



**Εθνικό Μετσόβιο Πολυτεχνείο**  
**Σχολή Μηχανολόγων Μηχανικών**  
**Τομέας Ρευστών**  
**Εργαστήριο Θερμικών Στροβιλομηχανών**  
**Μονάδα Παράλληλης Υπολογιστικής Ρευστοδυναμικής**  
**& Βελτιστοποίησης**

**Συμβολή στην Επιτάχυνση Σύγκλισης των**  
**Εξελικτικών Αλγορίθμων μέσω Μεταπροτύπων και**  
**Ανάλυσης σε Κύριες Συνιστώσες. Εφαρμογές**  
**στην Αεροδυναμική.**

Διπλωματική Εργασία

Κουκούλης Ιωάννης

Επιβλέπων: Κυριάκος Χ. Γιαννάκογλου, Καθηγητής ΕΜΠ

Αθήνα, Ιούλιος 2017



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΤΟΜΕΑΣ ΡΕΥΣΤΩΝ  
ΕΡΓΑΣΤΗΡΙΟ ΘΕΡΜΙΚΩΝ ΣΤΡΟΒΙΛΟΜΗΧΑΝΩΝ  
ΜΟΝΑΔΑ ΠΑΡΑΛΛΗΛΗΣ ΥΠΟΛΟΓΙΣΤΙΚΗΣ  
ΡΕΥΣΤΟΔΥΝΑΜΙΚΗΣ & ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

**Συμβολή στην Επιτάχυνση Σύγκλισης των  
Εξελικτικών Αλγορίθμων μέσω Μεταπροτύπων και  
Ανάλυσης σε Κύριες Συνιστώσες. Εφαρμογές  
στην Αεροδυναμική.**

**Ιωάννης Κουκούλης**

**Επιβλέπων: Κυριάκος Χ. Γιαννάκογλου, Καθηγητής ΕΜΠ**

**Αθήνα, Ιούλιος 2017**

**Περίληψη**

Στόχος της διπλωματικής εργασίας είναι η παρουσίαση, ανάλυση και πιστοποίηση μεθόδων βελτιστοποίησης, οι οποίες βασίζονται στους Εξελικτικούς αλγορίθμους (ΕΑ), και αξιοποιούν Ανάλυση σε Κύριες Συνιστώσες (Principal Component Analysis), για να αυξήσουν την αποτελεσματικότητά τους.

Η βελτιστοποίηση με ΕΑ ερευνάται τα τελευταία χρόνια από τη Μονάδα Παράλληλης Υπολογιστικής Ρευστομηχανικής & Βελτιστοποίησης (ΜΠΥΡ&Β) του Εργαστηρίου Θερμικών Στροβιλομηχανών (ΕΘΣ) του ΕΜΠ. Στο πλαίσιο αυτής της έρευνας, διαμορφώθηκε και αναπτύσσεται το λογισμικό *EASY*, το οποίο παρέχει τη δυνατότητα επίλυσης σύνθετων προβλημάτων βελτιστοποίησης με χρήση ΕΑ. Βασικό πλεονέκτημα του *EASY* είναι η ευελιξία που το χαρακτηρίζει. Σε συνδυασμό με το απαραίτητο, κάθε φορά, πρόγραμμα αξιολόγησης, είναι δυνατή η εύρεση των βέλτιστων λύσεων για μεγάλο εύρος μηχανολογικών προβλημάτων.

Η βασική αδυναμία των ΕΑ βρίσκεται στον μεγάλο αριθμό αξιολογήσεων υποψήφιων λύσεων που απαιτείται για τον εντοπισμό των βέλτιστων λύσεων. Κάτι τέτοιο είναι ιδιαίτερα εμφανές σε προβλήματα μεγάλης κλίμακας της βιομηχανίας, όπου η αξιολόγηση μίας υποψήφιας λύσης έχει σημαντικό υπολογιστικό και χρονικό κόστος.

Την αδυναμία αυτή επιχειρούν να αντιμετωπίσουν οι Εξελικτικοί Αλγόριθμοι Υποβοηθούμενοι με Μεταπρότυπα (Metamodel-Assisted Evolutionary

Algorithms - MAEAs). Οι MAEA κατασκευάζουν προσεγγιστικά μοντέλα της αντικειμενικής συνάρτησης, τα οποία απαιτούν ελάχιστο υπολογιστικό κόστος, και επιτρέπουν τον καθορισμό των υποσχόμενων λύσεων ενός προβλήματος, χωρίς να καταφεύγουν σε κλήσεις του ακριβούς λογισμικού αξιολόγησης. Με αυτόν τον τρόπο, μπορεί να μειωθεί σημαντικά ο χρόνος που χρειάζεται για την επίλυση ενός προβλήματος βελτιστοποίησης.

Η χρήση μεταπροτύπων βελτιώνει την επίδοση των EA, δεν αναιρεί ωστόσο το βασικό τους μειονέκτημα. Για το λόγο αυτό, αναζητώνται διαρκώς νέες τεχνικές που θα επιτρέψουν στους EA να ελαχιστοποιήσουν όσο το δυνατόν περισσότερο τη χρήση του ακριβούς λογισμικού αξιολόγησης.

Στην εργασία αυτή, αυτό επιτυγχάνεται με χρήση μεθόδων Ανάλυσης σε Κύριες Συνιστώσες (ΑΣΚΣ). Οι στατιστικές αυτές μέθοδοι χρησιμοποιούνται για την ανάλυση μεγάλων συνόλων δεδομένων. Ουσιαστική λειτουργία τους είναι να ορίζουν ένα καινούριο σύνολο μεταβλητών, στατιστικά ασυσχέτιστων μεταξύ τους, επιτρέποντας έτσι τον εντοπισμό σημαντικών χαρακτηριστικών του σετ δεδομένων και την οικονομικότερη αναπαράστασή τους. Παρουσιάζονται και χρησιμοποιούνται δύο ξεχωριστές μέθοδοι ΑΣΚΣ: η γραμμική ΑΣΚΣ (linear PCA), στην οποία το νέο σύνολο μεταβλητών προκύπτει από γραμμικό μετασχηματισμό των αρχικών, και η ΑΣΚΣ με χρήση συναρτήσεων (μη-γραμμικού) πυρήνα (Kernel PCA). Η δεύτερη αποτελεί μια γενίκευση της πρώτης, η οποία εκμεταλλεύεται τις ιδιότητες ενός μη-γραμμικού μετασχηματισμού για να εντοπίσει πιο σύνθετα χαρακτηριστικά του σετ δεδομένων.

Κάθε μία από τις δύο μεθόδους ΑΣΚΣ, εφαρμόζεται στο σύνολο απογόνων κάθε γενιάς του EA, και παρέχει πληροφορία για τα χαρακτηριστικά του χώρου σχεδιασμού του προβλήματος. Η απόδοση των EA μπορεί να βελτιωθεί με δύο διαφορετικούς τρόπους για κάθε μέθοδο: με την εφαρμογή των τελεστών εξέλιξης σε γονείς ορισμένους ως προς το νέο σύνολο μεταβλητών που προκύπτει, ευνοώντας έτσι την παραγωγή ποιοτικότερων υποψήφιας λύσεων (εκδοχή EA(PCA)), και με αποκοπή διαστάσεων κατά την εκπαίδευση των μεταπροτύπων (εκδοχή M(PCA)AEA). Η εκπαίδευση μεταπροτύπων χαμηλότερης διάστασης, η οποία γίνεται αποκόπτοντας τις λιγότερο σημαντικές μεταβλητές, είναι φθηνότερη υπολογιστικά, ενώ μπορεί να αυξήσει και την ακρίβειά τους.

Οι τρόποι αξιοποίησης που περιγράφηκαν, μπορούν να εφαρμοστούν μαζί ή και ξεχωριστά σε έναν EA. Έτσι προκύπτουν πολλές διαφορετικές εκδοχές των συμβατικών EA και MAEA, που μπορούν να ονομαστούν ανάλογα με τη μέθοδο ΑΣΚΣ που χρησιμοποιείται, και τον τρόπο αξιοποίησής της:

EA(L), EA(K), M(L)AEA, M(K)AEA, M(L)AEA(L), M(K)AEA(K), όπου το γράμμα L (Linear) ή K (Kernel) προσδιορίζει τη μέθοδο ΑσκΣ, ενώ η θέση του στο ακρωνύμιο τη χρήση της.

Οι προτεινόμενες εκδοχές, οι οποίες υλοποιούνται από το λογισμικό EASY, χρησιμοποιούνται για την επίλυση προβλημάτων βελτιστοποίησης. Συγκεκριμένα, γίνεται μελέτη τριών μαθηματικών προβλημάτων, δύο προβλημάτων μηχανολογικού σχεδιασμού, και ενός προβλήματος αεροδυναμικής βελτιστοποίησης. Κάθε πρόβλημα επιλύεται με διάφορες από τις εκδοχές που αναφέρονται παραπάνω και, στη συνέχεια, γίνεται σύγκριση της αποτελεσματικότητάς τους με τους συμβατικούς EA και MAEA.



NATIONAL TECHNICAL UNIVERSITY  
OF ATHENS  
SCHOOL OF MECHANICAL ENGINEERING  
FLUIDS DEPARTMENT  
LABORATORY OF THERMAL TURBOMACHINES  
PARALLEL CFD & OPTIMIZATION UNIT

**Acceleration of Evolutionary Algorithms using Metamodels and  
Principal Component Analysis. Applications in Aerodynamics.**

**Koukoulis Ioannis**

Advisor: Kyriakos C. Giannakoglou, Professor NTUA

Athens, July 2017

Abstract

The aim of this diploma thesis is to present and validate optimization methods based on Evolutionary Algorithms (EAs), which utilize Principal Component Analysis (PCA) in order to increase their efficiency.

EA-based optimization has been adequately developed over the last years at the Parallel CFD & Optimization Unit (PCOpt) of the Laboratory of Thermal Turbomachines (LTT) of the NTUA. During the last decades, the EASY optimization software has been developed. EASY is an EA-based, flexible all-purpose optimization tool. Coupled with the appropriate problem-specific evaluation program, it can compute the optimal solution to a wide range of optimization problems.

The major weakness of EA's is the large number of candidate solutions that must be evaluated, in order for the evolution process to take place. This difficulty is accentuated in large-scale industrial optimization problems, where the evaluation of a single candidate solution is very expensive, in terms of both wall-clock time and computational cost. For this purpose, Metamodel-Assisted Evolutionary Algorithms (MAEAs) have been developed. The role of MAEAs is to construct approximate models of the objective function(s) to be optimized. These models, called metamodels or surrogate evaluation models, are computationally cheap, and allow for the detection of the most promising candidate solutions, without resorting to the expensive problem-specific evaluation program. The time and computational resources required for the optimization process can be greatly reduced in this way.

The use of metamodels significantly improves the performance of EAs, however their major setback remains; therefore, new techniques are being researched that will allow for an EA with minimal usage of the costly evaluation software.

In this diploma thesis, this is accomplished through the use of Principal Component Analysis (PCA), applied to an appropriate dataset. PCA is a statistical tool used for analyzing large sets of data. Its function is to define a new set of variables, which describe data in a statistically uncorrelated way. This can facilitate the separation between noise and signal and the detection of patterns in the dataset. Two separate PCA methods are presented and used in this diploma thesis: Linear PCA, in which the new set of variables is the result of an orthogonal transformation of the original variables, and Kernel PCA. The latter is a generalization of the former, utilizing properties of a non-linear transformation in order to detect a great variety of characteristics in the dataset, a technique known as “the kernel trick”.

Principal component analysis is performed on the set of offspring produced by the EA in each generation, providing information about the nature of the problem’s design space. The results of the analysis can be used to increase the EA’s performance in two different ways: by applying the evolutionary operators of mutation and crossover to parents expressed in terms of the new variables, thus producing candidate solutions of higher quality (a variant that is called EA(PCA) ), or by training metamodels of lesser dimensionality (a variant called M(PCA)AEA). This reduction in dimensionality produces metamodels of improved predictive ability that are also computationally cheaper to train.

The above mentioned methods can be used separately or in conjunction with each other. This leads to several variants of the conventional EAs and MAEAs, which can be named accordingly, depending on the chosen PCA method and its use: EA(L), EA(K), M(L)AEA, M(K)AEA, M(L)AEA(L), M(K)AEA(K), where the letter L or K is used for a linear or kernel method respectively.

The proposed methods are implemented using the EASY software and applied to optimization problems: specifically three mathematical problems, two cases of pseudo-engineering optimization exercises, and a case of aerodynamic optimization are studied. Each problem is solved using the variants mentioned above and their performances are compared to those of the conventional EA and MAEA.

## Ακρωνύμια

ΕΜΠ	Εθνικό Μετσόβιο Πολυτεχνείο
ΕΘΣ	Εργαστήριο Θερμικών Στροβιλομηχανών
ΜΠΥΡ&Β	Μονάδα Παράλληλης Υπολογιστικής Ρευστοδυναμικής & Βελτιστοποίησης
ΥΡΔ	Υπολογιστική Ρευστοδυναμική
ΑσΚΣ	Ανάλυση σε Κύριες Συνιστώσες
ΔΣΑΒ	Δίκτυα Συναρτήσεων Ακτινικής Βάσης
ΠΠΑ	Προσεγγιστική Προ-Αξιολόγηση

## Nomenclature

CFD	Computational Fluid Dynamics
NTUA	National Technical University of Athens
PCopt	Parallel CFD & Optimization Unit
EA	Evolutionary Algorithm
MAEA	Metamodel-Assisted Evolutionary Algorithm
PCA	Principal Component Analysis
EA(L)	EA with linear-PCA-Driven Evolution Operators
MAEA(L)	MAEA with linear-PCA-Driven Evolution Operators
M(L)AEA	MAEA with linear-PCA-Driven Metamodels
M(L)AEA(L)	MAEA(L) with linear-PCA-Driven Metamodels
EA(K)	EA with kernel-PCA-Driven Evolution Operators
MAEA(K)	MAEA with kernel-PCA-Driven Evolution Operators
M(K)AEA	MAEA with kernel-PCA-Driven Metamodels
M(K)AEA(K)	MAEA(K) with kernel-PCA-Driven Metamodels



## Ευχαριστίες

Με την ολοκλήρωση αυτής της διπλωματικής εργασίας, φτάνει στο τέλος της η μεγάλη περιπέτεια των σπουδών μου στη σχολή Μηχανολόγων Μηχανικών του ΕΜΠ.

Στο σημείο αυτό θα ήθελα να ευχαριστήσω θερμά τον καθηγητή κ. Κυριάκο Γιαννάκογλου, για τη δυνατότητα που μου έδωσε να ασχοληθώ με ένα σύγχρονο και ενδιαφέρον επιστημονικό θέμα, για την ουσιαστική βοήθεια και καθοδήγησή που μου παρείχε, αλλά και την υπομονή που επέδειξε καθόλη τη διάρκεια εκπόνησης της εργασίας.

Επίσης, θα ήθελα να ευχαριστήσω ιδιαίτερα τον υποψήφιο διδάκτορα Δημήτρη Καψούλη για τις γνώσεις που μου μετέδωσε και τη συνεχή και άμεση υποστήριξη του σε όλα τα στάδια της εργασίας. Επιπλέον, θα ήθελα να ευχαριστήσω τη διδάκτορα Βαρβάρα Ασούτη, για τις πολύ χρήσιμες συμβουλές που μου προσέφερε.

Ευχαριστώ ακόμη, τους φίλους μου αλλά και όλα τα εξαίσια άτομα που είχα την τύχη να γνωρίσω κατά τη διάρκεια των σπουδών μου στο ΕΜΠ.

Τέλος, ένα τεράστιο ευχαριστώ οφείλω στους αγαπημένους μου γονείς, για όσα μου έχουν προσφέρει όλα αυτά τα χρόνια.

# Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>1</b>
1.1	Πρόλογος . . . . .	1
1.2	Σκοπός της Διπλωματικής Εργασίας . . . . .	2
1.3	Δομή της Εργασίας . . . . .	3
<b>2</b>	<b>Βελτιστοποίηση Με Χρήση Εξελικτικών Αλγορίθμων</b>	<b>5</b>
2.1	Γενικά για τη Βελτιστοποίηση . . . . .	5
2.2	Εισαγωγή στους Εξελικτικούς Αλγορίθμους . . . . .	8
2.3	Βασική Δομή Ενός Εξελικτικού Αλγορίθμου . . . . .	8
2.4	ΕΑ Υποβοηθούμενοι από Μεταπρότυπα . . . . .	12
2.4.1	Γενική Περιγραφή Λειτουργίας των ΜΑΕΑ . . . . .	13
2.4.2	Τεχνητά Νευρωνικά Δίκτυα . . . . .	14
2.4.3	Πολυωνυμική Προσέγγιση . . . . .	18
2.5	Λογισμικό EASY . . . . .	18
<b>3</b>	<b>Ανάλυση σε Κύριες Συνιστώσες και Εξελικτικοί Αλ- γόριθμοι</b>	<b>20</b>
3.1	“Κακώς Τοποθετημένα” Προβλήματα και ΕΑ . . . . .	20
3.2	Μέθοδος ΑσκΣ και Εξελικτικοί Αλγόριθμοι . . . . .	23
3.2.1	Η εκδοχή ΕΑ(L) . . . . .	25
3.2.2	Η εκδοχή Μ(L)ΑΕΑ . . . . .	26
3.3	Η Μέθοδος Kernel PCA . . . . .	26
3.3.1	Η εκδοχή ΕΑ(K) . . . . .	36
3.3.2	Η εκδοχή Μ(K)ΑΕΑ . . . . .	38
<b>4</b>	<b>Επίλυση Προβλημάτων Βελτιστοποίησης</b>	<b>39</b>
4.1	Εισαγωγή . . . . .	39
4.2	Συνάρτηση Rastrigin & Παραλλαγές της . . . . .	40

4.3	Συνάρτηση Ackley & Παραλλαγές της . . . . .	58
4.4	Συνάρτηση Fonseca - Fleming . . . . .	73
4.5	Πρόβλημα Σχεδιασμού Υποστηρίγματος Δύο Ράβδων . . . . .	78
4.6	Πρόβλημα Συγκόλλησης Δοκού . . . . .	84
4.7	Πρόβλημα Βελτιστοποίησης Αεροτομής Τεσσάρων Στοιχείων	89
<b>5</b>	<b>Ανακεφαλαίωση - Συμπεράσματα</b>	<b>96</b>
<b>A</b>	<b>Προβλήματα Πολλών Στόχων</b>	<b>99</b>
A.1	Σχέσεις Κυριαρχίας Μετώπων . . . . .	99
A.2	Μετρικές Ποιότητας-Δείκτης Υπερόγκου . . . . .	101
	<b>Βιβλιογραφία</b>	<b>104</b>

# Κεφάλαιο 1

## Εισαγωγή

### 1.1 Πρόλογος

Οι εξελικτικοί αλγόριθμοι (Evolutionary Algorithms EA) αποτελούν διαδεδομένη μέθοδο βελτιστοποίησης λόγω της απλότητας και ευελιξίας που τους χαρακτηρίζει. Βασικό στοιχείο λειτουργίας τους είναι η οργανωμένη παραγωγή και δοκιμή πληθυσμών λύσεων. Με αυτόν τον τρόπο αναδεικνύονται διαρκώς ολοένα και καλύτερες λύσεις του προβλήματος, οδεύοντας έτσι προς το βέλτιστο.

Οι EA είναι εξαιρετικά ευπροσάρμοστοι σε οποιοδήποτε πρόβλημα, αφού το μόνο που απαιτούν είναι η ύπαρξη ενός λογισμικού αξιολόγησης που θα τους επιτρέψει να πραγματοποιήσουν αυτή τη διαδικασία εξέλιξης προς τις βέλτιστες λύσεις. Ταυτόχρονα, εκεί βρίσκεται και η αδυναμία τους: τα λογισμικά αξιολόγησης λύσεων στα προβλήματα βελτιστοποίησης μεγάλης κλίμακας, (λ.χ. στα προβλήματα υπολογιστικής ρευστοδυναμικής - ΥΡΔ ή CFD) απαιτούν μεγάλο κόστος, τόσο υπολογιστικό, όσο και χρονικό. Συνεπώς η δοκιμή πολλών λύσεων σε ένα τέτοιο πρόβλημα κρίνεται ασύμφορη.

Για το λόγο αυτό, αναπτύσσονται διαρκώς νέες τεχνικές που τροποποιούν τους EA και επιχειρούν να αντισταθμίσουν αυτό το μειονέκτημα. Τέτοιες είναι η δυνατότητα παράλληλης αξιολόγησης των λύσεων ενός πληθυσμού[10], ο συνδυασμός (υβριδισμός) τους με αιτιοκρατικές μεθόδους βελτιστοποίησης [23] καθώς και η χρήση μεταπρότυπων[2]. Τα μεταπρότυπα αποτελούν μεθόδους προσεγγιστικού υπολογισμού των τιμών των αντικειμενικών συναρτήσεων ενός προβλήματος, οι οποίες επιτρέπουν τον καθορισμό των περισσότερων υποσχόμενων λύσεων, χωρίς να καταφεύγουν στο ακριβό και χρονοβόρο

πραγματικό λογισμικό αξιολόγησης.

Η χρήση εξελικτικών αλγορίθμων υποβοηθούμενων από μεταπρότυπα (Meta-model Assisted Evolutionary Algorithms - MAEA) έχει επεκτείνει τη χρήση τους σε πολύ μεγαλύτερο εύρος προβλημάτων βελτιστοποίησης. Ωστόσο, το βασικό μειονέκτημά τους παραμένει, και οποιαδήποτε απόπειρα περαιτέρω βελτίωσής τους πρέπει να στοχεύει στην επιπλέον εξοικονόμηση κλήσεων του ακριβούς και, συνήθως ακριβού, λογισμικού αξιολόγησης.

## 1.2 Σκοπός της Διπλωματικής Εργασίας

Σκοπός της διπλωματικής εργασίας είναι η παρουσίαση και πιστοποίηση εκδοχών των EA και MAEA, οι οποίες χρησιμοποιούν μεθόδους Ανάλυσης σε Κύριες Συνιστώσες (ΑΣΚΣ - Principal Component Analysis - PCA).

Η ΑΣΚΣ είναι μια μέθοδος στατιστικής ανάλυσης, η οποία εφαρμόζεται σε κάποιο χαρακτηριστικό σύνολο του εξελικτικού αλγορίθμου, εντοπίζει γραμμικά διαχωρίσιμα χαρακτηριστικά του χώρου σχεδιασμού και τα διατάσσει ως προς τη σημαντικότητά τους. Με αυτόν τον τρόπο πραγματοποιείται ένας επανακαθορισμός του συνόλου των μεταβλητών σχεδιασμού, ώστε να έχει ιδιότητες που βοηθούν στην ταχύτερη επίλυση του προβλήματος.

Τα αποτελέσματα της ανάλυσης μπορούν να αξιοποιηθούν με δύο τρόπους από τον εξελικτικό αλγόριθμο: με εφαρμογή των τελεστών εξέλιξης σε γονείς και απογόνους εκπεφρασμένους ως προς τις νέες μεταβλητές σχεδιασμού, αυξάνοντας έτσι την πιθανότητα παραγωγής ποιοτικότερων λύσεων, ή με την εκπαίδευση αξιόπιστων μεταπροτύπων χαμηλότερης διάστασης, η οποία συμβαίνει με αποκοπή των λιγότερο σημαντικών μεταβλητών, όπως αυτές υπολογίστηκαν.

Η ιδέα “ενίσχυσης” των EA και MAEA από τη μέθοδο ΑΣΚΣ έχει αναπτυχθεί και στο παρελθόν από τη Μονάδα Παράλληλης Υπολογιστικής Ρευστομηχανικής & Βελτιστοποίησης (ΜΠΥΡ&Β) του Εργαστηρίου Θερμικών Στροβιλομηχανών (ΕΘΣ) του ΕΜΠ [6, 7]. Στην εργασία αυτήν, η εφαρμογή της ανάλυσης για την εκπαίδευση των μεταπροτύπων γίνεται με διαφορετική λογική (αποκόπτονται οι μεταβλητές που αντιστοιχούν σε χαμηλότερη διακύμανση, θεωρώντας ότι παρέχουν λιγότερο χρήσιμη πληροφορία), ενώ επιπλέον προτείνεται και αξιολογείται και η χρήση Ανάλυσης σε Κύριες Συνιστώσες με συναρτήσεις πυρήνα (Kernel PCA). Πρόκειται ουσιαστικά για μια γενίκευση της απλής μεθόδου ΑΣΚΣ, με επιπλέον υπολογιστικά βήματα, που επι-

τρέπουν τον εντοπισμό και μη-γραμμικά διαχωρίσιμων χαρακτηριστικών του χώρου σχεδιασμού.

Οι δύο τρόποι αξιοποίησης που αναφέρθηκαν παραπάνω μπορούν να χρησιμοποιηθούν ξεχωριστά, αλλά και μαζί για την κάθε μέθοδο, δημιουργώντας έτσι διάφορες εκδοχές:  $M(L)AEA$ ,  $EA(L)$ ,  $MAEA(L)$ ,  $M(L)AEA(L)$ ,  $M(K)AEA$ ,  $EA(K)$ ,  $MAEA(K)$ ,  $M(K)AEA(K)$ , όπου το γράμμα  $L$  ή  $K$  αντιστοιχεί στη μέθοδο ΑσκΣ που πραγματοποιείται, ενώ η θέση του στην αρχή ή στο τέλος του ακρωνυμίου υποδηλώνει τη χρήση του στο μεταπρότυπο ή στους εξελικτικούς τελεστές, αντίστοιχα. Όλες αυτές οι διαφορετικές εκδοχές υλοποιούνται από το λογισμικό βελτιστοποίησης *EASY*.

Στη συνέχεια, πραγματοποιείται επίλυση προβλημάτων βελτιστοποίησης χρησιμοποιώντας τις εκδοχές αυτές, προκειμένου να διαπιστωθεί η χρησιμότητά τους και να εξαχθούν συμπεράσματα από τη σύγκρισή τους.

### 1.3 Δομή της Εργασίας

- Το κεφάλαιο 2 αποτελεί μια εισαγωγή στη βελτιστοποίηση με ΕΑ. Αρχικά δίνεται το μαθηματικό υπόβαθρο των προβλημάτων βελτιστοποίησης και αναφέρονται οι βασικές έννοιες της μονοκριτηριακής και πολυκριτηριακής βελτιστοποίησης. Επίσης γίνεται μια σύντομη αντιπαράθεση μεταξύ στοχαστικών και αιτιοκρατικών μεθόδων. Στη συνέχεια, εισάγεται η έννοια των ΕΑ, και περιγράφεται η δομή και λειτουργία του συμβατικού  $(\mu, \lambda)EA$ . Εξηγείται η σημασία και λειτουργία των μεταπρότυπων γενικά και, κατόπιν, αναλύονται τα τεχνητά νευρωνικά δίκτυα. Τέλος, παρουσιάζεται το λογισμικό *EASY*, το οποίο χρησιμοποιήθηκε ως βάση δόμησης των μεθόδων για την επίλυση των μελετούμενων προβλημάτων.
- Στο κεφάλαιο 3, ορίζεται η έννοια του “κακώς τοποθετημένου” προβλήματος και εξηγείται πώς η ιδιότητα αυτή, μαζί με το μεγάλο πλήθος διαστάσεων που χαρακτηρίζουν τα περισσότερα προβλήματα μηχανολογικού σχεδιασμού, επιβαρύνει σημαντικά την επίδοση των ΕΑ και ΜΑΕΑ. Αναλύεται η θεωρία της γραμμικής ανάλυσης σε κύριες συνιστώσες, και οι τρόποι με τους οποίους αυτή εφαρμόζεται στον ΕΑ, ώστε να βελτιώσει τη λειτουργία του στην αντιμετώπιση “κακώς τοποθετημένων” προβλημάτων υψηλής διάστασης.

Στη συνέχεια, αναπτύσσεται η θεωρία της ΑσΚΣ με χρήση συναρτήσεων πυρήνα. Σύμφωνα με αυτήν, το σύνολο των απογόνων απεικονίζεται μέσω μιας μη-γραμμικής συνάρτησης σε ένα χώρο πολύ μεγάλης (πρακτικά άπειρης) διάστασης, ώστε να διεξαχθεί εκεί μια “ισχυρότερη” γραμμική ΑσΚΣ.

Φυσικά, εάν οι υπολογισμοί πραγματοποιούνταν όντως σε έναν τέτοιο χώρο τεράστιας διάστασης θα επέφεραν και απαγορευτικό υπολογιστικό κόστος. Κάνοντας όμως χρήση των ιδιοτήτων των συναρτήσεων πυρήνα οι υπολογισμοί μεταφέρονται τελικά σε ένα χώρο διάστασης ίσης με το πλήθος του συνόλου απογόνων. Η διαδικασία αυτή, η οποία είναι γνωστή με τον όρο “τέχνασμα πυρήνα” (kernel trick), είναι εκείνη που καθιστά τελικά δυνατό τον προσδιορισμό σημαντικών μη-γραμμικών χαρακτηριστικών του χώρου σχεδιασμού. Τα αποτελέσματα αυτής της μεθόδου μπορούν να χρησιμοποιηθούν με τρόπο όμοιο αυτού της γραμμικής ΑσΚΣ για να αναβαθμίσουν τις επιδόσεις των EA & MAEA.

- Στο κεφάλαιο 4, γίνεται παρουσίαση έξι προβλημάτων βελτιστοποίησης: πρόκειται για τρία μαθηματικά προβλήματα, δύο απλοποιημένα προβλήματα μηχανολογικού σχεδιασμού και ένα πρόβλημα βελτιστοποίησης αεροτομής. Κάθε ένα από αυτά επιλύθηκε με χρήση του απλού EA ή και MAEA, καθώς και με χρήση των παραλλαγών που αναφέρονται στο προηγούμενο κεφάλαιο. Συνολικά δοκιμάστηκαν όλες οι παραλλαγές, και έμφαση δόθηκε σε αυτές που κάνουν εφαρμογή ΑσΚΣ στους εξελικτικούς τελεστές μόνο (EA(L), EA(K)), καθώς και στο συνδυασμό σε τελεστές εξέλιξης και μεταπρότυπα (M(L)AEA(L), M(K)AEA(K)). Τα αποτελέσματα από τις πορείες σύγκλισης συγκεντρώνονται σε διαγράμματα, στα οποία γίνονται παρατηρήσεις και σχολιασμός.
- Τέλος, στο κεφάλαιο 5, γίνεται ανακεφαλαίωση και εξαγωγή συμπερασμάτων για το σύνολο των αποτελεσμάτων.

## Κεφάλαιο 2

# Βελτιστοποίηση Με Χρήση Εξελικτικών Αλγορίθμων

### 2.1 Γενικά για τη Βελτιστοποίηση

Η βελτιστοποίηση είναι μια επιστήμη με ευρύ φάσμα εφαρμογών σε διάφορους τομείς όπως, η επιστημονική έρευνα, τα χρηματοπιστωτικά, η βιομηχανική παραγωγή και η οργάνωσή της. Όπως υποδηλώνει το όνομά της, η διαδικασία δεν στοχεύει απλώς στη λύση ενός προβλήματος, αλλά στην αναζήτηση και εύρεση των βέλτιστων λύσεων.

Από τη σκοπιά του μηχανικού, βελτιστοποίηση εφαρμόζεται κατά το σχεδιασμό μηχανολογικών προϊόντων, από τα πολύ απλά (λ.χ. ένας οδοντωτός τροχός ή μια μορφή αεροτομής), έως και τα πλέον σύνθετα (λ.χ. ένα κιβώτιο ταχυτήτων ή ένα αεροσκάφος). Συνήθως τα κριτήρια ως προς τα οποία διεξάγεται είναι η βελτίωση της ποιότητας (για παράδειγμα αύξηση ενός βαθμού απόδοσης ή του συντελεστή άνωσης), σε συνδυασμό με τη μεγαλύτερη δυνατή εξοικονόμηση κόστους. Ταυτόχρονα, οι προδιαγραφές αλλά και τα φυσικά μοντέλα που περιγράφουν τα μηχανολογικά προβλήματα θέτουν ένα σύνολο από περιορισμούς εντός των οποίων πρέπει να βρισχεται μια αποδεκτή (feasible) λύση. Μια βέλτιστη λύση λοιπόν οφείλει να ισορροπεί ανάμεσα σε όλες αυτές τις αντίρροπες τάσεις.



## Μαθηματική Διατύπωση Προβλημάτων Βελτιστοποίησης

Ένα πρόβλημα βελτιστοποίησης  $M_T$  στόχων εκφράζεται μαθηματικά στη γενική μορφή:

$$\min \vec{f}(\vec{x}) = \min [f_1(\vec{x}), \dots, f_{M_T}(\vec{x})] \quad (2.1)$$

$$\vec{c} \leq \vec{d} \quad (2.2)$$

όπου το διάνυσμα  $\vec{x}$  συγκεντρώνει τις  $N$  μεταβλητές σχεδιασμού, το διάνυσμα  $\vec{f}$  ονομάζεται διάνυσμα αντικειμενικών συναρτήσεων (objective functions) ή διάνυσμα συναρτήσεων στόχου, και αντιπροσωπεύει την ποιότητα της υποψήφιας λύσης, ενώ τα  $\vec{c}$  και  $\vec{d}$  εκφράζουν τους περιορισμούς. Κάθε συνιστώσα  $x_i$  του διανύσματος  $\vec{x}$  φράσσεται κάτω και πάνω από τους πραγματικούς αριθμούς  $x_i^{low}, x_i^{upp}$ , οι οποίοι ορίζουν το χώρο σχεδιασμού (design space).

Στα προβλήματα όπου στόχος είναι η μεγιστοποίηση κάποιας ποσότητας  $q$  (η οποία μπορεί να είναι λ.χ. ένας βαθμός απόδοσης), το πρόβλημα μπορεί να μετατραπεί σε πρόβλημα ελαχιστοποίησης θεωρώντας  $f = -q$ .

## Προβλήματα Ενός και Πολλών Στόχων

Σημαντική είναι η διάκριση ανάμεσα στα προβλήματα μονοκριτηριακής και πολυκριτηριακής βελτιστοποίησης (single-objective και multi-objective optimization - SOO & MOO). Ο καθορισμός της βέλτιστης λύσης των πρώτων μπορεί να γίνει άμεσα, εφόσον αυτή ορίζεται ως  $\vec{x}^* : f(\vec{x}^*) \leq f(x) \forall x$ , και είναι συνήθως μοναδική. Στην περίπτωση πολλών συναρτήσεων-στόχων όμως, οι οποίες πιθανότατα θα είναι και αντικρουόμενες μεταξύ τους (προφανώς καλύτερη ποιότητα σημαίνει και υψηλότερο κόστος), ο καθορισμός του βέλτιστου δεν είναι μονοσήμαντος. Για το λόγο αυτό εισάγονται οι έννοιες της κατά Pareto κυριαρχίας (Pareto dominance) και του μετώπου Pareto (Pareto front).

Μία υποψήφια λύση  $\vec{x}$  λέγεται ότι κυριαρχεί κατά Pareto μιάς άλλης λύσης  $\vec{y}$ , αν και μόνο αν, όλες οι τιμές των συναρτήσεων στόχων  $f_i(\vec{x})$  είναι μικρότερες ή ίσες από τις αντίστοιχες  $f_i(\vec{y})$ , ενώ ταυτόχρονα υπάρχει τουλάχιστον μία συνάρτηση στόχου για την οποία η λύση  $\vec{x}$  είναι καλύτερη από την  $\vec{y}$ . Ο ορισμός διατυπώνεται μαθηματικά ως:

$$\vec{x} \prec \vec{y} \iff f_i(\vec{x}) \leq f_i(\vec{y}) \forall i \in \{1, \dots, M\} \wedge \exists k \in \{1, \dots, M\} : f_k(\vec{x}) < f_k(\vec{y}) \quad (2.3)$$

Μια λύση  $\bar{x}^*$  χαρακτηρίζεται ως βέλτιστη κατά Pareto, όταν δεν υπάρχει άλλη λύση  $\bar{y}$  τέτοια ώστε  $y \prec x^*$ . Το σύνολο των κατά Pareto βέλτιστων λύσεων ορίζει μια υπερεπιφάνεια στον χώρο των  $M_T$  συναρτήσεων στόχων, η οποία ονομάζεται μέτωπο μη-κυριαρχούμενων λύσεων ή μέτωπο Pareto. Περισσότερα για τα προβλήματα πολλών στόχων αναφέρονται στο αντίστοιχο παράρτημα.

## Αιτιοκρατικές και Στοχαστικές Μέθοδοι

Γενικά, οι μέθοδοι βελτιστοποίησης διακρίνονται σε αιτιοκρατικές (deterministic) και στοχαστικές (stochastic). Η διάκριση έχει να κάνει με τον τρόπο εξερεύνησης του χώρου σχεδιασμού. Οι πρώτες ξεκινούν από μια υποψήφια λύση  $\bar{x}$  και επιχειρούν να τη βελτιώσουν ωθώντας την διαδοχικά προς μια “ευνοϊκή” κατευθυνση που προκύπτει κάθε φορά από υπολογισμό ή προσέγγιση των παραγώγων των αντικειμενικών συναρτήσεων. Οι δεύτερες στηρίζονται στην “οργανωμένα τυχαία” ή τυχηματική παραγωγή πολλών υποψήφιων λύσεων, διαδικασία που οδηγεί στατιστικά στην προσέγγιση των βέλτιστων λύσεων  $\bar{x}^*$ .

Βασικό πλεονέκτημα των αιτιοκρατικών μεθόδων βελτιστοποίησης είναι η μεγάλη ταχύτητα σύγκλισης στη βέλτιστη λύση. Ταυτόχρονα όμως, η σύγκλιση αυτή μπορεί να καταλήξει σε κάποιο τοπικό ακρότατο διαφορετικό από το ολικό βέλτιστο του προβλήματος, ανάλογα με το σημείο εκκίνησης της μεθόδου. Για την εφαρμογή μιας τέτοιας μεθόδου απαιτείται η δυνατότητα υπολογισμού ή προσέγγισης των παραγώγων των συναρτήσεων στόχων, κάτι που δεν μπορεί πάντα να γίνει με ακρίβεια ή οικονομικά. Μια άλλη αδυναμία των αιτιοκρατικών μεθόδων είναι η δυσκολία επέκτασης ή γενίκευσής τους σε διαφορετικά προβλήματα.

Από την άλλη, στις στοχαστικές μεθόδους βελτιστοποίησης η σύγκλιση είναι γενικά πιο αργή, αλλά και με πολύ μικρότερο κίνδυνο εγκλωβισμού σε τοπικό ακρότατο. Επίσης, επειδή χρησιμοποιούν μόνο την τιμή της συνάρτησης στόχου για τη λύση που δοκιμάζεται, είναι πολύ πιο ευέλικτες και μπορούν να χρησιμοποιηθούν για μεγάλο εύρος προβλημάτων. Η βασική αδυναμία των στοχαστικών μεθόδων είναι το μεγάλο πλήθος αξιολογήσεων της αντικειμενικής συνάρτησης που απαιτείται για την εφαρμογή τους.

## 2.2 Εισαγωγή στους Εξελικτικούς Αλγόριθμους

Οι Εξελικτικοί Αλγόριθμοι (Evolutionary Algorithms - EA ) αποτελούν μία από τις δημοφιλέστερες στοχαστικές μεθόδους βελτιστοποίησης [25, 27]. Η αρχή λειτουργίας τους βασίζεται στις παρατηρήσεις του Δαρβίνου για την εξέλιξη των ειδών[17]. Τα άτομα ενός φυσικού πληθυσμού συναγωνίζονται για την απόκτηση πεπερασμένων πόρων. Τα πιο επιτυχημένα από αυτά εξασφαλίζουν την αναπαραγωγή τους και μεταφέρουν τα χαρακτηριστικά τους στους απογόνους τους. Η διαδικασία εμπλουτίζεται και από την επίδραση της μετάλλαξης, η οποία συμβαίνει τυχαία προκαλώντας μικρές αλλαγές στα άτομα του πληθυσμού. Μακροπρόθεσμο αποτέλεσμα όλων αυτών είναι η διαρκής εξέλιξη του πληθυσμού και η επικράτηση χαρακτηριστικών ευνοϊκότερα προσαρμοσμένων στο περιβάλλον. Κατ' αναλογία, ένας εξελικτικός αλγόριθμος παράγει πληθυσμούς λύσεων, από τις οποίες οι καλύτερες επιλέγονται για να παραχθούν και να αναδειχθούν, μέσω κάποιων τελεστών εξέλιξης, νέες λύσεις καλύτερα προσαρμοσμένες στις συναρτήσεις στόχου του προβλήματος [25].

## 2.3 Βασική Δομή Ενός Εξελικτικού Αλγορίθμου

Η λειτουργία ενός EA στην πιο απλή του μορφή στηρίζεται στην αλληλεπίδραση μεταξύ τριών συνόλων: των πληθυσμών των απογόνων  $P_\lambda^g$ , των γονέων  $P_\mu^g$ , και των επιλέκτων (elite) $P_e^g$  κάθε γενιάς ( $g$ ). Ένας τέτοιος αλγόριθμος αναφέρεται και με τη μορφή (μ,λ)EA, υποδηλώνοντας έτσι δύο από τις πιο χαρακτηριστικές παραμέτρους που χρησιμοποιεί.

Από τα  $\lambda$  άτομα του πληθυσμού της γενιάς  $g$  επιλέγονται  $\mu$  υποψήφιας λύσεις ορίζοντας το σύνολο  $P_\mu^g$  γονέων, στο οποίο εφαρμόζονται οι τελεστές εξέλιξης (διασταύρωση και μετάλλαξη) , παράγοντας έτσι το σύνολο των απογόνων της νέας γενιάς  $P_\lambda^{g+1}$ . Το δε σύνολο των επιλέκτων συγκεντρώνει τις  $e$  καλύτερες υποψήφιας λύσεις που έχουν βρεθεί καθ'όλη τη διάρκεια της εξέλιξης.

Ένας μικρός αριθμός λύσεων από το σύνολο των επίλεκτων επανεισάγεται στον πληθυσμό κάθε γενιάς, διατηρώντας με αυτό τον τρόπο καλά ή υποσχόμενα χαρακτηριστικά του συνόλου λύσεων. Η στρατηγική αυτή ονομάζεται ελιτισμός (elitism) ή επιλεκτικότητα.

Για την εκκίνηση του αλγορίθμου είναι απαραίτητη η ρύθμιση από το χρήστη βασικών παραμέτρων, όπως τα  $\mu, \lambda, e$ , τα όρια των μεταβλητών σχεδιασμού και των περιορισμών, το κριτήριο τερματισμού, κ.α. Σημαντική είναι επίσης η επιλογή του τρόπου κωδικοποίησης των λύσεων από τον αλγόριθμο. Κατά τη χρήση δυαδικής κωδικοποίησης (binary encoding), κάθε μεταβλητή αντιπροσωπεύεται από μια συστοιχία bits προκαθορισμένου μήκους (καλούνται και “χρωμοσώματα”), ενώ στην περίπτωση της πραγματικής κωδικοποίησης (real encoding), κάθε μεταβλητή αντιμετωπίζεται ως πραγματικός αριθμός, σύμφωνα με την αναπαράσταση κινητής υποδιαστολής (floating point representation). Σε όλα τα προβλήματα που μελετά η παρούσα εργασία χρησιμοποιείται πραγματική κωδικοποίηση.

Μετά τον καθορισμό των βασικών παραμέτρων, ο  $(\mu, \lambda)$ EA ακολουθεί τα παρακάτω βήματα [1]:

1. Αρχικοποίηση του πληθυσμού της μηδενικής γενιάς ( $g=0$ ). Ο αλγόριθμος παράγει μέσω γεννήτριας τυχαίων αριθμών (RNG - random number generator) τα  $\lambda$  άτομα της γενιάς  $P_\lambda^0$ , δίνοντας τυχαίες τιμές (εντός των προκαθορισμένων ορίων) στις μεταβλητές σχεδιασμού. Είναι δυνατή και η εισαγωγή προκαθορισμένων από το χρήστη λύσεων ως τμήμα του πληθυσμού.
2. Αξιολόγηση και απόδοση βαθμωτού κόστους σε κάθε μέλος του συνόλου των απογόνων  $P_\lambda^g$ . Κάθε άτομο του πληθυσμού ισοδυναμεί με ένα διάνυσμα τιμών των μεταβλητών σχεδιασμού  $\vec{x}$  και εισάγεται στο καθορισμένο πρόγραμμα αξιολόγησης του προβλήματος, το οποίο με τη σειρά του επιστρέφει το διάνυσμα τιμών των συναρτήσεων-στόχων  $\vec{f}(\vec{x})$ .

Το διάνυσμα τιμών  $\vec{f}(\vec{x})$  μετασχηματίζεται σε ένα βαθμωτό μέγεθος κόστους  $\Phi$ , το οποίο σε περίπτωση ενός στόχου ταυτίζεται με τη συνάρτηση  $f$ , ενώ σε περίπτωση πολλών στόχων αποτελεί μέτρο των σχέσεων κυριαρχίας του ατόμου με τα υπόλοιπα άτομα του συνόλου  $P_\mu^g \cup P_\lambda^g \cup P_e^g$ .

Εάν το πρόβλημα περιλαμβάνει περιορισμούς το πρόγραμμα αξιολόγησης παράγει και ένα διάνυσμα  $\vec{C}$ , το οποίο εκφράζει τη θέση του συγκεκριμένου ατόμου στο χώρο των αποδεκτών λύσεων (feasible region). Για κάθε περιορισμό  $c_k$  ορίζεται ένα “τυπικό” ( $d_k$ ) και ένα “χαλαρωμένο” ( $d_k^* > d_k$ ) όριο παραβίασής του. Εάν η  $c_k$  υπερβεί το  $d_k^*$ , αυτόματα αποδίδεται “ποινή θανάτου” (death penalty) στο συγκεκριμένο άτομο, δηλαδή μία πολύ μεγάλη (πρακτικά άπειρη) τιμή κόστους  $\Phi$ .

Τελικά, το συνολικό κόστος που αποδίδεται σε κάθε υποψήφια λύση δίνεται από τη σχέση:

$$\Phi(\vec{x}) = \Phi(\vec{x}) + \prod_{k=1}^K \begin{cases} 1, & c_k \leq d_k \\ \exp(a_k \frac{c_k(x) - d_k}{d_k^* - d_k}), & c_k \geq d_k \end{cases} \quad (2.4)$$

3. Ανανέωση του συνόλου επιλέκτων. Ο πληθυσμός των επιλέκτων  $P_e^g$  σχηματίζεται από την επιλογή των  $e$  καλύτερων (με κριτήριο το χαμηλότερο βαθμωτό κόστος) ατόμων από το σύνολο  $P_\lambda^g \cup P_e^{g-1}$ .
4. Εφαρμογή του τελεστή επιλεκτικότητας ή ελιτισμού. Ένα μικρό πλήθος ατόμων, επιλεγμένο από το σύνολο των elite αντικαθιστά τα χειρότερα μέλη του πληθυσμού.
5. Επιλογή γονέων με εφαρμογή ενός τελεστή επιλογής γονέων στο σύνολο  $P_\lambda^g \cup P_\mu^{g-1}$ . Γενικότερα, άτομα με χαμηλή συνάρτηση κόστους είναι πιθανότερο να επιλεγούν για το σύνολο  $P_\mu^g$ . Εδώ χρησιμοποιείται ο τελεστής πιθανοτικής επιλογής με διαγωνισμό (probabilistic tournament selection): διεξάγεται διαγωνισμός ανάμεσα σε  $k$  τυχαία άτομα, με την καλύτερη λύση να έχει 80% πιθανότητα να επιλεγεί. Η διαδικασία αυτή επαναλαμβάνεται  $\mu$  φορές.
6. Αναπαραγωγή του πληθυσμού. Καθένα από τα  $\lambda$  άτομα του συνόλου  $P_\lambda^{g+1}$  σχηματίζεται από την εφαρμογή του τελεστή διασταύρωσης (crossover recombination operator) σε  $\rho$  γονείς τυχαία επιλεγμένους από το σύνολο  $P_\mu^g$ . Στη συνέχεια, εφαρμόζεται σε κάθε απόγονο ο τελεστής μετάλλαξης (mutation operator). Οι δύο αυτοί τελεστές πραγματοποιούν τις συμπληρωματικές λειτουργίες εξερεύνησης και εκμετάλλευσης του χώρου σχεδιασμού (exploration and exploitation).
7. Έλεγχος κριτηρίου σύγκλισης. Κριτήριο για τη σύγκλιση του αλγορίθμου μπορεί να είναι ο μέγιστος αριθμός κλήσεων του λογισμικού αξιολόγησης ή το μέγιστο πλήθος των γενεών εξέλιξης. Σε περίπτωση που το κριτήριο δεν ικανοποιείται η εξέλιξη συνεχίζεται, δηλαδή ο αλγόριθμος επιστρέφει στο βήμα 3.

## Σχήματα Διασταύρωσης

Τα σχήματα διασταύρωσης που χρησιμοποιούνται στην εργασία αυτήν είναι δύο: το σχήμα ενδιάμεσου ανασυνδυασμού και το σχήμα προσομοιούμενης δυαδικής διασταύρωσης[6].

Κατά το σχήμα του ενδιάμεσου ανασυνδυασμού (intermediate recombination), η κάθε συνιστώσα του απογόνου  $\vec{x}$  είναι προϊόν γραμμικού συνδυασμού του πρώτου γονέα  $\vec{x}^1$  και ενός άλλου  $\vec{x}^{random}$ , τυχαία επιλεγμένου από τους εναπομείναντες  $\rho-1$  γονείς. Συνεπώς, ο απόγονος μπορεί να γραφεί ως:

$$\vec{x} = \vec{x}^1 + r(\vec{x}^{random} - \vec{x}^1), \quad r \in [0,1]$$

όπου τα  $r$  και  $\vec{x}^{random}$  αλλάζουν για κάθε μεταβλητή σχεδιασμού. Από τα όρια της μεταβλητής  $r$  φαίνεται ότι οι συνιστώσες των απογόνων παίρνουν τιμές μέσα στο διάστημα που ορίζουν οι συνιστώσες των γονέων.

Κατά το σχήμα της προσομοιούμενης δυαδικής διασταύρωσης (simulated binary crossover - SBX)[20], για τον προσδιορισμό κάθε συνιστώσας του απογόνου υπολογίζεται πρώτα ένας παράγοντας εξάπλωσης (spread factor):

$$\beta = \begin{cases} (2r_1)^n & , r_1 \leq 0.5 \\ \left(\frac{1}{2r_1}\right)^{n+2} & , r_1 > 0.5 \end{cases} \quad (2.5)$$

όπου το  $r_1$  λαμβάνει τυχαία τιμές στο διάστημα  $[0, 1]$ , ενώ το  $n$  είναι μια θετική σταθερά (όπου χρησιμοποιείται SBX στην εργασία αυτήν το  $n$  έχει την τιμή 1).

Ο παράγοντας εξάπλωσης καθορίζει εάν οι απόγονοι θα μείνουν εντός των ορίων των γονέων ( $\beta < 1$ ) ή θα επεκταθούν ( $\beta > 1$ ). Σε αντίθεση με το σχήμα του ενδιάμεσου ανασυνδυασμού, οι συνιστώσες των απογόνων μπορούν να πάρουν τιμές σε όλο το διάστημα των μεταβλητών σχεδιασμού. Η συνιστώσα του απογόνου  $x_i$  προκύπτει ως γραμμικός συνδυασμός μεταξύ των συνιστωσών των γονέων σύμφωνα με τη σχέση:

$$x_i = \begin{cases} \bar{x}_i - \frac{\beta}{2}(x_i^{random} - x_i^1), & r_2 < 0.5 \\ \bar{x}_i + \frac{\beta}{2}(x_i^{random} - x_i^1), & r_2 \geq 0.5 \end{cases} \quad (2.6)$$

όπου το  $x_i^{random}$  ορίζεται όμοια με παραπάνω, ενώ με  $\bar{x}_i$  συμβολίζεται η μέση τιμή των δύο γονέων.

Το συγκεκριμένο σχήμα διασταύρωσης παράγει απογόνους οι οποίοι τείνουν να βρίσκονται κοντά στους γονείς (spread factor property). Η τάση αυτή ενισχύεται με αύξηση της παραμέτρου  $n$ . Επίσης, η μέση τιμή σε κάθε συνιστώσα παραμένει ίδια για τους πιθανούς απογόνους και τους γονείς (average property).

## Το Σχήμα Μετάλλαξης

Το σχήμα μετάλλαξης εφαρμόζεται αμέσως μετά το σχήμα διασταύρωσης, σε κάθε έναν απόγονο που παράγεται. Οι μικρές, τυχαίες αλλαγές που προκαλούνται από τη μετάλλαξη συνεισφέρουν στη διατήρηση της ποικιλομορφίας του πληθυσμού του ΕΑ. Κατ' αυτόν τον τρόπο αποφεύγεται γενικά η πρόωρη ομογενοποίηση του πληθυσμού, η οποία ευθύνεται συχνά για το τον εγκλωβισμό σε τοπικά ακρότατα της συνάρτησης-στόχου.

Για τους ΕΑ με πραγματική κωδικοποίηση, η μετάλλαξη λειτουργεί ως εξής[6, 24]: για κάθε συνιστώσα του απογόνου παράγεται με τη γεννήτρια τυχαίων αριθμών ένα  $r \in [0, 1]$ . Εάν αυτό είναι μεγαλύτερο από την προκαθορισμένη πιθανότητα μετάλλαξης  $P_m$  (η οποία συνήθως έχει τιμή 3%- 5%), η συνιστώσα μένει ως έχει. Εάν είναι μικρότερο ή ίσο εφαρμόζεται ένας τελεστής  $M(\vec{x}, D)$  ο οποίος τη "σπρώχνει" τυχαία προς το πάνω ή κάτω όριο της. Το μέγεθος των μεταβολών αυτών μειώνεται όσο αυξάνονται οι γενιές της εξέλιξης. Η παραπάνω διαδικασία συνοψίζεται με τις σχέσεις:

$$\vec{x} = \begin{cases} \vec{x}, & r > P_m \\ M(\vec{x}, D), & r \leq P_m \end{cases}$$

$$M(\vec{x}, D) = \begin{cases} \vec{x} + D(g, \vec{U} - \vec{x}), & r_1 > 0,5 \\ \vec{x} + D(g, \vec{L} + \vec{x}), & r_1 \leq 0,5 \end{cases} \quad (2.7)$$

$$D(g, y) = yr_2 \left(1 - \frac{g}{g_{max}}\right)^{0.2}$$

## 2.4 ΕΑ Υποβοηθούμενοι από Μεταπρότυπα

Στα περισσότερα προβλήματα μηχανολογικού σχεδιασμού, το λογισμικό επίλυσης που χρησιμοποιείται για την αξιολόγηση μίας υποψήφιας λύσης (πχ αριθμητική επίλυση των εξισώσεων ροής) είναι ιδιαίτερα χρονοβόρο και ακριβό υπολογιστικά. Αυτό, μαζί με το γεγονός ότι ο αλγόριθμος παράγει μεγάλο

πλήθος υποψήφιας λύσεων, οι οποίες πρέπει να αξιολογηθούν, καθιστά ασύμφορη την εφαρμογή τους.

Η δυσκολία αυτή μπορεί να αντιμετωπιστεί με τη χρήση μεταπροτύπων (metamodels ή surrogate evaluation models), δηλαδή μοντέλων που κατασκευάζονται από την παρατήρηση και επεξεργασία προηγούμενων αξιολογήσεων και υποκαθιστούν το ακριβές λογισμικό αξιολόγησης [2]. Τα μεταπρότυπα απαιτούν μικρό υπολογιστικό κόστος και παρέχουν τη δυνατότητα προσέγγισης των αντικειμενικών συναρτήσεων, μέσω της οποίας γίνεται ο προσδιορισμός των πιο “υποσχόμενων” λύσεων. Οι λύσεις αυτές είναι εκείνες που θα αξιολογηθούν τελικά από τον ακριβή επιλύτη του προβλήματος, εξοικονομώντας έτσι το κόστος “περιττών” κλήσεων.

Η τροποποίηση του ΕΑ που περιγράφηκε παραπάνω αναφέρεται ως Εξελικτικός Αλγόριθμος Υποβοηθούμενος με Μεταπρότυπα (Metamodel Assisted Evolutionary Algorithm - MAEA) [2, 4].

### 2.4.1 Γενική Περιγραφή Λειτουργίας των MAEA

Γενικά οι MAEA κατασκευάζουν τα μεταπρότυπα αντλώντας πληροφορία από τη βάση δεδομένων (database-DB), η οποία περιέχει άτομα που έχουν ήδη αξιολογηθεί από το λογισμικό αξιολόγησης. Τα άτομα από τα οποία εκπαιδεύεται το κάθε μεταπρότυπο ονομάζονται δείγματα εκπαίδευσης (training patterns).

Ένα μεταπρότυπο μπορεί να είναι καθολικό, δηλαδή να λειτουργεί σε όλο το εύρος του χώρου σχεδιασμού, ή να είναι τοπικό, δηλαδή να κατασκευάζεται ειδικά για μια περιοχή γύρω από μία λύση. Επίσης τα μεταπρότυπα μπορούν να είναι αποκομμένα από την εξέλιξη ή συνδεδεμένα με αυτή (off-line και on-line training).

Στην πρώτη περίπτωση, αρχικά αξιολογούνται κάποιες υποψήφιας λύσεις που εκλέγονται βάσει μιας μεθόδου σχεδιασμού πειραμάτων. Οι υποψήφιας λύσεις που εκλέγονται, χρησιμοποιούνται για την εκπαίδευση μεταπροτύπων και, στη συνέχεια, πραγματοποιείται ένας κύκλος βελτιστοποίησης, με συνάρτηση-στόχο την πρόβλεψη του μεταπροτύπου. Το σύνολο λύσεων μετά τη σύγκλιση τροφοδοτείται στο πραγματικό λογισμικό αξιολόγησης, και με τα αποτελέσματα πραγματοποιείται εκ νέου η εκπαίδευση του μεταπροτύπου. Η διαδικασία επαναλαμβάνεται με συνάρτηση-στόχο το νέο μεταπρότυπο[5].

Στην περίπτωση σύνδεσης με την εξέλιξη ο αλγόριθμος εκπαιδεύει άμεσα τα μεταπρότυπα με δείγματα που έχει αποθηκεύσει από την μέχρι τώρα πορεία

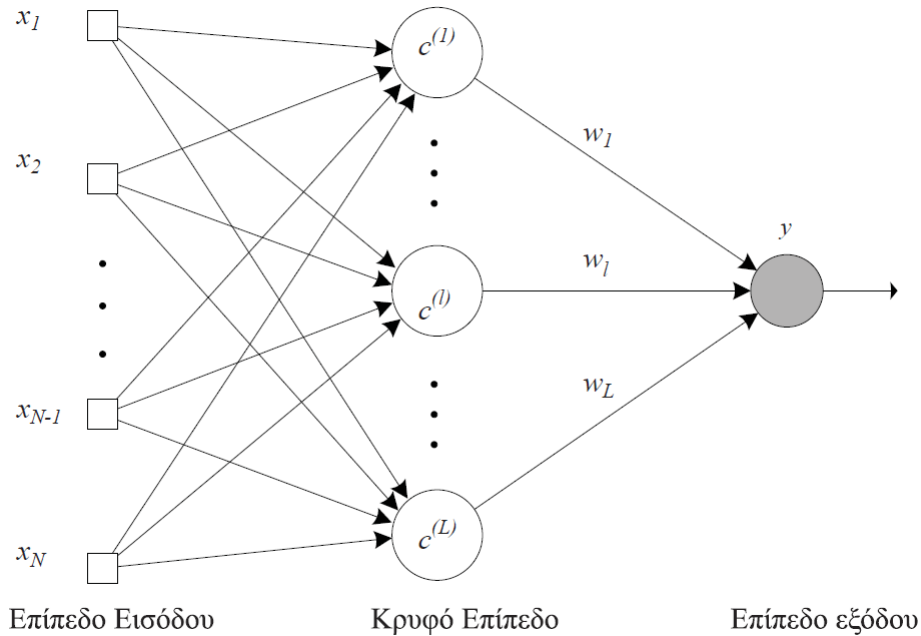


του. Κατά τη διάρκεια της εξέλιξης, γίνεται εναλλασσόμενη χρήση των προσεγγιστικών προτύπων και του ακριβούς λογισμικού αξιολόγησης. Η εναλλαγή αυτή μπορεί να γίνεται είτε από γενιά σε γενιά, είτε εντός της ίδιας γενιάς, όπου όλοι οι απόγονοι αξιολογούνται από τα μεταπρότυπα, και ύστερα ένα μικρό υποσύνολο με τις καλύτερες λύσεις, όπως αυτές προσεγγίστηκαν, αξιολογείται από το πραγματικό λογισμικό αξιολόγησης. Η τεχνική αυτή καλείται προσεγγιστική προ-αξιολόγηση (ΙΠΠΑ - Inexact Pre-Evaluation - IPE) [4].

Τα μεταπρότυπα που χρησιμοποιούν οι ΜΑΕΑ στην εργασία αυτή είναι κυρίως τεχνητά νευρωνικά δίκτυα, ενώ σε ορισμένες περιπτώσεις έγιναν δοκιμές και με μεταπρότυπα πολυωνυμικής προσέγγισης.

#### **2.4.2 Τεχνητά Νευρωνικά Δίκτυα**

Βασική ιδέα πίσω από τα τεχνητά νευρωνικά δίκτυα (ΤΝΔ - Artificial Neural Networks- ANNs) είναι η χρήση πολλών απλών υπολογιστικών μονάδων (νευρώνων) συνδεδεμένων μεταξύ τους για την προσέγγιση ή παρεμβολή της επιθυμητής συνάρτησης. Τα ΤΝΔ που χρησιμοποιούνται στα προβλήματα που μελετώνται είναι Δίκτυα Συναρτήσεων Ακτινικής Βάσης (ΔΣΑΒ -Radial Basis Function Networks - RBFNs)[3, 6].



Σχήμα 2.1: Βασική δομή ΔΣΑΒ μίας εξόδου [3].

Τα δίκτυα συναρτήσεων ακτινικής βάσης αποτελούνται από τρία επίπεδα:

- Το επίπεδο εισόδου, το οποίο αποτελείται από νευρώνες που αντιστοιχούν στις μεταβλητές εισόδου.
- Το κρυφό επίπεδο, του οποίου οι  $L$  νευρώνες συνδέονται με όλους τους νευρώνες εισόδου και εκτελούν ένα μη-γραμμικό μετασχηματισμό. Για τα ΔΣΑΒ του λογισμικού EASY χρησιμοποιείται η συνάρτηση Gauss :

$$G(\vec{x}, r) = \exp\left(-\frac{\|\vec{x} - \vec{c}_i\|}{r^2}\right) \quad (2.8)$$

Το  $r$  είναι η ακτίνα , ενώ σε κάθε νευρώνα αντιστοιχεί διαφορετικό  $\vec{c}_i$ , το οποίο αποκαλείται κέντρο της συνάρτησης ακτινικής βάσης. Το αποτέλεσμα του μη-γραμμικού μετασχηματισμού εξαρτάται μόνο από την απόσταση του διανύσματος εισόδου από το κέντρο  $\vec{c}_i$ .

- Το επίπεδο εξόδου, το οποίο πραγματοποιεί ένα γραμμικό συνδυασμό  $y$  μεταξύ των σημάτων που παράγονται από τους νευρώνες του κρυφού

επιπέδου:

$$y = \sum_{l=1}^L w_l G_l(\vec{x}, r) \quad (2.9)$$

Ο συνδυασμός αυτός είναι και η εκτίμηση της  $f(\vec{x})$  από το  $\Delta\Sigma\text{AB}$ .

Η διαδικασία κατασκευής του μεταπρότυπου για το σημείο πρόβλεψης  $x^*$  από το λογισμικό *EASY*, ξεκινά με την επιλογή των  $T$  πλησιέστερων στο  $x^*$  εγγραφών της *DB*, οι οποίες θα χρησιμοποιηθούν για την εκπαίδευση.

Στη συνέχεια, καθορίζεται εάν το μεταπρότυπο θα εκπαιδευτεί για παρεμβολή ή προσέγγιση. Η παρεμβολή αποτελεί καλύτερη στρατηγική, όταν το σημείο πρόβλεψης μαζί με τα δείγματα εκπαίδευσης συνιστούν ένα “ομοίμορφο” σύνολο. Αντίθετα, όταν το σημείο πρόβλεψης “ξεχωρίζει” σημαντικά από τα δείγματα εκπαίδευσης, προτιμάται η προσέγγιση των συναρτήσεων στόχων. Η επιλογή πραγματοποιείται με χρήση της αλλοτριότητας (strangeness), ενός μεγέθους που δείχνει κατά πόσο ένα σημείο είναι συναφές με το σύνολο (κλάση) στο οποίο ανήκει. Συγκεκριμένα, αν ένα σημείο  $\vec{x}$  ανήκει σε μια κλάση  $K$ , η αλλοτριότητά του ορίζεται ως [13]:

$$\alpha(\vec{x}) = \sum_{k \in K} |\vec{x} - \vec{x}_k| \quad (2.10)$$

ισούται δηλαδή με το άθροισμα των αποστάσεων του από τα υπόλοιπα μέλη της κλάσης  $K$ . Κατά τον αλγόριθμο επιλογής μεταξύ παρεμβολής ή προσέγγισης, κάθε σημείο (δείγμα εκπαίδευσης ή το σημείο πρόβλεψης) θεωρείται ότι ανήκει σε μια κλάση που απαρτίζεται από τα  $\sqrt{T}$  πλησιέστερα σε αυτό δείγματα εκπαίδευσης. Για κάθε σημείο υπολογίζεται η αλλοτριότητά του, βάσει της σχέσης 2.10. Στη συνέχεια, υπολογίζεται η ποσότητα

$$t^* = \frac{\#\{\vec{x}_i : \alpha(\vec{x}^*) > \alpha_i\}}{T + 1} \quad (2.11)$$

δηλαδή το πλήθος των σημείων που έχουν χαμηλότερη αλλοτριότητα από αυτή του σημείου πρόβλεψης  $\vec{x}^*$ , προς το συνολικό πλήθος των σημείων,  $T+1$ . Εάν η ποσότητα αυτή είναι μικρότερη από μια παράμετρο  $t_{crit}$ , τότε πραγματοποιείται παρεμβολή, ενώ σε διαφορετική περίπτωση επιλέγεται η προσέγγιση, θεωρώντας ότι το σημείο πρόβλεψης δεν είναι τοπολογικά συγγενές με το σύνολο των δειγμάτων εκπαίδευσης.

Σε περίπτωση παρεμβολής από  $T$  δείγματα  $(\vec{x}^{(i)}, \hat{y}^{(i)})$ ,  $i = 1, T$  το κρυφό επίπεδο φτιάχνεται από  $T$  νευρώνες, με το κέντρο κάθε νευρώνα να ταυτίζεται με το διάνυσμα  $x^{(i)}$ . Στη συνέχεια γίνεται ο υπολογισμός των κατάλληλων βαρών  $w_i$  με λύση του γραμμικού συστήματος:

$$H\vec{w} = \vec{\hat{y}} \quad (2.12)$$

όπου

$$H = G(\|\vec{x}_i - \vec{x}_j\|, r), (i, j) = 1, T \quad (2.13)$$

Η επιλογή της συνάρτησης Gauss για το μη-γραμμικό μετασχηματισμό του κρυφού επιπέδου (σχέση 2.8) εξασφαλίζει την ύπαρξη του  $H^{-1}$  και άρα την ύπαρξη λύσης της εξ. 2.13.

Κατά την εκπαίδευση του  $\Delta\Sigma\text{AB}$  για προσέγγιση, το κρυφό επίπεδο αποτελείται από  $L$  νευρώνες με  $L < T$ . Τα κέντρα των  $\Sigma\text{AB}$  δεν ταυτίζονται πλέον με τις συντεταγμένες των δειγμάτων εκπαίδευσης, αλλά τοποθετούνται σε περιοχές του χώρου όπου η βάση δεδομένων παρουσιάζει αυξημένη πυκνότητα δειγμάτων [4, 3]. Με αυτόν τον τρόπο εξασφαλίζεται μείωση του σφάλματος. Ο προσδιορισμός των βαρών γίνεται με λύση του προβλήματος

$$\min \epsilon = \frac{1}{T} \sum_{t=1}^T \|\hat{y}^{(t)} - y^{(t)}\| \quad (2.14)$$

το οποίο λύνεται με τη μέθοδο των ελαχίστων τετραγώνων.

Η κατασκευή των προσεγγιστικών μεταπροτύπων μπορεί να γίνει ακόμα πιο στιβαρή με την εισαγωγή των συντελεστών σημαντικότητας κάθε μεταβλητής (Importance Factors) [4, 2]. Οι συντελεστές αυτοί, που υπολογίζονται ούτως ή άλλως κατά τη λύση του προβλήματος 2.14, τροποποιούν τη νόρμα που χρησιμοποιεί το κρυφό επίπεδο για τον υπολογισμό της απόκρισης, και έχουν τη μορφή

$$I_p = \frac{\frac{|\partial y|}{\partial x_p}}{\sum_{n=1}^N \frac{|\partial y|}{\partial x_n}} \quad (2.15)$$

Μέσω της αντικατάστασης της νόρμας  $\|\vec{x} - \vec{c}_l\| = \sqrt{\sum_{n=1}^N (x_n - c_{nl})^2}$  από την

$$\|\vec{x} - \vec{c}_l\|_w = \sqrt{\sum_{n=1}^N N I_n (x_n - c_{nl})^2} \quad (2.16)$$

στη σχέση 2.8, εισάγεται πληροφορία για τις περισσότερό ή λιγότερο σημαντικές κατευθύνσεις στη συγκεκριμένη περιοχή, γεγονός που συμβάλλει στη μείωση του θορύβου. Η μέθοδος αυτή βελτιώνει την ικανότητα γενίκευσης του μεταπροτύπου, ειδικά για περιπτώσεις προβλημάτων όπου ο χώρος σχεδιασμού έχει μεγάλη διάσταση.

### 2.4.3 Πολυωνυμική Προσέγγιση

Ένα πιο απλός και άμεσος τρόπος κατασκευής μεταπροτύπων είναι μέσω πολυωνυμικής προσέγγισης. Συνήθως προτιμάται προσέγγιση με πολυώνυμο πρώτου ή δεύτερου βαθμού [8]. Η προβλεπόμενη απόκριση μιας συνάρτησης στόχου έχει δηλαδή τη μορφή

$$y = \beta_0 + \sum_{i=1}^N \beta_i x_i \quad , \quad \text{ή} \quad (2.17)$$

$$y = \beta_0 + \sum_{i=1}^N \beta_i x_i + \sum_{i=1}^N \alpha_{ii} x_i^2 + \sum_{i=1}^N \sum_{j=i+1}^N \alpha_{ij} x_i x_j \quad (2.18)$$

Για  $T$  δείγματα εκπαίδευσης, η εκπαίδευση του μεταπροτύπου πραγματοποιείται με ελαχιστοποίηση του σφάλματος:

$$\min \epsilon^2 = \sum_{t=1}^T (\hat{y}(\vec{x}_t) - y(\vec{x}_t))^2 \quad (2.19)$$

η οποία επίσης επιλύεται με τη μέθοδο των ελαχίστων τετραγώνων.

## 2.5 Λογισμικό EASY

Τα προβλήματα που μελετώνται στην εργασία αυτήν επιλύθηκαν με χρήση του λογισμικού EASY (Evolutionary Algorithm SYstem)[24]. Πρόκειται για ένα γενικής χρήσης λογισμικό βελτιστοποίησης, που αναπτύχθηκε στη Μονάδα Παράλληλης Υπολογιστικής Ρευστοδυναμικής & Βελτιστοποίησης (ΜΠΥΡ & Β) του ΕΜΠ. Το λογισμικό EASY μπορεί να χρησιμοποιηθεί για την επίλυση προβλημάτων μονοκριτηριακής και πολυκριτηριακής βελτιστοποίησης, με ή χωρίς περιορισμούς, μέσω στοχαστικών, αιτιοκρατικών ή συνδυασμένων μεθόδων.

Παρέχει επίσης τη δυνατότητα χρήσης μεταπροτύπων προκειμένου να περιορίσει τις κλήσεις στο, ενδεχομένως ακριβό υπολογιστικά, ακριβές λογισμικό αξιολόγησης. Τα μεταπρότυπα μπορούν να λειτουργούν με τεχνητά νευρωνικά δίκτυα (με ή χωρίς συντελεστές σημαντικότητας), μεθόδους απόκρισης επιφανείας πολυωνύμων, μηχανές διανυσματικής υποστήριξης (support vector machines), μοντέλα Kriging, ή και να διατίθενται στο πρόγραμμα από εξωτερικό λογισμικό.

Για τη διεξαγωγή βελτιστοποίησης με εξελικτικούς αλγορίθμους στο λογισμικό EASY, το μόνο που απαιτείται είναι η ύπαρξη ενός εξωτερικού προγράμματος αξιολόγησης, και ενός απλού λογισμικού επεξεργασίας που εξασφαλίζει την μεταξύ τους επικοινωνία : για κάθε υποψήφια λύση ο EASY δημιουργεί ένα αρχείο της μορφής *task.dat*, που περιέχει το πλήθος και την τιμή των μεταβλητών σχεδιασμού. Το αρχείο αυτό, υφίσταται κατάλληλη επεξεργασία και τροφοδοτείται στο λογισμικό αξιολόγησης. Τα αποτελέσματα του προγράμματος αξιολόγησης, μέσω μετεπεξεργασίας, επιστρέφονται στον EASY στη μορφή των αρχείων *task.res* και *task.cns*, με τα αποτελέσματα για τις συναρτήσεις στόχου και τους περιορισμούς αντίστοιχα. Με τον τρόπο αυτό υλοποιείται η διαδικασία εξέλιξης που περιγράφηκε στις προηγούμενες ενότητες.

## Κεφάλαιο 3

# Ανάλυση σε Κύριες Συνιστώσες και Εξελικτικοί Αλγόριθμοι

### 3.1 “Κακώς Τοποθετημένα” Προβλήματα και ΕΑ

Όπως προαναφέρθηκε, βασικό μειονέκτημα των ΕΑ είναι ο μεγάλος αριθμός κλήσεων του ακριβούς λογισμικού αξιολόγησης που απαιτούν. Οι ΜΑΕΑ επιχειρούν να αντισταθμίσουν το μειονέκτημα αυτό με την εισαγωγή της προσεγγιστικής προ-αξιολόγησης λύσεων, και γενικά το επιτυγχάνουν. Παρ’όλα αυτά, η χρήση τους στα μεγάλης κλίμακας προβλήματα βελτιστοποίησης που προκύπτουν στο πεδίο της βιομηχανίας συχνά παρουσιάζει δυσκολίες. Ο λόγος για αυτή τη δυσκολία είναι διπλός: πρώτον, τα προβλήματα αυτά χαρακτηρίζονται από πολύ μεγάλο αριθμό μεταβλητών σχεδιασμού (μπορεί να είναι της τάξεως των εκατοντάδων) και δεύτερον, η συνάρτηση στόχου που τα αντιπροσωπεύει είναι συνήθως “κακώς τοποθετημένη”. Τα δύο αυτά χαρακτηριστικά μαζί συνιστούν την “κατάρρα της μεγάλης διάστασης” (curse of dimensionality) [6].

Η “καλή” ή “κακή” τοποθέτηση ενός προβλήματος εξαρτάται σε μεγάλο βαθμό από την παραμετροποίησή του και το πώς αυτή επηρεάζει την αντικειμενική συνάρτηση. Απαραίτητοι για την διευκρίνιση της έννοιας είναι οι ορισμοί που ακολουθούν:

**Ανισοτροπία της Συνάρτησης Στόχου:** Ως ανισότροπη ορίζεται μία συνάρτηση στόχου  $f$  εάν η ίδια μεταβολή  $\Delta x$  στις διάφορες συνιστώσες  $x_i$  του διανύσματος  $\vec{x}$  προκαλεί διαφορετικές μεταβολές  $\Delta f$  στην τιμή της. Ένα μέτρο που μπορεί να χρησιμοποιηθεί για τον προσδιορισμό της ανισοτροπίας της  $f$  είναι η διακύμανση της τιμής της μερικής παραγώγου  $\frac{\partial f}{\partial x_i}$ , για τις διαφορετικές τιμές του δείκτη  $i$ , στην περιοχή της βέλτιστης λύσης της συνάρτησης. Η ύπαρξη ανισοτροπίας σε μια συνάρτηση υποδηλώνει απλά ότι υπάρχουν περισσότερο και λιγότερο σημαντικές μεταβλητές που επηρεάζουν την τιμή της.

**Διαχωρισιμότητα της Συνάρτησης Στόχου:** Μία συνάρτηση στόχου  $f(\vec{x})$ ,  $\vec{x} = (x_1, \dots, x_n)$  θα ονομάζεται διαχωρίσιμη (separable) ως προς τη μεταβλητή  $x_i$  εάν, η βέλτιστη τιμή  $x_i^*$  της συνιστώσας  $x_i$  είναι ανεξάρτητη από την τιμή των υπόλοιπων μεταβλητών σχεδιασμού. Η συνάρτηση θεωρείται διαχωρίσιμη εάν είναι διαχωρίσιμη για όλες τις μεταβλητές  $x_i$ .

Προβλήματα των οποίων η συνάρτηση στόχου είναι μη-διαχωρίσιμη και ανισότροπη θεωρούνται “κακώς τοποθετημένα”. Θα μπορούσε να ειπωθεί ότι, στη διατύπωση ενός “κακώς τοποθετημένου” προβλήματος, όχι μόνον οι μεταβλητές αυτές καθαυτές, αλλά και η αλληλεπίδραση μεταξύ τους επηρεάζει την τελική τιμή της  $f$ . Έτσι, δε γίνεται πλέον λόγος για σημαντικότερες μεταβλητές, αλλά για σημαντικότερες διευθύνσεις στο χώρο σχεδιασμού. Η ιδιότητα αυτή ονομάζεται και επίσταση (epistasis), ένας όρος που στο πεδίο της γενετικής χρησιμοποιείται για να περιγράψει την αλληλεπίδραση μη-αλληλόμορφων γονιδίων [27].

Η μη-διαχωρισιμότητα της συνάρτησης στόχου επιδρά αρνητικά στη σύγκλιση ενός τυπικού ΕΑ. Αυτό μπορεί να γίνει κατανοητό από το εξής απλουστευμένο παράδειγμα:

Έστω ένα πρόβλημα μονοκριτηριακής βελτιστοποίησης τριών μεταβλητών  $x_1, x_2, x_3$ , με την κάθε μία να παίρνει 100 διακριτές τιμές. Εάν το πρόβλημα είναι διαχωρίσιμο ως προς τις  $x_1, x_2, x_3$ , και ο αλγόριθμος μπορεί να αξιοποιήσει αυτή την πληροφορία (όπως συμβαίνει πρακτικά για τους τελεστές εξέλιξης του τυπικού ΕΑ [31]) τότε: αρκεί να βρεθεί το βέλτιστο  $x$  για το πρόβλημα, έστω  $x_1^*$ , κάνοντας δοκιμές με σταθερά τα  $x_2, x_3$ . Όμοια μπορεί να βρεθεί το  $x_2^*$ , κρατώντας σταθερά τα  $x_1^*, x_3$ , και αντίστοιχα το  $x_3^*$ . Τελικά η βέλτιστη λύση  $(x_1^*, x_2^*, x_3^*)$  θα ευρεθεί μετά από το πολύ 300 αξιολογήσεις της συνάρτησης στόχου. Εάν όμως το πρόβλημα είναι μη-διαχωρίσιμο, τότε δεν



νοείται βέλτιστη τιμή μιας μεταβλητής ανεξάρτητη των άλλων δύο, και επομένως ο αλγόριθμος είναι αναγκασμένος να εξερευνά ολόκληρο τον τριδιάστατο χώρο ταυτοχρονως, που στο συγκεκριμένο παράδειγμα σημαίνει έναν αριθμό σημείων της τάξης των  $100^3$ .

Από το παράδειγμα φαίνεται η επίδραση που μπορεί να έχει η διατύπωση ενός προβλήματος  $N$  διαστάσεων στο υπολογιστικό κόστος που απαιτείται για την επίλυσή του με ΕΑ: σε περίπτωση διαχωρίσιμου προβλήματος το υπολογιστικό κόστος αυξάνει γραμμικά με τον αριθμό διαστάσεων  $N$ , ενώ σε περίπτωση “κακώς τοποθετημένου” προβλήματος η αύξηση του κόστους είναι υπεργραμμική. Επιπρόσθετα, η δυσκολία που συναντά ο ΕΑ κατά τη λύση ενός μη-διαχωρίσιμου προβλήματος οξύνεται με αύξηση του βαθμού ανισοτροπίας του[6].

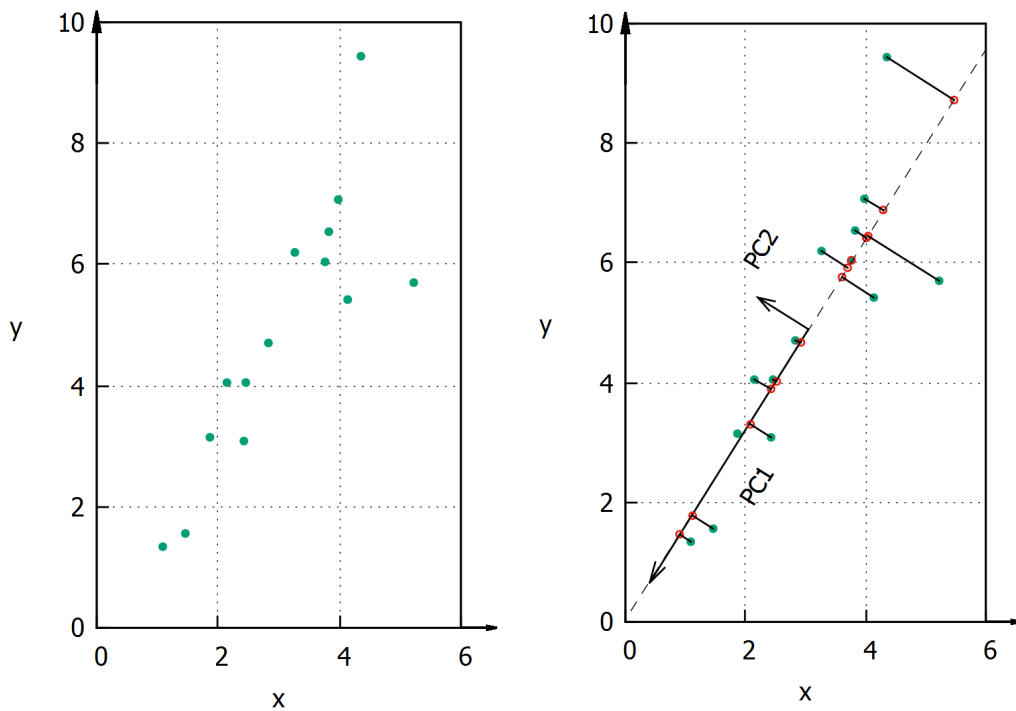
Το μεγάλο πλήθος διαστάσεων έχει αρνητική επίπτωση και στους ΜΑΕΑ. Πέρα από την άμεση επιβάρυνση στην εξέλιξη, η οποία όπως εξηγήθηκε παραπάνω είναι ακόμα πιο έντονη στα μη-διαχωρίσιμα προβλήματα, υπάρχει και η έμμεση επιβάρυνση λόγω εκπαίδευσης των μεταπροτύπων: τα μεταπρότυπα, για να διατηρήσουν την ακρίβειά τους σε ένα πρόβλημα με μεγάλο αριθμό διαστάσεων, απαιτούν αυξημένο αριθμό δειγμάτων εκπαίδευσης. Αυτό σημαίνει ότι θα απαιτούνται και περισσότερες εγγραφές στη βάση δεδομένων και, άρα, ότι η εκκίνηση της προσεγγιστικής προ-αξιολόγησης θα έρχεται πιο αργά στην πορεία της εξέλιξης. Ταυτόχρονα, περισσότερα δείγματα εκπαίδευσης καθιστούν και ακριβότερη υπολογιστικά την εκπαίδευση του κάθε μεταπροτύπου.

Οι δυσκολίες που αναφέρθηκαν ισχύουν και για μη-διαχωρίσιμα προβλήματα πολυκριτηριακής βελτιστοποίησης. Μάλιστα, από τον ορισμό της συνάρτησης απόδοσης βαθμωτού κόστους  $\Phi$ , η οποία σχηματίζεται από το μέτωπο μη-κυριαρχούμενων λύσεων κάθε γενιάς αποδίδοντας χαμηλή τιμή (κοντά στο 0) στα μέλη του και πολύ μεγαλύτερη στα κυριαρχούμενα μέλη, αυτόματα ορίζονται περισσότερο ή λιγότερο σημαντικές διευθύνσεις για τη  $\Phi$ , εκείνες που αντιστοιχούν σε κίνηση κάθετη στο μέτωπο και παράλληλη σε αυτό αντίστοιχα[6]. Συνεπώς, η συνάρτηση  $\Phi$  αναμένεται να είναι πολύ ανισότροπη.

Το εμπόδιο αυτό της “κατάρας της μεγάλης διάστασης”, όπως περιγράφηκε παραπάνω, μπορεί να αντιμετωπιστεί με την εφαρμογή μιας μεθόδου στατιστικής ανάλυσης, η οποία θα εντοπίζει τις σημαντικότερες διευθύνσεις του χώρου σχεδιασμού και θα “επαναπροσανατολίζει” τις λειτουργίες του εξελικτικού αλγορίθμου βάσει αυτών. Η μέθοδος αυτή είναι η ανάλυση σε κύριες συνιστώσες (ΑσΚΣ - Principal Component Analysis - PCA )

## 3.2 Μέθοδος ΑσκΣ και Εξελικτικοί Αλγόριθμοι

Η ανάλυση σε κύριες συνιστώσες είναι μια πολύ διαδεδομένη μέθοδος στατιστικής ανάλυσης, που χρησιμοποιείται συνήθως για τον εντοπισμό χαρακτηριστικών σε μεγάλα σετ δεδομένων. Τα δεδομένα αυτά μπορούν να αντιπροσωπεύουν οτιδήποτε, από τμήματα μιας ψηφιακής εικόνας μέχρι αλληλουχίες DNA [26, 34]. Η ουσιαστική λειτουργία της είναι να ορίζει, μέσω ενός γραμμικού μετασχηματισμού, ένα καινούριο σύνολο μεταβλητών, οι οποίες θα “εξηγούν” με αποδοτικότερο τρόπο τη διακύμανση των δεδομένων, διευκολύνοντας έτσι την εξαγωγή συμπερασμάτων και το διαχωρισμό της ουσιαστικής πληροφορίας από τον ενδεχόμενο θόρυβο [29].



Σχήμα 3.1: Παράδειγμα γραμμικής ΑσΚΣ σε δείγμα δεδομένων δύο διαστάσεων. Η προβολή στον άξονα  $PC1$  αποτελεί την κυριότερη συνιστώσα κάθε σημείου.

Το σχήμα 3.1 αποτελεί ένα απλό παράδειγμα εφαρμογής της μεθόδου ανάλυσης σε κύριες συνιστώσες, σε ένα δείγμα που αποτελείται από 14 σημεία στον διδιάστατο χώρο. Με την εφαρμογή της ΑσΚΣ, τα δεδομένα εκφράζονται από το αρχικό σύστημα συντεταγμένων  $x, y$  σε ένα νέο σύστημα συντεταγμένων  $PC1, PC2$ .

Στο νέο αυτό σύστημα, τα δεδομένα εμφανίζουν πολύ μεγάλη διακύμανση ως προς τον άξονα  $PC1$ , ενώ πολύ μικρή διακύμανση ως προς τον άξονα  $PC2$ . Για την ακρίβεια, οι άξονας  $PC1$  είναι η διεύθυνση ως προς την οποία τα δεδομένα έχουν τη μεγαλύτερη δυνατή διακύμανση και, υπο αυτήν την έννοια, εκείνη που φέρει και την περισσότερη πληροφορία [29].

Έτσι, γίνεται δυνατή μια αδρή περιγραφή του δείγματος, ως ένα σύνολο σημείων πάνω σε μία ευθεία (συνιστώσα  $PC1$ ), με κάποιες μικρές παρεκκλίσεις από αυτή (συνιστώσα  $PC2$ ), κάτι που δεν ήταν προφανές στο αρχικό σύστημα συντεταγμένων.

Κατά την εφαρμογή της στους εξελικτικούς αλγόριθμους, στόχος της ΑσΚΣ είναι η επαναδιατύπωση των μεταβλητών σχεδιασμού, ώστε η αντικειμενική συνάρτηση να αποκτήσει “διαχωρίσιμη” μορφή (όσο αυτό είναι δυνατό). Αυτό διευκολύνει τον ταχύτερο εντοπισμό των ποιοτικότερων λύσεων από τον ΕΑ, ενώ μπορεί να βοηθήσει και στην κατασκευή αποδοτικότερων μεταπροτύπων, όπως θα εξηγηθεί στη συνέχεια. Στη βασική της μορφή, η επαναδιατύπωση πραγματοποιείται με ένα γραμμικό μετασχηματισμό από το χώρο σχεδιασμού στον χώρο των κυρίων διευθύνσεων (principal directions).

Ο προσδιορισμός του κατάλληλου μετασχηματισμού πραγματοποιείται αξιοποιώντας πληροφορία από κάποιο χαρακτηριστικό σύνολο (πλήθους  $M$ ) του πληθυσμού. Το σύνολο αυτό αποτελεί το δείγμα από το οποίο εξάγονται οι κύριες κατευθύνσεις, θεωρώντας ότι αυτές αντιπροσωπεύουν τη συσχέτιση μεταξύ συναρτήσεων στόχων και μεταβλητών σχεδιασμού [6]. Στην εργασία αυτή, ο ΕΑ υπολογίζει τις κύριες κατευθύνσεις σε κάθε γενιά χρησιμοποιώντας το σύνολο των απογόνων, δηλαδή  $M=\lambda$ . Συγκεκριμένα ακολουθείται η παρακάτω διαδικασία [29]:

- Προσδιορισμός του συνόλου των απογόνων  $P_\lambda^g$ , και κανονικοποίηση του, έτσι ώστε να έχει μέση τιμή  $\mu_i$  μηδενική και διακύμανση  $\sigma_{ii}$ , ( $i = 1, \dots, N$ ) ίση με ένα. Τα δεδομένα καταχωρούνται στο  $N \times M$  μητρώο  $X$ , και αποτελούν το δείγμα από το οποίο θα προκύψουν οι κύριες κατευθύνσεις.

- Υπολογισμός του  $N \times N$  μητρώου συνδιακύμανσης του δείγματος (covariance matrix)

$$P = \frac{1}{M} X X^T \quad (3.1)$$

- Εύρεση των ιδιοτιμών και ιδιοδιανυσμάτων του μητρώου συνδιακύμανσης  $P$ . Το πρόβλημα που επιλύεται έχει τη μορφή:

$$P \vec{u}^r = \lambda \vec{u}^r \quad (3.2)$$

- Σύμφωνα με το φασματικό θεώρημα, ο συμμετρικός πίνακας  $P$  ξαναγράφεται στη μορφή  $P = U \Lambda U^{-1}$ , όπου  $U$  είναι το ορθοκανονικό μητρώο των ιδιοδιανυσμάτων, ενώ  $\Lambda$  το διαγώνιο μητρώο με τις ιδιοτιμές που τους αντιστοιχούν[35]. Το μητρώο  $U$  περιλαμβάνει τις κύριες κατευθύνσεις του δείγματος, ενώ οι ιδιοτιμές  $\lambda$  εκφράζουν τη διακύμανση που αναλογεί στις κατευθύνσεις αυτές[29].

Ο πίνακας  $U$  μπορεί να αξιοποιηθεί με δύο διαφορετικούς τρόπους από τον εξελικτικό αλγόριθμο: στην εφαρμογή των τελεστών εξέλιξης στον χώρο των κύριων κατευθύνσεων αντί του αρχικού χώρου σχεδιασμού (εκδοχή EA(L)), ή, στην περίπτωση προσεγγιστικής προ-αξιολόγησης, στην εκπαίδευση μετα-προτύπων μικρότερης διάστασης (εκδοχή M(L)EA).

### 3.2.1 Η εκδοχή EA(L)

Στην εκδοχή EA(L) οι υποψήφιοι γονείς  $\vec{x}$  κωδικοποιούνται σε ένα νέο, “στραμμένο” σύστημα συντεταγμένων σύμφωνα με τη σχέση:

$$\vec{x}^* = U(\vec{x} - \vec{\mu}) \quad (3.3)$$

Στο νέο αυτό σύστημα συντεταγμένων όπου η συνάρτηση στόχου είναι, κατά το δυνατόν, διαχωρίσιμη, λαμβάνει χώρα η εφαρμογή του τελεστή διασταύρωσης. Επίσης, το σχήμα μετάλλαξης τροποποιείται ώστε να δίνει βάρος στις λιγότερο εξερευνηθείσες κατευθύνσεις, δηλαδή εκείνες που αντιστοιχούν σε χαμηλότερη ιδιοτιμή (και άρα διακύμανση) [6]. Συγκεκριμένα, η πιθανότητα μετάλλαξης της  $i$  κύριας συνιστώσας δίνεται από τον τύπο

$$p_m^i = 0.1 p_m + \frac{0.9 p_m N}{D_V} \cdot \frac{V_{max} - V_i}{V_{max} - V_{min}} \quad (3.4)$$

όπου  $D_v = \sum \frac{V_{max} - V_i}{V_{max} - V_{min}}$ , και  $V_i$  η διακύμανση των δεδομένων κατά την ιδιοκατεύθυνση  $i$ .

Όπως θα φανεί και στη συνέχεια από τα προβλήματα που μελετώνται, η δράση αυτή των εξελικτικών τελεστών στο χώρο των κύριων κατευθύνσεων βελτιώνει σημαντικά την απόδοσή τους. Οι απόγονοι  $\vec{y}$  που προκύπτουν “μεταφράζονται” στον αρχικό χώρο σχεδιασμού με την εφαρμογή του αντίστροφου μετασχηματισμού:

$$y = U^{-1}y^* + \mu \quad (3.5)$$

### 3.2.2 Η εκδοχή M(L)ΑΕΑ

Κατα την εκδοχή M(L)ΑΕΑ, ο υπολογισθείς χώρος των ιδιοκατευθύνσεων χρησιμοποιείται στην εκπαίδευση των μεταπροτύπων που θα προ-αξιολογήσουν τις υποψήφιες λύσεις. Συγκεκριμένα, πριν την εκπαίδευση του εκάστοτε ΔΣΑΒ γίνεται στροφή των δειγμάτων με χρήση της σχέσης 3.3. Στη συνέχεια, αποκόπτονται από το δίκτυο οι νευρώνες εισόδου που αντιστοιχούν στις ιδιοκατευθύνσεις με χαμηλότερες ιδιοτιμές. Έτσι το δίκτυο εκπαιδεύεται για ένα πρόβλημα με πλήθος διαστάσεων μικρότερο του αρχικού (καθορισμένο από το χρήστη).

Αυτό βελτιώνει την απόδοση του μεταπροτύπου με δύο τρόπους: πρώτον, σε χώρο χαμηλότερης διάστασης απαιτούνται λιγότερα δείγματα για την εκπαίδευση ενός ακριβούς μεταπροτύπου, και δεύτερον, κρατώντας τις ιδιοκατευθύνσεις με μεγαλύτερη διακύμανση, το μεταπρότυπο εκπαιδεύεται σε διαστάσεις που έχουν εξερευνηθεί περισσότερο και, άρα, με ποιοτικότερη πληροφορία. Επιπλέον, λιγότερος αριθμός δειγμάτων εκπαίδευσης για κάθε μεταπρότυπο σημαίνει και ότι γενικά θα απαιτούνται λιγότερες εγγραφές στη βάση δεδομένων, επομένως η διαδικασία προσεγγιστικής προ-αξιολόγησης μπορεί να ξεκινήσει νωρίτερα στην πορεία της εξέλιξης.

## 3.3 Η Μέθοδος Kernel PCA

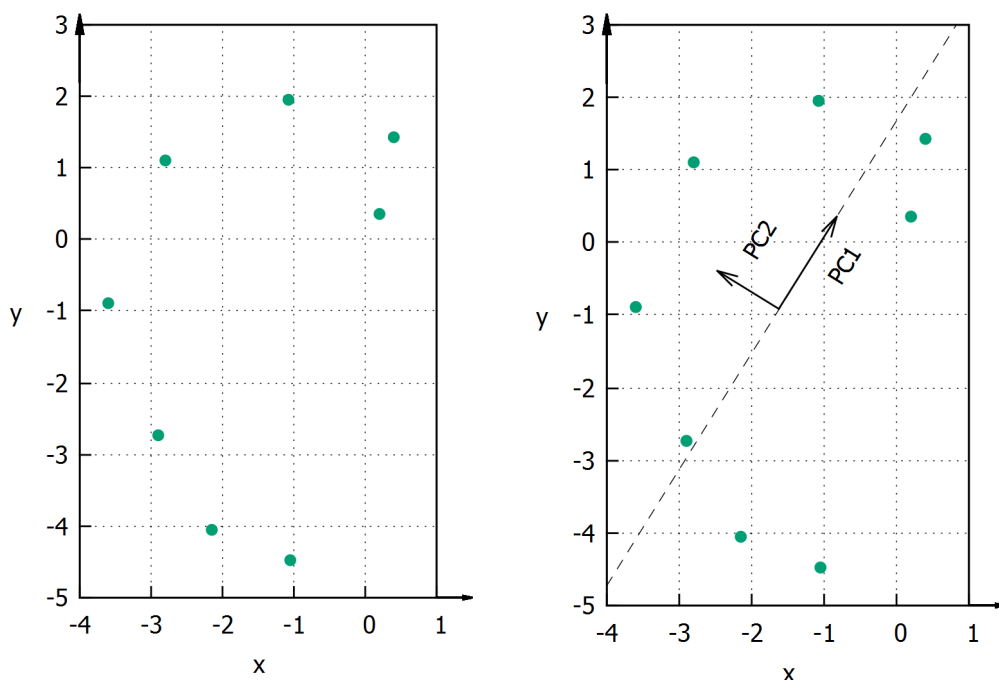
Η γραμμική μέθοδος που αναλύθηκε παραπάνω μπορεί να γενικευθεί, κάνοντας δυνατό τον εντοπισμό και μη-γραμμικών χαρακτηριστικών ενός δείγματος.

Την ιδέα αυτή υλοποιεί η μέθοδος ανάλυσης σε κύριες συνιστώσες με συνάρτηση πυρήνα (Kernel PCA), η οποία βασίζεται στην παρατήρηση ότι

ένα σύνολο μη-γραμμικά διαχωρίσιμων δεδομένων πλήθους  $M$  και διάστασης  $N$  μπορεί να μετατραπεί σε διαχωρίσιμο μέσω μιας μη-γραμμικής απεικόνισης  $\phi : x \in \mathbb{R}^N \rightarrow \phi(x) \in \mathbb{R}^L$  όπου  $L \gg N$  (θεώρημα του Cover[16]). Ο χώρος των  $\phi$  καλείται χώρος χαρακτηριστικών (feature space).

Η ίδια παρατήρηση εξηγεί και την αποτελεσματικότητα των ΔΣΑΒ στην προσέγγιση συναρτήσεων, που αναφέρθηκε στο προηγούμενο κεφάλαιο.

Το πλεονέκτημα αλλά και μια διαισθητική εξήγηση της μεθόδου απεικονίζονται στα σχήματα 3.2, 3.3 και 3.4:

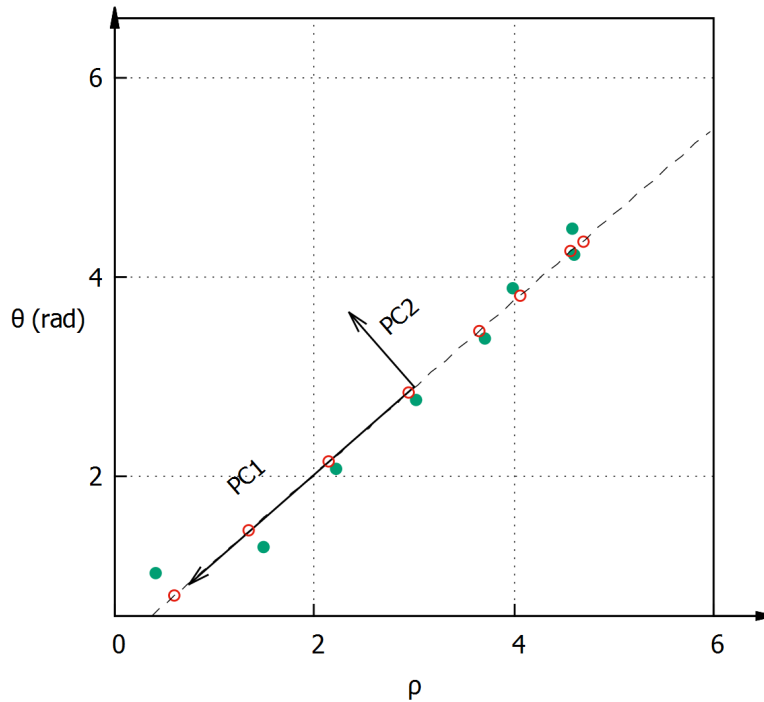


Σχήμα 3.2: Παράδειγμα δείγματος με μη-γραμμικά διαχωρίσιμα χαρακτηριστικά.

Εάν στο δείγμα του σχήματος 3.2 εφαρμοστεί η γραμμική μέθοδος ΑΣΚΣ, οι κύριες κατευθύνσεις που προκύπτουν δεν παρέχουν καλύτερη αναπαράσταση των δεδομένων. Αυτό συμβαίνει επειδή δεν υπάρχει κάποια ευθεία ως προς την οποία τα σημεία “ταιριάζουν” ιδιαίτερα.

Ωστόσο, εάν στο δείγμα εφαρμοστεί ο μη-γραμμικός μετασχηματισμός που περιγράφει τις πολικές του συντεταγμένες, δηλαδή κάθε σημείο  $(x, y)$  σταλεί

στο  $(\rho = \sqrt{x^2 + y^2}, \theta = \arctan2(x, y))$ , το μετασχηματισμένο δείγμα που προκύπτει είναι γραμμικά διαχωρίσιμο.

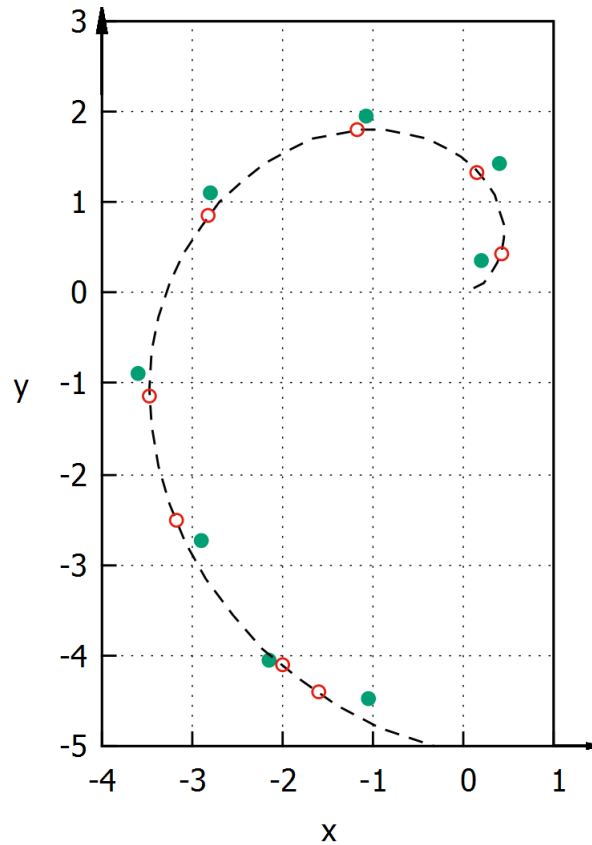


Σχήμα 3.3: Το δείγμα όπως απεικονίζεται στο χώρο  $\rho, \theta$ , όπου και εφαρμόζεται η γραμμική μέθοδος ΑσΚΣ.

Στον χώρο που ορίζουν οι διευθύνσεις  $\rho, \theta$  μπορεί να εφαρμοστεί η γραμμική μέθοδος ΑσΚΣ (σχήμα 3.3). Τα μετασχηματισμένα δεδομένα εμφανίζουν τη μέγιστη διακύμανση κατά μήκος της ευθείας που ορίζει το διάνυσμα  $PC1$ . Με εφαρμογή του αντίστροφου μετασχηματισμού  $x = \rho \cos(\theta), y = \rho \sin(\theta)$  η ευθεία αυτή μετασχηματίζεται στον αρχικό χώρο του δείγματος ως ένα μη-γραμμικό χαρακτηριστικό, το οποίο έχει τη μορφή σπείρας (σχήμα 3.4). Η δομή αυτή της σπείρας περιγράφει πολύ ικανοποιητικά το συγκεκριμένο δείγμα.

Βασική διαφορά του συγκεκριμένου παραδείγματος από τη μεθοδολογία που αναλύεται παρακάτω είναι στη διάσταση του μετασχηματισμού: εδώ χρησιμοποιείται ένας απλός μετασχηματισμός  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ , που επιλέχθηκε ώστε να ταιριάζει ειδικά στη μορφή του συγκεκριμένου δείγματος. Αντιθέτως, ο με-

τασχηματισμός που χρησιμοποιεί η ΑσΚΣ με συνάρτηση πυρήνα γίνεται στο χώρο  $\mathbb{R}^L$ , που έχει πρακτικά άπειρη διάσταση, επιτρέποντας έτσι τον εντοπισμό πολύ περισσότερων και πιο σύνθετων χαρακτηριστικών σε ένα δείγμα.



Σχήμα 3.4: Το μη-γραμμικό χαρακτηριστικό που εντοπίστηκε από το μετασχηματισμό.

Όπως και στην περίπτωση της απλής ΑσΚΣ, η διαδικασία ξεκινά με τον προσδιορισμό του  $N \times M$  συνόλου δεδομένων  $X$ . Στη συνέχεια, θεωρείται ότι εφαρμόζεται ο μη-γραμμικός μετασχηματισμός  $\phi : \mathbb{R}^N \rightarrow \mathbb{R}^L$  στα δεδομένα. Το μετασχηματισμένο δείγμα  $\phi(x_i)$  συγκεντρώνεται στον πίνακα  $Y$ , διάστασης  $L \times M$ .



Αυτό είναι το σετ δεδομένων στο οποίο θα γίνει ο υπολογισμός των κύριων κατευθύνσεων. Υποθέτοντας ότι τα δεδομένα του συνόλου  $\Upsilon$  έχουν μέσο όρο  $\vec{0}$ , το μητρώο συνδιακύμανσης ισούται με:

$$P = \frac{1}{M} Y Y^T \quad (3.6)$$

ή εκφράζοντας την παραπάνω σχέση για κάθε στοιχείο του  $P$ :

$$P_{ij} = P_{ji} = \frac{1}{M} \sum_{z=1}^M \phi_i(\vec{x}^z) \phi_j(\vec{x}^z) \quad i, j = 1 \dots L. \quad (3.7)$$

Το πρόβλημα ιδιοτιμών που καλείται να λύσει η μέθοδος είναι

$$P \vec{u} = \lambda \vec{u}, \vec{u} \in \mathbb{R}^L. \quad (3.8)$$

Όπως αναφέρθηκε, η διάσταση του χώρου στον οποίο απεικονίζεται το  $\phi$  εννοείται ως πολύ μεγαλύτερη αυτής του  $x$ , και ο υπολογισμός του καθώς και η επίλυση του ιδιοπρόβληματος, απαιτούν γενικά τεράστιο υπολογιστικό κόστος. Το εμπόδιο αυτό μπορεί να παρακαμφθεί εντελώς, κάνοντας χρήση μίας συνάρτησης πυρήνα (kernel function)  $k(\vec{x}_1, \vec{x}_2)$ ,  $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^N$ . Πιο αναλυτικά:

Από τη μορφή του πίνακα  $P$  αποδεικνύεται ότι υπάρχουν το πολύ  $M$  ιδιοδιανύσματα που αντιστοιχούν σε μη-μηδενική ιδιοτιμή [14]. Τα ιδιοδιανύσματα αυτά είναι που φέρουν πληροφορία για τα μη-γραμμικά χαρακτηριστικά του δείγματος  $X$ , και μπορούν να αναπαρασταθούν ως γραμμικός συνδυασμός μελών του  $\Upsilon$ .

Αυτό σημαίνει, ότι θα υπάρχουν  $a_r$  τέτοια ώστε  $\vec{u}^r = \sum_{\tau=1}^M a_\tau^r \vec{\phi}(x)_\tau$ ,  $r=1, \dots, M$ . Οι συντελεστές  $a_\tau^r$  συγκεντρώνονται στο  $M \times M$  μητρώο  $A$ . Επομένως, για κάθε τέτοιο ιδιοδιάνυσμα  $\vec{u}^r$ , η εξίσωση 3.8 διατυπώνεται ως :

$$\sum_{j=1}^L P_{ij} \sum_{\tau=1}^M a_\tau^r \phi_j(\vec{x}^\tau) = \lambda^r \sum_{\tau=1}^M a_\tau^r \phi_i(\vec{x}^\tau), i = 1, \dots, L \quad (3.9)$$

Η εξ. 3.9 αποτελεί ουσιαστικά σύστημα  $L$  γραμμικών εξισώσεων με αγνώστους τις τιμές των  $a$ . Υπολογίζοντας τα διαδοχικά εσωτερικά γινόμενα του ιδιοδιανύσματος  $\vec{u}^r$  με τα διανύσματα  $\phi(\vec{x}^q)$ ,  $q = 1, \dots, M$  του συνόλου  $\Upsilon$ , η εξ. 3.9 μετασχηματίζεται σε σύστημα  $M$  γραμμικών εξισώσεων:

$$\sum_{i=1}^L \sum_{j=1}^L \sum_{\tau=1}^M a_{\tau}^r P_{ij} \phi_i(\vec{x}^q) \phi_j(\vec{x}^{\tau}) = \lambda^r \sum_{i=1}^L \sum_{\tau=1}^M a_{\tau}^r \phi_i(\vec{x}^{\tau}) \phi_i(\vec{x}^q), \quad q = 1, \dots, M \quad (3.10)$$

Αντικαθιστώντας τη σχέση 3.7 στην εξίσωση και ορίζοντας

$$K_{ij} = k(\vec{x}^i, \vec{x}^j) = \vec{\phi}^i \cdot \vec{\phi}^j \quad (3.11)$$

η εξίσωση 3.10 καταλήγει:

$$\begin{aligned} \sum_{i=1}^L \sum_{j=1}^L \sum_{z=1}^M \sum_{\tau=1}^M a_{\tau}^r \phi_{\tau}^z \phi_j^z \phi_i^q \phi_j^{\tau} &= \lambda^r M \sum_{i=1}^L \sum_{\tau=1}^M a_{\tau}^r \phi_i^{\tau} \phi_i^q \implies \\ \sum_{\tau=1}^M \sum_{z=1}^M \underbrace{K_{qz} K_{z\tau}}_{=(KK)_{q\tau}} a_{\tau}^r &= \lambda^r M \sum_{\tau=1}^M a_{\tau}^r K_{\tau q} \implies \\ \sum_{\tau=1}^M (KK)_{q\tau} a_{\tau}^r &= \lambda^r M \sum_{\tau=1}^M K_{\tau q} a_{\tau}^r \end{aligned}$$

Εξαιρώντας μηδενικές ιδιοτιμές, μπορεί να γίνει διαγραφή του ενός  $K$ , καταλήγοντας έτσι στο σύστημα:

$$\sum_{\tau=1}^M K_{q\tau} a_{\tau}^r = \lambda^r M a_q^r \quad (3.12)$$

Η εξίσωση 3.12 αποτελεί το ισοδύναμο πρόβλημα ιδιοτιμών που επιλύει τελικά η μέθοδος [32, 14]. Το κέρδος που αποδίδουν αυτοί οι αλγεβρικοί χειρισμοί είναι ότι με αυτόν τον τρόπο αποφεύγεται ο ρητός υπολογισμός των  $\vec{\phi}$  και  $\vec{u}$ , που απαιτεί η εξίσωση 3.8. Μάλιστα, η μη-γραμμική απεικόνιση  $\vec{\phi}$  δεν χρειάζεται να προσδιοριστεί. Αρκεί να επιλεγεί μια συνάρτηση πυρήνα  $k(\vec{x}_1, \vec{x}_2)$ , η οποία υπολογίζει το εσωτερικό γινόμενο των εικόνων των  $\vec{x}_1$  και  $\vec{x}_2$ .

Η συνάρτηση πυρήνα μπορεί να ερμηνευθεί ως ένα μέτρο ομοιότητας στον χώρο των χαρακτηριστικών. Η διαδικασία βάσει της οποίας οι υπολογισμοί μεταφέρονται από το χώρο των  $\phi$  στο χώρο χαμηλότερης διάστασης  $M$  είναι γνωστή ως «τέχνασμα πυρήνα» (kernel trick).

Μερικές συναρτήσεις πυρήνα που μπορούν να χρησιμοποιηθούν είναι η συνάρτηση Laplace  $k(\vec{x}_1, \vec{x}_2) = \exp(-\frac{\|\vec{x}_1 - \vec{x}_2\|}{2\sigma^2})$ , η σιγμοειδής συνάρτηση όπου  $k(\vec{x}_1, \vec{x}_2) = \tanh(\alpha\vec{x}_1 \cdot \vec{x}_2 + c)$  ή πολυωνυμικές συναρτήσεις πυρήνα της μορφής  $k(\vec{x}_1, \vec{x}_2) = (\alpha\vec{x}_1 \cdot \vec{x}_2 + c)^d$  [14]. Εάν ως συνάρτηση πυρήνα ληφθεί το εσωτερικό γινόμενο στο χώρο των  $N$  διαστάσεων  $k(\vec{x}_1, \vec{x}_2) = \vec{x}_1 \cdot \vec{x}_2$ , τότε τα αποτελέσματα της ανάλυσης θα ισοδυναμούν με αυτά της απλής, γραμμικής ΑσΚΣ[30].

Η πιο δημοφιλής συνάρτηση πυρήνα, που χρησιμοποιείται και στο πλαίσιο αυτής της εργασίας είναι η συνάρτηση Gauss:

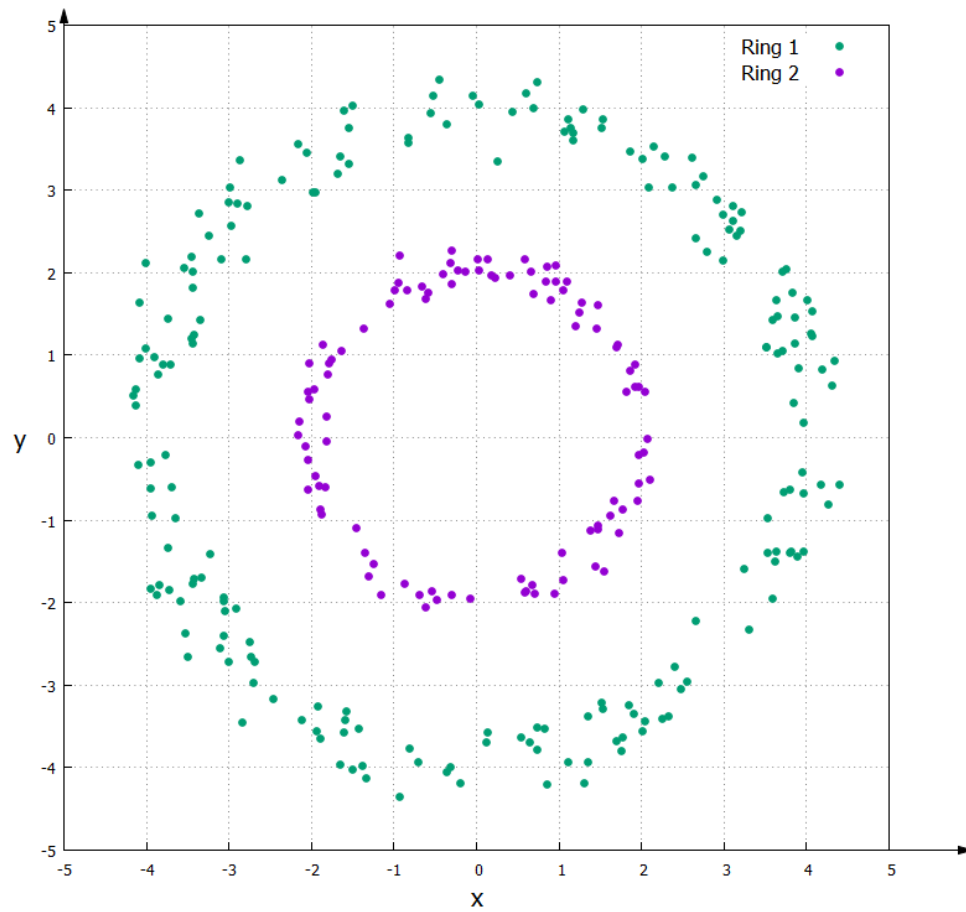
$$k(\vec{x}_1, \vec{x}_2) = \exp(-\frac{\|\vec{x}_1 - \vec{x}_2\|^2}{2\sigma^2}) \quad (3.13)$$

Η τιμή της παραμέτρου  $\sigma$ , δηλαδή της ακτίνας συνάρτησης ακτινικής βάσης έχει μεγάλη σημασία για την ποιότητα του αποτελέσματος: πολύ μεγάλο  $\sigma$  μπορεί να προκαλέσει εξίσωση ανόμοιων χαρακτηριστικών, ενώ με χρήση πολύ μικρού  $\sigma$  υπάρχει ο κίνδυνος της “υπερεξειδίκευσης”, δηλαδή της αδυναμίας γενίκευσης των ευρεθέντων χαρακτηριστικών από το συγκεκριμένο δείγμα στο χώρο σχεδιασμού[21].

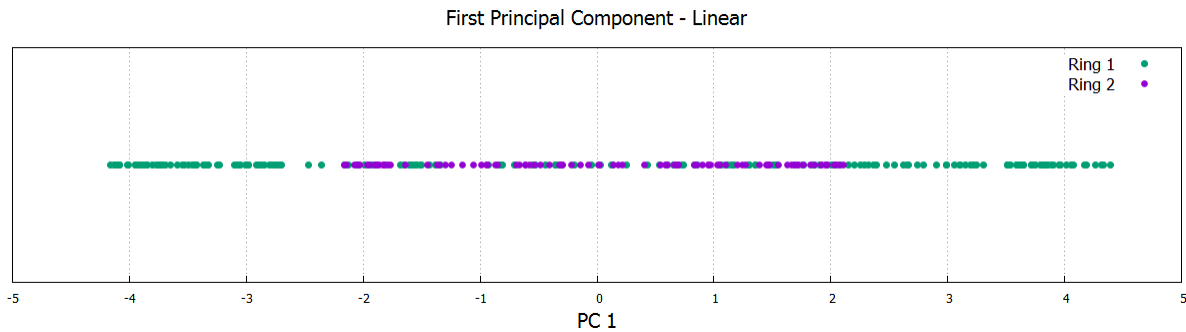
Για την εύρεση των  $M$  κύριων συνιστωσών της εικόνας μιάς υποψήφιας λύσης  $\vec{\phi}(x^\beta)$ , δεν μπορεί να γίνει χρήση μιας σχέσης όπως η 3.3, εφόσον δεν είναι γνωστός ο πίνακας  $U$ . Αντ’ αυτού αξιοποιούνται και πάλι οι ιδιότητες της συνάρτησης πυρήνα. Έτσι, οι κύριες συνιστώσες  $c_m^\beta$  γράφονται [32]:

$$c_m^\beta = \bar{u}^m \vec{\phi}(x^\beta) = \sum_{j=1}^M a_j^m \underbrace{\phi(\vec{x}^\beta)\phi(\vec{x}^j)}_{=K_{j\beta}} = \sum_{j=1}^M a_j^m K_{j\beta} \quad (3.14)$$

Τα σχήματα που ακολουθούν προσφέρουν ένα παράδειγμα του πώς η μέθοδος Kernel PCA μπορεί να εκφράσει αποδοτικότερα την πληροφορία σε ένα δείγμα. Στο σχήμα 3.5 απεικονίζεται ένα δείγμα δύο κύκλων, αποτελούμενο από 200 σημεία. Είναι προφανές, ότι ο γραμμικός μετασχηματισμός που ορίζει η απλή ΑσΚΣ δεν αλλάζει πολλά στην περιγραφή του δείγματος: στο νέο σύστημα συντεταγμένων που προκύπτει (το οποίο είναι πρακτικά ίδιο με τους άξονες  $xOy$  και για αυτό δεν απεικονίζεται), οι κύριοι άξονες δε συγκεντρώνουν καλύτερα τη διακύμανση. Εάν δε το δείγμα παρασταθεί χρησιμοποιώντας μόνο τη σημαντικότερη συνιστώσα, (σχήμα 3.6) τότε για τα περισσότερα σημεία δεν είναι δυνατό να καθοριστεί εάν ανήκουν στον κύκλο 1 ή 2.

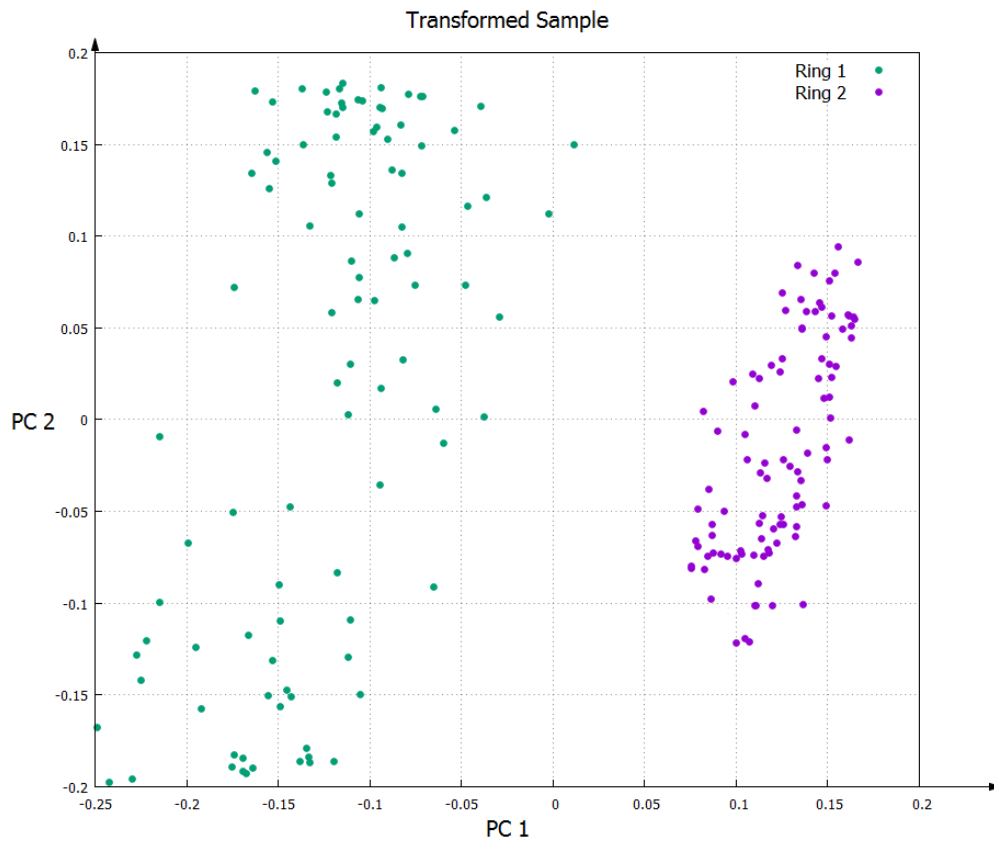


Σχήμα 3.5: Δείγμα 200 σημείων. Τα 100 από αυτά τα σημεία βρίσκονται πάνω σε έναν κύκλο ακτίνας 2 (με μικρές, τυχαίες διακυμάνσεις), ενώ τα υπόλοιπα 100 σε έναν κύκλο ακτίνας 4.

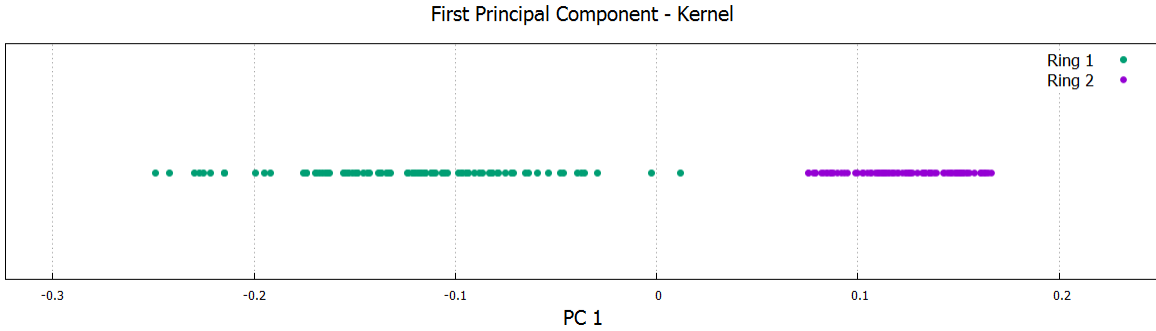


Σχήμα 3.6: Απεικόνιση της σημαντικότερης κύριας συνιστώσας των σημείων του δείγματος, όπως αυτή προέκυψε από γραμμική ΑσΚΣ. Για τα περισσότερα σημεία δεν είναι ξεκάθαρο εάν ανήκουν στον εσωτερικό ή τον εξωτερικό κύκλο.

Αντιθέτως, εάν στο δείγμα εφαρμοστεί η μέθοδος Kernel PCA (σχήμα 3.7), μπορεί να γίνει εντοπισμός των δύο βασικών χαρακτηριστικών του, δηλαδή των δύο κύκλων. Απεικονίζοντας τα δεδομένα χρησιμοποιώντας μόνο τη σημαντικότερη κύρια συνιστώσα  $PC1$ , ο διαχωρισμός ανάμεσα στους δύο κύκλους είναι σαφής, όπως φαίνεται στο σχήμα 3.8.



Σχήμα 3.7: Απεικόνιση του δείγματος ως προς τις δύο σημαντικότερες συνιστώσες του, όπως αυτές υπολογίστηκαν με τη μέθοδο Kernel PCA. Ως συνάρτηση πυρήνα χρησιμοποιείται η συνάρτηση Gauss , με  $\sigma^2 = 0.1$  .



Σχήμα 3.8: Απεικόνιση των δεδομένων χρησιμοποιώντας μόνον τη σημαντικότερη συνιστώσα. Η συνιστώσα περιέχει αρκετή πληροφορία ώστε να είναι δυνατός ο διαχωρισμός ανάμεσα στους δύο κύκλους.

Τα αποτελέσματα της μεθόδου Kernel PCA μπορούν να χρησιμοποιηθούν με όμοιο τρόπο στην εφαρμογή των εξελικτικών τελεστών και στην εκπαίδευση των μεταπροτύπων. Έτσι προκύπτουν αντίστοιχα οι εκδοχές EA(K) και M(K)ΑΕΑ.

### 3.3.1 Η εκδοχή EA(K)

Κατά την εκδοχή EA(K) οι κύριες συνιστώσες των υποψήφιων γονέων στο χώρο των χαρακτηριστικών υπολογίζονται από τη σχέση 3.14. Στη συνέχεια, εφαρμόζονται κατά τα γνωστά οι εξελικτικοί τελεστές για να δημιουργήσουν τον (μετασχηματισμένο και στραμμένο) απόγονο  $\vec{c}^* \in \mathbb{R}^M$ .

Σε αντίθεση με την περίπτωση της γραμμικής ανάλυσης σε κύριες συνιστώσες, ο προσδιορισμός του  $\vec{x}^* \in \mathbb{R}^N$ , δηλαδή του απογόνου στο χώρο σχεδιασμού, απαιτεί τη λύση μη-γραμμικών εξισώσεων, και είναι ακριβότερος υπολογιστικά (το κόστος του ωστόσο παραμένει ασήμαντο σε σχέση με αυτό που απαιτεί συνήθως μια κλήση του ακριβούς λογισμικού αξιολόγησης). Συγκεκριμένα, ανάζητάται κάποιο  $\vec{x}^*$  τέτοιο ώστε  $c_m^* = \vec{u}^m \vec{\phi}(\vec{x}^*)$  ή ισοδύναμα

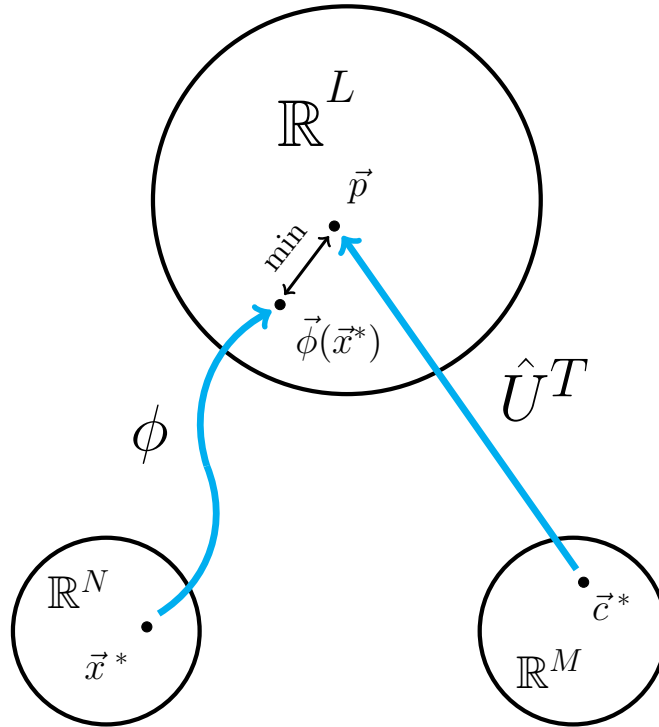
$$\vec{c}^* = \hat{U} \vec{\phi}(\vec{x}^*) \quad (3.15)$$

όπου  $\hat{U}$  ο  $M \times L$  πίνακας των ιδιοδιανυσμάτων που αντιστοιχούν σε μη-μηδενική ιδιοτιμή. Εφόσον ο  $\hat{U}$  περιέχει τα διανύσματα μιας ορθοκανονικής βάσης του χώρου χαρακτηριστικών, έπεται ότι ο ανάστροφος του είναι και ο

αντίστροφος. Συνεπώς, πολλαπλασιάζοντας με  $\hat{U}^T$  την εξ. 3.15:

$$\hat{U}^T c_m^* = \vec{\phi}(\vec{x}^*) \implies \sum_{m=1}^M c_m^* \vec{u}^m = \phi(\vec{x}^*) \implies \underbrace{\sum_{m=1}^M c_m^* \sum_{\tau=1}^M a_\tau^m \phi(x)_\tau}_{\vec{p}} = \phi(\vec{x}^*)$$

Το διάνυσμα  $\vec{p}$  είναι η προβολή του  $\vec{c}^*$  στον χώρο των  $\phi$ , όπως αυτή ορίζεται από τα (έμμεσα) υπολογισθέντα ιδιοδιανύσματα  $\hat{U}$ . Αρκεί, λοιπόν, να βρεθεί το  $\vec{x}^*$  του οποίου η απεικόνιση  $\phi(\vec{x}^*)$  είναι όσο το δυνατόν πιο κοντά στο  $\vec{p}$  [33]. Η διαδικασία ανεύρεσης λύσης, κατά την οποία οι χώροι  $\mathbb{R}^N$  και  $\mathbb{R}^M$  επικοινωνούν μέσω του χώρου των χαρακτηριστικών απεικονίζεται στο σχήμα 3.3.1.



Σχήμα 3.9: Σχηματική αναπαράσταση επίλυσης του προβλήματος εύρεσης του  $\vec{x}^*$ .

Η λύση της εξ. 3.15 αντιστοιχεί τελικά στην ελαχιστοποίηση της

$$[\vec{p} - \vec{\phi}(\vec{x}^*)]^2 = \vec{p} \cdot \vec{p} + \vec{\phi}(\vec{x}^*) \cdot \vec{\phi}(\vec{x}^*) - 2\vec{p} \cdot \vec{\phi}(\vec{x}^*) \quad (3.16)$$



Ο όρος  $\vec{p} \cdot \vec{p}$  είναι σταθερός και ανεξάρτητος του  $\vec{x}^*$ , ενώ για τη συνάρτηση πυρήνα που χρησιμοποιείται ισχύει  $\phi(\vec{x}^*) \cdot \phi(\vec{x}^*) = 1$ . Παραγωγίζοντας την παραπάνω σχέση:

$$\begin{aligned} \frac{\delta}{\delta \vec{x}^*} [-2\vec{p} \cdot \phi(\vec{x}^*)] &= \frac{\delta}{\delta \vec{x}^*} [-2 \sum_{m=1}^M c_m^* \sum_{\tau=1}^M a_\tau^m \underbrace{\phi(x)_\tau \phi(\vec{x}^*)}_{k(\vec{x}^*, \vec{x}^\tau)}] = \\ -2 \sum_{m=1}^M \sum_{\tau=1}^M c_m^* a_\tau^m \exp(-\frac{\|\vec{x}^\tau - \vec{x}^*\|^2}{2\sigma^2}) &(-\frac{1}{2\sigma^2}) 2(\vec{x}^\tau - \vec{x}^*)(-1) \end{aligned} \quad (3.17)$$

Απαιτώντας το μηδενισμό της παραγώγου προκύπτει η εξίσωση

$$\sum_{m=1}^M \sum_{\tau=1}^M c_m^* a_\tau^m \exp(-\frac{\|\vec{x}^\tau - \vec{x}^*\|^2}{2\sigma^2}) \vec{x}^\tau = \sum_{m=1}^M \sum_{\tau=1}^M c_m^* a_\tau^m \exp(-\frac{\|\vec{x}^\tau - \vec{x}^*\|^2}{2\sigma^2}) \vec{x}^*$$

η οποία μπορεί να επιλυθεί από τον αλγόριθμο σταθερού σημείου[33]:

$$\vec{x}^{*new} = \frac{\sum_{m=1}^M \sum_{\tau=1}^M c_m^* a_\tau^m \exp(-\frac{\|\vec{x}^\tau - \vec{x}^{*old}\|^2}{2\sigma^2}) \vec{x}^\tau}{\sum_{m=1}^M \sum_{\tau=1}^M c_m^* a_\tau^m \exp(-\frac{\|\vec{x}^\tau - \vec{x}^{*old}\|^2}{2\sigma^2})} \quad (3.18)$$

### 3.3.2 Η εκδοχή M(K)ΑΕΑ

Όπως και με τη γραμμική περίπτωση, τα αποτελέσματα της ΑΣκΣ μέσω συνάρτησης πυρήνα μπορούν να επεκταθούν και στα μεταπρότυπα. Αμέσως μετά τη διαδικασία επιλογής δειγμάτων εκπαίδευσης, πραγματοποιείται ο μετασχηματισμός τους στο χώρο διάστασης M. Η διαφορά εδώ είναι ότι υπάρχει, ενδεχομένως, η δυνατότητα εκπαίδευσης μεταπρωτύπων διάστασης μεγαλύτερης από αυτήν του χώρου σχεδιασμού. Στη συνέχεια, πραγματοποιείται η αποκοπή νευρώνων και το ΔΣΑΒ εκπαιδεύεται με βάση τις σημαντικότερες διευθύνσεις, όπως αυτές εκτιμήθηκαν.

## Κεφάλαιο 4

# Επίλυση Προβλημάτων Βελτιστοποίησης

### 4.1 Εισαγωγή

Για τη διαπίστωση της αποτελεσματικότητας των μεθόδων linear PCA και Kernel PCA, όπως αναπτύχθηκαν παραπάνω, πραγματοποιήθηκε μελέτη έξι προβλημάτων βελτιστοποίησης. Συγκεκριμένα έγινε μελέτη παραλλαγών των μαθηματικών συναρτήσεων Ackley και Rastrigin, το μαθηματικό πρόβλημα δύο στόχων Fonseca-Fleming, καθώς και δύο προβλημάτων “ψευδο-μηχανολογικού” σχεδιασμού (pseudo-engineering problems), ένα που αφορά στο σχεδιασμό ενός υποστηρίγματος αποτελούμενου από δύο δοκούς, και ένα που αφορά σε σχεδιασμό προδιαγραφών μιας συγκόλλησης. Επίσης, πραγματοποιήθηκε βελτιστοποίηση αεροτομής αποτελούμενης από τέσσερα στοιχεία (Four-Element Airfoil).

Οι συναρτήσεις Ackley και Rastrigin έχουν αυξημένη πολυπλοκότητα σε σχέση με τα συνήθη προβλήματα που συναντώνται, και χρησιμοποιούνται γενικά ως δοκιμασίες (test functions ή benchmark functions) για την ανάδειξη της αξιοπιστίας ή των αδυναμιών μιας μεθόδου βελτιστοποίησης. Χαρακτηρίζονται από πολλά τοπικά ακρότατα (πολυτροπική συμπεριφορά), ενώ τροποποιήθηκαν ώστε να παρουσιάζουν λιγότερες συμμετρίες και μικρότερη διαχωρισιμότητα[37]. Το μαθηματικό πρόβλημα των Fonseca και Fleming είναι ένα απλό παράδειγμα που δείχνει την αξία των προτεινόμενων μεθόδων και για περιπτώσεις πολυκριτηριακής βελτιστοποίησης.

Τα “ψευδο-μηχανολογικά” προβλήματα είναι χρήσιμα εκπαιδευτικά παρα-

δείγματα για την εισαγωγή στη βελτιστοποίηση μηχανολογικών εφαρμογών. Περιλαμβάνουν σχετικά μικρό πλήθος μεταβλητών, ενώ χρησιμοποιούν απλά φυσικά μοντέλα για τη διατύπωση της συνάρτησης στόχου και των περιορισμών.

Τέλος, η περίπτωση αεροτομής τεσσάρων στοιχείων αποτελεί ένα πρακτικό παράδειγμα αεροδυναμικής βελτιστοποίησης. Μια βασική διαφορά του με τα προηγούμενα παραδείγματα είναι η πολύ ακριβότερη (υπολογιστικά και χρονικά) διαδικασία αξιολόγησης μιας υποψήφιας λύσης, που περιλαμβάνει επίλυση των εξισώσεων ροής. Παρ' ότι η μοντελοποίηση είναι διδιάστατη, τα αποτελέσματα που παράγονται έχουν εφαρμογή και ανταποκρίνονται σε ένα πραγματικό πρόβλημα αεροδυναμικής.

Σε όλα τα προβλήματα που μελετώνται, επιλέχθηκαν πρώτα οι βασικές παράμετροι του ΕΑ (πλήθος γονέων  $\mu$ , πλήθος απογόνων  $\lambda$ , αριθμός γονέων που σχηματίζουν ένα απόγονο  $\rho$ , σχήμα διασταύρωσης, πιθανότητα μετάλλαξης, πλήθος δειγμάτων εκπαίδευσης για το μεταπρότυπο, πλήθος εγγραφών στη βάση δεδομένων που απαιτούνται για την εκκίνηση του μεταπρότυπου). Στην περίπτωση εφαρμογής ΑσκΣ στους εξελικτικούς τελεστές απαραίτητος είναι ο καθορισμός της γενιάς κατά την οποία αυτή ξεκινά να εφαρμόζεται, ενώ όταν η ΑσκΣ εφαρμόζεται στα μεταπρότυπα πρέπει να καθοριστεί και ο αριθμός των διαστάσεων με τον οποίο γίνεται η εκπαίδευσή τους.

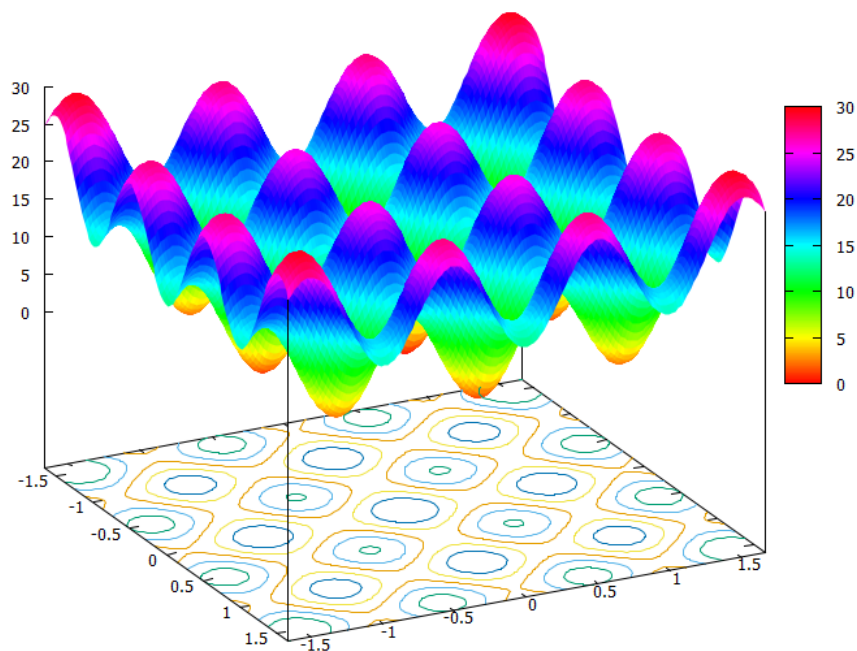
Στη συνέχεια, πραγματοποιήθηκαν τρεξίματα με διαφορετικές τιμές της γεννήτριας τυχαίων αριθμών για τις διάφορες εκδοχές. Τα αποτελέσματα όλων των δοκιμών συγκεντρώθηκαν σε διαγράμματα, ενώ κατασκευάστηκε για κάθε εκδοχή και η μέση πορεία της σύγκλισης. Με βάση αυτά, γίνεται σχολιασμός και σύγκριση μεταξύ των διάφορων εκδοχών.

## 4.2 Συνάρτηση Rastrigin & Παραλλαγές της

Στη βασική της μορφή, η συνάρτηση Rastrigin δίνεται από τη σχέση [36]:

$$f(\vec{x}) = 10N + \sum_{i=1}^N [x_i^2 - 10\cos(2\pi x_i)] , \quad x_i \in [-5.12, 5.12] \quad (4.1)$$

Η εξ. 4.1 είναι διαχωρίσιμη, αφού μπορεί να παρασταθεί ως άθροισμα  $N$  συναρτήσεων, καθεμιά από τις οποίες εμπλέκει μόνο μία μεταβλητή βελτιστοποίησης. Στο σχήμα 4.1 απεικονίζεται η διδιάστατη εκδοχή αυτής της συνάρτησης.



Σχήμα 4.1: Η συνάρτηση Rastrigin για  $N=2$ .

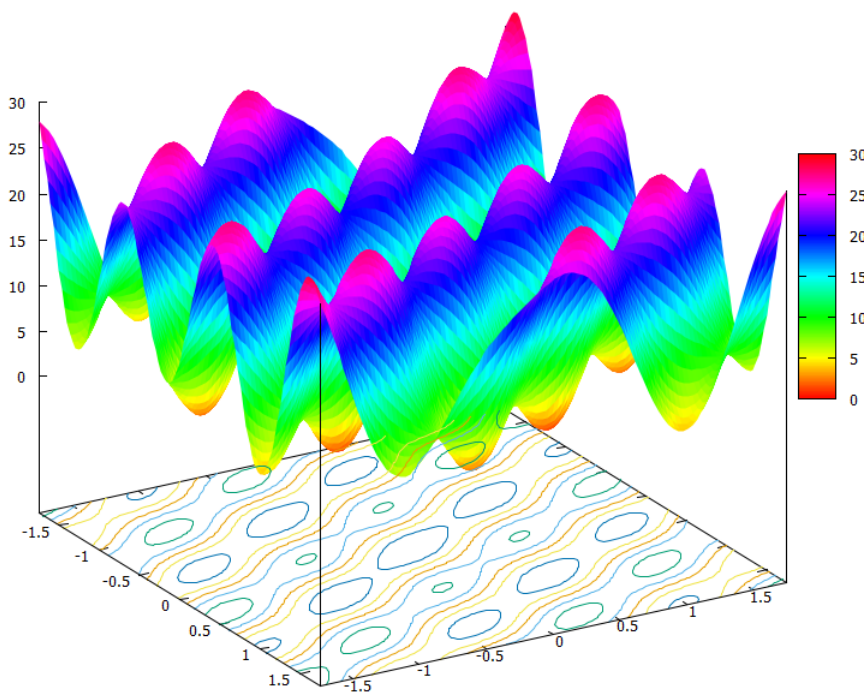
Έχει διαπιστωθεί ότι η απόδοση ενός τυπικού ΕΑ στο συγκεκριμένο πρόβλημα εξαρτάται σε μεγάλο βαθμό από την επιλογή των μεταβλητών βελτιστοποίησης.[31]. Μη-διαχωρίσιμες εκδοχές του προβλήματος παρουσιάζουν σημαντική επιβράδυνση στην πορεία σύγκλισης.

Για το λόγο αυτό, και για τη διερεύνηση της επίδρασης που μπορεί να έχουν οι προτεινόμενες μέθοδοι ΑσκΣ, δημιουργήθηκαν και μελετήθηκαν τροποποιημένες εκδοχές της συνάρτησης, οι οποίες χαρακτηρίζονται από μη-διαχωρισιμότητα ή και ανισοτροπία της συνάρτησης στόχου. Τα δύο αυτά χαρακτηριστικά συναντώνται συχνά σε προβλήματα μεγάλης κλίμακας της βιομηχανίας, και είναι εκείνα που καθιστούν ιδιαίτερα δύσκολη την επίλυσή τους από τους συμβατικούς ΕΑ και ΜΑΕΑ.

Τα προβλήματα που μελετήθηκαν είναι της μορφής:

$$f(\vec{x}) = 10\left(\sum_{i=1}^N w_i\right) + \sum_{i=1}^N w_i[\beta_i^2(\vec{x}) - 10\cos(2\pi\beta_i(\vec{x}))], \quad x_i \in [-5.12, 5.12] \quad (4.2)$$

όπου  $\vec{\beta} = B(\vec{x} - \vec{\sigma})$ . Το  $B$  είναι ένα μητρώο στροφής, το οποίο εμπλέκει τις μεταβλητές βελτιστοποίησης μεταξύ τους, καθιστώντας έτσι τη συνάρτηση μη-διαχωρίσιμη, ενώ  $\vec{w}$  ένα διάνυσμα βάρους που εισάγει ανισοτροπία, κάνοντας έτσι το πρόβλημα ακόμα πιο δύσκολο να λυθεί από τον ΕΑ. Το ολικό ακρότατο βρίσκεται στη θέση  $\vec{\sigma}$ , στην οποία η συνάρτηση παίρνει την τιμή 0. Η μορφή που μπορεί να έχει μια 2Δ εκδοχή αυτής της παραλλαγής δίνεται στο σχήμα 4.2.



Σχήμα 4.2: Παράδειγμα μιας διδιάστατης στραμμένης (κατά  $30^\circ$ ) και ανισότροπης συνάρτησης Rastrigin, με  $\vec{w} = [1, 4]$

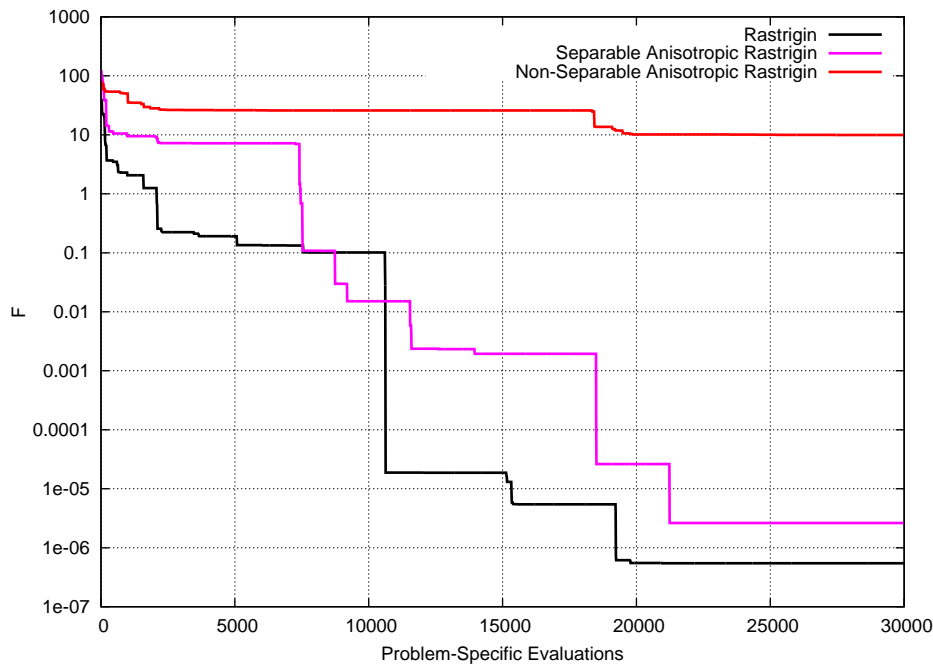
Οι μελέτες που ακολουθούν έγιναν για δύο περιπτώσεις, με  $N=5$  και  $N=20$  διαστάσεις.

## Περίπτωση $N=5$

Σε όλες τις δοκιμές που παρουσιάζονται σε αυτήν την υποενότητα, κάθε γενιά του ΕΑ αποτελείται από  $\mu=8$  γονείς και  $\lambda=16$  απογόνους, ενώ χρησιμοποιείται το σχήμα διασταύρωσης γενικευμένου ανασυνδυασμού. Δοκιμάστηκαν και άλλοι συνδυασμοί των  $\mu, \lambda$ , οι οποίοι δεν παρουσιάζονται, καθώς η επιλογή των συγκεκριμένων παραμέτρων είναι αυτή που έδωσε συνολικά τα καλύτερα αποτελέσματα.

Το βασικό πρόβλημα που μελετάται έχει τη μορφή της εξίσωσης 4.2: Στο διάνυσμα  $\vec{\sigma}$  δόθηκαν οι τιμές  $[1, 2, 3, 4, 0]$ . Πραγματοποιούνται 3 στροφές  $45^\circ$  στα επίπεδα που ορίζονται από τους άξονες  $x_1 - x_2$ ,  $x_2 - x_3$ , και  $x_3 - x_4$ , ενώ το διάνυσμα ανισοτροπίας έχει την τιμή  $\vec{w} = [1, 1, 1, 8, 1]$ .

Με τις τροποποιήσεις αυτές, η λειτουργία του ΕΑ δυσχεραίνεται σε πολύ μεγάλο βαθμό. Αυτό φαίνεται στο σχήμα 4.3, όπου απεικονίζεται η επίδοση για μία δοκιμή του απλού ΕΑ σε τρία διαφορετικά προβλήματα: την απλή συνάρτηση Rastrigin, μια ανισότροπη αλλά διαχωρίσιμη παραλλαγή και τη βασική μη-διαχωρίσιμη και ανισότροπη παραλλαγή, όλες για την περίπτωση των πέντε διαστάσεων.



Σχήμα 4.3: Πορείες σύγκλισης του απλού EA για τη συνάρτηση Rastrigin πέντε διαστάσεων και δύο παραλλαγές της.

Στην περίπτωση του μη-διαχωρίσιμου προβλήματος είναι εμφανής η δυσκολία του αλγορίθμου να συγκλίνει στη βέλτιστη λύση, ακόμη και για πολύ μεγάλο αριθμό κλήσεων του λογισμικού αξιολόγησης. Αντίθετα, για τα δύο πρώτα προβλήματα, τα οποία είναι διαχωρίσιμα, η σύγκλιση σε ικανοποιητική λύση επιτυγχάνεται αρκετά γρήγορα.

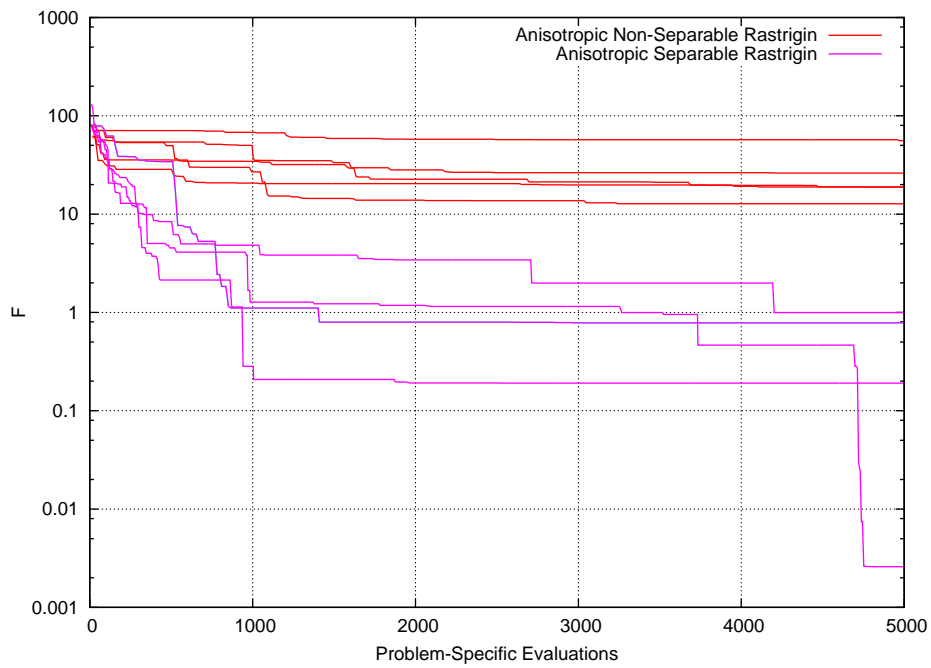
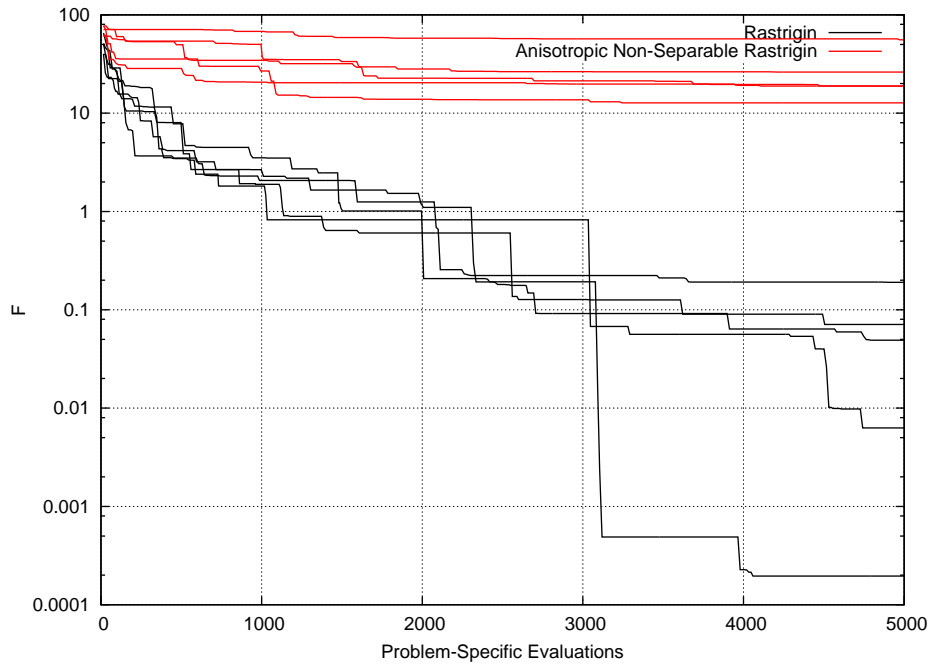
Η συμπεριφορά αυτή είναι σταθερή, όπως δείχνουν τα διαγράμματα του σχήματος 4.4, τα οποία απεικονίζουν πορείες σύγκλισης των προβλημάτων για περισσότερες τιμές της γεννήτριας τυχαίων αριθμών, με κοινό κριτήριο τερματισμού τις 5000 ακριβείς αξιολογήσεις.

Στο σχήμα 4.5, απεικονίζονται οι πορείες σύγκλισης των εκδοχών EA(L) και EA(K) για το μη-διαχωρίσιμο πρόβλημα και συγκριτικά με την αντίστοιχη πορεία του απλού EA. Στο σχήμα 4.6, απεικονίζεται η μέση πορεία της σύγκλισης για το σύνολο των δοκιμών: σε όλες σχεδόν τις περιπτώσεις, η εφαρμογή της ΑσΚΣ στους εξελικτικούς τελεστές, η οποία ξεκινά από την τρίτη γενιά, προκαλεί σημαντική βελτίωση της απόδοσης του αλγορίθμου. Είναι λοιπόν εμφανές το όφελος που μπορεί να αποκομίσει ο EA, εάν πρώτα φέρει το πρό-

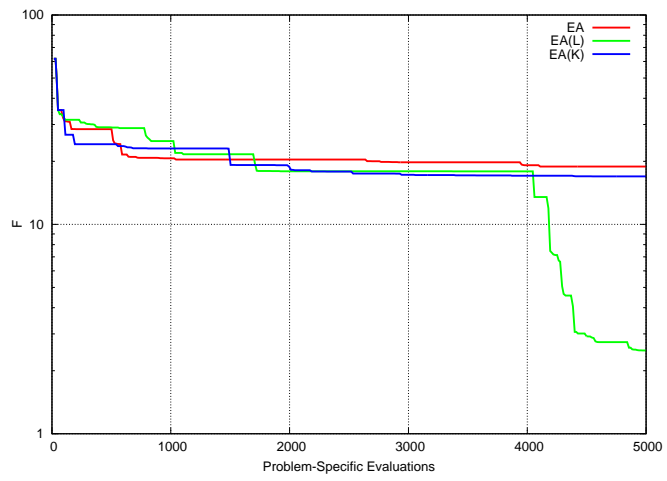
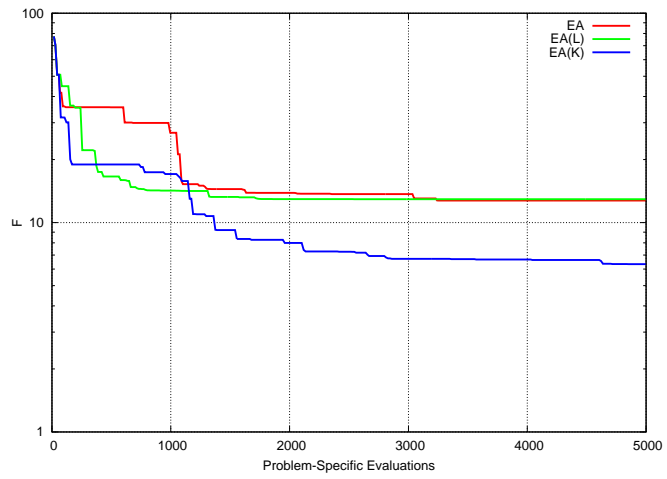
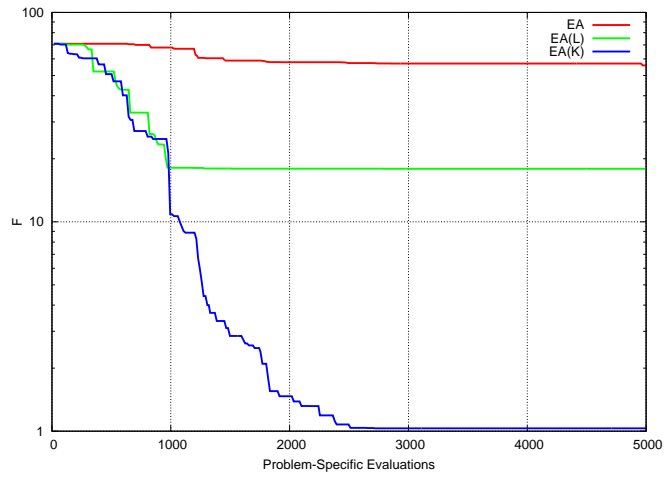
βλημα σε μια “περισσότερο” διαχωρίσιμη μορφή.

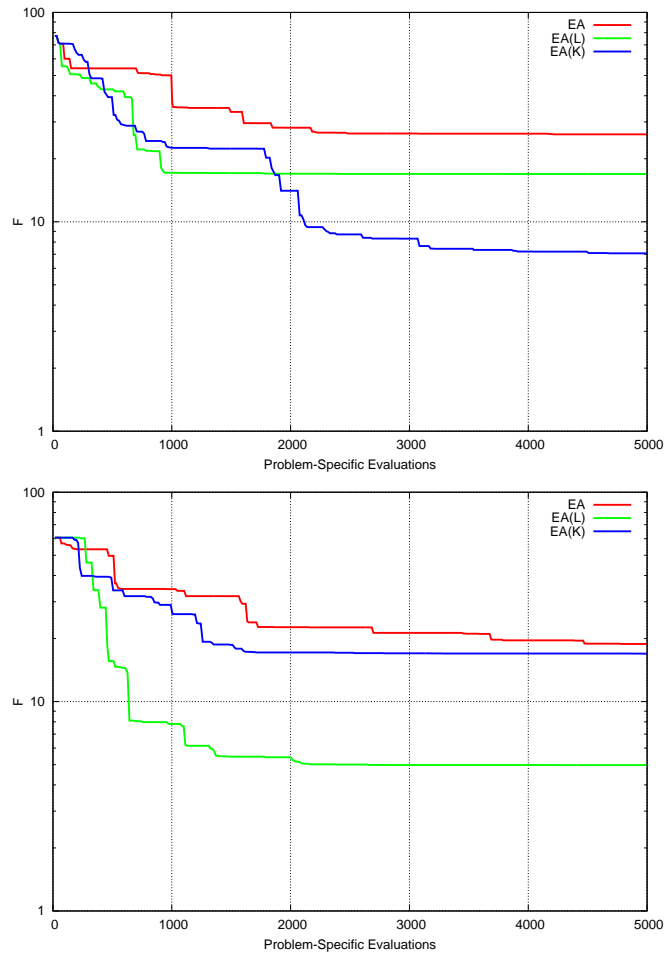
Από την άλλη, στην περίπτωση που το πρόβλημα είναι ήδη διαχωρίσιμο και δεν υπάρχει αλληλεπίδραση μεταξύ των μεταβλητών σχεδιασμού, δεν αναμένεται κάποια βελτίωση με εφαρμογή της ΑσΚΣ στους εξελικτικούς τελεστές. Αυτό επιβεβαιώνεται στα διαγράμματα του σχήματος 4.7, όπου απεικονίζονται οι πορείες σύγκλισης των τριών εκδοχών για δύο διαφορετικές αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών στη διαχωρίσιμη εκδοχή του προβλήματος.



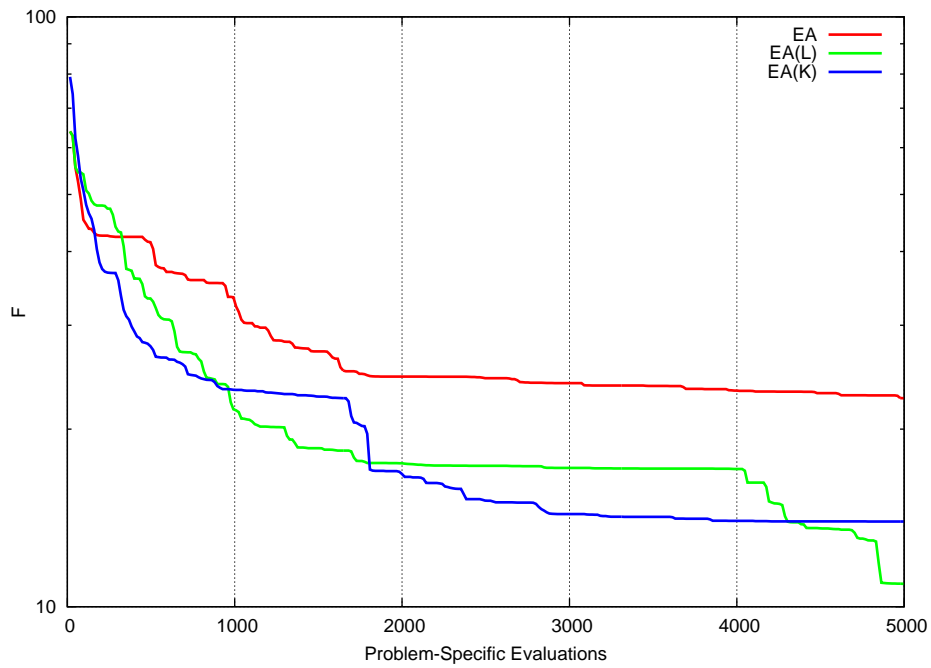


Σχήμα 4.4: 5Δ Συνάρτηση Rastrigin & Παραλλαγές της: Πορείες σύγκλισης για πέντε διαφορετικές τιμές της γεννήτριας τυχαίων αριθμών.

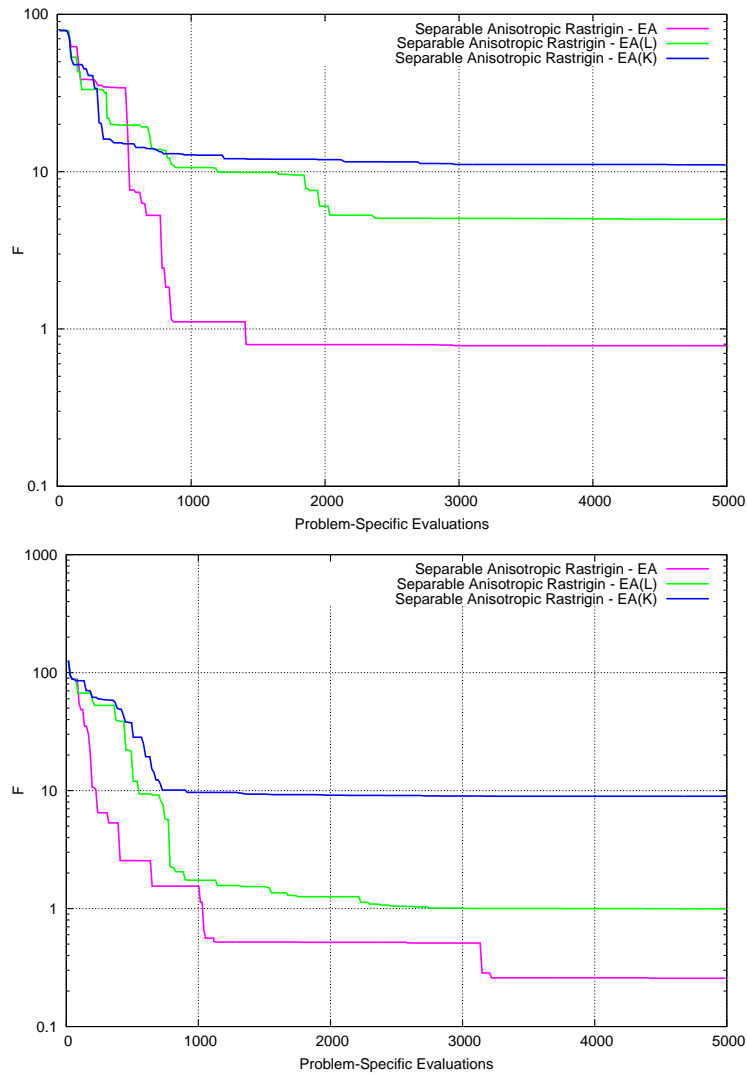




Σχήμα 4.5: Μη-Διαχωρίσιμη Παραλλαγή 5Δ Συνάρτησης Rastrigin: Πορείες σύγκλισης για τις μεθόδους EA(L) και EA(K), για πέντε αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.



Σχήμα 4.6: Μη-Διαχωρίσιμη Παραλλαγή 5Δ Συνάρτησης Rastrigin: Μέση πορεία σύγκλισης για τις μεθόδους EA, EA(L) και EA(K).



Σχήμα 4.7: 5Δ Διαχωρίσιμη Παραλλαγή Συνάρτησης Rastrigin: Πορείες σύγκλισης για τις μεθόδους EA(L) και EA(K), για δύο αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.

### Περίπτωση $N=20$ διαστάσεων

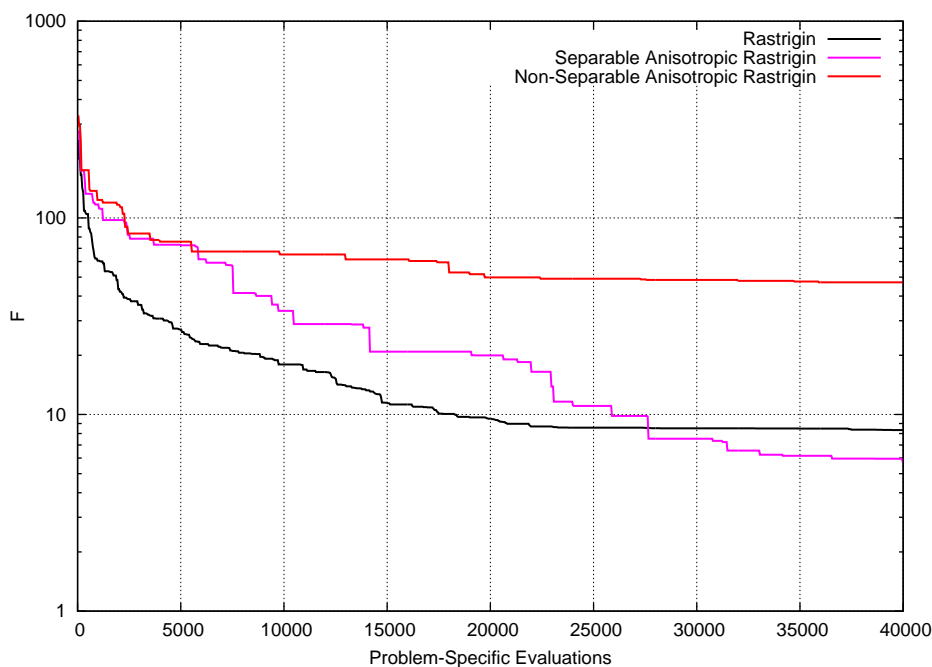
Το ολικό ακρότατο βρίσκεται στο σημείο  $\vec{0}$ . Στον 20-διάστατο χώρο πραγματοποιούνται τέσσερις στροφές  $45^\circ$  και μία  $30^\circ$ , οι οποίες εμπλέκουν τις πρώτες έξι μεταβλητές καθώς και την τελευταία. Το διάνυσμα βάρους δίνει

λαμβάνει τις τιμές  $\vec{w} = [1, 1, 1, \dots, 1, 20]$ . Οι μεταβλητές βελτιστοποίησης αυξήθηκαν από 5 σε 20, επομένως χρειάζεται να αυξηθούν και οι πληθυσμοί που χειρίζεται ο ΕΑ, προκειμένου να διατηρηθεί η αποτελεσματικότητά του. Και σε αυτήν την περίπτωση, έγιναν δοκιμές με διάφορες τιμές των παραμέτρων μ,λ. Παρουσιάζονται τα αποτελέσματα για τον (10,20)ΕΑ καθώς αυτός ο συνδυασμός έδωσε, σε γενικές γραμμές, τις καλύτερες πορείες σύγκλισης.

Κωδικοποίηση	Σχήμα Διασταύρωσης	Πιθανότητα Μετάλλαξης $P_m$
Πραγματική	Intermediate Recombination	0.05

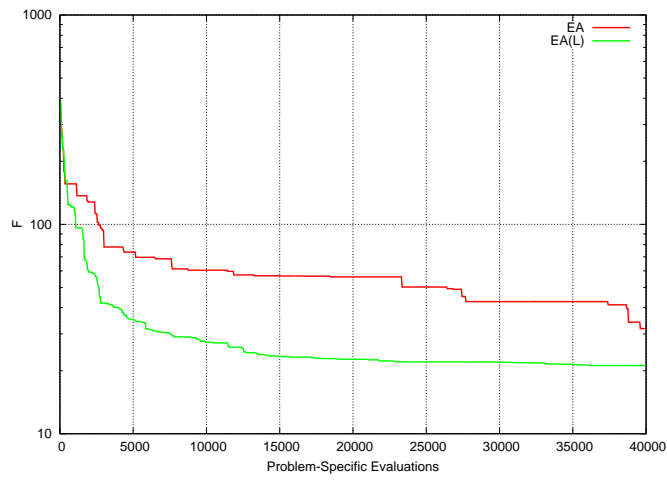
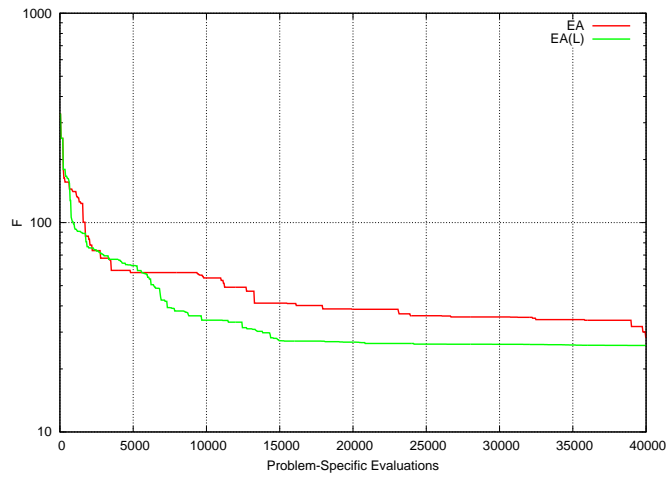
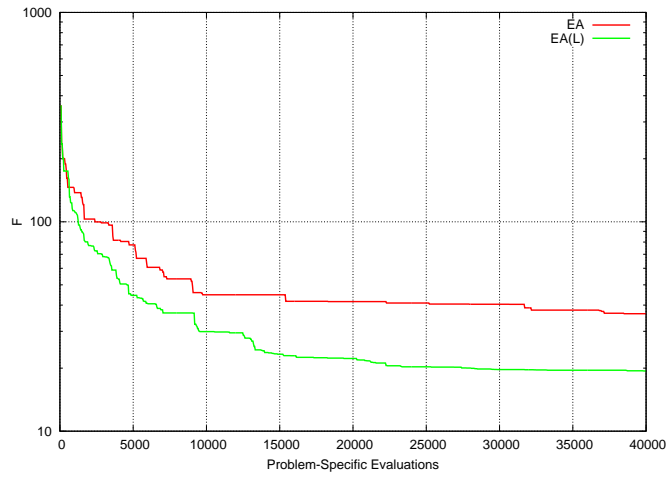
Πίνακας 4.1: Επιμέρους χαρακτηριστικά του τρεξίματος.

Συγκρίνοντας την επίδοση του συγκεκριμένου προβλήματος με το αντίστοιχο διαχωρίσιμο, καθώς και την απλή συνάρτηση Rastrigin 20 διαστάσεων (σχήμα 4.8), παρατηρείται η ίδια δυσκολία του ΕΑ, η οποία οφείλεται στην αλληλεπίδραση μεταξύ των μεταβλητών.

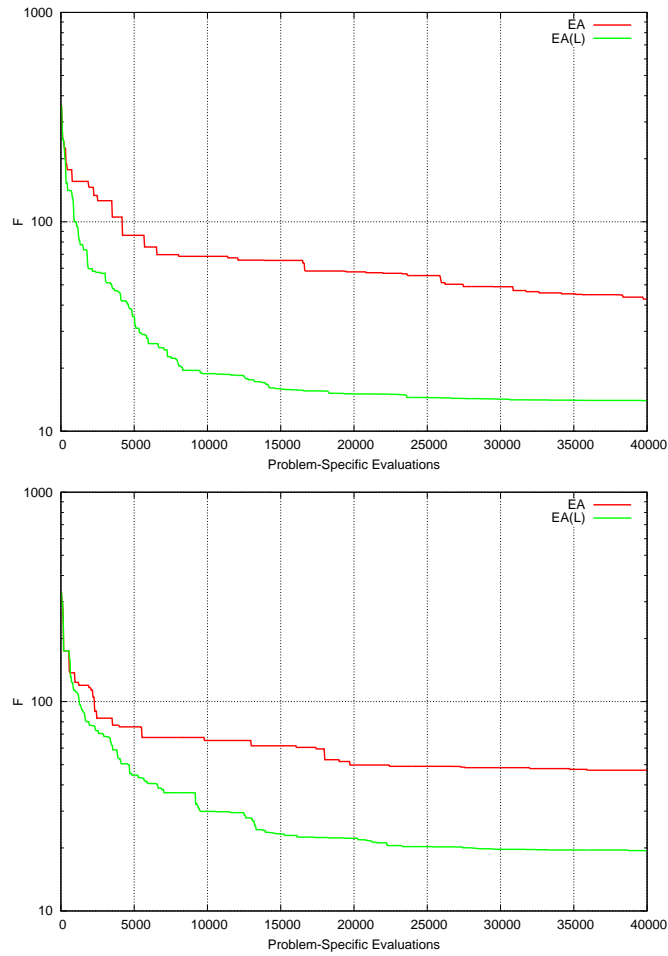


Σχήμα 4.8: 20Δ Συνάρτηση Rastrigin & παραλλαγές της. Πορείες σύγκλισης του απλού ΕΑ.

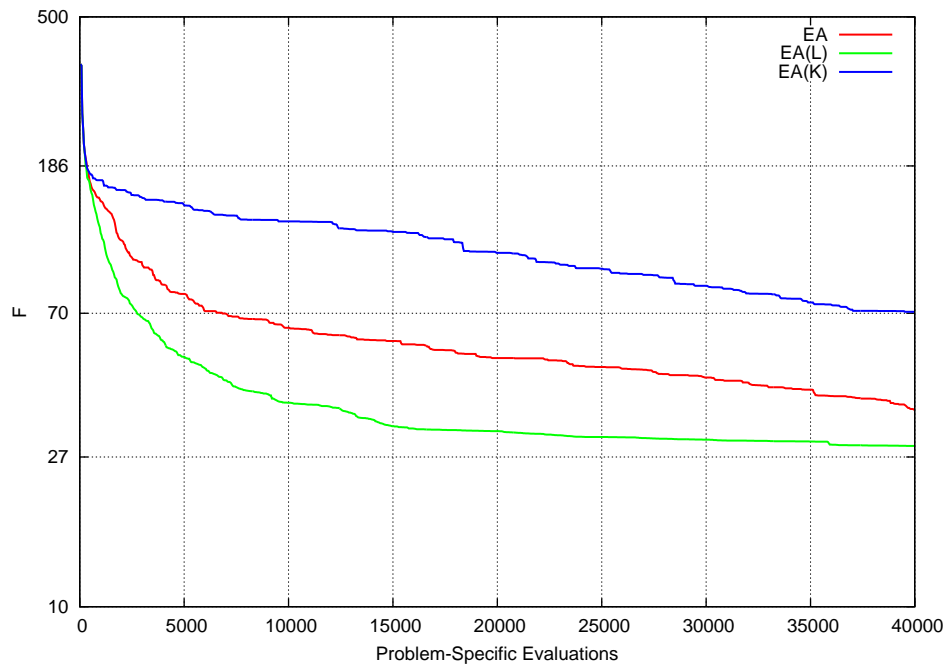
Στο σχήμα 4.9 απεικονίζονται, για τις μεθόδους EA και EA(L), οι πορείες σύγκλισης για πέντε διαφορετικές αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών. Στο σχήμα 4.10 απεικονίζεται η μέση πορεία της σύγκλισης. Σε αυτήν την περίπτωση, η εφαρμογή της γραμμικής ΑσκΣ στους εξελικτικούς τελεστές βελτιώνει την επίδοση του EA. Αντίθετα, η μέθοδος EA(K) δεν βελτιώνει καθόλου τη σύγκλιση (απεικονίζεται μόνο η μέση πορεία της σύγκλισης στο σχήμα 4.10).







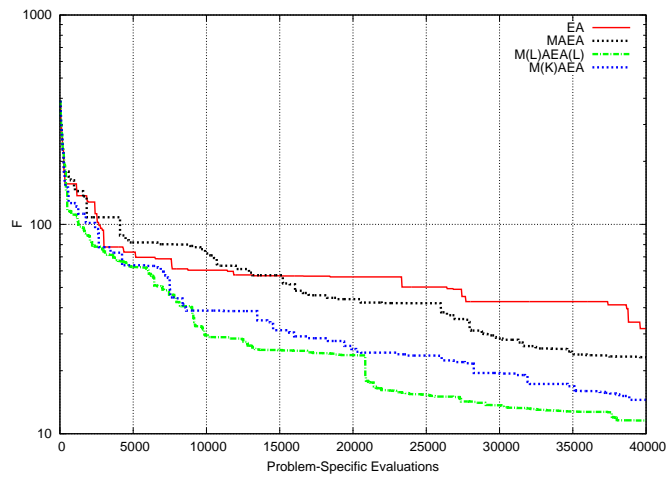
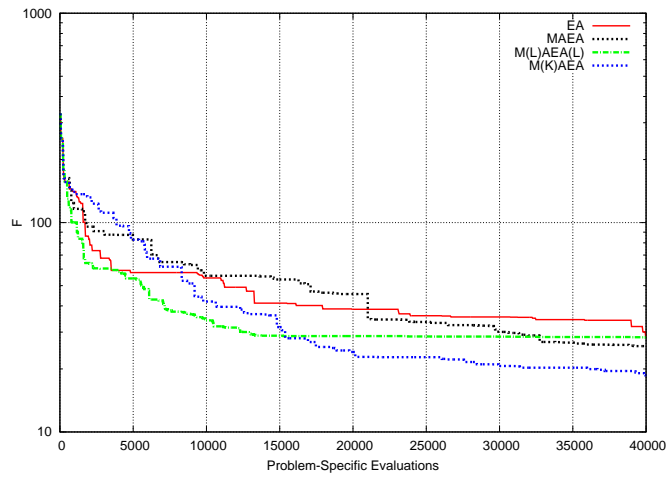
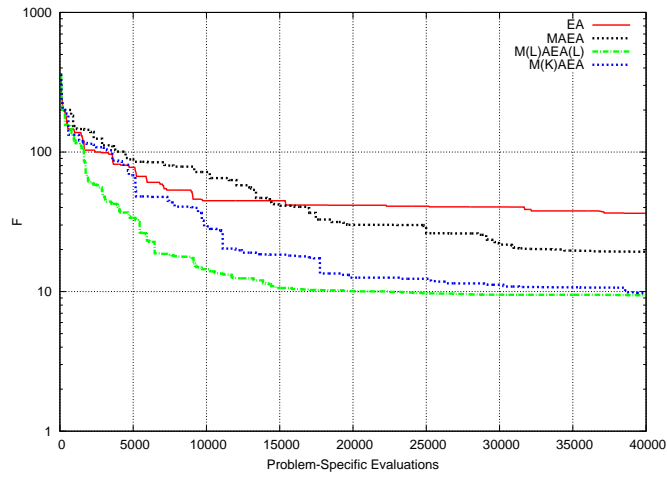
Σχήμα 4.9: Μη-Διαχωρίσιμη Παραλλαγή 20Δ Συνάρτησης Rastrigin: Πορείες σύγκλισης για τις μεθόδους και EA(L) για πέντε αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.

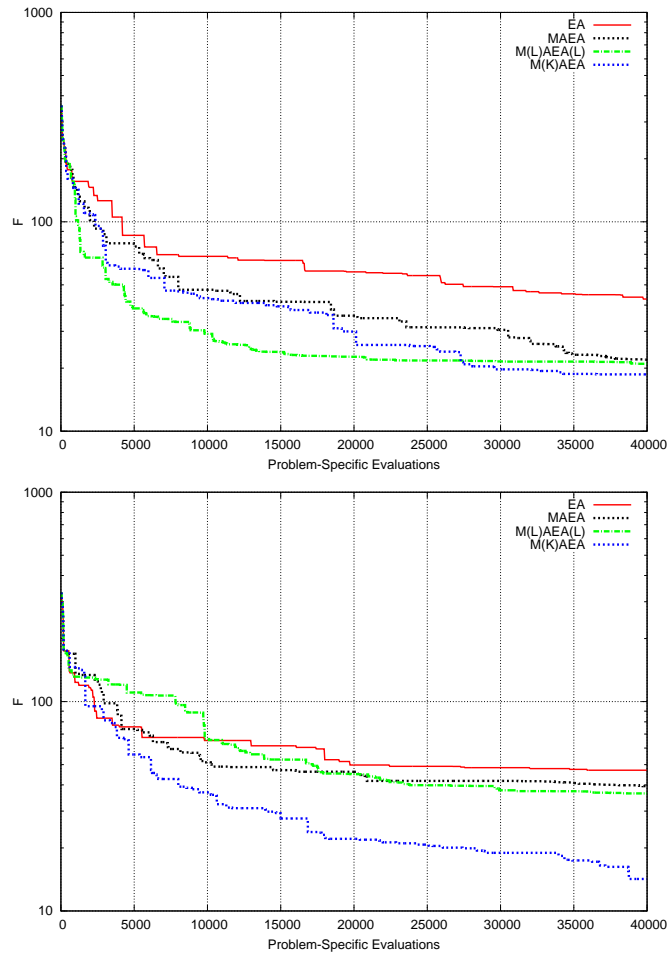


Σχήμα 4.10: Μη-Διαχωρίσιμη Παραλλαγή 20Δ Συνάρτησης Rastrigin: Μέση πορεία σύγκλισης για τις μεθόδους EA, EA(L) και EA(K).

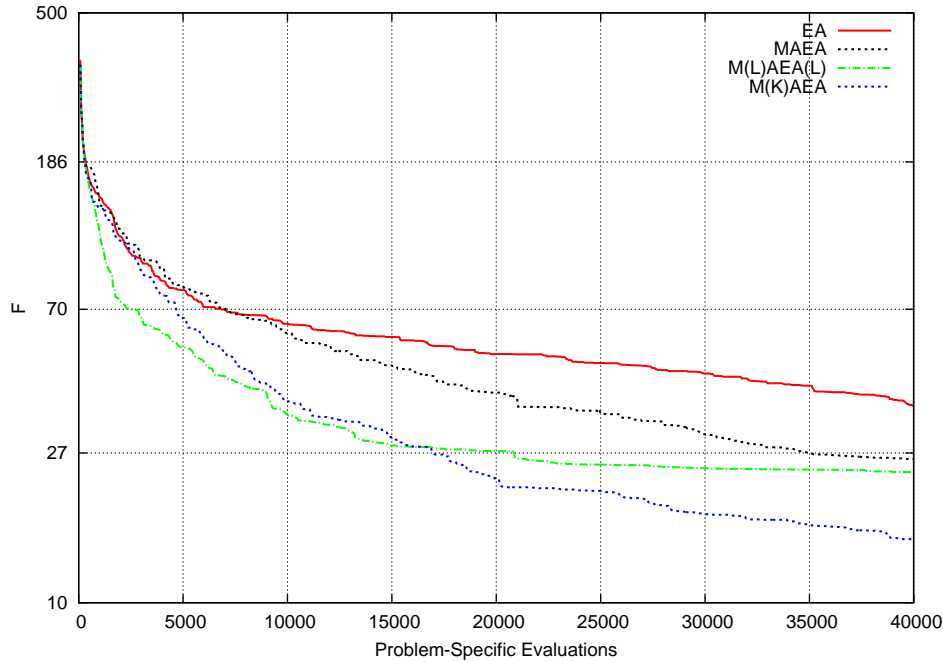
Πραγματοποιήθηκε επίσης μελέτη για την επιρροή των μεθόδων στα μεταπρότυπα τοπικής εμβέλειας. Τα μεταπρότυπα που χρησιμοποιήθηκαν είναι τεχνητά νευρωνικά δίκτυα, τα οποία χρησιμοποιούν 15-25 δείγματα εκπαίδευσης. Η λειτουργία τους ξεκινά μετά από 300 εγγραφές στη βάση δεδομένων, και σε κάθε γενιά γίνονται 5-7 ακριβείς αξιολογήσεις. Κατά τη χρήση των μεθόδων ΑσκΣ στο μεταπρότυπο, αποκόπτονται 6 διαστάσεις, συνεπώς τα μεταπρότυπα εκπαιδεύονται στο χώρο των 14 διαστάσεων.

Παρατίθενται οι πορείες σύγκλισης για τις μεθόδους MAEA, M(L)AEA(L) και M(K)AEA. Η εκδοχή M(K)AEA είχε την καλύτερη πορεία σύγκλισης κατά μέσο όρο. Η καλύτερη λύση στο πρόβλημα, εντός των μέγιστων αξιολογήσεων που ορίστηκαν βρέθηκε από την εκδοχή M(L)AEA(L).





Σχήμα 4.11: Μη-Διαχωρίσιμη Παραλλαγή 20Δ Συνάρτησης Rastrigin: Πορείες σύγκλισης για τις μεθόδους EA, MAEA, M(L)MAEA(L) και M(K)MAEA για πέντε αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.



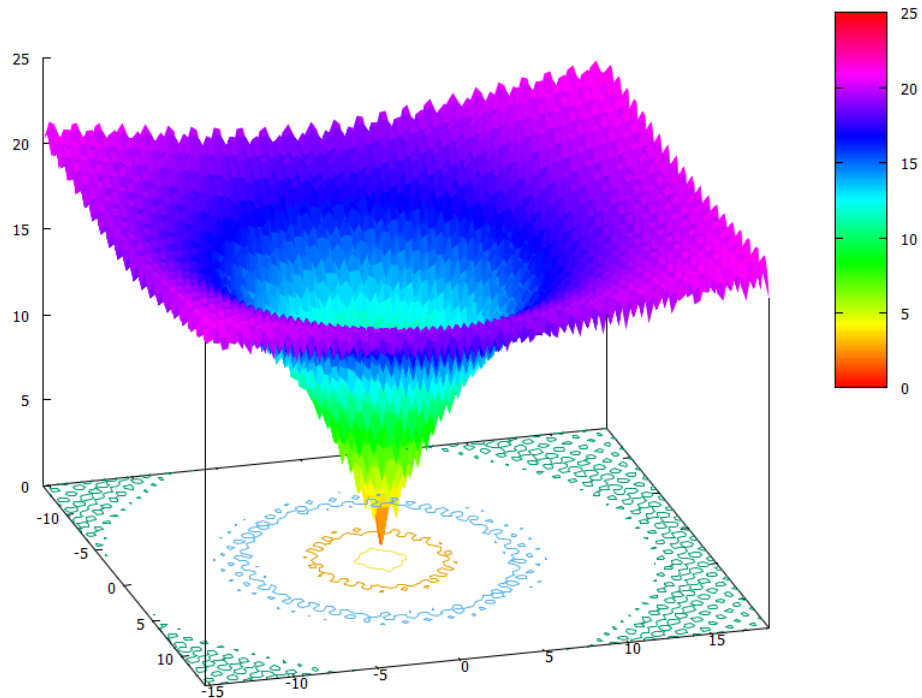
Σχήμα 4.12: Μη-Διαχωρίσιμη Παραλλαγή 20Δ Συνάρτησης Rastrigin : Μέση πορεία σύγκλισης για τις μεθόδους EA, MAEA M(L)AEA(L) και M(K)AEA.

### 4.3 Συνάρτηση Ackley & Παραλλαγές της

Η συνάρτηση Ackley δίνεται από τη σχέση [36]:

$$f(\vec{x}) = -20 \exp \left[ 0.2 \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \right] - \exp \left[ \sum_{i=1}^N \frac{1}{N} \cos(2\pi x_i) \right] + 20 + e, \quad x_i \in [-5.0, 5.0]. \quad (4.3)$$

Παρότι η συνάρτηση Ackley δεν μπορεί να παρασταθεί ως άθροισμα συναρτήσεων που εμπλέκουν μόνο μία μεταβλητή, όπως στην περίπτωση της Rastrigin, είναι διαχωρίσιμη, αφού η επιλογή του βέλτιστου  $x_i$  είναι ανεξάρτητη από την τιμή των υπόλοιπων  $x$ . Η συνάρτηση χαρακτηρίζεται επίσης από ισοτροπία, όπως δείχνει η συμμετρία της σχέσης 4.3.

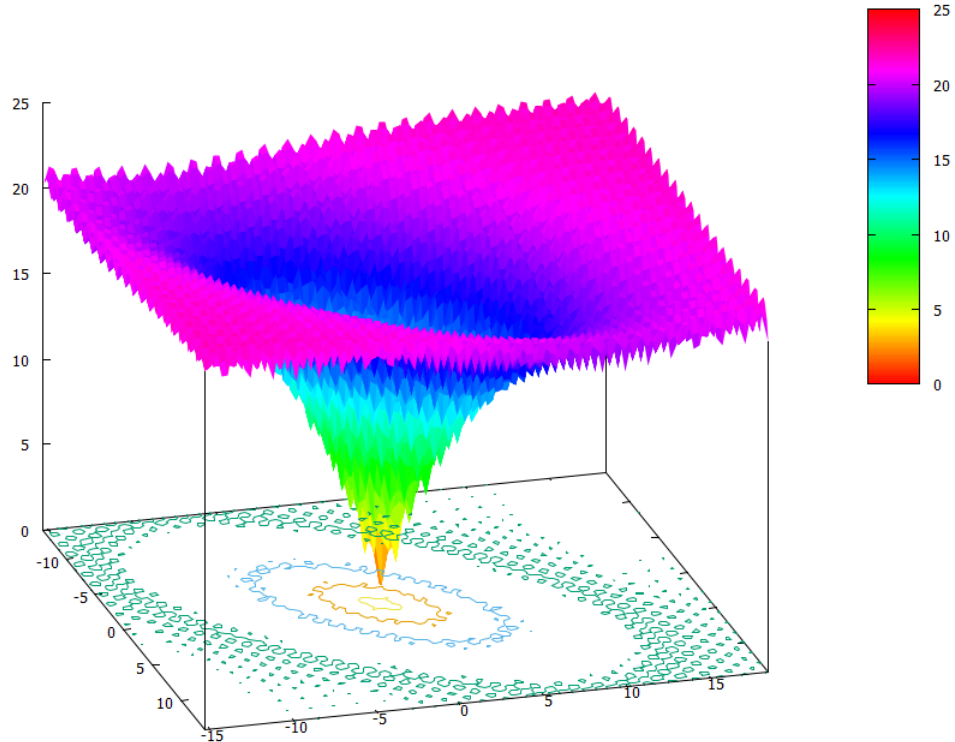


Σχήμα 4.13: Η συνάρτηση Ackley για N=2.

Όπως και με τη συνάρτηση Rastrigin , δημιουργήθηκαν και μελετήθηκαν μη-διαχωρίσιμες και ανισότροπες παραλλαγές του προβλήματος, της μορφής:

$$f(\vec{x}) = -20 \exp \left[ 0.2 \sqrt{\frac{1}{N} \sum_{i=1}^N w_i^2 \beta_i^2(\vec{x})} \right] - \exp \left[ \sum_{i=1}^N \frac{1}{N} \cos(2\pi \beta_i(\vec{x})) \right] + 20 + e \quad (4.4)$$

Το σχ. 4.14 απεικονίζει μια 2Δ τέτοια παραλλαγή.



Σχήμα 4.14: Παράδειγμα μιας 2Δ μη-διαχωρίσιμης (στραμμένης κατά  $45^\circ$ ) ανισότροπης συνάρτησης Ackley, με  $\vec{w} = [1, 4]$ .

Και εδώ η μελέτη έγινε για περιπτώσεις 5 και 20 διαστάσεων:

### Περίπτωση N=5 διαστάσεων

Στο  $\vec{d}$  δόθηκαν οι τιμές  $[1, 2, 3, 4, 0]$ . Η στροφή αποτελείται από 2 στροφές  $45^\circ$  στα επίπεδα που ορίζονται από τους άξονες  $x_1 - x_2$  και  $x_2 - x_3$ , και μία στροφή  $30^\circ$  στο επίπεδο που ορίζεται από τους  $x_3 - x_4$ . Το διάνυσμα ανισοτροπίας ισούται με  $\vec{w} = [1, 1, 1, 10, 1]$ .

Για αυτό το πρόβλημα χρησιμοποιήθηκε ένας (10,20)EA, με κριτήριο τερματισμού τις 5000 αξιολογήσεις. Εκτός από τον απλό EA, έγινε επίλυση με τις εκδοχές EA(L) και EA(K), με την εφαρμογή της ΑσΚΣ στους τελεστές εξέλιξης να ξεκινά από την τρίτη γενιά.

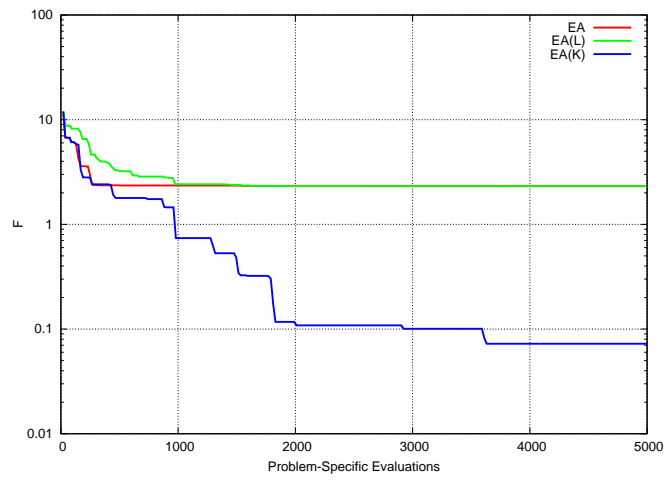
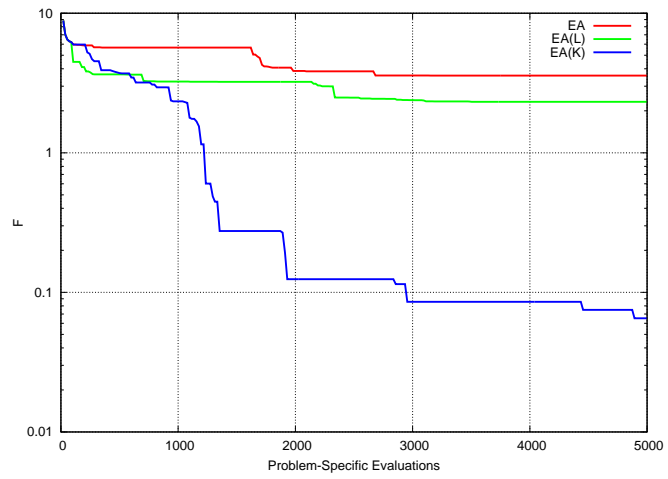
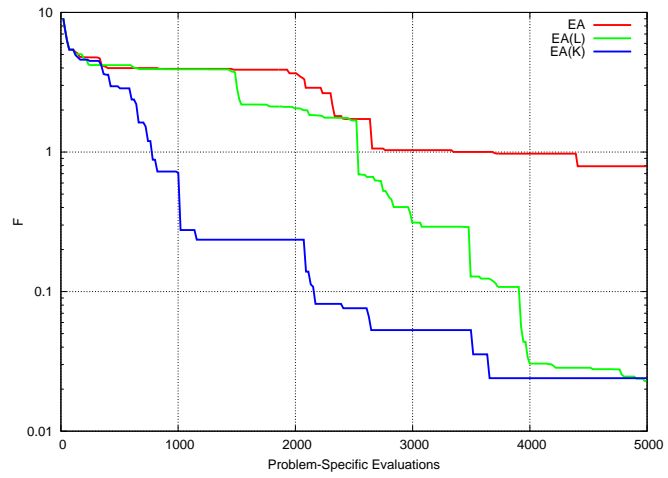
Κωδικοποίηση	Σχήμα Διασταύρωσης	Πιθανότητα Μετάλλαξης $P_m$
Πραγματική	Intermediate Recombination	0.05

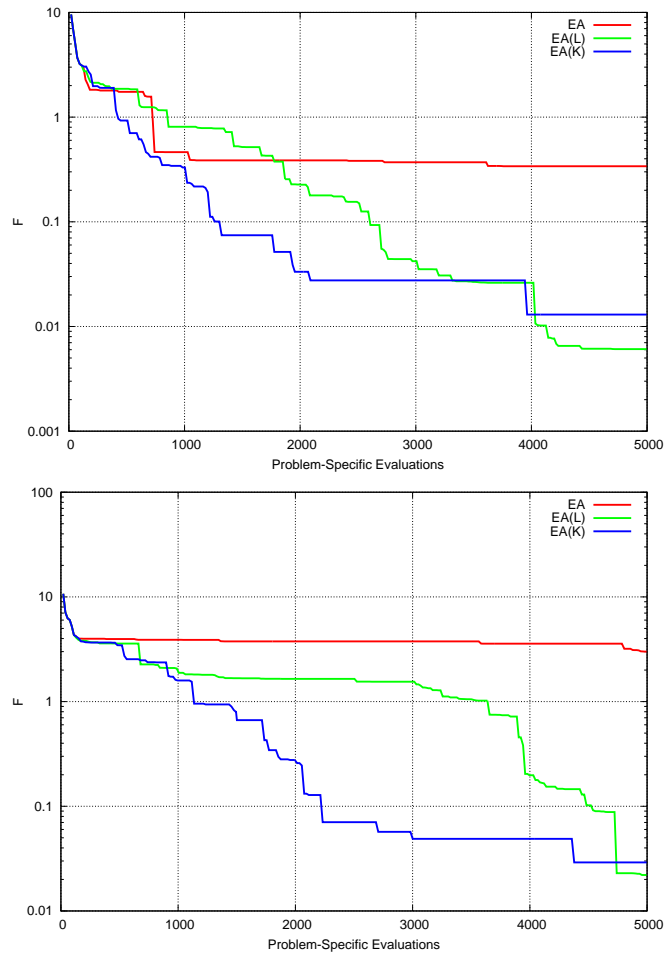
Πίνακας 4.2: Επιπλέον ρυθμίσεις των δοκιμών.

Στο σχήμα 4.15 απεικονίζονται οι πορείες σύγκλισης για πέντε διαφορετικές τιμές της γεννήτριας τυχαίων αριθμών. Στο σχήμα 4.16 απεικονίζεται η μέση πορεία σύγκλισης για τις μεθόδους. Όπως φαίνεται από τα διαγράμματα, είναι σημαντική η βελτίωση της σύγκλισης με την εφαρμογή της ΑΣκΣ στους εξελικτικούς τελεστές.

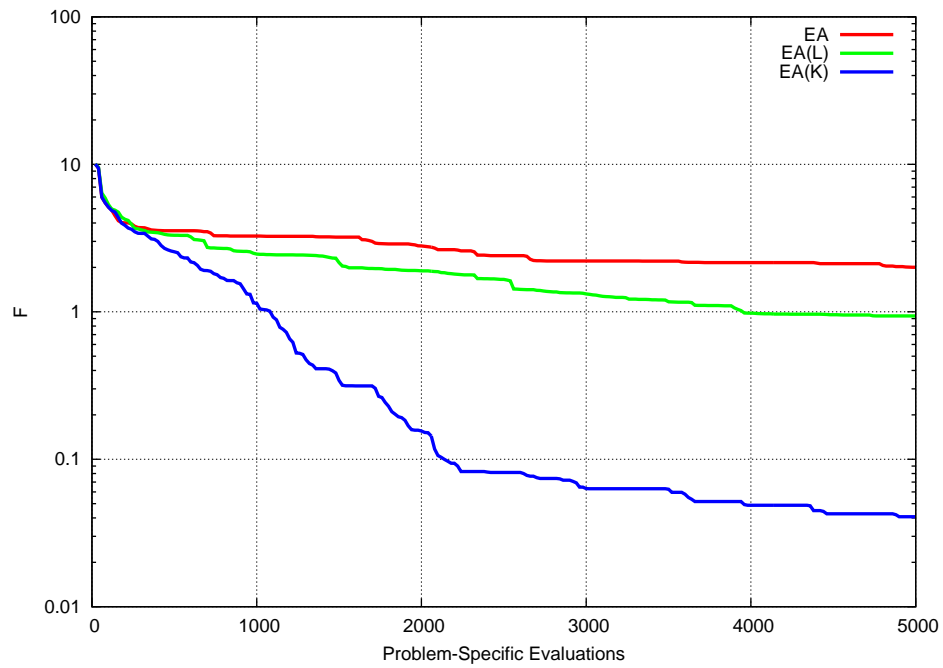
Αντίθετα, για δοκιμές που έγιναν με το αντίστοιχο διαχωρίσιμο πρόβλημα, οι εκδοχές EA(L) και EA(K) δεν βοηθούν ιδιαίτερα τη σύγκλιση. Κάτι τέτοιο όμως δεν είναι ανησυχητικό, εφόσον η διαχωρίσιμη εκδοχή του προβλήματος μπορεί ευκολα να λυθεί από το συμβατικό EA. Αυτό απεικονίζεται στο σχήμα 4.17.



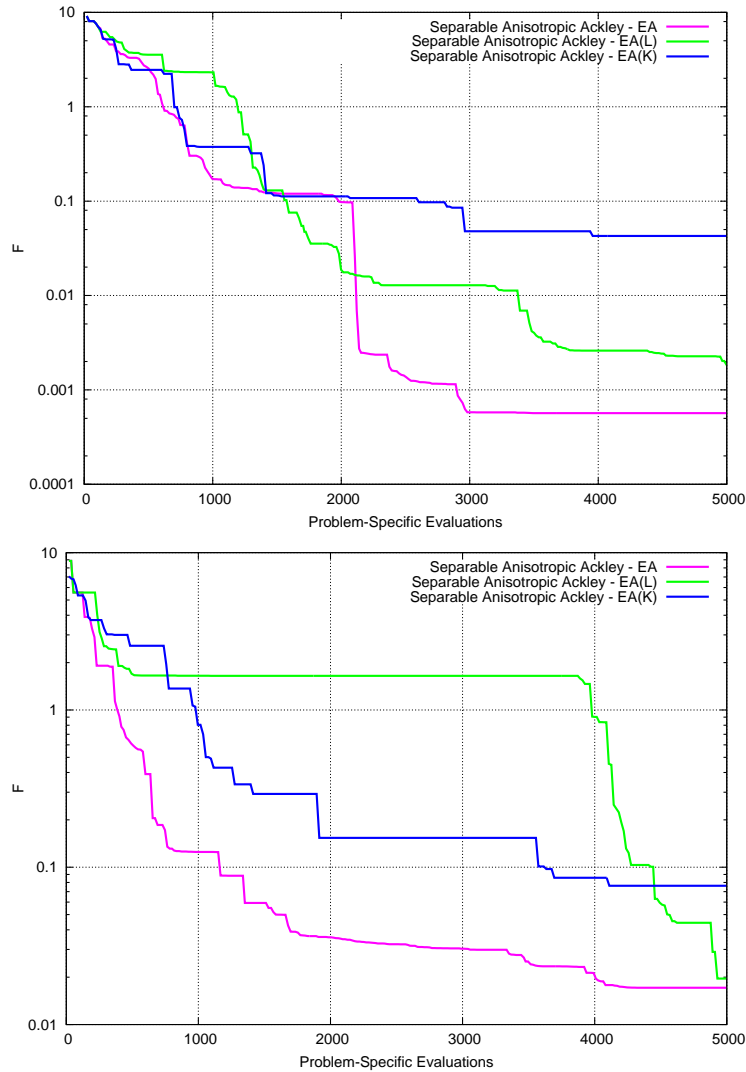




Σχήμα 4.15: Μη-Διαχωρίσιμη Παραλλαγή 5Δ Συνάρτησης Ackley: Πορείες σύγκλισης για τις μεθόδους και EA, EA(L) και EA(K) για πέντε αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.



Σχήμα 4.16: Μη-Διαχωρίσιμη Παραλλαγή 5Δ Συνάρτησης Ackley: Μέση πορεία σύγκλισης για τις μεθόδους EA, EA(L) και EA(K).



Σχήμα 4.17: 5Δ Διαχωρίσιμη Παραλλαγή Συνάρτησης Ackley: Πορείες σύγκλισης για τις μεθόδους EA(L) και EA(K), για δύο αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.

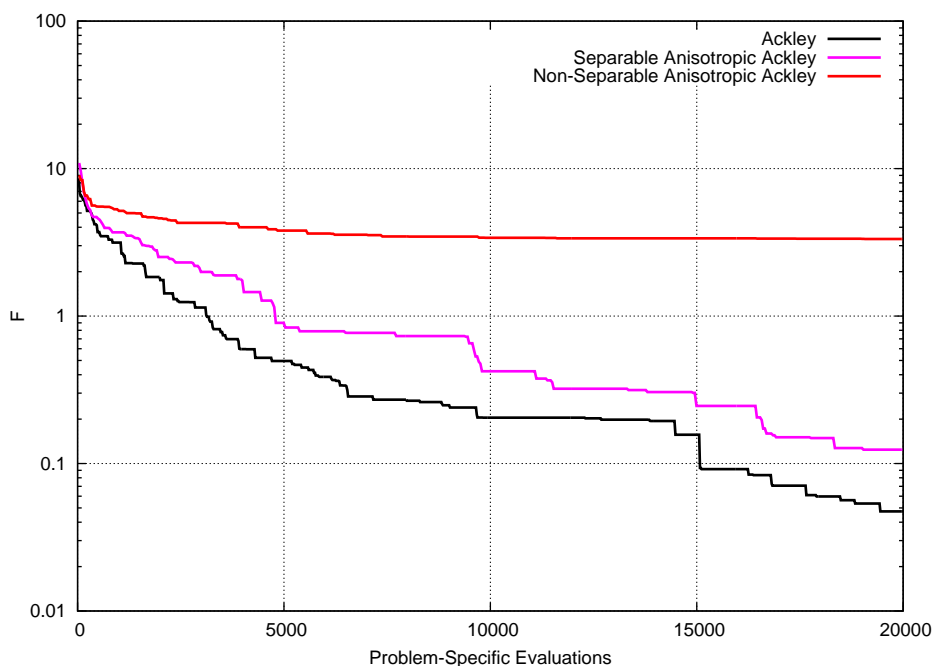
### Περίπτωση $N=20$ διαστάσεων

Η συνάρτηση παρουσιάζει ελάχιστο στην αρχή των αξόνων. Το μητρώο στροφών αποτελείται από πέντε περιστροφές  $45^\circ$ , οι οποίες εμπλέκουν τις πρώτες τέσσερις μεταβλητές σχεδιασμού, ενώ το διάνυσμα ανισοτροπίας ισού-

ται με  $\vec{w} = [1, 100, 1, 1, \dots, 1, 1]$ . Για την επίλυση του προβλήματος χρησιμοποιήθηκε ένας (20,40)ΕΑ, με κριτήριο τερματισμού του αλγορίθμου τις 20000 αξιολογήσεις.

Κωδικοποίηση	# Τρεξιμάτων	Σχήμα Διασταύρωσης	Πιθανότητα Μετάλλαξης $P_m$
Πραγματική	5	Intermediate Recombination	0.05

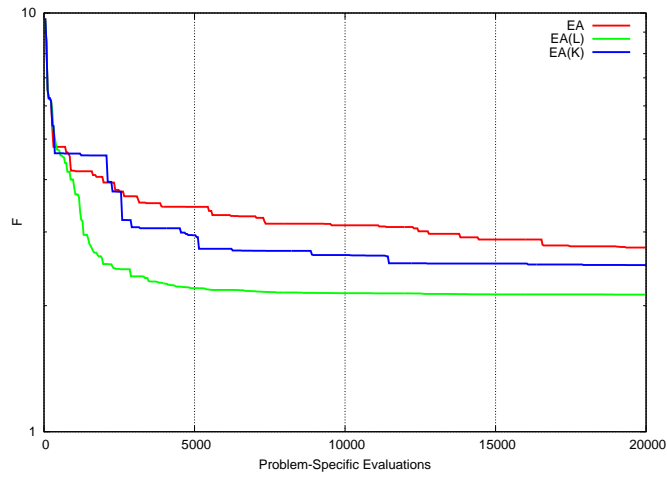
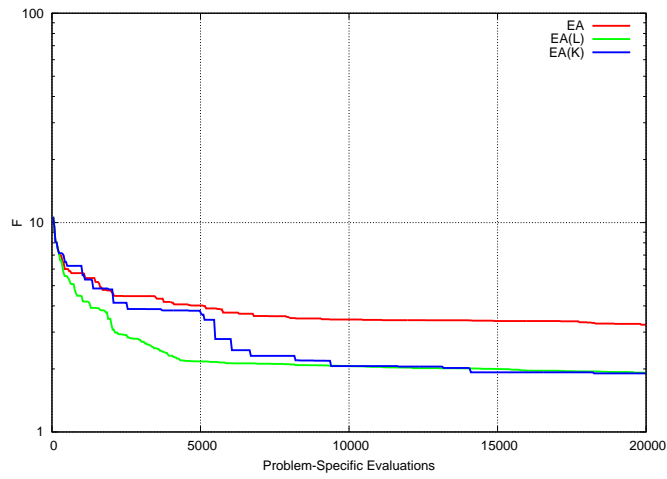
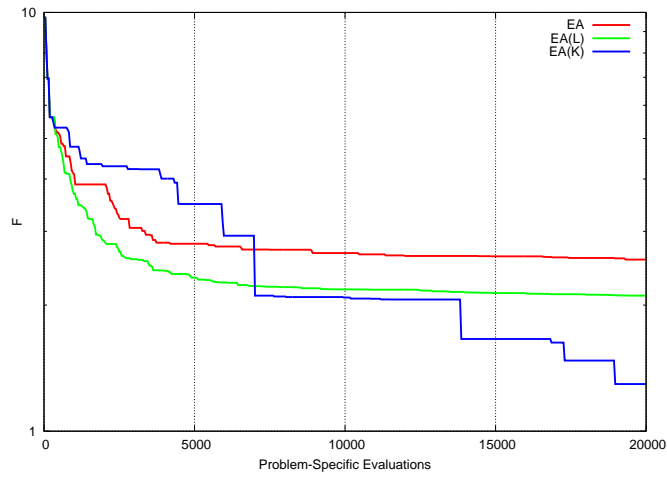
Πίνακας 4.3: Επιπλέον χαρακτηριστικά των δοκιμών.

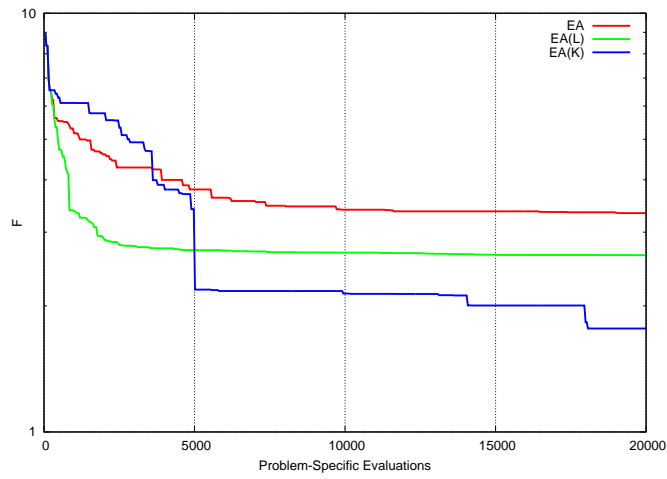
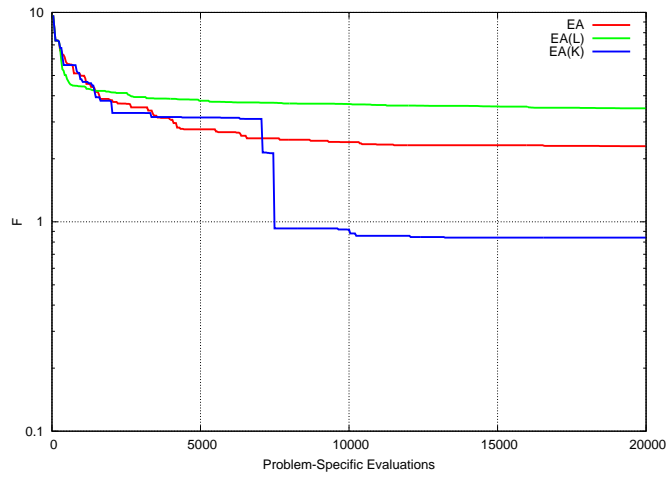


Σχήμα 4.18: Πορείες σύγκλισης του απλού ΕΑ για τη συνάρτηση Ackley είκοσι διαστάσεων και δύο παραλλαγές της.

Στο σχήμα 4.18 απεικονίζονται οι πορείες σύγκλισης, για μία δοκιμή, των διάφορων παραλλαγών της 20Δ συνάρτησης Ackley. Και σε αυτήν την περίπτωση, είναι εμφανής η δυσκολία του απλού ΕΑ να συγκλίνει γρήγορα στη βέλτιστη λύση, όταν το πρόβλημα είναι μη-διαχωρίσιμο και ανισότροπο. Στα διαγράμματα των σχημάτων 4.19, 4.20, απεικονίζονται οι πορείες σύγκλισης των εκδοχών EA(L), EA(K), για πέντε δοκιμές της γεννήτριας τυχαίων αριθ-

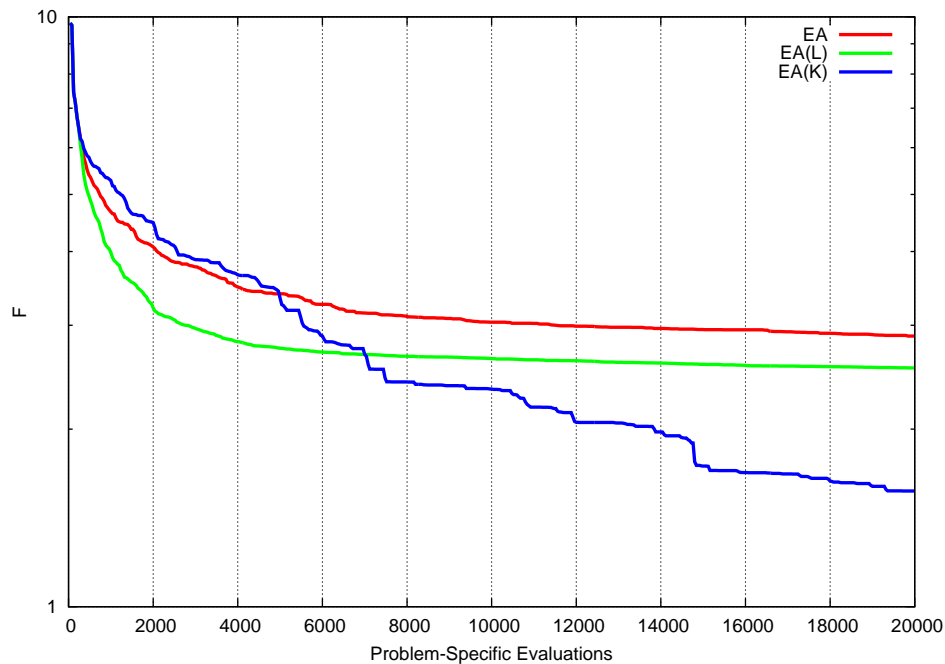
μών. Όπως φαίνεται από το σύνολο των διαγραμμάτων, και στην 20Δ παρα-  
λλαγή της συνάρτησης οι εκδοχές υπερτερούν, κατά πολύ, του απλού ΕΑ.





Σχήμα 4.19: Μη-Διαχωρίσιμη Παραλλαγή 20Δ Συνάρτησης Ackley: Πορείες σύγκλισης για τις μεθόδους EA, EA(L) και EA(K) για πέντε αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.

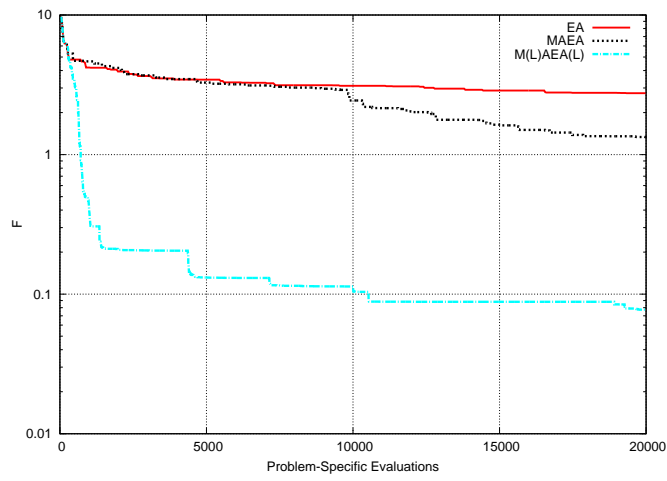
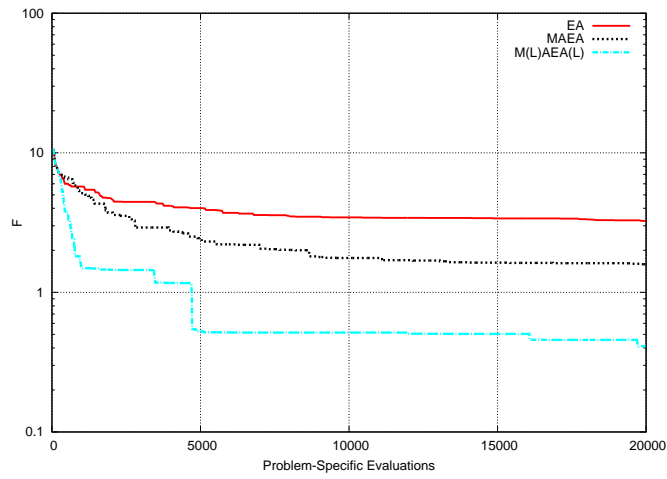
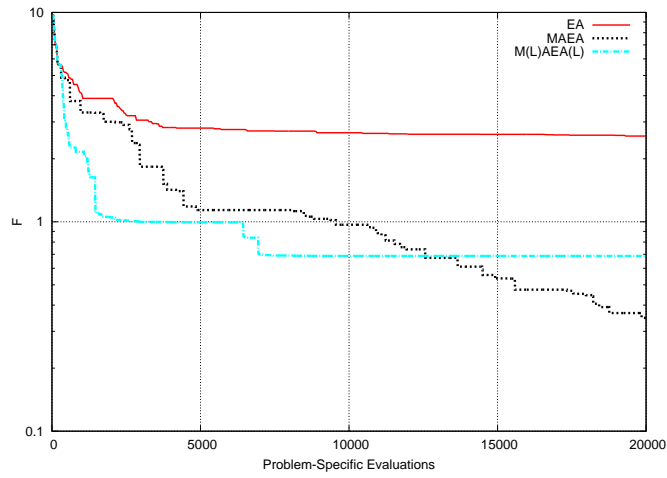


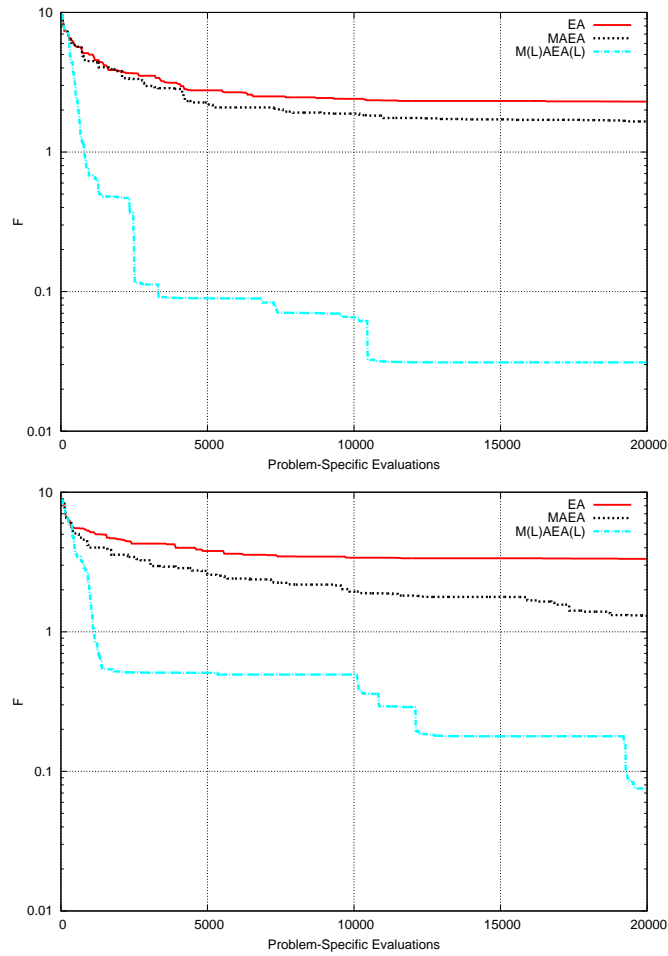


Σχήμα 4.20: Μη-Διαχωρίσιμη Παραλλαγή 20Δ Συνάρτησης Ackley: Μέση πορεία σύγκλισης για τις μεθόδους EA, EA(L) και EA(K) για πέντε αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.

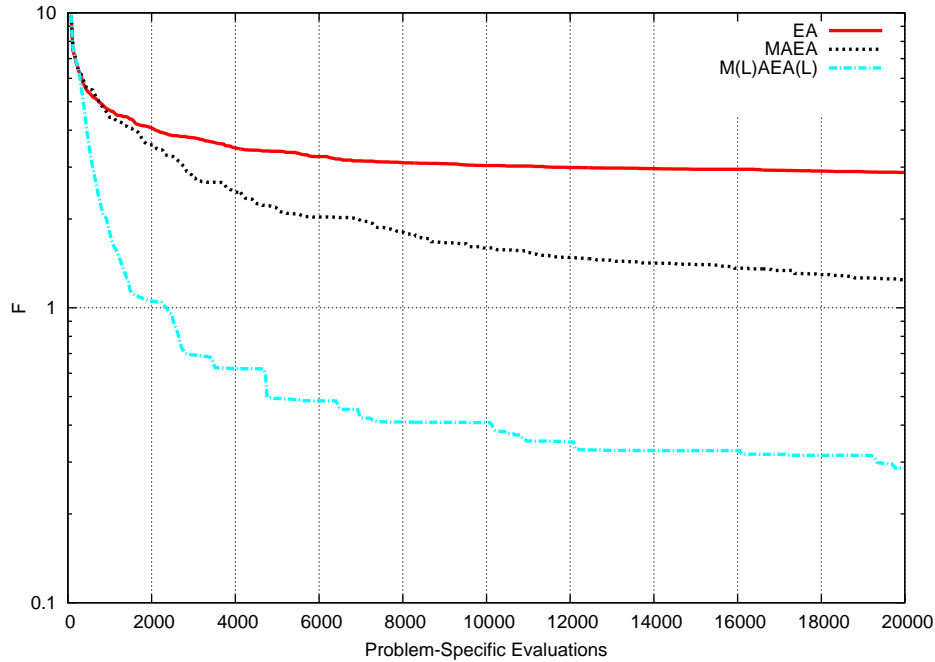
Έγινε επίσης επίλυση του προβλήματος, για τις ίδιες πέντε τιμές της γεννήτριας τυχαίων αριθμών, με τις εκδοχές MAEA και M(L)AEA(L). Οι αξιολογήσεις με μεταπρότυπα ενεργοποιούνται αφού συμπληρωθούν 250 εγγραφές στη βάση δεδομένων. Κάθε μεταπρότυπο εκπαιδεύεται με 10-20 δείγματα εκπαίδευσης, ενώ κατά την εφαρμογή της ΑσκΣ στο μεταπρότυπο αποκόπτονται 6 διαστάσεις. Έτσι τα μεταπρότυπα εκπαιδεύονται με 14 διαστάσεις, έναντι των 20 αρχικών. Τα αποτελέσματα απεικονίζονται στα διαγράμματα του σχ.4.22.

Τα μεταπρότυπα ενισχύουν τη σύγκλιση, ενώ ο συνδυασμός της γραμμικής ΑσκΣ στην εκπαίδευση τους και στους τελεστές εξέλιξης ακόμα περισσότερο. Η εκδοχή M(L)AEA(L), δίνει το καλύτερο αποτέλεσμα, συγκλίνοντας από πολύ νωρίς σε πολύ χαμηλή τιμή. Η ίδια μέθοδος βρήκε και την καλύτερη λύση του προβλήματος με τιμή 0.03, πολύ κοντά στο ολικό βέλτιστο της συνάρτησης.





Σχήμα 4.21: Μη-Διαχωρίσιμη Παραλλαγή 20Δ Συνάρτησης Ackley: Πορείες σύγκλισης για τις μεθόδους EA, MAEA, και M(L)MAEA(L) για πέντε αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.



Σχήμα 4.22: 20Δ Παραλλαγή Συνάρτησης Ackley: Μέση πορεία σύγκλισης για τις μεθόδους EA, MAEA, και M(L)MAEA(L).

## 4.4 Συνάρτηση Fonseca - Fleming

Το τελευταίο μαθηματικό πρόβλημα που μελετήθηκε είναι η συνάρτηση Fonseca-Fleming [22, 18]. Πρόκειται για πρόβλημα βελτιστοποίησης δύο στόχων:

$$\begin{aligned} \text{minimize} \begin{cases} f_1(\vec{x}) &= 1 - \exp[-\sum_{i=1}^n (x_i - \frac{1}{\sqrt{n}})^2] \\ f_2(\vec{x}) &= 1 - \exp[-\sum_{i=1}^n (x_i + \frac{1}{\sqrt{n}})^2] \end{cases} \quad (4.5) \\ x_i &\in [-4, 4] \end{aligned}$$

Οι συναρτήσεις στόχου  $f_1$  και  $f_2$  έχουν τη μορφή της συνάρτησης Gauss. Το βέλτιστο της  $f_1$  βρίσκεται στο σημείο  $\vec{x}_I^* = [\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}]$ , ενώ αυτό της  $f_2$  βρίσκεται στο σημείο  $\vec{x}_{II}^* = [-\frac{1}{\sqrt{n}}, -\frac{1}{\sqrt{n}}, \dots, -\frac{1}{\sqrt{n}}]$ .

Το σύνολο των μη-κυριαρχούμενων λύσεων για το πρόβλημα δύο στόχων αποτελείται από τα σημεία  $\vec{x}_{FON}^* = [x^*, x^*, \dots, x^*]$ ,  $x^* \in [-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}]$ , πρόκειται δηλαδή για το ευθύγραμμο τμήμα που ενώνει τα βέλτιστα των δύο συναρτήσεων  $f_1$  και  $f_2$ .

Κάθε μία από τις δύο συναρτήσεις  $f_1(\vec{x})$  και  $f_2(\vec{x})$  είναι διαχωρίσιμη και ισότροπη ως προς όλες τις μεταβλητές βελτιστοποίησης  $x_i$ . Ωστόσο, όπως αναφέρθηκε στο προηγούμενο κεφάλαιο, η διαδικασία βάσει της οποίας το διάνυσμα τιμών των συναρτήσεων στόχων (εδώ το  $[f_1, f_2]$ ) μετατρέπεται σε βαθμωτό κόστος, αυτόματα ορίζει σημαντικές διευθύνσεις στο χώρο σχεδιασμού. Συνεπώς, η εφαρμογή της ΑσκΣ μπορεί να βελτιώσει την επίδοση του αλγορίθμου.

Στο χώρο των συναρτήσεων στόχων το μέτωπο Pareto περιγράφεται από τη σχέση:

$$f_2^* = 1 - \exp\{-[2 - \sqrt{-\ln(1 - f_1^*)}]^2\} \quad (4.6)$$

για  $0 \leq f_1^* \leq 1 - \exp(-4)$ .

Το πρόβλημα μελετήθηκε για  $n = 20$  διαστάσεις, χρησιμοποιώντας τις εκδοχές EA, EA(L) EA(K), MAEA, M(L)AEA(L), M(K)AEA(K). Το πλήθος των γονέων και των απογόνων κάθε γενιάς τέθηκε ίσο προς  $\mu=10$ ,  $\lambda=20$  ενώ το σύνολο των επίλεκτων αποτελείται από  $e=20$  άτομα.

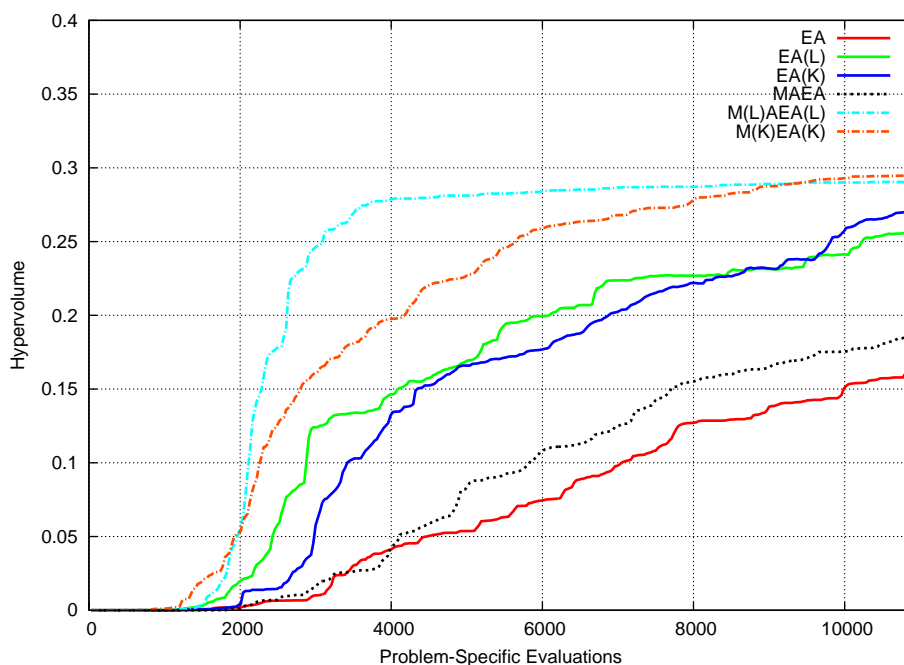
Για την απόδοση βαθμωτού κόστους σε κάθε μέλος του πληθυσμού χρησιμοποιείται η μέθοδος SPEA2 [38]. Ως κριτήριο για την ποιότητα των συνόλων λύσεων που παράγονται χρησιμοποιείται ο δείκτης υπερόγκου-ΔΥ (Hypervolume Indicator - HV)[15]. Ο ΔΥ αποτελεί μετρική η οποία αντιστοιχίζει τα σύνολα λύσεων σε ένα βαθμωτό μέγεθος, και αναλύεται στο παράρτημα Α.

Τα μεταπρότυπα που χρησιμοποιούνται είναι ΔΣΑΒ, τα οποία εκπαιδεύονται με 10-15 δείγματα εκπαίδευσης. Η λειτουργία των μεταπρότυπων ξεκινά μετά από 90 εγγραφές στη βάση δεδομένων. Κατά τη χρήση της ΑσκΣ στο μεταπρότυπο, αποκόπτονται 10 διαστάσεις.

Κωδικοποίηση	# Τρεξιμάτων	Σχήμα Διασταύρωσης	Πιθανότητα Μετάλλαξης $P_m$
Πραγματική	5	Intermediate Recombination	0.03

Πίνακας 4.4: Επιμέρους χαρακτηριστικά των δοκιμών.

Κάθε εκδοχή δοκιμάστηκε με πέντε διαφορετικές αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών, όπου ο αλγόριθμος έτρεξε για 12000 ακριβείς αξιολογήσεις. Μετά από κάθε τρέξιμο, συγκεντρώνονται τα σύνολα των επίλεκτων κάθε γενιάς  $P_e^g$ , και υπολογίζεται η πορεία εξέλιξης του δείκτη υπερόγκου βάσει του σημείου ναδίρ  $\Omega = (1, 1)$ , και του σημείου ζενίθ  $A = (0, 0)$ . Στη συνέχεια, υπολογίζεται η μέση πορεία της σύγκλισης του δείκτη υπερόγκου για όλα τα τρέξιμα. Τα αποτελέσματα απεικονίζονται στο σχήμα 4.23.



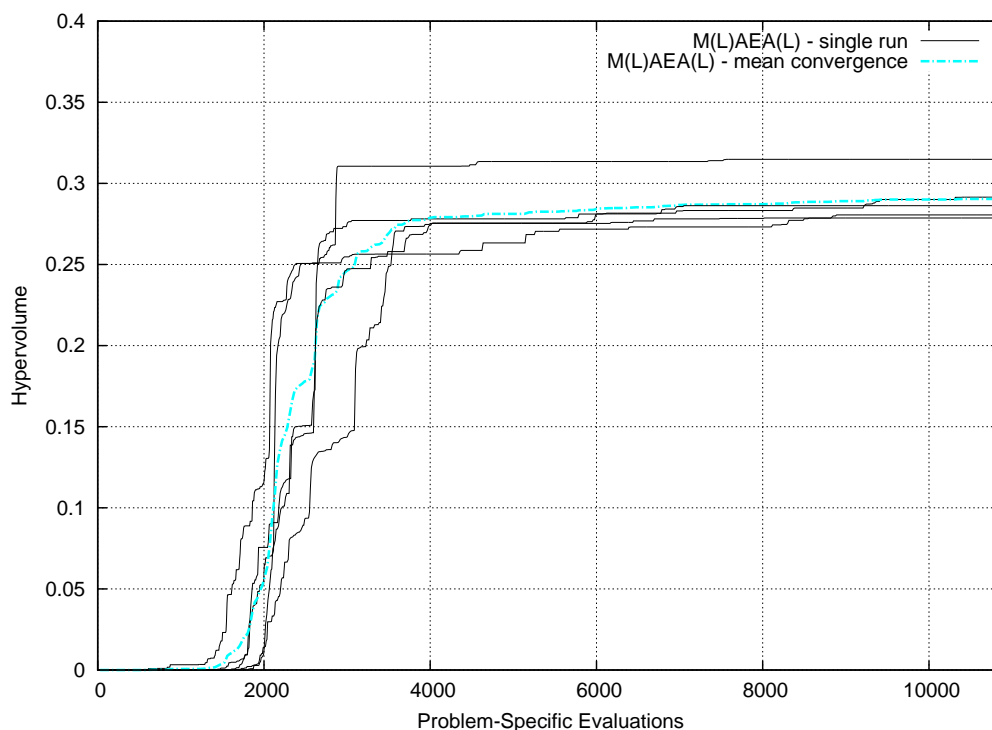
Σχήμα 4.23: Συνάρτηση Fonseca - Fleming: Μέση πορεία σύγκλισης του δείκτη υπερόγκου για τις διάφορες εκδοχές.

Από τα αποτελέσματα είναι εμφανές το πλεονέκτημα που προσφέρει η ανάλυση σε κύριες συνιστώσες, τόσο στη γραμμική όσο και στη μη-γραμμική εκδοχή της. Και για τις δύο μεθόδους, η εφαρμογή τους στους εξελικτικούς τελεστές βελτιώνει σημαντικά την απόδοση του EA. Σε συνδυασμό δε με τα μεταπρότυπα, η σύγκλιση βελτιώνεται σε ακόμα μεγαλύτερο βαθμό.

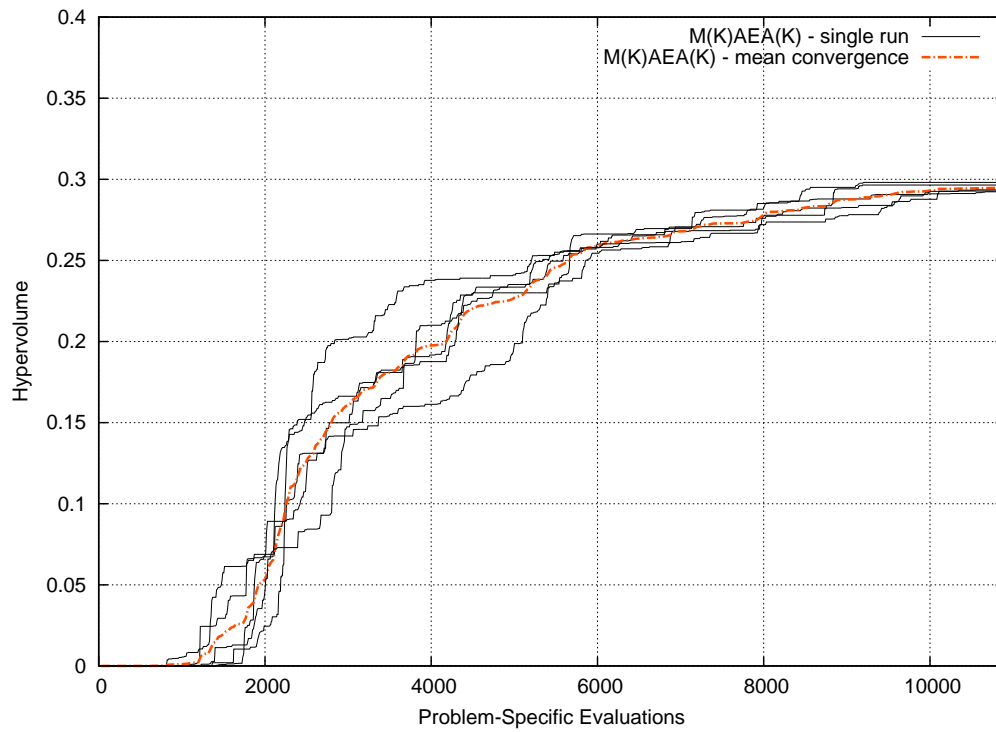
Το τελικό σύνολο λύσεων που παράγουν οι μέθοδοι M(L)AEA(L) και M(K)AEA(K) πρακτικά ταυτίζεται με το μέτωπο Pareto του προβλήματος. Ενδεικτικά, στο σχήμα 4.26, απεικονίζονται τα σύνολα των επίλεκτων για ένα

τρέξιμο της εκδοχής MAEA, και το αντίστοιχο για την M(K)AEA(K), με το δεύτερο να είναι εμφανώς ποιοτικότερο, αφού κυριαρχεί καθολικά του πρώτου.

Τα σχήματα 4.24 και 4.25 συγκεντρώνουν τις πορείες σύγκλισης για μεμονωμένα τρεξίματα μαζί με τη μέση πορεία σύγκλισης, για τις δύο εκδοχές που έδωσαν τα καλύτερα αποτελέσματα, τις M(L)AEA(L) και M(K)AEA(K). Όπως δείχνουν τα διαγράμματα, τα οποία έχουν μορφή θυσάνου, η συμπεριφορά των εκδοχών παρουσιάζει μικρή διακύμανση για διαφορετικές τιμές της γεννήτριας τυχαίων αριθμών. Συνεπώς, η βελτίωση που προκαλείται από το συνδυασμό των μεθόδων είναι σταθερή, και η μέση πορεία σύγκλισης που παρουσιάζεται αντιπροσωπεύει ικανοποιητικά τις δύο εκδοχές.

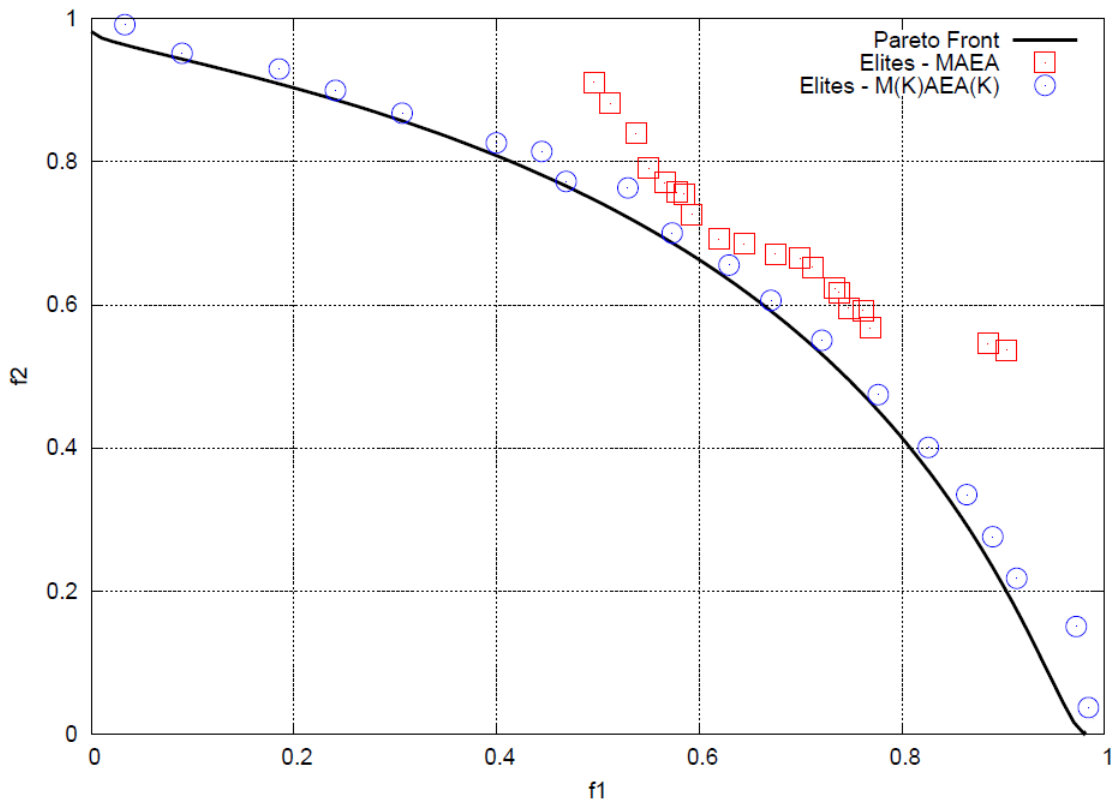


Σχήμα 4.24: Συνάρτηση Fonseca - Fleming: Πορείες σύγκλισης και μέση πορεία σύγκλισης για την εκδοχή M(L)AEA(L).



Σχήμα 4.25: Συνάρτηση Fonseca - Fleming: Πορείες σύγκλισης και μέση πορεία σύγκλισης για την εκδοχή M(K)AEA(K).





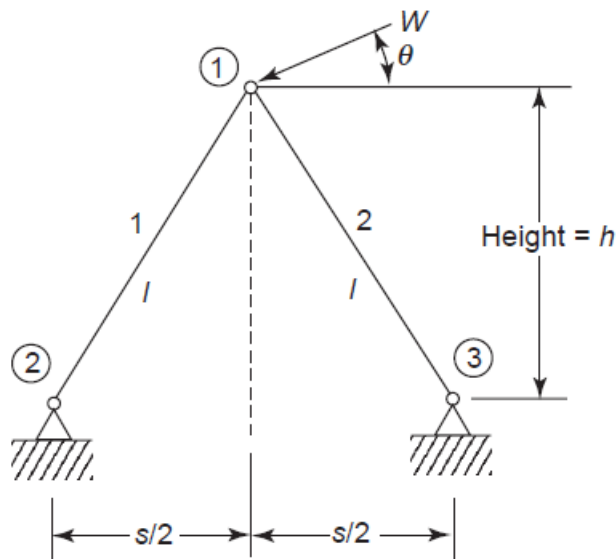
Σχήμα 4.26: Συνάρτηση Fonseca - Fleming: Μέτωπα μη-κυριαρχούμενων λύσεων για ένα τρέξιμο των εκδοχών MAEA και M(K)AEA(K).

## 4.5 Πρόβλημα Σχεδιασμού Υποστηρίγματος Δύο Ράβδων

Το τρίτο πρόβλημα που μελετήθηκε αφορά στο σχεδιασμό υποστηρίγματος δύο δοκών οι οποίες πρέπει να παραλάβουν φορτίο  $W$  υπό δεδομένη γωνία  $\theta$  (Two-Bar Bracket) [9]. Πρόκειται για μονοκριτηριακή βελτιστοποίηση, με στόχο την ελαχιστοποίηση της συνολικής μάζας της κατασκευής.

min		max	Περιγραφή
1.5	$x_1$	7	Συνολικό ύψος $h$ της κατασκευής.( $m$ )
1	$x_2$	4	Οριζόντια απόσταση μεταξύ των εδράσεων(span)( $m$ )
0.08	$x_3$	0.8	Εξωτερική διάμετρος της δοκού 1 ( $m$ )
0.08	$x_4$	0.8	Εσωτερική διάμετρος της δοκού 1 ( $m$ )
0.08	$x_5$	0.8	Εξωτερική διάμετρος της δοκού 2 ( $m$ )
0.08	$x_6$	0.8	Εσωτερική διάμετρος της δοκού 2 ( $m$ )

Πίνακας 4.5: Μεταβλητές σχεδιασμού του προβλήματος δύο δοκών.



Σχήμα 4.27: Σχηματική αναπαράσταση και βασικά μεγέθη του προβλήματος δύο δοκών.

Θεωρείται ότι οι δοκοί έχουν κυλινδρικό σχήμα, εσωτερικής διαμέτρου  $d_i$ , και εξωτερικής  $d_o$ . Το υλικό της κατασκευής θεωρήθηκε ατσάλι πυκνότητας  $\rho = 7800\text{kg/m}^3$  και αντοχής  $\sigma_a = 300\text{MPa}$ . Το μέτρο και η γωνία της δύναμης  $W$  τέθηκαν ίσα με  $10\text{kN}$  και  $\frac{\pi}{6}\text{rad}$  αντίστοιχα. Συνολικά, οι μεταβλητές σχεδιασμού είναι οι εξής έξι:

Η συνάρτηση στόχου που υπολογίζει τη συνολική μάζα της κατασκευής για δακτυλιοειδή διατομή δίνεται από τη σχέση:

$$F = M = [\rho \sqrt{x_1^2 + (0.5x_2)^2}] \frac{\pi}{4} (x_3^2 - x_4^2 + x_5^2 - x_6^2) \quad (kg) \quad (4.7)$$

Ταυτόχρονα, απαιτείται η ικανοποίηση περιορισμών, οι οποίοι προκύπτουν από τη γεωμετρία και τη στατική του προβλήματος. Συγκεκριμένα, η μέγιστη τάση που εμφανίζεται στις διατομές πρέπει να είναι μικρότερη από αυτήν που επιτρέπει η αντοχή του υλικού, μαζί με έναν παράγοντα ασφαλείας  $k$  (τέθηκε ίσος προς 1.8):

$$g_1 = \frac{2W \sqrt{x_1^2 + (0.5x_2)^2}}{\pi(x_3^2 - x_4^2)} \left[ \frac{\sin\theta}{x_1} + \frac{2\cos\theta}{x_2} \right] \leq \frac{\sigma_a}{k} \quad (4.8)$$

$$g_2 = \frac{2W \sqrt{x_1^2 + (0.5x_2)^2}}{\pi(x_5^2 - x_6^2)} \left[ \frac{\sin\theta}{x_1} - \frac{2\cos\theta}{x_2} \right] \leq \frac{\sigma_a}{k} \quad (4.9)$$

$$g_3 = \frac{-2W \sqrt{x_1^2 + (0.5x_2)^2}}{\pi(x_5^2 - x_6^2)} \left[ \frac{\sin\theta}{x_1} - \frac{2\cos\theta}{x_2} \right] \leq \frac{\sigma_a}{k} \quad (4.10)$$

Η πρώτη σχέση αναφέρεται στη μέγιστη τάση που εμφανίζεται στη διατομή της ράβδου 1, ενώ οι δύο επόμενες στην τάση που εμφανίζεται στη ράβδο 2, η οποία ανάλογα με τη γεωμετρία που επιλέγεται μπορεί να έχει θλιπτικό ή εφελκυστικό χαρακτήρα.

Τέλος, απαιτούνται δύο ακόμα περιορισμοί για το ελάχιστο πάχος που μπορεί να έχει η ράβδος:

$$g_4 = 0.02 - x_3 + x_4 \leq 0 \quad (4.11)$$

$$g_5 = 0.02 - x_5 + x_6 \leq 0 \quad (4.12)$$

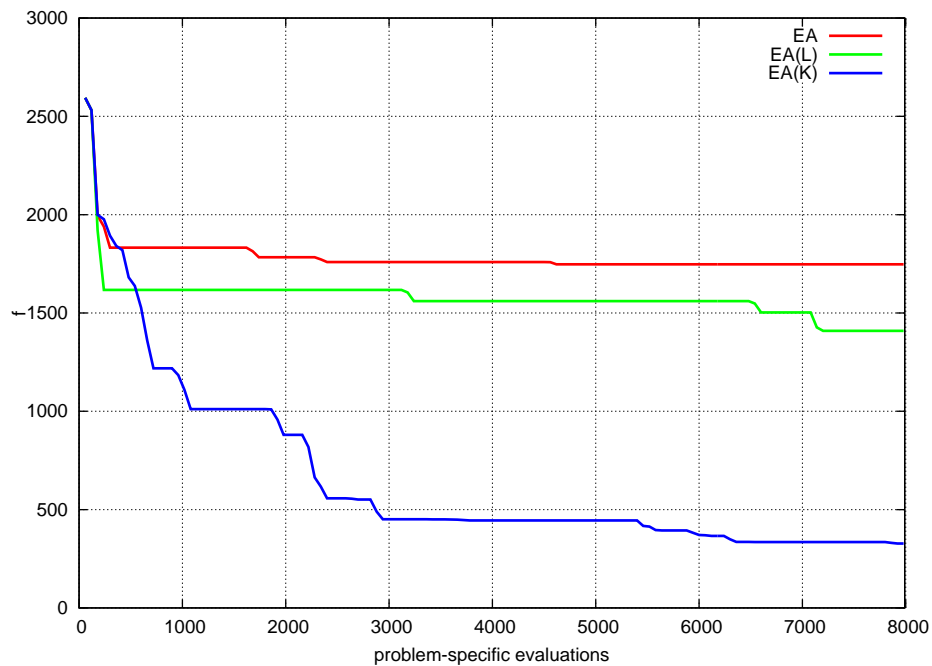
Με βάση τα παραπάνω διαμορφώθηκε λογισμικό αξιολόγησης των υποψήφιων λύσεων, το οποίο υπολογίζει την αντικειμενική συνάρτηση και κατά πόσο ικανοποιούνται όλοι οι περιορισμοί. Οι παράμετροι του εξελικτικού αλγορίθμου ρυθμίστηκαν ως εξής: Πλήθος γονέων  $\mu=30$ , πλήθος απογόνων  $\lambda=60$ . Χρησιμοποιήθηκε το σχήμα προσομοιούμενης δυαδικής διασταύρωσης (Simulated Binary Crossover), ενώ ο μέγιστος αριθμός αξιολογήσεων τέθηκε ίσος με 8000. Η εφαρμογή της ΑσκΣ στους εξελικτικούς τελεστές ξεκινά από τη δεύτερη γενιά.

Η λειτουργία των μεταπροτύπων ξεκινά μετά από 300 εγγραφές στη βάση δεδομένων, ενώ τα δείγματα εκπαίδευσης που χρησιμοποιούνται είναι 15-20.

Κωδικοποίηση	# Τρεξίμάτων	Σχήμα Διασταύρωσης	Πιθανότητα Μετάλλαξης $P_m$
Πραγματική	10	Simulated Binary	0.03

Πίνακας 4.6: Επιπλέον χαρακτηριστικά των δοκιμών.

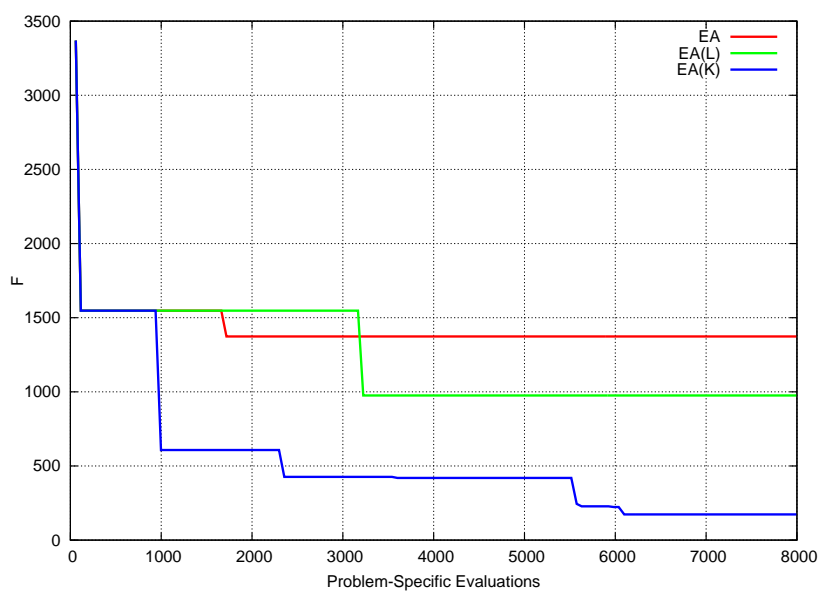
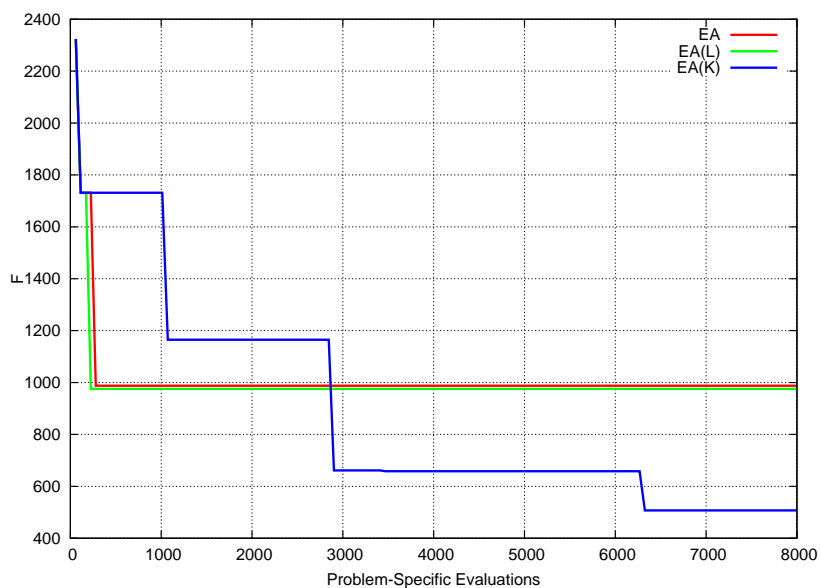
Κατά την αποκοπή διαστάσεων που συμβαίνει μετά το μετασχηματισμό του συστήματος συντεταγμένων, αποκόπτονται οι τρεις λιγότερο σημαντικές διαστάσεις και το εκπαιδευόμενο μεταπρότυπο είναι 3Δ. Τα σχήματα 4.28 και 4.30 συγκεντρώνουν τη μέση πορεία σύγκλισης, για δέκα τρεξίματα, των διάφορων εκδοχών που δοκιμάστηκαν.

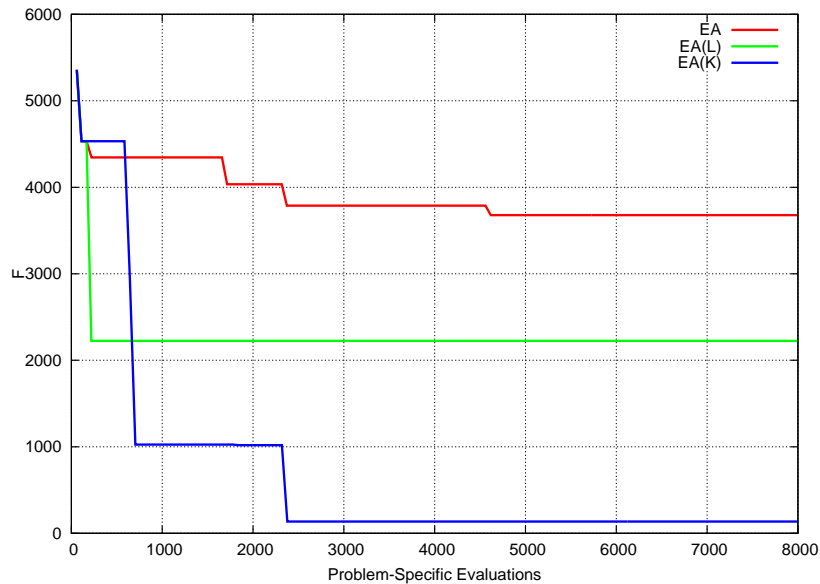


Σχήμα 4.28: Απόδοση των μεθόδων EA, EA(L) και EA(K) για το πρόβλημα του υποστηρίγματος δύο ράβδων.

Στο συγκεκριμένο πρόβλημα, ο απλός EA συναντά δυσκολία στο να βελτιώσει τις ήδη υπάρχουσες λύσεις. Αυτό φαίνεται στα διαγράμματα του σχήματος 4.29, όπου παρουσιάζονται ενδεικτικά οι πορείες σύγκλισης των εκδοχών EA, EA(L) και EA(K) για τρεις διαφορετικές τιμές της γεννήτριας τυχαίων αριθμών.

Όπως δείχνουν τα διαγράμματα για τα συγκεκριμένα τρεξίματα, αλλά και τα διαγράμματα μέσης πορείας σύγκλισης, η εφαρμογή γραμμικής ΑσΚΣ στους τελεστές εξέλιξης προκαλεί μικρή βελτίωση στην πορεία σύγκλισης στο συγκεκριμένο πρόβλημα. Αντιθέτως, η εφαρμογή ΑσΚΣ μέσω συνάρτησης πυρήνα βελτιώνει σημαντικά τα αποτελέσματα.

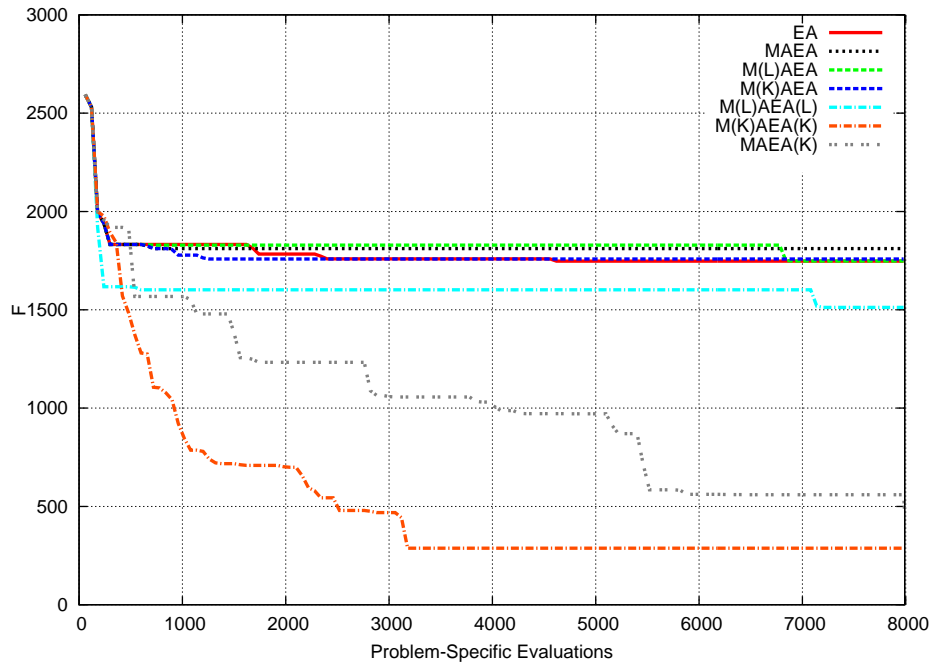




Σχήμα 4.29: Πρόβλημα υποστηρίγματος δύο ράβδων: Πορείες σύγκλισης των εκδοχών EA, EA(L) και EA(K) για τρεις διαφορετικές αρχικοποιήσεις της γεννήτριας τυχαίων αριθμών.

Η δυσκολία που αναφέρθηκε παραπάνω παρουσιάζεται και κατά την επίλυση με μεταπρότυπα. Μάλιστα, για το συγκεκριμένο πρόβλημα, η εκδοχή MAEA δίνει ελαφρώς χειρότερη πορεία σύγκλισης, όπως φαίνεται στο σχήμα 4.30. Πιθανότατα για αυτό ευθύνεται το σύνολο των περιορισμών, οι οποίοι “αποπροσανατολίζουν” τις προβλέψεις των μεταπροτύπων. Η ίδια δυσκολία παρατηρήθηκε και με διαφορετικές ρυθμίσεις των παραμέτρων του λογισμικού EASY (διαφορετικός αριθμός δειγμάτων εκπαίδευσης, επιπλέον χαλάρωση των περιορισμών, μεταπρότυπα πολυωνυμικής προσέγγισης).

Από την άλλη, η εφαρμογή της ΑσκΣ με χρήση συνάρτησης πυρήνα βελτιώνει και την απόδοση του μεταπροτύπου. Τα καλύτερα αποτελέσματα προκύπτουν από την εκδοχή M(K)AEA(K), τόσο από άποψη υπολογιστικού χρόνου, όσο και ποιότητας των τελικών λύσεων. Η εκδοχή αυτή εντόπισε και την καλύτερη συνολικά λύση, την [1.9134, 1.3349, 0.10425, 0.08, 0.100598, 0.08] για την οποία η μάζα της κατασκευής παίρνει την τιμή  $F = 108$ .

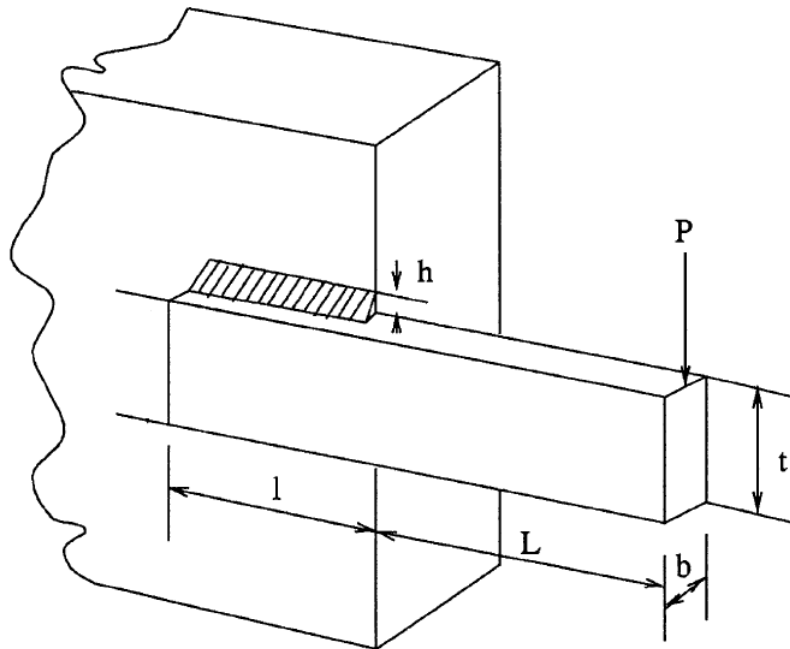


Σχήμα 4.30: Πρόβλημα υποστηρίγματος δύο ράβδων: Απόδοση των μεθόδων EA, MAEA, M(L)AEA, M(L)AEA(L) M(K)AEA και M(K)AEA(K).

## 4.6 Πρόβλημα Συγκόλλησης Δοκού

Το τέταρτο πρόβλημα που μελετήθηκε αφορά στη βελτιστοποίηση των προδιαγραφών μιας συγκόλλησης δοκού [19]. Η δοκός παραλαμβάνει φορτίο μέτρου  $6000 \text{ lbf}$ , σε σταθερή απόσταση  $L = 14 \text{ in}$  από το στέλεχος στήριξης. Στόχος είναι η ελαχιστοποίηση του κόστους (μονοκριτηριακή βελτιστοποίηση), ενώ ταυτόχρονα πρέπει να ικανοποιούνται περιορισμοί για τη διατμητική τάση, τις ορθές τάσεις λόγω κάμψης, το φορτίο λυγισμού καθώς και το μέγιστο βέλος μετατόπισης. Οι μεταβλητές σχεδιασμού είναι τέσσερις:

min		max	Περιγραφή
0.125	$x_1$	10	Πάχος συγκόλλησης $h$ (in)
0.1	$x_2$	10	Συνολικό μήκος συγκόλλησης $l$ (in)
0.1	$x_3$	10	Ύψος διατομής της δοκού $t$ (in)
0.1	$x_4$	10	Πλάτος διατομής δοκού $b$ (in)



Σχήμα 4.31: Βασικά μεγέθη του προβλήματος συγκόλλησης.

Η συνάρτηση που δίνει το κόστος της συγκόλλησης δίνεται από τον τύπο:

$$f(x) = 1.10471x_1^2x_2 + 0.04811x_3x_4(14 + x_2) \text{ (\$)} \quad (4.13)$$

Η βελτιστοποίηση υπόκειται στους παρακάτω έξι περιορισμούς:

$$g_1 = \tau(x) - 13000 \leq 0 \text{ (διατμητική τάση)} \quad (4.14)$$

$$g_2 = \sigma(x) - 30000 \leq 0 \text{ (ορθή τάση)} \quad (4.15)$$

$$g_3 = x_1 - x_4 \leq 0 \quad (4.16)$$

(πάχος δοκού μεγαλύτερο από αυτό της συγκόλλησης)

$$g_4 = 0.10471x_1^2 + 0.04811x_3x_4(14 + x_2) - 5 \leq 0 \quad (4.17)$$

(επιπλέον περιορισμός για το κόστος υλικών)

$$g_5 = \delta(x) - 0.25 \leq 0 \text{ (βέλος μετατόπισης)} \quad (4.18)$$

$$g_6 = 6000 - P_c(x) \leq 0 \text{ (φορτίο λυγισμού)} \quad (4.19)$$



Τα ενδιάμεσα μεγέθη δίνονται από τις σχέσεις:

$$\tau(x) = \sqrt{\tau'^2 + 2\tau'\tau''\frac{x_2}{2R} + \tau''^2} \quad (\text{συνολική διατμητική τάση}) \quad (4.20)$$

$$\tau' = \frac{6000}{\sqrt{2x_1x_2}} \quad (\text{διατμητική τάση λόγω κάθετης δύναμης}) \quad (4.21)$$

$$\tau'' = \frac{MR}{J} \quad (\text{διατμητική τάση λόγω ροπής}) \quad (4.22)$$

$$M = 6000(14 + \frac{x_2}{2}) \quad (\text{ροπή ως προς το μεσο της συγκόλλησης}) \quad (4.23)$$

$$R = \sqrt{\frac{x_2^2}{4} + (\frac{x_1 + x_3}{2})^2} \quad (4.24)$$

$$J = 2\sqrt{2x_1x_2}[\frac{x_2^2}{12} + (\frac{x_1 + x_3}{2})^2] \quad (\text{πολική ροπή αδρανείας}) \quad (4.25)$$

$$\sigma(x) = \frac{504000}{x_4x_3^2} \quad (4.26)$$

$$\delta(x) = \frac{2.1952}{x_3^2x_4} \quad (4.27)$$

$$P_c(x) = 64746.022(1 - 0.0282346x_3)x_3x_4^3 \quad (4.28)$$

Το μεγάλο πλήθος περιορισμών καθιστά αρκετά δύσκολο για τον αλγόριθμο να βρει έστω και μία αποδεκτή λύση ώστε να ξεκινήσει η διαδικασία της εξέλιξης. Για το λόγο αυτό έγινε χαλάρωση στα όρια παράβασης των περιορισμών.

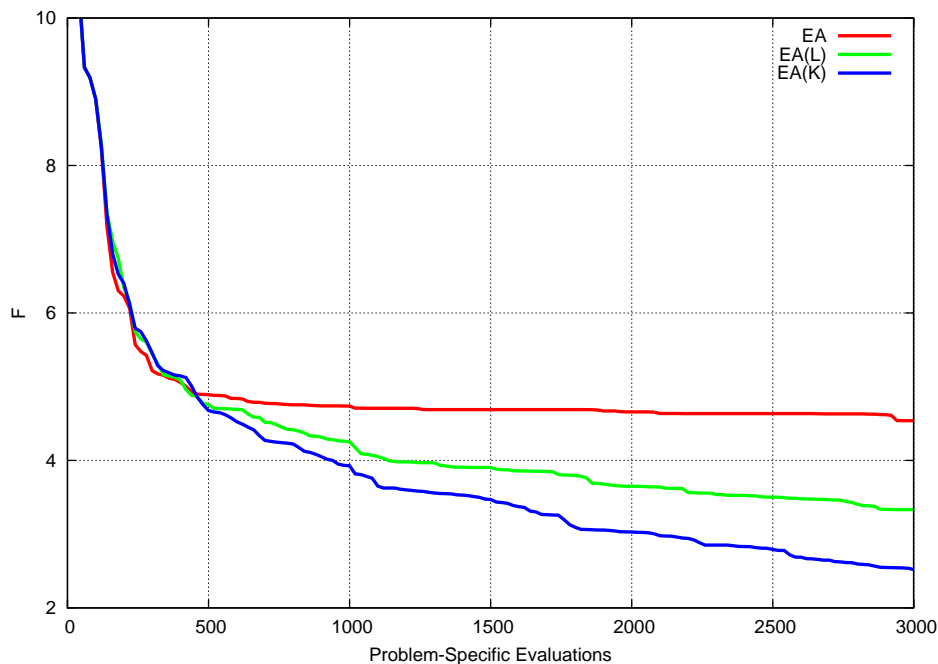
Όπως αναλύθηκε στο κεφάλαιο 2, για μία υποψήφια λύση που σέβεται τους περιορισμούς ισχύει  $c_k \leq d_k$ . Με  $d_k$  συμβολίζονται τα “τυπικά” όρια των περιορισμών (όπου στο συγκεκριμένο πρόβλημα είναι  $d_k=0$ ), και υποψήφιες λύσεις που παραβιάζουν αυτά τα όρια επιβαρύνονται με ένα “προστιμο” κατά τη διαδικασία απόδοσης βαθμωτού κόστους. Η επιβάρυνση αυξάνει εκθετικά με το μέγεθος του  $c_k$ , σύμφωνα με τη σχέση 2.4. Αυτό συμβαίνει μέχρι κάποιο “χαλαρωμένο” όριο  $d_k^*$ . Λύσεις των οποίων οι περιορισμοί υπερβαίνουν το όριο  $d_k^*$ , τιμωρούνται με ποινή θανάτου.

Εδώ, το μέγιστο όριο στο οποίο επιβάλλεται η ποινή θανάτου σε μια υποψήφια λύση τέθηκε ίσο με 100, ενώ ο παράγοντας εκθετικής επιβάρυνσης τέθηκε ίσος με 3. Με αυτό τον τρόπο δίνεται ευκαιρία σε πολλές λύσεις που θα απορρίπτονταν κατευθείαν από τον αλγόριθμο να συμμετάσχουν στην εξέλιξη.

Επίσης, για να είναι εγγυημένη η εκκίνηση της εξέλιξης και να είναι δυνατή η ταυτόχρονη εκκίνηση της ΑσκΣ σε όλα τα τρεξίματα, εισάγονται στην

πρώτη γενιά δύο σχετικά “κακές” αλλά αποδεκτές λύσεις, οι [1.0276, 5.9783, 4.79378, 1.03066] [1.49465, 2.39945, 3.2551, 1.838737] με κόστος 11.723 και 10.648 αντίστοιχα.

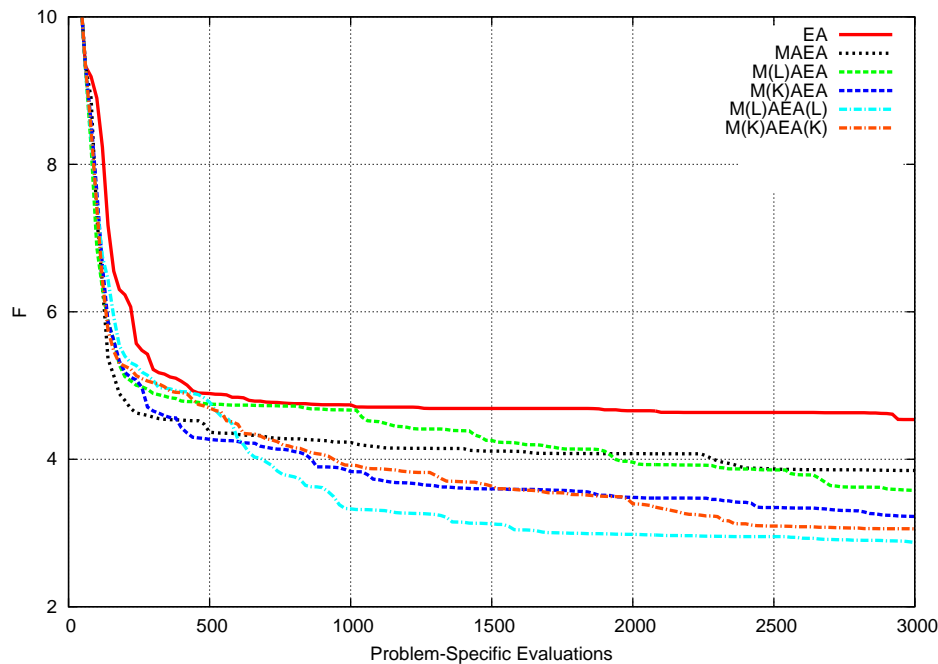
Η επίλυση του προβλήματος έγινε με έναν (10,20)EA, ενώ όταν πραγματοποιείται εφαρμογή της ΑσκΣ στους εξελικτικούς τελεστές, αυτή ξεκινά από την τρίτη γενιά. Το μεταπρότυπο ξεκινά όταν συμπληρωθούν 50 εγγραφές στη βάση δεδομένων (εκ των οποίων τουλάχιστον 15 ικανοποιούν τους περιορισμούς). Αποκόπτονται δύο διαστάσεις, επομένως κάθε μεταπρότυπο εκπαιδεύεται σε χώρο δύο διαστάσεων.



Σχήμα 4.32: Πρόβλημα Συγκόλλησης Δοκού: Απόδοση των μεθόδων EA, EA(L) και EA(K).

Κωδικοποίηση	# Τρεξιμάτων	Σχήμα Διασταύρωσης	Πιθανότητα Μετάλλαξης $P_m$
Πραγματική	10	Intermediate Recombination	0.03

Πίνακας 4.7: Επιπλέον χαρακτηριστικά των δοκιμών.



Σχήμα 4.33: Πρόβλημα Συγκόλλησης Δοκού: Απόδοση των μεθόδων EA, MAEA, M(L)AEA, M(L)AEA(L) M(K)AEA και M(K)AEA(K).

Στα σχήματα 4.32 και 4.33 απεικονίζεται η μέση πορεία σύγκλισης για όλες τις εκδοχές που δοκιμάστηκαν. Για την κατασκευή των καμπυλών αυτών λήφθηκαν υπόψη μόνο αποδεκτές λύσεις. Η απόδοση του τυπικού EA στο συγκεκριμένο πρόβλημα είναι αρκετά κακή, αφού προκαλεί μικρή βελτίωση των αρχικών λύσεων, ενώ η σύγκλιση γίνεται πολύ αργή από πολύ νωρίς. Αντιθέτως, οι εκδοχές EA(L), EA(K) συνεχίζουν να βρίσκουν καλύτερες λύσεις καθόλη τη διάρκεια της εξέλιξης. Η χρήση μεταπροτύπων βελτιώνει την εξελικτική πορεία. Η γραμμική ΑσκΣ στην εκπαίδευση των μεταπροτύπων δε βοηθά ιδιαίτερα στην καλύτερη πορεία της εξέλιξης, κάτι που συμβαίνει με την ΑσκΣ μέσω συνάρτησης πυρήνα. Καλύτερα αποτελέσματα δίνουν οι εκδοχές M(L)AEA(L), M(K)AEA(K). Η καλύτερη λύση βρέθηκε από τη εκδοχή M(K)AEA(K), στο σημείο  $[0.19419, 3.7268, 9.488, 0.20358]$ , με κόστος 1.803.

## 4.7 Πρόβλημα Βελτιστοποίησης Αεροτομής Τεσσάρων Στοιχείων

Το συγκεκριμένο πρόβλημα αεροδυναμικής βελτιστοποίησης αφορά στη διαμόρφωση διάταξης αεροτομής αποτελούμενης από τέσσερα στοιχεία. Διατάξεις τέτοιου είδους χρησιμοποιούνται στα περισσότερα επιβατικά αεροπλάνα κατά τις φάσεις προσγείωσης και απογείωσης παρέχοντας επιπλέον άνωση, για αυτό και ονομάζονται υπεραντωτικές (high-lift devices).

Η διάταξη αποτελείται από μια μεγάλη (κύρια/main) αεροτομή και τρεις μικρότερες, η μία μπροστά (slat) και οι άλλες δύο πίσω από την κύρια (flaps).

Το πρόβλημα μοντελοποιείται ως 2Δ μόνιμη ροή μη-συνεκτικού συμπιεστού ρευστού. Η γωνία της επ' άπειρο ροής θεωρείται ίση με  $\alpha = 10^\circ$  και ο αριθμός Mach  $M_\infty = 0.2$ . Στόχος είναι η μεγιστοποίηση του συντελεστή άνωσης  $C_L$  (επειδή το λογισμικό *EASY* πραγματοποιεί ελαχιστοποίηση συναρτήσεων, η συνάρτηση στόχου γράφεται ως  $F = -C_L$ ), ενώ οι μεταβλητές του προβλήματος είναι εννιά: οι μετατοπίσεις κατά x και y και οι γωνίες στροφής των τριών μικρότερων πτερυγίων σε σχέση με μία αρχική θέση. Η θέση και η στροφή της κύριας αεροτομής παραμένουν σταθερές. Σταθερό παραμένει και το σχήμα όλων των αεροτομών.

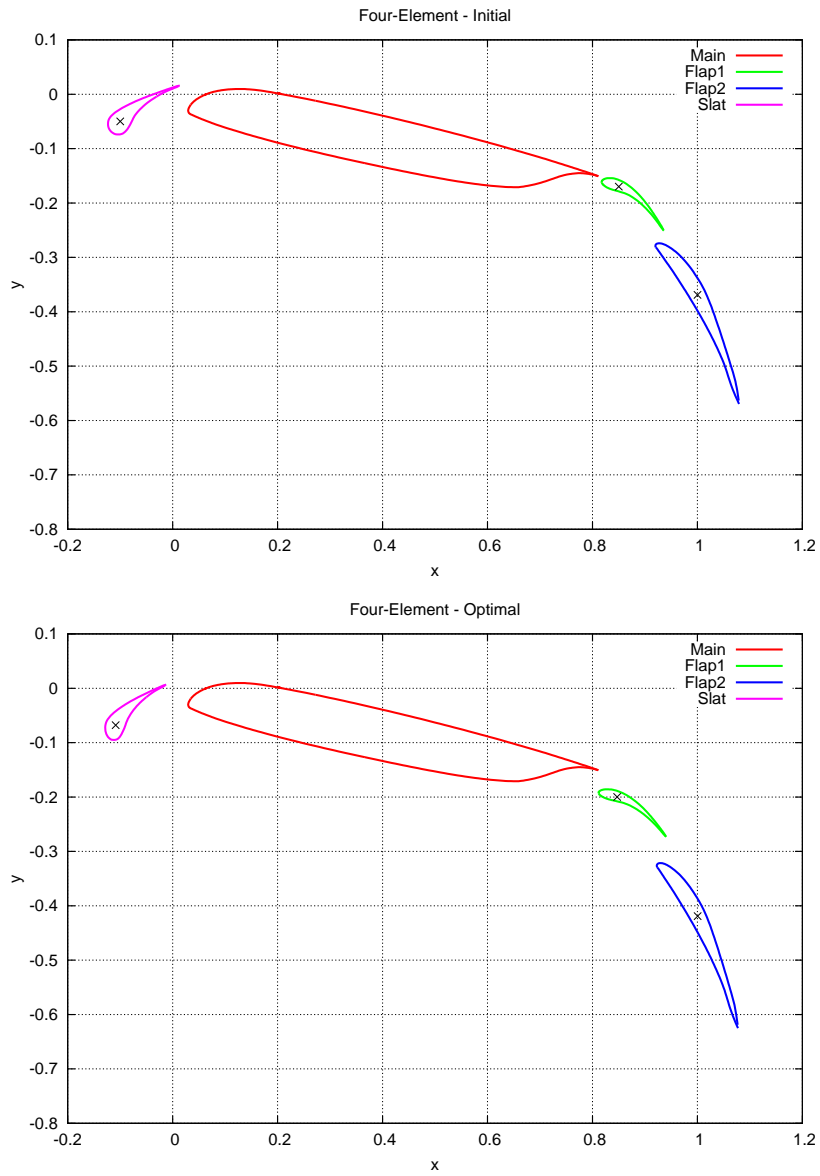
min		max	Περιγραφή
-0.15	$\theta_1$	0.15	Στροφή του flap 1(rad)
-0.05	$x_1$	0.05	Οριζόντια μετατόπιση flap 1(m)
-0.05	$y_1$	0.05	Κάθετη μετατόπιση flap 1(m)
-0.15	$\theta_2$	0.15	Στροφή του flap 2(rad)
-0.05	$x_2$	0.05	Οριζόντια μετατόπιση flap 2(m)
-0.05	$y_2$	0.05	Κάθετη μετατόπιση flap 2(m)
-0.15	$\theta_3$	0.15	Στροφή του slat(rad)
-0.05	$x_3$	0.05	Οριζόντια μετατόπιση slat(m)
-0.05	$y_3$	0.05	Κάθετη μετατόπιση slat(m)

Πίνακας 4.8: Μεταβλητές του προβλήματος.

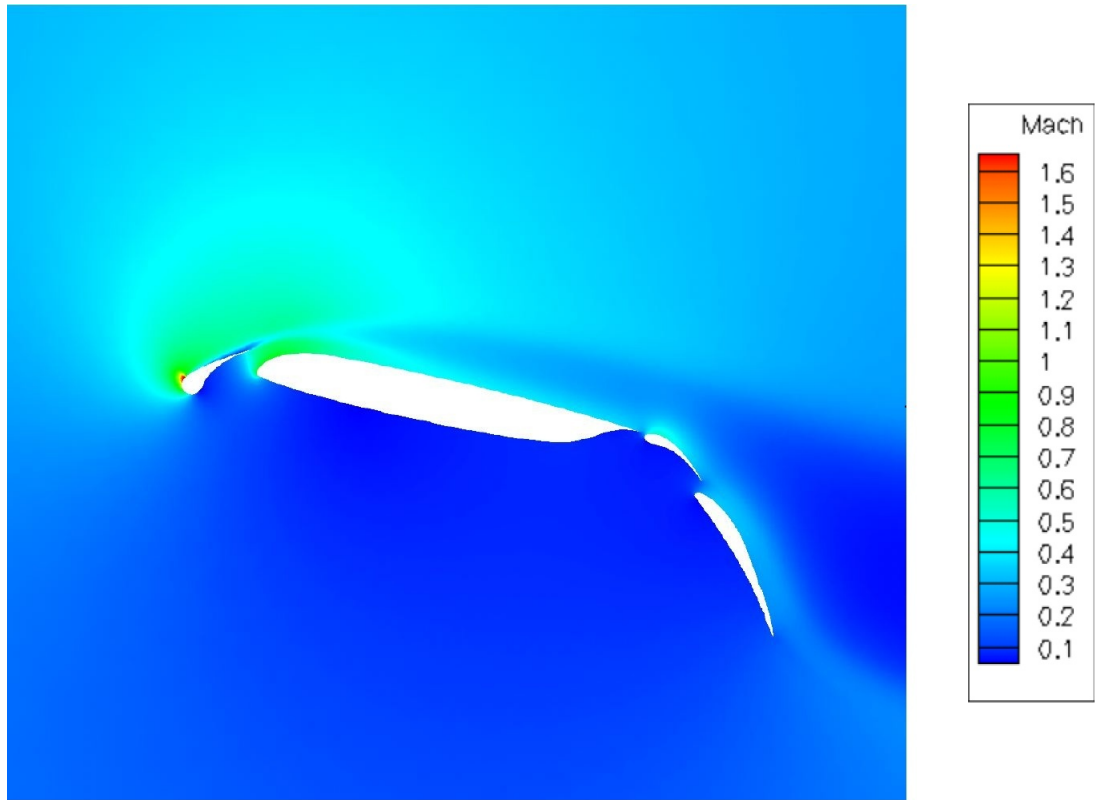
Για την αξιολόγηση μιας υποψήφιας λύσης πρέπει να λυθούν οι εξισώσεις πεδίου ατρίβους ροής (εξισώσεις Euler). Το λογισμικό που χρησιμοποιήθηκε

έχει αναπτυχθεί στη ΜΠΥΡ&Β, και είναι κώδικας υψηλής παράλληλης απόδοσης, ο οποίος τρέχει σε κάρτες γραφικών (Graphics Processing Units - GPUs)[11] [12]. Ο κώδικας δημιουργεί πλέγμα τριγωνικών στοιχείων γύρω από τις αεροτομές, και, στη συνέχεια, επιλύει τις αντίστοιχες εξισώσεις πεπερασμένων όγκων για κάθε κόμβο.

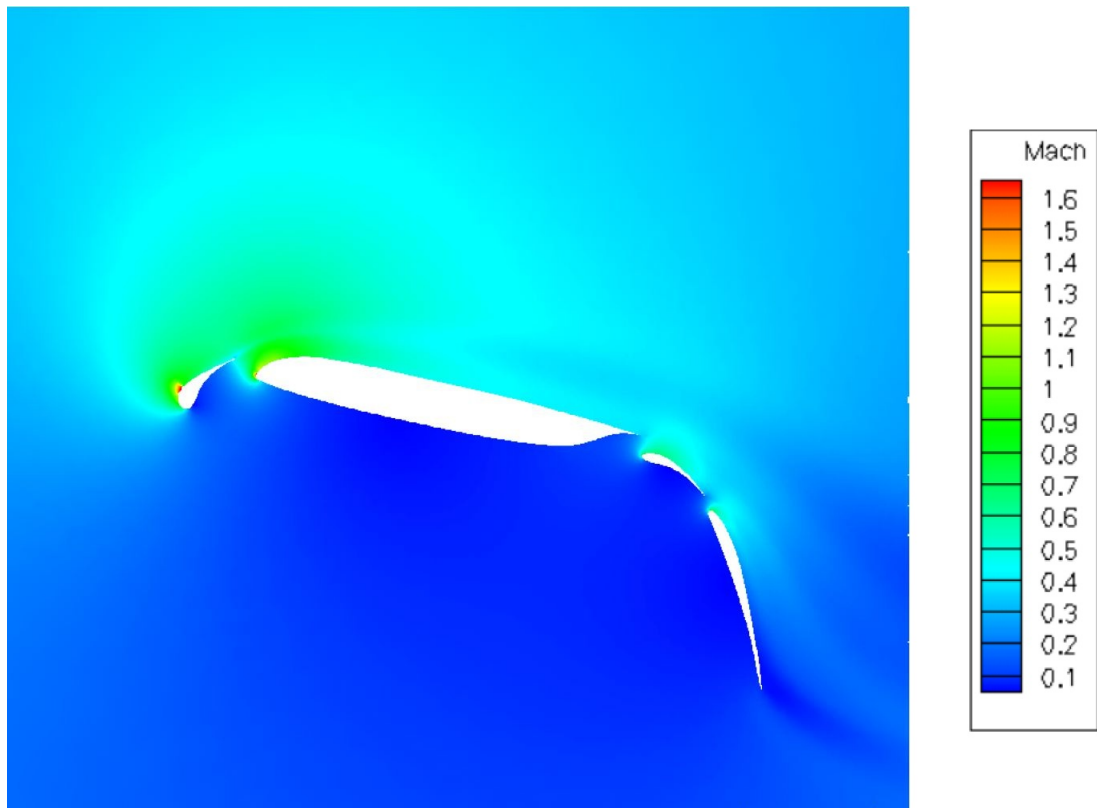
Το πλήθος των κόμβων που αποτελούν το πλέγμα κυμαίνεται από 39000-71000, με 55000 κόμβους κατά μέσο όρο. Αντίστοιχα, το πλήθος των τριγώνων του πλέγματος κυμαίνεται από 77000-142000 με 110000 τρίγωνα κατά μέσο όρο. Το μεγάλο εύρος του πλήθους των τριγώνων οφείλεται στον τρόπο που διαμορφώνεται το πλέγμα: ο κώδικας επιλέγει κάθε φορά να αυξήσει την πυκνότητα των τριγώνων κοντά στα στερεά όρια, όπου τα μεγέθη της ροής παρουσιάζουν μεγαλύτερη διακύμανση. Με αυτόν τον τρόπο, επιτυγχάνεται εξοικονόμηση υπολογιστικών πόρων ενώ διατηρείται η ακρίβεια του αποτελέσματος.



Σχήμα 4.34: Διατάξεις των αεροτομών για την αρχική ( $\vec{x} = \vec{0}$ ,  $C_L = 3.6761$ ), και τη βέλτιστη ευρεθείσα λύση ( $C_L = 4.01559$ ). Με X σημειώνονται τα σημεία ως προς τα οποία γίνεται η στροφή των πτερυγίων.



Σχήμα 4.35: Πεδίο ροής για την αρχική λύση.



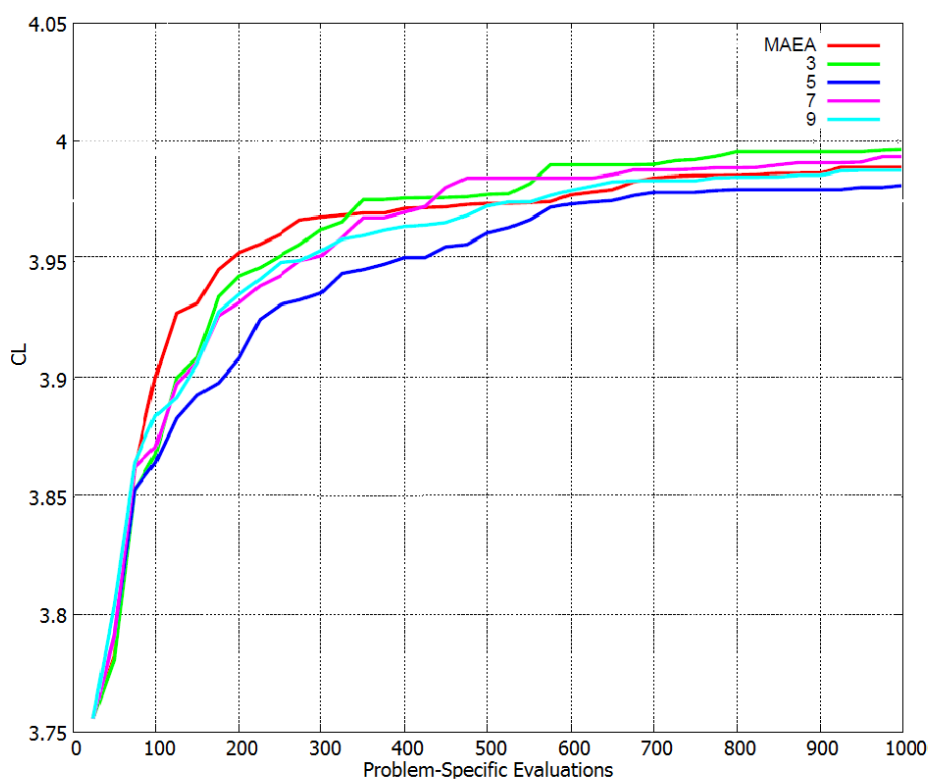
Σχήμα 4.36: Πεδίο ροής για τη βέλτιστη ευρεθείσα λύση.

Πραγματοποιήθηκε βελτιστοποίηση του προβλήματος κάνοντας χρήση των μεθόδων ΜΑΕΑ και Μ(Κ)ΑΕΑ(Κ). Για την τελευταία έγιναν τρεξίματα με μεταπρότυπα 3,5,7 και 9 διαστάσεων, προκειμένου να διαπιστωθεί το πώς επηρεάζει αυτή η παράμετρος την επίδοση του αλγορίθμου.

Χρησιμοποιήθηκε μία παραλλαγή της μεθόδου ΑσΚΣ με χρήση συναρτήσεων πυρήνα. Συγκεκριμένα, το  $\sigma$ , δηλαδή η ακτίνα της συνάρτησης πυρήνα που χρησιμοποιείται για την ανάλυση του δείγματος, μπορεί να μην είναι σταθερό, αλλά να μεταβάλλεται από γενιά σε γενιά. Όταν συμβαίνει αυτό, το  $\sigma$  τίθεται ίσο με ένα προς την τυπική απόκλιση των αποστάσεων στο σύνολο των απογόνων (ενώ η τιμή που χρησιμοποιείται κανονικά είναι  $\sigma^2 = 0.001$ ). Όπως αναφέρθηκε στο κεφάλαιο 3, η υπερπαράμετρος (hyperparameter)  $\sigma$  έχει μεγάλη σημασία για τον εντοπισμό χαρακτηριστικών στο δείγμα. Με τη ρύθμιση αυτή γίνεται μια απόπειρα προσαρμογής της συνάρτησης πυρήνα, ώστε αυτή να περιγράφει αποδοτικότερα τα σύνολα δεδομένων που εξετάζονται.



Τα τρεξίματα έγιναν για τρεις διαφορετικές τιμές της γεννήτριας τυχαίων αριθμών, και τα αποτελέσματα συνδυάστηκαν ώστε να φτιαχτεί η μέση πορεία σύγκλισης κάθε μεθόδου. Οι βασικές παράμετροι που τέθηκαν είναι:  $\mu=15$   $\lambda=25$ . Η εφαρμογή της μεθόδου Kernel PCA στους εξελικτικούς τελεστές ξεκινά από την πρώτη γενιά. Για την εκκίνηση της προσεγγιστικής προ-αξιολόγησης απαιτούνται 60 εγγραφές στη βάση δεδομένων, και κάθε μεταπρότυπο εκπαιδεύεται με 15-30 δείγματα εκπαίδευσης. Οι ακριβείς αξιολογήσεις που πραγματοποιούνται σε κάθε γενιά είναι 2 έως 5.

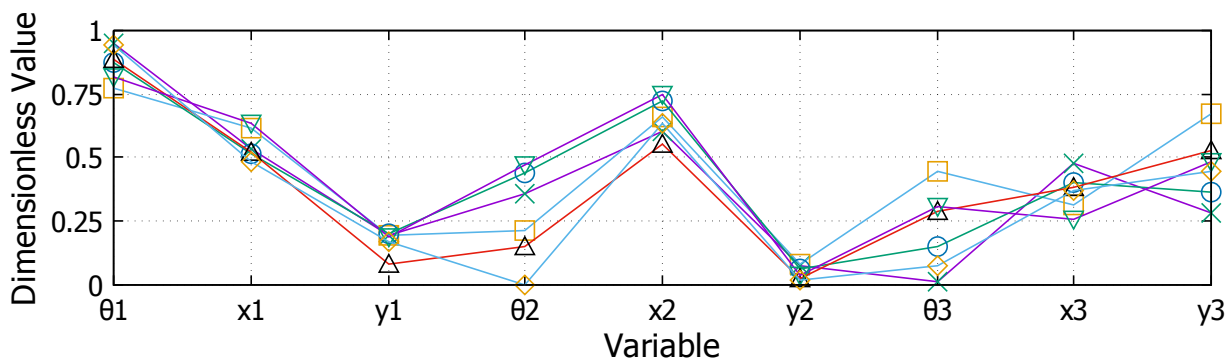


Σχήμα 4.37: Μέση πορεία σύγκλισης των μεθόδων MAEA και M(K)AEA(K), με μεταπρότυπα διάστασης 3, 5, 7 και 9.

Κωδικοποίηση	Σχήμα Διασταύρωσης	Πιθανότητα Μετάλλαξης $P_m$
Πραγματική	Intermediate Recombination	0.05

Πίνακας 4.9: Επιπλέον χαρακτηριστικά των δοκιμών.

Σε γενικές γραμμές, η μέθοδος M(K)AEA(K) δίνει οριακά καλύτερα αποτελέσματα. Η μικρή διαφορά των καμπυλών πιθανότατα οφείλεται στο μικρό περιθώριο βελτιστοποίησης του συγκεκριμένου προβλήματος. Ενδιαφέρον παρουσιάζει το γεγονός ότι καλύτερη μέση πορεία σύγκλισης είχε το μεταπρότυπο που απέκοπτε το μεγαλύτερο αριθμό διαστάσεων. Από αυτό βρέθηκε και η καλύτερη λύση  $\vec{x}=[0.09145, -0.00304, -0.03005, -0.0283, 0.00028, -0.05, -0.1414, -0.00879, -0.0179]$ , με  $C_L = 4.01559$ , και απεικονίζεται στο σχήμα 4.34, ενώ στα σχήματα 4.35 και 4.36 απεικονίζονται οι κατανομές του αριθμού Mach για την αρχική και την καλύτερη λύση που βρέθηκε.



Σχήμα 4.38: Οι έξι καλύτερες λύσεις, σε αδιαστατοποιημένη μορφή, όπως προέκυψαν από το σύνολο των τρεξίμάτων.

Το σχήμα 4.38 συγκεντρώνει σε μορφή διαγράμματος τις έξι καλύτερες λύσεις για τα διάφορα τρεξίματα, οι οποίες αντιστοιχούν σε συντελεστή άνωσης  $C_L \approx 4$ . Οι μεταβλητές έχουν αδιαστατοποιηθεί, ώστε η ελάχιστη και μέγιστη τιμή τους να ισούται με  $b_i = 0$  και  $b_i = 1$  αντίστοιχα. Αυτός ο τρόπος αναπαράστασης των λύσεων διευκολύνει την μεταξύ τους σύγκριση και τον εντοπισμό κοινών στοιχείων.

Παρατηρείται ότι για τις μεταβλητές  $\theta_1, x_1, y_1, y_2$ , οι λύσεις συγκεντρώνονται κοντά σε μια συγκεκριμένη τιμή, και παρουσιάζουν μικρή διακύμανση. Συνεπώς, οι μεταβλητές αυτές, που περιγράφουν τη σχετική θέση του flap 1 και την κατα  $y$  μετατόπιση του flap 2, είναι οι πιο σημαντικές, και οι τιμές τους συνιστούν βασικά χαρακτηριστικά που πρέπει να διαθέτουν οι καλύτερες λύσεις του προβλήματος. Από την άλλη, παρατηρείται μεγαλύτερη διακύμανση ως προς τις μεταβλητές  $\theta_2, x_2, \theta_3, x_3, y_3$ , δηλαδή για μια καλή λύση η σχετική θέση του slat, και η στροφή και κατά  $x$  μετατόπιση του flap 2 είναι λιγότερο αυστηρά καθορισμένες.

## Κεφάλαιο 5

# Ανακεφαλαίωση - Συμπεράσματα

Αντικείμενο της διπλωματικής εργασίας ήταν η παρουσίαση και πιστοποίηση εκδοχών των συμβατικών ΕΑ και ΜΑΕΑ, οι οποίες αξιοποιούν μεθόδους Ανάλυσης σε Κύριες Συνιστώσες για να αυξήσουν την αποτελεσματικότητά τους.

Στο πλαίσιο αυτό, έγινε αρχικά παρουσίαση των ΕΑ και εξηγήθηκαν οι βασικές λειτουργίες τους. Έμφαση δόθηκε στους αλγορίθμους διασταύρωσης και μετάλλαξης, δηλαδή στους τελεστές που οδηγούν τη διαδικασία της εξέλιξης.

Στη συνέχεια, αναλύθηκε η λειτουργία και χρησιμότητα των μεταπρότυπων. Τα μεταπρότυπα, τα οποία είναι προσεγγιστικά μοντέλα των συναρτήσεων στόχων που κατασκευάζονται από τις υπάρχουσες αξιολογήσεις, επιτρέπουν την προ-αξιολόγηση των υποψήφιων λύσεων. Με αυτόν τον τρόπο, είναι δυνατή η ανάδειξη των καλύτερων λύσεων, εξοικονομώντας ταυτόχρονα κλησεις στο, ακριβές αλλά ακριβό, λογισμικό αξιολόγησης.

Κατόπιν, έγινε παρουσίαση της Ανάλυσης σε Κύριες Συνιστώσες: όταν η μέθοδος αυτή εφαρμοστεί σε κάποιο δείγμα, ορίζεται ένα νέο σύνολο μεταβλητών, οι οποίες συγκεντρώνουν με αποδοτικότερο τρόπο τη διακύμανση των δεδομένων. Αναλύονται δύο διαφορετικές μέθοδοι ΑσΚΣ, η γραμμική και η ΑσΚΣ με χρήση συναρτήσεων (μη-γραμμικού)πυρήνα. Σύμφωνα με την πρώτη, το νέο σύνολο μεταβλητών παράγεται από έναν απλό ορθογώνιο μετασχηματισμό των αρχικών, ενώ το σύνολο μεταβλητών που ορίζει η δεύτερη προκύπτει από έναν μη-γραμμικό μετασχηματισμό σε ένα χώρο υψηλότερης

διάστασης, ο οποίος ονομάζεται χώρος χαρακτηριστικών. Η εξαγωγή πληροφορίας από τον χώρο χαρακτηριστικών, ο οποίος έχει πρακτικά άπειρη διάσταση, γίνεται δυνατή μέσω μιας αλγεβρικής τεχνικής που ονομάζεται “τέχνασμα πυρήνα” (kernel trick).

Κάθε μία από τις μεθόδους αυτές, μπορεί να εφαρμοστεί στο σύνολο των απογόνων κάθε γενιάς ενός εξελικτικού αλγορίθμου, εντοπίζοντας σημαντικά χαρακτηριστικά του χώρου σχεδιασμού. Τα αποτελέσματα είναι δυνατόν να αξιοποιηθούν με δύο διαφορετικούς τρόπους για κάθε μέθοδο: με εφαρμογή των τελεστών εξέλιξης σε γονείς που ορίζονται ως προς το νέο σύνολο μεταβλητών που προκύπτει, αυξάνοντας έτσι την πιθανότητα εμφάνισης ποιοτικότερων απογόνων, καθώς και με αποκοπή των λιγότερο σημαντικών διαστάσεων κατά την εκπαίδευση των μεταπροτύπων, βελτιώνοντας έτσι την ακρίβειά τους.

Σύμφωνα με τα παραπάνω, ορίζονται πολλές παραλλαγές των βασικών ΕΑ και ΜΑΕΑ, οι οποίες μπορούν να ονομαστούν αναλόγως με τη μέθοδο ΑσκΣ που χρησιμοποιείται, και τον τρόπο αξιοποίησής της: ΕΑ(L), ΕΑ(K), Μ(L)ΑΕΑ, Μ(L)ΑΕΑ(L), Μ(K)ΑΕΑ, Μ(K)ΑΕΑ(K).

Στη συνέχεια, οι μέθοδοι αυτές χρησιμοποιήθηκαν για την επίλυση έξι προβλημάτων βελτιστοποίησης: τριών μαθηματικών προβλημάτων, δύο προβλημάτων ψεύδο-μηχανολογικού σχεδιασμού και ενός προβλήματος αεροδυναμικής βελτιστοποίησης.

Σχεδόν σε όλα τα προβλήματα που παρουσιάστηκαν, η υποβοήθηση των εξελικτικών τελεστών με χρήση ΑσκΣ επιφέρει σημαντική βελτίωση στην πορεία σύγκλισης του αλγορίθμου. Μοναδική εξαίρεση σε αυτό, αποτελεί η εκδοχή ΕΑ(K) στην περίπτωση της παραλλαγής της 20Δ συνάρτησης Rastrigin. Όπως φάνηκε και από τη μελέτη των μαθηματικών συναρτήσεων, η βελτίωση αυτή είναι ιδιαίτερα σημαντική στην περίπτωση των “κακώς τοποθετημένων” προβλημάτων, τα οποία δυσκολεύουν ιδιαίτερα τους συμβατικούς ΕΑ, ενώ είναι αυτά που συναντώνται συχνότερα στα προβλήματα βελτιστοποίησης μεγάλης κλίμακας.

Η αποκοπή διαστάσεων χαμηλότερης σημασίας κατά την εκπαίδευση των μεταπροτύπων βελτιώνει γενικά την απόδοσή τους. Ειδικά στις περιπτώσεις του προβλήματος υποστηρίγματος δύο ράβδων, καθώς και στην περίπτωση της 20Δ παραλλαγής της συνάρτησης Rastrigin παρατηρείται ότι τα μεταπρότυπα που εκπαιδεύονται μόνο ως προς τις σημαντικότερες διαστάσεις (όπως αυτές υπολογίστηκαν), έχουν πολύ αυξημένη ακρίβεια σε σχέση με τα συμβατικά μεταπρότυπα της εκδοχής ΜΑΕΑ.

Ο συνδυασμός των μεθόδων, σχεδόν πάντα παράγει την καλύτερη πορεία

σύγκλισης. Κάτι τέτοιο ισχύει για την περίπτωση της συνάρτησης Ackley, το πρόβλημα Fonseca-Fleming, τα δύο παραδείγματα μηχανολογικού σχεδιασμού, αλλά και την περίπτωση της υπεραντωτικής διάταξης. Μάλιστα, στην περίπτωση του προβλήματος υποστηρίγματος δύο ράβδων, φαίνεται να υπάρχει ιδιαίτερη συνέργεια μεταξύ των δύο μεθόδων, γιατί ενώ η μέθοδος  $M(K)AEA$  από μόνη της δεν βελτιώνει ιδιαίτερα την επίδοση του αλγορίθμου (τουλάχιστον η εκδοχή του αλγορίθμου που δοκιμάστηκε), η εκδοχή  $M(K)AEA(K)$  δίνει την καλύτερη πορεία σύγκλισης.

Είναι λοιπόν εμφανές, ότι οι προτεινόμενες μέθοδοι επιταχύνουν σημαντικά τη σύγκλιση των συμβατικών  $EA$  και  $MAEA$  σε μεγάλο εύρος προβλημάτων βελτιστοποίησης. Η παρατήρηση αυτή ισχύει για προβλήματα ενός αλλά και πολλών στόχων, όπως φάνηκε από τη μελέτη του προβλήματος Fonseca - Fleming, όπου οι προτεινόμενες μέθοδοι κατέληξαν πολύ πιο γρήγορα σε ποιοτικότερα σύνολα λύσεων. Επίσης, η υποβοήθηση των  $EA$  και  $MAEA$  με  $AsKS$  αποδίδει και σε προβλήματα με περιορισμούς, όπως έδειξαν οι περιπτώσεις των “ψευδο-μηχανολογικών” προβλημάτων βελτιστοποίησης.

Συγκρίνοντας τους δύο τρόπους διεξαγωγής  $AsKS$ , είναι εμφανές ότι η θετική επίδραση στην απόδοση του αλγορίθμου τονίζεται ακόμα περισσότερο με τη χρήση συναρτήσεων πυρήνα. Κάτι τέτοιο είναι αναμενόμενο, εφόσον πρόκειται για γενίκευση της γραμμικής μεθόδου, με επιπλέον υπολογιστικά βήματα ώστε να εντοπίζει τα μη-γραμμικά χαρακτηριστικά του χώρου σχεδιασμού. Σε όλα σχεδόν τα μελετώμενα προβλήματα, η χρήση  $AsKS$  με συναρτήση πυρήνα υπερέχει της αντίστοιχης γραμμικής, βρίσκοντας καλύτερες λύσεις ταχύτερα.

# Παράρτημα Α

## Προβλήματα Πολλών Στόχων

Όπως αναφέρθηκε, στις λύσεις προβλημάτων πολυκριτηριακής βελτιστοποίησης δεν αντιστοιχεί μία βαθμωτή ποσότητα κόστους, αλλά ένα διάνυσμα που εκφράζει τις τιμές των διάφορων συναρτήσεων στόχων. Αποτέλεσμα της διαδικασίας βελτιστοποίησης είναι ένα μέτωπο από μη-κυριαρχούμενες λύσεις, το οποίο ιδανικά θα ταυτίζεται με το σύνολο Pareto .

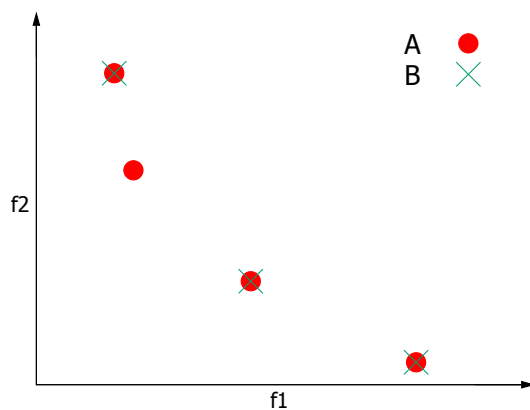
Συνεπώς, για να είναι δυνατή η σύγκριση μεταξύ διαφορετικών μεθόδων βελτιστοποίησης πρέπει να εισαχθούν έννοιες που κάνουν δυνατή τη σύγκριση μετώπων λύσεων[28].

### A.1 Σχέσεις Κυριαρχίας Μετώπων

**Ασθενής Κυριαρχία:** Ορίζεται ότι ένα μέτωπο λύσεων  $A$ , κυριαρχεί ασθενώς ενός μετώπου  $B$  ( $A O_W B$ ), αν και μόνο αν υπάρχει τουλάχιστον ένα στοιχείο του  $A$  που δεν ανήκει στο  $B$ , ενώ ταυτόχρονα όλα τα στοιχεία του  $B$  είναι ίσα ή κυριαρχούνται από αυτά του  $A$ . Ο ορισμός διατυπώνεται μαθηματικά ως:

$$A O_W B \iff ND(A \cup B) = A \wedge A \neq B \quad (A.1)$$

Η σχέση αυτή απεικονίζεται στο σχήμα A.1.

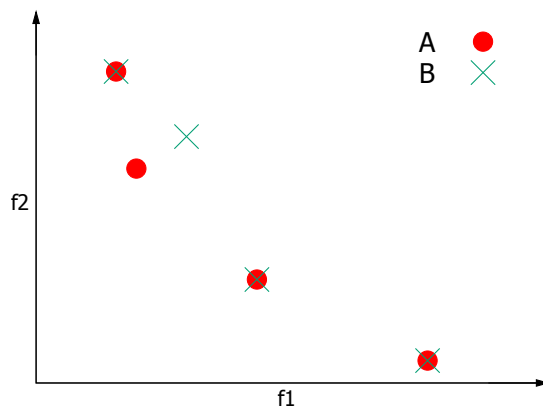


Σχήμα A.1: Γραφική απεικόνιση περίπτωσης ασθενούς κυριαρχίας ( $A O_W B$ ).

**Ισχυρή Κυριαρχία:** Το μέτωπο A κυριαρχεί ισχυρά του μετώπου B ( $A O_S B$ ), εάν, όλα τα στοιχεία του B είναι ίσα ή κυριαρχούνται από αυτά του A, και, ταυτόχρονα, υπάρχει τουλάχιστον ένα στοιχείο του A το οποίο κυριαρχεί ενός στοιχείου του B:

$$A O_S B \iff ND(A \cup B) = A \cap B \setminus ND(A \cup B) \neq \emptyset \quad (A.2)$$

Η σχέση απεικονίζεται στο σχήμα A.1.

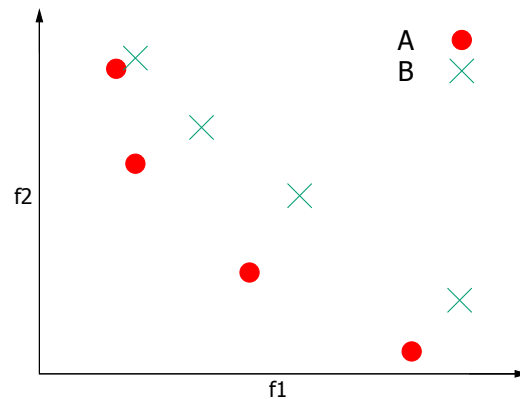


Σχήμα A.2: Γραφική απεικόνιση περίπτωσης ισχυρής κυριαρχίας ( $A O_S B$ ).

**Καθολική Κυριαρχία:** Το μέτωπο A κυριαρχεί καθολικά του μετώπου B ( $A O_C B$ ), εάν, όλα τα στοιχεία του B κυριαρχούνται από αυτά του A:

$$A O_C B \iff ND(A \cup B) = A \wedge B \cap (A \cup B) = \emptyset \quad (A.3)$$

Η σχέση απεικονίζεται στο σχήμα A.1.



Σχήμα A.3: Γραφική απεικόνιση περίπτωσης καθολικής κυριαρχίας ( $A O_C B$ ).

Οι έννοιες κυριαρχίας (outperformance) μεταξύ μετώπων που ορίστηκαν παραπάνω μπορούν να χρησιμοποιηθούν για τη σύγκριση και αξιολόγηση συνόλων λύσεων, ωστόσο δεν ποσοτικοποιούν με κάποιο τρόπο την ανωτερότητα ενός μετώπου σε σχέση με κάποιο άλλο. Για το σκοπό αυτό χρησιμοποιούνται οι μετρικές ποιότητας.

## A.2 Μετρικές Ποιότητας-Δείκτης Υπερόγκου

Με τον όρο μετρικές ποιότητας εννοούνται συναρτήσεις, οι οποίες αντιστοιχίζουν έναν πραγματικό αριθμό σε κάθε μέτωπο μη-κυριαρχούμενων λύσεων. Ο αριθμός αυτός μπορεί να εκφράζει το πόσο καλά προσεγγίζει το συγκεκριμένο μέτωπο το μέτωπο Pareto του προβλήματος, αλλά και κατά πόσο οι μη-κυριαρχούμενες λύσεις καλύπτουν ομοιόμορφα όλο το εύρος του μετώπου. Στην παρούσα εργασία ως μετρική ποιότητας των συνόλων λύσεων χρησιμοποιείται ο δείκτης υπερόγκου ( $\Delta Y$ ).

**Δείκτης Υπερόγκου** Ο δείκτης υπερόγκου (Hypervolume Indicator - HV) [18, 15] είναι μία από τις πιο δημοφιλείς μετρικές ποιότητας. Βασικό πλε-



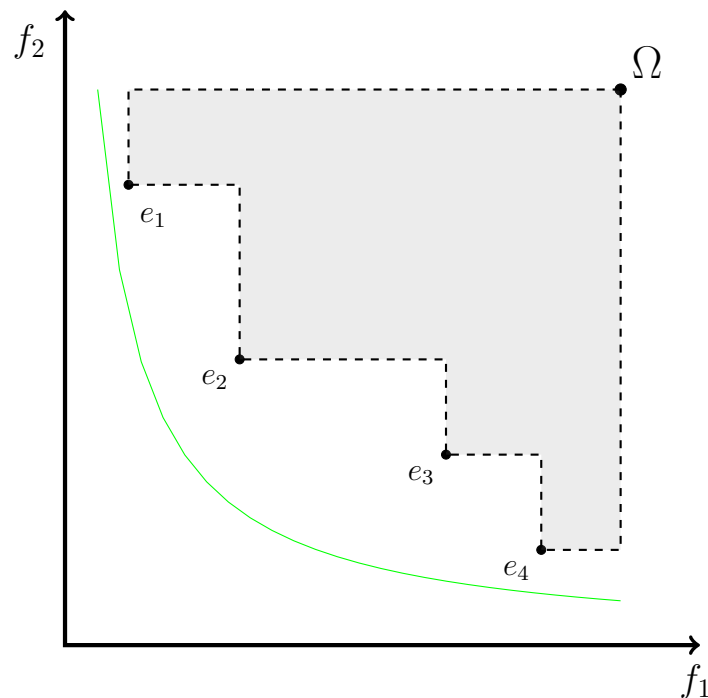
ονέκτημά της είναι ότι εκφράζει την έννοια της ασθενούς κυριαρχίας, δηλαδή εάν το μέτωπο A κυριαρχεί ασθενώς του μετώπου B, τότε ο δείκτης υπερόγκου του A θα είναι αυστηρά μεγαλύτερος και αντιστρόφως:

$$(A O_W B) \iff HV(A) > HV(B) \quad (A.4)$$

Ο υπολογισμός του  $\Delta Y$  γίνεται θεωρώντας ένα σημείο ναδίρ  $\Omega$  στο χώρο των  $M_T$  συναρτήσεων στόχων, που έχει ως συντεταγμένες τις χειρότερες (δηλαδή μεγαλύτερες) τιμές που μπορούν να πάρουν οι συναρτήσεις-στόχοι. Στη συνέχεια, για κάθε σημείο  $e_i$  του μετώπου μη-κυριαρχούμενων λύσεων φτιάχνεται ένα υπερρθογώνιο (hyperrectangle)  $Q_i$ , που έχει ως διαγώνιο το ευθύγραμμο τμήμα που ορίζουν τα σημεία  $\Omega$  και  $e_i$ . Ο δείκτης υπερόγκου υπολογίζεται τότε ως

$$HV = \lambda\{\bigcup Q_i\} \quad (A.5)$$

όπου  $\lambda$  το μέτρο Lebesgue του συνόλου.



Σχήμα A.4: Δείκτης υπερόγκου για ένα σύνολο μη κυριαρχούμενων λύσεων προβλήματος δύο στόχων. Η καμπύλη ορίζει το πραγματικό μέτωπο Pareto του προβλήματος.

Η συγκεκριμένη μετρική ποιότητας μπορεί να δοθεί και σε αδιάστατη μορφή, αν γίνει χρήση ενός σημείου ζενίθ  $A$ , το οποίο έχει συντεταγμένες τις καλύτερες (δηλαδή μικρότερες) τιμές των συναρτήσεων στόχων. Τότε ο  $\Delta Y$  εκφράζεται ως ο λόγος:

$$HV = \frac{\lambda\{U Q_i\}}{\lambda\{Q_{A\Omega}\}} \quad (A.6)$$

όπου με  $Q_{A\Omega}$  συμβολίζεται το μεγαλύτερο θεωρητικά υπερρθογώνιο, που έχει ως διαγώνιο το ευθύγραμμο τμήμα  $A\Omega$ .

Η παράσταση του δείκτη υπερόγκου συναρτήσει του πλήθους των ακριβών αξιολογήσεων που πραγματοποιούνται παρέχει μια καλή εικόνα για το πώς εξελίσσεται η ποιότητα του συνόλου λύσεων, και σε συνδυασμό με τα τελικά σύνολα λύσεων που παράγονται μπορεί να χρησιμοποιηθεί για τη σύγκριση διαφορετικών μεθόδων βελτιστοποίησης.

# Βιβλιογραφία

- [1] Γιαννάκογλου, Κ.Χ.: Μέθοδοι Βελτιστοποίησης στην Αεροδυναμική. Εθνικό Μετσόβιο Πολυτεχνείο, 2006.
- [2] Γιώτης, Α.: “Χρήση εξελικτικών τεχνικών, υπολογιστικής ευφυίας και μεθόδων υπολογιστικής ρευστομηχανικής στη βελτιστοποίηση-αντίστροφη σχεδίαση πτερυγώσεων στροβιλομηχανών, μέσω παράλληλης επεξεργασίας.” Διδακτορική διατρ. Αθήνα: Εργαστήριο Θερμικών Στροβιλομηχανών Ε.Μ.Π., 2003.
- [3] Καμπόλης, Ι.: “Πολυεπίπεδοι, Πολυεπεξεργαστικοί Αλγόριθμοι Αεροδυναμικής Βελτιστοποίησης στις Στροβιλομηχανές”. Διδακτορική διατρ. Εργαστήριο Θερμικών Στροβιλομηχανών Ε.Μ.Π., 2009.
- [4] Καρακάσης, Μ.: “Ιεραρχικοί, Κατανεμημένοι Εξελικτικοί Αλγόριθμοι και Υπολογιστική Ευφυία στη Σχεδίαση Βέλτιστων Αεροδυναμικών Μορφών”. Διδακτορική διατρ. Ε.Μ.Π., 2006.
- [5] Καψούλης, Δ.: “Σχεδιασμός Πειραμάτων και Εξωτερικά Μεταπρότυπα για την Υποβοήθηση της Διαδικασίας Βελτιστοποίησης με Εξελικτικούς Αλγορίθμους”. Διπλωματική Εργασία. Εργαστήριο Θερμικών Στροβιλομηχανών Ε.Μ.Π., 2013.
- [6] Κυριάκου, Σ.: “Μέθοδοι Σχεδιασμού-Βελτιστοποίησης στις Στροβιλομηχανές βασισμένες στους Εξελικτικούς Αλγόριθμους”. Διδακτορική διατρ. Εργαστήριο Θερμικών Στροβιλομηχανών Ε.Μ.Π., 2013.
- [7] Χούντα, Π.: “Βελτιστοποίηση με Εξελικτικούς Αλγορίθμους Υποβοηθούμενους από τη Μέθοδο Ανάλυσης σε Κύριες Συνιστώσες - Εφαρμογές στην Ενέργεια και στις Μεταφορές”. Διπλωματική Εργασία. Εργαστήριο Θερμικών Στροβιλομηχανών Ε.Μ.Π., 2013.

- [8] Anderson-Cook, C.M., Montgomery, D.C., Myers R.H.: Response Surface Methodology: Process and Product Optimization Using Designed Experiments. fourth. Wiley, 2016.
- [9] Arora, J.: Introduction to Optimum Design, third. Academic Press, 2011.
- [10] Asouti, V.G. and Giannakoglou, K.C.: “Aerodynamic optimization using a parallel asynchronous evolutionary algorithm controlled by strongly interacting demes”. Engineering Optimization 41 (3), pp. 241–257, 2009.
- [11] Asouti V.G., Giannakoglou K. C., Trompoukis X.S. and Tsiakas K. T.: “Design-optimization of a compressor blading on a gpu cluster.” EUROGEN 2013, Evolutionary and Deterministic Methods for Design, Optimization and Control with Applications to Industrial and Societal Problems. 2013.
- [12] Asouti V.G., Giannakoglou K.C., Kampolis, I.C. and Tsiakas K.T.: “CUDA implementation of Vertex-Centered, Finite Volume CFD methods on Un- structured Grids with Flow Control Applications”. Morgan Kaufmann, 2011, pp. 207–223.
- [13] Barbará, D. and Domeniconi, C.: “Detecting outliers using transduction and statistical testing”. In Proceedings of the 12th Annual SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM Press, 2006, pp. 54–60.
- [14] Bishop, C.: Pattern Recognition and Machine Learning. Springer, 2006.
- [15] Brockhoff, D.: “Optimal  $\mu$ -Distributions for the Hypervolume Indicator for Problems with Linear Bi-objective Fronts: Exact and Exhaustive Results”. Simulated Evolution and Learning: 8th International Conference, SEAL 2010, Kanpur, India, December 1-4, 2010. Proceedings. Springer Berlin Heidelberg, 2010, pp. 24–34.
- [16] Cover, T.: “Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition”. IEEE Transactions On Electronic Computers Vol. 14 (3), pp. 326–334, 1965.
- [17] Darwin, C.: On the Origin of Species by Means of Natural Selection. Murray, 1859.

- [18] Deb, K.: *Multi-Objective Optimization Using Evolutionary Algorithms*. Wiley, 2001.
- [19] Deb, K.: “Optimal design of a welded beam structure via genetic algorithms,” *AIAA Journal* vol. 29.11 (1991), pp. 2013–2015.
- [20] Deb, K.: “Simulated Binary Crossover for Continuous Search Space”. *Complex Systems* 9 (1995), pp. 115–148.
- [21] Dhahbi S., Barhoumi W., Zagrouba E.: “Multiscale Kernel PCA and Its Application to Curvelet Based Feature Extraction for Mammographic Mass Characterization”. *Advances in Intelligent Data Analysis XV: 15th International Symposium, IDA 2016, October 13-15, 2016, Proceedings*. Stockholm, Sweden, 2016.
- [22] Fonseca, C. and Fleming, P.J.: “An Overview of Evolutionary Algorithms in Multiobjective Computation.” *Evolutionary Computation* 3 (1), pp. 1–16, 1995.
- [23] Georgopoulou, C. and Giannakoglou, K.C.: “A multi-objective metamodel-assisted memetic algorithm with strength-based local refinement”. *Engineering Optimization* 41 (2009), pp. 909–923.
- [24] Giannakoglou, K.C.: *The EASY (Evolutionary Algorithm SYstem) Software*. <http://velos0.ltt.mech.ntua.gr/EASY/>.
- [25] Goldberg, D. E.: *Genetic Algorithms in Search, Optimization, and Machine Learning*. first. Addison-Wesley, 1989.
- [26] Gorban, A.N., Zinovyev, A.Y., Popova, T.G.: “Seven Clusters in Genomic Triplet Distributions”. In *Silico Biology* 3 (2003), pp. 471–482.
- [27] Haupt, R.L. and Haupt, S.E.: *Practical Genetic Algorithms*. second. Wiley-Interscience, 2004.
- [28] Knowles, J. and Corne, D.: “On metrics for comparing nondominated sets”. *Evolutionary Computation*, 2002. CEC '02. Proceedings of the 2002 Congress. Vol. 1, pp. 711–716.
- [29] Jolliffe, I. T.: *Principal Component Analysis (Springer Series in Statistics)*. second. Springer, 2002.
- [30] Postma, E.O., Van Der Herik, H.J. and Van der Maaten, L.J.P.: *Dimensionality Reduction: A Comparative Review*. Technical Report. TiCC, Tilburg University. 2008.

- [31] Salomon, R.: “Re-evaluating genetic algorithm performance under coordinate rotation of benchmark functions. a survey of some theoretical and practical aspects of genetic algorithms.” *BioSystems* 39 (3), pp. 263–278,1996.
- [32] Schölkopf, B., Smola, A. and Müller, K.R.: “Kernel principal component analysis”. *Artificial Neural Networks — ICANN’97: 7th International Conference Lausanne, Switzerland, October 8–10, 1997 Proceedings* (1997), pp. 583–588.
- [33] Schölkopf, B., Smola, A. and Müller, K.R.: “Nonlinear Component Analysis as a Kernel Eigenvalue Problem”. *Neural Computation* (1998), pp. 1299–1319.
- [34] Schowengerdt, R. A.: *Remote Sensing: Models and Methods for Image Processing*. third. Academic Press, 2006.
- [35] Strang, G.: *Introduction to Linear Algebra*. fourth. Wellesley Cambridge Press, 2009.
- [36] Suganthan P. N., Hansen N. Deb K. Chen Y.P. Auger A. Tiwari S.: *Problem Definitions and Evaluation Criteria for the CEC 2005 Special Session on Real-Parameter Optimization*. Technical Report. Nanyang Technological University, Singapore and Kanpur Genetic Algorithms Laboratory, IIT Kanpur. 2005.
- [37] Whitley, D, Mathias, K, Rana, S and Dzubera, J.: “Evaluating Evolutionary Algorithms”. *Artificial Intelligence Journal* 85 (1-2), pp. 245–276, 1996.
- [38] Zitzler, E., Laumanns, M. and Thiele, L.: “SPEA2: Improving the Strength Pareto Evolutionary Algorithm”. *TIK - Report* 103 (2001).