατομές εισόδου (inlet) και εξόδου (outlet) αντίστοιχα,  $\vec{U}_i$  είναι το μοναδιαίο διάνυσμα της ταχύτητας  $U$  και  $\vec{n}_i$  είναι το μοναδιαίο διάνυσμα κάθετο στη διατομή  $S$ . Η Εξίσωση (6.3) υπολογίζει τις συνολικές απώλειες κατά μήκος ολόκληρου του αγωγού, στις οποίες φυσικά υπεισέρχονται και οι απώλειες που οφείλονται αποκλειστικά στην καμπή.

Για τις μεταβλητές σχεδιασμού ορίζονται 12 ισαπέχοντα σημεία κατά μήκος κάθε τοιχώματος της καμπής, τα οποία αθροίζουν σε 24 σημεία συνολικά όπως απεικονίζεται στο Σχήμα 6.3. Η μετακίνηση οποιουδήποτε σημείου από αυτά έχει ως αποτέλεσμα την αλλαγή της γεωμετρίας της καμπής, οπότε και την αλλαγή των χαρακτηριστικών της ροής στο εσωτερικό της. Τα ακραία, κίτρινα του Σχήματος 6.3, σημεία δεν πρέπει σε καμία περίπτωση να μετακινηθούν, διαφορετικά θα επηρεαστεί η οριοθέτηση της καμπής και η ασυνέχεια με τα γειτονικά ευθέα τμήματα. Το ίδιο ισχύει και για τα αμέσως επόμενα, μπλε σημεία, τα οποία εξασφαλίζουν συνέχεια κλίσης (παραγώγου) του στερεού τοιχώματος στις περιοχές αυτές. Τα εναπομείναντα 16 σημεία όμως αφορούν αποκλειστικά την καμπή, οπότε τους επιτρέπεται να μετακινηθούν ακτινικά έως το 20% του εύρους του αγωγού ως προς το κέντρο της καμπής. Τα νέα τοιχώματα της καμπής δημιουργούνται παρεμβάλλοντας τα μετατοπισμένα σημεία με splines, διαδικασία που τελείται αυτόματα από το OpenFOAM. Συνοψίζοντας όλα τα παραπάνω, το πρόβλημα ελαχιστοποίησης εμπεριέχει 16 μεταβλητές σχεδιασμού  $x_i \in [-0.04, 0.04]$ , οι οποίες είναι οι ακτινικές μετακινήσεις των κόκκινων σημείων ως προς το κέντρο της καμπής, μετρούμενες από την αρχική τους θέση. Αν και οι μεταβλητές σχεδιασμού δεν εμφανίζονται άμεσα στη συνάρτηση-στόχο της Εξίσωσης (6.3) όπως συνηθίζεται, επηρεάζουν έμμεσα τις τιμές της στατικής πίεσης  $p$  και του μέτρου της ταχύτητας  $U$  στο εσωτερικό της καμπής κατά την επίλυση.



**Σχήμα 6.3:** Διαχωρισμός της καμπής  $90^\circ$  σε 24 ισαπέχοντα σημεία. Επιτρέπεται μόνο στα κόκκινα σημεία να μετακινούνται ακτινικά ως προς το κέντρο των ομόκεντρων κύκλων, το οποίο σημειώνεται με τον σταυρό στο  $(0.6, 0.6)$ .

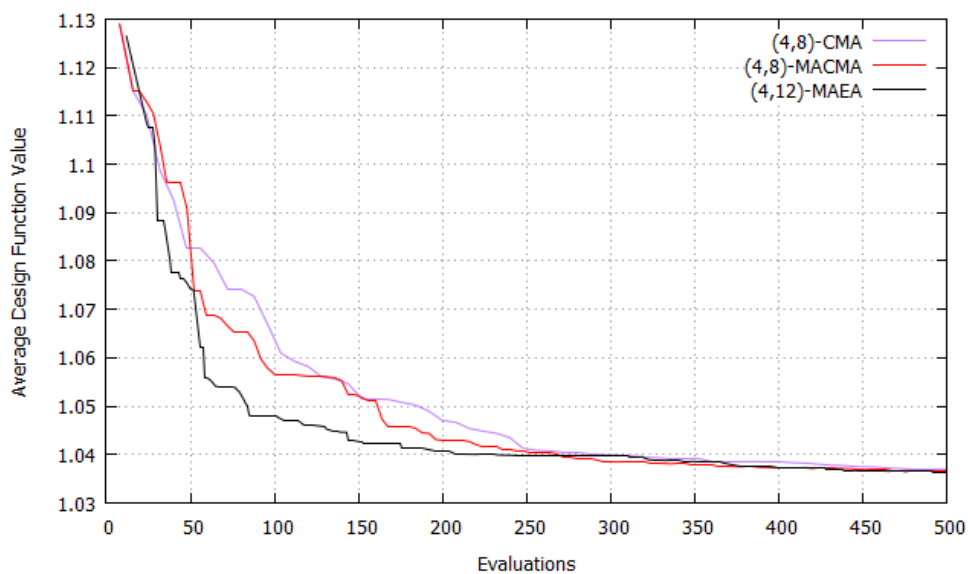
Το πρόβλημα της βελτιστοποίησης είναι άνευ περιορισμών και επιλύεται με όλους τους ΕΑ που εξετάστηκαν και προηγουμένως. Λόγω του αυξημένου υπολογιστικού φόρτου που διέπει τη χρήση ενός λογισμικού CFD, πραγματοποιούνται 2 ενδεικτικά τρεξίματα από κάθε αλγόριθμο με 500 επιτρεπόμενες αξιολογήσεις. Η απουσία περιορισμών επιτρέπει σε όλους τους ΕΑ να αντιμετωπίσουν το πρόβλημα με μικρούς πληθυσμούς και μάλιστα, όπως προκύπτει από τα αποτελέσματα, ο MAEA του EASY μπορεί να ανταγωνιστεί και να ξεπεράσει την επίδοση του CMA χωρίς τη χρήση δόμων, αφού κατάφερε να εντοπίσει την καλύτερη τιμή της συνάρτησης-στόχου. Για τους αλγόριθμους CMA αναφέρεται ότι, λόγω των πολυάριθμων μεταβλητών σχεδιασμού, το αρχικό μέγεθος βήματος  $\sigma^{(0)}$  τίθεται σημαντικά μειωμένο σε 0.2 και η πιθανότητα μετάλλαξης μηδενίζεται. Ακολουθούν οι πίνακες με τις ρυθμίσεις των ΕΑ και των βέλτιστων τιμών σχεδιασμού, το μέσο διάγραμμα σύγκλισης του προβλήματος βελτιστοποίησης και τα αποτελέσματα της επίλυσης του φυσικού προβλήματος σχεδιασμένα πάνω στη νέα, βέλτιστη γεωμετρία της καμπής, όπως αυτή τελικά προέκυψε από τον αλγόριθμο MAEA του EASY.



**Πίνακας 6.1:** Ρυθμίσεις των αλγορίθμων CMA και MAEA για τη βελτιστοποίηση του σχεδιασμού καμπής 90°.

CMA & MACMA			
Γονείς $\mu$	4	Χαλάρωση $c_m^{(0)}$	1.0
Απόγονοι $\lambda$	8	Αρχικό βήμα $\sigma^{(0)}$	0.2
Τελεστής επιλογής γονέων	Truncation selection		
Τελεστής δημιουργίας απογόνων	Weighted multi-recombination		
Πιθανότητα μετάλλαξης	0%		

MAEA	
Γονείς $\mu$	4
Απόγονοι $\lambda$	12
Τελεστής επιλογής γονέων	Tournament selection
Μέγεθος διαγωνισμού $t$	2
Πιθανότητα επιλογής $p$	90%
Τελεστής δημιουργίας απογόνων	2-point per variable crossover
Πιθανότητα διασταύρωσης	90%
Πιθανότητα μετάλλαξης	5%



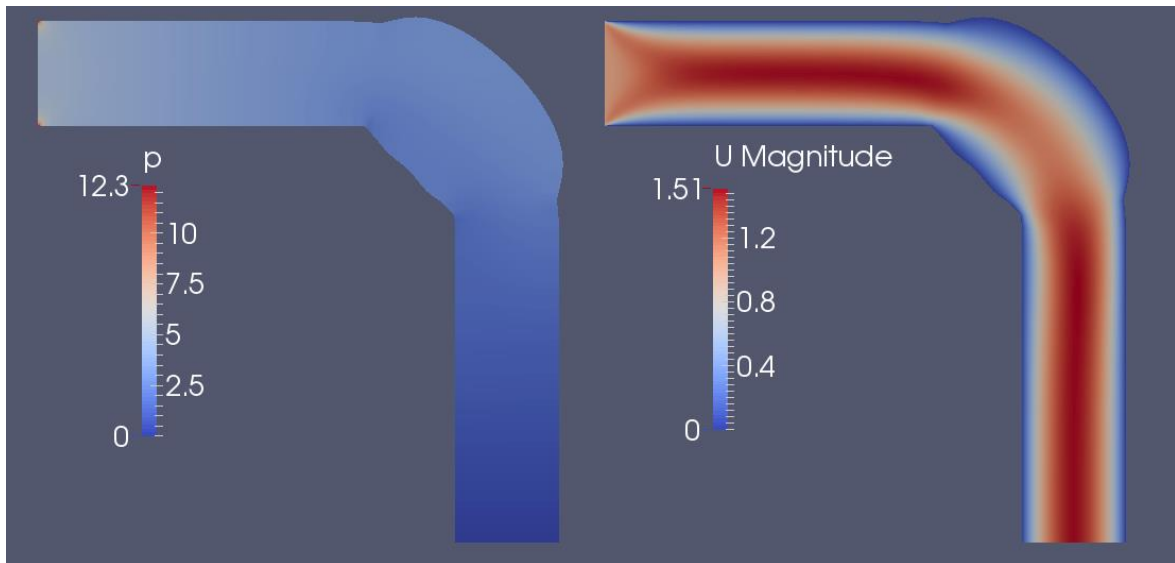
**Σχήμα 6.4:** Μέσο διάγραμμα σύγκλισης των 2 τρεξιμάτων (ανά μέθοδο) που πραγματοποιήθηκαν για τη βελτιστοποίηση του σχεδιασμού κλειστού αγωγού με καμπή 90°.

**Πίνακας 6.2:** Βέλτιστες τιμές της συνάρτησης-στόχου και αντίστοιχες τιμές των μεταβλητών σχεδιασμού για τη βελτιστοποίηση του σχεδιασμού κλειστού αγωγού με καμπή 90°. Οι μεταβλητές αριθμούνται με αύξουσα σειρά, από τα ανάντη προς τα κατάντη και από κάτω προς τα πάνω όπως φαίνεται στο Σχήμα 6.3, επομένως οι μεταβλητές από  $x_1$  έως  $x_8$  αφορούν τη μετατόπιση του κάτω τοιχώματος και από  $x_9$  έως  $x_{16}$  του επάνω.

Αλγόριθμος	CMA	MACMA	ΜΑΕΑ
f	1.036678116	1.036370428	1.036324772
$x_1$	-0.017530895	-0.015245467	-0.020136852
$x_2$	-0.039232390	-0.039107525	-0.034838709
$x_3$	-0.039707753	-0.039802876	-0.04
$x_4$	-0.039805197	-0.039640533	-0.039374389
$x_5$	-0.039442099	-0.037751450	-0.04
$x_6$	-0.039752409	-0.037876995	-0.039843597
$x_7$	-0.036661662	-0.036881999	-0.034995112
$x_8$	-0.014781881	-0.017357096	-0.014271749
$x_9$	0.026408550	0.038487956	0.030928641
$x_{10}$	0.037582802	0.038889808	0.038826979
$x_{11}$	0.035946142	0.039991554	0.039765395
$x_{12}$	0.039479085	0.039765447	0.038279569
$x_{13}$	0.039454142	0.039446614	0.04
$x_{14}$	0.038755397	0.038878787	0.039843597
$x_{15}$	0.036996486	0.039789568	0.039921798
$x_{16}$	0.023856541	0.026934829	0.029990224

Στο Σχήμα 6.5 διακρίνεται ότι η διεύρυνση της καμπής οδηγεί σε σημαντικά μειωμένες απώλειες ολικής πίεσης στο δίκτυο του αγωγού. Παρατηρώντας την Εξίσωση (6.3), είναι εμφανές ότι η μείωση της ταχύτητας  $U$  οδηγεί και σε πτώση των απωλειών της ολικής πίεσης. Ο τρόπος για να μειωθεί η ταχύτητα ενός ρευστού σε οποιοδήποτε αγωγό είναι η αύξηση της διατομής  $D$ , η οποία συνεπάγεται με αύξηση του εμβαδού διατομής  $A$  [ $m^2$ ], αφού η ογκομετρική παροχή  $Q = UA$  [ $m^3/s$ ] παραμένει πάντα σταθερή λόγω της αρχής διατήρησης της μάζας. Ως αποτέλεσμα, ενώ οι απώλειες ολικής πίεσης στην εφαρμογή της απλής, αρχικής καμπής του Σχήματος 2 ισούνται περίπου με  $1.134 m^5/s^3$ ,

με το πέρας της βελτιστοποίησης βρίσκονται σημαντικά μειωμένες κατά 8.6% σε  $1.036 \text{ m}^5/\text{s}^3$ .



**Σχήμα 6.5:** Πεδία της στατικής πίεσης  $p$  και του μέτρου της ταχύτητας  $U$  στον κλειστό αγωγό με τη νέα, βελτιωμένη καμπή  $90^\circ$ , σύμφωνα με τις τιμές σχεδιασμού του αλγόριθμου MAEA του λογισμικού EASY.

Αναφορικά με τη διαδικασία της βελτιστοποίησης, ο MAEA κατάφερε να εντοπίσει τον καλύτερο σχεδιασμό της καμπής με διαφορά στο 5<sup>ο</sup> δεκαδικό ψηφίο της συνάρτησης-στόχου, συγκριτικά με την αμέσως καλύτερη τιμή του MACMA. Στο Σχήμα 6.4 φαίνεται ότι ο τροποποιημένος CMA χωρίς μεταμοντέλα έχει εμφανώς τη χειρότερη σύγκλιση, αφού σε κανένα σημείο της βελτιστοποίησης δεν κατάφερε να προηγηθεί. Ο MAEA έχει εμφανή κυριαρχία κατά το πρώτο μισό της βελτιστοποίησης και ανταγωνίζεται τον MACMA κατά το δεύτερο μισό, καταφέροντας τελικά να τον ξεπεράσει. Συγκρίνοντας τις τελικές τιμές σχεδιασμού για κάθε αλγόριθμο στον Πίνακα 6.2, παρατηρούνται σημαντικές διακυμάνσεις στις τιμές των περισσότερων μεταβλητών σχεδιασμού με διαφορές ακόμη και στο 2<sup>ο</sup> δεκαδικό. Κανένας αλγόριθμος δεν φαίνεται να έχει συγκλίνει πλήρως, πράγμα που εκμαιεύεται και από τα διαγράμματα σύγκλισης, όπου η πτωτική πορεία δείχνει να συνεχίζεται έως και το μέγιστο αριθμό αξιολογήσεων από όλους τους αλγόριθμους. Αν και 2 τρεξίματα ενός προβλήματος δεν αποτελούν ικανοποιητικό δείγμα και σαφή συμπεράσματα αναφορικά με το ποιος αλγόριθμος μπορεί εν τέλει να εντοπίσει την καλύτερη λύση δεν μπορούν να ληφθούν, η έντονη υπεροχή του MAEA έναντι του CMA στην αρχή της βελτιστοποίησης είναι ένα στοιχείο που παρατηρήθηκε και στο πρόβλημα του μειωτήρα στροφών, οπότε μπορεί να θεωρηθεί αντιπροσωπευτική.

## **ΚΕΦΑΛΑΙΟ 7**

### **ΣΥΜΠΕΡΑΣΜΑΤΑ**

Ανακεφαλαιώνοντας, στην εργασία εξετάστηκαν όλες οι πτυχές ενός εξελικτικού αλγορίθμου που εφαρμόζει μια τροποποιημένη εκδοχή της προσαρμογής του μητρώου συνδιακύμανσης, όπως αυτή ορίστηκε από τον κλασικό αλγόριθμο CMA των Hansen και Ostermeier. Έγινε αναφορά σε βασικές έννοιες της εξελικτικής βελτιστοποίησης, στις διεργασίες που χρησιμοποιούν οι αλγόριθμοι της κατηγορίας αυτής και στις βασικές διαφορές μεταξύ τους. Θίχτηκαν οι συνηθέστερες επιπλοκές που συναντώνται κατά την επίλυση προβλημάτων βελτιστοποίησης και πώς αυτές μπορούν να αντιμετωπιστούν με αποτελεσματικότητα από έναν EA. Στη συνέχεια, παρατέθηκαν οι ιδιότητες της κανονικής κατανομής, η οποία αποτελεί το κύριο εργαλείο της υποκατηγορίας ES των EA, καθώς και η φυσική σημασία του μητρώου συνδιακύμανσης. Έχοντας, πλέον, τελειώσει με τις προαπαιτούμενες γνώσεις, αναλύθηκαν εξονυχιστικά οι αρχές λειτουργίας του CMA, στον οποίο προστέθηκαν ένας νέος τελεστής μετάλλαξης εμπνευσμένος από τους GA, νέοι τρόποι αρχικοποίησης και ελέγχου των παραμέτρων του, η δυνατότητα αντιμετώπισης προβλημάτων βελτιστοποίησης με περιορισμούς μέσω εκθετικά αυξανόμενων ποινών, καθώς και μεταμοντέλα RBFN που τελούν προ-αξιολόγηση χαμηλού κόστους. Ο εν λόγω τροποποιημένος CMA δοκιμάστηκε τελικά, με και χωρίς την υποβοήθηση μεταμοντέλων, σε 3 προβλήματα ψευδομηχανικής με περιορισμούς, καθώς και σε 1 πραγματική εφαρμογή υπολογιστικής ρευστοδυναμικής. Τα εν λόγω προβλήματα λύθηκαν παράλληλα και από τον γενικευμένο EA του λογισμικού EASY, ώστε να διαπιστωθεί αν είναι σκόπιμη η αφομοίωση στοιχείων του νέου CMA.

Κατά την επίλυση των προβλημάτων ψευδομηχανικής, ο τροποποιημένος CMA κατάφερε να αναδειχθεί ως μια πολύ συνεπής μέθοδος βελτιστοποίησης. Η στατιστική ανάλυση που πραγματοποιήθηκε για 10 τρεξίματα του κάθε προβλήματος αποδεικνύει ότι ο αλγόριθμος ανιχνεύει λύσεις εξαιρετικής ποιότητας, απαιτώντας κατά κανόνα σημαντικά λιγότερες αξιολογήσεις από άλλες στοχαστικές μεθόδους της βιβλιογραφίας. Η χρήση μεταμοντέλων φαίνεται να δυσχεραίνει ελαφρώς τα στατιστικά των αποτελεσμάτων σε εύκολα προβλήματα, αλλά από πρακτικής άποψης ο βαθμός αυτός κρίνεται αμελητέος. Αντιθέτως, τα μεταμοντέλα διαδραματίζουν έναν καθοριστικό ρόλο στο απαιτητικότερο πρόβλημα του μειωτήρα στροφών, όπου η προσαρμογή του μητρώου συνδιακύμανσης

δυσχεραίνεται από τους αυστηρούς περιορισμούς. Το ίδιο φαινόμενο παρατηρείται και στο πρόβλημα CFD, όπου οι διαστάσεις του προβλήματος είναι σημαντικά αυξημένες και ο απλός CMA έχει αισθητά μειωμένη απόδοση. Λαμβάνοντας υπόψη όλα τα παραπάνω, συμπεραίνεται με ασφάλεια ότι τα μεταμοντέλα είναι ικανά να ενισχύσουν περαιτέρω την ήδη παρατηρούμενη αποτελεσματικότητα του αλγόριθμου CMA.

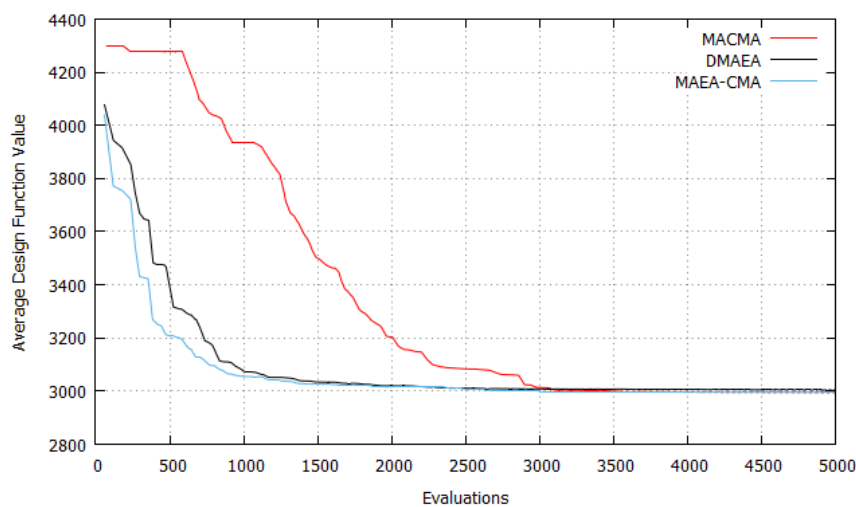
Το μόνο σημείο όπου ο CMA φαίνεται να υστερεί συγκριτικά με το γενικευμένο EA του EASY είναι κατά τις πρώτες γενιές της βελτιστοποίησης, στο ονομαζόμενο στάδιο της εξερεύνησης. Στα απλούστερα προβλήματα ο CMA μπορεί να ανταγωνιστεί τον DMAEA καθ' όλη τη διάρκεια της βελτιστοποίησης, έχοντας εντοπίσει λύση ίσης ή καλύτερης ποιότητας σε οποιοδήποτε στάδιο. Σε δυσκολότερες όμως εφαρμογές, όπου ο αριθμός των μεταβλητών σχεδιασμού ή και των περιορισμών είναι υψηλός, ο DMAEA είναι σημαντικά καλύτερος από τον CMA στην εύρεση καλών λύσεων με μικρό αριθμό απαιτούμενων αξιολογήσεων. Αυτό γίνεται διότι ο CMA απαιτεί έναν αυξημένο αριθμό δειγμάτων για να ανταπεξέλθει στις πρώτες γενιές, πράγμα που τον κάνει εκ των πραγμάτων λιγότερο αποτελεσματικό. Το θέμα λύνεται μερικώς με τη χρήση μεταμοντέλων, αλλά σε καμία περίπτωση δεν απαλείφεται. Τα μονοπάτια εξέλιξης και η προσαρμογή του μητρώου συνδιακύμανσης είναι τα ισχυρότερα προσόντα του CMA, τα οποία του προσδίδουν έναν αιτιοκρατικό χαρακτήρα όπως προαναφέρθηκε, αλλά χρειάζονται αρκετές γενιές ώστε να μπορούν να αξιοποιηθούν στο έπακρο. Όταν υπάρχουν αρκετά δεδομένα για τη σωστή ρύθμισή τους, ο CMA φτάνει και σε βάθος χρόνου ξεπερνά τον DMAEA λόγω της εκμετάλλευσης ήδη υπάρχουσών πληροφοριών από τις προηγούμενες γενιές.

Εν κατακλείδι, ο τροποποιημένος αλγόριθμος CMA με μεταμοντέλα κρίνεται ως μια ισχυρή μέθοδος που διακρίνεται στην αξιοποίηση πληροφοριών με σκοπό τη βελτίωση των χαρακτηριστικών ενός πληθυσμού σε βάθος χρόνου, παρουσιάζοντας κάποιες αρχικές δυσκολίες όταν το χωρίο σχεδιασμού είναι είτε εξαιρετικά πολυδιάστατο, είτε παρουσιάζει ασυνέχειες σε μεγάλο βαθμό. Ο DMAEA έχει αντιθέτως πολύ καλή αρχική σύγκλιση ανεξαρτήτων συνθηκών, η οποία τείνει να μειώνεται απότομα καθώς η λύση οδεύει προς το βέλτιστο. Ως λογικό επακόλουθο προκύπτει ότι το λογισμικό EASY μπορεί να ευνοηθεί από την προσαρμογή του μητρώου συνδιακύμανσης στα τέλη της βελτιστοποίησης, χωρίς απαραίτητα να την εφαρμόζει εξ αρχής.

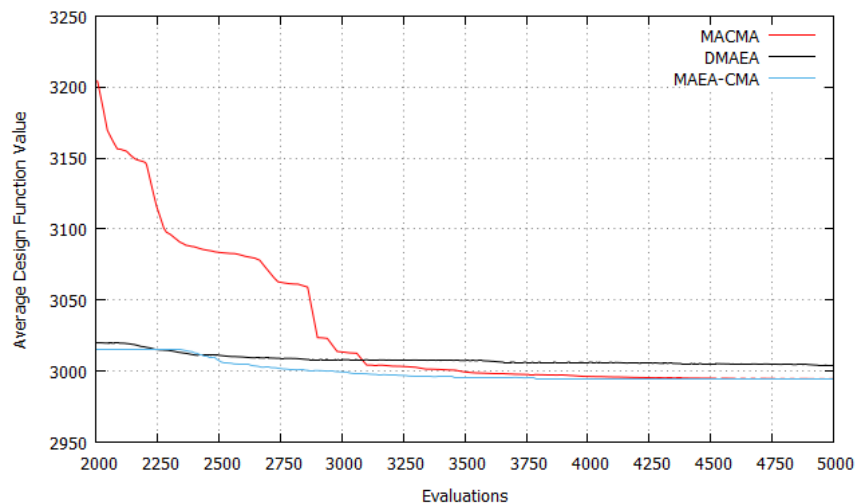
Ο απλούστερος τρόπος ένταξης του CMA στο EASY είναι χρησιμοποιώντας τον κατά τα τελικά στάδια της βελτιστοποίησης. Όταν ο γενικευμένος EA φτάσει σε ένα

σημείο όπου αρχίζει να συγκλίνει, το οποίο μπορεί να ελέγχεται π.χ. με τον αριθμό στάσιμων γενιών ή των μέγιστων επιτρεπτών αξιολογήσεων, μπορεί να ενεργοποιείται αυτόματα η λειτουργία του τροποποιημένου CMA. Ο αρχικός μέσος όρος μπορεί να προκύψει με πολλούς τρόπους, ακόμη και να τεθεί απλά ίσος με τον τρέχοντα καλύτερο απόγονο, ενώ για το μέγεθος του αρχικού βήματος απαιτείται μια έξυπνη, πρώτη πρόβλεψη ανάλογα με τις διαθέσιμες πληροφορίες. Το αρνητικό είναι ότι τα μονοπάτια εξέλιξης και το μητρώο συνδιακύμανσης πρέπει να δημιουργηθούν από το μηδέν, αλλά θεωρείται ότι ο CMA μπορεί να ξεπεράσει την απόδοση του EASY σε περίπτωση που το ολικό ακρότατο βρίσκεται στην εγγύς περιοχή της μέσης τιμής και η δειγματοληψία αρχικοποιηθεί αποτελεσματικά.

Ο υβριδικός MAEA-CMA που προτείνεται δοκιμάζεται στο πρόβλημα σχεδιασμού του μειωτήρα στροφών που εξετάστηκε στην Ενότητα 5.3, όπου αρχικά χρησιμοποιείται ο DMAEA για 2000 αξιολογήσεις και μετά ενεργοποιείται ο αρχικοποιημένος MACMA για άλλες 3000. Οι ρυθμίσεις των μεμονωμένων αλγορίθμων παραμένουν αναλλοίωτες, όπως δίνονται στον Πίνακα 5.9. Γίνονται 5 ανεξάρτητα τρεξίματά του MAEA-CMA και το μέσο διάγραμμα σύγκλισης που προκύπτει συγκρίνεται με τα αντίστοιχα των απλών MACMA και DMAEA στα Σχήματα 7.1 και 7.2. Στο Σχήμα 7.1 διακρίνεται ότι, στην αρχή της βελτιστοποίησης, τυχαίνει ο DMAEA του MAEA-CMA να έχει καλύτερη επίδοση με τις νέες αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών, αλλά από τις 1000 έως και τις 2000 αξιολογήσεις δεν υπάρχει σημαντική διαφορά μεταξύ των δύο αλγορίθμων.



**Σχήμα 7.1:** Μέσο διάγραμμα σύγκλισης των MACMA, DMAEA και του υβριδικού MAEA-CMA για το πρόβλημα σχεδιασμού του μειωτήρα στροφών με 5000 επιτρεπόμενες αξιολογήσεις.



**Σχήμα 7.2:** Μέσο διάγραμμα σύγκλισης των MACMA, DMAEA και του υβριδικού MAEA-CMA για το πρόβλημα σχεδιασμού του μειωτήρα στροφών από 2000 ως 5000 αξιολογήσεις.

Στο Σχήμα 7.2, όπου έχει ενεργοποιηθεί πλέον ο MACMA στον υβριδικό αλγόριθμο, φαίνεται πως επιτυγχάνεται σημαντικά μεγαλύτερη μείωση της τιμής της συνάρτησης-στόχου έως το τέλος της βελτιστοποίησης, συγκριτικά με τον απλό DMAEA. Το αρχικό προβάδισμα του MAEA-CMA έναντι του απλού MACMA φθίνει έως και το τέλος της βελτιστοποίησης, όπου ο νέος αλγόριθμος τελικά καταφέρνει να εντοπίσει την καλύτερη μέση τιμή σχεδιασμού με διαφορά στο 1<sup>ο</sup> δεκαδικό ψηφίο, έχοντας παράλληλα απαλείψει τις δυσμενείς περιοχές κάθε αλγορίθμου. Αναφέρεται ότι ο MAEA-CMA έχει βρει ως μέση τελική λύση 2994.4722 με 5000 αξιολογήσεις, ενώ ο απλός MACMA για τον ίδιο αριθμό αξιολογήσεων έχει μέση τιμή συνάρτησης-στόχου 2994.6020, την οποία βελτιώνει σε 2994.4717 με 2000 ακόμη αξιολογήσεις. Επομένως, ο νέος αλγόριθμος βρήκε παρόμοια ποιότητα τελικής λύσης με διαφορά μόλις στο 3<sup>ο</sup> δεκαδικό, ενώ απαιτεί σημαντικά λιγότερες αξιολογήσεις.

Κλείνοντας, αναφέρεται ότι η ένταξη μεμονωμένων τμημάτων του CMA στο EASY αποτελεί μία μεγαλύτερη πρόκληση. Το θέμα που εντοπίζεται στη συμβατότητα των δύο αλγορίθμων είναι ότι το λογισμικό EASY δεν χρησιμοποιεί τους κλασικούς τελεστές ανασυνδυασμού σε πραγματική κωδικοποίηση, δηλαδή δεν υπολογίζει έναν απόγονο τον οποίο τελικά μεταλλάσσει. Δημιουργείται έτσι ένα αδιέξοδο, διότι το μητρώο συνδιακύμανσης βοηθά στη συλλογή όσο το δυνατόν καλύτερων δειγμάτων γύρω από μια μέση τιμή, τον «μέσο απόγονο». Το μητρώο συνδιακύμανσης πρέπει επομένως να αξιοποιηθεί με διαφορετικό τρόπο, ενδεχομένως για να καταγράφει τα αποτελέσματα των

επιτυχημένων μεταλλάξεων που ασκούνται στους ήδη γεννημένους απογόνους και να τις βελτιστοποιεί. Όσο για το EASY με δυαδική κωδικοποίηση, το μητρώο συνδιακύμανσης μπορεί να χρησιμοποιηθεί ως ένας διαφορετικός τρόπος γένεσης απογόνων, χωρίς ή σε συνεργασία με τους κλασικούς τελεστές διασταύρωσης, αφού μπορεί να εμπεριέχει τις συνδιακυμάνσεις των bit και να δημιουργεί στατιστικά καλούς απογόνους βάσει των υπάρχουσών πληροφοριών.



## ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Rao S S (2009). *Engineering Optimization: Theory and Practice, Fourth Edition*. John Wiley & Sons, Inc., Hoboken, New Jersey. United States of America.
- [2] Γιαννάκογλου Κ Χ (2012). *Μέθοδοι Αιτιοκρατικής και Στοχαστικής Βελτιστοποίησης και Εφαρμογές*. Σημειώσεις για το Μεταπτυχιακό Μάθημα (ΔΠΜΣ). Μονάδα Παράλληλης Υπολογιστικής Μηχανικής & Βελτιστοποίησης. Τομέας Ρευστών. Σχολή Μηχανολόγων Μηχανικών. Εθνικό Μετσόβιο Πολυτεχνείο. Αθήνα.
- [3] Van Veldhuizen D A, Lamont G B (2000). *Multiobjective Evolutionary Algorithms: Analyzing the State-of-the-Art*. In: *Evolutionary Computation* 8(2), pp. 125 – 147. Massachusetts Institute of Technology.
- [4] Miettinen K M (1999). *Nonlinear Multiobjective Optimization*. University of Jyväskylä. Kluwer Academic Publishers. United States of America.
- [5] Hansen N (2011). *The CMA Evolution Strategy: A Tutorial*. DOI:  
<http://www.cmap.polytechnique.fr/~nikolaus.hansen/cmatutorial110628.pdf>
- [6] De Jong K, Fogel D B, Schwefel H-P (1997). *A History of Evolutionary Computation*. In: *Handbook of Evolutionary Computation*. IOP Publishing LTD and Oxford University Press.
- [7] Hansen N, Ostermeier A (1996). *Adapting Arbitrary Normal Mutation Distributions in Evolution Strategies: The Covariance Matrix Adaptation*. In: *Proceedings of the 1996 IEEE Intern. Conf. on Evolutionary Computation (ICEC '96)*, pp. 312 – 317.
- [8] *EASY v2.0 User's Manual*. Parallel CFD & Optimisation Unit. School of Mechanical Engineering. National Technical University of Athens. 2008. DOI:  
<http://velos0.ltt.mech.ntua.gr/EASY/>
- [9] Hansen N, Arnold D V, Auger A (2015). *Evolution Strategies*. In: Janusz Kacprzyk, Witold Pedrycz. *Handbook of Computational Intelligence*, Springer. Berlin. Germany.
- [10] Goldberg D E (1989). *Genetic Algorithms in Search, Optimization & Machine Learning*. University of Alabama. Addison-Wesley Publishing Company, Inc. United States of America.

- [11] Beyer H-G, Schwefel H-P (2002). *Evolution Strategies. A Comprehensive Introduction*. In: Natural Computing 1, pp. 3 – 52. Kluwer Academic Publishers. Netherlands.
- [12] Umbarkar A J, Sheth P D (2015). *Crossover Operators in Genetic Algorithms. A review*. In: Ictact Journal on Soft Computing, Volume: 06, Issue: 01, pp. 1083 – 1092.
- [13] Hansen N, Ostermeier A (2001). *Completely Derandomized Self-Adaptation in Evolution Strategies*. In: Evolutionary Computation 9(2), pp. 159 – 195.
- [14] Zitzler E, Thiele L (1999). *Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach*. In: IEEE Transactions on Evolutionary Computation 3(4), pp. 257 – 271.
- [15] Zitzler E, Laumanns M, Thiele L (2001). *SPEA 2: Improving the Strength Pareto Evolutionary Algorithm*. Computer Engineering and Networks Laboratory. Department of Electrical Engineering. Swiss Federal Institute of Technology. Zurich.
- [16] Giannakoglou K C, Asouti V G, Kapsoulis D (2020). *Low-Cost Evolutionary Algorithms for Engineering Applications*. Parallel CFD & Optimisation Unit. Laboratory of Thermal Turbomachines. School of Mechanical Engineering. National Technical University of Athens.
- [17] Kampolis I C, Zymaris A S, Asouti V G, Giannakoglou K C (2007). *Multilevel Optimization Strategies Based on Metamodel-Assisted Evolutionary Algorithms, for Computationally Expensive Problems*. In: 2007 IEEE, Proceedings of the Congress on Evolutionary Computation (CEC). Singapore.
- [18] Giannakoglou K C, Karakasis M K (2006). *Evolutionary Algorithms with Surrogate Modeling for Computationally Expensive Optimization Problems*. In: ERCOFTAC 2006, Proceedings of the Design Optimization International Conference. Spain.
- [19] Giotis A P, Emmerich M, Naujoks B, Giannakoglou K C, Bäck T (2001). *Low-Cost Stochastic Optimization for Engineering Applications*. In: EUROGEN 2001, Proceedings of the International Conference on Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems. Athens. Greece.
- [20] Karakasis M K, Giotis A P, Giannakoglou K C (2001). *Efficient Genetic Optimization Using Inexact Information and Sensitivity Analysis. Application in Shape*

- Optimization Problems*. In: ECCOMAS, Proceedings of the Computational Fluid Dynamics Conference 2001. Swansea. Wales. United Kingdom.
- [21] Kampolis I C, Papadimitriou D I, Giannakoglou K C (2004). *Evolutionary Optimization Using a New Radial Basis Function Network and the Adjoint Formulation*. In: Proceedings of the Inverse Problems, Design and Optimization (IPDO) Symposium. Rio de Janeiro. Brazil.
- [22] Karakasis M K, Giannakoglou K C (2005). *Metamodel-Assisted Multi-Objective Evolutionary Optimization*. In: EUROGEN 2005, Evolutionary and Deterministic Methods for Design, Optimization and Control with Applications to Industrial and Societal Problems. Munich.
- [23] Hansen N (2008). *Tutorial – The Covariance Matrix Adaptation Evolution Strategy (CMA-ES)*. Lorentz Center. Leiden.
- [24] Στείρος Σ Κ (2010). *Θεωρία Μετρήσεων και Σφαλμάτων*. Εκδόσεις Συμμετρία. Αθήνα.
- [25] Wang B, Shi W, Miao Z (2015). *Confidence Analysis of Standard Deviation Ellipse and Its Extension into Higher Dimensional Euclidean Space*. PLoS ONE 10(3): e0118537. DOI:10.1371/journal.pone.0118537.
- [26] Box G E P and Muller M E (1958). *A Note on the Generation of Random Normal Deviates*. In: The Annals of Mathematical Statistics 29(2), pp. 610 – 611.
- [27] Παπαδοπούλου Μ (2020). *Μέθοδος Προσαρμογής του Μητρώου Συνδιακύμανσης (CMA) στην Εξελικτική Βελτιστοποίηση*. Διπλωματική Εργασία.. Μονάδα Παράλληλης Υπολογιστικής Μηχανικής & Βελτιστοποίησης. Τομέας Ρευστών. Σχολή Μηχανολόγων Μηχανικών. Εθνικό Μετσόβιο Πολυτεχνείο. Αθήνα.
- [28] Ostermeier A, Gawelczyk A, Hansen N (1994). *Step-size Adaptation Based on Non-Local Use of Selection Information*. In: Proceedings of PPSN IV: Parallel Problem Solving from Nature. pp. 189 – 198. Springer. Berlin. Germany.
- [29] Hansen N (2006). *The CMA Evolution Strategy: A Comparing Review*. In: Lozano J A, Larranaga P, Inza I, Bengoetxea E. Towards a New Evolutionary Computation. Studies in Fuzziness and Soft Computing, vol 192. Springer, Berlin. pp. 75 – 102.
- [30] Yildirim A E, Karci A (2018). *Application of Three Bar Truss Problem among Engineering Design Optimization Problems using Artificial Atom Algorithm*. DOI: 10.1109/IDAP.2018.8620762.

- [31] Ray T, Liew K M (2003). *Society and Civilization: An Optimization Algorithm Based on the Simulation of Social Behavior*. In: IEEE Transactions of Evolutionary Computation 7(4), pp. 386 – 396.
- [32] Sadollah A, Bahreininejad A, Eskandar H, Hamdi M (2012). *Mine Blast Algorithm: A New Population Based Algorithm for Solving Constrained Engineering Optimization Problems*. In: Applied Soft Computing 13(2013), pp. 2592 – 2612.
- [33] Kunjie Y, Xin W, Zhenlei W (2014). *An Improved Teaching-Learning-Based Optimization Algorithm for Numerical and Engineering Optimization Problems*. In: Journal of Intelligent Manufacturing. Springer Science + Business Media. New York.
- [34] Zengqiang M, Yikun X, Yang Y, Tong Z, Bo Z, Liqing L (2014). *Hybrid Biogeography Based Optimization for Constrained Numerical and Engineering Optimization*. In: Mathematical Problems in Engineering, Volume 2015, DOI: <https://doi.org/10.1155/2015/423642>.
- [35] Hao Z, Yumei H, Weidong Z (2019). *A Dynamic Adaptive Particle Swarm Optimization and Genetic Algorithm for Different Constrained Engineering Design Optimization Problems*. In: Advances in Mechanical Engineering 11(3), pp. 1 – 27.
- [36] Garg H (2009). *Solving Structural Engineering Design Optimization Problems Using an Artificial Bee Colony Algorithm*. In: Journal of Industrial and Management Optimization 10(3), pp. 777 – 794.
- [37] Vivek K M, Bhaskar D (2012). *A Constrained Optimization Algorithm Based on the Simplex Search Method*. In: Engineering Optimization 44(5), pp. 537 – 550.
- [38] David C N D, Stephen S E A (2018). *Cost Minimization of Welded Beam Design Problem Using Nontraditional Optimization through MATLAB and Validation through ANSYS Simulation*. In: International Journal of Mechanical Engineering and Technology (IJMET) 9(8), pp. 180 – 192.
- [39] Stephen S E A, David C N D, Joe A A (2016). *Weight Minimization of Speed Reducer Design Problem using PSO, SA, PS, GODLIKE, CUCKOO, FF, FP, ALO, GSA and MVO*. In: International Journal of Scientific & Engineering Research 7(7), pp. 1356 – 1365.

- [40] Δημητρακόπουλος Α Κ (2017). *Στοιχεία Υδραυλικής Κλειστών και Ανοιχτών Αγωγών*. Πανεπιστημιακές Σημειώσεις. Τμήμα Πολιτικών Μηχανικών. Πανεπιστήμιο Πατρών. Πάτρα.
- [41] Giannakoglou K C, Papoutsis-Kiachagias E M, Gkaragkounis K T (2020). *adjointOptimisationFoam, an OpenFOAM-based Optimization Tool*. Parallel CFD & Optimisation Unit. School of Mechanical Engineering. National Technical University of Athens.