



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ

ΕΠΙΣΤΗΜΩΝ

ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

---

**Κατασκευές και Στατιστική Ανάλυση  
Παραγοντικών και Εύρωστων  
Παραμετρικών Σχεδιασμών**

---

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

**ΚΡΥΣΤΑΛΛΕΝΙΑ Π. ΔΡΟΣΟΥ**

Διπλωματούχος Εθνικού Μετσόβιου Πολυτεχνείου

**Επιβλέπων:**

Χ. ΚΟΥΚΟΥΒΙΝΟΣ, Καθηγητής Ε.Μ.Π.

**ΑΘΗΝΑ, 2019**





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ

ΕΠΙΣΤΗΜΩΝ

ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

**Κατασκευές και Στατιστική Ανάλυση  
Παραγοντικών και Εύρωστων  
Παραμετρικών Σχεδιασμών**

Διδακτορική Διατριβή

**Κρυσταλλένια Π. Δρόσου**

**ΑΘΗΝΑ, 2019**





NATIONAL TECHNICAL UNIVERSITY OF ATHENS  
SCHOOL OF APPLIED MATHEMATICAL AND PHYSICAL SCIENCES  
DEPARTMENT OF MATHEMATICS

**Construction and Statistical Analysis of  
Factorial and Robust Parameter Designs.**

Doctoral Dissertation  
**Krystallenia P. Drosou**

**ATHENS, 2019**



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ  
ΕΠΙΣΤΗΜΩΝ  
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

**Κατασκευές και Στατιστική Ανάλυση  
Παραγοντικών και Εύρωστων  
Παραμετρικών Σχεδιασμών**

Διδακτορική Διατριβή  
**ΚΡΥΣΤΑΛΛΕΝΙΑ Π. ΔΡΟΣΟΥ**  
Αθήνα, 2019

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Χρήστος Κουκουβίνος

Καθηγητής Ε.Μ.Π. (Επιβλέπων Καθηγητής)

Φιλία Βόντα

Αναπληρώτρια Καθηγήτρια Ε.Μ.Π. (Μέλος της Τριμελούς Επιτροπής)

Χρυσή Καρώνη-Ρίτσαρντσον

Καθηγήτρια Ε.Μ.Π. (Μέλος της Τριμελούς Επιτροπής)

Χαράλαμπος Ευαγγελάρας

Επίκουρος Καθηγητής Πανεπιστημίου Πειραιώς

Αλέξανδρος Καραγρηγορίου

Καθηγητής Πανεπιστημίου Αιγαίου

Μάρκος Κούτρας

Καθηγητής Πανεπιστημίου Πειραιώς

Άγγελος Τσακανίκας

Επίκουρος Καθηγητής Ε.Μ.Π.





Στον δάσκαλό μου, στους γονείς μου και στον αδερφό μου, για τη στήριξη και την υπομονή τους.



---

# Περίληψη

---

Η παρούσα διατριβή ασχολείται με την κατασκευή και τη στατιστική ανάλυση των παραγοντικών και των εύρωστων παραμετρικών σχεδιασμών.

Η μελέτη των πειραματικών σχεδιασμών είναι εξαιρετικά σημαντική στη σύγχρονη βιομηχανία, στην επιστήμη και στη μηχανική. Στο Κεφάλαιο 1 παρέχουμε μερικές βασικές έννοιες στους παραγοντικούς σχεδιασμούς, στους βέλτιστους σχεδιασμούς και στα πειράματα υπολογιστών. Επίσης παρέχουμε μια συγκεντρωτική παρουσίαση των σημαντικότερων σημείων όσον αφορά στους στόχους των πειραμάτων, τα βήματα για να είναι επιτυχημένο ένα πείραμα καθώς επίσης και τα στοιχεία που επηρεάζουν την καλή επιλογή ενός πειραματικού σχεδιασμού.

Το Κεφάλαιο 2 αφιερώνεται σε μια σύντομη εισαγωγή για τα πειράματα υπολογιστών, τα οποία είναι εξαιρετικά σημαντικά στη σύγχρονη βιομηχανία, στην επιστήμη και στη μηχανική.

Στο Κεφάλαιο 3 μελετάμε τους U-type, ορθογώνιους κατά στήλη (column-orthogonal) και τους σχεδόν ορθογώνιους κατά στήλη (nearly column-orthogonal) σχεδιασμούς. Οι σχεδόν ορθογώνιοι κατά στήλη σχεδιασμοί είναι πολύ χρήσιμοι σε περιπτώσεις όπου δεν είναι γνωστοί οι ορθογώνιοι κατά στήλη σχεδιασμοί. Στο παρόν κεφάλαιο παρουσιάζουμε την κατασκευή νέων σχεδιασμών οι οποίοι είναι κατάλληλοι για πειράματα κρησαρίσματος (screening experiments). Σε κάποιες περιπτώσεις, οι σχεδιασμοί που κατασκευάστηκαν είναι βέλτιστοι όσον αφορά στη δομή των ταυτόσημων επιδράσεων (aliased structure). Οι δομές των ταυτόσημων επιδράσεων των U-type, ορθογώνιων και σχεδόν ορθογώνιων κατά στήλη σχεδιασμών, υπολογίζονται και παρουσιάζονται σε κλειστή μορφή. Το γεγονός αυτό κάνει αυτή την καινούργια προσέγγιση ιδιαίτερα καινοτόμα και καθιστά ικανή την κατασκευή σχεδιασμών που είναι διαφορετικοί από τους ήδη υπάρχοντες στη βιβλιογραφία. Επίσης παρουσιάζονται ένα εκτεταμένο πολλαπλασιαστικό θεώρημα (multiplication theorem), καθώς και νέες άπειρες οικογένειες (infinite families) ορθογώνιων κατά στήλη σχεδιασμών

χρησιμοποιώντας ζεύγη περιοδικών ακολουθιών Golay (periodic Golay pairs). Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [III].

Στο επόμενο κεφάλαιο (Κεφάλαιο 4) παρουσιάζονται κάποιες βασικές έννοιες για τους παραγοντικούς σχεδιασμούς. Ιδιαίτερη βαρύτητα δίνεται σε μία ειδική κλάση κλασματικών παραγοντικών σχεδιασμών, τους υπερκορεσμένους σχεδιασμούς για τους οποίους παρουσιάζονται βασικές έννοιες σχετικά με τα κριτήρια βελτιστότητας, την κατασκευή και την ανάλυσή τους.

Το Κεφάλαιο 5 παρουσιάζει μία μέθοδο κατασκευής υπερκορεσμένων σχεδιασμών πολλαπλών και μικτών-επιπέδων. Πιο συγκεκριμένα, παρέχουμε ένα νέο κάτω φράγμα για την βελτιστοποίηση τέτοιων σχεδιασμών. Αυτό το φράγμα αποδεικνύεται ότι είναι ισχυρό και αρκετά γενικό αφού μπορεί να εφαρμοστεί σε ισορροπημένους και μη ισορροπημένους σχεδιασμούς. Επιπλέον, κατασκευάζονται και αξιολογούνται νέες οικογένειες ισορροπημένων και μη ισορροπημένων υπερκορεσμένων σχεδιασμών. Οι νέοι σχεδιασμοί κατασκευάζονται εφαρμόζοντας τη μεθοδολογία των  $k$ -κυκλικών σχεδιασμών για τους σχεδιασμούς πολλαπλών και μικτών επιπέδων. Επιπλέον, αναπτύσσονται και αποδεικνύονται πρόσθετες συνθήκες που είναι απαραίτητες για τους γεννήτορες και οι οποίες εξασφαλίζουν ότι οι αντίστοιχοι  $k$ -κυκλικοί υπερκορεσμένοι σχεδιασμοί είναι βέλτιστοι. Η παρούσα μεθοδολογία οδηγεί σε μια σειρά νέων, βελτιωμένων, οικογενειών υπερκορεσμένων σχεδιασμών και παρέχει εργαλεία για την άμεση κατασκευή βέλτιστων ή σχεδόν βέλτιστων  $k$ -κυκλικών σχεδιασμών, ελέγχοντας απλά τον αντίστοιχο γεννήτορα. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [IV].

Στα επόμενα δύο κεφάλαια μελετάμε δύο μεθόδους για την ανάλυση υπερκορεσμένων σχεδιασμών. Πιο συγκεκριμένα, στο Κεφάλαιο 6 παρουσιάζουμε μία μέθοδο ανάλυσης υπερκορεσμένων σχεδιασμών δύο, πολλαπλών και μικτών-επιπέδων. Η ανάπτυξη νέων στατιστικών μεθόδων που είναι εμπνευσμένες από τους αλγόριθμους μηχανικής μάθησης αυξάνεται ταχύτατα, ιδιαίτερα στην εποχή μας. Μία από αυτές τις μεθόδους βασίζεται στις μηχανές διανυσμάτων υποστήριξης (SVM), οι οποίες καταφέρνουν να εξαγάγουν με μεγάλη επιτυχία τα σημαντικά γονίδια σε προβλήματα ταξινόμησης, επιτυγχάνοντας εξαιρετικά υψηλή απόδοση. Σε αυτό το κεφάλαιο,

λοιπόν, προτείνουμε μία νέα μέθοδο επιλογής μεταβλητών σε προβλήματα παλινδρόμησης, που την ονομάζουμε SVR-RFE, για την ανίχνευση ενεργών επιδράσεων τόσο σε σχεδιασμούς δύο επιπέδων όσο και σε σχεδιασμούς μικτών επιπέδων. Οι μελέτες προσομοίωσης που παρουσιάζουμε δείχνουν ότι αυτή η διαδικασία είναι αρκετά αποτελεσματική, ειδικά όσον αφορά τη στατιστική ισχύ. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [III]. Στη συνέχεια, στο Κεφάλαιο 7 παρουσιάζουμε μία μέθοδο ανάλυσης υπερκορεσμένων σχεδιασμών δύο επιπέδων. Πιο συγκεκριμένα, παρουσιάζουμε μια διαδικασία δύο σταδίων για την ανάλυση δύο επιπέδων ΥΣ θεωρώντας ένα μοντέλο κύριων επιδράσεων, χωρίς την παρουσία αλληλεπιδράσεων. Η προτεινόμενη διαδικασία συνδυάζει τη μέθοδο του σίγουρου κρησαρίσματος (sure independence screening, SIS) με τις ποινικοποιημένες μεθόδους, όπως είναι η SCAD, η LASSO και η MC. Με τη διαδικασία αυτή επιτυγχάνουμε ταυτόχρονα τόσο την επιλογή όσο και την εκτίμηση των σημαντικών επιδράσεων. Η απόδοση της μεθόδου παρουσιάζεται μέσω διαφόρων σεναρίων προσομοίωσης όπου πραγματοποιούνται συγκρίσεις με υπάρχουσες μεθοδολογίες. Οι μέθοδοι που χρησιμοποιήσαμε στην παρούσα μελέτη είναι η παλινδρόμηση κατά βήματα σε συνδυασμό με τη SCAD, καθώς επίσης και ο DS. Τα αποτελέσματα από τη μελέτη προσομοιώσεων, και από ένα παράδειγμα με πραγματικά δεδομένα, αποκαλύπτουν ότι η προτεινόμενη διαδικασία μπορεί να θεωρηθεί ως ένα επωφελές εργαλείο λόγω της εξαιρετικά καλής απόδοσής του για τον εντοπισμό ενεργών παραγόντων. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [VI].

Το Κεφάλαιο 8 πραγματεύεται κάποιες βασικές έννοιες για τους εύρωστους παραμετρικούς σχεδιασμούς. Ιδιαίτερη βαρύτητα δίνεται τόσο στα είδη των σχεδιασμών που απαρτίζουν τη μεθοδολογία των εύρωστων παραμετρικών σχεδιασμών όσο και στις τεχνικές μοντελοποίησης, που συμπεριλαμβάνουν τη μοντελοποίηση θέσης και διασποράς άρα και την αποκριτική μοντελοποίηση.

Στο Κεφάλαιο 9 προτείνουμε μια κατασκευή αποτελεσματικών ΥΣ μαζί με μία μέθοδο ανάλυσης, προκειμένου να αντιμετωπιστεί το ιδιαίτερα σημαντικό πρόβλημα που αφορά στη μεθοδολογία των εύρωστων παραμετρικών σχεδιασμών (RPDM). Συνδυάζοντας την επαναληπτική μέθοδο σίγουρου κρησαρίσματος (iterative sure independence screening, iter-SIS) για την επιλογή των μεταβλητών με μία ποινικο-

ποιημένη μέθοδο, την αποκαλούμενη SCAD, προχωρήσαμε στην ανάλυση των ΥΣ που κατασκευάστηκαν στο παρόν κεφάλαιο. Η προτεινόμενη μεθοδολογία εφαρμόστηκε σε διαφορετικά μοντέλα με σκοπό να γίνει ορατή η αποτελεσματικότητά της σε πολλά διαφορετικά σενάρια, υποθέτοντας τόσο μοντέλα πρώτης όσο και μοντέλα δεύτερης τάξης υπό την έννοια των σχεδιασμών της επιφάνειας απόκρισης. Στο παρόν κεφάλαιο παρουσιάζουμε δύο επεξηγηματικά παραδείγματα καθώς και πολυάριθμα αριθμητικά πειράματα προσομοίωσης για πολλές διαφορετικές περιπτώσεις. Τα αποτελέσματα υποδηλώνουν ότι η προτεινόμενη μέθοδος είναι ιδιαίτερα αποτελεσματική για τον προσδιορισμό των ενεργών επιδράσεων των κύριων παραγόντων, των αλληλεπιδράσεων δύο παραγόντων, των αλληλεπιδράσεων τριών παραγόντων καθώς επίσης και των καθαρά τετραγωνικών όρων, υπό την προϋπόθεση της σποραδικότητας των επιδράσεων. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [V].

Η παρουσίαση του ερευνητικού έργου για αυτή τη διατριβή ολοκληρώνεται με το Κεφάλαιο 10, όπου προτείνουμε μια κατασκευή για υπερκορεσμένους split-plot σχεδιασμούς (SSSPD) κατάλληλους για RPD πειράματα. Εφαρμόσαμε τη μεθοδολογία των σύνθετων ορθογώνιων σχηματισμών χρησιμοποιώντας ορθογώνιους σχηματισμούς ισχύος μεγαλύτερης από 1 για τις split-plot μονάδες και ισορροπημένους ΥΣ για όλες τις whole-plot μονάδες. Η προτεινόμενη μεθοδολογία υποδεικνύει ότι τα πειράματα που διεξάγονται υπό το πρίσμα των split-plot σχεδιασμών μπορούν να έχουν τεράστια αξία στη μεθοδολογία των RPD, αφού επιτρέπουν την εκτίμηση των επιδράσεων που μας ενδιαφέρουν με πολύ πιο εύκολο τρόπο από τους σχεδιασμούς που προτάθηκαν από τον Taguchi. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [IV].

Η παρουσίαση της διατριβής ολοκληρώνεται με τα τέσσερα Παραρτήματα. Στο Παράρτημα Α παρουσιάζουμε συνοπτικά τα αποτελέσματα που προέκυψαν από το Κεφάλαιο 5. Στο Παράρτημα Β παρουσιάζουμε τα αποτελέσματα των προσομοιώσεων που προέκυψαν από το Κεφάλαιο 6. Στο Παράρτημα Γ παρουσιάζουμε τα αποτελέσματα των προσομοιώσεων που προέκυψαν από το Κεφάλαιο 7 και τέλος στο Παράρτημα Δ δίνουμε την απόδειξη του Λήμματος 1 καθώς επίσης και τα αποτελέσματα των προσομοιώσεων που προέκυψαν από το Κεφάλαιο 9.

---

# Abstract

---

This dissertation deals with the construction and statistical analysis of factorial and robust parameter designs. The study of experimental designs is extremely important in modern industry, science and engineering. In Chapter 1 we provide some basic concepts in factorial designs, optimal designs, and computer experiments. We also provide a concise presentation of the most important points regarding the objectives of the experiments, the steps for an experiment to be successful as well as the elements that influence the good choice of an experimental design.

Chapter 2 is devoted to a brief introduction to computer experiments, which are extremely important in modern industry, science and engineering.

In Chapter 3 we study U-type, column-orthogonal and nearly column-orthogonal designs. The nearly column-orthogonal designs are very useful in cases where column-orthogonal designs are unknown. In this chapter we present the construction of new designs suitable for screening experiments. In some cases, the constructed designs are optimal in respect to the alias structure. The structure of the alias effects of the U-type, column-orthogonal and nearly column-orthogonal designs are calculated and presented in closed form. This fact makes this new approach highly innovative and makes it capable of constructing designs that are different from the existing designs in the literature. In addition, we present an extended multiplication theorem, as well as new infinite families of column-orthogonal designs using periodic golay pairs. The research results presented in this chapter were published in the scientific paper [II].

The following chapter (Chapter 4) introduces some basic concepts for factorial designs. Particular emphasis is given in a special class of fractional factorial designs, the supersaturated designs for which basic concepts are presented regarding their optimization criteria, their construction as well as their analysis.

Chapter 5 discloses a method of constructing multi and mixed-level supersaturated designs. In particular, we provide a new lower bound for the optimality of such designs. This bound is proved to be tight and quite general since it can be applied to both balanced and unbalanced designs. In addition, new families of balanced and nearly balanced supersaturated designs are both constructed and evaluated. The new designs are constructed by applying the  $k$ -circulant type methodology for multi-level and mixed-level designs. Moreover, additional conditions, that are needed on the generator vectors, are developed and proved in this paper to pre-ensure the optimality of the corresponding  $k$ -circulant supersaturated designs. This methodology leads to a series of new families of improved supersaturated designs and provides tools for directly constructing optimal or nearly-optimal  $k$ -circulant designs by just checking the corresponding generator vector. The research results presented in this chapter were published in the scientific paper [V].

In the next two chapters we study two methods for analyzing supersaturated designs. More specifically, in Chapter 6 we present a method for analyzing both two-level and mixed-level designs. The development of new statistical methods inspired by machine learning algorithms is increasing rapidly, especially nowadays. One of such methods is based on support vector machines (SVM), which manage to successfully extract important genes in classification problems, achieving extremely high performance. In this chapter, we propose a new method of selecting variables in regression problems, called SVR-RFE (Support Vector Regression - Recursive Feature Elimination), for detecting active effects in both two-level and mixed-level designs. The simulation studies that we present demonstrate that this process is quite effective, especially in terms of statistical power. The research results presented in this chapter were published in the scientific paper [III]. Then, in Chapter 7, we present a method of analyzing two-level supersaturated designs. In particular, we present a two-stage procedure for analyzing two-level supersaturated designs assuming a main-effect only model without including interaction terms. The proposed method combines sure independence screening (SIS) with different penalty functions; such as Smoothly Clipped Absolute Deviation (SCAD), Lasso and MC penalty achieving both the down-selection and the estimation of the



significant effects, simultaneously. Insights on using the proposed methodology are provided through various simulation scenarios and several comparisons with existing approaches, such as stepwise in combination with SCAD and Dantzig Selector (DS) are presented as well. Results of the numerical study and real data analysis reveal that the proposed procedure can be considered as an advantageous tool due to its extremely good performance for identifying active factors. The research results presented in this chapter were published in the scientific paper [VI].

Chapter 8 addresses some basic concepts for robust parameter designs. Particular emphasis is given on both types of designs that make up the robust parameter design methodology and the modelling techniques, including location and dispersion modelling as well as response modelling.

In Chapter 9, we propose a construction of an effective SSD along with an analysis method, in order to deal with the significant problem of the robust parameter design methodology (RPDM). Combining iterative SIS variable selection and a penalized method, namely SCAD, we perform the analysis of the SSDs developed in the present work. The proposed methodology is applied in different models so as to show its effectiveness in many different scenarios, assuming both first and second-order models in a sense of a response surface design. Two illustrative examples as well as numerous numerical experiments are conducted for plenty cases. The results imply that the proposed method is highly effective for identifying the active effects of main factors, two-factor interactions, three-factor interactions as well as the pure quadratic ones, under the assumption of effect sparsity. The research results presented in this chapter were published in the scientific paper [V].

The presentation of the research project for this dissertation is completed with Chapter 10, where we propose a construction for supersaturated split-plot designs (SSSPD) suitable for RPD experiments. We applied the methodology of compound orthogonal arrays using orthogonal arrays of strength greater than 1, for the split-plot units and any relevant balanced SSD for the whole plot units. The proposed methodology suggests that experiments conducted in a split-plot point of view can be of tremendous value in RPD since they enable the estimation of the contrasts

of interest in considerably easier way than the designs proposed by Taguchi. The research results presented in this chapter were published in the scientific paper en [IV].

The presentation of this dissertation is completed with the four Appendixes. In Appendix A we summarize the results obtained from Chapter 5. In Appendix B we present the results of the simulations that emerged from Chapter 6. In Appendix  $\Gamma$  we present the results of the simulations resulting from Chapter 7 and finally in Appendix  $\Delta$  we give the proof of Lemma 1 as well as the results of the simulations resulting from Chapter 9.

---

## Ευχαριστίες

---

Η παρούσα διατριβή δε θα μπορούσε να εκπονηθεί χωρίς τη βοήθεια και τη συμπαράσταση πολλών ανθρώπων. Αισθάνομαι πρωτίστως την ανάγκη να ευχαριστήσω θερμά τον Δάσκαλό μου, Καθηγητή Ε.Μ.Π. κ. Χ. Κουκουβίνο στον οποίο οφείλεται κατά έναν πολύ μεγάλο βαθμό η υλοποίηση της παρούσας διατριβής. Η ουσιαστική καθοδήγησή του στο ξεπέρασμα των ποικίλων δυσκολιών που συνάντησα κατά τη διάρκεια της έρευνας, οι πολύτιμες συμβουλές και η ηθική του συμπαράσταση με βοήθησαν τα μέγιστα. Παράλληλα, θα ήθελα να τον ευχαριστήσω θερμά για την ευκαιρία που μου έδωσε να συνεργαστώ με διακεκριμένους καθηγητές, τον Καθηγητή του Πανεπιστημίου Visva Bharati της Ινδίας κ. Kashinath Chatterjee, τον Καθηγητή του Πανεπιστημίου RMIT University της Μελβούρνης στην Αυστραλία κ. Στέλιο Γεωργίου, την Καθηγήτρια του Πανεπιστημίου RMIT University της Μελβούρνης στην Αυστραλία κ. Στέλλα Στυλιανού, τον Καθηγητή του Πανεπιστημίου Sichuan της Κίνας κ. Yong-Dao Zhou, τον Καθηγητή του Γεωπονικού Πανεπιστημίου κ. Αναστάσιο Κατσιλέρο και τον Καθηγητή του Πανεπιστημίου Cardiff της Ουαλίας κ. Ανδρέα Αρτεμίου. Ήταν ιδιαίτερη χαρά και τιμή οι προαναφερθείσες συνεργασίες και θα ήθελα να τους ευχαριστήσω θερμά για τις πολύτιμες υποδείξεις τους σε θέματα Πειραματικών Σχεδιασμών καθώς και σε θέματα Εύρωστων Παραμετρικών Σχεδιασμών.

Ευχαριστώ επίσης και τα άλλα δύο μέλη της Τριμελούς Συμβουλευτικής Επιτροπής μου, τις καθηγήτριες κ. Χ. Καρώνη, Καθηγήτρια Ε.Μ.Π. και την κ. Φ. Βόντα, αναπληρώτρια Καθηγήτρια Ε.Μ.Π., των οποίων η βοήθεια ήταν επίσης καθοριστική. Θα ήθελα επίσης να εκφράσω τις ευχαριστίες μου και στα υπόλοιπα μέλη της Επιταμελούς Εξεταστικής Επιτροπής για το χρόνο που αφιέρωσαν τόσο για την ανάγνωση όσο και για τις χρήσιμες υποδείξεις τους σχετικά με την παρούσα διδακτορική διατριβή.

Ευχαριστίες επίσης απευθύνονται στον Τομέα Μαθηματικών της Σ.Ε.Μ.Φ.Ε. του

Ε.Μ.Π. για τη χορήγηση υποτροφίας ΕΛΚΕ κατά τη διάρκεια εκπόνησης της διδακτορικής διατριβής μου, υποστηρίζοντας οικονομικά την έρευνά μου.

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες σε όλους τους φίλους και συναδέλφους που με στήριξαν όλα αυτά τα χρόνια. Τέλος θα ήθελα εκ βαθέων να ευχαριστήσω τους γονείς μου Παύλο και Καίτη και τον αδερφό μου Νίκο, για την αμέριστη ηθική συμπαράσταση και ενθάρρυνσή τους καθόλη τη διάρκεια των διδακτορικών μου σπουδών.

Κρυσταλλένια Π. Δρόσου

Αθήνα, 2019.

---

## Ερευνητικό έργο

---

Κατά τη διάρκεια της παρούσας διδακτορικής διατριβής προέκυψαν οι παρακάτω δημοσιευμένες ή προς δημοσίευση επιστημονικές εργασίες:

- I. Construction of orthogonal and nearly orthogonal designs for computer experiments (with Georgiou, S.D., Stylianos, S., Koukouvinos, C.), *Biometrika*, 101, Vol. 3 (2014) 741-747.
- II. Column-orthogonal and nearly column-orthogonal designs for models with second-order terms (with Stylianos, S., Georgiou, S.D., Koukouvinos, C.), *Journal of Statistical Planning and Inference*, Vol. 161 (2015), 81-90.
- III. Response Modelling Approach to Robust Parameter Design Methodology Using Supersaturated Designs (with Chatterjee, K., Georgiou, S.D., Koukouvinos, C.), *Journal of Quality Technology*, Vol. 50 (2018), 66-75.
- IV. Multi-level and mixed-level k-circulant supersaturated designs (with Chatterjee, K., Georgiou, S., Koukouvinos, C.), *Metrika*, Vol. 81 (2018), 337-355.
- V. A new variable selection method based on SVM for analyzing supersaturated designs (with Koukouvinos, C.), *Journal of Quality Technology*, 51 (2019), 21-36.
- VI. Sure independence screening for analyzing supersaturated designs (with Koukouvinos, C.), *Communications in Statistics-Simulation and Computation*, (accepted for publication).
- VII. Supersaturated Split-Plot Designs for Robust Parameter Experiments, (with Koukouvinos, C.), *Journal of Quality Technology*, (accepted for publication).



---

# Περιεχόμενα

---

<b>Περίληψη</b>	<b>v</b>
<b>Ευχαριστίες</b>	<b>xiii</b>
<b>1 Εισαγωγή</b>	<b>1</b>
1.1 Πειράματα και οι αντίστοιχοι στατιστικοί σχεδιασμοί τους . . . . .	1
1.2 Μερικές βασικές έννοιες για τους πειραματικούς σχεδιασμούς . . . . .	3
1.2.1 Συγκεντρωτικά . . . . .	12
<b>I Πειράματα Υπολογιστών</b>	<b>15</b>
<b>2 Εισαγωγικά στοιχεία και βασικές έννοιες για τα πειράματα υπολογιστών</b>	<b>17</b>
2.1 Εισαγωγή . . . . .	17
2.2 Πειράματα υπολογιστών . . . . .	19
2.3 Σχεδιασμοί για πειράματα υπολογιστών . . . . .	21
<b>3 Ορθογώνιοι κατά στήλη και σχεδόν ορθογώνιοι κατά στήλη σχεδιασμοί για μοντέλα με όρους δεύτερης τάξης</b>	<b>23</b>
3.1 Ερευνητικό πρόβλημα-Λύση . . . . .	24
3.2 Εισαγωγή και βασικές έννοιες . . . . .	25
3.3 U-type και ορθογώνιοι κατά στήλη σχεδιασμοί από GODs . . . . .	29
3.4 U-type και σχεδόν ορθογώνιοι κατά στήλη σχεδιασμοί από GODs . . . . .	38
3.5 Κατασκευές χρησιμοποιώντας περιοδικά ζεύγη Golay . . . . .	40
3.5.1 Νέα αποτελέσματα στα περιοδικά ζεύγη Golay . . . . .	40
3.5.2 Κατασκευή ορθογώνιων σχεδιασμών . . . . .	44
3.5.3 Κατασκευές χρησιμοποιώντας περιοδικά ζεύγη Golay . . . . .	49

3.6	Δομή ταυτόσημων επιδράσεων των σχεδιασμών που δημιουργήθηκαν	51
3.7	Συζήτηση . . . . .	56
<b>II</b>	<b>Κατασκευή και Ανάλυση Υπερκορεσμένων Σχεδιασμών</b>	<b>59</b>
<b>4</b>	<b>Εισαγωγικά στοιχεία και βασικές έννοιες για τους υπερκορεσμένους σχεδιασμούς</b>	<b>61</b>
4.1	Εισαγωγή . . . . .	61
4.2	Κατασκευή και κριτήρια αξιολόγησης υπερκορεσμένων σχεδιασμών .	63
4.2.1	Κριτήρια αξιολόγησης και βελτιστοποίησης υπερκορεσμένων σχεδιασμών . . . . .	65
4.3	Ανάλυση υπερκορεσμένων σχεδιασμών . . . . .	68
<b>5</b>	<b>Κατασκευή πολλαπλών και μικτών επιπέδων <math>k</math>-κυκλικών υπερκορεσμένων σχεδιασμών</b>	<b>71</b>
5.1	Ερευνητικό πρόβλημα-Λύση . . . . .	72
5.2	Συμβολισμός και βασικές έννοιες . . . . .	73
5.3	Νέα κάτω φράγματα του $E(f_{\text{NOD}})$ . . . . .	77
5.4	Κατασκευή βέλτιστων πολλαπλών-επιπέδων $k$ -κυκλικών ΥΣ . . . . .	82
5.5	Κατασκευή βέλτιστων μικτών-επιπέδων $k$ -κυκλικών ΥΣ . . . . .	89
5.6	Συμπεράσματα . . . . .	93
<b>6</b>	<b>Ανάλυση υπερκορεσμένων σχεδιασμών με μία νέα μέθοδο που βασίζεται στις SVM</b>	<b>95</b>
6.1	Ερευνητικό πρόβλημα-Λύση . . . . .	95
6.2	Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines, SVM)	97
6.3	Νέα μέθοδος επιλογής μεταβλητών που βασίζεται στις SVR . . . . .	99
6.3.1	Αλγόριθμος . . . . .	100
6.4	Αξιολόγηση της προτεινόμενης μεθόδου . . . . .	101
6.4.1	Μέθοδοι ανάλυσης . . . . .	101
6.4.2	Παράδειγμα . . . . .	103
6.4.3	Μελέτη προσομοιώσεων . . . . .	104
6.4.4	Αποτελέσματα προσομοιώσεων . . . . .	105
6.5	Συμπεράσματα . . . . .	113



<b>7</b>	<b>Ανάλυση υπερκορεσμένων σχεδιασμών με τη μέθοδο του σίγουρου κρησαρίσματος</b>	<b>115</b>
7.1	Ερευνητικό πρόβλημα-Λύση . . . . .	115
7.2	Ανάλυση ΥΣ με τη μέθοδο του Σίγουρου Κρησαρίσματος (SIS) . . . . .	116
7.2.1	Ποινικοποιημένα ελάχιστα τετράγωνα . . . . .	118
7.2.2	Μέθοδος Σίγουρου Κρησαρίσματος . . . . .	121
7.3	Νέα μέθοδος επιλογής μεταβλητών που βασίζεται στη SIS . . . . .	124
7.4	Αξιολόγηση της προτεινόμενης μεθόδου . . . . .	125
7.4.1	Μελέτη προσομοιώσεων . . . . .	126
7.4.2	Αποτελέσματα προσομοιώσεων . . . . .	128
7.4.3	Παραδείγματα . . . . .	134
7.4.4	Γενικές παρατηρήσεις . . . . .	136
7.5	Συμπεράσματα . . . . .	138
<b>III</b>	<b>Εύρωστοι Παραμετρικοί Σχεδιασμοί</b>	<b>139</b>
<b>8</b>	<b>Εισαγωγικά στοιχεία και βασικές έννοιες για τους εύρωστους παραμετρικούς σχεδιασμούς</b>	<b>141</b>
8.1	Εισαγωγή . . . . .	141
8.2	Διασταυρωμένοι και Συνδυασμένοι Σχηματισμοί . . . . .	143
8.3	Αρχή σημαντικότητας των επιδράσεων στον Παραμετρικό Σχεδιασμό	144
8.4	Στρατηγικές μοντελοποίησης . . . . .	145
8.4.1	Μοντελοποίηση θέσης και διασποράς (Location and Dispersion Modeling) . . . . .	146
8.4.2	Μοντελοποίηση Taguchi . . . . .	147
8.4.3	Αποκριτική Μοντελοποίηση (Response Modeling) . . . . .	148
<b>9</b>	<b>Προσέγγιση με τη μοντελοποίηση της απόκρισης στη μεθοδολογία των εύρωστων παραμετρικών σχεδιασμών με τη χρήση υπερκορεσμένων σχεδιασμών</b>	<b>151</b>
9.1	Ερευνητικό πρόβλημα-Λύση . . . . .	152
9.2	Θεωρητικό υπόβαθρο . . . . .	153
9.2.1	Μεθοδολογία Εύρωστων Παραμετρικών Σχεδιασμών . . . . .	153

9.2.2	Επιλογή μεταβλητών μέσω της μη κοίλης ποινικοποιημένης πιθανοφάνειας . . . . .	155
9.2.3	Επαναληπτική μέθοδος σίγουρου κρησαρίσματος στους Υπερκορεσμένους Σχεδιασμούς . . . . .	155
9.3	Προτεινόμενη μέθοδος . . . . .	156
9.3.1	Κατασκευή . . . . .	158
9.3.2	Ανάλυση του προτεινόμενου σχεδιασμού . . . . .	161
9.3.3	Παραδείγματα . . . . .	164
9.4	Πειράματα προσομοίωσης . . . . .	168
9.5	Συμπεράσματα . . . . .	171

## **10 Υπερκορεσμένοι Split-Plot Σχεδιασμοί**

	<b>για Εύρωστα Παραμετρικά Πειράματα</b>	<b>173</b>
10.1	Ερευνητικό πρόβλημα-Λύση . . . . .	174
10.2	Θεωρητικό υπόβαθρο . . . . .	176
10.2.1	Μεθοδολογία Εύρωστων Παραμετρικών Σχεδιασμών . . . . .	176
10.2.2	Σύνθετοι ορθογώνιοι σχεδιασμοί . . . . .	177
10.3	Κατασκευή σχεδιασμών και δομή ταυτόσημων επιδράσεων . . . . .	178
10.3.1	Κατασκευή . . . . .	179
10.3.2	Δομή ταυτόσημων επιδράσεων . . . . .	182
10.3.3	Γενικές παρατηρήσεις . . . . .	184
10.3.4	Ειδική περίπτωση 1 . . . . .	185
10.3.5	Ειδική περίπτωση 2 . . . . .	187
10.3.6	Αλγόριθμος δημιουργίας σχεδιασμών . . . . .	189
10.4	Ανάλυση των προτεινόμενων σχεδιασμών . . . . .	194
10.4.1	Στάδιο I: Whole-plot (N) ανάλυση . . . . .	195
10.4.2	Στάδιο II: Ανάλυση split-plot (C) και whole-plot επί split-plot (C × N) αλληλεπιδράσεων. . . . .	196
10.4.3	Στάδιο III: Αποκριτική Μοντελοποίηση . . . . .	197
10.5	Παραδείγματα . . . . .	197
10.6	Συμπεράσματα . . . . .	207

<b>IV Παραρτήματα – Βιβλιογραφία</b>	<b>209</b>
<b>Παράρτημα Α</b>	<b>211</b>
<b>Παράρτημα Β</b>	<b>219</b>
<b>Παράρτημα Γ</b>	<b>223</b>
<b>Παράρτημα Δ</b>	<b>241</b>
<b>Βιβλιογραφία</b>	<b>263</b>



---

## Κατάλογος Σχημάτων

---

2.1	Πειράματα Υπολογιστών. . . . .	20
6.1	Σφάλματα Τύπου I (Type I) και Τύπου II (Type II) για το SVR-RFE στους ΥΣ δύο επιπέδων. Το αριστερό πλαίσιο αναφέρεται στους σχεδιασμούς « $K$ » ενώ το δεξιό πλαίσιο αναφέρεται στους σχεδιασμούς « $MW$ ». . .	106
6.2	Μέσες τιμές για τα σφάλματα Τύπου I (Type I) και Τύπου II (Type II) για όλες τις μεθόδους στους εξεταζόμενους σχεδιασμούς δύο επιπέ- δων. Τα δύο πάνω πλαίσια αναφέρονται στους μέσους των σφαλμάτων (α) Τύπου I και (β) Τύπου II των σχεδιασμών « $K$ », ενώ τα δύο κάτω πλαίσια αναφέρονται στους μέσους των σφαλμάτων (γ) Τύπου I και (δ) Τύπου II των σχεδιασμών « $MW$ ». Τα σφάλματα του SVR-RFE ανα- παριστώνται από μπλε έντονες ράβδους, τα έντονα κόκκινα γράμματα αναπαριστούν τα σφάλματα για το DS, οι μωβ έντονες ράβδοι την παλιν- δρόμηση κατά βήματα και οι πράσινες έντονες ράβδοι αναπαριστούν την ποινή $L_q$ . . . . .	107
6.3	Σφάλματα Τύπου II (Type II) για τους δύο επιπέδων σχεδιασμούς « $K$ » και « $MW$ ». Κάθε πάνελ αναφέρεται σε κάθε ξεχωριστό σχεδιασμό. Ο άξονας x αντιπροσωπεύει τους ενεργούς παράγοντες που ανήκουν στο σύνολο $\{0, 1, 2, \dots, n/2\}$ . Ο άξονας y αντιπροσωπεύει τα σφάλματα Τύπου II για τις τέσσερις διαφορετικές μεθόδους. . . . .	108
6.4	Σφάλματα Τύπου I (Type I) και Τύπου II (Type II) για τον SVR-RFE στους σχεδιασμούς μικτών επιπέδων. . . . .	110

- 6.5 Μέσες τιμές για τα σφάλματα Τύπου I (Type I) και Τύπου II (Type II) για όλες τις μεθόδους στους εξεταζόμενους μικτών επιπέδων σχεδιασμούς. Το αριστερό πλαίσιο αναφέρεται (α) στους μέσους των σφαλμάτων Τύπου I και το δεξί πλαίσιο (β) στους μέσους των σφαλμάτων Τύπου II των σχεδιασμών μικτών επιπέδων. Τα σφάλματα του SVR-RFE αναπαριστώνται από μπλε έντονες ράβδους, τα έντονα κόκκινα γράμματα αναπαριστούν τα σφάλματα για το DS, οι μωβ έντονες ράβδοι την παλινδρόμηση κατά βήματα και οι πράσινες έντονες ράβδοι αναπαριστούν την ποινή  $L_q$ . . . . . 111
- 6.6 Σφάλματα Τύπου II (Type II) για τους σχεδιασμούς μικτών επιπέδων. Κάθε πλαίσιο αναφέρεται σε κάθε ξεχωριστό σχεδιασμό. Ο άξονας  $x$  αναπαριστά τους ενεργούς παράγοντες οι οποίοι ανήκουν στο σύνολο  $\{0, 1, 2, \dots, n/2\}$ . Ο άξονας  $y$  αναπαριστά τα σφάλματα Τύπου II για τις τέσσερις διαφορετικές μεθόδους. . . . . 112
- 6.7 Τιμές της ισχύος (Power) για τον αλγόριθμο SVR-RFE. Το πλαίσιο (α) αναφέρεται στους σχεδιασμούς δύο επιπέδων, « $K$ » και « $MW$ » και το πλαίσιο (β) αναφέρεται στους σχεδιασμούς «μικτών επιπέδων». Κάθε κουκίδα αναπαριστά κάθε ξεχωριστό σχεδιασμό. Ο άξονας  $x$  αναπαριστά το όνομα του σχεδιασμού, ενώ ο άξονας  $y$  την ισχύ. . . . . 113
- 7.1 Μέθοδοι επιλογής μοντέλων σε ΥΣ. Η SIS ή η Iter-SIS χρησιμοποιούνται για να μειώσουν τη διάσταση του προβλήματος από τις αρχικές  $p$  επιδράσεις σε μία μέτρια κλίμακα, έστω  $d$ , κάτω από το μέγεθος του δείγματος. Στο δεύτερο στάδιο για να υπολογίσουμε τους συντελεστές των επιδράσεων χρησιμοποιήσαμε μία μέθοδο ποινής όπως είναι η SCAD, η LASSO και η MC. . . . . 124
- 7.2 Μέσο συνολικό σφάλμα για τις μεθόδους που βασίζονται στη SIS χρησιμοποιώντας διαφορετικές ποινές. Υπάρχουν συγκρίσεις χρησιμοποιώντας είτε το κριτήριο AIC (μαύρες ράβδοι) είτε το κριτήριο BIC (γκρι ράβδοι). Το πάνω αριστερό πλαίσιο (α) αναφέρεται στο Σενάριο I, το πάνω δεξί πλαίσιο (β) αναφέρεται στο Σενάριο II, και το κάτω αριστερό πλαίσιο (c) αναφέρεται στο Σενάριο III. . . . . 129

7.3	Μέσο συνολικό σφάλμα για το Σενάριο I. Το αριστερό πλαίσιο (a) αναφέρεται στο Μοντέλο 1, το δεξί πλαίσιο (b) αναφέρεται στο Μοντέλο 2. . . . .	130
7.4	Μέσο συνολικό σφάλμα για το Σενάριο II. Το αριστερό πλαίσιο (a) αναφέρεται στο Μοντέλο 1, το δεξί πλαίσιο (b) αναφέρεται στο Μοντέλο 2. . . . .	131
7.5	Μέσο συνολικό σφάλμα για το Σενάριο III. Το αριστερό πλαίσιο (a) αναφέρεται στο Μοντέλο 1, το δεξί πλαίσιο (b) αναφέρεται στο Μοντέλο 2. . . . .	132
7.6	Στατιστική ισχύς στο Μοντέλο 1 (μαύρες γραμμές) και στο Μοντέλο 2 (γκρι γραμμές). Το επάνω αριστερό πλαίσιο (a) αναφέρεται στο Σενάριο I, το επάνω δεξιό πλαίσιο (b) αναφέρεται το Σενάριο II και το κάτω αριστερό πλαίσιο (c) αναφέρεται στο Σενάριο III. . . . .	133





---

## Κατάλογος Πινάκων

---

1.1	Ορθογώνιος σχεδιασμός $L_9(3^4)$ και ο σχετικός σχεδιασμός. . . . .	8
3.1	Παραδείγματα ξένων, φίλιων κυκλικών πινάκων . . . . .	35
3.2	Τα ζεύγη δύο μεταβλητών κατευθυνόμενων ακολουθιών μήκους $p$ και τύπου $(p, p)$ με μηδενική περιοδική συνάρτηση αυτοσυσχέτισης . . .	42
3.3	Ορισμένοι ορθογώνιοι σχεδιασμοί $OD(n; q^m)$ κατασκευασμένοι από δύο κυκλικούς πίνακες. Όλοι οι σχεδιασμοί για $q = 4, 8, 16$ ή $q = n$ είναι επιπρόσθετα U-type. Οι σχεδιασμοί που σημειώνονται με αστέρισκο κατασκευάζονται χρησιμοποιώντας περιοδικά ζεύγη Golay που αναπτύσσονται όπως στο κεφάλαιο 3.5.1. . . . .	47
5.1	Ένας ισορροπημένος 5-επιπέδων 3-κυκλικός ΥΣ με $m = 27$ και $N = 10$ πειραματικές εκτελέσεις. Αυτός ο σχεδιασμός είναι ένας $E(f_{NOD})$ -βέλτιστος ΥΣ. . . . .	75
5.2	Ένας σχεδόν ισορροπημένος μικτών επιπέδων 2-κυκλικός ΥΣ με $m = 22$ παράγοντες και $N = 12$ πειραματικές εκτελέσεις. Αυτός ο σχεδιασμός είναι ένας $E(f_{NOD})$ -βέλτιστος ΥΣ. . . . .	76
5.3	Ένας σχεδόν ισορροπημένος 5-επιπέδων 3-κυκλικός ΥΣ με $m = 27$ παράγοντες και $N^* = 9$ πειραματικές εκτελέσεις που κατασκευάστηκε από το Πρόσιμα 8 (διαγράφοντας την πρώτη γραμμή). . . . .	85
5.4	Ένας σχεδόν-ισορροπημένος 5-επιπέδων 3-κυκλικός υπερκορεσμένος σχεδιασμός με $m = 24$ παράγοντες και $N = 8$ πειραματικές εκτελέσεις που κατασκευάστηκε από το Θεώρημα 3 (διαγράφοντας τις πρώτες δύο γραμμές και τρεις στήλες). Αυτός ο σχεδιασμός είναι βέλτιστος (Θεώρημα 4). . . . .	88

5.5	Coincidences ( $c_{i_1 i_2}$ ) μεταξύ όλων των ζευγών των επιπέδων της $i_1$ -οστής και της $i_2$ -οστής γραμμών του σχεδιασμού που παρουσιάζονται στον Πίνακα 5.4. . . . . .	89
5.6	Ένας σχεδόν ισορροπημένος μικτών επιπέδων 2-κυκλικός ΥΣ με $m = 22$ παράγοντες και $N = 11$ πειραματικές εκτελέσεις. Αυτός ο σχεδιασμός είναι βέλτιστος. . . . .	91
5.7	Ένας σχεδόν ισορροπημένος μικτών επιπέδων 2-κυκλικός ΥΣ με $m^* = 16$ παράγοντες και $N^* = 10$ πειραματικές εκτελέσεις. . . . .	92
6.1	Δύο επιπέδων υπερκορεσμένος σχεδιασμός (Lin, [91]) και τα δεδομένα Rubber του Williams [141] (Απόκριση). . . . .	103
6.2	Μεταβλητές που έχουν επιλεγεί από κάθε μέθοδο για τα δεδομένα Rubber του Williams [141]. . . . .	103
6.3	Δύο επιπέδων σχεδιασμοί που χρησιμοποιήθηκαν στη μελέτη προσομοίωσης. . . . .	106
6.4	Σχεδιασμοί μικτών επιπέδων που χρησιμοποιήθηκαν στη μελέτη προσομοίωσης. Η πρώτη στήλη αναφέρεται στο όνομα του σχεδιασμού όπου το $n$ συμβολίζει τις πειραματικές εκτελέσεις, το $k$ την κυκλικότητα και το $p$ τον αριθμό των παραγόντων. Η τέταρτη στήλη παρουσιάζει τον αριθμό των παραγόντων 2-επιπέδων, ενώ η πέμπτη συμβολίζει τον αριθμό των παραγόντων 3-επιπέδων. Η τελευταία στήλη συμβολίζει το διάνυσμα των στάθμεων του αντίστοιχου $k$ -κυκλικού σχεδιασμού. . . . .	110
7.1	Σχεδιασμοί που εξετάζονται στη μελέτη προσομοίωσης . . . . .	127
7.2	Δύο επιπέδων υπερκορεσμένος σχεδιασμός και τα δεδομένα Rubber του Williams. . . . .	134
7.3	Μοντέλα που αποκομήθηκαν από την ανάλυση των δεδομένων Rubber. . . . .	134
7.4	Μοντέλα που αποκομήθηκαν από την ανάλυση των δεδομένων Rubber χρησιμοποιώντας τις μεθόδους που βασίζονται στη SIS. . . . .	135
8.1	Κατάταξη των επιδράσεων στους παραγοντικούς έναντι των εύρωστων παραμετρικών σχεδιασμών μέχρι τρίτης τάξης σημαντικότητας των παραγόντων. . . . .	144

9.1	Δομή συσχετίσεων του πίνακα σχεδιασμού των 24 πειραματικών εκτελέσεων που κατασκευάζεται από έναν πίνακα Hadamard τάξεως 12.	162
9.1	(συνέχεια)	163
9.2	Πίνακες Hadamard που χρησιμοποιήθηκαν στη μελέτη.	164
10.1	Εύρωστος Παραμετρικός Split-plot σχεδιασμός.	181
10.2	Σχεδιασμός $D_1$ ο οποίος αποτελείται από $m_1 = 10$ whole-plot παράγοντες και $n_1 = 6$ πειραματικές εκτελέσεις.	192
10.3	Σχεδιασμός $D$ ο οποίος αποτελείται από $m_1 + m_2 = 21$ παράγοντες και $n_1 n_2 = 24$ πειραματικές εκτελέσεις.	193
10.4	Whole-plot ανάλυση διασποράς (Στάδιο I) για το Παράδειγμα 1. Παλινδρόμηση κατά βήματα με $\alpha = 0.2$ .	198
10.5	Split-plot ανάλυση διασποράς (Στάδιο II) για το Παράδειγμα 1. Παλινδρόμηση κατά βήματα με $\alpha = 0.2$ .	199
10.6	Whole-plot ανάλυση διασποράς (Στάδιο I) για το Παράδειγμα 2. Παλινδρόμηση κατά βήματα με $\alpha = 0.2$ .	202
10.7	Split-plot ανάλυση διασποράς (Στάδιο II) για το Παράδειγμα 2. Παλινδρόμηση κατά βήματα με $\alpha = 0.2$ .	202
10.8	Whole-plot ανάλυση διασποράς (Στάδιο II) για το Παράδειγμα 3. Παλινδρόμηση κατά βήματα με $\alpha = 0.2$ .	205
10.9	Split-plot ανάλυση διασποράς (Στάδιο II) για το Παράδειγμα 3. Παλινδρόμηση κατά βήματα με $\alpha = 0.2$ .	206
A.1	$E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ.	211
A.1	$E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ (συνέχεια) - A * δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.	212
A.1	$E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ (συνέχεια) - A * δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.	213
A.1	$E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ (συνέχεια) - A * δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.	214

- A.1  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ (συνέχεια) - A \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί. . . . . 215
- A.2  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί μικτών επιπέδων ΥΣ - A \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί. 216
- A.2  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί μικτών επιπέδων ΥΣ (συνέχεια) - A \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί. . . . . 217
- A.2  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί μικτών επιπέδων ΥΣ (συνέχεια) - A \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί. . . . . 218
- B.1 Αποτελέσματα συγκρίσεων προσομοίωσης μεταξύ των μεθόδων επιλογής μεταβλητών για τους δύο-επιπέδων « $K$ »,  $E(s^2)$ -βέλτιστους, σχεδιασμούς. Η πρώτη στήλη παρουσιάζει το σχεδιασμό, η δεύτερη στήλη αναφέρεται στον αριθμό των πραγματικά ενεργών παραγόντων  $d$  ενώ οι άλλες στήλες παρουσιάζουν τα σφάλματα που αντιστοιχούν σε κάθε μέθοδο. Οι σειρές που αναφέρονται ως **mean**, παρουσιάζουν το συνολικό μέσο όρο των αντίστοιχων σφαλμάτων για κάθε σχεδιασμό. 219
- B.2 Αποτελέσματα συγκρίσεων προσομοίωσης μεταξύ των μεθόδων επιλογής μεταβλητών για τους δύο-επιπέδων « $K$ »,  $E(s^2)$ -βέλτιστους, σχεδιασμούς. Η πρώτη στήλη παρουσιάζει το σχεδιασμό, η δεύτερη στήλη αναφέρεται στον αριθμό των πραγματικά ενεργών παραγόντων  $d$  ενώ οι άλλες στήλες παρουσιάζουν τα σφάλματα που αντιστοιχούν σε κάθε μέθοδο. Οι σειρές που αναφέρονται ως **mean**, παρουσιάζουν το συνολικό μέσο όρο των αντίστοιχων σφαλμάτων για κάθε σχεδιασμό. 220
- B.3 Αποτελέσματα συγκρίσεων προσομοίωσης μεταξύ των μεθόδων επιλογής μεταβλητών για τους δύο-επιπέδων « $MW$ », Bayesian  $D$ -βέλτιστους σχεδιασμούς. Η πρώτη στήλη παρουσιάζει το σχεδιασμό, η δεύτερη στήλη αναφέρεται στον αριθμό των πραγματικά ενεργών παραγόντων  $d$  ενώ οι άλλες στήλες παρουσιάζουν τα σφάλματα που αντιστοιχούν σε κάθε μέθοδο. Οι σειρές που αναφέρονται ως **mean**, παρουσιάζουν τον συνολικό μέσο όρο των αντίστοιχων σφαλμάτων για κάθε σχεδιασμό. 221

B.4	Αποτελέσματα συγκρίσεων προσομοίωσης μεταξύ των μεθόδων επιλογής μεταβλητών για τους σχεδιασμούς μικτών επιπέδων. Η πρώτη στήλη παρουσιάζει τον σχεδιασμό, η δεύτερη στήλη αναφέρεται στον αριθμό των πραγματικά ενεργών παραγόντων $d$ ενώ οι άλλες στήλες παρουσιάζουν τα σφάλματα που αντιστοιχούν σε κάθε μέθοδο. Οι σειρές που αναφέρονται ως <b>mean</b> , παρουσιάζουν τον συνολικό μέσο όρο των αντίστοιχων σφαλμάτων για κάθε σχεδιασμό. . . . .	222
Γ.1	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	223
Γ.2	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	224
Γ.3	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	225
Γ.4	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους σχεδιασμούς των Koukouninos et al. [76]. . .	225
Γ.5	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους σχεδιασμούς των Koukouninos et al. [76]. . .	226
Γ.6	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	226

Γ.7	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	227
Γ.8	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	227
Γ.9	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	228
Γ.10	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	228
Γ.11	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	229
Γ.12	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	229
Γ.13	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	230
Γ.14	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	230

Γ.15	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	231
Γ.16	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 1, για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	232
Γ.17	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 1, για τους $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	233
Γ.18	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 1, για τους $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	233
Γ.19	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	234
Γ.20	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	234
Γ.21	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	235
Γ.22	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	235
Γ.23	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 2, για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	236
Γ.24	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 2, για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	236

Γ.25	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 2, για τους $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	237
Γ.26	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 2, για τους $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	237
Γ.27	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 2, για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	238
Γ.28	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 2, για τους σχεδιασμούς των Koukouninos et al. [76]. . . . .	238
Γ.29	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 2, για τους $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	239
Γ.30	Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101]. . . . .	239
Δ.1	Διανύσματα δεδομένων $y$ που δημιουργήθηκαν από το πραγματικό αποκριτικό μοντέλο για το Παράδειγμα 1 και το Παράδειγμα 2, αντίστοιχα. . . . .	246
Δ.2	Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται τρεις παράγοντες ελέγχου και τρεις παράγοντες θορύβου για το Μοντέλο 1. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση II-I (Case III) περιγράφονται στην ενότητα 4 και το $\beta = 1, 2, 3$ υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες $E(\beta_i \beta_i \neq 0)$ και $V(\beta_i \beta_i \neq 0)$ υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων $\beta_i$ αντίστοιχα, δεδομένου ότι $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . .	247



- Δ.3 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται τρεις παράγοντες ελέγχου και τρεις παράγοντες θορύβου για το Μοντέλο 1. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 248
- Δ.4 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 1. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III), περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$ , αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 249
- Δ.5 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 1. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III), περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$ , αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 250

- Δ.6 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται τρεις παράγοντες ελέγχου και τρεις παράγοντες θορύβου για το Μοντέλο 2. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 251
- Δ.7 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται τρεις παράγοντες ελέγχου και τρεις παράγοντες θορύβου για το Μοντέλο 2. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 252
- Δ.8 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 2. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 253

- Δ.9 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 2. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 254
- Δ.10 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 3. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 255
- Δ.11 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 3. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 256

- Δ.12 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 4. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 257
- Δ.13 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 4. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 258
- Δ.14 Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 4 χρησιμοποιώντας ένα CCD. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης. . . . . 259
- Δ.15 Όλοι (δύο) κατάλληλοι πίνακες Hadamard τάξεως 24. . . . . 259
- Δ.16 Τέσσερις κατάλληλοι πίνακες Hadamard τάξεως 40. . . . . 260
- Δ.16 (συνέχεια) . . . . . 260

Δ.16 (συνέχεια) . . . . .	261
Δ.16 (συνέχεια) . . . . .	261



## Εισαγωγή

Όταν κάποιος ανακαλύψει την αλήθεια για κάτι με μεγάλο κόπο, τότε, επιθεωρώντας πιο προσεκτικά την ανακάλυψή του, συχνά διαπιστώνει ότι αυτό που τον κούρασε πολύ για να βρεθεί θα μπορούσε να παρατηρηθεί με τη μεγαλύτερη ευκολία.

—Galileo Galilei (1564-1642)

### 1.1 Πειράματα και οι αντίστοιχοι στατιστικοί σχεδιασμοί τους

Η μελέτη των πειραματικών σχεδιασμών είναι εξαιρετικά σημαντική στη σύγχρονη βιομηχανία, στην επιστήμη και στη μηχανική. Σήμερα τα πειράματα εκτελούνται σχεδόν παντού ως εργαλείο μελέτης και βελτιστοποίησης διαδικασιών και συστημάτων. Ο σκοπός ενός πειράματος στη βιομηχανική μηχανική είναι:

- να βελτιώσει την απόδοση των διαδικασιών
- να βελτιώσει την ποιότητα των προϊόντων (να μειωθεί η μεταβλητότητα και να αυξηθεί η αξιοπιστία)
- να μειώσει το χρόνο ανάπτυξης
- να μειώσει το συνολικό κόστος.

Ένας καλός πειραματικός σχεδιασμός θα πρέπει να ελαχιστοποιεί τον αριθμό των απαιτούμενων εκτελέσεων που χρειάζονται για να αποκτήσουμε όσο το δυνατόν περισσότερες πληροφορίες. Οι πειραματικοί σχεδιασμοί αποτελούν έναν κλάδο της

στατιστικής που έχει μια μακρά ιστορία τόσο από άποψη θεωρητικής ανάπτυξης, όσο και εφαρμογών. Περιεκτικές ανασκοπήσεις στη βιβλιογραφία για διάφορα είδη σχεδιασμών μπορούν να βρεθούν στο Handbook of Statistics, Vol. 13, που εκδόθηκε από τους S. Ghosh και C.R. Rao όπως επίσης και στο Handbook of Statistics, Vol. 22: Statistics in Industry, Edited by R. Khattree and C.R. Rao. Τα πειράματα μπορούν να ταξινομηθούν ως

- φυσικά πειράματα, ή
- πειράματα υπολογιστών (ή προσομοιώσεις).

Παραδοσιακά, ένα πείραμα υλοποιείται σε ένα εργαστήριο, ένα εργοστάσιο ή στον γεωργικό τομέα. Αυτό ονομάζεται φυσικό πείραμα ή πραγματικό πείραμα, όπου ο πειραματιστής πραγματοποιεί με φυσικό τρόπο όλη τη διαδικασία του πειράματος. Στα φυσικά πειράματα, υποθέτουμε ότι υπάρχουν τυχαία σφάλματα και λόγω αυτού αποκτούμε διαφορετικές τιμές εξόδου ακόμη και κάτω από τις ίδιες πειραματικές συνθήκες. Η ύπαρξη τυχαίων σφαλμάτων δημιουργεί πολυπλοκότητα τόσο στην ανάλυση δεδομένων όσο και στη μοντελοποίηση. Η πολυπλοκότητα αυξάνεται ακόμη περισσότερο στην περίπτωση που υπάρχει μία πληθώρα παραγόντων, ως εκ τούτου, ο πειραματιστής μπορεί να επιλέξει τους λίγους, αλλά σημαντικούς, παράγοντες που επηρεάζουν το πείραμα έτσι ώστε να είναι ευκολότερο να εξερευνηθεί τη σχέση μεταξύ της μεταβλητής εξόδου (απόκρισης) και των μεταβλητών εισόδου. Παράλληλα, ο πειραματιστής, για να αυξήσει την ακρίβεια ως προς την ανίχνευση αλλά και την εκτίμηση των σημαντικών παραγόντων μπορεί να χρησιμοποιήσει διάφορους ισχυρούς στατιστικούς πειραματικούς σχεδιασμούς. Η στατιστική προσέγγιση στο σχεδιασμό πειραμάτων βασίζεται συνήθως σε ένα στατιστικό μοντέλο. Ένας καλός σχεδιασμός είναι αυτός ο οποίος είναι βέλτιστος σε σχέση με το στατιστικό μοντέλο που έχουμε θεωρήσει. Υπάρχουν πολλοί στατιστικοί σχεδιασμοί για φυσικά πειράματα, μεταξύ των οποίων είναι ο κλασματικός παραγοντικός σχεδιασμός - που βασίζονται σε ένα μοντέλο ANOVA - και ο βέλτιστος σχεδιασμός - με βάση ένα μοντέλο παλινδρόμησης. Αυτοί είναι οι πιο ευρέως χρησιμοποιούμενοι σχεδιασμοί στην πράξη. Αυτά τα μοντέλα αφορούν άγνωστες παραμέτρους όπως είναι οι κύριες επιδράσεις, οι αλληλεπιδράσεις, οι συντελεστές παλινδρόμησης και η διακύμανση του τυχαίου σφάλματος. Καλοί σχεδιασμοί (π.χ. ορθογώνιοι σχεδιασμοί και διάφοροι βέλτιστοι



σχεδιασμοί) μπορεί να παρέχουν αμερόληπτες εκτιμήτριες των παραμέτρων με μικρότερο ή ακόμη και τον πιο μικρό πίνακα μεταβλητότητας-συνδιακύμανσης κάτω από μια ορισμένη έννοια. Κάποιες χρήσιμες έννοιες για τους σχεδιασμούς (φυσικών) πειραμάτων καθώς επίσης και κάποιες βασικές γνώσεις των ορθογώνιων σχεδιασμών και των βέλτιστων σχεδιασμών παρουσιάζονται στο επόμενο κεφάλαιο.

Πολλά φυσικά πειράματα μπορεί να είναι δαπανηρά και χρονοβόρα επειδή οι φυσικές διεργασίες είναι συχνά δύσκολο ή ακόμη και αδύνατον να μελετηθούν με συμβατικές πειραματικές μεθόδους. Καθώς η υπολογιστική ισχύς αυξάνεται ταχέως και είναι πλέον προσιτή σε όλους, έχει γίνει δυνατός ο υπολογισμός ορισμένων από αυτές τις διεργασίες από εξελιγμένους κώδικες υπολογιστών [118]. Τις τελευταίες δεκαετίες τα πειράματα υπολογιστών ή οι υπολογιστικές προσομοιώσεις αποτελούν κυρίαρχα θέματα τόσο στη στατιστική όσο και στη μηχανική και έχουν λάβει εκρηκτικές διαστάσεις εφιστώντας την προσοχή τόσο των επαγγελματιών όσο και από της ακαδημαϊκής κοινότητας. Το υποκείμενο μοντέλο σε πειράματα υπολογιστών είναι ντετερμινιστικό και δεδομένο, αλλά συχνά είναι πολύ περίπλοκο να το διαχειριστούμε και να το αναλύσουμε. Ένας από τους στόχους των πειραμάτων υπολογιστών (που παρουσιάζονται στο Μέρος I της παρούσας διδακτορικής διατριβής) είναι να βρούμε ένα κατά προσέγγιση μοντέλο «metamodel» που είναι πολύ απλούστερο από το πραγματικό. Τα πειράματα προσομοίωσης μελετούν την υποκείμενη διαδικασία με προσομοίωση της συμπεριφοράς της σε έναν υπολογιστή. Το πραγματικό μοντέλο είναι ντετερμινιστικό και δίνεται όπως σε ένα πείραμα υπολογιστή, αλλά τα σφάλματα στις τιμές εισόδου θεωρούνται δεδομένα και χρειάζεται να τα υποθέσουμε. Η προσομοίωση της τυχαίας διαδικασίας διεξάγεται με την ενσωμάτωση τυχαίων εισόδων στο ντετερμινιστικό μοντέλο.

## **1.2 Μερικές βασικές έννοιες για τους πειραματικούς σχεδιασμούς**

Αυτή η ενότητα εισάγει μερικές βασικές έννοιες στους παραγοντικούς σχεδιασμούς, στους βέλτιστους σχεδιασμούς, και στα πειράματα υπολογιστών. Για λεπτομέρειες ο αναγνώστης μπορεί να ανατρέξει στις ερευνητικές εργασίες [15] και [32] για τους παραγοντικούς σχεδιασμούς, στην [4] για τους βέλτιστους σχεδιασμούς και

στην [118] για τα πειράματα υπολογιστών.

**Παράγοντας:** Ένας παράγοντας είναι μια ελεγχόμενη μεταβλητή που μας ενδιαφέρει στο πείραμα. Ένας παράγοντας μπορεί να είναι ποσοτικός ή ποιοτικός. Ποσοτικός παράγοντας είναι αυτός του οποίου οι τιμές μπορούν να μετρηθούν σε αριθμητική κλίμακα και να ανήκουν σε ένα διάστημα, π.χ. θερμοκρασία, πίεση, ποσοστό δύο πρώτων υλών, μήκος χρόνου αντίδρασης, κτλ. Ποιοτικός παράγοντας είναι αυτός του οποίου οι τιμές είναι κατηγορίες, όπως για παράδειγμα ο διαφορετικός εξοπλισμός του ίδιου τύπου, οι διαφορετικοί χειριστές, τα διάφορα είδη υλικού κλπ. Ο ποιοτικός παράγοντας ονομάζεται επίσης και κατηγορικός παράγοντας. Σε πειράματα υπολογιστών, οι παράγοντες ονομάζονται συνήθως μεταβλητές εισόδου. Οι μεταβλητές που δεν μελετήθηκαν στο πείραμα δε θεωρούνται παράγοντες και ορίζονται με σταθερές τιμές. Οι μεταβλητές που δεν μπορούν να ελεγχθούν καθόλου θεωρούνται ως τυχαία σφάλματα. Σε έναν εύρωστο σχεδιασμό, ορισμένοι μη-ελεγχόμενοι παράγοντες κατά τη διάρκεια της κανονικής λειτουργίας ελέγχονται σκόπιμα και συμπεριλαμβάνονται κατά τη διάρκεια του πειράματος, έτσι ώστε το προϊόν να μπορεί να σχεδιαστεί για να μπορεί να λειτουργεί καλά σε διαφορετικά περιβάλλοντα. Αυτοί είναι γνωστοί ως παράγοντες θορύβου. Εκτενής συζήτηση για τέτοιου είδους σχεδιασμούς γίνεται στο Μέρος III της παρούσας διδακτορικής διατριβής.

**Πειραματική περιοχή, επίπεδα και συνδυασμός επιπέδων:** Η πειραματική περιοχή είναι ο χώρος όπου οι παράγοντες (μεταβλητές εισόδου) λαμβάνουν τις τιμές τους. Σε πειράματα υπολογιστών η πειραματική περιοχή ονομάζεται επίσης χώρος μεταβλητών εισόδου. Ένας παράγοντας μπορεί να επιλεγεί για να έχει μερικές συγκεκριμένες τιμές στην πειραματική περιοχή, στην οποία ο παράγοντας δοκιμάζεται. Αυτές οι επιλεγμένες τιμές ονομάζονται επίπεδα του παράγοντα. Τα επίπεδα χρησιμοποιούνται στο μοντέλο ANOVA όπου ο πειραματιστής θέλει να ελέγξει εάν η απόκριση  $y$  έχει σημαντική διαφορά μεταξύ των διάφορων επιπέδων. Ένας συνδυασμός επιπέδων είναι ένας από τους δυνατούς συνδυασμούς των επιπέδων των παραγόντων. Ένας συνδυασμός επιπέδων μπορεί να θεωρηθεί ένα σημείο στο χώρο μεταβλητών εισόδου και να ονομάζεται πειραματικό σημείο. Σε πειράματα υπολογιστών η έννοια του επιπέδου χάνει την αρχική στατιστική σημασία που ορίζεται στους παραγοντικούς σχεδιασμούς, αλλά στην παρούσα διατριβή διατηρούμε αυτή την έννοια για να διευκολύνουμε την κατασκευή των σχεδιασμών (βλ. Μέρος I).

**Εκτέλεση, δοκιμή:** Μια εκτέλεση ή δοκιμή είναι η εφαρμογή ενός συνδυασμού επιπέδων στο πειραματικό περιβάλλον. Μια εκτέλεση μπορεί να είναι φυσική ή πείραμα υπολογιστή. Επειδή ένα πείραμα υπολογιστή είναι ντετερμινιστικό, πολλαπλές δοκιμές θα παράγουν ταυτόσημες εξόδους. Έτσι, μια δοκιμή (με την έννοια ενός ή αρκετών εκτελέσεων στο ίδιο πειραματικό σημείο) έχει νόημα μόνο στα φυσικά πειράματα.

**Απόκριση:** Η απόκριση είναι το αποτέλεσμα μιας εκτέλεσης που βασίζεται στον σκοπό του πειράματος. Η απόκριση μπορεί να είναι αριθμητική τιμή ή ποιοτική ή κατηγορηματική και μπορεί να είναι μια συνάρτηση που ονομάζεται συναρτησιακή απόκριση. Στα πειράματα υπολογιστών οι αποκρίσεις ονομάζονται επίσης και έξοδοι.

**Τυχαίο Σφάλμα:** Σε οποιοδήποτε βιομηχανικό ή εργαστηριακό πείραμα υπάρχουν τυχαία σφάλματα. Μπορεί να έχουμε διαφορετικά αποτελέσματα σε δύο εκτελέσεις του ίδιου περιβάλλοντος λόγω του τυχαίου σφάλματος. Το τυχαίο σφάλμα μπορεί συχνά να θεωρηθεί ότι ακολουθεί την κανονική κατανομή  $N(0, \sigma^2)$  στα περισσότερα πειράματα. Η διασπορά  $\sigma^2$  μετράει το μέγεθος του τυχαίου σφάλματος. Σε πειράματα υπολογιστών, ωστόσο, δεν υπάρχει τυχαίο σφάλμα. Επομένως, η στατιστική θεωρία και οι μέθοδοι που έχουν κατασκευαστεί για την αντιμετώπιση τυχαίων σφαλμάτων δεν μπορούν να εφαρμοστούν άμεσα για την ανάλυση δεδομένων από πειράματα υπολογιστών. Ωστόσο, οι ιδέες και οι αρχές πολλών στατιστικών μεθόδων μπορούν να επεκταθούν σε αυτές τις περιπτώσεις χωρίς τυχαία σφάλματα.

**Παραγοντικός σχεδιασμός:** Σε έναν παραγοντικό σχεδιασμό, που αποτελείται από ένα σύνολο πειραματικών εκτελέσεων, έχουμε ως κύριο σκοπό την εκτίμηση των κύριων επιδράσεων και ορισμένων αλληλεπιδράσεων των παραγόντων. Ένας παραγοντικός σχεδιασμός ονομάζεται συμμετρικός αν όλοι οι παράγοντες έχουν τον ίδιο αριθμό επιπέδων, διαφορετικά, ονομάζεται ασυμμετρικός. Υπάρχουν αρκετές ειδικές περιπτώσεις του γενικού παραγοντικού σχεδιασμού που είναι ιδιαίτερα σημαντικές λόγω του ότι αφενός χρησιμοποιούνται ευρύτατα σε ερευνητικές εργασίες και αφετέρου αποτελούν τη βάση άλλων σχεδιασμών με μεγάλη πρακτική αξία. Η σημαντικότερη αυτών, είναι η περίπτωση που έχουμε  $m$  παράγοντες, καθέναν σε δύο μόνο στάθμες. Μία πλήρης επανάληψη ενός τέτοιου σχεδιασμού απαιτεί  $2^m$  παρατηρήσεις και λέγεται  $2^m$  παραγοντικός σχεδιασμός.

**Πλήρης παραγοντικός σχεδιασμός:** Ένας σχεδιασμός όπου όλοι οι συνδυα-

ομοί επιπέδων των παραγόντων εμφανίζονται εξίσου συχνά, ονομάζεται πλήρης παραγοντικός σχεδιασμός ή πλήρης σχεδιασμός. Με τον όρο παραγοντικοί σχεδιασμοί εννοούμε λοιπόν ότι σε κάθε πλήρη δοκιμή ή επανάληψη του πειράματος εξετάζονται όλοι οι δυνατοί συνδυασμοί των επιπέδων (στάθμεων) των παραγόντων. Είναι σαφές ότι ο αριθμός των εκτελέσεων σε έναν πλήρη παραγοντικό σχεδιασμό θα πρέπει να είναι  $k \prod_{j=1}^m q_j$ , όπου το  $q_j$  είναι ο αριθμός των επιπέδων του παράγοντα  $j$  και  $k$  είναι ο αριθμός των επαναλήψεων για όλους τους συνδυασμούς επιπέδων. Όταν όλοι οι παράγοντες έχουν τον ίδιο αριθμό επιπέδων,  $q$ , το  $n = kq^m$ , όπου  $n$  ο αριθμός των πειραματικών εκτελέσεων. Συμπεραίνουμε λοιπόν ότι, ο αριθμός των εκτελέσεων ενός πλήρους παραγοντικού σχεδιασμού αυξάνεται εκθετικά με τον αριθμό των παραγόντων. Αντιλαμβανόμαστε λοιπόν, ότι πρόκειται για ιδιαίτερα κοστοβόρα πειράματα και στην περίπτωση αυτή ιδιαίτερα όταν υπάρχει αδυναμία εκτέλεσης του πλήρους πειραματικού σχεδιασμού, λόγω έλλειψης χρόνου ή πόρων, επιλέγουμε να εξετάζουμε την εφαρμογή ενός αντιπροσωπευτικού υποσυνόλου των πειραματικών εκτελέσεων.

**Κλασματικός Παραγοντικός Σχεδιασμός:** Ένα κλάσμα ενός πλήρους παραγοντικού σχεδιασμού (Fractional factorial design, FFD) είναι ένα υποσύνολο όλων των πειραματικών εκτελέσεων [32]. Ένας FFD μπορεί να εκφραστεί ως ένας  $n \times m$  πίνακας, όπου  $n$  είναι ο αριθμός των εκτελέσεων και  $m$  ο αριθμός των παραγόντων και η  $j$ -οστή στήλη του πίνακα έχει  $q_j$  επίπεδα (εγγραφές). Για παράδειγμα, ο σχεδιασμός στο Παράδειγμα 1 που ακολουθεί μπορεί να εκφραστεί ως

$$D = \begin{bmatrix} 80 & 90 & 5 & a \\ 80 & 120 & 7 & b \\ 80 & 150 & 9 & c \\ 90 & 90 & 7 & c \\ 90 & 120 & 9 & a \\ 90 & 150 & 5 & b \\ 100 & 90 & 9 & b \\ 100 & 120 & 5 & c \\ 100 & 150 & 7 & a \end{bmatrix}. \quad (1.1)$$

όπου υπάρχουν τέσσερις παράγοντες, καθένας από τους οποίους έχει τρία επίπεδα.

Συνολικά, υπάρχουν  $81 = 3^4$  δυνατοί συνδυασμοί επιπέδων, αλλά ο σχεδιασμός  $D$  επιλέγει μόνο εννέα από αυτούς. Το πώς να επιλέξουμε ένα καλό υποσύνολο είναι το πιο σημαντικό θέμα στους FFD. Στη βιβλιογραφία συνιστάται η προσεκτική επιλογή συνδυασμών, γνωστοί ως ορθογώνιοι σχεδιασμοί, και αυτή η επιλογή χρησιμοποιείται ευρέως και στην πράξη.

**Ορθογώνιος σχηματισμός:** Ένας ορθογώνιος σχηματισμός (orthogonal array, OA) ισχύος  $t$  με  $n$  εκτελέσεις και  $s$  παράγοντες που συμβολίζεται ως  $OA(n, s, q, t)$  είναι ένας FFD όπου κάθε υποσχεδιασμός  $n$  εκτελέσεων και  $m$  ( $m \leq t$ ) παραγόντων είναι ένας πλήρης σχεδιασμός. Οι ορθογώνιοι σχηματισμοί ισχύος δύο, χρησιμοποιούνται εκτενώς για τον σχεδιασμό πειραμάτων σε διάφορους τομείς και συχνά εκφράζονται ως πίνακες ορθογώνιων σχεδιασμών. Ο αναγνώστης μπορεί να ανατρέξει στους Hedayat et al. [62] για περισσότερες λεπτομέρειες.

**Πίνακες ορθογώνιων σχηματισμών:** Ένας πίνακας ορθογώνιου σχηματισμού, που συμβολίζεται με  $L_n(q_1 \times \dots \times q_s)$ , είναι ένας  $n \times s$  πίνακας με εισόδους  $1, 2, \dots, q_j$  στην  $j$ -οστή στήλη έτσι ώστε:

1. κάθε επίπεδο σε κάθε στήλη εμφανίζεται εξίσου συχνά.
2. κάθε συνδυασμός επιπέδων σε οποιοδήποτε  $t$  στήλες εμφανίζεται εξίσου συχνά.

Όταν όλοι οι παράγοντες έχουν τον ίδιο αριθμό επιπέδων, ο σχηματισμός δηλώνεται με  $L_n(q^s)$  που είναι ένας ορθογώνιος σχηματισμός  $OA(n, s, q, 2)$ . Οι πρώτες πέντε στήλες του Πίνακα 1.1 δίνουν έναν ορθογώνιο πίνακα  $L_9(3^4)$ , όπου ο σχεδιασμός χρησιμεύει για ένα πείραμα εννέα εκτελέσεων και τεσσάρων παραγόντων καθένας σε τρία επίπεδα: 1, 2 και 3. Στην πραγματικότητα μπορούμε να χρησιμοποιήσουμε άλλα τρία σύμβολα όπως « $\alpha$ », « $\beta$ » και « $\gamma$ » για να αντικαταστήσουμε το 1, 2 και 3 στο  $L_9(3^4)$ .

**Παράδειγμα 1** Υποθέτουμε ότι υπάρχουν τέσσερις παράγοντες σε ένα πείραμα χημικής μηχανικής καθένας σε τρία επίπεδα ως εξής:

A (η θερμοκρασία της αντίδρασης):  $80^\circ C, 90^\circ C, 100^\circ C$

B (το χρονικό μήκος της αντίδρασης):  $90min, 120min, 150min$

C (το ποσοστό αλκαλίων):  $5\%, 7\%, 9\%$

D (χειριστής):  $a, b, c$ .

Υπάρχουν  $3^4 = 81$  συνδυασμοί επιπέδων. Εάν χρησιμοποιούμε έναν  $L_9(3^4)$  για να κατασκευάσουμε ένα κλασματικό παραγοντικό σχεδιασμό, οι στήλες του  $L_9(3^4)$  μπορούν να συμπληρωθούν για τους παράγοντες  $A, B, C$  και  $D$  με οποιοσδήποτε επιλογές από τις  $4! = 24$  δυνατές. Ας υποθέσουμε ότι οι τέσσερις παράγοντες τοποθετούνται σε στήλες με φυσική σειρά. Στη συνέχεια, αντικαθιστούμε τα «αφηρημένα» τρία επίπεδα των τεσσάρων στηλών από τα επίπεδα των παραγόντων για να αποκτήσουμε έναν κλασματικό παραγοντικό σχεδιασμό στη δεξιά πλευρά του Πίνακα 1.1. Υπάρχουν εννέα επιλεγμένοι συνδυασμοί επιπέδων. Ο πρώτος είναι  $(80^\circ C, 90min., 5\%, a)$  και ο τελευταίος είναι  $(100^\circ C, 150min., 7\%, a)$ . Σημειώνουμε ότι ο παράγοντας  $D$  (χειριστής) δεν μας ενδιαφέρει, αλλά πρέπει να εξετάσουμε το αποτέλεσμα του καθώς ο χειριστής μπορεί να επηρεάσει την έξοδο του πειράματος. Ένας τέτοιος παράγοντας ονομάζεται παράγοντας θορύβου.

**Πίνακας 1.1:** Ορθογώνιος σχεδιασμός  $L_9(3^4)$  και ο σχετικός σχεδιασμός.

No	1	2	3	4	A	B	C	D
1	1	1	1	1	$80^\circ C$	90min	5%	a
2	1	2	2	2	$80^\circ C$	120min	7%	b
3	1	3	3	3	$80^\circ C$	150min	9%	c
4	2	1	2	3	$90^\circ C$	90 min	7%	c
5	2	2	3	1	$90^\circ C$	120 min	9%	a
6	2	3	1	2	$90^\circ C$	150 min	5%	b
7	3	1	3	2	$100^\circ C$	90 min	9%	b
8	3	2	1	3	$100^\circ C$	120min	5%	c
9	3	3	2	1	$100^\circ C$	150 min	7%	a

**Ισομορφισμός:** Η εναλλαγή στηλών και γραμμών του ορθογώνιου πίνακα σχεδιασμού εξακολουθεί να δίνει έναν ορθογώνιο πίνακα σχεδιασμού. Δύο ορθογώνιοι σχεδιασμοί ονομάζονται ισόμορφοι εάν ο ένας μπορεί να ληφθεί από τον άλλο με εναλλαγή σειρών και στηλών και μετάθεση των επιπέδων μιας ή περισσότερων στηλών του πίνακα σχεδιασμού. Δύο ισόμορφοι παραγοντικοί σχεδιασμοί θεωρούνται ισοδύναμοι.

**Μοντέλα ANOVA:** Ένας παραγοντικός σχεδιασμός βασίζεται σε ένα στατιστικό μοντέλο. Για ένα πείραμα με επίπεδα παραγόντων  $x_1, \dots, x_q$ , το μοντέλο μπορεί να εκφραστεί ως

$$y_{ij} = \mu + \alpha_j + \epsilon_{ij} = \mu_j + \epsilon_{ij}, j = 1, \dots, q, i = 1, \dots, n_j, \quad (1.2)$$

όπου  $\mu$  είναι ο γενικός μέσος της μεταβλητής  $y$ ,  $\mu_j$  είναι η πραγματική τιμή της απόκρισης  $y$  στο  $x_j$ , και  $\epsilon_{ij}$  είναι το τυχαίο σφάλμα στην  $i$ th επανάληψη στο  $j$ th επίπεδο του  $x_j$ . Όλα τα  $\epsilon_{ij}$  υποθέτουμε ότι είναι ανεξάρτητα και ομοιόμορφα κατανεμημένα σύμφωνα με την  $N(0, \sigma^2)$ . Τότε ο μέσος  $\mu_j$  μπορεί να διασπαστεί σε  $\mu_j = \mu + \alpha_j$ , όπου  $\alpha_j$  καλείται η κύρια επίδραση του  $y$  στο  $x_j$  και ικανοποιεί τη σχέση  $\alpha_1 + \dots + \alpha_q = 0$ . Ο αριθμός των εκτελέσεων σε αυτό το πείραμα είναι  $n = n_1 + \dots + n_q$ . Το μοντέλο 1.2 καλείται μοντέλο ANOVA. Ένα μοντέλο ANOVA για ένα πείραμα δύο παραγόντων, παράγοντας  $A$  και παράγοντας  $B$ , μπορεί να εκφραστεί ως:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}, i = 1, \dots, p; j = 1, \dots, q; k = 1, \dots, n_{ij}, \quad (1.3)$$

όπου  $\mu$  = είναι ο γενικός μέσος,

$\alpha_i$  = η κύρια επίδραση του παράγοντα  $A$  στο επίπεδο  $\alpha_i$ ,

$\beta_j$  = η κύρια επίδραση του παράγοντα  $B$  στο επίπεδο  $\beta_j$ ,

$\epsilon_{ijk}$  = τυχαίο σφάλμα στην  $k$ th εκτέλεση στο συνδυασμό επιπέδου  $\alpha_i\beta_j$ ,

$(\alpha\beta)_{ij}$  = αλληλεπίδραση μεταξύ  $A$  και  $B$  στο συνδυασμό επιπέδου  $\alpha_i\beta_j$  υπό τους περιορισμούς:

$$\sum_{i=1}^p \alpha_i = 0; \quad \sum_{j=1}^q \beta_j = 0;$$

$$\sum_{i=1}^p (\alpha\beta)_{ij} = \sum_{j=1}^q (\alpha\beta)_{ij} = 0.$$

Υπάρχουν  $(p - 1)$  ανεξάρτητες παράμετροι  $\alpha_i$ ,  $(q - 1)\beta_j$ , και  $(p - 1)(q - 1)(\alpha\beta)_{ij}$ . Ο συνολικός αριθμός αυτών των παραμέτρων είναι  $pq - 1$ . Για ένα παραγοντικό πείραμα  $m$  παραγόντων με  $q_1, \dots, q_m$  επίπεδα, μπορούμε να εξετάσουμε τις αλληλεπιδράσεις μεταξύ τριών παραγόντων, τεσσάρων παραγόντων κλπ. και ο συνολικός αριθμός των κύριων επιδράσεων και αλληλεπιδράσεων γίνεται  $(\prod_{j=1}^m q_j - 1)$ , το οποίο αυξάνεται εκθετικά όσο αυξάνεται το  $m$ .

**Αρχή σποραδικότητας:** Ο αριθμός των σημαντικών επιδράσεων και αλληλεπιδράσεων σε έναν παραγοντικό σχεδιασμό είναι μικρός.

**Αρχή ιεραρχίας:** Οι επιδράσεις χαμηλότερης τάξης είναι πιθανότερο να είναι σημαντικές από τις επιδράσεις υψηλότερης τάξης. Οι κύριες επιδράσεις είναι πιθανότερο να είναι σημαντικές από τις αλληλεπιδράσεις, και οι επιδράσεις της ίδιας τάξης είναι εξίσου πιθανόν να είναι σημαντικές.

**Υπερκορεσμένοι σχεδιασμοί:** Οι υπερκορεσμένοι σχεδιασμοί είναι κλασματικοί παραγοντικοί σχεδιασμοί στους οποίους ο αριθμός των εκτιμώμενων επιδράσεων (κύριων ή αλληλεπιδράσεων) είναι μεγαλύτερος από τον αριθμό των πειραματικών εκτελέσεων. Τόσο σε βιομηχανικά όσο και επιστημονικά πειράματα, ειδικά στα αρχικά στάδιά τους, είναι ιδιαίτερα συχνό φαινόμενο να υπάρχει ένας μεγάλος αριθμός παραγόντων που είναι απαραίτητο να μελετηθούν αλλά το μέγεθος των πειραματικών εκτελέσεων να είναι περιορισμένο λόγω του υψηλού κόστους. Σύμφωνα με την αρχή σποραδικότητας των επιδράσεων, οι υπερκορεσμένοι σχεδιασμοί μπορούν να χρησιμοποιηθούν αποτελεσματικά σε τέτοια προβλήματα για να εντοπίσουν τους κυρίαρχους παράγοντες. Σύμφωνα με αυτή την αρχή, όπως τονίσαμε και προηγουμένως, οι παράγοντες που επηρεάζουν την απόκριση του πειράματος είναι λιγυσοστοί, συγκριτικά με το αρχικό σύνολο όλων των παραγόντων, επομένως στόχος αυτών των πειραμάτων είναι ο προσδιορισμός των λίγων αλλά κυρίαρχων ενεργών παραγόντων. Σκεφτείτε ένα σχεδιασμό με  $n$  εκτελέσεις και  $m$  παράγοντες, κάθε έναν με  $q$  επίπεδα. Ο σχεδιασμός ονομάζεται μη-κορεσμένος εάν  $n - 1 > m(q - 1)$ , κορεσμένος εάν  $n - 1 = m(q - 1)$  και υπερκορεσμένος εάν  $n - 1 < m(q - 1)$ . Περισσότερες λεπτομέρειες σχετικά με τους υπερκορεσμένους σχεδιασμούς παρέχονται στο Μέρος II της παρούσας διδακτορικής διατριβής.

**Βέλτιστοι σχεδιασμοί:** Υποθέτουμε ότι η υποκείμενη σχέση μεταξύ της απόκρισης  $y_k$  και των παραγόντων εισόδου  $x_k = (x_{k1}, \dots, x_{ks})$  μπορούν να εκφράζονται σε ένα μοντέλο παλινδρόμησης ως:

$$y_k = \sum_{j=1}^m \beta_j g_j(x_{k1}, \dots, x_{ks}) + \epsilon_k = \sum_{j=1}^m \beta_j g_j(x_k) + \epsilon_k, k = 1, \dots, n, \quad (1.4)$$

όπου  $g_j(\cdot)$  είναι προκαθορισμένες ή γνωστές συναρτήσεις,  $\epsilon$  είναι το τυχαίο σφάλμα με  $E(\epsilon) = 0$ , και  $Var(\epsilon) = \sigma^2$ . Το μοντέλο 1.4 περιλαμβάνει πολλά χρήσιμα μοντέλα, όπως είναι το γραμμικό μοντέλο,

$$y_k = \beta_0 + \beta_1 x_1 + \dots + \beta_s x_s + \epsilon_k \quad (1.5)$$



και το τετραγωνικό μοντέλο,

$$y_k = \beta_0 + \sum_{i=1}^s \beta_i x_i + \sum_{i=1}^s \beta_i x_i^2 + \sum_{1 \leq i < j \leq s} \beta_{ij} x_i x_j + \epsilon_k. \quad (1.6)$$

Σημειώνουμε ότι οι συναρτήσεις  $g_j$  μπορεί να είναι μη γραμμικές στο  $x$ , όπως για παράδειγμα  $\exp(-x_i)$ ,  $\log(x_j)$ ,  $1/(a + x_i x_j)$ , κ.λπ. Στην παρούσα διδακτορική διατριβή θεωρούμε τόσο γραμμικά όσο και πολυωνυμικά μοντέλα και τετραγωνικά μοντέλα. Το μοντέλο 1.4 μπορεί να εκφραστεί με τη μορφή διανυσμάτων και πινάκων ως

$$\mathbf{y} = \mathbf{G}\boldsymbol{\beta} + \boldsymbol{\epsilon}. \quad (1.7)$$

όπου

$$\mathbf{G} = \begin{bmatrix} g_1(x_1) & \dots & g_m(x_1) \\ \vdots & & \vdots \\ g_1(x_n) & \dots & g_m(x_n) \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \quad (1.8)$$

Ο πίνακας  $\mathbf{G}$ , που ονομάζεται πίνακας σχεδιασμού, αντιπροσωπεύει τόσο τα δεδομένα όσο και το μοντέλο.  $M = \mathbf{G}'\mathbf{G}$  ονομάζεται ο πίνακας πληροφορίας, όπου  $\mathbf{G}'$  είναι ο ανάστροφος του  $\mathbf{G}$ . Αυτός μπορεί να γραφεί ως  $\mathbf{M}(D_n)$  για να τονίσει ότι είναι μία συνάρτηση του σχεδιασμού  $D_n = x_k = (x_1, \dots, x_s)$ . Ο πίνακας συνδιακύμανσης του εκτιμητή ελαχίστων τετραγώνων δίνεται από τη σχέση

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = \frac{\sigma^2}{n} \mathbf{M}^{-1}.$$

Σαφώς, θέλουμε το  $\text{Cov}(\hat{\boldsymbol{\beta}})$  να είναι όσο το δυνατό μικρότερο με μια ορισμένη έννοια. Αυτό σημαίνει τη μεγιστοποίηση του  $\mathbf{M}(D_n)$  σε σχέση με το  $D_n$ . Ωστόσο,  $\mathbf{M}$  είναι ένας πίνακας  $m \times m$ , οπότε πρέπει να βρούμε ένα κριτήριο κλίμακας για τη μεγιστοποίηση του  $\mathbf{M}$ , που υποδηλώνεται με  $\phi(\mathbf{M}(D_n))$ , και τότε μπορούμε να βρούμε τον  $\phi$ -βέλτιστο σχεδιασμό που μεγιστοποιεί το  $\phi(\mathbf{M})$  πάνω στο χώρο του σχεδιασμού. Πολλοί συγγραφείς έχουν προτείνει πολλά κριτήρια, όπως για παράδειγμα τα:

D-βελτιστότητα: μεγιστοποίηση της ορίζουσας  $\mathbf{M}$ ,

A-βελτιστότητα: ελαχιστοποίηση του ίχνους του  $\mathbf{M}^{-1}$ ,

E-βελτιστότητα: ελαχιστοποίηση της μεγαλύτερης ιδιοτιμής του  $\mathbf{M}^{-1}$ .

Στην πολυμεταβλητή στατιστική ανάλυση η ορίζουσα του πίνακα συνδιακύμανσης ονομάζεται γενικευμένη διακύμανση. Η D-βελτιστότητα είναι ισοδύναμη με την ελαχιστοποίηση της γενικευμένης διακύμανσης και επίσης ισοδυναμεί με την ελαχιστοποίηση του όγκου του ελλειψοειδούς  $(\beta - \hat{\beta})' \mathbf{M} (\beta - \hat{\beta}) \leq \alpha^2$  για κάθε  $\alpha^2 > 0$ . Η A-βελτιστότητα είναι ισοδύναμη με την ελαχιστοποίηση του αθροίσματος των διακυμάνσεων  $\hat{\beta}_1, \dots, \hat{\beta}_m$ . Όταν το πραγματικό μοντέλο είναι γνωστό, οι βέλτιστοι σχεδιασμοί είναι οι καλύτεροι υπό τη δεδομένη βελτιστότητα. Οι βέλτιστοι σχεδιασμοί έχουν πολλές ελκυστικές ιδιότητες, αλλά έχουν έλλειψη ευρωστίας στο μη προσδιορισμένο μοντέλο. Όταν το υποκείμενο μοντέλο είναι άγνωστο, οι βέλτιστοι σχεδιασμοί ενδέχεται να έχουν κακή απόδοση. Αν μπορεί κανείς να συνδυάσει την αποτελεσματικότητα του βέλτιστου σχεδιασμού και της ευρωστίας του ομοιόμορφου σχεδιασμού, ο νέος σχεδιασμός μπορεί να έχει μία καλή απόδοση.

### 1.2.1 Συγκεντρωτικά

Στην παρούσα ενότητα κάνουμε μια συγκεντρωτική παρουσίαση των σημαντικότερων σημείων όσον αφορά στους στόχους των πειραμάτων, τα βήματα για να είναι επιτυχήμενο ένα πείραμα, καθώς επίσης και τα στοιχεία που επηρεάζουν την καλή επιλογή ενός πειραματικού σχεδιασμού.

#### **Οι στόχοι των πειραμάτων περιλαμβάνουν τα ακόλουθα:**

- ποιες μεταβλητές έχουν μεγαλύτερη επιρροή στην απόκριση  $y$  (κρησάρισμα).
- που θα θέσουμε τα σημαντικά  $x$ , δηλαδή ποια θα είναι τα βέλτιστα επίπεδα των  $x$ , έτσι ώστε το  $y$  να είναι σχεδόν πάντα κοντά στην επιθυμητή τιμή-στόχο (βελτιστοποίηση).
- που θα θέσουμε τα σημαντικά  $x$ , δηλαδή ποια θα είναι τα βέλτιστα επίπεδα των  $x$ , έτσι ώστε η μεταβλητότητα στο  $y$  να είναι η μικρότερη δυνατή.
- που θα θέσουμε τα σημαντικά  $x$ , δηλαδή ποια θα είναι τα βέλτιστα επίπεδα των  $x$ , έτσι ώστε να ελαχιστοποιούνται οι επιδράσεις των μη ελεγχόμενων μεταβλητών  $z$ .
- ένα μαθηματικό μοντέλο για να προβλέψουμε τις μελλοντικές αποκρίσεις.

**Σχεδιασμός του πειράματος:** Τα βήματα που πρέπει να ακολουθήσουμε έτσι ώστε να διεξάγουμε ένα επιτυχημένο πείραμα είναι:

- Αναγνώριση και δήλωση του προβλήματος.
- Επιλογή παραγόντων, επιπέδων και εύρους.
- Επιλογή της μεταβλητής απόκρισης.
- Επιλογή πειραματικού σχεδιασμού.
- Εκτέλεση του πειράματος.
- Στατιστική ανάλυση των δεδομένων.
- Συμπεράσματα και συστάσεις.

**Επιλογή του πειραματικού σχεδιασμού:** Σε πολλές περιπτώσεις, υπάρχουν περισσότερες από μία «καλές» επιλογές ενός πειραματικού σχεδιασμού. Όταν επιλέγουμε έναν σχεδιασμό, πάντα πρέπει να έχουμε κατά νου τους πειραματικούς στόχους. Συχνά, η επιλογή σχεδιασμού εξαρτάται από το μοντέλο που χρησιμοποιείται. Ορισμένοι σχεδιασμοί απαιτούν ανάλυση δύο (ή περισσότερων) σταδίων για να προσαρμόσουμε το μοντέλο. Για παράδειγμα, για τους σχεδιασμούς κρησαρίσματος χρειάζονται συμπληρωματικά πειράματα για να προσαρμόσουμε πολύπλοκα μοντέλα. Επίσης, διάφοροι περιορισμοί στο πείραμα θα πρέπει πάντα να λαμβάνονται υπόψη (όπως ο αριθμός των παραγόντων, τα επίπεδα του κάθε παράγοντα). Τέλος, το κόστος κάθε πειράματος (επεξεργασία) και ο διαθέσιμος προϋπολογισμός εκτός από το μοντέλο που χρησιμοποιείται, θα καθορίσουν τον αριθμό των πειραματικών εκτελέσεων στο σχεδιασμό.



# **Μέρος Ι**

## **Πειράματα Υπολογιστών**



---

## Εισαγωγικά στοιχεία και βασικές έννοιες για τα πειράματα υπολογιστών

---

Statisticians, like artists,  
have the bad habit of falling in love with their models.

—George Box ( 1985-2013)

$\Sigma$  το παρόν κεφάλαιο παρουσιάζεται μια σύντομη εισαγωγή για τα πειράματα υπολογιστών.

### 2.1 Εισαγωγή

Με την έλευση της τεχνολογίας υπολογιστών και των αριθμητικών μεθόδων, οι μηχανικοί και οι επιστήμονες συχνά χρησιμοποιούν προσομοιώσεις υπολογιστών για να μελετήσουν πραγματικά ή θεωρητικά φυσικά συστήματα. Για να προσομοιώσουμε ένα φυσικό σύστημα, πρέπει να δημιουργήσουμε μαθηματικά μοντέλα για να αναπαραστήσουμε τις φυσικές συμπεριφορές. Τα μοντέλα μπορούν να λάβουν πολλές μορφές και διαφορετικά επίπεδα λεπτομερούς εκπροσώπησης του φυσικού συστήματος. Τα μοντέλα είναι συχνά πολύ περίπλοκα και κατασκευασμένα με διαφορετικά επίπεδα ακρίβειας, όπως το λεπτομερές μοντέλο βασισμένο στη φυσική, καθώς επίσης και πιο αφηρημένα και υψηλότερου επιπέδου μοντέλα με λιγότερο λεπτομερή αναπαράσταση. Ένα μοντέλο βασισμένο στη φυσική μπορεί να αντιπροσωπεύεται από ένα σύνολο εξισώσεων που περιλαμβάνουν γραμμικές, μη γραμμικές, συνήθεις και μερικές διαφορικές εξισώσεις. Ένα λιγότερο λεπτομερές και ένα

## 18 Εισαγωγικά στοιχεία και βασικές έννοιες για τα πειράματα υπολογιστών

πιο αφηρημένο μοντέλο μπορεί να χρησιμοποιεί στοχαστικές αναπαραστάσεις όπως αυτές που χρησιμοποιούνται συνήθως για την προσομοίωση διαδικασιών με βάση εμπειρικές παρατηρήσεις (π.χ. προσομοιώσεις διακριτών συμβάντων για δίκτυα αναμονής). Λόγω της πολυπλοκότητας των πραγματικών φυσικών συστημάτων, δεν υπάρχει συνήθως ένας απλός αναλυτικός τύπος για την επαρκή περιγραφή των φαινομένων και απαιτούνται συστηματικά πειράματα για την ανάπτυξη των μοντέλων.

Συγκεκριμένα, συστηματικά πειράματα είναι χρήσιμα στις ακόλουθες περιπτώσεις: 1) το μοντέλο είναι πολύ σύνθετο όπου οι υποκείμενες σχέσεις είναι μη γραμμικές με πολλές μεταβλητές να αλληλεπιδρούν μεταξύ τους, 2) το μοντέλο παρουσιάζει στοχαστική συμπεριφορά, ή 3) το μοντέλο παράγει πολλαπλές και σύνθετες εξόδους, όπως η συναρτησιακή απόκριση ή ακόμα και η 3D απεικόνιση. Σε τέτοιες καταστάσεις, είναι συχνά δύσκολο, ακόμη και αδύνατο να μελετηθεί η συμπεριφορά των μοντέλων υπολογιστών που χρησιμοποιούν παραδοσιακές μεθόδους πειραμάτων. Μια προσέγγιση για τη μελέτη περίπλοκων σχέσεων εισόδου-εξόδου που εμφανίζεται από το μοντέλο προσομοίωσης είναι η κατασκευή ενός προσεγγιστικού μοντέλου (επίσης γνωστό ως μεταμοντέλο-metamodel) με βάση ένα σύνολο περιορισμένων δεδομένων παρατήρησης που αποκτήθηκαν με την εκτέλεση του μοντέλου προσομοίωσης σε προσεκτικά επιλεγμένα σημεία του σχεδιασμού. Αυτό το προσεγγιστικό μοντέλο έχει μια αναλυτική μορφή και είναι απλούστερο στη χρήση και στην εφαρμογή. Η διαθεσιμότητα ενός μεταμοντέλου ως υποκατάστατου του αρχικού πολύπλοκου προσομοιωμένου μοντέλου είναι ιδιαίτερα χρήσιμη όταν η εφαρμογή του αρχικού μοντέλου είναι μια εντατική διαδικασία (π.χ. από την άποψη της προετοιμασίας του μοντέλου, του υπολογισμού και της ανάλυσης της εξόδου), όπως συνήθως συμβαίνει στα πολύπλοκα μοντέλα προσομοίωσης. Η ανάγκη για ένα «φτηνό σε υπολογισμό» μεταμοντέλο είναι ακόμη μεγαλύτερη όταν κάποιος θα ήθελε να χρησιμοποιήσει το μοντέλο για πιο απαιτητικά υπολογιστικά ζητήματα όπως η ανάλυση ευαισθησίας, η βελτιστοποίηση ή η πιθανοτική ανάλυση. Για να εκτελέσουμε με επιτυχία τα προαναφερθέντα ζητήματα, ένα μεταμοντέλο υψηλής ποιότητας που αναπτύχθηκε από ένα περιορισμένο μέγεθος δείγματος είναι μια πολύ σημαντική προϋπόθεση. Για να βρεθεί ένα μεταμοντέλο υψηλής ποιότητας, η επιλογή ενός σωστού συνόλου δεδομένων «εκπαίδευσης» γίνεται ένα σημαντικό ζήτημα για την προσομοίωση. Ο λεγόμενος space-filling σχεδιασμός παρέχει τρόπους για τη δημιουργία αποτελεσματικών δεδο-



μένων εκπαίδευσης. Με την αποδοτική έννοια εννοούμε ότι είναι σε θέση να παράγει ένα σύνολο σημείων δειγματοληψίας που απεικονίζουν τη μέγιστη πληροφορία μεταξύ των σχέσεων εισόδου-εξόδου. Υπάρχουν πολλοί space-filling σχεδιασμοί, όπως η δειγματοληψία λατινικών υπερκύβων και οι τροποποιήσεις τους, οι βέλτιστες δειγματοληψίες λατινικών υπερκύβων και οι ομοιόμορφοι σχεδιασμοί. Οι McKay, Conover και Beckman [103] εισήγαγαν τη δειγματοληψία λατινικών υπερκύβων (Latin hypercube sampling, LHS) ως μία εναλλακτική της iid δειγματοληψίας η οποία οδηγεί στη μείωση της διασποράς της μεταβλητής  $y$ .

## 2.2 Πειράματα υπολογιστών

Πολλές επιστημονικές περιοχές συμπεριλαμβάνουν την μοντελοποίηση φυσικών φαινομένων χρησιμοποιώντας μαθηματικά μοντέλα

$$y = f(x_1, \dots, x_s) = f(x), x = (x_1, \dots, x_s)' \in T$$

Όπου το  $x$  αποτελείται από επεξηγηματικές μεταβλητές που εισάγουμε στο μοντέλο,  $y$  είναι η μεταβλητή απόκρισης, η συνάρτηση  $f$  είτε είναι είτε μπορεί να μην είναι σε αναλυτική μορφή, και  $T$  είναι ο χώρος των μεταβλητών που εισάγουμε. Το παραπάνω μοντέλο μπορεί να θεωρηθεί ως μία λύση ενός συνόλου εξισώσεων, συμπεριλαμβανομένων γραμμικών, μη γραμμικών, συνήθων ή μερικών διαφορικών εξισώσεων, όπου είναι συνήθως απίθανο να αποκτήσουμε μία αναλυτική λύση για αυτές τις εξισώσεις. Μηχανικοί και επιστήμονες χρησιμοποιούν αυτά τα μοντέλα για να πάρουν αποφάσεις εφαρμόζοντας το μοντέλο για την εκτίμηση της συμπεριφοράς συστημάτων υπό τον καθορισμό διαφορετικών μεταβλητών εισόδου. Επομένως, το μοντέλο είναι ένα κρίσιμο στοιχείο στην επιστημονική έρευνα και στον μηχανικό σχεδιασμό. Η διαθεσιμότητα ενός μοντέλου είναι συνήθως κρίσιμη σε πολλές περιπτώσεις διότι μπορεί να είναι είτε πολύ ακριβό είτε ιδιαίτερα χρονοβόρο να πραγματοποιηθούν φυσικά πειράματα για να κατανοήσουμε πλήρως τη σχέση μεταξύ της μεταβλητής απόκρισης  $y$  και των επεξηγηματικών μεταβλητών  $x_j$ . Έτσι τα μοντέλα υπολογιστών έχουν γίνει ιδιαίτερα σημαντικά για να διερευνηθούν περίπλοκα φυσικά φαινόμενα. Οι επιστήμονες και οι μηχανικοί κάνουν χρήση των προσομοιώσεων με υπολογιστές για να διερευνήσουν πολύπλοκες σχέσεις μεταξύ των μεταβλητών εισόδου και εξόδου. Έ-

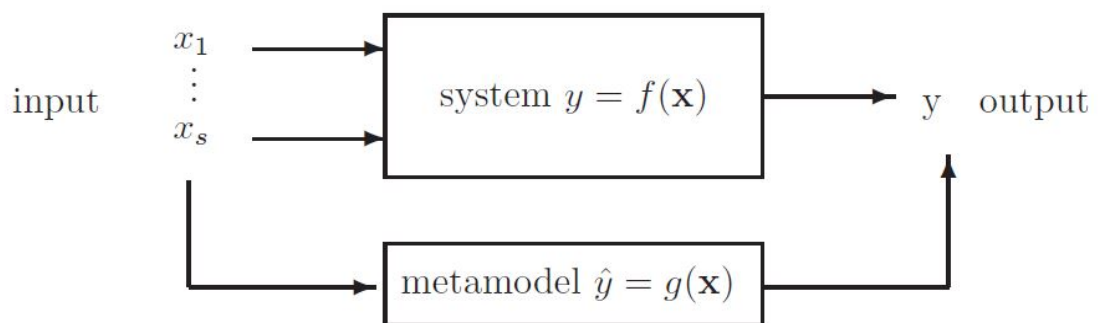
## 20 Εισαγωγικά στοιχεία και βασικές έννοιες για τα πειράματα υπολογιστών

νας από τους στόχους των πειραμάτων υπολογιστών είναι αρχικά να βρουν ένα κατά προσέγγιση μοντέλο που είναι πολύ απλούστερο από το πραγματικό (αλλά περίπλοκο) μοντέλο προσομοιώνοντας την συμπεριφορά μίας συσκευής/διαδικασίας στον υπολογιστή. Η ξεχωριστή φύση των πειραμάτων υπολογιστή θέτουν μία μοναδική πρόκληση, η οποία απαιτεί μία ιδιαίτερη προσέγγιση σχεδιασμού πειραμάτων.

Τις τελευταίες δεκαετίες ο σχεδιασμός πειραμάτων υπολογιστών χαίρουν ιδιαίτερης προσοχής. Εκατοντάδες ερευνητικές εργασίες έχουν δημοσιευτεί αναφερόμενες σε αυτό το θέμα. Ανάμεσα σε αυτές και των [5, 72, 116] δίνουν μία διεξοδική ανασκόπηση σε αυτή τη ραγδαία αναπτυσσόμενη περιοχή και παρέχουν αρκετές αναφορές για το παρόν αντικείμενο. Αξίζει να σημειωθούν τα μοναδικά χαρακτηριστικά των πειραμάτων υπολογιστή σε σύγκριση με τα φυσικά πειράματα:

- Τα πειράματα υπολογιστών συχνά περιέχουν έναν μεγάλο αριθμό μεταβλητών σε σύγκριση με τα τυπικά φυσικά πειράματα.
- Μεγαλύτερη πειραματική περιοχή ή χώρος σχεδιασμού συχνά απασχολείται για τη διεξαγωγή πολύπλοκων μη γραμμικών συναρτήσεων.
- Τα πειράματα υπολογιστών είναι ντετερμινιστικά. Αυτό σημαίνει ότι δείγματα με ακριβώς τις ίδιες επεξηγηματικές μεταβλητές θα παράξουν πανομοιότυπες μεταβλητές εξόδου-μεταβλητές απόκρισης.

Για να γίνουμε πιο σαφείς στα προαναφερθέντα χαρακτηριστικά, τα πειράματα υπολογιστών απαιτούν διαφορετικές προσεγγίσεις από τα φυσικά πειράματα.



Σχήμα 2.1: Πειράματα Υπολογιστών.

## 2.3 Σχεδιασμοί για πειράματα υπολογιστών

Για τα πειράματα υπολογιστών, η επιλογή ενός πειραματικού σχεδιασμού είναι ένα βασικό ζήτημα στην οικοδόμηση ενός αποτελεσματικού μοντέλου. Ο καλύτερος τρόπος να κατασκευαστεί ένας σχεδιασμός εξαρτάται από το υποκείμενο στατιστικό μοντέλο. Το συνολικό μοντέλο του μέσου αποτελεί τη βάση για το κίνητρο της δειγματοληψίας των λατινικών υπερκύβων (Latin hypercube sampling) και των ομοιόμορφων σχεδιασμών. Μια καλή μέθοδος κατασκευής για space-filling σχεδιασμούς θα πρέπει να πληρεί τα ακόλουθα κριτήρια :

- (α) να είναι βέλτιστη κάτω από το υποκείμενο στατιστικό μοντέλο,
- (β) να μπορεί να εξυπηρετήσει διάφορους αριθμούς πειραματικών εκτελέσεων και μεταβλητών εισόδου,
- (γ) να μπορεί εύκολα να παραχθεί σε ένα λογικό υπολογιστικό χρόνο.

Δύο space-filling σχεδιασμοί ονομάζονται ισοδύναμοι αν ο ένας μπορεί να ληφθεί από τον άλλο με μετατόπιση της τάξεως των παραγόντων ή/και των πειραματικών εκτελέσεων του άλλου. Ένας space-filling σχεδιασμός μπορεί να δημιουργηθεί τυχαία, όπως στη δειγματοληψία λατινικών υπερκύβων. Σε αυτή την περίπτωση, ένας σχεδιασμός είναι ένα δείγμα από μια σειρά προτεινόμενων σχεδιασμών. Εμείς πρέπει να χρησιμοποιούμε τον ίδιο συμβολισμό για αυτή την κλάση σχεδιασμών. Για παράδειγμα, το  $LHS(n, s)$  μπορεί να είναι ένα δείγμα Λατινικού υπερκύβου ή το σύνολο όλων αυτών των δειγμάτων. Ένας space-filling σχεδιασμός μπορεί επίσης να είναι ντετερμινιστικός, όπως ο ομοιόμορφος σχεδιασμός. Επομένως, μπορεί να ισχύει ένας συμβολισμός είτε για έναν σχεδιασμό είτε για το σύνολο του ίδιου τύπου σχεδιασμών χωρίς κάποια σύγχυση. Η βελτιστοποίηση διαδραματίζει σημαντικό ρόλο στον σχεδιασμό και τη μοντελοποίηση των πειραμάτων υπολογιστών. Υπάρχουν γνωστοί ευρετικοί συνδυαστικοί αλγόριθμοι βελτιστοποίησης για τη δημιουργία βέλτιστων σχεδιασμών, συμπεριλαμβανομένων της τοπικής αναζήτησης, της προσομοιωτικής δακτυλίωσης (simulated annealing), της αποδοχής κατωφλίου (threshold accepting) και των στοχαστικών εξελικτικών μεθόδων (stochastic evolutionary methods).

## **22 Εισαγωγικά στοιχεία και βασικές έννοιες για τα πειράματα υπολογιστών**

Στην παρούσα διδακτορική διατριβή θα ασχοληθούμε με την επιλογή και τον σχεδιασμό πειραμάτων υπολογιστών και όχι με τη δημιουργία του μεταμοντέλου. Για περισσότερες πληροφορίες σχετικά με τα μεταμοντέλα (metamodels) παραπέμπουμε στην εργασία των Fang et al. [46]. Στο επόμενο κεφάλαιο προτείνουμε γενικούς σχεδιασμούς για πειράματα υπολογιστών. Οι προτεινόμενοι βέλτιστοι σχεδιασμοί έχουν πολλά πλεονεκτήματα. Είναι πολύ εύκολο να κατασκευαστούν, έχουν ευέλικτα μεγέθη και μπορούν να φιλοξενούν πολλά επίπεδα. Αυτά τα πλεονεκτήματα καθιστούν τους σχεδιασμούς αυτού του κεφαλαίου μία καλή εναλλακτική για πειράματα υπολογιστών.

---

## Ορθογώνιοι κατά στήλη και σχεδόν ορθογώνιοι κατά στήλη σχεδιασμοί για μοντέλα με όρους δεύτερης τάξης

---

**Σ**το παρόν κεφάλαιο μελετάμε τους U-type, ορθογώνιους κατά στήλη (column-orthogonal) και τους σχεδόν ορθογώνιους κατά στήλη (nearly column-orthogonal) σχεδιασμούς. Οι σχεδόν ορθογώνιοι κατά στήλη σχεδιασμοί είναι πολύ χρήσιμοι σε περιπτώσεις όπου δεν είναι γνωστοί οι ορθογώνιοι κατά στήλη σχεδιασμοί. Στην παρούσα ενότητα παρουσιάζουμε την κατασκευή νέων σχεδιασμών οι οποίοι είναι κατάλληλοι για πειράματα κρησαρίσματος (screening experiments). Σε κάποιες περιπτώσεις, οι σχεδιασμοί που κατασκευάστηκαν είναι βέλτιστοι όσον αφορά στη δομή των ταυτόσημων επιδράσεων (aliased structure). Οι δομές των ταυτόσημων επιδράσεων των U-type, ορθογώνιων και σχεδόν ορθογώνιων κατά στήλη σχεδιασμών υπολογίζονται και παρουσιάζονται σε κλειστή μορφή. Το γεγονός αυτό κάνει αυτή την καινούργια προσέγγιση καινοτόμα και καθιστά ικανή την κατασκευή σχεδιασμών που είναι διαφορετικοί από τους ήδη υπάρχοντες στη βιβλιογραφία. Επίσης παρουσιάζονται, ένα εκτεταμένο πολλαπλασιαστικό θεώρημα (multiplication theorem), καθώς και νέες άπειρες οικογένειες (infinite families) ορθογώνιων κατά στήλη σχεδιασμών χρησιμοποιώντας ζεύγη περιοδικών ακολουθιών Golay (periodic Golay pairs). Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [III].

### 3.1 Ερευνητικό πρόβλημα-Λύση

Από μία σύντομη ανασκόπηση στην βιβλιογραφία αντιλαμβανόμαστε ότι οι σχεδιασμοί για τα πειράματα υπολογιστών έχουν αναπτυχθεί σε μεγάλο βαθμό τα τελευταία χρόνια. Μία μεγάλη κλάση σχεδιασμών για πειράματα υπολογιστών περιλαμβάνουν τους ιδιαίτερος γνωστούς και συχνά χρησιμοποιούμενους σχεδιασμούς λατινικού υπερκύβου (Latin hypercube designs, LHDs). Ωστόσο, οι σχεδιασμοί που είναι κατάλληλοι για πειράματα υπολογιστών είναι συνήθως περίπλοκοι ενώ είναι υπολογιστικά κοστοβόρο τις περισσότερες φορές να καθοριστεί η δομή των ταυτώσεων επιδράσεων.

Στην παρούσα ενότητα παρουσιάζουμε την κατασκευή νέων σχεδιασμών οι οποίοι είναι κατάλληλοι για πειράματα κρησαρίσματος (screening experiments). Σε κάποιες περιπτώσεις, οι σχεδιασμοί που κατασκευάστηκαν είναι βέλτιστοι όσον αφορά στη δομή των ταυτώσεων επιδράσεων (aliased structure). Οι δομές των ταυτώσεων επιδράσεων των U-type, ορθογώνιων και σχεδόν ορθογώνιων κατά στήλη σχεδιασμών υπολογίζονται και παρουσιάζονται σε κλειστή μορφή κάτι που κάνει αυτή την καινούργια προσέγγιση καινοτόμα και καθιστά ικανή την κατασκευή σχεδιασμών που είναι διαφορετικοί από τους ήδη υπάρχοντες στη βιβλιογραφία. Η κατασκευή των νέων ορθογώνιων κατά στήλη σχεδιασμών (COD) και σχεδόν ορθογώνιων σχεδιασμών (NCOD) πραγματοποιήθηκε συνδυάζοντας γενικευμένους ορθογώνιους σχεδιασμούς, ξένους κυκλικούς πίνακες και περιοδικά ζεύγη Golay. Επίσης παρέχουμε ένα νέο πολλαπλασιαστή με την παρουσίαση νέων κατευθυνόμενων ακολουθιών μήκους 72. Αυτή η προσέγγιση οδηγεί σε μια νέα πολλαπλασιαστική τεχνική και μια νέα άπειρη οικογένεια COD που περιλαμβάνει βέλτιστους σχεδιασμούς πολλών νέων τάξεων. Σημειώνουμε ότι οι βέλτιστοι ορθογώνιοι κατά στήλη σχεδιασμοί σε αυτήν την άπειρη οικογένεια είναι νέων τάξεων και δεν μπορούσαν να κατασκευαστούν με οποιαδήποτε γνωστή μέθοδο στη βιβλιογραφία. Οι προτεινόμενοι βέλτιστοι σχεδιασμοί έχουν πολλά πλεονεκτήματα. Είναι πολύ εύκολο να κατασκευαστούν, έχουν ευέλικτα μεγέθη και μπορούν να φιλοξενούν πολλά επίπεδα. Αυτά τα πλεονεκτήματα καθιστούν τους σχεδιασμούς αυτούς ιδιαίτερα αποδοτικούς για πειράματα υπολογιστών.

## 3.2 Εισαγωγή και βασικές έννοιες

Σχεδιασμοί για τα πειράματα υπολογιστών έχουν εκτενέστατα αναπτυχθεί στην πρόσφατη βιβλιογραφία: για παράδειγμα παραπέμπουμε στις ακόλουθες εργασίες [53, 63, 93, 94, 126, 127, 153, 154]. Μία μεγάλη κλάση σχεδιασμών για πειράματα υπολογιστών, όπως έχουμε ήδη αναφέρει, περιλαμβάνουν τους ιδιαίτερος γνωστούς και συχνά χρησιμοποιούμενους σχεδιασμούς λατινικού υπερκύβου (Latin hypercube designs, LHDs). Αυτοί οι σχεδιασμοί έχουν  $n$  ομοιόμορφα κατανομημένα επίπεδα σε  $n$  πειραματικές εκτελέσεις. Πρόσφατα, οι Sun et al. [128], ακολουθώντας τους Bingham et al. [11], χαλάρωσαν τους περιορισμούς που ισχύουν στα LHDs σχετικά με το μέγεθος των πειραματικών εκτελέσεων ότι πρέπει να ισούται με τον αριθμό των επιπέδων, και εισήγαγαν μία εναλλακτική κλάση σχεδιασμών, τους επονομαζόμενους ορθογώνιους κατά στήλη (column-orthogonal) σχεδιασμούς για πειράματα υπολογιστών. Όταν όλα τα επίπεδα επαναλαμβάνονται τον ίδιο αριθμό φορές και κατανέμονται ομοιόμορφα σε κάθε στήλη, ο σχεδιασμός καλείται U-type σχεδιασμός. Όταν κάθε στήλη είναι ορθογώνια στην επίδραση που αντιστοιχεί στο μέσο (mean effect), ο σχεδιασμός καλείται ορθογώνιος στο μέσο (mean orthogonal).

Ένας σχεδιασμός με  $n$  πειραματικές εκτελέσεις και  $m$  παράγοντες που έχει  $z$  μηδενικά και  $q - 1$  μη μηδενικά διακριτά επίπεδα που επαναλαμβάνονται τον ίδιο αριθμό φορές σε κάθε παράγοντα συμβολίζεται με  $D(n; z; q^m)$ .

- Όταν το  $z = 0$ , τότε ο σχεδιασμός είναι ένας U-type σχεδιασμός και το  $z$  μπορεί να παραληφθεί από το συμβολισμό, δηλαδή  $D(n; (q - 1)^m)$ .
- Όταν το  $z = n/q$ , τότε ο σχεδιασμός είναι ένας U-type σχεδιασμός και το  $z$  μπορεί να παραληφθεί από το συμβολισμό, για παράδειγμα  $D(n; q^m)$ .
- Όταν το  $z \neq 0$  και το  $z \neq n/q$ , τότε οι παραγόμενοι σχεδιασμοί είναι ορθογώνιοι στο μέσο, αλλά όχι U-type, και τέτοιοι σχεδιασμοί είναι κατάλληλοι μόνο για ποσοτικούς παράγοντες.

Ένας σχεδιασμός καλείται ορθογώνιος κατά στήλη σχεδιασμός και συμβολίζεται με  $COD(n; z; q^m)$ , εάν το εσωτερικό γινόμενο οποιωνδήποτε δύο στηλών είναι μηδέν. Ένας σχεδόν ορθογώνιος κατά στήλη σχεδιασμός συμβολίζεται με  $NCOD(n; z; q^m)$ . Σημειώνουμε ότι ένας ορθογώνιος κατά στήλη σχεδιασμός  $COD(n; z; q^m)$  δεν είναι

πάντοτε U-type. Έναν ορθογώνιο κατά στήλη σχεδιασμό  $COD(n; z; q^m)$  που είναι επίσης U-type, θα τον συμβολίζουμε ως U-type  $COD(n; z; q^m)$ . Ένας ορθογώνιος κατά στήλη σχεδιασμός καλείται  $l$ -ορθογώνιος εάν το άθροισμα των γινομένων των στοιχείων (sum of the elementwise products) οποιωνδήποτε  $l$  στηλών (είτε αυτές είναι διακριτές είτε όχι) είναι μηδέν. Όταν  $\ell = 3$ , τότε οι σχεδιασμοί λέγονται 3-ορθογώνιοι [11]. Ο σχεδιασμός  $T$  λέγεται ότι είναι ένας πλήρης fold-over του σχεδιασμού  $D$  εάν  $T = \begin{pmatrix} D \\ -D \end{pmatrix}$ . Οι Georgiou et al. [56] έδειξαν ότι οποιοσδήποτε πλήρης fold-over σχεδιασμός  $T$  είναι 3-ορθογώνιος.

Ένας πίνακας λέγεται ότι είναι κυκλικός εάν κάθε διάνυσμα γραμμή περιστρέφεται κατά ένα στοιχείο δεξιά σε σύγκριση με το προηγούμενο διάνυσμα γραμμή. Ένας κυκλικός πίνακας  $A = circ(B)$  καθορίζεται πλήρως από ένα διάνυσμα  $B$ , το οποίο εμφανίζεται ως η πρώτη γραμμή του πίνακα. Οι υπολοιπούμενες γραμμές του  $A$  είναι η καθεμία κυκλικές μεταθέσεις του διανύσματος  $B$  με μία offset να ισούται με το δείκτη της γραμμής. Οι κυκλικοί πίνακες με μία αντίθετη κατεύθυνση μετακίνησης των στοιχείων ονομάζονται πίσω-κυκλικοί. Εάν δεν προκαλείται σύγχυση, ο κυκλικός πίνακας και ο αντίστοιχος γεννήτορας διάνυσμα γραμμή (generator row vector) θα γραφούν εν συντομία με το ίδιο γράμμα, δηλαδή  $A = circ(A)$ . Τα στοιχεία του πίνακα πληροφορίας ενός κυκλικού πίνακα μπορούν να υπολογιστούν χρησιμοποιώντας την περιοδική συνάρτηση αυτοσυσχέτισης της πρώτης γραμμής του. Ένα σύνολο πινάκων  $\bigcup_{j=1}^{\ell} \{B_j\}$  λέγεται ότι είναι ξένοι αν  $B_i * B_j = 0$  για όλα τα  $i \neq j$ ,  $i, j = 1, 2, \dots, \ell$  όπου  $*$  συμβολίζει το γινόμενο Hadamard.

Έστω  $A = \{A_j : A_j = \{a_{j,0}, a_{j,1}, \dots, a_{j,n-1}\}, j = 1, \dots, \ell\}$ , ένα σύνολο από  $\ell$  διανύσματα μήκους  $n$ . Αυτά τα διανύσματα λέγονται ότι είναι ένα σύνολο από ξένα (δεν έχουν κοινά σημεία, εκεί που έχει μονάδες το ένα, το άλλο έχει μηδενικά) διανύσματα εάν το σύνολο των αντίστοιχων κυκλικών πινάκων  $B_j = \cup_j \{circ(A_j)\}$ ,  $j = 1, \dots, \ell$  είναι ξένα. Η περιοδική συνάρτηση αυτοσυσχέτισης  $P_A(s)$  (εν συντομία PAF) και η μη περιοδική συνάρτηση αυτοσυσχέτισης  $N_A(s)$  (εν συντομία NPAF) ορίζεται, θεωρώντας το δείκτη  $i + s$  modulo  $n$ , ως εξής

$$P_A(s) = \sum_{j=1}^{\ell} \sum_{i=0}^{n-1} a_{j,i} a_{j,i+s}, \quad \text{and} \quad N_A(s) = \sum_{j=1}^{\ell} \sum_{i=0}^{n-s-1} a_{j,i} a_{j,i+s}, \quad s = 0, \dots, n-1$$

αντίστοιχα. Το σύνολο των διανυσμάτων γραμμής (row vectors)  $A$  λέγεται ότι έχει



μηδενική περιοδική συνάρτηση αυτοσυσχέτισης (μηδενικό  $PAF$ ) εάν  $P_A(s) = 0$  και μηδενική μη-περιοδική συνάρτηση αυτοσυσχέτισης (μηδενικό  $NPAF$ ) εάν  $N_A(s) = 0$  ( $s = 1, \dots, n-1$ ).

Ένα ζεύγος πινάκων  $(A, B)$  λέγεται ότι είναι φίλιοι (αντίστοιχα αντί-φίλιοι) εάν  $AB^T - BA^T = 0$  (αντίστοιχα  $AB^T + BA^T = 0$ ). Σύμφωνα με τον Kharaghani [71], ένα σύνολο  $\{B_1, B_2, \dots, B_n\}$  από τετραγωνικούς, πραγματικούς πίνακες λέγεται ότι είναι φίλιοι εάν

$$\sum_{i=1}^n (B_{2i-1}B_{2i}^T - B_{2i}B_{2i-1}^T) = 0. \quad (3.1)$$

Εκκάθαρα, ένα σύνολο από αμοιβαία φίλιους πίνακες είναι φίλιο αλλά το αντίστροφο δεν ισχύει γενικά. Ένα σύνολο πινάκων  $\{B_1, B_2, \dots, B_n\}$  τάξης  $m$  ικανοποιεί την προσθετική ιδιότητα (additive property) εάν

$$\sum_{i=1}^n B_i B_i^T = f I_m. \quad (3.2)$$

### Συμβολισμός

Ο ακόλουθος συμβολισμός είναι απαραίτητος για την παρουσίαση των αποτελεσμάτων.

- $\{.\}$  χρησιμοποιείται για να υποδηλώσουμε ένα σύνολο, ενώ  $[.]$  χρησιμοποιείται για να υποδηλώσουμε ένα πολλαπλό-σύνολο (το ίδιο στοιχείο επιτρέπεται να υπάρχει πολλαπλές φορές σε ένα πολλαπλό σύνολο). Για παράδειγμα το σύνολο  $A = \{a_1, a_2, a_3\}$  είναι το ίδιο με το σύνολο  $B = \{a_1, a_1, a_2, a_2, a_2, a_3\}$  και τα δύο έχουν τρία διακριτά στοιχεία  $a_1, a_2, a_3$ . Από την άλλη μεριά, το πολλαπλό σύνολο  $C = [a_1, a_2, a_3]$  δεν είναι το ίδιο με το πολλαπλό σύνολο  $D = [a_1, a_1, a_2, a_2, a_2, a_3]$  αφού το πρώτο έχει τρία στοιχεία  $a_1, a_2, a_3$  ενώ το δεύτερο έχει έξι στοιχεία  $a_1, a_1, a_2, a_2, a_2, a_3$ , ακόμη και αν και τα δύο έχουν τα ίδια τρία διακριτά στοιχεία  $a_1, a_2, a_3$ .
- $\#(x, S)$  χρησιμοποιείται για να υποδηλώσει τον αριθμό των  $x$  στο πολλαπλό σύνολο  $S$ . Για παράδειγμα  $\#(a_2, D) = 3$  και  $\#(a_3, D) = 1$ .
- $\{S\}$  υποδηλώνει το σύνολο των διακριτών στοιχείων του πολλαπλού συνόλου  $S$ . Για παράδειγμα  $\{C\} = \{D\} = A$ .

- $|x|$  υποδηλώνει την απόλυτη τιμή των στοιχείων  $x$ .

Σε αυτό το σημείο υπενθυμίζουμε κάποιους βασικούς ορισμούς και ιδιότητες των γενικευμένων ορθογώνιων σχεδιασμών. Οι γενικευμένοι ορθογώνιοι σχεδιασμοί και οι ιδιότητές τους εισήχθησαν από τους Georgiou et al. [57].

**Ορισμός 1** Έστω  $D$  ένας  $n \times m$  πίνακας στις αντιμεταθετικές μεταβλητές  $x_1, \dots, x_t$  όπου κάθε μεταβλητή εμφανίζεται σε κάθε στήλη σε μία από τις δύο μορφές  $\pm a_{ij}x_i$ , για κάθε  $i = 1, \dots, t, j = 1, \dots, u_i$ , και  $\sum_{i=0}^t u_i = n$ , όπου  $u_0$  είναι ο αριθμός των μηδενικών σε κάθε στήλη. Θέτουμε  $s_i = \sum_{j=1}^{u_i} a_{ij}^2$ . Τότε  $D$  είναι ένας γενικευμένος ορθογώνιος σχεδιασμός (GOD) εάν και μόνο εάν  $D^T D = (\sum_{i=1}^t s_i x_i^2) I_m$ . Ο σχεδιασμός  $D$  θα συμβολίζεται ως  $D = GOD(n; m; a_{1,1}, \dots, a_{1,u_1}; a_{2,1}, \dots, a_{2,u_2}; \dots; a_{t,1}, \dots, a_{t,u_t})$ . Ένας εναλλακτικός συμβολισμός ενός γενικευμένου ορθογώνιου σχεδιασμού είναι  $D = GOD(n; m; \langle k_{1,1}, a_{1,1} \rangle, \dots, \langle k_{1,u_1}, a_{1,u_1} \rangle; \dots \langle k_{t,1}, a_{t,1} \rangle, \dots, \langle k_{t,u_t}, a_{t,u_t} \rangle)$  όπου  $k_{i,j}$  δηλώνει πόσες φορές η μεταβλητή  $x_i$  έχει συντελεστή  $a_{i,j}$ . Εάν  $k_{i,j} = 1$ , συνήθως γράφουμε  $(\dots, a_{ij}, \dots)$ , αλλιώς γράφουμε  $(\dots, \langle k_{i,j}, a_{ij} \rangle, \dots)$ .

Έστω  $A_i$  ( $i = 1, \dots, 8$ ) οι κυκλικοί πίνακες τάξης  $n$  και  $R_n$  ο πίσω διαγώνιος μοναδιαίος πίνακας τάξης  $n$ . Οι ακόλουθες μορφές είναι γνωστές:

$$C_2 = \begin{pmatrix} A_1 & A_2 R_n \\ -A_2 R_n & A_1 \end{pmatrix}, \quad C_4 = \begin{pmatrix} A_1 & A_2 R_n & A_3 R_n & A_4 R_n \\ -A_2 R_n & A_1 & A_4^T R_n & -A_3^T R_n \\ -A_3 R_n & -A_4^T R_n & A_1 & A_2^T R_n \\ -A_4 R_n & A_3^T R_n & -A_2^T R_n & A_1 \end{pmatrix},$$

$$C_8 = \begin{pmatrix} A_1 & A_2 & A_4 R_n & A_3 R_n & A_6 R_n & A_5 R_n & A_8 R_n & A_7 R_n \\ -A_2 & A_1 & A_3 R_n & -A_4 R_n & A_5 R_n & -A_6 R_n & A_7 R_n & -A_8 R_n \\ -A_4 R_n & -A_3 R_n & A_1 & A_2 & -A_8^T R_n & A_7^T R_n & A_6^T R_n & -A_5^T R_n \\ -A_3 R_n & A_4 R_n & -A_2 & A_1 & A_7^T R_n & A_8^T R_n & -A_5^T R_n & -A_6^T R_n \\ -A_6 R_n & -A_5 R_n & A_8^T R_n & -A_7^T R_n & A_1 & A_2 & -A_4^T R_n & A_3^T R_n \\ -A_5 R_n & A_6 R_n & -A_7^T R_n & -A_8^T R_n & -A_2 & A_1 & A_3^T R_n & A_4^T R_n \\ -A_8 R_n & -A_7 R_n & -A_6^T R_n & A_5^T R_n & A_4^T R_n & -A_3^T R_n & A_1 & A_2 \\ -A_7 R_n & A_8 R_n & A_5^T R_n & A_6^T R_n & -A_3^T R_n & -A_4^T R_n & -A_2 & A_1 \end{pmatrix}.$$

Σημειώνουμε ότι υπάρχει άλλη μία, εναλλακτική, μορφή της  $C_2$  η οποία είναι η ακόλουθη:

$$C_2 = \begin{pmatrix} A_1 & A_2 \\ -A_2^T & A_1^T \end{pmatrix},$$

Οι σχεδιασμοί  $C_2$  και  $C_4$  είναι ορθογώνια όταν οι κυκλικοί πίνακες είναι προσθετικοί [58]. Ο σχεδιασμός  $C_8$  είναι ορθογώνιος όταν οι κυκλικοί πίνακες που χρησιμοποιούνται είναι ταυτόχρονα προσθετικοί και φίλιοι [71].

### 3.3 U-type και ορθογώνιοι κατά στήλη σχεδιασμοί από GODs

Στην ενότητα αυτή παρουσιάζουμε μερικές μεθόδους για την κατασκευή ορθογώνιων κατά στήλη σχεδιασμών χρησιμοποιώντας γενικευμένους ορθογώνιους σχεδιασμούς (GODs), κατάλληλη αντικατάσταση όλων των μεταβλητών και μία fold-over δομή. Το ακόλουθο Θεώρημα περιγράφει τις ιδιότητες που απαιτούνται για έναν GOD έτσι ώστε να είναι χρήσιμο για την κατασκευή των CODs.

**Θεώρημα 1** Έστω

$$D = GOD(n; m; a_{1,1}, \dots, a_{1,u_1}; a_{2,1}, \dots, a_{2,u_2}; \dots; a_{t,1}, \dots, a_{t,u_t})$$

είναι ένας γενικευμένος ορθογώνιος σχεδιασμός στις αντιμεταθετικές μεταβλητές

$$x_1, x_2, \dots, x_t.$$

Ορίζουμε το πολλαπλό σύνολο

$$M = [|a_{1,1}x_1|, \dots, |a_{1,u_1}x_1|, |a_{2,1}x_2|, \dots, |a_{2,u_2}x_2|, \dots, |a_{t,1}x_t|, \dots, |a_{t,u_t}x_t|]$$

και το σύνολο  $W = \{M\} = \{w_1, w_2, \dots, w_q\}$ . Ο πίνακας  $X = [D^T - D^T]^T$  είναι ένας  $2n \times m$  πίνακας και επιλέγοντας οποιοσδήποτε  $\ell \leq m$  από τις στήλες του, δημιουργούμε έναν πίνακα σχεδιασμού για έναν  $COD(2n; (2q)^\ell)$ ,  $\ell \leq m$ . Επιπλέον εάν υπάρχουν κατάλληλοι πραγματικοί αριθμοί  $a_{i,j}$  και  $x_i$ ,  $i = 1, 2, \dots, t$ ,  $j = 1, 2, \dots, u_i$ , τέτοιοι ώστε

(i)  $\#(w_i, M) = \#(w_{i'}, M) = p$  για όλα τα  $i, i' = 1, 2, \dots, q$  και

(ii) τα στοιχεία του  $W$  είναι ομοιόμορφα κατανεμημένα στο  $[\min W, \max W]$ ,

τότε ο σχεδιασμός  $X$  που κατασκευάστηκε είναι ένας  $U$ -type σχεδιασμός.

### **Απόδειξη Θεωρήματος.**

Από την fold-over δομή αυτής της κατασκευής, έχουμε ότι ο παραγόμενος σχεδιασμός  $Q$  είναι πάντοτε 3-ορθογώνιος  $COD(2n; (2q)^m)$  και έτσι οποιεσδήποτε επιλεγόμενες στήλες  $\ell \leq m$  θα συνιστούν έναν 3-ορθογώνιο σχεδιασμό  $COD(2n; (2q)^\ell)$ . Είναι εύκολο να δούμε ότι η συνθήκη (i) επιβάλλει όλες οι στήλες του πίνακα σχεδιασμού  $X$  να είναι ισορροπημένες (κάθε ένα από τα  $2q$  επίπεδα θα εμφανίζεται  $p$  φορές σε κάθε στήλη), ενώ οι συνθήκες (i) και (ii), μαζί, επιβεβαιώνουν την  $U$ -type ιδιότητα του σχεδιασμού. **Τέλος απόδειξης**

**Παράδειγμα 1** Έστω  $A = (3x_1, -2x_2)$  και  $B = (3x_1, 2x_2)$ , όπου  $x_1$  και  $x_2$  είναι αντιμεταθετικές μεταβλητές. Είναι εύκολο να επαληθεύσουμε ότι αυτά τα δύο διανύσματα έχουν μηδενική περιοδική συνάρτηση αυτοσυσχέτισης (PAF). Χρησιμοποιώντας τους αντίστοιχους κυκλικούς πίνακες στον πίνακα  $C_2$ , μπορούμε να κατασκευάσουμε έναν  $GOD(4; 3, 3; 2, 2)$

$$D = \begin{pmatrix} 3x_1 & -2x_2 & 3x_1 & 2x_2 \\ -2x_2 & 3x_1 & 2x_2 & 3x_1 \\ -3x_1 & -2x_2 & 3x_1 & -2x_2 \\ -2x_2 & -3x_1 & -2x_2 & 3x_1 \end{pmatrix}$$

Αντικαταστήσει τη μεταβλητή  $x_1$  με 1 και τη μεταβλητή  $x_2$  με  $1/2$ . Τότε, ο πίνακας  $X = [D^T - D^T]^T$  είναι ένας 3-ορθογώνιος  $U$ -type  $COD(8; 4^\ell)$ ,  $\ell \leq m$  (δείτε Θεώρημα 1) αφού  $M = [1, 1, 3, 3]$  και  $W = \{M\} = \{1, 3\}$ , και έτσι οι συνθήκες (i) και (ii), του Θεωρήματος 1, ικανοποιούνται.

Η ακόλουθη παρατήρηση δείχνει ότι υπάρχουν πολλοί κατάλληλοι GODs που μπορούν να χρησιμοποιηθούν με το Θεώρημα 1 για την παραγωγή των επιθυμητών CODs.

**Παρατήρηση 1** Τα αποτελέσματα του θεωρήματος 1 είναι αρκετά γενικά αφού χρησιμοποιώντας τους αντίστοιχους κυκλικούς πίνακες από οποιαδήποτε 2, 4 ή 8 διανύσματα, που έχουν μηδενικό PAF, στις δομές  $C_2, C_4, C_8$  αντίστοιχα, λαμβάνουμε αρκετούς νέους ορθογώνιους πίνακες και GODs. Με τους περιορισμούς που περιγράψαμε στο θεώρημα 1, αυτά τα GODs θα δημιουργήσουν νέους U-type ορθογώνιους κατά στήλη σχεδιασμούς.

**Πόρισμα 1** Έστω ότι τα  $A = (a_0, a_1, \dots, a_{t-1})$ ,  $B = (b_0, b_1, \dots, b_{t-1})$  είναι δύο ξένα διανύσματα από πραγματικά στοιχεία, μήκους  $t$  που έχουν μηδενικό PAF. Θέτουμε

$$M = [|a_0 + b_0|, |a_1 + b_1|, \dots, |a_{t-1} + b_{t-1}|]$$

και

$$W = \{M\} = \{w_1, w_2, \dots, w_q\}.$$

Εάν  $0 \in W$ , τότε υπάρχει ένας  $COD(4t; (2q - 1)^\ell)$ ,  $\ell \leq 2t$  και ένας  $COD(8t; (2q - 1)^\ell)$ ,  $\ell \leq 4t$ . Εάν  $0 \notin W$ , τότε υπάρχει ένας  $COD(4t; (2q)^\ell)$ ,  $\ell \leq 2t$  και ένας  $COD(8t; (2q)^\ell)$ ,  $\ell \leq 4t$ . Επιπλέον εάν

$$(i) \#(w_i, M) = \#(w_{i'}, M) = p \text{ για όλα τα } i, i' = 1, 2, \dots, q \text{ και}$$

$$(ii) \text{ τα στοιχεία του } W \text{ είναι ομοιόμορφα κατανομημένα στο } [\min W, \max W],$$

τότε ο σχεδιασμός  $X$  που κατασκευάστηκε είναι ένας U-type σχεδιασμός.

### Απόδειξη

Χρησιμοποιώντας τα διανύσματα  $A$  και  $B$  του Θεωρήματος 3.14 των Georgiou et al. [57], αποκτούμε τους GODs που χρειάζεται στο Θεώρημα 1. **Τέλος απόδειξης**

**Παράδειγμα 2** Έστω  $A = (2, 0, 2, 0, -1, 0)$  και  $B = (0, 4, 0, 4, 0, -2)$  είναι δύο ξένα διανύσματα από πραγματικά στοιχεία, μήκους  $t = 6$  που έχουν μηδενικό PAF. Χρησιμοποιώντας τα διανύσματα  $A$  και  $B$  του Θεωρήματος 3.14 [57], αποκτούμε τον δύο-μεταβλητών

$$GODD_1 = GOD(12; 1, 2, 2, 2, 4, 4; 1, 2, 2, 2, 4, 4)$$

και των τεσσάρων μεταβλητών

$$GODD_2 = GOD(24; 1, 2, 2, 2, 4, 4; 1, 2, 2, 2, 4, 4; 1, 2, 2, 2, 4, 4; 1, 2, 2, 2, 4, 4).$$

Χρησιμοποιώντας τον  $D_1$  στο Θεώρημα 1

- με τις δύο μεταβλητές να αντικαθιστούνται με 1, λαμβάνουμε έναν ορθογώνιο κατά στήλη σχεδιασμό  $COD(24; 6^\ell)$ ,  $\ell \leq 12$ .
- με τη μία μεταβλητή να αντικαθίσταται με 1 και την άλλη με 2, λαμβάνουμε έναν ορθογώνιο κατά στήλη σχεδιασμό  $COD(24; 8^\ell)$ ,  $\ell \leq 12$ .
- με τη μία μεταβλητή να αντικαθίσταται με 1 και την άλλη με 4, λαμβάνουμε έναν ορθογώνιο κατά στήλη σχεδιασμό  $COD(24; 10^\ell)$ ,  $\ell \leq 12$ .
- Με τη μία μεταβλητή να αντικαθίσταται με 1 και την άλλη από οποιονδήποτε αριθμό μεγαλύτερο του 4 (5 για παράδειγμα), λαμβάνουμε έναν ορθογώνιο κατά στήλη σχεδιασμό  $COD(24; 12^\ell)$ ,  $\ell \leq 12$ .

Χρησιμοποιώντας τον  $D_2$  στο Θεώρημα 1 με διαφορετική αντικατάσταση των μεταβλητών, μπορούμε να κατασκευάσουμε CODs με 48 εκτελέσεις, 24 παράγοντες σε  $2s$  επίπεδα,  $s = 3, 4, \dots, 12$ . Για παράδειγμα:

- με όλες τις τέσσερις μεταβλητές να αντικαθίστανται με 1, λαμβάνουμε έναν ορθογώνιο κατά στήλη σχεδιασμό  $COD(48; 6^\ell)$ ,  $\ell \leq 24$ .
- με τις τέσσερις μεταβλητές να αντικαθίστανται με  $(1, 5, 25, 105)$ , λαμβάνουμε έναν ορθογώνιο κατά στήλη σχεδιασμό  $COD(48; 24^\ell)$ ,  $\ell \leq 24$ .

Επιπλέον, το πολλαπλό σύνολο  $M = [|2+0|, |0+4|, |2+0|, |0+4|, |-1+0|, |0-2|] = [2, 4, 2, 4, 1, 2]$ . Έχουμε ότι  $\#(1, M) = 1$ ,  $\#(2, M) = 3$ , και  $\#(4, M) = 2$ . Επομένως, η προϋπόθεση (i) του Πορίσματος 1 δεν ικανοποιείται, οπότε ο παραγόμενος COD δεν είναι  $U$ -type.

**Θεώρημα 2** Έστω  $T_1$  και  $T_2$  δύο ξένοι και φίλιοι  $\{0, 1, -1\}$  κυκλικοί πίνακες τάξης  $t$ , που ικανοποιούν την προσθετική ιδιότητα  $T_1 T_1^T + T_2 T_2^T = kI_t$ , τότε

- i) για  $k = t$  υπάρχει ένας U-type  $COD(4t; (2q)^\ell)$ ,  $q = 1, 2$  και ένας  $COD(4t; 3^\ell)$ , ενώ για  $k < t$ , υπάρχει ένας  $COD(4t; (2q + 1)^\ell)$ ,  $q = 1, 2$ ,  $\ell \leq 2t$ .
- ii) για  $k = t$  υπάρχει ένας  $COD(8t; (2q)^\ell)^*$ ,  $q = 1, 2, 3, 4$  και ένας  $COD(8t; (2q - 1)^\ell)$ ,  $q = 2, 3, 4$ , ενώ για  $k < t$ , υπάρχει ένας  $COD(8t; (2q + 1)^\ell)$ ,  $q = 1, 2, 3, 4$ ,  $\ell \leq 4t$ . Σχεδιασμοί που έχουν επισημανθεί με \* είναι U-type όταν  $q = 1, 2, 4$ .
- iii) για  $k = t$  υπάρχει ένας  $COD(16t; (2q)^\ell)^*$ ,  $q = 1, \dots, 8$  και ένας  $COD(16t; (2q - 1)^\ell)$ ,  $q = 2, \dots, 8$ , ενώ για  $k < t$ , υπάρχει ένας  $COD(16t; (2q + 1)^\ell)$ ,  $q = 1, \dots, 8$ ,  $\ell \leq 8t$ . Σχεδιασμοί που έχουν επισημανθεί με \* είναι U-type όταν  $q = 1, 2, 4, 8$ .

### Απόδειξη

Σε κάθε περίπτωση, πρέπει να κατασκευάσουμε τον απαιτούμενο GOD.

- i) Θέτουμε  $A_1 = x_1T_1 + a_1x_1T_2$  και  $A_2 = -a_1x_1T_1 + x_1T_2$ . Χρησιμοποιώντας τους αντίστοιχους κυκλικούς πίνακες των  $A_1$  και  $A_2$  στη δομή  $C_2$ , παίρνουμε έναν γενικευμένο ορθογώνιο σχεδιασμό μιας μεταβλητής  $D = GOD(2t; \underbrace{1, \dots, 1}_{k \text{ times}}, \underbrace{a_1, \dots, a_1}_{k \text{ times}})$ . Φτιάχνουμε τη μεταβλητή του  $D$  σε 1 ( $x_1 = 1$ ). Ένας επιθυμητός σχεδιασμός θα κατασκευαστεί επιλέγοντας μια κατάλληλη τιμή για τη σταθερά  $a_1$  και στη συνέχεια χρησιμοποιώντας τον προκύπτοντα πίνακα στο Θεώρημα 1. Για  $a_1 = 2q - 1$  και  $k = t$ , το Θεώρημα 1 δίνει έναν U-type  $COD(4t; (2q)^\ell)$ ,  $q = 1, 2$ . Για  $a_1 = q$  και  $k < t$ , το Θεώρημα 1 δίνει έναν  $COD(4t; (2q + 1)^\ell)$ ,  $q = 1, 2$ . Για  $a_1 = 0$ , το Θεώρημα 1 θα δίνει πάντα έναν  $COD(4t; 3^\ell)$ .
- ii) Θέτουμε  $A = a_1x_1T_1 + a_2x_2T_2$ ,  $B = -a_2x_2T_1 + a_1x_1T_2$ ,  $C = -a_3x_1T_1 + a_4x_2T_2$ , και  $D = a_4x_2T_1 + a_3x_1T_2$ . Χρησιμοποιώντας τους αντίστοιχους κυκλικούς πίνακες των διανυσμάτων  $A, B, C$ , και  $D$  στη δομή  $C_4$ , αποκτάμε έναν γενικευμένο ορθογώνιο σχεδιασμό  $X = GOD(4t; \langle k, a_1 \rangle, \langle k, a_3 \rangle; \langle k, a_2 \rangle, \langle k, a_4 \rangle)$ . Ο επιθυμητός  $COD$  κατασκευάζεται χρησιμοποιώντας το  $X$  στο Θεώρημα 1 αφού ορίσουμε και τις δύο μεταβλητές ( $x_1$  και  $x_2$ ) ίσες με 1 και

	$k = t$		$k < t$
$q$	$COD(8t; (2q)^\ell)$	$COD(8t; (2q - 1)^\ell)$	$COD(8t; (2q + 1)^t)$
1	$a_1 = a_2 = a_3 = a_4 = 1$	—	$a_1 = a_2 = a_3 = a_4 = 1$
2	$a_1 = a_2 = 1, a_3 = a_4 = 3$	$a_1 = a_2 = a_3 = 1, a_4 = 0$	$a_1 = a_2 = 1, a_3 = a_4 = 2$
3	$a_i = 2i - 1, i = 1, 2, 3,$ $a_4 = 1$	$a_i = i - 1, i = 1, 2, 3,$ $a_4 = 1$	$a_i = i - 1, i = 1, \dots, 4$
4	$a_i = 2i - 1, i = 1, \dots, 4$	$a_i = i - 1, i = 1, \dots, 4$	$a_i = i, i = 1, \dots, 4$

Σημειώνουμε ότι οι παραπάνω συντελεστές των μεταβλητών είναι απλώς ενδεικτικοί και υπάρχουν πολλοί εναλλακτικοί συντελεστές.

- iii) Θέτουμε  $A_1 = a_1x_1T_1 + a_2x_2T_2$ ,  $A_2 = -a_2x_2T_1 + a_1x_1T_2$ ,  $A_3 = -a_3x_1T_1 + a_4x_2T_2$ ,  $A_4 = a_4x_2T_1 + a_3x_1T_2$ ,  $A_5 = -a_5x_1T_1 + a_6x_2T_2$ ,  $A_6 = a_6x_2T_1 + a_5x_1T_2$ ,  $A_7 = -a_7x_1T_1 + a_8x_2T_2$ , και  $A_8 = a_8x_2T_1 + a_7x_1T_2$ . Χρησιμοποιώντας τους αντίστοιχους κυκλικούς πίνακες των διανυσμάτων  $A_1, A_2, A_3, A_4, A_5, A_6, A_7$ , και  $A_8$  στη δομή  $C_8$ , παίρνουμε έναν γενικευμένο ορθογώνιο σχεδιασμό  $D = GOD(8t; \langle k, a_1 \rangle, \langle k, a_3 \rangle, \langle k, a_5 \rangle, \langle k, a_7 \rangle; \langle k, a_2 \rangle, \langle k, a_4 \rangle, \langle k, a_6 \rangle, \langle k, a_8 \rangle)$ . Ο επιθυμητός  $COD$  κατασκευάζεται χρησιμοποιώντας τον  $D$  στο Θεώρημα 1 αφού ορίσουμε και τις δύο μεταβλητές του ( $x_1$  και  $x_2$ ) να ισούνται με 1 και

	$k = t$		$k < t$
$q$	$COD(16t; (2q)^\ell)$	$COD(16t; (2q - 1)^\ell)$	$COD(16t; (2q + 1)^t)$
1	$a_i = 1, i = 1, \dots, 8$	—	$a_i = 1, i = 1, \dots, 8$
2,3	$a_{2i} = a_{2i-1} = 2i - 1,$ $i = 1, \dots, q$ και $a_{2i} =$ $a_{2i-1} = 2i - 2q - 1, i =$ $q + 1, \dots, 4$	$a_{2i} = a_{2i-1} = i - 1,$ $i = 1, \dots, q$ και $a_{2i} =$ $a_{2i-1} = i - q - 1, i =$ $q + 1, \dots, 4$	$a_{2i} = a_{2i-1} = i,$ $i = 1, \dots, q$ και $a_{2i} =$ $a_{2i-1} = i - q, i = q +$ $1, \dots, 4$
$\geq 4$	$a_i = 2i - 1, i = 1, \dots, q$ και $a_i = 2i - 2q - 1, i =$ $q + 1, \dots, 8$	$a_i = i - 1, i = 1, \dots, q$ και $a_i = i - q - 1, i =$ $q + 1, \dots, 8$	$a_i = i, i = 1, \dots, q$ και $a_i = i - q, i = q + 1, \dots, 8$

Σημειώνουμε ότι οι παραπάνω συντελεστές των μεταβλητών είναι απλώς ενδεικτικοί και υπάρχουν πολλοί εναλλακτικοί συντελεστές.



### Τέλος απόδειξης

Όταν το  $k < t$ , ο  $COD$  που δημιουργείται από το Θεώρημα 2 μπορεί να είναι ή να μην είναι U-type. Για να είναι τέτοιοι σχεδιασμοί U-type, πρέπει να διασφαλίσουμε ότι όλα τα επίπεδα (συμπεριλαμβανομένου του μηδενικού επιπέδου) εμφανίζονται εξίσου συχνά. Οι περιορισμοί του Θεωρήματος 2 δεν είναι αυστηροί. Στην πραγματικότητα, υπάρχουν πολλές άπειρες οικογένειες ζευγών διανυσμάτων που πληρούν τις προϋποθέσεις που απαιτούνται από το Θεώρημα 2. Μερικά μικρά παραδείγματα παρέχονται στον Πίνακα 3.1. Η πρώτη στήλη του Πίνακα 3.1 δίνει τα μήκη των διανυσμάτων που δημιουργήθηκαν. Για καθένα από τα μήκη ακολουθήστε την κατασκευή που περιγράφεται στην Παρατήρηση 2.

**Πίνακας 3.1:** Παραδείγματα ξένων, φίλων κυκλικών πινάκων

Τάξη $t =$ ( $n \geq 0$ )	Βάρος $k =$	Διανύσματα Γεννήτορα $T_1 \cdot T_2$	Μηδενικό
$n + 1$	1	1 ; 0	<i>NPAF</i>
$n + 2$	2	1,0 ; 0,1	<i>NPAF</i>
$n + 4$	4	0,1,0,1 ; 1,0,-1,0	<i>NPAF</i>
$n + 6$	4	0,0,1,0,0,1 ; 0,-1,0,0,1,0	<i>NPAF</i>
$n + 6$	5	0,1,0,-1,0,1 ; 0,0,1,0,1,0	<i>NPAF</i>
$7(n + 1)$	4	0,0,1,0,1,1,-1 ; 0,0,0,0,0,0,0	<i>PAF</i>
$n + 8$	8	1,1,1,0,-1,1,-1,0 ; 0,0,0,1,0,0,0,1	<i>NPAF</i>
$10(n + 1)$	9	0,1,0,1,0,-1,0,1,0,1 ; 0,0,1,0,-1,0,-1,0,1,0	<i>PAF</i>
$12(n + 1)$	8	0,1,1,0,1,0,0,-1,1,0,-1,0 ; 0,0,0,0,0,1,0,0,0,0,0,1	<i>PAF</i>
$13(n + 1)$	9	0,0,1,0,1,1,1,-1,-1,0,1,-1,1 ; 0,0,0,0,0,0,0,0,0,0,0,0,0	<i>PAF</i>
$14(n + 1)$	8	0,0,1,0,1,0,-1,0,0,0,0,0,1,0 ; 0,0,0,0,0,1,0,0,0,1,0,1,0,-1	<i>PAF</i>
$n + 14$	10	1,0,0,0,0,0,0,0,1,0,0,0,1,0 ; 0,1,0,-1,0,1,0,1,0,-1,0,-1,0,-1	<i>NPAF</i>
$14(n + 1)$	13	0,1,0,1,0,-1,0,1,0,-1,0,1,0,1 ; 0,0,1,0,1,0,-1,0,-1,0,1,0,1,0	<i>PAF</i>

**Παρατήρηση 2** Θέτουμε  $T_1$  και  $T_2$  να είναι ένα ζευγάρι διανυσμάτων που δίνονται στον Πίνακα 3.1. Εάν τα διανύσματα έχουν μηδενικό *NPAF* τότε μπορούν να επεκταθούν σε μήκος με τη συνένωση ενός διανύσματος  $0_n$  μήκους  $n$  με όλες τις καταχωρήσεις του μηδενός. Για παράδειγμα, για  $k = 2$  μπορούμε να έχουμε  $T_1 = circ(1, 0, 0_n)$  και  $T_2 = circ(0, 1, 0_n)$  για κάθε  $n \geq 0$ . Εάν τα διανύσματα έχουν μηδενικό *PAF* τότε μπορούν να επεκταθούν σε μήκος εισάγοντας ένα διάνυσμα  $0_n$  μήκους  $n$  με όλες τις

καταχωρήσεις του μηδενός μεταξύ των στοιχείων και των δύο ακολουθιών. Για παράδειγμα, για  $k = 9$  έχετε δύο κυκλικούς πίνακες τάξεως  $10(n + 1)$  για κάθε  $n \geq 0$  χρησιμοποιώντας

$$T_1 = \text{circ}(0, 0_n, 1, 0_n, 0, 0_n, 1, 0_n, 0, 0_n, -1, 0_n, 0, 0_n, 1, 0_n, 0, 0_n, 1, 0_n),$$

$$T_2 = \text{circ}(0, 0_n, 0, 0_n, 1, 0_n, 0, 0_n, -1, 0_n, 0, 0_n, -1, 0_n, 0, 0_n, 1, 0_n, 0, 0_n)$$

**Παράδειγμα 3** Για  $n = 2$  και  $k = 2$ , επιλέγουμε  $T_1 = (1, 0, 0)$  και  $T_2 = (0, 1, 0)$  από τον Πίνακα 3.1. Είναι εύκολο να δούμε ότι τα  $T_1$  και  $T_2$  μπορούν να δημιουργήσουν δύο ξένους και φίλιους  $\{0, 1, -1\}$  κυκλικούς πίνακες που ικανοποιούν την προσθετική ιδιότητα  $T_1 T_1^T + T_2 T_2^T = 2I_3$ . Έχουμε ότι  $k = 2 < 3 = t$ . Επιλέγουμε  $a_1 = 2$  για να κατασκευάσουμε έναν 5-επιπέδων  $COD(12; 5^6)$  και  $a_1 = 1$  για να κατασκευάσουμε έναν τριών επιπέδων  $COD(12; 3^6)$ . Τα δύο ζεύγη κυκλικών πινάκων που χρειάζονται για τους δύο σχεδιασμούς θα είναι:  $A_1 = x_1 T_1 + 2x_1 T_2 = (x_1, 2x_1, 0)$ ,  $A_2 = -2x_1 T_1 + x_1 T_2 = (-2x_1, x_1, 0)$  και  $A_1 = x_1 T_1 + x_1 T_2 = (x_1, x_1, 0)$ ,  $A_2 = -x_1 T_1 + x_1 T_2 = (-x_1, x_1, 0)$ . Χρησιμοποιώντας τους αντίστοιχους κυκλικούς πίνακες των κατάλληλων διανυσμάτων  $A_1$  και  $A_2$  στη δομή  $C_2$ , αποκτάμε έναν γενικευμένο ορθογώνιο σχεδιασμό μιας μεταβλητής

$$\left( \begin{array}{cccccc} \underline{D_5 = GOD(6; 1, 1, 2, 2)} & & \underline{D_3 = GOD(6; 1, 1, 1, 1)} \\ x_1 & 2x_1 & 0 & -2x_1 & x_1 & 0 \\ 0 & x_1 & 2x_1 & 0 & -2x_1 & x_1 \\ 2x_1 & 0 & x_1 & x_1 & 0 & -2x_1 \\ 2x_1 & 0 & -x_1 & x_1 & 0 & 2x_1 \\ -x_1 & 2x_1 & 0 & 2x_1 & x_1 & 0 \\ 0 & -x_1 & 2x_1 & 0 & 2x_1 & x_1 \end{array} \right) \left( \begin{array}{cccccc} x_1 & x_1 & 0 & -x_1 & x_1 & 0 \\ 0 & x_1 & x_1 & 0 & -x_1 & x_1 \\ x_1 & 0 & x_1 & x_1 & 0 & -x_1 \\ x_1 & 0 & -x_1 & x_1 & 0 & x_1 \\ -x_1 & x_1 & 0 & x_1 & x_1 & 0 \\ 0 & -x_1 & x_1 & 0 & x_1 & x_1 \end{array} \right).$$

Τα επιθυμητά  $CODs$  κατασκευάζονται στη συνέχεια αντικαθιστώντας τη μοναδική μεταβλητή των  $D_5$  και  $D_3$  με 1 και στη συνέχεια χρησιμοποιώντας τον παραγόμενο

πίνακα στο Θεώρημα 1. Αυτά θα είναι:

$$\left( \begin{array}{cccccc} \text{COD}(12; 5^6) & & \text{COD}(12; 3^6) & & & \\ \left( \begin{array}{cccccc} 1 & 2 & 0 & -2 & 1 & 0 \\ 0 & 1 & 2 & 0 & -2 & 1 \\ 2 & 0 & 1 & 1 & 0 & -2 \\ 2 & 0 & -1 & 1 & 0 & 2 \\ -1 & 2 & 0 & 2 & 1 & 0 \\ 0 & -1 & 2 & 0 & 2 & 1 \\ -1 & -2 & 0 & 2 & -1 & 0 \\ 0 & -1 & -2 & 0 & 2 & -1 \\ -2 & 0 & -1 & -1 & 0 & 2 \\ -2 & 0 & 1 & -1 & 0 & -2 \\ 1 & -2 & 0 & -2 & -1 & 0 \\ 0 & 1 & -2 & 0 & -2 & -1 \end{array} \right) & \left( \begin{array}{cccccc} 1 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 \\ 1 & 0 & 1 & 1 & 0 & -1 \\ 1 & 0 & -1 & 1 & 0 & 1 \\ -1 & 1 & 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 & -1 & 0 \\ 0 & -1 & -1 & 0 & 1 & -1 \\ -1 & 0 & -1 & -1 & 0 & 1 \\ -1 & 0 & 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 & -1 & -1 \end{array} \right) & \end{array} \right).$$

Είναι απλό να επαληθεύσετε ότι και οι δύο σχεδιασμοί είναι CODs. Ο σχεδιασμός τριών επιπέδων είναι U-type, αφού οποιοδήποτε από τα επίπεδα  $(1, 0, -1)$  εμφανίζεται τέσσερις φορές σε κάθε στήλη. Καταλήγουμε στο συμπέρασμα ότι ακόμα και όταν  $k < t$ , αυτή η μέθοδος μπορεί να παράγει U-type CODs.

**Παρατήρηση 3** Με την προσθήκη ενός κεντρικού σημείου (μηδενική γραμμή  $0_n^T$ ) στους προκύπτοντες πίνακες σχεδιασμού του Θεωρήματος 2, η ορθογωνικότητα κατά στήλη δεν επηρεάζεται. Έτσι, από το Θεώρημα 2, μπορούμε να κατασκευάσουμε

- i)  $COD(4t + 1; (2q + 1)^\ell)$ ,  $q = 1, 2, \ell \leq 2t$ .
- ii)  $COD(8t + 1; (2q + 1)^\ell)$ ,  $q = 1, 2, 3, 4, \ell \leq 4t$ .
- iii)  $COD(16t + 1; (2q + 1)^\ell)$ ,  $q = 1, \dots, 8, \ell \leq 8t$ .

### 3.4 U-type και σχεδόν ορθογώνιοι κατά στήλη σχεδιασμοί από GODs

Σε αυτή την ενότητα, δείχνουμε ότι οι μέθοδοι της προηγούμενης ενότητας μπορούν να χρησιμοποιηθούν για την κατασκευή σχεδόν ορθογώνιων κατά στήλη σχεδιασμών στην περίπτωση που το πρόβλημα της ύπαρξης ορθογώνιων κατά στήλη σχεδιασμών για συγκεκριμένα μεγέθη πειραματικών εκτελέσεων είναι ανοιχτό. Το ακόλουθο θεώρημα περιγράφει τις ιδιότητες που απαιτούνται για να είναι ένας GOD χρήσιμος για την κατασκευή σχεδόν ορθογώνιων κατά στήλη σχεδιασμών.

**Θεώρημα 3** Έστω

$$D = GOD(n; m; a_{1,1}, \dots, a_{1,u_1}; a_{2,1}, \dots, a_{2,u_2}; \dots; a_{t,1}, \dots, a_{t,u_t})$$

είναι ένας γενικευμένος ορθογώνιος σχεδιασμός στις αντιμεταθετικές μεταβλητές

$$x_1, x_2, \dots, x_t.$$

Ορίζουμε τα πολλαπλά σύνολα

$$M_1 = [ |a_{1,1}x_1|, \dots, |a_{1,u_1}x_1|, |a_{2,1}x_2|, \dots, |a_{2,u_2}x_2|, \dots, |a_{t,1}x_t|, \dots, |a_{t,u_t}x_t|, 0, 1 ]$$

,

$$M_2 = [ |a_{1,1}x_1|, \dots, |a_{1,u_1}x_1|, |a_{2,1}x_2|, \dots, |a_{2,u_2}x_2|, \dots, |a_{t,1}x_t|, \dots, |a_{t,u_t}x_t|, 0, 1, 2 ]$$

και τα αντίστοιχα σύνολα

$$W_1 = \{M_1\} = \{w_1^{(1)}, w_2^{(1)}, \dots, w_{q_1}^{(1)}\}$$

,

$$W_2 = \{M_2\} = \{w_1^{(2)}, w_2^{(2)}, \dots, w_{q_2}^{(2)}\}.$$

Ο πίνακας σχεδιασμού  $L_1 = [D^T \ 1_n \ 0_n \ -1_n \ -D^T]^T$  είναι ένας NCOD( $2n + 3; (2q_1 - 1)^\ell$ ), ενώ ο πίνακας σχεδιασμού  $L_2 = [D^T \ 2_n \ 1_n \ 0_n \ -1_n \ -2_n \ -D^T]^T$  είναι

ένος  $NCOD(2n + 5; (2q_2 - 1)^\ell)$ ,  $\ell \leq m$ . Επιπλέον, για  $z \in \{1, 2\}$ , εάν υπάρχουν κατάλληλοι πραγματικοί αριθμοί  $a_{i,j}^{(z)}$  και  $x_i^{(z)}$ ,  $i = 1, 2, \dots, t$ ,  $j = 1, 2, \dots, u_i$ , έτσι ώστε (i)  $\#(w_i^{(z)}, M_z) = \#(w_{i'}^{(z)}, M_z) = p$  για όλα τα  $i, i' = 1, 2, \dots, q_z$  και (ii) τα στοιχεία του  $W_z$  είναι ομοιόμορφα κατανεμημένα στο  $[\min W_z, \max W_z]$ , τότε ο παραγόμενος σχεδιασμός  $L_z$  είναι U-type.

Η ακόλουθη παρατήρηση δείχνει ότι υπάρχουν πολλοί κατάλληλοι GODs που μπορούν να χρησιμοποιηθούν με το Θεώρημα 3 για τη δημιουργία των επιθυμητών CODs.

**Παρατήρηση 4** Χρησιμοποιώντας κυκλικούς πίνακες από 2, 4, ή 8 διανύσματα με μηδέν PAF στη δομή  $C_2$ ,  $C_4$ , ή  $C_8$ , αντίστοιχα, λαμβάνουμε ορθογώνιους πίνακες και GODs. Με τους περιορισμούς που περιγράφονται στο Θεώρημα 3 και την προσθήκη σειρών από μηδενικά, μονάδες και δύο, ανάλογα με τις ανάγκες, αυτοί οι GODs θα δημιουργήσουν σχεδόν ορθογώνιους κατά στήλη σχεδιασμούς.

**Θεώρημα 4** Έστω  $T_1$  και  $T_2$  δύο ξένοι και φίλιοι  $\{0, 1, -1\}$  κυκλικοί πίνακες τάξης  $t$ , ικανοποιώντας την προσθετική ιδιότητα  $T_1 T_1^T + T_2 T_2^T = kI_t$ , τότε υπάρχει

i) ένας  $NCOD(4t + 3; (2q + 3)^\ell)$  και ένας  $NCOD(4t + 5; (2q + 5)^\ell)$ ,  $q = 1, 2$ ,  
 $\ell \leq 2t$ .

ii) ένας  $NCOD(8t + 3; (2q + 3)^\ell)$  και ένας  $NCOD(8t + 5; (2q + 5)^\ell)$ ,  $q = 1, 2, 3, 4$ ,  
 $\ell \leq 4t$ .

iii) ένας  $NCOD(16t + 3; (2q + 3)^\ell)$  και ένας  $NCOD(16t + 5; (2q + 5)^\ell)$ ,  $q = 1, \dots, 8$ ,  
 $\ell \leq 8t$ .

### Απόδειξη

Σε κάθε περίπτωση, πρέπει να κατασκευάσουμε τον απαιτούμενο GOD. Η κατασκευή είναι όπως στο Θεώρημα 1, αλλά η αντικατάσταση της μεταβλητής και των συντελεστών είναι ελαφρώς διαφορετική και πρέπει να προσδιοριστεί.

i) Κατασκευάζουμε το  $D$  όπως στο Θεώρημα 2. Για να πάρουμε έναν  $NCOD(4t + 3; (2q + 3)^\ell)$ , απλά ορίζουμε τη μεταβλητή του  $D$  σε  $2q - 1$  ( $x_1 = 2q - 1$ ),

$a_1 = q/(2q - 1)$  και χρησιμοποιούμε τον  $L_1$  στο Θεώρημα 3, ενώ για τον  $NCOD(4t+5; (2q+5)^\ell)$  χρησιμοποιούμε το  $x_1 = q+2$  και  $a_1 = (2q-1)/(q+2)$  με  $L_2$  στο Θεώρημα 3.

ii) και iii) Η απόδειξη είναι όπως στο Θεώρημα 2 με τη διαφορά ότι χρησιμοποιούμε τις δομές  $L_1$  και  $L_2$ , που περιγράφονται στο Θεώρημα 3. Στη συνέχεια, και οι δύο μεταβλητές ( $x_1$  και  $x_2$ ) είναι ίσες με 1 και όλοι οι συντελεστές αντικαθίστανται με τιμές ίσες με εκείνες στην απόδειξη του Θεωρήματος 2 (περίπτωση  $k < t$ ) συν 1 (για  $L_1$ ) και συν 2 (για  $L_2$ ).

Σημειώνουμε ότι όλοι οι παραπάνω συντελεστές των μεταβλητών είναι απλώς ενδεικτικοί και υπάρχουν πολλοί εναλλακτικοί συντελεστές. **Τέλος απόδειξης**

## **3.5 Κατασκευές χρησιμοποιώντας περιοδικά ζεύγη Golay**

Σε μια πρόσφατη δημοσίευση των Georgiou et al. [54], οι ίδιοι συγγραφείς πρότειναν για πρώτη φορά μια πολλαπλασιαστική μεθοδολογία για τα περιοδικά ζεύγη Golay (periodic Golay pairs, PGPs) με στόχο τις κατευθυνόμενες (directed) ακολουθίες. Στη συνέχεια κατασκεύασαν πολλές άπειρες οικογένειες PGP καθώς και ορθογώνιους σχεδιασμούς και γενικευμένους ορθογώνιους σχεδιασμούς. Αυτά χρησιμοποιούνται για την παραγωγή πολλών χρήσιμων ορθογώνιων κατά στήλη σχεδιασμών για πειράματα υπολογιστών με καλές στατιστικές ιδιότητες.

### **3.5.1 Νέα αποτελέσματα στα περιοδικά ζεύγη Golay**

Οι ακολουθίες Golay, επίσης γνωστές ως ζεύγη Golay, είναι δύο διανύσματα  $(A_1, A_2)$  μήκους  $n$ , με στοιχεία από το σύνολο  $\{1, -1\}$ , που έχουν μηδενική μη περιοδική συνάρτηση αυτοσυσχέτισης. Αυτές οι ακολουθίες είναι γνωστό ότι υπάρχουν για τα μήκη  $g = 2^a 10^b 26^c$ , όπου  $g = 2^a 10^b 26^c$  είναι μη αρνητικοί ακέραιοι αριθμοί. Περισσότερες λεπτομέρειες μπορείτε να βρείτε στους Borwein και Ferguson [14]. Τα περιοδικά ζεύγη Golay (Bomer και Antweiler, 1990) είναι η φυσική γενίκευση των ακολουθιών Golay με μηδενική περιοδική συνάρτηση αυτοσυσχέτισης. Οποιοσδήποτε θετικός ακέραιος  $g$  για τον οποίο υπάρχουν ακολουθίες Golay

ονομάζεται αριθμός Golay . Ομοίως, κάθε θετικός ακέραιος  $p$  για τον οποίο υπάρχει ένα περιοδικό ζεύγος Golay ονομάζεται περιοδικός αριθμός Golay. Είναι προφανές ότι οποιοσδήποτε αριθμός Golay είναι επίσης ένας περιοδικός αριθμός Golay, αλλά το αντίστροφο δεν είναι γενικά αληθές. Τόσο τα περιοδικά ζεύγη Golay όσο και οι ακολουθίες Golay μπορούν να χρησιμοποιηθούν για την παραγωγή ορθογώνιων κατά μπλοκ κυκλικών πινάκων, παρόλο που οι τελευταίες μπορούν να παρέχουν μερικές χρήσιμες πολλαπλασιαστικές μεθόδους και να δώσουν πιο γενικά αποτελέσματα. Η ύπαρξη περιοδικών ζευγών Golay είναι σημαντική ειδικά όταν δεν υπάρχουν ακολουθίες Golay, δεν μπορούν να κατασκευαστούν ή είναι άγνωστες.

Τα περιοδικά ζεύγη Golay μήκους  $n$  είναι επίσης δύσκολο να κατασκευαστούν. Τα περιοδικά ζεύγη Golay που δεν είναι ακολουθίες Golay είναι γνωστές μόνο για μήκη  $n = 34, 50, 58, 74, 82, 122, 136, 164, 202, 226$  και όλες αυτές έχουν κατασκευαστεί με εντατική αναζήτηση μέσω υπολογιστών [33]. Με τη χρήση του Λήμματος 1 και του Θεωρήματος 5 κατασκευάζουμε νέα περιοδικά ζεύγη Golay, άπειρης ποικιλίας διαφορετικών μηκών, χωρίς αναζήτηση μέσω υπολογιστών.

**Λήμμα 1** *Υπάρχουν κατευθυνόμενες ακολουθίες με μηδενική περιοδική συνάρτηση αυτοσυσχέτισης μήκους  $p$  και τύπου  $(p, p)$ , για  $p \in P = \{34, 50, 58, 74, 82, 122, 136, 202, 226\}$ .*

### Απόδειξη

Για την απόκτηση των απαιτούμενων κατευθυνόμενων ακολουθιών εφαρμόζουμε τις τεχνικές κατασκευής από τους Koukouvinos and Seberry [78] σε περιοδικά ζεύγη Golay. Οι προκύπτουσες κατευθυνόμενες ακολουθίες παρέχονται στον Πίνακα 3.2.

### Τέλος απόδειξης







### 3.5.2 Κατασκευή ορθογώνιων σχεδιασμών

Έστω  $A_i$  ( $i = 1, \dots, 8$ ) κυκλικοί πίνακες τάξεως  $n$  και  $R_n$  ο πίσω διαγώνιος μοναδιαίος πίνακας τάξεως  $n$ . Οι ακόλουθες κατασκευές είναι γνωστές.

$$C_2 = \begin{pmatrix} A_1 & A_2 R_n \\ -A_2 R_n & A_1 \end{pmatrix}, \quad C_4 = \begin{pmatrix} A_1 & A_2 R_n & A_3 R_n & A_4 R_n \\ -A_2 R_n & A_1 & A_4^T R_n & -A_3^T R_n \\ -A_3 R_n & -A_4^T R_n & A_1 & A_2^T R_n \\ -A_4 R_n & A_3^T R_n & -A_2^T R_n & A_1 \end{pmatrix},$$

$$C_8 = \begin{pmatrix} A_1 & A_2 & A_4 R_n & A_3 R_n & A_6 R_n & A_5 R_n & A_8 R_n & A_7 R_n \\ -A_2 & A_1 & A_3 R_n & -A_4 R_n & A_5 R_n & -A_6 R_n & A_7 R_n & -A_8 R_n \\ -A_4 R_n & -A_3 R_n & A_1 & A_2 & -A_8^T R_n & A_7^T R_n & A_6^T R_n & -A_5^T R_n \\ -A_3 R_n & A_4 R_n & -A_2 & A_1 & A_7^T R_n & A_8^T R_n & -A_5^T R_n & -A_6^T R_n \\ -A_6 R_n & -A_5 R_n & A_8^T R_n & -A_7^T R_n & A_1 & A_2 & -A_4^T R_n & A_3^T R_n \\ -A_5 R_n & A_6 R_n & -A_7^T R_n & -A_8^T R_n & -A_2 & A_1 & A_3^T R_n & A_4^T R_n \\ -A_8 R_n & -A_7 R_n & -A_6^T R_n & A_5^T R_n & A_4^T R_n & -A_3^T R_n & A_1 & A_2 \\ -A_7 R_n & A_8 R_n & A_5^T R_n & A_6^T R_n & -A_3^T R_n & -A_4^T R_n & -A_2 & A_1 \end{pmatrix}.$$

Οι σχεδιασμοί  $C_2$  και  $C_4$  είναι ορθογώνιοι όταν οι κυκλικοί πίνακες που χρησιμοποιούνται είναι προσθετικοί [58]. Ο σχεδιασμός  $C_8$  είναι ορθογώνιος όταν οι κυκλικοί πίνακες που χρησιμοποιούνται είναι ταυτόχρονα προσθετικοί και φίλιοι [71].

Από τα περιοδικά ζεύγη Golay  $T_1, T_2$  μήκους  $t$  και τους αντίστοιχους κυκλικούς πίνακες  $Z_1 = circ(T_1), Z_2 = circ(T_2)$  μπορούμε να ορίσουμε

$$X = (Z_1 + Z_2)/2, \quad W = (Z_1 - Z_2)/2. \quad (3.3)$$

**Λήμμα 2** Οι πίνακες  $X$  και  $W$  στην (3.8) είναι κυκλικοί, ξένοι, προσθετικοί, φίλιοι πίνακες τάξεως  $t$ , με εισόδους από το σύνολο  $\{0, 1, -1\}$  και  $X \pm W$  είναι ένας  $\pm 1$  πίνακας.

#### Απόδειξη

Προκύπτει με απλούς υπολογισμούς. **Τέλος απόδειξης**

Έστω  $(T_1, T_2)$  το περιοδικό ζεύγος Golay μήκους  $t = 2^{a_1} 10^{a_2} 26^{a_3} p^{a_4}$ , όπου  $a_1, a_2, a_3$

είναι μη-αρνητικοί ακέραιοι,  $p \in \{34, 50, 58, 74, 82, 122, 202, 226\}$  και  $a_4 = 0$  ή  $1$ .

Χρησιμοποιώντας τους πίνακες  $X$  και  $W$  από την (3.8), θεωρούμε τις κατασκευές

$$A_i = (2i - 1)X + (2i + 1)W, \quad A_{i+1} = (2i + 1)X - (2i - 1)W, \quad (i = 1, 3, 5, 7), \quad (2)$$

$$A_{2i-1} = (2i - 1)X + 2iW, \quad A_{2i} = 2iX - (2i - 1)W, \quad (i = 1, 2, 3, 4), \quad (3)$$

$$A_i = (2i + 1)X + (2i + 3)W, \quad A_{i+1} = (2i + 3)X - (2i + 1)W, \quad (i = 1, 3, 5, 7), \quad (4)$$

$$A_{2i-1} = 2iX + (2i + 1)W, \quad A_{2i} = (2i + 1)X - 2iW, \quad (i = 1, 2, 3, 4), \quad (5)$$

$$A_i = (2i + 3)X + (2i + 5)W, \quad A_{i+1} = (2i + 5)X - (2i + 3)W, \quad (i = 1, 3, 5, 7), \quad (6)$$

$$A_{2i-1} = (2i + 1)X + 2(i + 1)W, \quad A_{2i} = 2(i + 1)X - (2i + 1)W, \quad (i = 1, 2, 3, 4). \quad (7)$$

**Θεώρημα 6** Έστω  $t = 2^{a_1}10^{a_2}26^{a_3}p^{a_4}$ , όπου  $a_1, a_2, a_3$  είναι μη-αρνητικοί ακέραιοι,  $p \in \{34, 50, 58, 74, 82, 122, 202, 226\}$  και  $a_4 = 0$  ή  $1$ . Υπάρχουν οι παρακάτω ορθογώνιοι σχεδιασμοί για πειράματα υπολογιστών: (i)  $OD(2^{b+1}t; q_1^\ell)$  και (ii)  $OD(2^{b+1}t + s; q_2^\ell)$ ,  $q_1 = 2^{b+1}$ ,  $q_2 = 2^{b+1} + 1$ ,  $\ell \leq 2^{bt}$ , για  $b = 1, 2, 3$  και για κάθε θετικό ακέραιο  $s$ .

### Απόδειξη

Για να αποδείξουμε την ύπαρξη των ορθογώνιων σχεδιασμών (i) και (ii) του θεωρήματος χρησιμοποιούμε τις κατασκευές που δίνονται στο (2) και στο (3) αντίστοιχα. Από το Λήμμα 4 έχουμε ότι οι πίνακες  $\{A_i\}_{i=1}^8$  είναι κυκλικοί, προσθετικοί και φίλιοι. Για  $b = 1, 2, 3$  και ένα συγκεκριμένο  $1 \leq \ell \leq 2^{bt}$ , ο επιθυμητός ορθογώνιος σχεδιασμός προκύπτει από την τυχαία επιλογή οποιωνδήποτε  $\ell$  στηλών από τον πίνακα  $(C_{2^b}^T, -C_{2^b}^T)^T$  ή  $(C_{2^b}^T, 0_{2^b t \times s}, -C_{2^b}^T)^T$  για (i) ή (ii) αντίστοιχα, όπου  $C_{2^b}$  είναι ο τετραγωνικός πίνακας τάξης  $2^{bt}$  που προκύπτει χρησιμοποιώντας τα  $\{A_1, \dots, A_{2^b}\}$ .

### Τέλος απόδειξης

Για να δείξουμε την κατασκευή που περιγράφεται στο Θεώρημα 8, παρέχουμε ένα λεπτομερές παράδειγμα. Παρόλο που ο επιθυμητός σχεδιασμός σε αυτό το παράδειγμα είναι αρκετά μεγάλος, είναι αρκετά εύκολο να κατασκευαστεί και να παρουσιαστεί λόγω της μπλοκ κυκλικής δομής του.

**Παράδειγμα 5** Ας υποθέσουμε ότι θέλουμε να κατασκευάσουμε ένα  $OD(272; 4^\ell)$ ,  $\ell = 1, \dots,$

136. Από το 5, κατασκευάζουμε το περιοδικό ζεύγος Golay,  $(T_1, T_2)$  μήκους  $t = 68$  όπως παρουσιάζεται στο Παράδειγμα 4. Ορίζουμε τους ξένους κυκλικούς πίνακες  $X$



**Πίνακας 3.3:** Ορισμένοι ορθογώνιοι σχεδιασμοί  $OD(n; q^m)$  κατασκευασμένοι από δύο κυκλικούς πίνακες. Όλοι οι σχεδιασμοί για  $q = 4, 8, 16$  ή  $q = n$  είναι επιπρόσθετα U-type. Οι σχεδιασμοί που σημειώνονται με αστερίσκο κατασκευάζονται χρησιμοποιώντας περιοδικά ζεύγη Golay που αναπτύσσονται όπως στο κεφάλαιο 3.5.1.

$q = 4$		$q = 5$		$q = 8$		$q = 9$		$q = 16$		$q = 17$	
$n$	$m$	$n$	$m$	$n$	$m$	$n$	$m$	$n$	$m$	$n$	$m$
4	2	5	2	8	4	9	4	16	8	17	8
8	4	9	4	16	8	17	8	32	16	33	16
16	8	17	8	32	16	33	16	64	32	65	32
32	16	33	16	64	32	65	32	128	64	129	64
40	20	41	20	80	40	81	40	160	80	161	80
64	32	65	32	128	64	129	64	256	128	257	128
80	40	81	40	160	80	161	80	320	160	321	160
104	52	105	52	208	104	209	104	416	208	417	208
128	64	129	64	256	128	257	128	512	256	513	256
136	68*	137	68*	272	136*	273	136*	544	272*	545	272*
160	80	161	80	320	160	321	160	640	320	641	320
200	100*	201	100*	400	200*	401	200*	800	400*	801	400*
208	104	209	104	416	208	417	208	832	416	833	416
232	116*	233	116*	464	232*	465	232*	928	464*	929	464*
256	128	257	128	512	256	513	256	1024	512	1025	512
272	136*	273	136*	544	272*	545	272*	1088	544*	1089	544*
296	148*	297	148*	592	296*	593	296*	1184	592*	1185	592*
320	160	321	160	640	320	641	320	1280	640	1281	640
328	164*	329	164*	656	328*	657	328*	1312	656*	1313	656*
400	200	401	200	800	400	801	400	1600	800	1601	800
416	208	417	208	832	416	833	416	1664	832	1665	832
464	232*	465	232*	928	464*	929	464*	1856	928*	1857	928*
488	244*	489	244*	976	488*	977	488*	1952	976*	1953	976*
512	256	513	256	1024	512	1025	512	2048	1024	2049	1024
544	272*	545	272*	1088	544*	1089	544*	2176	1088*	2177	1088*

Εδώ εξισώνοντας σημαίνει ότι θέτουμε το μεγαλύτερο επίπεδο ίσο με το προηγούμενο επίπεδο, με επανάληψη αν είναι απαραίτητο, και διαγράφοντας σημαίνει ότι ορίζουμε το μεγαλύτερο επίπεδο ίσο με μηδέν, με επανάληψη εάν είναι απαραίτητο.

### **Τέλος απόδειξης**

Η προτεινόμενη μέθοδος μπορεί επίσης να επεκταθεί και να χρησιμοποιηθεί για την κατασκευή σχεδόν ορθογώνιων σχεδιασμών, όπως στους Bingham et al. [11]. Σε αυτές τις περιπτώσεις, η συσχέτιση των προκύπτόντων σχεδιασμών υπολογίζεται εκ των προτέρων από την περιοδική συνάρτηση των χρησιμοποιημένων διανυσμάτων γραμμών και της δομής του σχεδιασμού.

**Θεώρημα 7** Έστω  $t = 2^{a_1} 10^{a_2} 26^{a_3} p^{a_4}$ , όπου  $a_1, a_2, a_3$  είναι μη αρνητικοί ακέραιοι αριθμοί,  $p \in \{34, 50, 58, 74, 82, 122, 202, 226\}$  και  $a_4 = 0$  ή 1. Τότε υπάρχουν οι ακόλουθοι σχεδόν ορθογώνιοι σχεδιασμοί για πειράματα υπολογιστών: (i)  $NOD(2(2^{bt} + 1); \{2(2^{b+1} + 1)\}^\ell)$ , (ii)  $NOD(2^{b+1}t + s + 2; (2^{b+1} + 3)^\ell)$ , (iii)  $NOD(2^2(2^{b-1}t + 1); \{2^2(2^{b-1} + 1)\}^\ell)$  και (iv)  $NOD(2^{b+1}t + s + 4; (2^{b+1} + 5)^\ell)$ ,  $\ell \leq 2^bt$ , για  $b = 1, 2, 3$  και για κάθε θετικό ακέραιο  $s$ .

### **Απόδειξη**

Για να αποδειχθεί η ύπαρξη των σχεδόν ορθογώνιων σχεδιασμών που περιγράφονται στα (i), (ii), (iii) και (iv), χρησιμοποιούμε την κατασκευή που δίνεται στα (4), (5), (6) και (7) αντίστοιχα. Από το Λήμμα 4 έχουμε ότι οι πίνακες  $\{A_i\}_{i=1}^8$  είναι κυκλικοί, προσθετικοί και φίλιοι. Για  $b = 1, 2, 3$  και ένα συγκεκριμένο  $1 \leq \ell \leq 2^bt$ , ο επιθυμητός σχεδόν ορθογώνιος σχεδιασμός (i), (ii), (iii) ή (iv) λαμβάνεται με τυχαία επιλογή οποιωνδήποτε  $\ell$  στηλών από τον πίνακα  $(C_{2^b}^T, 1^T, -1^T, -C_{2^b}^T)^T$ ,  $(C_{2^b}^T, 1^T, 0_{2^{bt} \times s}^T, -1^T, -C_{2^b}^T)^T$ ,  $(C_{2^b}^T, 3^T, 1^T, -1^T, -3^T, -C_{2^b}^T)^T$ , ή  $(C_{2^b}^T, 2^T, 1^T, 0_{2^{bt} \times s}^T, -1^T, -2^T, -C_{2^b}^T)^T$ , αντίστοιχα, όπου  $C_{2^b}$  είναι ο τετραγωνικός πίνακας τάξεως  $2^{bt}$  που λαμβάνεται χρησιμοποιώντας το  $\{A_1, \dots, A_{2^b}\}$ . **Τέλος απόδειξης**

Ένας σχεδιασμός ονομάζεται  $\ell$ -ορθογώνιος εάν το άθροισμα των γινομένων των στοιχείων οποιωνδήποτε  $\ell$  στηλών είναι μηδέν. Προφανώς, ένας  $\ell$ -ορθογώνιος σχεδιασμός είναι ορθογώνιος στο μέσο όταν  $\ell = 1$  και οι κύριες επιδράσεις του είναι

ορθογώνιες κατά ζεύγη όταν  $\ell = 2$ . Από την fold-over δομή των σχεδιασμών έχουμε το ακόλουθο Λήμμα.

**Λήμμα 3** Όλοι οι σχεδιασμοί που κατασκευάστηκαν από τα Θεωρήματα 8 και 9 είναι  $2\ell + 1$  ορθογώνιοι, όπου  $\ell$  οποιοσδήποτε μη-αρνητικός ακέραιος.

### 3.5.3 Κατασκευές χρησιμοποιώντας περιοδικά ζεύγη Golay

Ωστόσο, παρατηρήσαμε ότι η φίλια ιδιότητα (amicability property) στο Λήμμα 2, σελίδα 744 της αντίστοιχης ερευνητικής εργασίας, δεν ισχύει. Αυτό οδηγεί σε μια μικρή τροποποίηση ορισμένων αποτελεσμάτων που αναφέρονται στο Θεώρημα 2 σελίδα 744, της αντίστοιχης ερευνητικής εργασίας, στο Πόρισμα 2, σελίδα 745 και στο Θεώρημα 3, σελίδα 745. Δίνουμε εδώ τη διορθωμένη έκδοση αυτών των θεωρημάτων, συμπεριλαμβανομένων μερικών νέων αποτελεσμάτων που προκύπτουν από την επέκταση αυτών των θεωρημάτων πολλαπλασιασμού.

Τα παρακάτω είναι νέες κατευθυνόμενες ακολουθίες μήκους 72 και τύπου (72, 72) που μπορούν να χρησιμοποιηθούν για να επεκτείνουν το πολλαπλασιαστικό θεώρημα των Georgiou et al. [54]. Για λεπτομέρειες σχετικά με το συμβολισμό στις κατευθυνόμενες ακολουθίες και στη χρήση τους σας παραπέμπουμε στην εργασία των Georgiou et al. [54].

$$D_1 = [a, a, a, b, a, -a, -a, a, a, -b, -b, b, b, -b, b, -b, a, a, -a, a, -b, b, a, -a, a, b, -b, -a, a, -b, -b, -a, b, a, -b, -b, a, -a, b, -b, -b, -a, -b, -a, b, b, a, a, b, -a, -b, a, -b, b, -a, b, -a, -a, -b, -a, b, a, -a, b, -b, a, -b, -b, -b, -b, a, a]$$

$$D_2 = [b, b, a, a, a, a, b, a, -a, -b, b, -a, -b, a, -b, -b, -a, -b, -a, a, b, a, -b, -a, b, b, -a, -a, -b, a, -b, a, a, -a, -b, b, a, a, b, -a, -b, a, a, b, -b, a, -a, b, -b, b, -a, a, b, -b, b, b, a, -a, a, -a, -a, a, a, b, b, -b, -b, b, -a, b, b, b]$$

**Πόρισμα 5** Τα περιοδικά ζεύγη Golay μπορούν να κατασκευαστούν για όλα τα μήκη  $t = 2^{a_1} 10^{a_2} 26^{a_3} p^{a_4}$ , όπου τα  $a_1, a_2, a_3$  είναι μη αρνητικοί ακέραιοι,  $p \in \{34, 50, 58, 72, 74, 82, 122, 202, 226\}$  και  $a_4 = 0, 1$ .

Από τα περιοδικά ζεύγη Golay  $T_1, T_2$  μήκους  $t$  και τους αντίστοιχούς τους κυκλικούς πίνακες  $Z_1 = \text{circ}(T_1), Z_2 = \text{circ}(T_2)$  μπορούμε να ορίσουμε

$$X = (Z_1 + Z_2)/2, \quad W = (Z_1 - Z_2)/2. \quad (3.8)$$

Χρησιμοποιώντας τις νέες κατευθυνόμενες ακολουθίες που παρέχουμε παραπάνω μπορούμε να διορθώσουμε και να επεκτείνουμε τα αποτελέσματα των Georgiou et al. [54] ως εξής.

**Λήμμα 4** Οι πίνακες  $X$  και  $W$  στο (3.8) είναι κυκλικοί, ξένοι, και προσθετικοί (additive) πίνακες τάξης  $t$ , με εισόδους από το  $\{0, 1, -1\}$  και  $X \pm W$  είναι ένας  $\pm 1$  πίνακας.

**Θεώρημα 8** Έστω  $t = 2^{a_1} 10^{a_2} 26^{a_3} p^{a_4}$ , όπου  $a_1, a_2, a_3$  είναι μη-αρνητικοί ακέραιοι,  $p \in \{34, 50, 58, 72, 74, 82, 122, 202, 226\}$  και  $a_4 = 0$  ή  $1$ . Οι ακόλουθοι ορθογώνιοι κατά στήλη σχεδιασμοί για πειράματα υπολογιστών υπάρχουν:

(i)  $COD(2^{b+1}t; q_1^\ell)$  και

(ii)  $COD(2^{b+1}t + s; q_2^\ell)$ ,  $q_1 = 2^{b+1}$ ,  $q_2 = 2^{b+1} + 1$ ,  $\ell \leq 2^{bt}$ ,

για  $b = 1, 2$  και για κάθε θετικό ακέραιο  $s$ .

**Πόρισμα 6** Εάν  $t$  είναι ένας Golay ή ένας περιοδικός Golay αριθμός τότε οι ακόλουθοι ορθογώνιοι κατά στήλη σχεδιασμοί για πειράματα υπολογιστών υπάρχουν:  $COD(4t; 4^{2t})$ ,  $COD(4t + 1; 5^{2t})$ ,  $COD(8t; 8^{4t})$ , και  $COD(8t + 1; 9^{4t})$ .

**Πόρισμα 7** Έστω  $t = 2^{a_1} 10^{a_2} 26^{a_3} p^{a_4}$ , όπου  $a_1, a_2, a_3$  είναι μη-αρνητικοί ακέραιοι,  $p \in \{34, 50, 58, 72, 74, 82, 122, 202, 226\}$  και  $a_4 = 0$  ή  $1$ . Οι ακόλουθοι ορθογώνιοι κατά στήλη σχεδιασμοί για πειράματα υπολογιστών υπάρχουν:

(i)  $COD(2^{b+1}t; q_1^\ell)$  και

(ii)  $COD(2^{b+1}t + s; q_2^\ell)$ ,  $\ell \leq 2^{bt}$ ,

για  $b = 1, 2$ , για κάθε θετικό ακέραιο  $s$ ,  $q_1 = 2, \dots, 2^{b+1}$ , και  $q_2 = 3, \dots, 2^{b+1} + 1$ .

Η προτεινόμενη μέθοδος μπορεί επίσης να επεκταθεί και να χρησιμοποιηθεί για την κατασκευή σχεδόν ορθογώνιων σχεδιασμών, όπως στους Bingham et al. [11]. Σε τέτοιες περιπτώσεις, η συσχέτιση των παραγόμενων σχεδιασμών μπορεί να υπολογιστεί με την περιοδική συνάρτηση των χρησιμοποιούμενων διανυσμάτων γραμμών και της δομής του σχεδιασμού.



**Θεώρημα 9** Έστω  $t = 2^{a_1} 10^{a_2} 26^{a_3} p^{a_4}$ , όπου  $a_1, a_2, a_3$  είναι μη-αρνητικοί ακέραιοι,  $p \in \{34, 50, 58, 72, 74, 82, 122, 202, 226\}$  και  $a_4 = 0$  ή  $1$ . Οι ακόλουθοι σχεδόν ορθογώνιοι κατά στήλη σχεδιασμοί για πειράματα υπολογιστών υπάρχουν:

- (i)  $NCOD(2(2^b t + 1); [\{2(2^{b+1} + 1)\}^\ell])$ ,
- (ii)  $NCOD(2^{b+1} t + s + 2; (2^{b+1} + 3)^\ell)$ ,
- (iii)  $NCOD(2^2(2^{b-1} t + 1); [\{2^2(2^{b-1} + 1)\}^\ell])$  και
- (iv)  $NCOD(2^{b+1} t + s + 4; (2^{b+1} + 5)^\ell), \ell \leq 2^b t$ ,

για  $b = 1, 2$  για κάθε θετικό ακέραιο  $s$ .

### 3.6 Δομή ταυτόσημων επιδράσεων των σχεδιασμών που δημιουργήθηκαν

Όταν χρησιμοποιούμε σχεδιασμούς σε πειράματα υπολογιστών ή για σκοπούς κρησαρίσματος, ενδιαφερόμαστε για ιδιότητες που σχετίζονται με τις δομές των συσχετίσεων τους και την αποτελεσματικότητά τους κατά την προσαρμογή διαφορετικών μοντέλων. Σε αυτή την ενότητα ερευνούμε τέτοιες ιδιότητες και αξιολογούμε τους σχεδιασμούς που κατασκευάσαμε. Ακολουθώντας παρόμοιες τεχνικές όπως ο Georgiou [51], μπορούμε εύκολα να δείξουμε ότι οι κύριες επιδράσεις των ορθογώνιων κατά στήλη σχεδιασμών που κατασκευάζονται από τη μέθοδο fold-over που παρουσιάσαμε προηγουμένως είναι ορθογώνιοι σε οποιαδήποτε τετραγωνική επίδραση και σε οποιαδήποτε αλληλεπίδραση δύο παραγόντων.

Ο μέσος όρος και η μέγιστη τιμή για τις απόλυτες τιμές των στοιχείων εκτός διαγωνίου των πινάκων των ταυτόσημων επιδράσεων για τις τετραγωνικές επιδράσεις ( $E(|q|)$  και  $\max |q_{ij}|$ ) και τις επιδράσεις των αλληλεπιδράσεων ( $E(|t|)$  και  $\max |t_{ij}|$ ) έχουν οριστεί από τον Georgiou [51] για να εκτιμήσουν την απόδοση των LHD. Τα κάτω όρια για τα παραπάνω κριτήρια, όπως αποδείχθηκαν από τον Georgiou [51], δεν είναι κατάλληλα για τους ορθογώνιους κατά στήλη σχεδιασμούς που κατασκευάστηκαν σε αυτό το κεφάλαιο. Γενικεύουμε αυτά τα όρια έτσι ώστε να δουλεύουν για τα CODs και τα NCODs έτσι ώστε τα αποδεδειγμένα όρια για τα LHDs να γίνουν μια ειδική περίπτωση των νέων ορίων.

**Ορισμός 2** Ένας σχεδιασμός που ικανοποιεί οποιοδήποτε από τα κάτω όρια που δίνονται στο Λήμμα 6 λέγεται ότι είναι τετραγωνικά-βέλτιστος (quadratic-optimal) όσον αφορά σε αυτό το κριτήριο. Επιπλέον, κάθε σχεδιασμός με  $E(|t|) = 0$  ή  $\max|t_{ij}| = 0$  καλείται βέλτιστος-αλληλεπιδράσεων (interaction-optimal) όσον αφορά στο αντίστοιχο κριτήριο.

Είναι προφανές ότι, για ορθογώνιους σχεδιασμούς U-type που κατασκευάστηκαν σε αυτό το κεφάλαιο, το κάτω όριο για τα  $E(|t|)$  και τα  $\max|t_{ij}|$  είναι μηδέν, και έτσι αυτοί οι σχεδιασμοί είναι βέλτιστοι-αλληλεπιδράσεων για τα κριτήρια  $E(|t|)$  και  $\max|t_{ij}|$ .

Το Λήμμα 5 καθορίζει, σε ρητή μορφή, τη δομή των ταυτόσημων επιδράσεων του πίνακα πληροφορίας του U-type σχεδιασμού στο  $[-1, 1]^{n \times m}$  με  $n$  πειραματικές εκτελέσεις,  $m$  παράγοντες,  $q$  ομοιόμορφα κατανομημένα επίπεδα και  $s$  κεντρικά σημεία. Τα αποτελέσματα αυτού του λήμματος θα χρησιμοποιηθούν για τον προσδιορισμό των κάτω ορίων για τα κριτήρια  $E(|q|)$  και  $\max|q_{ij}|$  για τέτοιους σχεδιασμούς.

**Λήμμα 5** Έστω  $X$  ένας U-type ορθογώνιος κατά στήλη σχεδιασμός στο  $[-1, 1]^m$ , με  $n$  πειραματικές εκτελέσεις,  $m$  παράγοντες, και  $q$  επίπεδα με  $s$  κεντρικά (μηδενικά) σημεία. Έστω  $X_1 = [1_n \ X]$  ο πίνακας παλινδρόμησης για το μοντέλο πρώτης τάξης, συμπεριλαμβανομένης μίας στήλης με μονάδες και όλες τις στήλες του  $X$ . Τότε, ο πίνακας  $(X_1^T X_1)^{-1}$  είναι ένας διαγώνιος πίνακας τάξης  $m + 1$  και η διαγώνιός του είναι  $(n^{-1}, u^{-1}, u^{-1}, \dots, u^{-1})$ , όπου

$$u = \begin{cases} \frac{n}{3} \left( \frac{q+1}{q-1} \right), & \text{εάν } q \text{ ζυγός } (s = 0) \\ \frac{q(q+1)(n-s)}{3(q-1)^2}, & \text{εάν } q \text{ περιττός } (s \neq 0) \end{cases}.$$

### Απόδειξη

Το στοιχείο  $(i, j)$  του  $X_1^T X_1$  είναι ίσο με το εσωτερικό γινόμενο των στηλών  $i$  και  $j$  του  $X_1$ . Έτσι, το στοιχείο  $(1, 1)$  του  $X_1^T X_1$  θα ισούται με το  $n$ . Δεδομένου ότι ο σχεδιασμός έχει τις στήλες του κατά ζεύγος ορθογώνιες (ο μέσος είναι επίσης ορθογώνιος σε όλες τις άλλες στήλες επειδή ο σχεδιασμός είναι επίσης U-type αυτό σημαίνει ότι όλα τα μη διαγώνια στοιχεία του  $X_1^T X_1$ , δηλαδή τα στοιχεία στις θέσεις  $(i, j)$  για  $i \neq j$ , θα είναι μηδέν.

### 3.6 Δομή ταυτόσημων επιδράσεων των σχεδιασμών που δημιουργήθηκαν 53

$q = 2k$ : Κάθε στήλη του  $X$  έχει τα στοιχεία

$$\left\{ \frac{-2k+1}{2k-1}, \frac{-2k+3}{2k-1}, \dots, \frac{2k-3}{2k-1}, \frac{2k-1}{2k-1} \right\}$$

και μηδενικά κεντρικά σημεία ( $s = 0$ ). Κάθε ένα από αυτά τα στοιχεία εμφανίζεται  $n/q$  φορές σε κάθε στήλη. Για  $i > 1$ , το  $i_{th}$  διαγώνιο στοιχείο του  $X_1^T X_1$  είναι ίσο με

$$u = 2 \frac{n}{2k} \sum_{i=1}^k \left( \frac{2i-1}{2k-1} \right)^2 = \frac{n(2k+1)}{6k-3} = \frac{n}{3} \left( \frac{q+1}{q-1} \right).$$

$q = 2k + 1$ : Κάθε στήλη του  $X$  έχει τα στοιχεία

$$\left\{ \frac{-k}{k}, \frac{-k+1}{k}, \dots, \frac{k-1}{k}, \frac{k}{k} \right\}.$$

Κάθε ένα από τα μη μηδενικά στοιχεία εμφανίζεται  $(n-s)/(q-1)$  φορές σε κάθε στήλη. Για  $i > 1$ , το  $i_{th}$  διαγώνιο στοιχείο του  $X_1^T X_1$  είναι ίσο με

$$u = 2 \frac{(n-s)}{(q-1)} \sum_{i=1}^k \left( \frac{i}{k} \right)^2 = 8 \frac{(n-s)}{(q-1)^3} \sum_{i=1}^{(q-1)/2} i^2 = \frac{q(q+1)(n-s)}{3(q-1)^2}.$$

#### Τέλος απόδειξης

Τα παράγωγα γενικευμένα κάτω όρια για τα κριτήρια  $E(|q|)$  και  $\max |q_{ij}|$  δίνονται στο ακόλουθο Λήμμα.

**Λήμμα 6** Έστω  $X$  ένας  $U$ -type σχεδιασμός στο  $[-1, 1]^m$ , με  $n$  πειραματικές εκτελέσεις,  $m$  παράγοντες, και  $q$  ομοιόμορφα κατανομημένα επίπεδα. Τότε,

$$E(|q|) \geq \frac{u}{n(m+1)} = LB_{Eq} \text{ και } \max |q_{ij}| \geq \frac{u}{n} = LB_{maxq}, \quad (3.9)$$

όπου  $u$  είναι όπως δίνεται στο Λήμμα 5.

#### Απόδειξη

Για να αποδείξουμε τα όρια για οποιονδήποτε σχεδιασμό  $U$ -type, ακολουθούμε την απόδειξη που δίνεται στο Λήμμα 2 του Georgίου [51], αλλά χρησιμοποιούμε τα απο-

τελέσματα για τους U-type ορθογώνιους σχεδιασμούς που δίδονται στο Λήμμα 5. Για να εξοικονομήσουμε χώρο, η υπόλοιπη απόδειξη παραλείπεται. **Τέλος απόδειξης**

Σημειώνουμε ότι τα κάτω όρια που αποδείχτηκαν εδώ συμφωνούν με εκείνα του Georgίου [51] αν εφαρμοστούν σε LHDs.

**Λήμμα 7** Έστω  $X$  ένας σχεδόν ορθογώνιος κατά στήλη σχεδιασμός στο  $[-1, 1]^m$ , με  $n$  πειραματικές εκτελέσεις,  $m$  παράγοντες,  $q$  επίπεδα, και  $s$  κεντρικά σημεία, που κατασκευάστηκε χρησιμοποιώντας το Θεώρημα 3. Υποθέτουμε ότι η συνθήκη (i) του Θεωρήματος 3 ικανοποιείται για όλα τα επίπεδα εκτός από  $\pm 1, 0$  για  $L_1$  και για όλα τα επίπεδα εκτός από  $\pm 2, \pm 1, 0$  για  $L_2$ . Επιπλέον, θέτουμε τα επιπρόσθετα επίπεδα ( $\pm 1$  για  $L_1$  και  $\pm 2, \pm 1$  για  $L_2$ ) να εμφανιστούν μόνο μία φορά. Ο πίνακας  $X_1 = [1_n \ X]$  είναι ο πίνακας παλινδρόμησης για το μοντέλο πρώτης τάξης, που περιλαμβάνει μία στήλη με μονάδες και όλες τις στήλες του  $X$ . Τότε, ο πίνακας πληροφορίας ( $X_1^T X_1$ ) είναι ένας τετραγωνικός πίνακας τάξης  $m + 1$  της μορφής

$$X_1^T X_1 = \begin{pmatrix} n & 0 & 0 & \dots & 0 \\ 0 & u & b & \dots & b \\ 0 & b & u & \dots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & b & b & \dots & u \end{pmatrix} = \begin{pmatrix} n & \mathbf{0} \\ \mathbf{0} & (u - b)I + bJ \end{pmatrix},$$

όπου

$$(u, b) = \begin{cases} \left( \frac{[q(q+3)+8](n-s-2)}{3(q-1)^2} + \frac{8}{(q-1)^2}, \frac{8}{(q-1)^2} \right), & \text{για σχεδιασμούς από το } L_1. \\ \left( \frac{[q(q+5)+24](n-s-4)}{3(q-1)^2} + \frac{40}{(q-1)^2}, \frac{40}{(q-1)^2} \right), & \text{για σχεδιασμούς από το } L_2. \end{cases}$$

στο Θεώρημα 3.

### Απόδειξη

Το στοιχείο  $(i, j)$  του  $X_1^T X_1$  είναι ίσο με το εσωτερικό γινόμενο των στηλών  $i$  και  $j$  του  $X_1$ . Η απόδειξη για τα διαγώνια στοιχεία  $(u)$ , η πρώτη σειρά και η πρώτη στήλη είναι παρόμοιες με την απόδειξη του Λήμματος 5. Κάθε στοιχείο εκτός διαγώνιου, του  $X_1^T X_1$ , λαμβάνεται υπολογίζοντας το εσωτερικό γινόμενο των δύο αντίστοιχων διακριτών στηλών του σχεδόν ορθογώνιου σχεδιασμού.

### 3.6 Δομή ταυτόσημων επιδράσεων των σχεδιασμών που δημιουργήθηκαν 55

Περίπτωση 1: Σχεδιασμοί που κατασκευάστηκαν από το  $L_1$  στο Θεώρημα 3: Κάθε ένα από αυτά τα στοιχεία, εκτός από το 0, και το  $\pm 1$ , εμφανίζεται  $(ns - 2)/(q - 3)$  φορές σε κάθε στήλη. Έχουμε ότι  $q = 2k + 1$ , και έτσι κάθε στήλη του  $X$  έχει τα στοιχεία

$$\left\{ \frac{-k}{k}, \frac{-k+1}{k}, \dots, \frac{k-1}{k}, \frac{k}{k} \right\}.$$

Για  $i > 1$ , το  $i_{th}$  διαγώνιο στοιχείο του  $X_1^T X_1$ , είναι ίσο με

$$u = 2 \frac{n-s-2}{2k-2} \sum_{j=2}^k \left( \frac{j}{k} \right)^2 + \frac{2}{k^2} = \frac{[q(q+3)+8](n-s-2)}{3(q-1)^2} + \frac{8}{(q-1)^2},$$

$$b = \frac{2(1^2)}{k^2} = \frac{2}{k^2} = \frac{8}{(q-1)^2}.$$

Περίπτωση 2: Σχεδιασμοί που κατασκευάστηκαν από το  $L_2$  στο Θεώρημα 3: Κάθε ένα από αυτά τα στοιχεία, εκτός από το 0, το  $\pm 1$ , και το  $\pm 2$ , εμφανίζεται  $(n-s-4)/(q-5)$  φορές σε κάθε στήλη. Για  $i > 1$ , το  $i_{th}$  διαγώνιο στοιχείο του  $X_1^T X_1$ , είναι ίσο με

$$u = 2 \frac{n-s-4}{2k-4} \sum_{j=3}^k \left( \frac{j}{k} \right)^2 + \frac{10}{k^2} = \frac{[q(q+5)+24](n-s-4)}{3(q-1)^2} + \frac{40}{(q-1)^2},$$

$$b = \frac{2(1^2+2^2)}{k^2} = \frac{10}{k^2} = \frac{40}{(q-1)^2}.$$

#### Τέλος απόδειξης

Οι σχεδόν ορθογώνιοι σχεδιασμοί που κατασκευάστηκαν σε αυτό το κεφάλαιο είναι τετραγωνικά-βέλτιστοι αλλά όχι βέλτιστοι-αλληλεπιδράσεων όσον αφορά τα προτεινόμενα κριτήρια. Από την άλλη πλευρά, οι σχεδόν ορθογώνιοι σχεδιασμοί αυτού του κεφαλαίου έχουν πολύ καλές (μικρές) τιμές αυτών των κριτηρίων, όπως υποδεικνύεται από το Θεώρημα 10.

**Θεώρημα 10** Έστω  $X$ ,  $b$ , και  $u$  όπως ορίστηκαν στο Λήμμα 7. Τα κριτήρια αξιολό-

γησης για αυτό το σχεδιασμό είναι

$$\begin{aligned} \text{ave}(|t|) = E(|t|) &= \frac{2 \sum_{i=1}^{m+1} \sum_{j=1}^{m(m-1)/2} |t_{ij}|}{m(m^2-1)} = \frac{b}{n(m+1)}, \quad \max t = \max_{i,j} |t_{ij}| = b/n. \\ \text{ave}(|q|) = E(|q|) &= \frac{\sum_{i=1}^{m+1} \sum_{j=1}^m |q_{ij}|}{m(m+1)} = \frac{u}{n(m+1)}, \quad \max q = \max_{i,j} |q_{ij}| = u/n. \end{aligned}$$

### Απόδειξη

Τα αποτελέσματα ακολουθούν με απλούς υπολογισμούς χρησιμοποιώντας το Λήμμα

#### 7. Τέλος απόδειξης

Τα αποτελέσματα που παρουσιάζονται σε αυτή την ενότητα δείχνουν ότι ακόμα και όταν οι εκτιμήσεις των τετραγωνικών επιδράσεων και των αλληλεπιδράσεων δύο παραγόντων είναι ασυσχέτιστες με τις εκτιμήσεις των γραμμικών επιδράσεων, αυτό δεν εμποδίζει τις τετραγωνικές επιδράσεις και τις αλληλεπιδράσεις δύο παραγόντων από το να είναι συσχετισμένες μεταξύ τους. Επιπλέον, όλοι οι ορθογώνιοι U-type σχεδιασμοί που κατασκευάστηκαν με την προτεινόμενη μέθοδο fold-over είναι τετραγωνικοί και αλληλεπιδράσεων-βέλτιστοι μέσα στην κλάση των U-type σχεδιασμών. Τα παραπάνω αποτελέσματα παρέχουν κάποια εμπιστοσύνη και περιορισμένη ευρωστία σε πιθανές μη-γραμμικότητες στο πρώτο στάδιο (κρησαρίσματος) του πειράματος. Ως εκ τούτου, μπορούμε να χρησιμοποιήσουμε τους σχεδιασμούς που κατασκευάστηκαν για το κρησαρίσμα, ακόμη και όταν υποπτευόμαστε ότι ενδέχεται να είναι ενεργές κάποιες αλληλεπιδράσεις δύο παραγόντων ή/και τετραγωνικές επιδράσεις, αλλά πρέπει να είμαστε προσεκτικοί επειδή υπάρχει αβεβαιότητα.

## 3.7 Συζήτηση

Η κύρια συμβολή του παρόντος κεφαλαίου είναι η κατασκευή ορθογώνιων κατά στήλη σχεδιασμών (COD) και σχεδόν ορθογώνιων σχεδιασμών (NCOD) συνδυάζοντας γενικευμένους ορθογώνιους σχεδιασμούς, ξένους κυκλικούς πίνακες και περιοδικά ζεύγη Golay. Παρέχεται ένας νέος πολλαπλασιαστής με την παρουσίαση νέων κατευθυνόμενων (directed) ακολουθιών μήκους 72. Αυτή η προσέγγιση οδηγεί σε μια νέα πολλαπλασιαστική τεχνική και μια νέα άπειρη οικογένεια COD που περιλαμβάνει βέλτιστους σχεδιασμούς πολλών νέων τάξεων. Σημειώνουμε ότι οι βέλ-

τιστοι ορθογώνιοι κατά στήλη σχεδιασμοί σε αυτήν την άπειρη οικογένεια είναι νέων τάξεων και δεν μπορούσαν να κατασκευαστούν με οποιαδήποτε γνωστή μέθοδο στη βιβλιογραφία.

Μια σημαντική ιδιότητα των σχεδιασμών που παρουσιάστηκαν είναι ότι η σύνθεσή τους υπολογίζεται και δίνεται σε κλειστή μορφή, παρέχοντας έναν εκ των προτέρων υπολογισμό όλων των σχεδιασμών εντός αυτής της κατηγορίας. Τα νέα κατώτατα όρια που παρέχονται επαληθεύουν ότι οι σχεδιασμοί που κατασκευάστηκαν είναι βέλτιστοι, σε σχέση με τη δομή των ταυτόσημων επιδράσεων, στην κατηγορία των περιγραφέντων ορθογώνιων κατά στήλη σχεδιασμών. Η γνώση της δομής των ταυτόσημων επιδράσεων πρώτης και δεύτερης τάξης μπορεί να αποδειχθεί πολύ χρήσιμη όταν γίνεται κρησάρισμα με την παρουσία όρων δεύτερης τάξης. Αυτές οι ιδιότητες είναι πολύ χρήσιμες για την ανάλυση αυτών των σχεδιασμών κάτι που τους καθιστά ιδιαίτερα ελκυστικούς στην κατηγορία των *CODs*.

Οι προτεινόμενοι βέλτιστοι σχεδιασμοί έχουν πολλά πλεονεκτήματα. Είναι πολύ εύκολο να κατασκευαστούν, έχουν ευέλικτα μεγέθη και μπορούν να φιλοξενούν πολλά επίπεδα. Αυτά τα πλεονεκτήματα καθιστούν τους σχεδιασμούς αυτού του κεφαλαίου μία καλή εναλλακτική για πειράματα υπολογιστών.

Παρόλο που έχει γίνει πολύ καλή δουλειά σε αυτόν τον τομέα και έχουν κατασκευαστεί πολλές νέες τάξεις κατάλληλων ορθογώνιων κατά στήλη σχεδιασμών, εξακολουθούν να υπάρχουν ανοικτά προβλήματα που πρέπει να αντιμετωπιστούν και τάξεις για τις οποίες οι βέλτιστοι ή ορθογώνιοι σχεδιασμοί είναι ακόμη άγνωστοι. Για αυτές τις τάξεις απαιτείται περαιτέρω μελέτη και έρευνα για να ληφθούν βέλτιστοι ή ορθογώνιοι σχεδιασμοί. Σε τέτοιες περιπτώσεις και σε περίπτωση που αποδειχθεί ότι δεν υπάρχουν βέλτιστοι ή ορθογώνιοι σχεδιασμοί, προτείνουμε να χρησιμοποιηθούν εναλλακτικές λύσεις, όπως οι σχεδόν ορθογώνιοι κατά στήλη σχεδιασμοί.





## **Μέρος ΙΙ**

# **Κατασκευή και Ανάλυση Υπερκορεσμένων Σχεδιασμών**



---

## Εισαγωγικά στοιχεία και βασικές έννοιες για τους υπερκορεσμένους σχεδιασμούς

---

To find out what happens when you change something,  
it is necessary to change it.

—Box, Hunter, and Hunter, *Statistics for Experimenters* (1978)

Στο παρόν κεφάλαιο παρουσιάζονται κάποιες βασικές έννοιες για τους παραγοντικούς σχεδιασμούς. Ιδιαίτερη βαρύτητα δίνεται σε μία ειδική κλάση κλασματικών παραγοντικών σχεδιασμών, τους υπερκορεσμένους σχεδιασμούς για τους οποίους παρουσιάζονται βασικές έννοιες σχετικά με τα κριτήρια βελτιστότητας, την κατασκευή και την ανάλυσή τους.

### 4.1 Εισαγωγή

Οι παραγοντικοί σχεδιασμοί χρησιμοποιούνται ευρύτατα σε πειράματα που περιλαμβάνουν αρκετούς παράγοντες, και είναι αναγκαία η μελέτη της κοινής επίδρασής τους στην μεταβλητή απόκριση. Η αλλαγή που γίνεται στην μεταβλητή απόκρισης από την αλλαγή του επιπέδου του παράγοντα ονομάζεται επίδραση (effect) του παράγοντα. Αντιλαμβανόμαστε λοιπόν ότι ακόμη και για έναν μέτριο αριθμό παραγόντων, ο συνολικός αριθμός των παρατηρήσεων είναι αρκετά μεγάλος. Άλλοτε για να περιοριστεί το μέγεθος των πειραματικών εκτελέσεων και άλλοτε διότι δεν υπάρχουν τα απαραίτητα μέσα, οι πειραματιστές είναι αναγκασμένοι να εκτελέσουν μία μόνο επανάληψη τέτοιων σχεδιασμών (μη-επαναλαμβανόμενοι  $2^m$  παραγοντικοί σχεδια-

μοί). Σε άλλες περιπτώσεις υπάρχει η αναγκαιότητα να εξετάσουμε έναν μεγάλο αριθμό παραγόντων αλλά στη διάθεσή μας έχουμε έναν περιορισμένο αριθμό πειραματικών εκτελέσεων. Στην περίπτωση αυτή είναι απαραίτητη η χρήση κλασματικών παραγοντικών σχεδιασμών. Οι σχεδιασμοί αυτοί είναι μεταξύ των ευρύτερα χρησιμοποιούμενων τύπων σχεδιασμών σε πολλά πρακτικά προβλήματα. Για τη σωστή και ως εκ τούτου επιτυχή χρήση των κλασματικών παραγοντικών σχεδιασμών, ο πειραματιστής πρέπει να βασίζεται στις ακόλουθες τρεις θεμελιώδεις αρχές:

- 1 Αρχή σποραδικότητας των επιδράσεων (sparsity of effects principle). Στην περίπτωση ύπαρξης πολλών μεταβλητών-παραγόντων, η διαδικασία είναι πιθανό να οδηγείται αρχικά από μερικές από τις κύριες επιδράσεις και αλληλεπιδράσεις χαμηλής τάξης.
- 2 Προβολική ιδιότητα (projective property). Οι σχεδιασμοί αυτοί μπορούν να προβάλλονται σε μεγαλύτερους σχεδιασμούς με αντικείμενο τους σημαντικούς παράγοντες.
- 3 Ακολουθιακός πειραματισμός (sequential experimentation). Είναι δυνατόν να συνδυάσουμε τις εκτελέσεις δύο (ή περισσότερων) κλασματικών παραγοντικών σχεδιασμών έτσι ώστε να συγκεντρωθεί ακολουθιακά ένας μεγαλύτερος σχεδιασμός με σκοπό την εκτίμηση των επιδράσεων και αλληλεπιδράσεων των παραγόντων που ενδιαφέρουν.

Παράλληλα πρέπει να αναφερθεί ότι οι ιεραρχικές υποθέσεις που ισχύουν για τους παραγοντικούς σχεδιασμούς είναι

- i Οι επιδράσεις μικρότερης τάξης είναι σημαντικότερες από επιδράσεις μεγαλύτερης τάξης.
- ii Οι επιδράσεις ίδιας τάξης είναι το ίδιο σημαντικές.

Έστω  $D$  ένας  $N \times m$  ο πίνακας ενός παραγοντικού σχεδιασμού με στοιχεία από το σύνολο των παραγοντικών επιπέδων  $1, 2, \dots, s$ . Ο πίνακας  $D$  έχει την ιδιότητα των ίσων εμφανίσεων αν κάθε ένα από τα παραγοντικά επίπεδα εμφανίζεται τον ίδιο αριθμό φορές σε κάθε στήλη του  $D$  και αν ισχύει η ιδιότητα αυτή τότε ονομάζεται ισορροπημένος. Τότε το  $N$  είναι πολλαπλάσιο του μέγιστου αριθμού  $s$ . Ένας

ισορροπημένος σχεδιασμός ονομάζεται ορθογώνιος ισχύος 2 αν όλοι οι  $s^2$  συνδυασμοί επιπέδων εμφανίζονται τον ίδιο αριθμό φορών σε οποιοσδήποτε δύο στήλες του σχεδιασμού.

Η κυριότερη χρήση των κλασματικών παραγοντικών σχεδιασμών είναι σε πειράματα κρησαρίσματος (screening experiments) δηλαδή σε πειράματα στα οποία θεωρούμε ότι μετέχουν πολλοί παράγοντες με σκοπό την ανίχνευση των σημαντικών παραγόντων. Τα πειράματα κρησαρίσματος συνήθως εκτελούνται στα αρχικά στάδια μιας μελέτης, όπου είναι πιθανό πολλοί από τους παράγοντες που θεωρήθηκαν αρχικά να μην έχουν σημαντική επίδραση στη μεταβλητή απόκρισης και πρωταρχικός στόχος είναι να προσδιοριστούν οι λίγοι, αλλά κυρίαρχοι, ενεργοί παράγοντες, διατηρώντας το κόστος όσο το δυνατόν χαμηλότερο. Οι παράγοντες που αναγνωρίζονται ως σημαντικοί, ερευνώνται πιο λεπτομερειακά στα επόμενα πειράματα. Ιδιαίτερα αποδοτικοί σε τέτοιου είδους πειράματα είναι μία ειδική κλάση κλασματικών παραγοντικών σχεδιασμών, οι υπερκορεσμένοι σχεδιασμοί.

## **4.2 Κατασκευή και κριτήρια αξιολόγησης υπερκορεσμένων σχεδιασμών**

Οι υπερκορεσμένοι σχεδιασμοί (ΥΣ) είναι σχεδιασμοί των οποίων οι παράγοντες υπερβαίνουν το μέγεθος των πειραματικών εκτελέσεων και είναι ιδιαίτερα χρήσιμοι όταν θέλουμε να εξετάσουμε έναν μεγάλο αριθμό παραγόντων αλλά έχουμε στη διάθεσή μας έναν περιορισμένο αριθμό πειραματικών εκτελέσεων. Οι δύο μεγάλες κλάσεις που μπορούν να διαχωριστούν οι ΥΣ είναι οι σχεδιασμοί δύο επιπέδων, όπου όλοι οι παράγοντες βρίσκονται σε δύο και μόνο επίπεδα (two-level), και οι σχεδιασμοί μικτών επιπέδων (mixed-level) που περιλαμβάνουν παράγοντες με διαφορετικό αριθμό επιπέδων στον ίδιο πίνακα σχεδιασμού. Μία ειδική περίπτωση των σχεδιασμών μικτών επιπέδων είναι οι σχεδιασμοί πολλαπλών επιπέδων (multi-level) που περιλαμβάνουν παράγοντες που αποτελούνται όλοι από τον ίδιο αριθμό επιπέδων  $s > 2$ .

**Ορισμός 1** *Υπερκορεσμένοι σχεδιασμοί δύο επιπέδων ονομάζονται οι κλασματικοί παραγοντικοί σχεδιασμοί των οποίων ο αριθμός των παραγόντων  $m$  είναι μεγαλύτερος ή ίσος με τον αριθμό των πειραματικών εκτελέσεων, δηλαδή  $m \geq N$  ή  $m > N - 1$ .*

Στη γενικότερη περίπτωση, όπου έχουμε περισσότερα από 2 επίπεδα, υπερκορεσμένοι ονομάζονται οι σχεδιασμοί για τους οποίους ισχύει  $m(s - 1) > N - 1$ .

Από τα παραπάνω αντιλαμβανόμαστε ότι ένας ΥΣ δεν μπορεί να είναι ορθογώνιος, οπότε χρειάζεται να μετρήσουμε πόσο απέχει από την ορθογωνιότητα. Βάσει αυτής της απόστασης ορίζονται κάποια κριτήρια βελτιστότητας ΥΣ που συζητούνται στο κεφάλαιο 4.2.1. Το γεγονός ότι ένας ΥΣ δεν μπορεί να είναι ορθογώνιος έχει ως αποτέλεσμα να δημιουργούνται συσχετίσεις μεταξύ των στηλών του σχεδιασμού. Η δομή των συσχετίσεων ενός ΥΣ επηρεάζει σημαντικά τόσο την κατασκευή όσο και την ανάλυσή τους.

Ο Satterthwaite [119] ήταν ο πρώτος που εισήγαγε την ιδέα των ΥΣ ως ενός τυχαίου ισορροπημένου σχεδιασμού. Οι Booth και Cox [13] ήταν οι πρώτοι που εξέτασαν συστηματικά τους ΥΣ δύο επιπέδων και πρότειναν κάποια από τα βασικότερα κριτήρια βελτιστότητας, το  $ave(s^2)$  και το  $s_{max}$ . Το ενδιαφέρον για τους ΥΣ άρχισε να εμφανίζεται τριάντα χρόνια αργότερα από τον Lin [91] και από τότε η ανάπτυξη τους ήταν ραγδαία αφού αρκετοί ερευνητές ασχολήθηκαν με την κατασκευή και την εξέταση των ιδιοτήτων των ΥΣ, όπως στις ακόλουθες εργασίες [; 2, 18, 20, 24, 45, 47, 68, 70, 76, 83, 87–90, 92, 95, 97, 107, 108, 132, 133, 139, 142, 144–146, 149–151]. Υπάρχουν πολλές ερευνητικές εργασίες που αφορούν στην κατασκευή ΥΣ, όχι μόνο για την περίπτωση δύο επιπέδων, αλλά και για σχεδιασμούς πολλαπλών και μικτών επιπέδων. Οι Fang et al. [47] πρότειναν το κριτήριο  $E(f_{NOD})$  για τη σύγκριση ΥΣ πολλαπλών και μικτών επιπέδων και παρουσίασαν μια μέθοδο για την κατασκευή  $E(f_{NOD})$ -βέλτιστων ΥΣ. Οι Liu και Dean [95] πρότειναν μια μέθοδο για την κατασκευή δύο-επιπέδων ΥΣ, χρησιμοποιώντας  $k$ -κυκλικούς γεννήτορες. Οι Georgiou και Koukouninos [55] γενίκευσαν την  $k$ -κυκλική μέθοδο των [95] για την κατασκευή ΥΣ πολλαπλών επιπέδων. Οι Chen και Liu [26] γενίκευσαν περαιτέρω την  $k$ -κυκλική μέθοδο των [55] για ΥΣ πολλαπλών επιπέδων και πρότειναν μία μέθοδο κατασκευής βέλτιστων μικτών επιπέδων  $k$ -κυκλικών ΥΣ. Πιο πρόσφατα, οι Liu και Liu [89] πρότειναν τρεις μεθόδους, μία εκ των οποίων αφορά ισαπέχοντες σχεδιασμούς και οι άλλες δύο σχεδόν ισαπέχοντες σχεδιασμούς, για την κατασκευή  $E(f_{NOD})$ -βέλτιστων ΥΣ. Περαιτέρω λεπτομέρειες και μία εκτεταμένη βιβλιογραφική ανασκόπηση σχετικά με την κατασκευή και την ανάλυση ΥΣ γίνεται στην εργασία του Georgiou [52]. Οι Sun et al. [129] παρουσίασαν επίσης μια πλήρη ανασκόπηση των υφιστάμενων μεθόδων

ΥΣ μικτών επιπέδων. Μία ανασκόπηση της βιβλιογραφίας αποκαλύπτει ότι οι περισσότερες από τις υπάρχουσες μελέτες έχουν επικεντρωθεί στους ισορροπημένους ΥΣ όπου όλα τα επίπεδα εμφανίζονται τον ίδιο αριθμό φορών σε κάθε παράγοντα. Στην παρούσα διδακτορική διατριβή παρουσιάζουμε μία μέθοδο κατασκευής υπερκορεσμένων σχεδιασμών πολλαπλών και μικτών-επιπέδων τόσο για ισορροπημένους όσο και για μη ισορροπημένους σχεδιασμούς (κεφάλαιο 5).

### **4.2.1 Κριτήρια αξιολόγησης και βελτιστοποίησης υπερκορεσμένων σχεδιασμών**

#### **Υπερκορεσμένοι σχεδιασμοί δύο επιπέδων**

Για τη σύγκριση και την αξιολόγηση των υπερκορεσμένων σχεδιασμών δύο επιπέδων, αναπτύχθηκαν κατά καιρούς ποικίλα κριτήρια. Τα κριτήρια αναδεικνύουν έναν τρόπο μέτρησης της απόστασης του σχεδιασμού από τον ορθογώνιο, την απόσταση της πρόβλεψης της διασποράς, ή την προβολικότητα του σχεδιασμού σε διάσταση χαμηλότερης τάξης. Για να καθοριστεί η απόσταση από τον ορθογώνιο σχεδιασμό, τα κριτήρια βελτιστότητας αξιολογούν τα μη διαγώνια στοιχεία του πίνακα  $X^T X$ , όπου  $X$  είναι ο πίνακας του σχεδιασμού. Τα κριτήρια που αξιολογούν την εκτίμηση της διασποράς βασίζονται στην ορίζουσα ή το ίχνος του πίνακα  $X^T X$  ενώ τα κριτήρια που βασίζονται στην προβολικότητα του σχεδιασμού εστιάζουν την προσοχή τους στη συσχέτιση μεταξύ των στηλών. Έστω ότι  $s_{ij}$  είναι το στοιχείο της  $i$ -οστής γραμμής και της  $j$ -οστής στήλης του πίνακα  $X^T X$ . Οι Booth και Cox [13] πρότειναν σαν κριτήριο σύγκρισης και κατασκευής σχεδιασμών την ελαχιστοποίηση της μέσης τιμής του  $s_{ij}^2$  που συμβολίζεται με  $ave(s^2)$  ή  $E(s^2)$ , όπου

$$E(s^2) = \sum_{1 \leq i < j \leq m} s_{ij}^2 / [m(m-1)/2] \quad (4.1)$$

Όταν θέλουμε να δούμε την αποδοτικότητα ενός υπερκορεσμένου σχεδιασμού, βρίσκουμε την τιμή του  $E(s^2)$  και όσο πιο κοντά είναι στο μηδέν τόσο καλύτερο σχεδιασμό έχουμε. Αν το  $s_{ij}^2 = 0$ , οι παράγοντες  $i$  και  $j$  είναι ορθογώνιοι. Όταν το  $s_{ij}^2 = \pm N$ , τότε  $c_i = \pm c_j$  και οι παράγοντες  $i$  και  $j$  είναι πλήρως εξαρτημένοι. Οι

Nguyen [107] και Tang και Wu [133], απέδειξαν ότι το

$$E(s^2) \geq \frac{N^2(m - N + 1)}{(N - 1)(m - 1)}. \quad (4.2)$$

Ένας υπερκορεσμένος σχεδιασμός ονομάζεται  $E(s^2)$ -βέλτιστος όταν η  $E(s^2)$  τιμή του πίνακα σχεδιασμού στην 4.2 πετυχαίνει το κατώτερο φράγμα.

### Υπερκορεσμένοι σχεδιασμοί πολλών και μικτών επιπέδων

Υπάρχουν διάφορα κριτήρια βελτιστότητας για την περίπτωση υπερκορεσμένων σχεδιασμών πολλών και μικτών επιπέδων. Έστω  $c_1, \dots, c_m$  είναι οι στήλες ενός  $SSD(N, s_1 s_2 \dots s_m)$  και  $n_{ab}^{ij}$  είναι ο αριθμός των γραμμών με συνδυασμούς επιπέδων  $a, b$  στον  $N \times 2$  πίνακα  $[c_i, c_j]$ . Οι Yamada και Lin [148], Yamada, Ikebe, Hashiguchi και Niki [147] και Yamada και Matsui [150] όρισαν το

$$\chi^2(c_i, c_j) = \sum_{a=0}^{s_i-1} \sum_{b=0}^{s_j-1} [n_{ab}^{ij} - N/(s_i s_j)]^2 / (N/(s_i s_j))$$

για την αξιολόγηση της εξάρτησης των στηλών  $c_i$  και  $c_j$ . Δύο στήλες είναι απόλυτα εξαρτημένες όταν το  $\chi^2(c_i, c_j) = N(s_i - 1)(s_j - 1)$  και ανεξάρτητες όταν  $\chi^2(c_i, c_j) = 0$ . Επιπρόσθετα πρότειναν τα ακόλουθα δύο κριτήρια για το βαθμό εξάρτησης ολόκληρου του σχεδιασμού ως:

$$ave(\chi^2) = \sum_{1 \leq i < j \leq N} \chi^2(c_i, c_j) / [m(m - 1)/2],$$

$$max(\chi^2) = \max_{1 \leq i < j \leq N} \chi^2(c_i, c_j).$$

Ένα άλλο μέτρο για τον βαθμό εξάρτησης μεταξύ δύο στηλών  $c_i$  και  $c_j$  του πίνακα σχεδιασμού ορίστηκε από τους Fang, Lin και Ma [48] ως

$$f(c_i, c_j) = \sum_{a=0}^{s_i-1} \sum_{b=0}^{s_j-1} \left| [n_{ab}^{ij} - N/(s_i s_j)] \right|$$



Πιο συγκεκριμένα πρότειναν τα εξής κριτήρια που αφορούν στον βαθμό εξάρτησης όλου του σχεδιασμού και μπορούν να υπολογιστούν από τις σχέσεις:

$$ave(f) = \sum_{1 \leq i < j \leq m} f(c_i, c_j) / [m(m-1)/2],$$

$$max(f) = \max_{1 \leq i < j \leq m} f(c_i, c_j).$$

Όταν όλα τα  $s_i = s$  οι Lu και Sun [99] και Lu, Hu και Zheng [98] όρισαν το

$$d_{ij}^2 = \sum_{a=0}^{s-1} \sum_{b=0}^{s-1} [n_{ab}^{ij} - N/(s^2)]^2$$

για την μέτρηση της παρέκκλισης από την ορθογωνιότητα για τις στήλες  $c_i$  και  $c_j$ . Αυτοί πρότειναν την ελαχιστοποίηση των ακόλουθων μέτρων

$$E(d^2) = \sum_{1 \leq i < j \leq N} d_{ij}^2(c_i, c_j) / [m(m-1)/2]$$

$$max(d^2) = \max_{1 \leq i < j \leq N} d_{ij}^2.$$

Είναι προφανές ότι  $E(d^2) = (N/s^2)ave(\chi^2)$  και  $max(d^2) = (N/s^2)max(\chi^2)$ . Οι Fang *et al.* [47] πρότειναν το  $E(f_{\text{NOD}})$  κριτήριο για την αξιολόγηση και τη σύγκριση των ΥΣ με μικτά επίπεδα υπό το πρίσμα της ορθογωνιότητας και της ομοιομορφίας. Το κριτήριο αυτό ορίζεται ως

$$E(f_{\text{NOD}}) = \sum_{1 \leq j_1 \neq j_2 \leq m} f_{\text{NOD}}^{j_1 j_2} / m(m-1), \quad (4.3)$$

όπου

$$f_{\text{NOD}}^{j_1 j_2} = \sum_{\alpha_{j_1}=1}^{s_{j_1}} \sum_{\alpha_{j_2}=1}^{s_{j_2}} \left( n_{\alpha_{j_1} \alpha_{j_2}}^{j_1 j_2} - \frac{N}{s_{j_1} s_{j_2}} \right)^2 = \sum_{\alpha_{j_1}=1}^{s_{j_1}} \sum_{\alpha_{j_2}=1}^{s_{j_2}} \left( n_{\alpha_{j_1} \alpha_{j_2}}^{j_1 j_2} \right)^2 - \frac{N^2}{s_{j_1} s_{j_2}}$$

Ακόμη και αν αυτό το κριτήριο ορίζεται κυρίως για ισορροπημένους ΥΣ, μπορεί να εφαρμοστεί στην περίπτωση των σχεδόν ισορροπημένων σχεδιασμών ή ακόμη και

στην περίπτωση των μη ισορροπημένων σχεδιασμών. Περισσότερες πληροφορίες σχετικά με το κριτήριο  $E(f_{\text{NOD}})$  και την εφαρμογή του στην περίπτωση των σχεδόν και των μη ισορροπημένων σχεδιασμών θα βρείτε στο κεφάλαιο 5.

### 4.3 Ανάλυση υπερκορεσμένων σχεδιασμών

Η ανάλυση των ΥΣ βασίζεται στην αρχή σποραδικότητας των επιδράσεων [17] η οποία υποθέτει ότι μόνο λίγοι από τους παράγοντες επηρεάζουν τη μεταβλητή απόκριση. Παρόλο που έχουν το τεράστιο πλεονέκτημα να μειώνουν δραστικά το πειραματικό κόστος, ένα κρίσιμο μειονέκτημά τους είναι οι συσχετίσεις μεταξύ των παραγόντων που εμπλέκονται στη στατιστική ανάλυση. Η βασική δυσκολία στη στατιστική συμπερασματολογία είναι ότι ο αριθμός των παραμέτρων είναι πολύ μεγαλύτερος από τον αριθμό των πειραματικών εκτελέσεων που σημαίνει ότι κλασικές μέθοδοι όπως είναι η μέθοδος ελαχίστων τετραγώνων δεν μπορούν να χρησιμοποιηθούν.

Η ανάλυση των ΥΣ έχει προσελκύσει το ενδιαφέρον πολλών ερευνητών και για τον λόγο αυτό έχουν προταθεί αρκετές μέθοδοι για την αντιμετώπιση αυτού του προβλήματος. Τέτοιες μέθοδοι περιλαμβάνουν την κλασική κατά βήματα παλινδρόμηση και τη μέθοδο επιλογής του καλύτερου υποσυνόλου μεταβλητών (step-wise and all subset regression methods) [1, 87, 140, 142]. Οι Beattie et al. [6] πρότειναν μία μέθοδο επιλογής μοντέλου δύο σταδίων που συνδυάζει τη στοχαστική αναζήτηση επιλογής μεταβλητών (stochastic search variable selection, SSVS) με την εγγενή μέθοδο παραγόντων Bayes. Οι Holcomb et al. [64] επέκτειναν την απλή ιδέα της γραμμικής παλινδρόμησης σε μια μέθοδο πολλών μοντέλων ενώ ο Zhang et al. [158] πρότειναν τα μερικά ελάχιστα τετράγωνα και οι Li και Lin [82] πρότειναν τις μεθόδους συρρίκνωσης, όπως είναι η ποινή της ομαλά περικομμένης απόλυτης απόκλισης (smoothly clipped absolute deviation penalty, SCAD), για την ικανότητά τους να παρέχουν ταυτόχρονη επιλογή μεταβλητών και εκτίμηση των συντελεστών του μοντέλου. Παράλληλα οι Lu και Wu [100] παρουσίασαν μια τροποποιημένη κατά βήματα επιλογή, με βάση την ιδέα της σταδιακής μείωσης των διαστάσεων. Μετά από τους Candès και Tao [21] που εισήγαγαν τον γνωστό επιλογέα Dantzig Selector (DS), οι Phoa et al. [112] ήταν οι πρώτοι που χρησιμοποίησαν τον DS

για ΥΣ. Μερικά χρόνια αργότερα, ο Phoa [111] δημιούργησε άλλη μια μέθοδο, που ονομάζεται SRRS, για την ανίχνευση των σημαντικών επιδράσεων στους ΥΣ. Οι Koukouninos και Mylona [74] πρότειναν την εφαρμογή των μεθόδων ομαδικού κρησαρίσματος για την ανάλυση δεδομένων χρησιμοποιώντας  $E(f_{NOD})$ -βέλτιστους μικτών επιπέδων ΥΣ. Οι Marley και Woods [101] παρουσίασαν μια μέθοδο χρησιμοποιώντας έναν σταθμισμένο μέσο όρο των εκτιμήσεων ενώ, τον ίδιο χρόνο, οι Li et al. [85] πρότειναν μια νέα στρατηγική για την ανίχνευση ενεργών παραγόντων που βασίζεται στην ανάλυση συστάδων, γνωστή ως μέθοδος (contrast-orthogonality cluster analysis, COCA). Τέλος, οι Koukouninos et al. [75] πρότειναν μια άλλη μέθοδο για  $s$ -μπλοκ δύο επιπέδων  $E(f_{NOD})$ - βέλτιστους ΥΣ με  $N$  πειραματικές εκτελέσεις και  $p = s(N - 1)$  παράγοντες. Ορισμένες πρόσφατες μέθοδοι ανάλυσης περιλαμβάνουν τον δειγματολήπτη Gibbs [27], την πολυμεταβλητή μερικών ελαχίστων τετραγώνων κατά βήματα παλινδρόμηση [155], τη Μπεϋζιανή στρατηγική επιλογής μεταβλητών που πρότειναν οι Huang et al. [67], το κριτήριο cAIC των Das et al. [30] και τον αλγόριθμο MIC που προτάθηκε από τους Drosou et al. [36]. Μια περιεκτική επισκόπηση της ανάλυσης των ΥΣ δίνεται από τους Gupta και Kohli [59] και Georgίου [52]. Ωστόσο, οι περισσότερες από αυτές τις μεθόδους αποτυγχάνουν να βρουν τη λύση στο πρόβλημα όταν υπάρχουν πέραν του ενός ή των δύο ενεργών παραγόντων. Επιπρόσθετα, οι περισσότερες μελέτες περιορίζονται σε σχεδιασμούς δύο επιπέδων. Το πρόβλημα της ανάλυσης δεδομένων πολλαπλών και μικτών επιπέδων ΥΣ έχει μελετηθεί μόνο από μερικούς ερευνητές όπως οι Zhang et al. [158] με την PLSVS και οι Phoa et al. [112] με το Dantzig Selector.

Το πρόβλημα που υπάρχει στην ανάλυση των ΥΣ είναι η ύπαρξη μικρής μεροληψίας ανάμεσα στις εκτιμημένες επιδράσεις, αφού όπως ήδη έχουμε πει, οι σχεδιασμοί αυτοί δεν είναι ορθογώνιοι. Αυτό σημαίνει ότι για να αναγνωρίσουμε έναν σημαντικό παράγοντα, η επίδρασή του θα πρέπει να είναι τόσο μεγάλη, όσο να μην μπορεί να επισκιαστεί από το πειραματικό σφάλμα και την συνδυασμένη επίδραση των μη σημαντικών παραγόντων. Η μέθοδος ανάλυσης που θα επιλεγεί για έναν ΥΣ θα πρέπει να διατηρεί τις τιμές αυτού του σφάλματος στο χαμηλότερο δυνατό επίπεδο. Η επιτυχία έγκειται στην αναγνώριση των σημαντικών επιδράσεων με τα μικρότερα δυνατά σφάλματα Τύπου I (μέσος όρος μη ενεργών παραγόντων που εσφαλμένα αναγνωρίστηκαν ως ενεργοί) και Τύπου II (μέσος όρος ενεργών παραγόν-

ντων που εσφαλμένα αναγνωρίστηκαν ως μη ενεργοί).

Στην παρούσα διδακτορική διατριβή προτείνουμε δύο μεθόδους ανάλυσης ΥΣ, τη μέθοδο SVR-RFE (κεφάλαιο 6) που εφαρμόζεται τόσο σε ΥΣ δύο επιπέδων όσο και σε ΥΣ πολλαπλών και μικτών επιπέδων, και τη μέθοδο που βασίζεται στη SIS (κεφάλαιο 7) οι οποίες κρατούν το πειραματικό σφάλμα σε εξαιρετικά χαμηλό επίπεδο.

---

## Κατασκευή πολλαπλών και μικτών επιπέδων $k$ –κυκλικών υπερκορεσμένων σχεδιασμών

---

**Σ**το παρόν κεφάλαιο παρουσιάζουμε μία μέθοδο κατασκευής υπερκορεσμένων σχεδιασμών πολλαπλών και μικτών-επιπέδων. Πιο συγκεκριμένα, παρέχουμε ένα νέο κάτω φράγμα για την βελτιστοποίηση τέτοιων σχεδιασμών. Αυτό το φράγμα αποδεικνύεται ότι είναι ισχυρό και αρκετά γενικό αφού μπορεί να εφαρμοστεί σε ισορροπημένους και μη ισορροπημένους σχεδιασμούς. Επιπλέον, κατασκευάζονται και αξιολογούνται νέες οικογένειες ισορροπημένων και μη ισορροπημένων υπερκορεσμένων σχεδιασμών. Οι νέοι σχεδιασμοί κατασκευάζονται εφαρμόζοντας τη μεθοδολογία των  $k$ –κυκλικών σχεδιασμών για τους σχεδιασμούς πολλαπλών και μικτών επιπέδων. Επιπλέον, αναπτύσσονται και αποδεικνύονται πρόσθετες συνθήκες που είναι απαραίτητες για τους γεννήτορες και οι οποίες εξασφαλίζουν ότι οι αντίστοιχοι  $k$ –κυκλικοί υπερκορεσμένοι σχεδιασμοί είναι βέλτιστοι. Η παρούσα μεθοδολογία οδηγεί σε μια σειρά νέων, βελτιωμένων, οικογενειών υπερκορεσμένων σχεδιασμών και παρέχει εργαλεία για την άμεση κατασκευή βέλτιστων ή σχεδόν βέλτιστων  $k$ –κυκλικών σχεδιασμών, ελέγχοντας απλά τον αντίστοιχο γεννήτορα. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [IV].

## 5.1 Ερευνητικό πρόβλημα-Λύση

Από μία σύντομη ανασκόπηση στη βιβλιογραφία (κεφάλαιο 4.2) αντιλαμβανόμαστε ότι οι περισσότερες από τις υπάρχουσες μελέτες έχουν επικεντρωθεί στους ισορροπημένους ΥΣ όπου όλα τα επίπεδα εμφανίζονται τον ίδιο αριθμό φορών σε κάθε παράγοντα. Οι παραπάνω μελέτες θεωρούν ότι όλα τα επίπεδα του παράγοντα εμφανίζονται τον ίδιο αριθμό φορών σε κάθε στήλη (παράγοντα), δηλαδή ότι ο σχεδιασμός είναι ισορροπημένος. Ωστόσο, είναι συνηθισμένο στα προβλήματα της πραγματικής ζωής να αντιμετωπίζουμε καταστάσεις όπου ο αριθμός των πειραματικών εκτελέσεων ( $N$ ) δεν είναι πολλαπλάσιο του αριθμού των επιπέδων ( $s$ ). Αξίζει να σημειωθεί ότι οι μη ισορροπημένοι σχεδιασμοί μπορούν να χρησιμοποιηθούν στις περιπτώσεις όπου ορισμένες θεραπείες μπορεί να είναι πιο ακριβές ή πιο δύσκολο να εκτελεστούν από άλλες ή όταν ορισμένοι συνδυασμοί θεραπείας μπορεί να παρουσιάζουν ιδιαίτερο ενδιαφέρον, έτσι ο πειραματιστής επιλέγει να πειραματιστεί περισσότερο πάνω σ' αυτούς. Περιπτώσεις μη ισορροπημένων σχεδιασμών έχουν ήδη μελετηθεί αποτελεσματικά σε πολλά διαφορετικά πεδία έρευνας, συμπεριλαμβανομένων και των υπερκορεσμένων σχεδιασμών (βλέπε [60, 68]) και split-plot σχεδιασμών (βλέπε [110]). Πρόσφατα έχουν παρουσιαστεί μελέτες ΥΣ όπου το  $N$  δεν είναι πολλαπλάσιο του  $s$ . Σε αυτήν την κατεύθυνση κινήθηκαν οι Nguyen και Cheng [108], οι Bulutoglu και Ryan [19] και οι Suen και Das [125], οι οποίοι θεώρησαν δύο επιπέδων  $E(s^2)$ -ΥΣ με  $N$  περιττό. Ομοίως, οι Chen και Liu [25] και οι Chai et al. [22] θεώρησαν βέλτιστους ΥΣ για γενικό αριθμό πειραματικών εκτελέσεων. Σε τέτοιες περιπτώσεις, επιθυμούμε να έχουμε «σχεδόν» ορθογωνιότητα καθεμιάς από τις επιδράσεις των παραγόντων με τη μέση επίδραση.

Στην παρούσα ενότητα παρουσιάζουμε την κατασκευή υπερκορεσμένων σχεδιασμών πολλαπλών και μικτών-επιπέδων. Για την κατασκευή αυτή τροποποιήσαμε την  $k$ -κυκλική μέθοδο για την άμεση αναζήτηση «καλών» (βέλτιστων ή σχεδόν βέλτιστων), ισορροπημένων ή όχι, πολλαπλών επιπέδων ή μικτών επιπέδων, ΥΣ. Αποδεικνύουμε ένα νέο κάτω φράγμα για το  $E(f_{NOD})$  κριτήριο το οποίο εφαρμόζεται και σε μη ισορροπημένους ΥΣ. Στην πραγματικότητα το φράγμα αυτό αποτελεί μία γενίκευση άλλων, ήδη υπαρχόντων φραγμάτων. Εξασφαλίζουμε επίσης τις αναγκαίες προϋποθέσεις για τον γεννήτορα, έτσι ώστε οι αντίστοιχοι ΥΣ, που κατασκευάστηκαν από τον γεννήτορα αυτό και την  $k$ -κυκλική δομή, να είναι  $E(f_{NOD})$ -βέλτιστοι. Η πιο σημαν-

τική συνεισφορά αυτής της μεθοδολογίας είναι ότι οι σχεδιασμοί που κατασκευάστηκαν εφαρμόζονται σε κάθε ΥΣ ανεξάρτητα αν ο σχεδιασμός είναι ισορροπημένος ή όχι, πολλαπλών ή μικτών επιπέδων.

## 5.2 Συμβολισμός και βασικές έννοιες

Στο παρόν κεφάλαιο παρουσιάζονται ο σχετικός συμβολισμός και κάποιες θεωρητικές έννοιες που θα φανούν χρήσιμες στα επόμενα κεφάλαια. Θεωρούμε ένα πείραμα που περιλαμβάνει  $m$  παράγοντες  $F_1, F_2, \dots, F_m$  με επίπεδα  $s_1, s_2, \dots, s_m$ , αντίστοιχα. Ένας μικτών επιπέδων σχεδιασμός,  $d$ , που έχει  $N$  πειραματικές εκτελέσεις και  $m$  παράγοντες μπορεί να εκφραστεί ως ένας  $N \times m$  πίνακας τέτοιος ώστε η  $j$ -οστή στήλη έχει στοιχεία  $1, 2, \dots, s_j$ . Στην περίπτωση όπου το  $N-1 < \sum_{j=1}^m (s_j-1)$ , ο σχεδιασμός καλείται ΥΣ. Εδώ χρησιμοποιούμε το συμβολισμό  $D(N, s_1 \cdots s_m)$  για να δηλώσουμε μία κλάση μικτών επιπέδων ΥΣ με  $N$ -πειραματικές εκτελέσεις, που περιλαμβάνει  $m$  παράγοντες με επίπεδα  $s_1, s_2, \dots, s_m$ , αντίστοιχα. Στην περίπτωση που κάποια  $s_j$ 's είναι ίσα τότε ο σχεδιασμός αυτός συμβολίζεται με  $D(N, s_1^{r_1} \cdots s_l^{r_l})$ , όπου  $\sum_{j=1}^l r_j = m$ , ενώ όταν όλα τα  $s_j$ 's είναι ίσα ( $= s$ ), ο σχεδιασμός είναι ένας πολλαπλών-επιπέδων ΥΣ, και συμβολίζεται με  $D(N, s^m)$ .

Για έναν ισορροπημένο σχεδιασμό, οι συνδυασμοί των επιπέδων  $N$  θα πρέπει να είναι πολλαπλάσιο όλων των  $s_j$ . Στο παρόν κεφάλαιο θεωρούμε ότι ο αριθμός των φορών που κάθε επίπεδο εμφανίζεται σε κάθε στήλη διαφέρει το πολύ κατά ένα. Έτσι, για κάθε δοθέν  $N$ , είναι πιθανό να βρούμε μη αρνητικά  $l_j$  και  $u_j$  τέτοια ώστε, για  $1 \leq j \leq m$ ,  $N = l_j s_j + u_j$  όπου  $l_j = \lfloor N/s_j \rfloor$ ,  $u_j < s_j$ , και  $\lfloor N/s_j \rfloor$  δηλώνει το ακέραιο μέρος του  $N/s_j$ . Ως εκ τούτου, στην  $j$ -οστή στήλη έχουμε ότι  $u_j$  από τα  $s_j$  επίπεδα εμφανίζονται  $l_j + 1$  φορές, και τα υπόλοιπα  $s_j - u_j$  επίπεδα εμφανίζονται  $l_j$  φορές. Πρέπει να σημειωθεί ότι αυτή η επιλογή του  $N$  είναι γενικά κατάλληλη για μη ισορροπημένους σχεδιασμούς θεωρώντας τους ισορροπημένους ως μία ειδική περίπτωση όταν  $u_j = 0$ .

Δοθέντος ότι ο  $j$ -οστός παράγοντας κωδικοποιείται ως  $1, \dots, s_j$ , για  $1 \leq j \leq m$ , υποθέτουμε ότι οι  $v = \prod_{j=1}^m s_j$  συνδυασμοί επιπέδων εκπροσωπούνται από το  $m$ -οστό  $\alpha_1 \dots \alpha_m$  ( $1 \leq \alpha_j \leq s_j, 1 \leq j \leq m$ ), το οποίο υποτίθεται ότι είναι λεξικογραφικά διατεταγμένο. Επιπρόσθετα θεωρούμε ως  $n_{\alpha_j}^j$ , για  $1 \leq j \leq m$ , τον αριθμό των φορών

που τα επίπεδα  $\alpha_j$  εμφανίζονται στον παράγοντα  $F_j$ ,  $1 \leq \alpha_j \leq s_j$  και ως  $n_{\alpha_{j_1} \alpha_{j_2}}^{j_1 j_2}$ , για  $1 \leq j_1 < j_2 \leq m$ , τον αριθμό των φορών που τα επίπεδα  $\alpha_{j_1}$  και  $\alpha_{j_2}$  εμφανίζονται στους παράγοντες  $F_{j_1}$  και  $F_{j_2}$ , αντίστοιχα, για  $1 \leq \alpha_{j_i} \leq s_{j_i}, i = 1, 2$ .

**Παράδειγμα 6** Θεωρούμε ένα πείραμα που περιλαμβάνει 2 παράγοντες  $F_1, F_2$  στα επίπεδα  $s_1 = 2$  και  $s_2 = 3$  αντίστοιχα και με  $N = 4$  πειραματικές εκτελέσεις. Παρακάτω εξηγούμε τον συμβολισμό που περιγράφεται στην προηγούμενη παράγραφο. Το σύμβολο  $D(4, 2^1 3^1)$  αντιπροσωπεύει μια κλάση 4-πειραματικών εκτελέσεων ΥΣ που περιλαμβάνουν τους παράγοντες  $F_1$  και  $F_2$ . Θεωρούμε οποιοδήποτε σχεδιασμό  $d \in D(4, 2^1 3^1)$ . Εδώ το  $v = 6$  και όλοι οι δυνατοί συνδυασμοί επιπέδων  $\alpha_1 \alpha_2$  και ο σχεδιασμός  $d$  που παρουσιάζεται παρακάτω:

$$\begin{array}{|c|c|c|c|c|c|} \hline \alpha_1 \alpha_2 & 11 & 12 & 13 & 21 & 22 & 23 \\ \hline \end{array}, \quad d = \begin{array}{|c|c|c|c|} \hline 11 & 12 & 13 & 21 \\ \hline \end{array}$$

Επίσης, εδώ έχουμε  $l_1 = 2, u_1 = 0$  και  $l_2 = 1, u_2 = 1$ . Τέλος, για να εξηγήσουμε τα υπόλοιπα σύμβολα, παρουσιάζουμε τους παρακάτω πίνακες:

$$\begin{array}{|c|c|} \hline \alpha_1 & n_{\alpha_1}^1 \\ \hline 1 & 3 \\ \hline 2 & 1 \\ \hline \end{array}, \quad \begin{array}{|c|c|} \hline \alpha_2 & n_{\alpha_2}^2 \\ \hline 1 & 2 \\ \hline 2 & 1 \\ \hline 3 & 1 \\ \hline \end{array} \quad \text{και} \quad \begin{array}{|c|c|c|} \hline n_{11}^{12} = 1 & n_{12}^{12} = 1 & n_{13}^{12} = 1 \\ \hline n_{21}^{12} = 1 & n_{22}^{12} = 0 & n_{23}^{12} = 0 \\ \hline \end{array}$$

**Ορισμός 1** Ένας σχεδιασμός  $d \in D(N, \prod_{j=1}^m s_j)$  θα λέγεται ότι είναι ένας σχεδόν ισορροπημένος μικτών επιπέδων ΥΣ εάν  $N < \sum_{j=1}^m (s_j - 1)$  και για τα επίπεδα  $1 \leq j \leq m$ ,  $(s_j - u_j)$  του παράγοντα  $F_j$  εμφανίζονται  $l_j$  φορές και τα υπολειπόμενα  $u_j$  επίπεδα εμφανίζονται  $(l_j + 1)$  φορές στον παράγοντα  $F_j$ .

**Ορισμός 2** Ένας σχεδιασμός  $d \in D(N, \prod_{j=1}^m s_j)$  θα λέγεται ότι είναι  $k$ -κυκλικός εάν το  $d$  λαμβάνεται από τον γεννήτορα  $(g_1, g_2, \dots, g_m)$  με την κυκλική εναλλαγή  $k$  στοιχείων σε κάθε βήμα.

Βασίζομενοι στον ορισμό 2, για ένα δοθέν σύνολο παραμέτρων ενός σχεδιασμού,  $N, s_1, \dots, s_m$ , μπορεί κανείς να αποκτήσει αρκετούς  $k$ -κυκλικούς ΥΣ πολλαπλών είτε μικτών επιπέδων.



**Παράδειγμα 7** Θεωρούμε έναν 5-επιπέδων 3-κυκλικό ΥΣ  $d \in D(10, 5^{27})$ , με  $m = 27$  παράγοντες. Αυτός ο σχεδιασμός προκύπτει από τον γεννήτορα

$$(5, 4, 4, 2, 2, 5, 2, 5, 3, 1, 1, 4, 3, 3, 1, 4, 1, 3, 4, 2, 1, 3, 4, 2, 1, 3, 2),$$

με την κυκλική εναλλαγή 3 στοιχείων σε κάθε βήμα, είναι ένας  $E(f_{NOD})$ -βέλτιστος ΥΣ και αναπαριστάται στον Πίνακα 5.1.

**Πίνακας 5.1:** Ένας ισορροπημένος 5-επιπέδων 3-κυκλικός ΥΣ με  $m = 27$  και  $N = 10$  πειραματικές εκτελέσεις. Αυτός ο σχεδιασμός είναι ένας  $E(f_{NOD})$ -βέλτιστος ΥΣ.

5	4	4	2	2	5	2	5	3	1	1	4	3	3	1	4	1	3	4	2	1	3	4	2	1	3	2
1	3	2	5	4	4	2	2	5	2	5	3	1	1	4	3	3	1	4	1	3	4	2	1	3	4	2
3	4	2	1	3	2	5	4	4	2	2	5	2	5	3	1	1	4	3	3	1	4	1	3	4	2	1
4	2	1	3	4	2	1	3	2	5	4	4	2	2	5	2	5	3	1	1	4	3	3	1	4	1	3
4	1	3	4	2	1	3	4	2	1	3	2	5	4	4	2	2	5	2	5	3	1	1	4	3	3	1
3	3	1	4	1	3	4	2	1	3	4	2	1	3	2	5	4	4	2	2	5	2	5	3	1	1	4
1	1	4	3	3	1	4	1	3	4	2	1	3	4	2	1	3	2	5	4	4	2	2	5	2	5	3
2	5	3	1	1	4	3	3	1	4	1	3	4	2	1	3	4	2	1	3	2	5	4	4	2	2	5
2	2	5	2	5	3	1	1	4	3	3	1	4	1	3	4	2	1	3	4	2	1	3	2	5	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5

Ας θεωρήσουμε ένα 2-κυκλικό μικτών επιπέδων ΥΣ  $d \in D(12, 2^{11}6^{11})$  για  $m = 22$  παράγοντες. Αυτός ο σχεδιασμός προέρχεται από τον γεννήτορα

$$(2, 5, 1, 3, 2, 5, 1, 2, 1, 4, 1, 1, 2, 2, 2, 3, 2, 4, 1, 6, 2, 6),$$

με την κυκλική εναλλαγή 2 στοιχείων σε κάθε βήμα και παρουσιάζεται στον Πίνακα 5.2.

Λόγω του γεγονότος ότι μπορούν να κατασκευαστούν αρκετοί ΥΣ χρησιμοποιώντας την προαναφερθείσα μεθοδολογία καθώς και άλλες υπάρχουσες μεθόδους κατασκευής που έχουν προταθεί στη βιβλιογραφία, δημιουργείται η αναγκαιότητα σύγκρισης τέτοιων σχεδιασμών έτσι ώστε να καταλήξουμε στον βέλτιστο ΥΣ που τελικά θα χρησιμοποιήσουμε. Οι Fang *et al.* [47] πρότειναν το  $E(f_{NOD})$  κριτήριο για την αξιολόγηση και τη σύγκριση των ΥΣ με μικτά επίπεδα υπό το πρίσμα της ορθογωνιότητας

**Πίνακας 5.2:** Ένας σχεδόν ισορροπημένος μικτών επιπέδων 2-κυκλικός ΥΣ με  $m = 22$  παράγοντες και  $N = 12$  πειραματικές εκτελέσεις. Αυτός ο σχεδιασμός είναι ένας  $E(f_{NOD})$ -βέλτιστος ΥΣ.

2	5	1	3	2	5	1	2	1	4	1	1	2	2	2	3	2	4	1	6	2	6
2	6	2	5	1	3	2	5	1	2	1	4	1	1	2	2	2	3	2	4	1	6
1	6	2	6	2	5	1	3	2	5	1	2	1	4	1	1	2	2	2	3	2	4
2	4	1	6	2	6	2	5	1	3	2	5	1	2	1	4	1	1	2	2	2	3
2	3	2	4	1	6	2	6	2	5	1	3	2	5	1	2	1	4	1	1	2	2
2	2	2	3	2	4	1	6	2	6	2	5	1	3	2	5	1	2	1	4	1	1
1	1	2	2	2	3	2	4	1	6	2	6	2	5	1	3	2	5	1	2	1	4
1	4	1	1	2	2	2	3	2	4	1	6	2	6	2	5	1	3	2	5	1	2
1	2	1	4	1	1	2	2	2	3	2	4	1	6	2	6	2	5	1	3	2	5
2	5	1	2	1	4	1	1	2	2	2	3	2	4	1	6	2	6	2	5	1	3
1	3	2	5	1	2	1	4	1	1	2	2	2	3	2	4	1	6	2	6	2	5
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

και της ομοιομορφίας. Το κριτήριο αυτό ορίζεται ως

$$E(f_{NOD}) = \sum_{1 \leq j_1 \neq j_2 \leq m} f_{NOD}^{j_1 j_2} / m(m-1), \quad (5.1)$$

όπου

$$f_{NOD}^{j_1 j_2} = \sum_{\alpha_{j_1}=1}^{s_{j_1}} \sum_{\alpha_{j_2}=1}^{s_{j_2}} \left( n_{\alpha_{j_1} \alpha_{j_2}}^{j_1 j_2} - \frac{N}{s_{j_1} s_{j_2}} \right)^2 = \sum_{\alpha_{j_1}=1}^{s_{j_1}} \sum_{\alpha_{j_2}=1}^{s_{j_2}} \left( n_{\alpha_{j_1} \alpha_{j_2}}^{j_1 j_2} \right)^2 - \frac{N^2}{s_{j_1} s_{j_2}}$$

Ο ορισμός του  $E(f_{NOD})$  στην εξίσωση (5.1) υπολογίζει την απόσταση  $X^2$  μεταξύ των προβολών δύο παραγόντων και της ομοιόμορφης κατανομής, και επομένως αντιπροσωπεύει μια απόσταση μεταξύ του σχεδιασμού και ενός ορθογώνιου σχεδιασμού ισχύος 2. Έτσι, το κριτήριο  $E(f_{NOD})$  μετρά την έκταση της μη ορθογωνιότητας μεταξύ των στηλών του σχεδιασμού  $d$ . Ακόμη και αν αυτό το κριτήριο ορίζεται κυρίως για ισορροπημένους ΥΣ, μπορεί να εφαρμοστεί στην περίπτωση των σχεδόν ισορροπημένων σχεδιασμών ή ακόμη και στην περίπτωση των μη ισορροπημένων σχεδιασμών. Το  $E(f_{NOD})$  είναι αμετάβλητο κατά την μετάθεση των επιπέδων σε μία ή περισσότερες από τις στήλες του σχεδιασμού. Υπάρχουν πολλά κριτήρια βελτιστότητας για την αξιολόγηση των ΥΣ μικτών επιπέδων (4.2.1). Στο παρόν κεφάλαιο, θεωρούμε ως κύριο κριτήριο το  $E(f_{NOD})$  προσαρμοσμένο για σχεδιασμούς με γενικό μέγεθος πειραματικών εκτελέσεων. Για αυτό το κριτήριο παρέχουμε νέα κάτω φράγματα που ισχύουν σε όλες τις περιπτώσεις, συμπεριλαμβανομένης και της περίπτωσης των μη

ισορροπημένων σχεδιασμών. Επιπλέον, αναπτύσσουμε νέες μεθόδους κατασκευής και αποδεικνύουμε ότι οι σχεδιασμοί που δημιουργούνται από αυτές τις μεθόδους είναι βέλτιστοι βάσει του κριτηρίου  $E(f_{\text{NOD}})$  και του κάτω φράγματος που αποδείξαμε.

### 5.3 Νέα κάτω φράγματα του $E(f_{\text{NOD}})$

Από μία μικρή βιβλιογραφική ανασκόπηση, σημειώσαμε ότι είναι εξαιρετικά περίπλοκη και υπολογιστικά αδύνατη η αναζήτηση και κατασκευή ενός  $E(f_{\text{NOD}})$ -βέλτιστου ΥΣ. Ως εκ τούτου, υπάρχει η ανάγκη να αποδειχθεί ένα κάτω φράγμα (LB) για το κριτήριο  $E(f_{\text{NOD}})$  το οποίο θα χρησιμεύσει ως σημείο αναφοράς για την κατασκευή βέλτιστων ΥΣ. Τα ακόλουθα Λήμματα, Λήμμα 8 και Λήμμα 9, συμβάλλουν στην απόκτηση ενός γενικού κάτω φράγματος για το κριτήριο  $E(f_{\text{NOD}})$ , το οποίο παρουσιάζεται στο Θεώρημα 1. Κάτω από αυτό το LB, κατασκευάσαμε αρκετούς ισορροπημένους ή μη, πολλαπλών ή μικτών επιπέδων  $E(f_{\text{NOD}})$ -βέλτιστους ΥΣ.

Για να αποδείξουμε τα ακόλουθα Λήμματα είναι απαραίτητο να ορίσουμε τις ακόλουθες σχέσεις. Για  $1 \leq j_1 < j_2 \leq m$ ,

$$r_{j_1 j_2} = \frac{N}{s_{j_1} s_{j_2}}, \quad \eta_{j_1 j_2}^{(1)} = \lfloor r_{j_1 j_2} \rfloor, \quad \text{και} \quad \eta_{j_1 j_2}^{(2)} = r_{j_1 j_2} - \eta_{j_1 j_2}^{(1)}, \quad (5.2)$$

όπου  $\lfloor r_{j_1 j_2} \rfloor$  είναι το ακέραιο μέρος του  $r_{j_1 j_2}$  και συμβολίζεται με  $\eta_{j_1 j_2}^{(1)}$  ενώ το  $r_{j_1 j_2} - \eta_{j_1 j_2}^{(1)}$  είναι το κλασματικό μέρος του  $r_{j_1 j_2}$  και συμβολίζεται με  $\eta_{j_1 j_2}^{(2)}$ .

Τώρα μπορούμε να παρέχουμε το ακόλουθο λήμμα το οποίο δίνει ένα νέο κάτω όριο για το κριτήριο  $E(f_{\text{NOD}})$ .

**Λήμμα 8** Για κάθε σχεδιασμό  $d \in D(N, s_1 \cdots s_m)$ , έχουμε ότι,

$$E(f_{\text{NOD}}) \geq LB_1,$$

όπου

$$LB_1 = \sum_{1 \leq j_1 \neq j_2 \leq m} \left[ s_{j_1} s_{j_2} \left\{ r_{j_1 j_2}^2 + \eta_{j_1 j_2}^{(2)} (1 - \eta_{j_1 j_2}^{(2)}) \right\} - \frac{N^2}{s_{j_1} s_{j_2}} \right] / m(m-1).$$

ή ισοδύναμα

$$LB_1 = \sum_{1 \leq j_1 \neq j_2 \leq m} s_{j_1} s_{j_2} \left\{ \eta_{j_1 j_2}^{(2)} (1 - \eta_{j_1 j_2}^{(2)}) \right\} / m(m-1).$$

**Απόδειξη.** Για  $1 \leq j_1 < j_2 \leq m$ , έχουμε ότι,

$$\sum_{\alpha_{j_1}=1}^{s_{j_1}} \sum_{\alpha_{j_2}=1}^{s_{j_2}} n_{\alpha_{j_1} \alpha_{j_2}}^{j_1 j_2} = N.$$

Για  $1 \leq j_1 < j_2 \leq m$ , έστω  $t_1$  και  $t_2$  θετικοί ακέραιοι τέτοιοι ώστε

$$t_2 \eta_{j_1 j_2}^{(1)} + t_1 (1 + \eta_{j_1 j_2}^{(1)}) = s_{j_1} s_{j_2} r_{j_1 j_2} = N \quad \text{και} \quad t_1 + t_2 = s_{j_1} s_{j_2}. \quad (5.3)$$

Από την εξίσωση (5.3), προκύπτει αμέσως ότι

$$t_1 = s_{j_1} s_{j_2} \eta_{j_1 j_2}^{(2)} \quad \text{και} \quad t_2 = s_{j_1} s_{j_2} (1 - \eta_{j_1 j_2}^{(2)}),$$

όπου  $\eta_{j_1 j_2}^{(2)} = r_{j_1 j_2} - \eta_{j_1 j_2}^{(1)}$  δίνει το κλασματικό μέρος του  $r_{j_1 j_2}$ . Επιπλέον, προκύπτει ότι

$$\begin{aligned} \sum_{\alpha_{j_1}=1}^{s_{j_1}} \sum_{\alpha_{j_2}=1}^{s_{j_2}} \left( n_{\alpha_{j_1} \alpha_{j_2}}^{j_1 j_2} \right)^2 &\geq \left[ t_2 (\eta_{j_1 j_2}^{(1)})^2 + t_1 (1 + (\eta_{j_1 j_2}^{(1)}))^2 \right] \\ &= \left[ t_2 (r_{j_1 j_2} - \eta_{j_1 j_2}^{(2)})^2 + t_1 (1 - \eta_{j_1 j_2}^{(2)} + r_{j_1 j_2})^2 \right] \\ &= s_{j_1} s_{j_2} \left[ r_{j_1 j_2}^2 + \eta_{j_1 j_2}^{(2)} (1 - \eta_{j_1 j_2}^{(2)}) \right] \end{aligned}$$

Αυτό ολοκληρώνει την απόδειξη.

Ομοίως, μπορούμε να αποκτήσουμε ένα επιπλέον κάτω φράγμα του  $E(f_{NOD})$  όπως παρουσιάζεται στο Λήμμα 9. Για να αποδείξουμε το Λήμμα 9, θα χρειαστούμε τα ακόλουθα. Για  $1 \leq i_1, i_2 \leq N$ ,

$$c_{i_1 i_2} = \sum_{j=1}^m c_{i_1 i_2}^j,$$

όπου

$$c_{i_1 i_2}^j = \begin{cases} 1, & \text{εάν οι είσοδοι της } i_1 \text{ και της } i_2 \text{ γραμμής της } j\text{-οστής στήλης του } d \text{ συμπίπτουν,} \\ 0, & \text{αλλιώς.} \end{cases}$$

Υπενθυμίζουμε ότι ο σχεδιασμός  $d$  είναι ένας  $N \times m$  πίνακας. Για  $1 \leq i \leq N$  έχουμε ότι  $c_{ii} = m$ . Επιπρόσθετα,

$$\sum_{1 \leq i_1 \neq i_2 \leq N} c_{i_1 i_2} = \sum_{j=1}^m s_j l_j (l_j - 1) + \sum_{j=1}^m 2u_j l_j = c.$$

Ορίζουμε,

$$q = \frac{c}{N(N-1)}, \quad \zeta^{(1)} = \lfloor q \rfloor, \quad \text{και} \quad \zeta^{(2)} = q - \zeta^{(1)}.$$

Έστω  $p_1$  και  $p_2$  θετικοί ακέραιοι τέτοιοι ώστε

$$p_2 \zeta^{(1)} + p_1 (1 + \zeta^{(1)}) = c \quad \text{και} \quad p_1 + p_2 = N(N-1). \quad (5.4)$$

Σημειώνουμε ότι τόσο το Λήμμα 8 όσο και το Λήμμα 9 ισχύουν για οποιοδήποτε σχεδιασμό  $d = D(N; s_1, \dots, s_m)$ , είτε πρόκειται για ένα ισορροπημένο σχεδιασμό είτε όχι.

**Λήμμα 9** Για κάθε σχεδόν ισορροπημένο σχεδιασμό  $d \in D(N, s_1 \cdots s_m)$ ,

$$E(f_{\mathbf{NOD}}) \geq LB_2,$$

όπου,

$$LB_2 = \frac{N(N-1) [q^2 + \zeta^{(2)}(1 - \zeta^{(2)})]}{m(m-1)} + \frac{Nm}{m-1} - \frac{\sum_{j=1}^m [l_j^2 s_j + 2l_j u_j + u_j]}{m(m-1)} - \frac{N^2 \sum_{1 \leq j_1 \neq j_2 \leq m} (s_{j_1} s_{j_2})^{-1}}{m(m-1)}.$$

**Απόδειξη.** Για  $1 \leq j_1 < j_2 \leq m$ , έστω  $N_{j_1 j_2} = \binom{j_1 j_2}{\alpha_{j_1} \alpha_{j_2}}$ . Για κάθε στήλη  $d^j$  του

$d$ , έστω  $W^j = (w_{il}^j)$  ένας  $N \times s_j$  πίνακας, όπου για  $i = 1, \dots, N$ ,  $l = 1, \dots, s_j$ ,

$$w_{il}^j = \begin{cases} 1, & \text{αν } x_{ij} = l \\ 0, & \text{αλλιώς.} \end{cases}, \quad (5.5)$$

και έστω  $W = (W^1, W^2, \dots, W^m)$ . Από την εξίσωση (5.4), μπορούμε εύκολα να διαπιστώσουμε ότι το  $W$  πληροί την  $WW^T = (c_{i_1 i_2})$  και την

$$W^T W = \begin{pmatrix} A_1 & N_{12} & \cdots & N_{1m} \\ N_{21} & A_2 & \cdots & N_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ N_{m1} & N_{m2} & \cdots & A_m \end{pmatrix},$$

όπου, για  $1 \leq j \leq m$ ,  $A_j$  είναι ένας  $s_j \times s_j$  πίνακας του οποίου όλα τα στοιχεία εκτός-διαγωνίου είναι μηδενικά και ανάμεσα σε όλα τα διαγώνια στοιχεία  $s_j$ , τα στοιχεία  $u_j$  είναι ίσα με  $l_j$  και τα υπόλοιπα στοιχεία είναι  $l_j + 1$ . Αυτό προκύπτει από τον ορισμό του  $W^j$  και από τον Ορισμό 1 για τους σχεδόν ισορροπημένους μικτών επιπέδων ΥΣ. Σημειώνουμε ότι ο πίνακας  $W$  είναι πράγματι ένας induced πίνακας [128].

Ξανά, από τον ορισμό του  $f_{\text{NOD}}^{j_1 j_2}$ , παίρνουμε

$$\begin{aligned} f_{\text{NOD}}^{j_1 j_2} &= \sum_{\alpha_1=1}^{s_{j_1}} \sum_{\alpha_2=1}^{s_{j_2}} \left( n_{\alpha_{j_1} \alpha_{j_2}}^{j_1 j_2} \right)^2 - \frac{N^2}{s_{j_1} s_{j_2}} \\ &= \text{tr}(N_{j_1 j_2} N_{j_2 j_1}) - \frac{N^2}{s_{j_1} s_{j_2}}. \end{aligned}$$

Έτσι, μπορούμε να γράψουμε το  $E(f_{\text{NOD}})$  ως

$$\begin{aligned} E(f_{\text{NOD}}) &= \sum_{1 \leq j_1 \neq j_2 \leq m} \left[ \text{tr}(N_{j_1 j_2} N_{j_2 j_1}) - \frac{N^2}{s_{j_1} s_{j_2}} \right] / m(m-1) \\ &= \frac{\sum_{1 \leq j_1 \neq j_2 \leq m} \text{tr}(N_{j_1 j_2} N_{j_2 j_1})}{m(m-1)} - \frac{N^2 \sum_{1 \leq j_1 \neq j_2 \leq m} (s_{j_1} s_{j_2})^{-1}}{m(m-1)}. \end{aligned} \quad (5.6)$$

έχουμε ότι

$$\begin{aligned}
\sum_{1 \leq j_1 \neq j_2 \leq m} \text{tr}(N_{j_1 j_2} N_{j_2 j_1}) &= \text{tr}[(W^T W)^2] - \sum_{j=1}^m \text{tr}(A_j^T A_j) \\
&= \text{tr}[(W W^T)^2] - \sum_{j=1}^m [l_j^2 s_j + 2l_j u_j + u_j] \\
&= \sum_{1 \leq i_1 \neq i_2 \leq N} c_{i_1 i_2}^2 + Nm^2 - \sum_{j=1}^m [l_j^2 s_j + 2l_j u_j + u_j].
\end{aligned}$$

Επομένως,

$$\begin{aligned}
E(f_{\text{NOD}}) &= \frac{\sum_{1 \leq i_1 \neq i_2 \leq N} c_{i_1 i_2}^2}{m(m-1)} + \frac{Nm}{m-1} - \frac{\sum_{j=1}^m [l_j^2 s_j + 2l_j u_j + u_j]}{m(m-1)} - \\
&\quad \frac{N^2 \sum_{1 \leq j_1 \neq j_2 \leq m} (s_{j_1} s_{j_2})^{-1}}{m(m-1)}. \tag{5.7}
\end{aligned}$$

Από την εξίσωση (5.4), αμέσως προκύπτει ότι

$$p_1 = N(N-1)\zeta^{(2)} \quad \text{και} \quad p_2 = N(N-1)(1 - \zeta^{(2)}).$$

Έτσι,

$$\begin{aligned}
\sum_{1 \leq i_1 \neq i_2 \leq N} c_{i_1 i_2}^2 &\geq [p_2(\zeta^{(1)})^2 + p_1(1 + (\zeta^{(1)})^2)] \\
&= N(N-1) [q^2 + \zeta^{(2)}(1 - \zeta^{(2)})]. \tag{5.8}
\end{aligned}$$

Από τις εξισώσεις (5.7) και (5.8), η απόδειξη του Λήμματος 2 ακολουθεί αμέσως.

### Τέλος απόδειξης

**Παρατήρηση 1** Σημειώνουμε ότι το Λήμμα 1 των Chen και Liu [25] μπορεί να θεωρηθεί ως ειδική περίπτωση του Λήμματος 9 διαλέγοντας τα  $u_j = 0$  για όλα τα  $j$  (όταν ο σχεδιασμός είναι ισορροπημένος τότε  $u_j = 0$  για όλα τα  $j$ ).

Συνδυάζοντας το Λήμμα 8 και το Λήμμα 9 μπορούμε να διατυπώσουμε το ακόλουθο θεώρημα το οποίο παρέχει ένα κάτω φράγμα για το κριτήριο  $E(f_{\text{NOD}})$ .

**Θεώρημα 1** Για κάθε σχεδιασμό  $d \in D(N, \prod_{j=1}^m s_j)$ ,

$$E(f_{NOD}) \geq LB$$

όπου,  $LB = \max\{LB_1, LB_2\}$ .

Το Θεώρημα 1 παρέχει ένα γενικό κάτω φράγμα του κριτηρίου  $E(f_{NOD})$ . Το Θεώρημα αυτό αποτελεί μία γενίκευση και περιλαμβάνει άλλα κάτω φράγματα ως ειδικές περιπτώσεις, όπως αυτό των Chen και Liu [25] το οποίο είναι για ισορροπημένους σχεδιασμούς ( $u_j = 0$  για όλα τα  $j$ ). Αυτό το LB εφαρμόζεται τόσο σε ισορροπημένους όσο και σε μη ισορροπημένους σχεδιασμούς, καθώς και σε σχεδιασμούς πολλαπλών και μικτών επιπέδων. Έτσι, μπορεί να χρησιμοποιηθεί ως σημείο αναφοράς βελτιστότητας ενός σχεδιασμού για την περίπτωση των μη ισορροπημένων σχεδιασμών. Κάθε ΥΣ που επιτυγχάνει αυτή την τιμή του LB είναι ο καλύτερος σχεδιασμός στην κλάση που ανήκει ως προς το κριτήριο  $E(f_{NOD})$  και αυτός ο σχεδιασμός ονομάζεται  $E(f_{NOD})$ -βέλτιστος.

## **5.4 Κατασκευή βέλτιστων πολλαπλών-επιπέδων $k$ -κυκλικών ΥΣ**

Σε αυτή την ενότητα, παρέχουμε τις απαραίτητες προϋποθέσεις ενός γεννήτορα από τον οποίο μπορούμε να κατασκευάσουμε έναν  $E(f_{NOD})$ -βέλτιστο πολλαπλών επιπέδων, ισορροπημένο ή μη,  $k$ -κυκλικό υπερκορεσμένο σχεδιασμό. Οι Georgiou και Koukouninos [55], παρήγαγαν τις ακόλουθες αναγκαίες και ικανές συνθήκες (Λήμμα 10) για την ύπαρξη ενός ισορροπημένου πολλαπλών επιπέδων  $k$ -κυκλικού υπερκορεσμένου σχεδιασμού, με την ιδιότητα των ίσων εμφανίσεων (ισορροπημένοι σχεδιασμοί). Σε αυτή την ενότητα, παρέχουμε τις αναγκαίες και ικανές συνθήκες (Θεώρημα 2 και Θεώρημα 3) για την ύπαρξη των πολλαπλών επιπέδων, σχεδόν-ισορροπημένων,  $k$ -κυκλικών ΥΣ. Η κατασκευή τέτοιων βέλτιστων σχεδιασμών παρέχεται μέσω του Πορίσματος 8 και του Θεωρήματος 4.

**Λήμμα 10** (Georgiou και Koukouninos [55]). Έστω  $d \in D(N, s^m)$  ένας  $k$ -κυκλικός υπερκορεσμένος σχεδιασμός. Υποθέτουμε ότι ο σχεδιασμός  $d$  δημιουργήθηκε από τον



γεννήτορα  $(g_1, g_2, \dots, g_m)$  με την κυκλική εναλλαγή  $k$  στοιχείων σε κάθε βήμα και προσθέτοντας μία γραμμή με όλα τα στοιχεία να είναι ίσα με  $s$ . Οι αναγκαίες και ικανές συνθήκες για έναν  $s$ -επιπέδων,  $k$ -κυκλικό, ΥΣ  $d$ , με την ιδιότητα των ίσων εμφανίσεων, είναι:

(i)  $N = st$ ,  $m = (N - 1)k = (st - 1)k$ , για κάποιο θετικό ακέραιο  $t$ .

(ii) Ο γεννήτορας περιέχει ακριβώς  $kt$  στοιχεία ίσα με  $1, 2, \dots, (s - 1)$  και  $kt - k$  ίσα με  $s$ .

(iii) Για κάθε ένα από τα επίπεδα  $i = 1, 2, \dots, s - 1$  και (κάθε στήλη)  $j = 1, 2, \dots, m$  έχουμε

$$\sum_{u=0}^{st-2} \delta_{i, g_{uk+j}} - t = 0,$$

όπου  $\delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$  και εάν το  $uk + j > m$  αυτό αντικαθίσταται με το  $uk + j - m$ .

Κατά μήκος της γραμμής της εξίσωσης (5.2), ορίζουμε, για τα  $1 \leq j_1 < j_2 \leq m$ ,

$$r = \frac{N}{s^2}, \eta^{(1)} = \lfloor r \rfloor, \text{ και } \eta^{(2)} = r - \eta^{(1)},$$

όπου  $\lfloor r \rfloor$  είναι ο ακέραιος που περιέχεται στο  $r$ . Έστω  $t_1$  και  $t_2$  θετικοί ακέραιοι τέτοιοι ώστε

$$t_1 = s^2 \eta^{(2)} \text{ και } t_2 = s^2(1 - \eta^{(2)}).$$

Βασίζομενοι στο νέο LB, που παρουσιάζεται στο Θεώρημα 1 της ενότητας 5.3, αποδεικνύουμε το ακόλουθο, γενικότερο θεώρημα το οποίο παρέχει μια ικανή συνθήκη υπό την οποία το διάνυσμα του γεννήτορα παράγει έναν  $E(f_{NOD})$ -βέλτιστο, πολλαπλών επιπέδων, σχεδόν-ισορροπημένο,  $k$ -κυκλικό ΥΣ.

**Θεώρημα 2** Έστω  $d \in D(N, s^m)$  ένας ισορροπημένος  $k$ -κυκλικός ΥΣ που παράγεται από τον γεννήτορα  $(g_1, g_2, \dots, g_m)$ . Έστω  $d^* \in D(N^*, s^m)$  ένας σχεδόν-ισορροπημένος

πολλαπλών επιπέδων  $k$ -κυκλικός ΥΣ που λαμβάνεται από τον  $d$  μέσω της λειτουργίας διαγραφής μερικών γραμμών. Η ικανή συνθήκη για να είναι ο σχεδιασμός  $d^*$ ,  $E(f_{NOD})$ -βέλτιστος είναι ο γεννήτορας  $(g_1, g_2, \dots, g_m)$  να επιλεγεί με τέτοιο τρόπο ώστε να ισχύει ένα από τα ακόλουθα:

(i) για  $1 \leq \alpha_1, \alpha_2 \leq s, 1 \leq j_1 < j_2 \leq m,$

$$n_{\alpha_1 \alpha_2}^{j_1 j_2} = \sum_{\theta=0}^{N^*-1} \delta_{\alpha_1 g_{\theta k + j_1}} \delta_{\alpha_2 g_{\theta k + j_2}} = \eta^{(1)} \dot{\eta}(1 + \eta^{(1)}),$$

όπου οποιοδήποτε  $j_1 + \theta k > m$  ( $j_2 + \theta k > m$ ) αντικαθίσταται με  $j_1 + \theta k - m$  ( $j_2 + \theta k - m$ ). Επιπλέον,  $t_2$  από τις πιθανές  $n_{\alpha_1 \alpha_2}^{j_1 j_2}$  τιμές θα πάρουν την τιμή  $\eta^{(1)}$  και οι υπόλοιπες (π.χ.  $t_1$ ) θα πάρουν την τιμή  $(1 + \eta^{(1)})$ , ή,

(ii) οι τιμές των  $c_v$  για  $v = 1, \dots, [(N^* + 1)/2]$  διαφέρουν το πολύ κατά ένα, όπου, για  $v = 1, \dots, [(N^* - 1)/2], 1 \leq i_1 \neq i_2 \leq N^* - 1, |i_1 - i_2| = v$  ή  $|i_1 - i_2| = N^* - 1 - v,$

$$c_{i_1 i_2} = \sum_{j=1}^m \delta_{x_{i_1 j} x_{i_2 j}} = \sum_{j=1}^m \delta_{g_j g_{j+vk}} = c_v,$$

$$c_{[(N^*+1)/2]} = k(t-1),$$

όπου οποιαδήποτε  $j + vk > m$  αντικαθίστανται με  $j + vk - m$ .

**Απόδειξη.** Με τη χρήση των παρεχόμενων συνθηκών είναι εύκολο να αποδειχθεί ότι το  $E(f_{NOD})$  των παραγόμενων ΥΣ θα ικανοποιεί και το κάτω φράγμα (LB) που παρέχεται στο Θεώρημα 1. Ακολουθώντας το Θεώρημα 2, είναι εύκολο να αποδειχθεί η βελτιστότητα του σχεδιασμού που προέκυψε. Πιο συγκεκριμένα, εάν ο αρχικός σχεδιασμός  $d$  είναι  $E(f_{NOD})$ -βέλτιστος, τότε και ο σχεδιασμός  $d^*$  θα είναι επίσης  $E(f_{NOD})$ -βέλτιστος. Με αυτό τον τρόπο είμαστε σε θέση να αποκτήσουμε βέλτιστους σχεδόν ισορροπημένους, πολλών επιπέδων,  $k$ -κυκλικούς ΥΣ. Οι προαναφερθείσες παρατηρήσεις παρουσιάζονται στο Πρόρισμα 8. **Τέλος απόδειξης**

Από το Θεώρημα 2 αντλούμε άμεσα το ακόλουθο πόρισμα.

**Πόρισμα 8** Έστω  $d$  ένας σχεδιασμός που προκύπτει από το Λήμμα 10 βασιζόμενοι σ' έναν δοθέν γεννήτορα  $(g_1, g_2, \dots, g_m)$ . Έστω  $d^*$  ένας σχεδιασμός που προκύπτει από τον  $d$  μέσω της διαγραφής οποιασδήποτε γραμμής εκτός από τη γραμμή όπου όλα τα στοιχεία είναι ίσα με  $s$ . Τότε ο σχεδιασμός  $d^*$  είναι ένας σχεδόν ισορροπημένος πολλαπλών επιπέδων  $k$ -κυκλικός ΥΣ με  $N^* = st - 1, m = N^*k$ . Επιπρόσθετα, για  $1 \leq i \leq s$ , εάν το επίπεδο που αντιστοιχεί στην διεγραμμένη γραμμή και στην  $j$ -οστή στήλη,  $1 \leq j \leq m$ , είναι το  $i$ , τότε σε εκείνη τη στήλη του σχεδιασμού  $d^*$ , όλα τα επίπεδα εκτός από το  $i$  θα εμφανίζονται το ίδιο συχνά ( $t$  φορές) και το επίπεδο  $i$  θα εμφανίζεται  $(t - 1)$  φορές. Αν ο αρχικός σχεδιασμός  $d$  είναι  $E(f_{NOD})$ -βέλτιστος, τότε ο σχεδιασμός  $d^*$  θα είναι επίσης  $E(f_{NOD})$ -βέλτιστος.

**Παράδειγμα 8** Το διάνυσμα του γεννήτορα

$$(5, 4, 4, 2, 2, 5, 2, 5, 3, 1, 1, 4, 3, 3, 1, 4, 1, 3, 4, 2, 1, 3, 4, 2, 1, 3, 2)$$

παράγει έναν ισορροπημένο 5-επιπέδων 3-κυκλικό ΥΣ για  $m = 27$  παράγοντες και  $N = 10$  πειραματικές εκτελέσεις (που παρουσιάζεται στον Πίνακα 5.1). Σύμφωνα με το Πόρισμα 8, μετά τη διαγραφή της πρώτης γραμμής αποκτούμε έναν  $E(f_{NOD})$ -βέλτιστο, σχεδόν ισορροπημένο, 5-επιπέδων, 3-κυκλικό ΥΣ, με  $m = 27$  παράγοντες και  $N^* = 9$  πειραματικές εκτελέσεις. Ο σχεδιασμός παρουσιάζεται στον Πίνακα 5.3.

Λόγω του γεγονότος ότι ο αρχικός σχεδιασμός,  $d$ , (που παρουσιάζεται στον Πίνακα

**Πίνακας 5.3:** Ένας σχεδόν ισορροπημένος 5-επιπέδων 3-κυκλικός ΥΣ με  $m = 27$  παράγοντες και  $N^* = 9$  πειραματικές εκτελέσεις που κατασκευάστηκε από το Πόρισμα 8 (διαγράφοντας την πρώτη γραμμή).

1	3	2	5	4	4	2	2	5	2	5	3	1	1	4	3	3	1	4	1	3	4	2	1	3	4	2
3	4	2	1	3	2	5	4	4	2	2	5	2	5	3	1	1	4	3	3	1	4	1	3	4	2	1
4	2	1	3	4	2	1	3	2	5	4	4	2	2	5	2	5	3	1	1	4	3	3	1	4	1	3
4	1	3	4	2	1	3	4	2	1	3	2	5	4	4	2	2	5	2	5	3	1	1	4	3	3	1
3	3	1	4	1	3	4	2	1	3	4	2	1	3	2	5	4	4	2	2	5	2	5	3	1	1	4
1	1	4	3	3	1	4	1	3	4	2	1	3	4	2	1	3	2	5	4	4	2	2	5	2	5	3
2	5	3	1	1	4	3	3	1	4	1	3	4	2	1	3	4	2	1	3	2	5	4	4	2	2	5
2	2	5	2	5	3	1	1	4	3	3	1	4	1	3	4	2	1	3	4	2	1	3	2	5	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5

5.1) είναι  $E(f_{NOD})$ -βέλτιστος, συμπεραίνουμε ότι ο σχεδιασμός  $d^*$  (που παρουσιάζεται

στον Πίνακα 5.3) είναι ένας  $E(f_{NOD})$ -βέλτιστος σχεδιασμός αφού κατασκευάστηκε σύμφωνα με το Πόρισμα 8.

Ακολουθώντας την απόδειξη του Θεωρήματος 1 που αποδείχθηκε από τους Georgίου και Koukouninos (2006), παρέχουμε μία άλλη κατασκευή, που παρουσιάζεται στο Θεώρημα 3, για σχεδόν ισορροπημένους, πολλαπλών επιπέδων,  $k$ -κυκλικούς ΥΣ.

**Θεώρημα 3** Έστω  $d$  ένας σχεδιασμός που λαμβάνεται μέσω του Λήμματος 10 και βασίζεται σ' έναν δοθέν γεννήτορα  $(g_1, g_2, \dots, g_m)$ . Έστω  $d^*$  ένας σχεδιασμός που λαμβάνεται από τον  $d$  σύμφωνα με τον ακόλουθο κανόνα:

(i) διαγράφοντας οποιοσδήποτε δύο γραμμές εκτός από τη γραμμή που όλα τα στοιχεία είναι ίσα με  $s$ .

(ii) διαγράφοντας όλες αυτές τις στήλες, όπου και οι δύο έχουν τα ίδια επίπεδα στις σειρές που διαγράφηκαν.

Ο προκύπτων σχεδιασμός  $d^*$  είναι ένας σχεδόν ισορροπημένος πολλαπλών-επιπέδων  $k$ -κυκλικός ΥΣ με  $N^* = st - 2, m^* = (N^* + 1)k - f$ , όπου  $f$  είναι ο αριθμός των στηλών που διαγράφηκαν. Επιπρόσθετα, για  $1 \leq i_1 < i_2 \leq s$ , εάν τα επίπεδα που αντιστοιχούν στις γραμμές που διαγράφηκαν και στην  $j$ -οστή στήλη, είναι  $i_1$  και  $i_2$  αντίστοιχα, τότε σ' αυτή τη στήλη του σχεδιασμού  $d^*$ , όλα τα επίπεδα εκτός από τα  $i_1$  και  $i_2$  θα εμφανίζονται τον ίδιο αριθμό φορές ( $t$  φορές) και τα επίπεδα  $i_1$  και  $i_2$  θα εμφανίζονται  $t - 1$  φορές το καθένα.

**Απόδειξη.** Η απόδειξη αυτού του θεωρήματος προκύπτει από την απόδειξη του Θεωρήματος 1 των Georgίου και Koukouninos [55]. **Τέλος απόδειξης**

Οι ικανές συνθήκες για έναν σχεδιασμό που προέκυψε από το Θεώρημα 3 για να είναι  $E(f_{NOD})$ -βέλτιστος παρέχονται στο Θεώρημα 4.

**Θεώρημα 4** Έστω  $d^* \in D(N^*, s^{m^*})$  ένας σχεδόν ισορροπημένος πολλαπλών επιπέδων  $k$ -κυκλικός ΥΣ που προέκυψε όπως εκείνος από το Θεώρημα 3. Η ικανή συνθήκη για να είναι ο σχεδιασμός  $d^*$   $E(f_{NOD})$ -βέλτιστος είναι ο γεννήτορας  $(g_1, g_2, \dots, g_{m+f})$ , όπου  $f$  ορίζεται όπως στο Θεώρημα 3, πρέπει να επιλεγεί με τέτοιο τρόπο ώστε να ισχύει ένα από τα ακόλουθα.

(i) για  $1 \leq \alpha_1, \alpha_2 \leq s, 1 \leq j_1 < j_2 \leq m^*$ ,

$$n_{\alpha_1 \alpha_2}^{j_1 j_2} = \eta^{(1)} \dot{\eta} (1 + \eta^{(1)}),$$

όπου  $\eta^{(1)} = [N^*/s^2]$ . Επιπρόσθετα,  $t_2$  από τις πιθανές  $n_{\alpha_1 \alpha_2}^{j_1 j_2}$  τιμές θα λαμβάνουν την τιμή  $\eta^{(1)}$  και οι υπόλοιπες (π.χ.  $t_1$ ) θα λαμβάνουν την τιμή  $(1 + \eta^{(1)})$ , ή,

(ii) για  $1 \leq i_1 \neq i_2 \leq N^*$ ,

$$c_{i_1 i_2} = \sum_{j=1}^{m^*} \delta_{x_{i_1 j} x_{i_2 j}} = \zeta^{(1)} \dot{\zeta} (1 + \zeta^{(1)}).$$

Εδώ,  $p_2$  από τις πιθανές  $c_{i_1 i_2}$  τιμές θα λαμβάνουν την τιμή  $\zeta^{(1)}$  και οι υπόλοιπες (π.χ.  $p_1$ ) θα λαμβάνουν την τιμή  $(1 + \zeta^{(1)})$ . Επιπρόσθετα,  $p_1$  και  $p_2$  είναι θετικοί ακέραιοι τέτοιοι ώστε

$$\left. \begin{aligned} p_2 \zeta^{(1)} + p_1 (1 + \zeta^{(1)}) &= m^*(t-1)(st-4) \\ p_1 + p_2 &= N^*(N^* - 1). \end{aligned} \right\}$$

**Απόδειξη.** Για τον σχεδιασμό  $d^*$ , έχουμε

$$\sum_{1 \leq i_1 \neq i_2 \leq N^*} c_{i_1 i_2} = m^*(t-1)(st-4) = c.$$

Ορίζουμε

$$q = \frac{c}{N^*(N^* - 1)}, \quad \zeta^{(1)} = [q], \quad \text{και} \quad \zeta^{(2)} = q - \zeta^{(1)}.$$

Έστω  $p_1$  και  $p_2$  θετικοί ακέραιοι τέτοιοι ώστε

$$\left. \begin{aligned} p_2 \zeta^{(1)} + p_1 (1 + \zeta^{(1)}) &= c \\ p_1 + p_2 &= N^*(N^* - 1). \end{aligned} \right\}$$

Θεωρώντας τις συνθήκες που παρήχθησαν στο Θεώρημα 4, μετά από αλγεβρικές πράξεις, παρατηρούμε ότι η τιμή του  $E(f_{NOD})$  συμπίπτει με το κάτω όριο (LB) που παρέχεται στο Θεώρημα 1. Επομένως ο προκύπτων σχεδιασμός είναι  $E(f_{NOD})$ -βέλτιστος. **Τέλος απόδειξης**

**Παράδειγμα 9** Σε αυτό το παράδειγμα δείχνουμε πως να κατασκευάσουμε έναν σχεδόν ισορροπημένο  $E(f_{NOD})$ -βέλτιστο 5-επιπέδων 3-κυκλικό ΥΣ με 8 πειραματικές εκτελέσεις και 24 παράγοντες. Το διάνυσμα του γεννήτορα

$$(5, 4, 4, 2, 2, 5, 2, 5, 3, 1, 1, 4, 3, 3, 1, 4, 1, 3, 4, 2, 1, 3, 4, 2, 1, 3, 2)$$

παράγει έναν ισορροπημένο 5-επιπέδων 3-κυκλικό ΥΣ για  $m = 27$  παράγοντες και  $N = 10$  πειραματικές εκτελέσεις που παρουσιάζεται στον Πίνακα 5.1. Από το Θεώρημα 3, διαγράφουμε τις πρώτες δύο γραμμές και τις τρεις σχετικές στήλες (όπου και οι δύο έχουν τα ίδια επίπεδα στις δύο γραμμές που διαγράφηκαν) οι οποίες είναι οι στήλες 7, 19 και 27.

**Πίνακας 5.4:** Ένας σχεδόν-ισορροπημένος 5-επιπέδων 3-κυκλικός υπερκορεσμένος σχεδιασμός με  $m = 24$  παράγοντες και  $N = 8$  πειραματικές εκτελέσεις που κατασκευάστηκε από το Θεώρημα 3 (διαγράφοντας τις πρώτες δύο γραμμές και τρεις στήλες). Αυτός ο σχεδιασμός είναι βέλτιστος (Θεώρημα 4).

3	4	2	1	3	2	4	4	2	2	5	2	5	3	1	1	4	3	1	4	1	3	4	2
4	2	1	3	4	2	3	2	5	4	4	2	2	5	2	5	3	1	4	3	3	1	4	1
4	1	3	4	2	1	4	2	1	3	2	5	4	4	2	2	5	5	3	1	1	4	3	3
3	3	1	4	1	3	2	1	3	4	2	1	3	2	5	4	4	2	5	2	5	3	1	1
1	1	4	3	3	1	1	3	4	2	1	3	4	2	1	3	2	4	4	2	2	5	2	5
2	5	3	1	1	4	3	1	4	1	3	4	2	1	3	4	2	3	2	5	4	4	2	2
2	2	5	2	5	3	1	4	3	3	1	4	1	3	4	2	1	4	2	1	3	2	5	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5

Σύμφωνα με το Θεώρημα 4, παρατηρούμε ότι ο παραγόμενος σχεδιασμός είναι ένας  $E(f_{NOD})$ -βέλτιστος σχεδόν-ισορροπημένος 5-επιπέδων 3-κυκλικός ΥΣ με  $m^* = 24$  παράγοντες και  $N^* = 8$  πειραματικές εκτελέσεις, αφού ικανοποιείται η συνθήκη (ii) του Θεωρήματος 4.

Πιο συγκεκριμένα, βασιζόμενοι στον Πίνακα 5.5, έχουμε  $p_2 = 24$  ζεύγη έχοντας  $c_{i_1, i_2} = 2 = \zeta^{(1)}$  και  $p_1 = 32$  ζεύγη γραμμών έχοντας  $c_{i_1, i_2} = 3 = \zeta^{(1)} + 1$ . Έχουμε ότι  $N^* = st - 2 \Rightarrow t = 2$ ,  $p_1 + p_2 = 56$  και  $N^*(N^* - 1) = 56$ . Έτσι,  $p_1 + p_2 = N^*(N^* - 1)$ . Επιπρόσθετα,  $p_2\zeta^{(1)} + p_1(\zeta^{(1)} + 1) = 144$  και  $m^*(t-1)(st-4) = 144$ . Έτσι,  $p_2\zeta^{(1)} + p_1(\zeta^{(1)} + 1) = m^*(t-1)(st-4)$ . Οι προαναφερθείσες εξισώσεις δικαιολογούν ότι η συνθήκη (ii) του Θεωρήματος 4 ικανοποιείται, με αποτέλεσμα έναν  $E(f_{NOD})$ -βέλτιστο, σχεδόν ισορροπημένο, σχεδιασμό ο οποίος παρουσιάζεται στον Πίνακα 5.4.

**Πίνακας 5.5:** Coincidences ( $c_{i_1 i_2}$ ) μεταξύ όλων των ζευγών των επιπέδων της  $i_1$ -οστής και της  $i_2$ -οστής γραμμών του σχεδιασμού που παρουσιάζονται στον Πίνακα 5.4.

$i_1$	1	1	1	1	1	1	1	2	2	2	2	2	2	3
$i_2$	2	3	4	5	6	7	8	3	4	5	6	7	8	4
$c_{i_1 i_2}$	3	2	3	3	3	2	2	3	3	2	2	2	3	2
$i_1$	3	3	3	3	4	4	4	4	5	5	5	6	6	7
$i_2$	5	6	7	8	5	6	7	8	6	7	8	7	8	8
$c_{i_1 i_2}$	3	2	3	3	2	3	2	3	3	3	2	3	2	3

## 5.5 Κατασκευή βέλτιστων μικτών-επιπέδων $k$ -κυκλικών ΥΣ

Σε αυτή την ενότητα παρέχουμε τις απαραίτητες προϋποθέσεις (Θεώρημα 6) για τον γεννήτορα έτσι ώστε να κατασκευαστεί ένας  $E - (f_{NOD})$ -βέλτιστος, μικτών επιπέδων, σχεδόν ισορροπημένος,  $k$ -κυκλικός ΥΣ. Το Λήμμα 11 των Chen και Liu [25], παρέχει τις αναγκαίες και ικανές συνθήκες για την ύπαρξη μικτών-επιπέδων,  $k$ -κυκλικών ισορροπημένων ΥΣ. Στο Πρόρισμα 9 και στο Θεώρημα 5 παρέχουμε δύο μεθόδους κατασκευής μικτών-επιπέδων, σχεδόν ισορροπημένων,  $k$ -κυκλικών ΥΣ (Πρόρισμα 9, Θεώρημα 5) οι οποίοι είναι  $E(f_{NOD})$ -βέλτιστοι, εάν ο αρχικός σχεδιασμός,  $d$ , είναι επίσης  $E(f_{NOD})$ -βέλτιστος.

**Λήμμα 11** (Chen και Liu [25])

Έστω  $d \in D(N, s_1 \cdots s_m)$  ένας μικτών-επιπέδων  $k$ -κυκλικός ΥΣ, όπου  $m = qk$ , για έναν θετικό ακέραιο  $q$ . Θεωρούμε τον γεννήτορα  $(g_1, g_2, \dots, g_m)$ , για κάποιους θετικούς ακέραιους  $g_j, g_{j+k}, \dots, g_{j+(q-1)k}$  να παίρνουν τιμές από το σύνολο  $s_j$  συμβόλων  $\{1, 2, \dots, s_j\}, j = 1, 2, \dots, k$ . Υποθέτουμε ότι ο σχεδιασμός  $d^*$  δημιουργήθηκε από τον γεννήτορα  $(g_1, g_2, \dots, g_m)$  με την κυκλική εναλλαγή  $k$  στοιχείων προς τα δεξιά σε κάθε βήμα και προσθέτοντας μία γραμμή με όλα τα στοιχεία να είναι ίσα με 1. Οι αναγκαίες και ικανές συνθήκες για να είναι ο  $d$  ένας μικτών-επιπέδων ισορροπημένος  $k$ -κυκλικός ΥΣ, είναι:

(i)  $N = q + 1, N = s_j t_j$ , για κάποιους θετικούς ακέραιους  $t_j, j = 1, 2, \dots, k$ .

(ii) για κάθε  $j = 1, 2, \dots, k$ , και για κάθε ένα από τα επίπεδα  $l_j = 1, 2, \dots, s_j$ ,

$$\sum_{u=0}^{q-1} \delta_{l_j, g_{j+uk}} = t_j - \delta_{l_j, 1}.$$

Βασιζόμενοι στο Θεώρημα 1 μπορούμε να επιτύχουμε, παρόμοια με το Θεώρημα 2, ικανές συνθήκες υπό τις οποίες ο γεννήτορας παράγει έναν σχεδόν ισορροπημένο, μικτών-επιπέδων,  $k$ -κυκλικό ΥΣ. Μέσω του Πορίσματος 9 και του Θεωρήματος 5, παρέχουμε τον τρόπο κατασκευής σχεδόν ισορροπημένων μικτών-επιπέδων  $k$ -κυκλικών ΥΣ. Η τιμή  $m = qk$  αφορά στον αριθμό των παραγόντων ενώ οι θετικοί ακέραιοι  $q, g_j, g_{j+k}, \dots, g_{j+(q-1)k}$  παίρνουν τιμές από το σύνολο  $s_j$  συμβόλων  $\{1, 2, \dots, s_j\}, j = 1, 2, \dots, k$ .

Όμοια με το Θεώρημα 2, μπορούμε να αποδείξουμε το ακόλουθο Πρόγραμμα.

**Πόρισμα 9** Έστω  $d$  ένας σχεδιασμός που λαμβάνεται μέσω του Λήμματος 11 και βασίζεται σ' ένα δοθέν γεννήτορα  $(g_1, g_2, \dots, g_m)$ . Έστω  $d^*$  ένας σχεδιασμός που λαμβάνεται από τον  $d$  διαγράφοντας μία οποιαδήποτε γραμμή. Τότε ο  $d^*$  είναι ένας σχεδόν ισορροπημένος μικτών-επιπέδων  $k$ -κυκλικός ΥΣ με  $N^* = q, N^* = s_j t_j - 1, m = qk$ . Επιπρόσθετα, για  $j = 1, 2, \dots, k, u = 1, \dots, q$ , εάν τα σύμβολα της γραμμής που διαγράφηκε και αντιστοιχούν στη  $((u - 1)k + j)$ -οστή στήλη είναι  $l_j$ , τότε τα σύμβολα  $1, \dots, s_j$ , εκτός από το  $l_j$ , εμφανίζονται τον ίδιο αριθμό φορές  $((t_j)$  φορές) ενώ το σύμβολο  $l_j$  εμφανίζεται  $(t_j - 1)$  φορές. Αν ο αρχικός σχεδιασμός  $d$  είναι  $E(f_{NOD})$ -βέλτιστος, τότε οι σχεδιασμοί που προτείνονται στο Πρόγραμμα 9 θα είναι επίσης  $E(f_{NOD})$ -βέλτιστοι.

**Παράδειγμα 10** Σε αυτό το παράδειγμα παρουσιάζουμε έναν σχεδόν ισορροπημένο μικτών επιπέδων 2-κυκλικό ΥΣ  $d^* = D(11; 2^{11}6^{11})$  με  $N^* = 11$  πειραματικές εκτελέσεις και  $m = 22$  παράγοντες που κατασκευάστηκε από το Πρόγραμμα 9.

Ο πίνακας σχεδιασμού του  $d^*$  (βλέπε Πίνακα 5.6) παράγεται από τον αρχικό σχεδιασμό  $d$  που παρουσιάζεται στον Πίνακα 5.2, ο γεννήτορας του οποίου είναι

$$(2, 5, 1, 3, 2, 5, 1, 2, 1, 4, 1, 1, 2, 2, 2, 3, 2, 4, 1, 6, 2, 6).$$

Εργαζόμενοι με παρόμοιο τρόπο, όπως στα Θεωρήματα 3 και 4, για πολλαπλών επιπέδων σχεδιασμούς, μπορούμε να κατασκευάσουμε μικτών επιπέδων, σχεδόν ισορροπημένους,  $k$ -κυκλικούς ΥΣ (Θεώρημα 5) και ικανές συνθήκες (Θεώρημα 6) κάτω από τις οποίες το διάνυσμα του γεννήτορα παράγει έναν  $E(f_{NOD})$ -βέλτιστο σχεδόν ισορροπημένο, μικτών επιπέδων,  $k$ -κυκλικό ΥΣ. Οι αποδείξεις των Θεωρημάτων 5 και 6 παραλείπονται.



**Πίνακας 5.6:** Ένας σχεδόν ισορροπημένος μικτών επιπέδων 2-κυκλικός ΥΣ με  $m = 22$  παράγοντες και  $N = 11$  πειραματικές εκτελέσεις. Αυτός ο σχεδιασμός είναι βέλτιστος.

2	5	1	3	2	5	1	2	1	4	1	1	2	2	2	3	2	4	1	6	2	6
2	6	2	5	1	3	2	5	1	2	1	4	1	1	2	2	2	3	2	4	1	6
1	6	2	6	2	5	1	3	2	5	1	2	1	4	1	1	2	2	2	3	2	4
2	4	1	6	2	6	2	5	1	3	2	5	1	2	1	4	1	1	2	2	2	3
2	3	2	4	1	6	2	6	2	5	1	3	2	5	1	2	1	4	1	1	2	2
2	2	2	3	2	4	1	6	2	6	2	5	1	3	2	5	1	2	1	4	1	1
1	1	2	2	2	3	2	4	1	6	2	6	2	5	1	3	2	5	1	2	1	4
1	4	1	1	2	2	2	3	2	4	1	6	2	6	2	5	1	3	2	5	1	2
1	2	1	4	1	1	2	2	2	3	2	4	1	6	2	6	2	5	1	3	2	5
2	5	1	2	1	4	1	1	2	2	2	3	2	4	1	6	2	6	2	5	1	3
1	3	2	5	1	2	1	4	1	1	2	2	2	3	2	4	1	6	2	6	2	5

**Θεώρημα 5** Έστω  $d$  ένας σχεδιασμός που λαμβάνεται μέσω του Λήμματος 11 και βασίζεται σ' έναν δοθέν γεννήτορα  $(g_1, g_2, \dots, g_m)$ . Έστω  $d^*$  ένας σχεδιασμός που λαμβάνεται από τον  $d$  σύμφωνα με τον ακόλουθο κανόνα:

- (i) διαγράφοντας οποιοσδήποτε δύο γραμμές.
- (ii) διαγράφοντας όλες αυτές τις στήλες, όπου έχουμε τα ίδια επίπεδα στις γραμμές που διαγράφηκαν.

Ο προκύπτων σχεδιασμός  $d^*$  είναι ένας σχεδόν ισορροπημένος μικτών-επιπέδων  $k$ -κυκλικός ΥΣ με  $N^* = q - 1$ ,  $N^* = s_j t_j - 2$ ,  $m^* = qk - f$ , όπου  $f$  είναι ο αριθμός των στηλών που διαγράφηκαν. Επιπρόσθετα, για  $1 \leq j \leq m$ , εάν η  $j$ -οστή στήλη είναι  $1 \leq l_{j1} < l_{j2} \leq s_j$ , τότε σ' αυτή τη στήλη του σχεδιασμού  $d^*$ , όλα τα επίπεδα εκτός από τα  $l_{j1}$  και  $l_{j2}$  θα εμφανίζονται τον ίδιο αριθμό φορών ( $t_j$  φορές) και τα επίπεδα  $l_{j1}$  και  $l_{j2}$  θα εμφανίζονται  $t_j - 1$  φορές.

**Παράδειγμα 11** Σε αυτό το παράδειγμα παρουσιάζουμε έναν σχεδόν ισορροπημένο μικτών επιπέδων 2-κυκλικό ΥΣ  $d^* = D(10; 2^6 6^{10})$ , με  $N^* = 10$  πειραματικές εκτελέσεις και  $m^* = 16$  παράγοντες, που κατασκευάστηκε από το Θεώρημα 5. Ο αρχικός σχεδιασμός παρουσιάζεται στον Πίνακα 5.2 και κατασκευάστηκε από τον γεννήτορα

$$(2, 5, 1, 3, 2, 5, 1, 2, 1, 4, 1, 1, 2, 2, 2, 3, 2, 4, 1, 6, 2, 6).$$

Σύμφωνα με το Θεώρημα 5, αφού αφαιρέσουμε τις πρώτες δύο γραμμές, αποκτούμε τον σχεδόν ισορροπημένο ΥΣ του Πίνακα 5.7. Η αποδοτικότητα του απεικονιζόμενου σχεδιασμού (στον Πίνακα 5.7) ισούται με 0.9451 αφού η τιμή του κριτηρίου  $E(f_{NOD})$ ,

5.4667, δεν επιτυγχάνει την τιμή 5.1667 του κάτω φράγματος. Οι παραπάνω τιμές υπολογίστηκαν χρησιμοποιώντας τους μαθηματικούς τύπους που αποδείξαμε και το λογισμικό του προγράμματος matlab.

**Πίνακας 5.7:** Ένας σχεδόν ισορροπημένος μικτών επιπέδων 2-κυκλικός ΥΣ με  $m^* = 16$  παράγοντες και  $N^* = 10$  πειραματικές εκτελέσεις.

6	2	6	2	5	1	3	5	2	1	4	1	2	2	3	2
4	1	6	2	6	2	5	3	5	1	2	4	1	2	2	2
3	2	4	1	6	2	6	5	3	2	5	2	4	1	1	2
2	2	3	2	4	1	6	6	5	1	3	5	2	1	4	1
1	2	2	2	3	2	4	6	6	2	5	3	5	1	2	1
4	1	1	2	2	2	3	4	6	2	6	5	3	2	5	1
2	1	4	1	1	2	2	3	4	1	6	6	5	1	3	2
5	1	2	1	4	1	1	2	3	2	4	6	6	2	5	1
3	2	5	1	2	1	4	1	2	2	3	4	6	2	6	2
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Το ακόλουθο Θεώρημα παρέχει ικανές συνθήκες για έναν σχεδιασμό που επιτυγχάνεται από το Θεώρημα 5 για να είναι  $E(f_{NOD})$ -βέλτιστος.

**Θεώρημα 6** Έστω  $d^* \in D(N^*, s_1 \cdots s_m)$  ένας σχεδόν ισορροπημένος μικτών-επιπέδων  $k$ -κυκλικός ΥΣ που επιτυγχάνεται όπως στο Θεώρημα 5 με  $N^* = s_j t_j - 2, 1 \leq j \leq m$ . Ικανή συνθήκη για να είναι ο σχεδιασμός  $d^*$   $E(f_{NOD})$ -βέλτιστος είναι ο γεννήτορας  $(g_1, g_2, \dots, g_{m^*+f})$ , όπου  $f$  είναι όπως περιγράφεται στο Θεώρημα 5, να επιλεγεί με τέτοιο τρόπο ώστε να ισχύει κάποιο από τα ακόλουθα.

(i) Για  $1 \leq \alpha_1, \alpha_2 \leq s, 1 \leq j_1 < j_2 \leq m^*$ ,

$$n_{\alpha_1 \alpha_2}^{j_1 j_2} = \eta_{j_1 j_2}^{(1)} \quad \eta \left( 1 + \eta_{j_1 j_2}^{(1)} \right),$$

Εδώ,  $t_2$  από τις πιθανές  $n_{\alpha_1 \alpha_2}^{j_1 j_2}$  τιμές θα παίρνουν την τιμή  $\eta_{j_1 j_2}^{(1)}$  και οι υπόλοιπες (π.χ.  $t_1$ ) θα λαμβάνουν την τιμή  $(1 + \eta_{j_1 j_2}^{(1)})$  όπου  $\eta_{j_1 j_2}^{(1)}, t_1$  και  $t_2$  ορίζονται όπως στις εξισώσεις (5.2)-(10.5), ή,

(ii) Για  $1 \leq i_1 \neq i_2 \leq N^*$ ,

$$c_{i_1 i_2} = \sum_{j=1}^m \delta_{x_{i_1 j} x_{i_2 j}} = \zeta^{(1)} \quad \eta \left( 1 + \zeta^{(1)} \right).$$

Εδώ,  $p_2$  από τις πιθανές  $c_{i_1 i_2}$  τιμές θα πάρουν την τιμή  $\zeta^{(1)}$  και οι υπόλοιπες (π.χ.  $p_1$ )

θα πάρουν την τιμή  $(1 + \zeta^{(1)})$ . Επιπρόσθετα,  $p_1$  και  $p_2$  είναι θετικοί ακέραιοι τέτοιοι ώστε

$$\begin{aligned} p_2 \zeta^{(1)} + p_1 (1 + \zeta^{(1)}) &= \sum_{j=1}^{m^*} (t_j - 1)(s_j t_j - 4) \\ p_1 + p_2 &= N^*(N^* - 1). \end{aligned}$$

**Απόδειξη.** Για τον σχεδιασμό  $d^*$ , έχουμε

$$\sum_{1 \leq i_1 \neq i_2 \leq N^*} c_{i_1 i_2} = \sum_{j=1}^{m^*} (t_j - 1)(s_j t_j - 4) = c.$$

Ορίζουμε

$$q = \frac{c}{N^*(N^* - 1)}, \quad \zeta^{(1)} = [q], \quad \text{και} \quad \zeta^{(2)} = q - \zeta^{(1)}.$$

Έστω  $p_1$  και  $p_2$  θετικοί ακέραιοι τέτοιοι ώστε

$$\left. \begin{aligned} p_2 \zeta^{(1)} + p_1 (1 + \zeta^{(1)}) &= c \\ p_1 + p_2 &= N^*(N^* - 1). \end{aligned} \right\}$$

Τότε, χρησιμοποιώντας τις παρεχόμενες συνθήκες και εύκολους υπολογισμούς, μπορούμε να δούμε ότι η τιμή του  $E(f_{NOD})$  είναι η ίδια με το LB που παρέχεται στο Θεώρημα 1 και έτσι ο σχεδιασμός είναι  $E(f_{NOD})$ -βέλτιστος.

## 5.6 Συμπεράσματα

Οι ΥΣ καθίστανται μια ελκυστική τεχνική πειραμάτων κρησαρίσματος τόσο στη θεωρία όσο και στην πράξη. Οι περισσότερες μελέτες για ΥΣ επικεντρώθηκαν κυρίως στην περίπτωση ισορροπημένων σχεδιασμών. Ωστόσο, σε πραγματικά προβλήματα ζωής η κατάσταση που ο αριθμός των πειραματικών εκτελέσεων δεν είναι πολλαπλάσιο του  $N$  μπορεί να είναι πολύ ενδιαφέρουσα.

Στην ενότητα αυτή παρουσιάζονται νέες κλάσεις βέλτιστων ισορροπημένων και σχεδόν ισορροπημένων ΥΣ που κατασκευάστηκαν μέσω κατάλληλης τροποποίησης και εφαρμογής της  $k$ -κυκλικής μεθόδου στην περίπτωση πολλαπλών και μικτών επιπέδων σχεδιασμών. Παρέχονται επίσης νέα κάτω φράγματα για την βελτιστότητα

ως προς το  $E(f_{NOD})$ -κριτήριο, τα οποία ισχύουν τόσο σε ισορροπημένους όσο και σε μη ισορροπημένους σχεδιασμούς. Επιπλέον, παρέχονται οι αναγκαίες και ικανές συνθήκες για τον γεννήτορα έτσι ώστε να επιτυγχάνεται αυτό το κάτω φράγμα για τους μη ισορροπημένους σχεδιασμούς εξασφαλίζοντας ότι οι σχεδιασμοί αυτοί είναι  $E(f_{NOD})$ -βέλτιστοι. Είναι σημαντικό να αναφερθεί ότι η παρούσα μεθοδολογία μπορεί να εφαρμοστεί σε οποιονδήποτε ΥΣ, ανεξάρτητα αν είναι ισορροπημένος ή μη, πολλαπλών ή μικτών επιπέδων.

Παράλληλα, σε αυτό το κεφάλαιο παρουσιάζουμε μια σειρά από κατασκευές  $E(f_{NOD})$ -βέλτιστων σχεδόν ισορροπημένων  $k$ -κυκλικών πολλαπλών επιπέδων και μικτών επιπέδων ΥΣ. Ο Πίνακας 5 που παρουσιάζεται στο παράρτημα Α, περιέχει τους  $E(f_{NOD})$ -βέλτιστους  $k$ -κυκλικούς πολλαπλών επιπέδων ΥΣ και τους αντίστοιχους σχεδόν ισορροπημένους  $k$ -κυκλικούς πολλαπλών επιπέδων ΥΣ που κατασκευάστηκαν με την τροποποιημένη  $k$ -κυκλική μέθοδο που περιγράφηκε σε προηγούμενο κεφάλαιο (Πόρισμα 8 και Θεώρημα 3). Οι περισσότεροι από τους σχεδιασμούς που έχουν παρουσιαστεί πληρούν τις απαραίτητες προϋποθέσεις (Θεώρημα 4). Ακόμη και αν ορισμένοι από τους σχεδιασμούς που παρουσιάστηκαν δεν ικανοποιούν αυτές τις συνθήκες, έτσι ώστε να είναι βέλτιστοι, επιτυγχάνουν πραγματικά υψηλές τιμές όσον αφορά την αποδοτικότητά τους.

Με τον ίδιο τρόπο παρουσιάσαμε έναν μεγάλο αριθμό  $E(f_{NOD})$ -βέλτιστων  $k$ -κυκλικών ΥΣ μικτών επιπέδων και τους αντίστοιχους σχεδόν ισορροπημένους  $k$ -κυκλικούς ΥΣ μικτών επιπέδων. Αν και είναι πολύ δύσκολο να βρεθεί ο βέλτιστος σχεδιασμός, ιδίως στην περίπτωση των ΥΣ μικτών επιπέδων, κατασκευάσαμε και παρουσιάζουμε στο παρόν κεφάλαιο νέες κλάσεις βέλτιστων και σχεδόν βέλτιστων σχεδιασμών. Ο Πίνακας 6 επιβεβαιώνει όσα προαναφέρθηκαν, παρέχοντας έναν μεγάλο αριθμό σχεδιασμών με μικτά επίπεδα οι οποίοι κατασκευάστηκαν με βάση το Πόρισμα 9 και το Θεώρημα 5.

---

## Ανάλυση υπερκορεσμένων σχεδιασμών με μία νέα μέθοδο που βασίζεται στις SVM

---

**Σ**το παρόν κεφάλαιο παρουσιάζουμε μία μέθοδο ανάλυσης υπερκορεσμένων σχεδιασμών δύο, πολλαπλών και μικτών-επιπέδων. Η ανάπτυξη νέων στατιστικών μεθόδων που είναι εμπνευσμένες από τους αλγόριθμους μηχανικής μάθησης αυξάνεται ταχύτατα, ιδιαίτερα στην εποχή μας. Μία από αυτές τις μεθόδους βασίζεται στις μηχανές διανυσμάτων υποστήριξης (SVM-RFE) οι οποίες καταφέρνουν να εξαγάγουν με μεγάλη επιτυχία τα σημαντικά γονίδια σε προβλήματα ταξινόμησης, επιτυγχάνοντας εξαιρετικά υψηλή απόδοση. Σε αυτό το κεφάλαιο, λοιπόν, προτείνουμε μία νέα μέθοδο επιλογής μεταβλητών σε προβλήματα παλινδρόμησης, που την ονομάζουμε SVR-RFE, για την ανίχνευση ενεργών επιδράσεων τόσο σε σχεδιασμούς δύο επιπέδων όσο και σε σχεδιασμούς μικτών επιπέδων. Οι μελέτες προσομοίωσης που παρουσιάζουμε δείχνουν ότι αυτή η διαδικασία είναι αρκετά αποτελεσματική, ειδικά όσον αφορά τη στατιστική ισχύ. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [V].

### 6.1 Ερευνητικό πρόβλημα-Λύση

Η επιλογή των χαρακτηριστικών έχει εξελιχθεί σε ένα σημαντικό εργαλείο προεπεξεργασίας των δεδομένων για την ανάλυση ΥΣ. Έχουν προταθεί κατά καιρούς πολλοί αλγόριθμοι επιλογής χαρακτηριστικών για την εξάλειψη των μη σημαντικών χαρακτηριστικών. Αυτό είναι ένα σημαντικό στάδιο αφού η αφαίρεση των μη σημαντικών χαρακτηριστικών βελτιώνει την απόδοση των μοντέλων μάθησης ενώ παρέχει ταχύτερους, αλλά και οικονομικότερους παράγοντες πρόβλεψης. Οι μηχανές δια-

νυσμάτων υποστήριξης (SVM) που εισήγαγαν οι Cortes και Vapnik [28], αποτελούν έναν από τους πλέον ισχυρούς αλγόριθμους ταξινόμησης. Ωστόσο, οι SVM έχουν εφαρμοστεί με επιτυχία όχι μόνο στην ταξινόμηση αλλά και στις διαδικασίες επιλογής χαρακτηριστικών. Ένας από αυτούς τους αλγόριθμους είναι ο λεγόμενος SVM-RFE ο οποίος αποτελεί μία μέθοδο περιτυλίγματος που συνδυάζει το SVM με τη μέθοδο εξάλειψης επαναλαμβανόμενων χαρακτηριστικών (Support Vector Machines - Recursive Feature Elimination, SVM-RFE) [38, 61, 105, 134]. Η προαναφερθείσα μέθοδος είναι αρκετά δημοφιλής λόγω της αποτελεσματικότητάς της στην επιλογή γονιδίων για την ταξινόμηση δεδομένων του καρκίνου μέσω της ταξινόμησης όλων των χαρακτηριστικών σύμφωνα με μία συνάρτηση σκορ, εξαλείφοντας, σε κάθε βήμα, το χαρακτηριστικό με τη χαμηλότερη βαθμολογία. Έως τώρα έχουν προταθεί πολλές τροποποιήσεις και βελτιώσεις έχοντας ως θεμελιώδη έννοια το RFE, λόγω της αποδοτικότητάς του [49, 65, 96, 117, 152]. Ωστόσο, οι αλγόριθμοι επιλογής χαρακτηριστικών χρησιμοποιούνται συνήθως στην ταξινόμηση και όχι στην παλινδρόμηση. Μόνο οι Fang και Tai [49] αξιολόγησαν κάποιους τελευταίας τεχνολογίας αλγόριθμους σε μια ποσοτική σχέση δομής-δραστηριότητας (quantitative structure-activity relationship, QSAR), ένας από τους οποίους ήταν ο RFE, ως βασικού αλγόριθμου στην παλινδρόμηση των διανυσμάτων υποστήριξης.

Η αποτελεσματικότητα του αλγόριθμου SVM-RFE καθώς επίσης και η σχέση που υπάρχει μεταξύ των δεδομένων υψηλής διάστασης και των ΥΣ από την άποψη της διαστάσεως των προβλημάτων μας ώθησε στη δημιουργία μίας μεθόδου ανάλυσης ΥΣ κατάλληλα προσαρμοσμένης στην παλινδρόμηση. Αυτή η μέθοδος αποτελεί τροποποίηση του αλγορίθμου SVM-RFE, για προβλήματα γραμμικής παλινδρόμησης, κατάλληλα προσαρμοσμένα στις περιπτώσεις που ο πίνακας σχεδιασμού είναι ένας ΥΣ. Αναφερόμαστε σε αυτή την τροποποίηση ως Μηχανές διανυσμάτων υποστήριξης βασιζόμενες στην αναδρομική αφαίρεση μεταβλητών (Support vector regression based on Recursive Feature Elimination, SVR-RFE). Σε σύγκριση με την εφαρμογή των Fang και Tai [49], ο αλγόριθμός μας παρέχει ένα μέτρο πρόβλεψης μέσω του διανύσματος του βάρους που αποκτάται από το SVR, ενώ περιλαμβάνει δύο επιπλέον βήματα, τα βήματα 7 και 8, έτσι ώστε να εκτιμήσει αποτελεσματικά το τελικό μοντέλο στην περίπτωση ενός ΥΣ. Μια λεπτομερής περιγραφή του αλγορίθμου παρέχεται στην ενότητα 6.3.1. Τα αποτελέσματα που αποκτήθηκαν από τα πειράματα

προσομοίωσης αποκαλύπτουν ότι η μέθοδος είναι ιδιαίτερα αποτελεσματική, ιδίως όσον αφορά τα ποσοστά σφάλματος Τύπου II, για την ανάλυση ΥΣ με παράγοντες είτε μικτών είτε δύο επιπέδων.

## **6.2 Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines, SVM)**

Θεωρούμε ένα σύνολο εκπαίδευσης που αποτελείται από  $n$  παρατηρήσεις  $x_1, x_2, \dots, x_n \in \mathcal{X}$ , όπου  $\mathcal{X}$  συμβολίζει το χώρο των εσωτερικών προτύπων (π.χ.  $\mathcal{X} = \mathbb{R}^p$ ), και ένα σύνολο που αποτελείται από τις ετικέτες των σχετικών ετικετών κλάσης  $y_1, y_2, \dots, y_n \in \mathbb{R}$ . Οι μεταβλητές εισόδου  $\mathbf{x} \in \mathbb{R}^p$  είναι διανύσματα  $p$ -διαστάσεων και η έξοδος ή αλλιώς μεταβλητή απόκρισης  $y \in \mathbb{R}$  είναι μία συνεχής μεταβλητή. Στην  $\epsilon$ -SV παλινδρόμηση (SVR, [136]), ο κύριος στόχος μας είναι να βρούμε μία συνάρτηση  $f(\mathbf{x})$  η οποία απέχει το πολύ  $\epsilon$  από τους πραγματικούς στόχους  $y_i$  για όλα τα δεδομένα εκπαίδευσης, και ταυτόχρονα είναι όσο το δυνατόν πιο επίπεδη. Δηλαδή, ενδιαφερόμαστε για σφάλματα εφόσον είναι μεγαλύτερα από  $\epsilon$  αλλά δεν δεχόμαστε καμία απόκλιση μικρότερη από  $\epsilon$ . Ας εξετάσουμε την περίπτωση γραμμικών συναρτήσεων  $f$  ως

$$f(\mathbf{x}; w) = w^T \mathbf{x} + b, \quad w \in \mathcal{X}, b \in \mathbb{R}. \quad (6.1)$$

Για την προσέγγιση του σφάλματος, οι SVM χρησιμοποιούν την  $\epsilon$ -ευαίσθητη συνάρτηση του Vapnik, η οποία ορίζεται ως [66]

$$L_\epsilon(y(\mathbf{x})) = |y(\mathbf{x}) - f(\mathbf{x}; w)|_\epsilon = \begin{cases} 0 & \text{if } |y(\mathbf{x}) - f(\mathbf{x}; w)| \leq \epsilon, \\ |y(\mathbf{x}) - f(\mathbf{x}; w)| - \epsilon, & \text{otherwise,} \end{cases} \quad (6.2)$$

όπου  $\epsilon$  είναι μια ακτίνα ενός σωλήνα εντός του οποίου πρέπει να βρίσκεται η συνάρτηση παλινδρόμησης.

Με άλλα λόγια, το  $\epsilon$  αποτελεί την ακρίβεια με την οποία η συνάρτηση  $f$  προσεγγίζει όλα τα ζεύγη  $(x_i, y_i)$ . Η βασική ιδέα είναι να μειωθεί η πολυπλοκότητα του μοντέλου με μία ανοχή στα αντίστοιχα σφάλματα. Η ελαχιστοποίηση του σφάλματος γενίκευσης, δηλαδή η εύρεση της βέλτιστης λύσης, ισοδυναμεί με την ελαχιστο-

ποίηση της νόρμας  $\|w\|^2 = w^T w$ . Το αντίστοιχο κυρτό πρόβλημα βελτιστοποίησης μπορεί να θεωρηθεί ως

$$\text{minimize } \frac{1}{2} \|w\|^2, \quad (6.3a)$$

$$\text{s.t. } \begin{cases} y_i - w^T x_i - b \leq \epsilon, \\ w^T x_i + b - y_i \leq \epsilon, \end{cases} \quad (6.3b)$$

που είναι εφικτή με ακρίβεια  $\epsilon$ . Ωστόσο, στην περίπτωση που θέλουμε να επιτρέψουμε κάποια λάθη, ανάλογα με τη συνάρτηση απώλειας μαλακών περιθωρίων (soft margin loss function) που χρησιμοποιήθηκε στις SVMs [28], εισάγουμε τις χαλαρές μεταβλητές  $\xi_i, \xi_i^*$ , έτσι ώστε να αντιμετωπίσουμε τους μη εφικτούς περιορισμούς. Για να το θέσουμε διαφορετικά, οι χαλαρές μεταβλητές είναι μη μηδενικές μεταβλητές οι οποίες επιτρέπουν να υπάρχουν μεμονωμένες πληροφορίες-μεταβλητές στη λάθος πλευρά του περιθωρίου ή του υπερεπιπέδου. Έτσι, παίρνουμε τη μορφοποίηση για το  $\epsilon$ -ευαίσθητο SVR [37, 123]

$$\text{minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*), \quad (6.4a)$$

$$\text{s.t. } \begin{cases} y_i - w^T x_i - b \leq \epsilon + \xi_i, \\ w^T x_i + b - y_i \leq \epsilon + \xi_i^*, \\ \xi_i, \xi_i^* \geq 0, \end{cases} \quad (6.4b)$$

όπου  $w$  είναι το διάνυσμα των βαρών για τις μεταβλητές και  $C > 0$ . Αυτή η σταθερά καθορίζει την αντιστάθμιση (trade-off) μεταξύ της ποινής για τα σφάλματα τα οποία είναι μεγαλύτερα από  $\epsilon$  και της επιπεδότητας των βαρών. Στις περισσότερες περιπτώσεις, το πρόβλημα βελτιστοποίησης (6.4) μπορεί να λυθεί ευκολότερα στη διπλή του μορφοποίηση. Μια τυποποιημένη μέθοδος δυοποίησης (dualization method) είναι αυτή της χρήσης των πολλαπλασιαστών Lagrange. Η βασική ιδέα είναι να κατασκευάσουμε μία συνάρτηση Lagrange (πρωτεύουσα, primal) από την αντικειμενική συνάρτηση και τους αντίστοιχους περιορισμούς, εισάγοντας ένα διπλό σύνολο μεταβλητών, τους πολλαπλασιαστές Lagrange. Μπορεί να αποδειχθεί ότι αυτή η συνάρτηση έχει ένα σημείο σέλας ως προς τις αρχικές και δυϊκές μεταβλητές (primal and dual variables) στη λύση. Αν οι  $\hat{w}, \hat{b}$  είναι οι ελαχιστοποιητές της (6.4), η



συνάρτηση της λύσης μπορεί να αποδειχθεί ότι έχει τη μορφή

$$\hat{w} = \sum_{i=1}^n (\hat{\alpha}_i - \hat{\alpha}_i^*) x_i, \quad (6.5)$$

και, έτσι, η συνάρτηση της πρόβλεψης του SVR,  $f(\mathbf{x})$ , παίρνει τη μορφή

$$\hat{f}(\mathbf{x}; w) = \sum_{i=1}^n (\hat{\alpha}_i - \hat{\alpha}_i^*) x_i^T \mathbf{x} + b. \quad (6.6)$$

Αυτή η συνάρτηση πρόβλεψης βασίζεται μόνο σε ένα μικρό υποσύνολο δεδομένων κατάρτισης (δεδομένα εκπαίδευσης) με μη μηδενικούς πολλαπλασιαστές Lagrange  $\hat{\alpha}_i$  ή  $\hat{\alpha}_i^*$ , όπου  $\hat{\alpha}_i, \hat{\alpha}_i^* \in [0, C]$ . Αυτοί οι, μη μηδενικοί, συντελεστές είναι θετικοί και επιλύουν το τετραγωνικό πρόβλημα προγραμματισμού:

$$\text{maximize} \begin{cases} -\frac{1}{2} \sum_{i=1}^n (\alpha_i - \alpha_i^*) (\alpha_i - \alpha_i^*) x_i^T x_i \\ -\epsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*), \end{cases} \quad (6.7a)$$

$$\text{s.t.} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \quad \text{and} \quad \alpha_i, \alpha_i^* \in [0, C]. \quad (6.7b)$$

Λόγω της φύσης των προαναφερθέντων περιορισμών, τυπικά μόνο ένα υποσύνολο από τις τιμές της λύσης  $(\alpha_i - \alpha_i^*)$  είναι μη μηδενικές και οι σχετικές τιμές των δεδομένων καλούνται διανύσματα υποστήριξης. Η μηχανή διανυσμάτων υποστήριξης [136], που περιγράφεται παραπάνω, χρησιμοποιείται σε προβλήματα παλινδρόμησης και ονομάζεται (SVR).

## 6.3 Νέα μέθοδος επιλογής μεταβλητών που βασίζεται στις SVR

Στο παρόν κεφάλαιο θα παρουσιάσουμε μια νέα μέθοδο επιλογής μεταβλητών εμπνευσμένη από την παλινδρόμηση των διανυσμάτων υποστήριξης σε συνδυασμό με τον αλγόριθμο επαναλαμβανόμενης επιλογής χαρακτηριστικών (support vector regression-recursive feature elimination SVR-RFE). Αυτή η μέθοδος αφαιρεί τα μη σχετικά χαρακτηριστικά χρησιμοποιώντας το μέγεθος του βάρους. Πιο συγκεκριμένα, βασίζεται στο κριτήριο κατάταξης  $w^2$ , ενώ παρέχει ένα μέτρο προβλεπτικής

ικανότητας, δεδομένου ότι τα χαρακτηριστικά με μικρή τιμή αυτού του κριτηρίου έχουν μικρή επίδραση στη συνάρτηση πρόβλεψης. Ο αλγόριθμος ξεκινά με το σύνολο των χαρακτηριστικών και διαδοχικά αφαιρεί αυτά με τα μικρά βάρη. Μια λεπτομερής περιγραφή του SVR-RFE παρέχεται στην επαναληπτική διαδικασία της ενότητας 6.3.1.

### 6.3.1 Αλγόριθμος

Ένα σχεδιάγραμμα του αλγορίθμου μας αποτελείται από τα ακόλουθα βήματα: Αρχικοποίηση το υποσύνολο των χαρακτηριστικών ως  $SFI = [1, 2, \dots, m]$  (Surviving Features Indexes, Χαρακτηριστικά που έχουν επιβιώσει) και ως  $r$  έναν κενό πίνακα.

1. Ενημέρωσε τα νέα παραδείγματα εκπαίδευσης  $X = X(:, SFI)$
2. Εκπαίδευσε το μοντέλο SVR.
3. Υπολόγισε το διάνυσμα του βάρους  $w$  για κάθε χαρακτηριστικό.
4. Υπολόγισε το κριτήριο κατάταξης  $w^2$  για όλα τα χαρακτηριστικά.
5. Βρες το χαρακτηριστικό  $z$  με το μικρότερο κριτήριο κατάταξης.
6. Ενημέρωσε το  $r$  και εξάλειψε το χαρακτηριστικό  $z$  από το  $SFI$ .

$$r = [SFI(z), r], \quad (6.8)$$

$$SFI = SFI - SFI(z). \quad (6.9)$$

Επανάλαβε τα βήματα [1] - [6] μέχρι το  $SFI$  να γίνει ένας άδειος πίνακας. Εκτελώντας αυτή την επαναληπτική διαδικασία αποκτάμε ως έξοδο την ταξινομημένη λίστα των χαρακτηριστικών  $r$ . Στη συνέχεια, συνεχίζουμε τη διαδικασία με τα ακόλουθα βήματα

7. Εξάλειψε τα χαρακτηριστικά που είναι λιγότερο σημαντικά, κρατώντας τα  $n/2$  πιο σημαντικά.
8. Εκτίμησε τις μεταβλητές που παραμένουν, μέσω του  $\hat{\beta} = (X'X)^{-1}X'y$ . Οι μεταβλητές που εμφανίζονται ως μη σημαντικές θα λάβουν τιμές κοντά στο μηδέν.

**Παρατήρηση 1** Στο βήμα 2 το πακέτο "e1071" με βάση το εγχειρίδιο LibSVM [23], του λογισμικού της R χρησιμοποιείται για την εκπαίδευση του μοντέλου  $\epsilon$ -SVR με γραμμικό πυρήνα. Οι παράμετροι του αλγορίθμου προσδιορίστηκαν μέσω διασταυρωμένης επικύρωσης όπως και στους Scholkopf et al. [121]. Τέλος, οι επιλεγμένες τιμές για τα  $C$  και  $\epsilon$  ήταν  $C = 10$  και  $\epsilon = 0.1$ .

**Παρατήρηση 2** Στο βήμα 5, επιλέξαμε να κρατήσουμε τις  $n/2$  πιο σημαντικές μεταβλητές υπό την υπόθεση της σποραδικότητας των επιδράσεων. Πολλοί συγγραφείς υποθέτουν ότι υπάρχουν μέχρι και  $n/3$  ενεργοί παράγοντες. Στη μελέτη μας υποθέτουμε έως και  $n/2$  ενεργούς παράγοντες. Τα πλεονεκτήματα της μελέτης μας παρουσιάζονται στην ενότητα 6.4.

## 6.4 Αξιολόγηση της προτεινόμενης μεθόδου

Σε αυτή την ενότητα εξετάζουμε την απόδοση του SVR-RFE μέσω πειραμάτων προσομοίωσης. Στη μελέτη προσομοίωσης καθώς και στο παράδειγμα 6.4.2 συγκρίνεται η απόδοση του SVR-RFE με τρεις, γνωστές από τη βιβλιογραφία, προσεγγίσεις οι οποίες περιγράφονται εν συντομία στην ενότητα 6.4.1.

### 6.4.1 Μέθοδοι ανάλυσης

Παρόλο που έχουν δημοσιευθεί πολλές εργασίες όσον αφορά στην κατασκευή ΥΣ, πολύ λιγότερες ερευνητικές εργασίες έχουν γραφεί για την ανάλυσή τους [31]. Τρεις μέθοδοι που παρουσιάζουν ιδιαίτερο ενδιαφέρον είναι ο επιλογέας Dantzig (Dantzig selector), η παλινδρόμηση κατά βήματα και η ποινικοποιημένη μέθοδος  $L_q$ .

Ο επιλογέας Dantzig (DS, [21]) αποτελεί μία δημοφιλή μέθοδο συρρίκνωσης μεταξύ των μεθόδων επιλογής μεταβλητών. Οι Phoa et al. [112] πρότειναν μια μέθοδο ανάλυσης για ΥΣ μέσω του DS, στην οποία ο εκτιμητής  $\hat{\beta} \in R^p$  είναι η λύση στο ακόλουθο πρόβλημα  $l_1$ -κανονικοποίησης. Ο DS επιλέγει τους ενεργούς παράγοντες επιλύοντας το ακόλουθο κυριό πρόβλημα βελτιστοποίησης

$$\min_{\hat{\beta} \in R^p} \|\hat{\beta}\|_{l_1}, \quad \text{subject to } \|X'(y - X\hat{\beta})\|_{l_\infty} \leq \lambda \quad (6.10)$$

όπου  $\lambda$  είναι η παράμετρος συντονισμού,  $\|\beta\|_{l_1} = |\beta_1| + \dots + |\beta_p|$  είναι η νόρμα  $l_1$  και  $\|\beta\|_{l_\infty} = \max(|\beta_0|, \dots, |\beta_p|)$  είναι η νόρμα  $l_\infty$ , για το αντίστοιχο διάνυσμα. Για να επιλέξουμε τη βέλτιστη παράμετρο κανονικοποίησης/ρύθμισης χρησιμοποιήσαμε διασταυρωμένη επικύρωση. Η αποτελεσματικότητα αυτής της μεθόδου αποδείχθηκε επίσης από τους Marley και Woods [101] και από τους Errore et al. [40]. Στην παλινδρόμηση κατά βήματα, όπως ακριβώς και στην προς τα εμπρός επιλογή (forward selection), είναι σύνηθες τα ποσοστά σφαλμάτων Τύπου I να είναι αρκετά υψηλά [140]. Τέτοιες τεχνικές είναι πολύ δημοφιλείς και διατίθενται σε πολλά λογισμικά πακέτα, όπως στην R, τη SAS και την JMP. Μία από τις πιο συνηθισμένες επιλογές για το επίπεδο σημαντικότητας είναι το 0.1, και για το  $\alpha_{enter}$  και για το  $\alpha_{remove}$ , το οποίο και χρησιμοποιούμε στην παρούσα μελέτη. Έχουν προταθεί πολλές άλλες τεχνικές ανάλυσης μερικές από τις οποίες αποτελούν επεκτάσεις της κλασικής παλινδρόμησης κατά βήματα [100]. Παρά τις ελλείψεις της, η παλινδρόμηση κατά βήματα είναι απλή, γρήγορη και χρησιμοποιείται συνήθως στην πράξη λόγω της διαθεσιμότητάς της στα στατιστικά λογισμικά πακέτα.

Η ποινή  $L_q$  Lasso επιλύει το ακόλουθο πρόβλημα βελτιστοποίησης

$$\min_{\hat{\beta} \in R^p} n^{1/q} \|(y - X\hat{\beta})\|_{l_q} + \lambda \|\hat{\beta}\|_{l_1}, \quad (6.11)$$

όπου  $1 \leq q \leq 2$ . Η ποινή  $L_q$  Lasso είναι ισοδύναμη με την SQR Lasso, όταν το  $q = 2$ . Η παράμετρος κανονικοποίησης/ρύθμισης της SQR Lasso, είναι  $C\sqrt{\log(p)/n}$ , όπου  $C$  είναι μία σταθερά,  $p$  είναι η διάσταση του σχεδιασμού και  $n$  το αντίστοιχο μέγεθος των πειραματικών εκτελέσεων. Επιλέγουμε χειροκίνητα την ελάχιστη παράμετρο κανονικοποίησης να ισούται με  $\sqrt{\log(p)/n}$ , ενώ αναφερόμαστε σε αυτή η μέθοδος ως  $L_q$ .

Στην παρούσα μελέτη προσομοίωσης εφαρμόσαμε όλες τις μεθόδους χρησιμοποιώντας το στατιστικό λογισμικό της R. Ως συμπλήρωμα στους κώδικές μας, χρησιμοποιήθηκαν τα στατιστικά πακέτα "flare" [84] και "e1071". Στις επόμενες ενότητες παρουσιάζονται μια ολοκληρωμένη μελέτη προσομοίωσης, καθώς και ένα ιδιαίτερα επεξηγηματικό παράδειγμα.

### 6.4.2 Παράδειγμα

Αυτό το παράδειγμα είναι ιδιαίτερα δημοφιλές δεδομένου ότι έχει αναλυθεί από πολλούς συγγραφείς, μέχρι στιγμής. Θεωρούμε τον ΥΣ του Lin [91], ο οποίος βασίζεται στα δεδομένα του Williams [141]. Τα επίπεδα των παραγόντων, καθώς επίσης και η μεταβλητή απόκριση παρέχονται στον Πίνακα 7.2. Ο σχεδιασμός αποτελείται από 24 παράγοντες και 14 πειραματικές εκτελέσεις, ωστόσο περιέχει δύο όμοιες στήλες, τη 13 και τη 16, μία από τις οποίες πρέπει να διαγραφεί, οδηγώντας σε 23 παράγοντες. Διαγράψαμε τη στήλη 13, η οποία αντιστοιχεί στον παράγοντα  $X_{13}$ , αλλάζοντας την επισήμανση των παραγόντων 14 – 24 σε 13 – 23 (βλ. Πίνακα 1). Ο προκύπτων σχεδιασμός 23 παραγόντων και 14 πειραματικών εκτελέσεων είναι  $E(s^2)$ -βέλτιστος ΥΣ και αποτελεί ένα δημοφιλή σχεδιασμό για την εξέταση της απόδοσης διάφορων μεθόδων ανάλυσης σε ΥΣ.

**Πίνακας 6.1:** Δύο επιπέδων υπερκορεσμένος σχεδιασμός (Lin, [91]) και τα δεδομένα Rubber του Williams [141] (Απόκριση).

Εκτέλεση	Παράγοντες																							Απόκριση
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
1	1	1	1	-1	-1	-1	1	1	1	1	1	-1	-1	-1	1	1	-1	-1	1	-1	-1	-1	1	133
2	1	-1	-1	-1	-1	-1	1	1	1	-1	-1	-1	1	1	1	-1	1	-1	-1	1	1	-1	-1	62
3	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	1	1	1	1	-1	-1	-1	-1	-1	1	1	45
4	1	1	-1	1	-1	1	-1	-1	-1	1	1	-1	-1	1	1	-1	1	1	1	1	-1	-1	-1	52
5	-1	-1	1	1	1	1	-1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1	-1	1	1	1	56
6	-1	-1	1	1	1	1	1	-1	1	1	1	-1	1	1	-1	1	1	1	1	1	1	1	-1	47
7	-1	-1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1	1	1	1	1	1	-1	-1	1	88
8	-1	1	1	-1	-1	1	-1	1	-1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	193
9	-1	-1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1	-1	1	1	-1	-1	-1	-1	-1	1	32
10	1	1	1	1	-1	1	1	1	-1	-1	-1	1	1	1	-1	1	-1	1	-1	1	-1	-1	1	53
11	-1	1	-1	1	1	-1	-1	1	1	-1	1	-1	1	-1	-1	-1	1	1	-1	-1	-1	-1	1	276
12	1	-1	-1	-1	1	1	1	-1	1	1	1	1	-1	-1	1	-1	-1	-1	-1	1	1	1	1	145
13	1	1	1	1	1	-1	1	-1	1	-1	1	-1	1	-1	-1	-1	1	-1	1	1	1	-1	1	130
14	-1	-1	1	-1	-1	-1	-1	-1	-1	-1	1	1	1	-1	-1	-1	-1	-1	1	-1	1	-1	-1	127

Ο Πίνακας 6.2 εμφανίζει τις μεταβλητές που επελέγησαν από κάθε μέθοδο, τις μεταβλητές που ανιχνεύτηκαν από τους Westfall et al. [140], καθώς και τα αντίστοιχα  $R^2$  και  $R^2_{adjusted}$  που υπολογίστηκαν για κάθε μοντέλο. Αναμφισβήτητα, ο κυρίαρχος παράγοντας είναι ο  $X_{14}$  αφού εμφανίζεται σε κάθε ένα από τα επιλεγμένα μοντέλα.

**Πίνακας 6.2:** Μεταβλητές που έχουν επιλεγεί από κάθε μέθοδο για τα δεδομένα Rubber του Williams [141].

Μέθοδος	Επιλεγμένες μεταβλητές	Μέγεθος του μοντέλου	$R^2$	$R^2_{adjusted}$
Westfall [140]	14 12 19 4 10 11	6	0.987	0.975
SVR-RFE	14 19 7 9 15 8 11	7	0.960	0.913
Danzig Selector	14	1	0.632	0.601
$L_q$ penalty	14 16	2	0.694	0.639
stepwise regression	14 12 19 4 10 11 7 1 13 16 21	11	0.969	0.795

Ο συντελεστής  $R^2$  αυξάνεται αυτόματα και ψευδώς κάποιες φορές όταν προστίθενται επιπλέον επεξηγηματικές μεταβλητές στο μοντέλο. Ο  $R_{adjusted}^2$  είναι μια προσπάθεια να ληφθεί υπόψη αυτό το φαινόμενο. Βασιζόμενοι κυρίως στις τιμές  $R_{adjusted}^2$  καταλήγουμε στο συμπέρασμα ότι ο SVR-RFE επιλέγει ένα κατάλληλο σύνολο σημαντικών επιδράσεων ενώ διατηρεί το μέγεθος του μοντέλου σε χαμηλό επίπεδο σε σύγκριση με την παλινδρόμηση κατά βήματα (stepwise regression).

### 6.4.3 Μελέτη προσομοιώσεων

Για να συγκρίνουμε την απόδοση του αλγορίθμου SVR-RFE με τις άλλες τρεις μεθόδους θεωρούμε διάφορους ΥΣ και διάφορα σενάρια προσομοίωσης. Πιο συγκεκριμένα:

1. **Υπερκορεσμένοι Σχεδιασμοί.** Χρησιμοποιήθηκαν είκοσι ΥΣ. Δεκατέσσερις είναι σχεδιασμοί δύο επιπέδων οκτώ εξ αυτών  $E(s^2)$ -βέλτιστοι σχεδιασμοί που κατασκευάστηκαν από τους Koukouninos et al. [76], και τρεις εξ αυτών Μπεϋζιανοί  $D$ -βέλτιστοι σχεδιασμοί που κατασκευάστηκαν από τους Marley και Woods [101] (βλ. Πίνακα 6.4). Επίσης χρησιμοποιήθηκαν έξι σχεδιασμοί μικτών επιπέδων με παράγοντες που είχαν είτε δύο είτε τρεις στάθμες (βλ. Πίνακα Α.1). Πέντε από τους έξι ΥΣ είναι  $k$ -κυκλικοί μικτών επιπέδων ΥΣ ενώ ο έκτος ήταν αυτός που παρουσιάστηκε στο πείραμα σχετικά με τη γλυκόζη στο αίμα των Wu και Hamada [143].
2. **Σενάριο Προσομοιώσεων.** Ο αριθμός  $d$  των ενεργών παραγόντων επιλέχθηκε από το σύνολο  $\{0, 1, 2, \dots, n/2\}$  για κάθε έναν από τους πίνακες σχεδιασμού  $n \times p$ .
3. **Μέθοδοι Ανάλυσης.** Εκτός από τον SVR-RFE εφαρμόσαμε το DS, την παλινδρόμηση κατά βήματα και την ποινικοποιημένη μέθοδο  $L_q$ .

Τα σενάρια προσομοιώσεων πραγματοποιήθηκαν σύμφωνα με εκείνα που είχαν εφαρμόσει οι Marley και Woods [101]. Πιο συγκεκριμένα, σε κάθε μία από τις 1000 επαναλήψεις:

1. Από τις στήλες  $p$  του  $\mathbf{X}$ ,  $d$  στήλες επιλέγονται τυχαία ως οι ενεργοί παράγοντες.

Οι συντελεστές  $\beta$  για τους ενεργούς παράγοντες λαμβάνονται με δειγματοληψία από την κατανομή  $N(5, 1)$  και τυχαία εφαρμόζεται ένα πρόσημο (+ ή -).

2. Οι συντελεστές των μη ενεργών παραγόντων λαμβάνονται τυχαία από την κατανομή  $N(0, 0.2)$ .
3. Το διάνυσμα της απόκρισης παράγεται από το γραμμικό μοντέλο  $y = \mathbf{X}\beta + \epsilon$ , όπου  $\mathbf{X} = [x_1, x_2, \dots, x_p]$ ,  $\beta$  είναι οι συντελεστές που επελέγησαν και  $\epsilon \sim N(0, 1)$  το τυχαίο σφάλμα.
4. Οι πραγματικοί ενεργοί παράγοντες επιλέγονται χρησιμοποιώντας την εξεταζόμενη μέθοδο επιλογής μοντέλων.
5. Στο τέλος των 1000 προσομοιώσεων υπολογίστηκαν το μέσο ποσοστό των μη ενεργών επιδράσεων που εντοπίστηκαν ως ενεργοί (ποσοστό σφάλματος Τύπου I), καθώς και το μέσο ποσοστό των ενεργών επιδράσεων που εντοπίστηκαν ως μη ενεργοί (ποσοστό σφάλματος Τύπου II). Επίσης καθορίστηκε το ποσοστό των ενεργών επιδράσεων που ορθώς εντοπίστηκαν ως ενεργές (στατιστική ισχύς, statistical power).

Σημειώνουμε ότι στην περίπτωση των ΥΣ, ο έλεγχος του ποσοστού σφάλματος Τύπου II, διατηρώντας τη στατιστική ισχύ στο υψηλότερο επίπεδο, είναι ακόμη πιο κρίσιμος από τον έλεγχο του ποσοστού σφάλματος Τύπου I [140].

#### 6.4.4 Αποτελέσματα προσομοιώσεων

Σύμφωνα με τη διαδικασία προσομοιώσεων και με βάση τις μεθόδους ανάλυσης που περιγράφονται στις προηγούμενες υποενότητες, παρουσιάζουμε τα αποτελέσματα της μελέτης προσομοίωσης στα Σχήματα 6.1 – 6.6 καθώς και στους Πίνακες B.1 – B.4 για μια λεπτομερή περιγραφή των αποτελεσμάτων. Οι Πίνακες B.1 έως B.3 και τα Σχήματα 6.1, 6.2, 6.3 παρουσιάζουν τα αποτελέσματα που λαμβάνονται με την ανάλυση των ΥΣ δύο επιπέδων που περιγράφονται στον Πίνακα 6.3. Ο Πίνακας B.4, καθώς επίσης και τα Σχήματα 6.4, 6.5 και 6.6 παρουσιάζουν τα αποτελέσματα των σχεδιασμών μικτών επιπέδων, των οποίων η λεπτομερής περιγραφή δίδεται στον Πίνακα 6.4. Οι παραπάνω Πίνακες συνοψίζουν τα σφάλματα ως συνάρτηση του με-

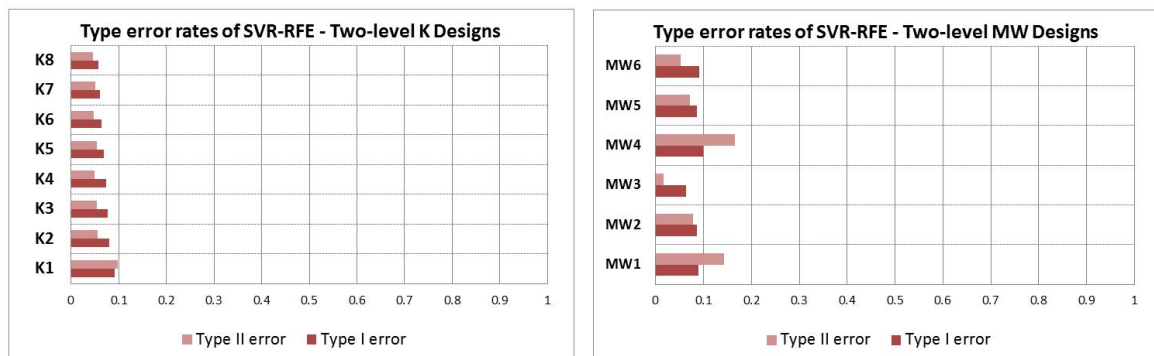
γέθους του σχεδιασμού, του τύπου του σχεδιασμού και του σεναρίου προσομοίωσης, για όλες τις εξεταζόμενες μεθόδους.

**Δύο επιπέδων υπερκορεσμένοι σχεδιασμοί**

Ο Πίνακας 6.3 παρουσιάζει τους ΥΣ δύο επιπέδων που χρησιμοποιήθηκαν στη μελέτη προσομοίωσης. Η πρώτη στήλη αντιπροσωπεύει το όνομα κάθε σχεδιασμού, όπου τα γράμματα « $K_i$ »,  $i = 1, \dots, 8$  αντιπροσωπεύουν τους σχεδιασμούς του Koukouvinos et al. [76] ενώ τα γράμματα « $MW_i$ »,  $i = 1, \dots, 6$  αντιπροσωπεύουν τους σχεδιασμούς του Marley και Woods [101], αντίστοιχα. Η δεύτερη και η τρίτη στήλη αντιπροσωπεύουν το μέγεθος των πειραματικών εκτελέσεων και των παραγόντων, αντίστοιχα, ενώ η τέταρτη στήλη αντιπροσωπεύει τον τύπο του σχεδιασμού.

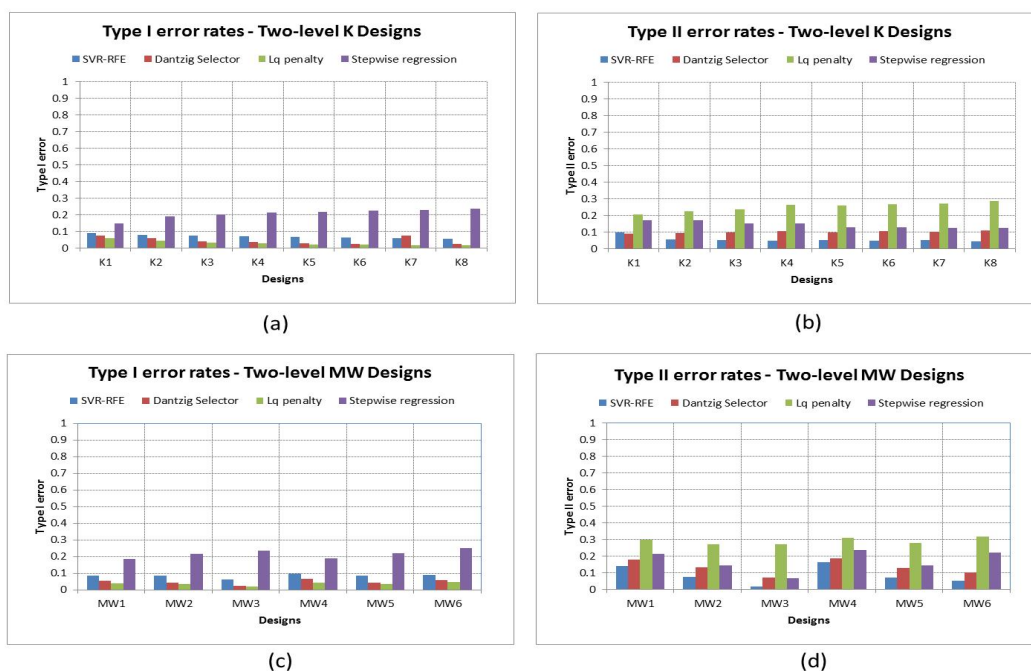
**Πίνακας 6.3:** Δύο επιπέδων σχεδιασμοί που χρησιμοποιήθηκαν στη μελέτη προσομοίωσης.

Συμβολισμός	Μέγεθος του σχεδιασμού		Τύπος Σχεδιασμού	
	Εκτελέσεις	Παράγοντες		
$K1$	8	14	$E(s^2)$ -optimal	Koukouvinos et al. [76]
$K2$	10	18	$E(s^2)$ -optimal	Koukouvinos et al. [76]
$K3$	12	22	$E(s^2)$ -optimal	Koukouvinos et al. [76]
$K4$	14	26	$E(s^2)$ -optimal	Koukouvinos et al. [76]
$K5$	16	30	$E(s^2)$ -optimal	Koukouvinos et al. [76]
$K6$	18	34	$E(s^2)$ -optimal	Koukouvinos et al. [76]
$K7$	20	38	$E(s^2)$ -optimal	Koukouvinos et al. [76]
$K8$	22	42	$E(s^2)$ -optimal	Koukouvinos et al. [76]
$MW1$	12	26	Bayesian $D$ -optimal	Marley and Woods [101]
$MW2$	14	24	Bayesian $D$ -optimal	Marley and Woods [101]
$MW3$	18	22	Bayesian $D$ -optimal	Marley and Woods [101]
$MW4$	12	26	$E(s^2)$ -optimal	Marley and Woods [101]
$MW5$	14	24	$E(s^2)$ -optimal	Marley and Woods [101]
$MW6$	18	22	$E(s^2)$ -optimal	Marley and Woods [101]



**Σχήμα 6.1:** Σφάλματα Τύπου I (Type I) και Τύπου II (Type II) για το SVR-RFE στους ΥΣ δύο επιπέδων. Το αριστερό πλαίσιο αναφέρεται στους σχεδιασμούς « $K$ » ενώ το δεξιό πλαίσιο αναφέρεται στους σχεδιασμούς « $MW$ ».



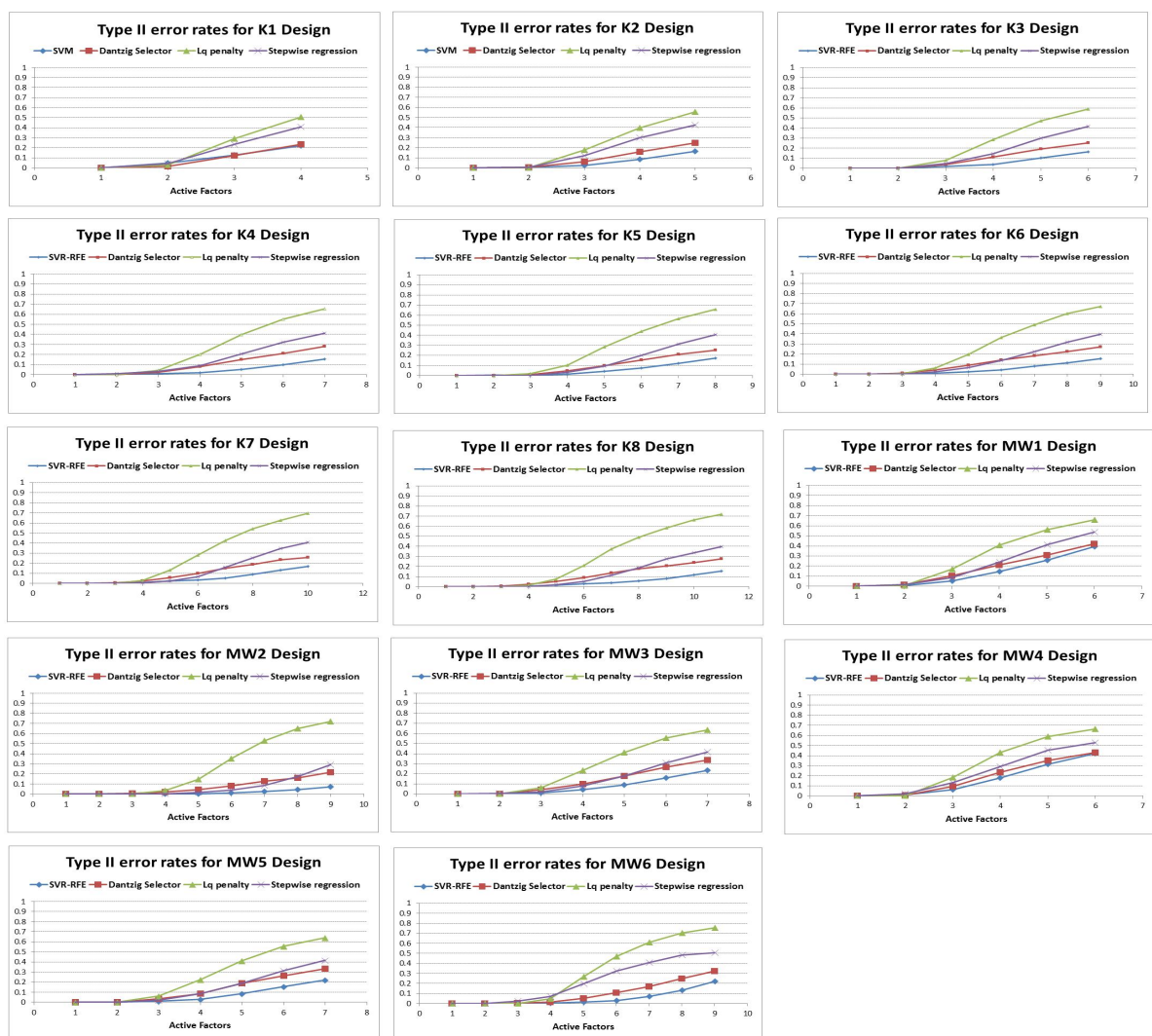


**Σχήμα 6.2:** Μέσες τιμές για τα σφάλματα Τύπου I (Type I) και Τύπου II (Type II) για όλες τις μεθόδους στους εξεταζόμενους σχεδιασμούς δύο επιπέδων. Τα δύο πάνω πλαίσια αναφέρονται στους μέσους των σφαλμάτων (α) Τύπου I και (β) Τύπου II των σχεδιασμών « $K$ », ενώ τα δύο κάτω πλαίσια αναφέρονται στους μέσους των σφαλμάτων (γ) Τύπου I και (δ) Τύπου II των σχεδιασμών « $MW$ ». Τα σφάλματα του SVR-RFE αναπαριστώνται από μπλε έντονες ράβδους, τα έντονα κόκκινα γράμματα αναπαριστούν τα σφάλματα για το DS, οι μωβ έντονες ράβδοι την παλινδρόμηση κατά βήματα και οι πράσινες έντονες ράβδοι αναπαριστούν την ποινή  $L_q$ .

Το Σχήμα 6.1 απεικονίζει την απόδοση του SVR-RFE όπου τόσο τα σφάλματα Τύπου I όσο και τα σφάλματα Τύπου II απεικονίζονται για τους σχεδιασμούς δύο επιπέδων, « $K$ » και « $MW$ », του Πίνακα 6.3. Είναι αξιοσημείωτο ότι ο αλγόριθμος SVR-RFE ελέγχει τη μέση τιμή τόσο των σφαλμάτων Τύπου I όσο και των σφαλμάτων Τύπου II, στο επίπεδο 0.10 για τους σχεδιασμούς « $K$ », και στο επίπεδο 0.16 για τους σχεδιασμούς « $MW$ ».

Το Σχήμα 6.2 απεικονίζει τη σαφή υπεροχή του SVR-RFE έναντι του DS, της  $L_q$  ποινικοποιημένης μεθόδου και της παλινδρόμησης κατά βήματα. Ιδιαίτερα, ο αλγόριθμος SVR-RFE επιτυγχάνει πολύ μικρότερα σφάλματα Τύπου II από αυτά που προέκυψαν από τις άλλες τρεις μεθόδους (Σχήμα 6.2(b) και 6.2(d)). Υπό αυτή την έννοια, το SVR-RFE είναι πιο αποδοτικό σε σύγκριση με τις άλλες τρεις διαδικασίες. Παρόλο που ο DS και η ποινή  $L_q$ , επιτυγχάνουν ελαφρώς καλύτερη επίδοση από το SVR-RFE σε σχέση με τα σφάλματα Τύπου I, αυτό συμβαίνει σε ιδιαίτερα μικρό ποσοστό της τάξεως του 5%. Έτσι, μπορούμε σίγουρα να υποστηρίξουμε ότι

δεν υπάρχουν αξιοσημείωτες διαφορές μεταξύ των διαφόρων διαδικασιών ανάλυσης (Σχήμα 6.2(a) και 6.2(c)). Σημειώνουμε ότι, για ΥΣ, το σφάλμα Τύπου II είναι πιο σημαντικό [139], δεδομένου του γεγονότος ότι αυτό ορίζεται ως η πιθανότητα να χάσουμε έναν ενεργό παράγοντα. Επομένως, όσο χαμηλότερα είναι τα σφάλματα Τύπου II, τόσο υψηλότερη είναι η πιθανότητα για την ανίχνευση ενός πραγματικά ενεργού παράγοντα. Εξάλλου, όπως σημειώνεται από τους Westfall et al. [140], τα σφάλματα Τύπου I είναι πολύ πιθανό να εμφανιστούν στην προσαρμογή μοντέλων όπου ισχύει η αρχή σποραδικότητας των επιδράσεων (effect sparsity models).



**Σχήμα 6.3:** Σφάλματα Τύπου II (Type II) για τους δύο επιπέδων σχεδιασμούς «*K*» και «*MW*». Κάθε πάνελ αναφέρεται σε κάθε ξεχωριστό σχεδιασμό. Ο άξονας x αντιπροσωπεύει τους ενεργούς παράγοντες που ανήκουν στο σύνολο  $\{0, 1, 2, \dots, n/2\}$ . Ο άξονας y αντιπροσωπεύει τα σφάλματα Τύπου II για τις τέσσερις διαφορετικές μεθόδους.

Το Σχήμα 6.3 δείχνει ότι, όταν υπάρχει μία ή ακόμα και δύο ενεργές επιδράσεις,

η επίδοση των εξεταζόμενων μεθόδων ποικίλει ελάχιστα. Ωστόσο, εάν υπάρχουν περισσότερες από δύο ενεργές επιδράσεις, η επίδοση των εξεταζόμενων μεθόδων επηρεάζεται έντονα, ενώ αυτή η επίδραση αυξάνεται όταν αυξάνεται ο αριθμός των παραγόντων. Αυτό το σημείο μπορεί να γίνει φανερό από τα ποσοστά ανίχνευσης των αληθώς ενεργών παραγόντων, δηλαδή των σφαλμάτων Τύπου II.

### Μικτών επιπέδων υπερκορεσμένοι σχεδιασμοί

Συνεχίζουμε κάνοντας κάποια πειράματα προσομοίωσης για να εξετάσουμε την απόδοση του SVR-RFE στην ανίχνευση των ενεργών επιδράσεων σε μικτών επιπέδων ΥΣ. Αυτός ο τύπος σχεδιασμών περιλαμβάνει ως ειδική περίπτωση τους ΥΣ πολλαπλών επιπέδων καθώς επίσης και τους ΥΣ δύο επιπέδων. Στους σχεδιασμούς μικτών επιπέδων, εκχωρούμε δύο αντιθέσεις στους παράγοντες με τρία επίπεδα, μία γραμμική και μία τετραγωνική. Η κωδικοποίηση των γραμμικών και των τετραγωνικών αντιθέσεων έχει ως εξής

$$\text{Linear contrast : } (0 \ 1 \ 2) \rightarrow (1 \ 0 \ -1)$$

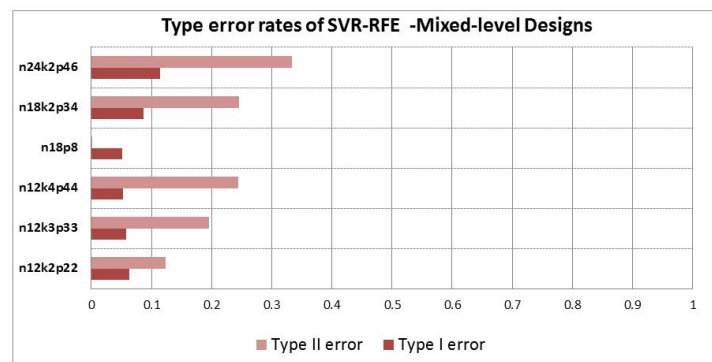
$$\text{Quadratic contrast : } (0 \ 1 \ 2) \rightarrow (1 \ -2 \ +1)$$

Οι σχεδιασμοί μικτών επιπέδων που χρησιμοποιούνται στη μελέτη προσομοίωσης απεικονίζονται στον Πίνακα 6.4. Η πρώτη στήλη αναφέρεται στο όνομα του σχεδιασμού που βασίζεται τόσο στον αριθμό των πειραματικών εκτελέσεων όσο και των παραγόντων. Η τέταρτη και η πέμπτη στήλη παρουσιάζουν τον αριθμό των παραγόντων 2-επιπέδων και των παραγόντων 3-επιπέδων, αντίστοιχα. Η τελευταία στήλη αντιπροσωπεύει το διάνυσμα των στάθμεων του αντίστοιχου  $k$ -κυκλικού σχεδιασμού.

Τα αποτελέσματα της προσομοίωσης συνοψίζονται στους Πίνακες Β.1 και Β.2 του Παραρτήματος Β. Το SVR-RFE επιτυγχάνει εξαιρετικά καλά αποτελέσματα στους ΥΣ μικτών επιπέδων, λαμβάνοντας υπόψη την τιμή του σφάλματος Τύπου I, ενώ παρουσιάζει μία αρκετά καλή απόδοση σε σχέση με τα σφάλματα Τύπου II. Πιο συγκεκριμένα, είναι προφανές ότι το σφάλμα Τύπου I είναι μικρότερο από 0.12 σε όλες τις εξεταζόμενες περιπτώσεις. Το γεγονός αυτό δικαιολογείται επίσης μέσω της γραφικής παράστασης που παρουσιάζεται στο Σχήμα 6.4. Ο αλγόριθμος SVR-RFE

**Πίνακας 6.4:** Σχεδιασμοί μικτών επιπέδων που χρησιμοποιήθηκαν στη μελέτη προσομοίωσης. Η πρώτη στήλη αναφέρεται στο όνομα του σχεδιασμού όπου το  $n$  συμβολίζει τις πειραματικές εκτελέσεις, το  $k$  την κυκλικότητα και το  $p$  τον αριθμό των παραγόντων. Η τέταρτη στήλη παρουσιάζει τον αριθμό των παραγόντων 2-επιπέδων, ενώ η πέμπτη συμβολίζει τον αριθμό των παραγόντων 3-επιπέδων. Η τελευταία στήλη συμβολίζει το διάνυσμα των στάθμεων του αντίστοιχου  $k$ -κυκλικού σχεδιασμού.

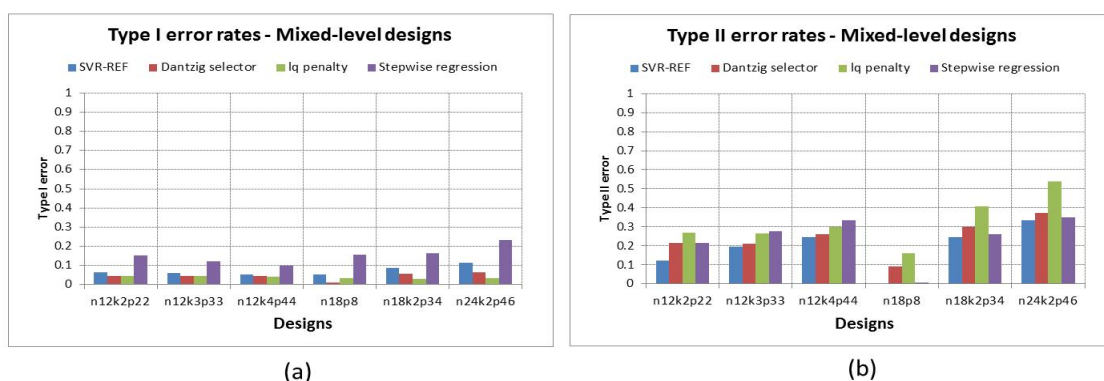
Όνομα	Μέγεθος Σχεδιασμού		Παράγοντες σε		Τελικές στήλες	Τύπος σχεδιασμού	Διάνυσμα επιπέδων
	Εκτελέσεις	Παράγοντες	2-επίπεδα	3-επίπεδα			
$n12k2p22$	12	22	11	11	33	$k$ -κυκλικός	(2, 3)
$n12k3p33$	12	33	22	11	44	$k$ -κυκλικός	(2, 2, 3)
$n12k4p44$	12	44	33	11	55	$k$ -κυκλικός	(2, 2, 2, 3)
$n18p8$	18	8	1	7	15		
$n18k2p34$	18	34	17	17	51	$k$ -κυκλικός	(2, 3)
$n24k2p46$	24	46	23	23	69	$k$ -κυκλικός	(2, 3)



**Σχήμα 6.4:** Σφάλματα Τύπου I (Type I) και Τύπου II (Type II) για τον SVR-RFE στους σχεδιασμούς μικτών επιπέδων.

επιτυγχάνει επίσης εξαιρετικά χαμηλά ποσοστά σφάλματος Τύπου II, ειδικά όταν συγκρίνεται με τις άλλες τρεις μεθόδους.

Το Σχήμα 6.5(α) απεικονίζει την έλλειψη διαφορών μεταξύ των εξεταζόμενων μεθόδων όσον αφορά τις μέσες τιμές των σφαλμάτων Τύπου I. Εκτός από την παλινδρόμηση κατά βήματα, η οποία αποφέρει τα χειρότερα αποτελέσματα, υπάρχουν ακόμα λιγότερο αξιοσημείωτες διαφορές μεταξύ των άλλων τριών μεθόδων ανάλυσης. Το Σχήμα 6.5 (β) απεικονίζει την ιδιαίτερα υψηλή απόδοση του SVR-RFE σε σύγκριση με τις διάφορες μεθόδους ανάλυσης, ακόμη και συγκριτικά με το DS. Πιο συγκεκριμένα, το SVR-RFE έχει περίπου 33% μέση μείωση στα ποσοστά σφαλμάτων Τύπου II έναντι του DS, περίπου 48% μέση μείωση στα ποσοστά σφαλμάτων Τύπου II έναντι της ποινής  $L_q$  καθώς και μια μέση μείωση στα ποσοστά σφαλμάτων Τύπου II περίπου κατά 40% στην παλινδρόμηση κατά βήματα. Το γεγονός αυτό επιβεβαιώνεται μέσω



**Σχήμα 6.5:** Μέσες τιμές για τα σφάλματα Τύπου I (Type I) και Τύπου II (Type II) για όλες τις μεθόδους στους εξεταζόμενους μικτών επιπέδων σχεδιασμούς. Το αριστερό πλαίσιο αναφέρεται (α) στους μέσους των σφαλμάτων Τύπου I και το δεξί πλαίσιο (β) στους μέσους των σφαλμάτων Τύπου II των σχεδιασμών μικτών επιπέδων. Τα σφάλματα του SVR-RFE αναπαριστώνται από μπλε έντονες ράβδους, τα έντονα κόκκινα γράμματα αναπαριστούν τα σφάλματα για το DS, οι μωβ έντονες ράβδοι την παλινδρόμηση κατά βήματα και οι πράσινες έντονες ράβδοι αναπαριστούν την ποινή  $L_q$ .

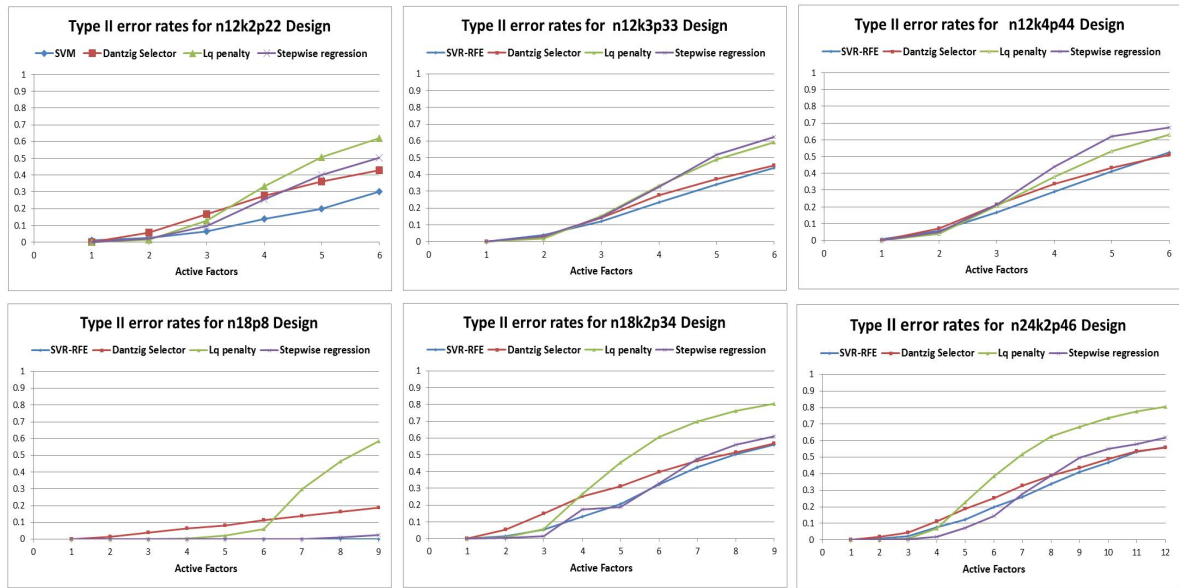
των Πινάκων B.1 και B.2 του Παραρτήματος B.

Επιπλέον, το SVR-RFE εμφανίζει εξαιρετική απόδοση όχι μόνο όταν το πραγματικό μοντέλο είναι πολύ αραιό, αλλά και όταν υπάρχει ένας μεγάλος αριθμός ενεργών παραγόντων. Είναι αξιοσημείωτο ότι επιτυγχάνει πολύ καλά ποσοστά ανεξάρτητα από το πρόσημο των εξεταζόμενων συντελεστών  $\beta$ . Το Σχήμα 6.6 δείχνει, μέσω μιας γραφικής απεικόνισης των σφαλμάτων Τύπου II, την υπεροχή του SVR-RFE έναντι των άλλων μεθόδων ανάλυσης, ειδικά όταν υπάρχει ένας μεγάλος αριθμός ενεργών παραγόντων. Συγκεκριμένα, ακόμη και αν η απόδοση των διαφόρων μεθόδων ποικίλλει λίγο στην περίπτωση ενός ή δύο ενεργών επιδράσεων, η βελτίωση που επιφέρει ο SVR-RFE αποκαλύπτεται όταν ο αριθμός των ενεργών παραγόντων αυξάνεται.

### Γενικές παρατηρήσεις

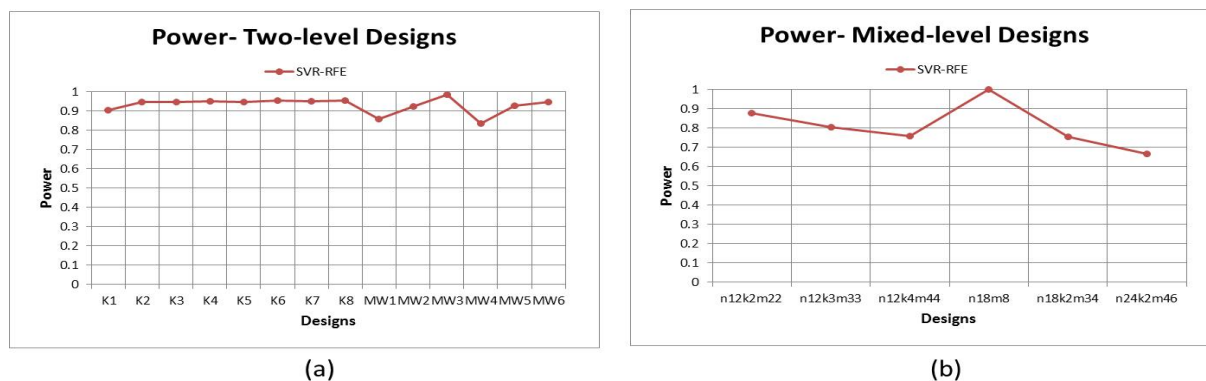
Προκειμένου να βοηθήσουμε τη διαμόρφωση των συμπερασμάτων, αναπτύσσουμε μερικές γενικές παρατηρήσεις αναφέροντας τα πιο σημαντικά σημεία της μελέτης μας:

1. Όλα τα σφάλματα Τύπου I για τον αλγόριθμο SVR-RFE είναι μικρότερα από 10% για τους σχεδιασμούς δύο επιπέδων,  $K$  και  $MW$  (Σχήμα 6.1), ομοίως τα σφάλματα Τύπου I είναι χαμηλότερα από 11% για τους σχεδιασμούς μικτών επιπέδων.



**Σχήμα 6.6:** Σφάλματα Τύπου II (Type II) για τους σχεδιασμούς μικτών επιπέδων. Κάθε πλαίσιο αναφέρεται σε κάθε ξεχωριστό σχεδιασμό. Ο άξονας  $x$  αναπαριστά τους ενεργούς παράγοντες οι οποίοι ανήκουν στο σύνολο  $\{0, 1, 2, \dots, n/2\}$ . Ο άξονας  $y$  αναπαριστά τα σφάλματα Τύπου II για τις τέσσερις διαφορετικές μεθόδους.

2. Το Σχήμα 6.7 απεικονίζει μια αξιοσημείωτα υψηλή στατιστική ισχύ για τον αλγόριθμο SVR-RFE . Ιδιαίτερα, λαμβάνοντας υπόψη έναν ΥΣ δύο επιπέδων (Σχήμα 6.7(a) ), η στατιστική ισχύς επιτυγχάνει ποσοστό υψηλότερο από 90% για τους σχεδιασμούς « $K$ », ενώ το ποσοστό αυτό φτάνει το 85% για τους σχεδιασμούς « $MW$ ».
3. Τα Σχήματα 6.3 και 6.6 δείχνουν πώς η απόδοση των μεθόδων υποβαθμίζεται όσο το επίπεδο της σποραδικότητας των επιδράσεων μειώνεται, σε σχέση με τα ποσοστά του σφάλματος Τύπου II, για τους σχεδιασμούς δύο επιπέδων και τους σχεδιασμούς μικτών επιπέδων, αντίστοιχα. Αυτά τα Σχήματα μαρτυρούν την κυριαρχία του SVR-RFE, ο οποίος επιτυγχάνει μια εξαιρετικά καλή απόδοση, ειδικά στην περίπτωση της σποραδικότητας των επιδράσεων, σε σύγκριση με τις άλλες τρεις μεθόδους.
4. Δεν υπάρχουν σημαντικές διαφορές μεταξύ των διαφόρων προσεγγίσεων όσον αφορά τα ποσοστά σφάλματος Τύπου I, δηλαδή μόνο περίπου 5%. Το γεγονός αυτό ενισχύει την κυριαρχία του αλγορίθμου SVR-RFE, δεδομένου ότι καταφέρνει να επιτύχει σημαντικά χαμηλά ποσοστά σφάλματος Τύπου II ενώ διατηρεί το σφάλμα Τύπου I στο ίδιο επίπεδο, σε σύγκριση με τις άλλες τρεις



**Σχήμα 6.7:** Τιμές της ισχύος (Power) για τον αλγόριθμο SVR-RFE. Το πλαίσιο (α) αναφέρεται στους σχεδιασμούς δύο επιπέδων, « $K$ » και « $MW$ » και το πλαίσιο (β) αναφέρεται στους σχεδιασμούς «μικτών επιπέδων». Κάθε κουκίδα αναπαριστά κάθε ξεχωριστό σχεδιασμό. Ο άξονας  $x$  αναπαριστά το όνομα του σχεδιασμού, ενώ ο άξονας  $y$  την ισχύ.

προσεγγίσεις.

5. Οι αριθμητικές συγκρίσεις με το DS, την ποινή  $L_q$  και την παλινδρόμηση κατά βήματα, υποδηλώνουν την υπεροχή του SVR-RFE. Ειδικά όσον αφορά τα σφάλματα Τύπου II, ο SVR-RFE επιτυγχάνει τα μικρότερα ποσοστά σε όλους τους εξεταζόμενους σχεδιασμούς, υποθέτοντας είτε παράγοντες με δύο είτε παράγοντες με μικτά επίπεδα.

## 6.5 Συμπεράσματα

Οι μηχανές διανυσμάτων υποστήριξης αποτελούν ένα ελκυστικό και ισχυρό εργαλείο για τη στατιστική μάθηση, με πληθώρα εφαρμογών. Ένα από τα πιο αποτελεσματικά είναι η επιλογή μεταβλητών. Σε αυτό το κεφάλαιο, προτείνουμε μια μέθοδο επιλογής μεταβλητών σε ΥΣ υποθέτοντας ένα γραμμικό μοντέλο παλινδρόμησης. Για την αποτελεσματική επιλογή των ενεργών επιδράσεων, αναπτύξαμε ένα νέο αλγόριθμο, τον SVR-RFE, χρησιμοποιώντας την ελαχιστοποίηση του διανύσματος του βάρους ως κριτήριο για τον εντοπισμό των σημαντικότερων μεταβλητών. Είναι αξιοσημείωτο ότι το SVR-RFE έχει ιδιαίτερα υψηλή απόδοση και για τους σχεδιασμούς δύο επιπέδων και για τους σχεδιασμούς μικτών επιπέδων.

Ο SVR-RFE επιτυγχάνει την επιλογή των σημαντικών επιδράσεων με υπερβολικά χαμηλά σφάλματα Τύπου I και Τύπου II ενώ εμφανίζει εξαιρετική απόδοση λαμβάνοντας υπόψη τη στατιστική ισχύ. Και τα δύο σφάλματα Τύπου I και Τύπου II είναι

σημαντικά και θα πρέπει να διατηρούνται όσο το δυνατόν χαμηλότερα. Ωστόσο, τα σφάλματα Τύπου II είναι πιο σημαντικά στην περίπτωση των ΥΣ. Ο SVR-RFE τείνει να επιτυγχάνει το χαμηλότερο ποσοστό σφάλματος Τύπου II σε σύγκριση με τον DS, την παλινδρόμηση και την ποινικοποιημένη μέθοδο  $L_q$ , διατηρώντας παράλληλα τα σφάλματα Τύπου I σε ένα αξιοσημείωτα χαμηλό επίπεδο. Αυτό είναι εξαιρετικά σημαντικό αφού οι ΥΣ χρησιμοποιούνται κυρίως για την επιλογή των παραγόντων που πρέπει να ληφθούν υπόψη για περαιτέρω έρευνα. Επομένως, αυτό που επιθυμούμε είναι υψηλές τιμές της στατιστικής ισχύος. Επιπλέον, ο SVR-RFE καταφέρνει να αποδίδει καλά σε σενάρια πραγματικής ζωής, κάτι που απεικονίζεται μέσω του παραδείγματος, όπου το υποσύνολο των χαρακτηριστικών που επιλέγεται από τον SVR-RFE παρέχει πολύ καλές επιδόσεις όσον αφορά το  $R^2$  και το προσαρμοσμένο  $R^2$  οδηγώντας στο συμπέρασμα ότι οι παράγοντες που επιλέγονται από τον SVR-RFE παρέχουν μια καλύτερη ερμηνεία των δεδομένων.



---

## Ανάλυση υπερκορεσμένων σχεδιασμών με τη μέθοδο του σίγουρου κρησαρίσματος

---

Στο παρόν κεφάλαιο παρουσιάζουμε μία μέθοδο ανάλυσης υπερκορεσμένων σχεδιασμών δύο επιπέδων. Πιο συγκεκριμένα, παρουσιάζουμε μια διαδικασία δύο σταδίων για την ανάλυση δύο επιπέδων ΥΣ υποθέτοντας ένα μοντέλο κύριων επιδράσεων, χωρίς την παρουσία αλληλεπιδράσεων. Η προτεινόμενη διαδικασία συνδυάζει τη μέθοδο του σίγουρου κρησαρίσματος (sure independence screening, SIS) με τις ποινικοποιημένες μεθόδους, όπως είναι η SCAD, η LASSO και η MC. Με τη διαδικασία αυτή επιτυγχάνουμε ταυτόχρονα τόσο την επιλογή όσο και την εκτίμηση των σημαντικών επιδράσεων. Η απόδοση της μεθόδου παρουσιάζεται μέσω διαφόρων σεναρίων προσομοίωσης όπου πραγματοποιούνται συγκρίσεις με υπάρχουσες μεθοδολογίες. Οι μέθοδοι που χρησιμοποιήσαμε στην παρούσα μελέτη είναι η παλινδρόμηση κατά βήματα σε συνδυασμό με τη SCAD, καθώς επίσης και ο DS. Τα αποτελέσματα από τη μελέτη προσομοιώσεων και από ένα παράδειγμα με πραγματικά δεδομένα αποκαλύπτουν ότι η προτεινόμενη διαδικασία μπορεί να θεωρηθεί ως ένα επωφελές εργαλείο λόγω της εξαιρετικά καλής απόδοσής του για τον εντοπισμό ενεργών παραγόντων. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [VI].

### 7.1 Ερευνητικό πρόβλημα-Λύση

Οι Fan και Lv [44] εισήγαγαν μια νέα τεχνική επιλογής μεταβλητών βασισμένη στη μάθηση των συσχετίσεων, που ονομάζεται μέθοδος του σίγουρου κρησαρίσματος (sure independence screening, SIS) καθώς και μία επέκταση της SIS, την επανα-

ληπτική μέθοδο του σίγουρου κρησαρίσματος (Iter-SIS), με σκοπό να μειώσουμε τη διάσταση του υπό εξέταση προβλήματος από υψηλή σε μέτριας κλίμακας που θα είναι δηλαδή κάτω από το μέγεθος του δείγματος.

Στην παρούσα ενότητα παρουσιάζουμε την ανάλυση υπερκορεσμένων σχεδιασμών δύο επιπέδων υποθέτοντας ένα μοντέλο κύριων επιδράσεων, χωρίς την παρουσία αλληλεπιδράσεων. Η προτεινόμενη μέθοδος επιλογής μεταβλητών βασίζεται στο συνδυασμό των διαδικασιών SIS με τα ποινικοποιημένα ελάχιστα τετράγωνα. Υποθέσαμε μόνο μοντέλα πρώτης τάξης που περιλαμβάνουν τις κύριες επιδράσεις, χωρίς την παρουσία αλληλεπιδράσεων. Η αποτελεσματικότητα της προτεινόμενης μεθοδολογίας αποδεικνύεται μέσω προσομοιώσεων θεωρώντας πολλά διαφορετικά σενάρια και περιπτώσεις έτσι ώστε να γίνει ορατή η ευρωστία της προτεινόμενης μεθόδου. Η πιο σημαντική συνεισφορά αυτής της μεθοδολογίας είναι ότι πρόκειται για μία εύκολη μέθοδο στην εφαρμογή, η οποία δίνει εξαιρετικά αποτελέσματα στην ανίχνευση των ενεργών παραγόντων σε προβλήματα που περιλαμβάνουν έναν πολύ μεγάλο αριθμό παραγόντων.

## 7.2 Ανάλυση ΥΣ με τη μέθοδο του Σίγουρου Κρησαρίσματος (SIS)

Στα πειράματα κρησαρίσματος, ο πειραματιστής μπορεί να υποθέσει ότι υπάρχουν μόνο μέχρι  $p'$  ενεργοί παράγοντες από το σύνολο των  $p$  παραγόντων που συμπεριλαμβάνονται στο πείραμα. Συνήθως, η διαδικασία είναι πιο ισχυρή όταν το  $p'$  είναι το πολύ το μισό του αριθμού των πειραματικών εκτελέσεων (effect sparsity, Marley και Woods, [101]). Έστω ότι το  $\mathbf{X}$  είναι ο  $n \times p$  πίνακας σχεδιασμού των  $+1$ 's και των  $-1$  για τους παράγοντες δύο επιπέδων και το  $\mathbf{1}_n$  είναι ένα  $n \times 1$  διάνυσμα με 1. Για ένα σχεδιασμό με  $p$  παράγοντες, το μοντέλο των κύριων επιδράσεων μπορεί να γραφεί ως

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (7.1)$$

όπου  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  είναι το διάνυσμα της απόκρισης  $n$  διαστάσεων,  $\boldsymbol{\beta}$  είναι το διάνυσμα στήλη  $(p+1) \times 1$  των άγνωστων παραμέτρων παλινδρόμησης,  $\mathbf{X}_1 = [\mathbf{1}_n, \mathbf{X}]$  είναι ο  $n \times (p+1)$  πίνακας του μοντέλου, όπου κάθε γραμμή είναι μια πειραματι-

κή εκτέλεση, συμπεριλαμβανομένης μίας στήλης με 1's. Οι στήλες  $p$  περιέχουν τα κωδικοποιημένα επίπεδα  $[-1, 1]$  των παραγόντων και  $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)$  είναι το διάνυσμα διάστασης  $n$  των τυχαίων σφαλμάτων τα οποία είναι ανεξάρτητα και κανονικά κατανομημένα (IID). Θα επικεντρωθούμε στον εντοπισμό των κύριων επιδράσεων, χωρίς να εξετάσουμε τους όρους αλληλεπιδράσεων, χρησιμοποιώντας ΥΣ. Εξ ορισμού, προκειμένου να χρησιμοποιηθούν τα συνήθη ελάχιστα τετράγωνα (OLS) η τάξη του πίνακα  $\mathbf{X}_1$  πρέπει να είναι μικρότερη ή ίση με το ελάχιστο των  $n$  και  $p + 1$ . Αυτά τα προβλήματα, ειδικά όταν η διάσταση είναι πολύ μεγαλύτερη από το μέγεθος του δείγματος, απαιτούν νέες στατιστικές μεθοδολογίες [34, 43, 44].

Τα τελευταία χρόνια [44, 59, 64, 82], έχουν προταθεί πολλές μέθοδοι ανάλυσης για την επίλυση του προαναφερθέντος προβλήματος, εστιάζοντας στα σφάλματα Τύπου I και Τύπου II. Αυτά τα σφάλματα δεν πρέπει να αντιμετωπίζονται με τον ίδιο τρόπο. Ειδικά κάτω από το σενάριο της σποραδικότητας των επιδράσεων, το επιπλέον κόστος από μεγαλύτερες τιμές των σφαλμάτων Τύπου I μπορεί να είναι ανεκτό για την επίτευξη υψηλής τιμής της ισχύος (Power). Εξάλλου, όπως σημειώνεται από τους Westfall et al. [140], τα σφάλματα Τύπου I είναι πολύ πιθανά σε μοντέλα όπου ισχύει η αρχή σποραδικότητας των επιδράσεων. Επομένως, η ισχύς είναι ακόμα πιο κρίσιμη στους ΥΣ δεδομένου ότι είναι η ικανότητα ανίχνευσης των ενεργών παραγόντων. Έτσι, όσο μεγαλύτερη η τιμή της ισχύος, τόσο μικρότερη είναι η πιθανότητα να λείπει ένας πραγματικά ενεργός παράγοντας.

Στο παρόν κεφάλαιο προτείνουμε μια εναλλακτική μέθοδο για την ανάλυση ΥΣ χρησιμοποιώντας είτε τη SIS είτε την επέκτασή της, την επαναληπτική SIS. Αυτή η διαδικασία είναι μια μέθοδος ανίχνευσης που χρησιμοποιεί μόνο μία περιθωριακή σχέση μεταξύ των επεξηγηματικών μεταβλητών και της απόκρισης, έτσι ώστε να ανιχνεύσει τις σημαντικές μεταβλητές. Όταν ο αριθμός  $d$  των επιλεγμένων μεταβλητών είναι αρκετά μεγάλος, έχει υψηλή πιθανότητα επιλογής όλων των σημαντικών επεξηγηματικών μεταβλητών. Για την αξιολόγηση και τη σύγκριση της SIS με τις υπόλοιπες μεθόδους ανάλυσης χρησιμοποιήθηκαν τα ακόλουθα μέτρα: (i) Η ισχύς, δηλαδή το ποσοστό των ενεργών κύριων επιδράσεων που εντοπίστηκαν με επιτυχία (ii) το ποσοστό σφάλματος Τύπου I, δηλαδή το ποσοστό των μη ενεργών κύριων επιδράσεων που έχουν δηλωθεί σωστά ως ενεργές και (iii) το ποσοστό σφάλματος Τύπου II, δηλαδή το κόστος δήλωσης μιας ενεργής επίδρασης ως μη ενεργής. Κυρί-

ως επικεντρωθήκαμε στη σύγκριση των εξεταζόμενων μεθόδων ως προς την Ισχύ του ελέγχου, η οποία ισούται με (1-Τύπου II), καθώς αυτό το μέτρο είναι και το πιο κρίσιμο στους ΥΣ. Επιπλέον, εξετάσαμε τις συγκρίσεις ως προς το μέσο συνολικό σφάλμα (που περιγράφεται στο κεφάλαιο 7.4), προκειμένου να ελέγξουμε ταυτόχρονα και τα δύο σφάλματα, Τύπου I και Τύπου II.

### 7.2.1 Ποινικοποιημένα ελάχιστα τετράγωνα

Με χαρακτηριστικά που έχουν επιλεγεί κατά προσέγγιση από την SIS, η μετέπειτα επιλογή μεταβλητών και η εκτίμηση των παραμέτρων μπορούν να διεξαχθούν ταυτόχρονα χρησιμοποιώντας μια ποινικοποιημένη μέθοδο. Χωρίς απώλεια της γενικότητας υποθέτουμε ότι τα  $X_1, \dots, X_d$  είναι τα χαρακτηριστικά που επιλέγονται από τη SIS. Οι Fan και Li [42] πρότειναν μια κλάση μεθόδων επιλογής μεταβλητών μέσω της μη κοίλης ποινικοποιημένης πιθανοφάνειας. Οι Li και Lin [82] εισήγαγαν μια επέκταση αυτής της μεθόδου στα μη κοίλα ποινικοποιημένα τετράγωνα εστιάζοντας στην κατάσταση όπου ο πίνακας σχεδιασμού δεν είναι πλήρης. Επαναπροσδιορίζοντας τα  $\beta = \beta_1, \dots, \beta_d$ , οι Fan και Li [42] θεώρησαν μια μορφή ποινικοποιημένων ελαχίστων τετραγώνων που μπορούν να γραφτούν ως

$$Q(\beta) = \frac{1}{2n} \sum_{i=1}^n (Y_i - x_i^T \beta)^2 + \sum_{j=1}^d p_{\lambda_n}(|\beta_j|), \quad (7.2)$$

όπου  $p_{\lambda_n}(\cdot)$  είναι μια συνάρτηση ποινής και  $\lambda_n > 0$  είναι μια ρυθμιστική παράμετρος που ελέγχει την πολυπλοκότητα του μοντέλου. Για παράδειγμα, θα θέλαμε να επιλέξουμε να μην ποινικοποιήσουμε τους συντελεστές για ορισμένους σημαντικούς παράγοντες πρόβλεψης που θέλουμε να διατηρήσουμε στο μοντέλο. Αυτές οι ρυθμιστικές παράμετροι  $\lambda_n$  μπορούν να επιλεγούν από προσεγγίσεις οδηγούμενες από δεδομένα, όπως π.χ. η γενικευμένη διασταυρωμένη επικύρωση [29] (GCV) ή ακόμη και με την ελαχιστοποίηση των κριτηρίων πληροφορίας Akaike (AIC) ή Bayesian (BIC). Για να ελαχιστοποιηθεί το πρόβλημα των ποινικοποιημένων ελαχίστων τετραγώνων της εξίσωσης (9.3), οι Fan και Li [42] πρότειναν την τοπική τετραγωνική προσέγγιση που είναι ένας ενοποιημένος και αποτελεσματικός αλγόριθμος για τη βελτιστοποίηση της ποινικοποιημένης πιθανοφάνειας. Στην παρούσα εργασία επιλέξαμε  $\lambda_n = \lambda$ , διαλέγοντας το  $\lambda_n$  χρησιμοποιώντας και τα δύο AIC και BIC, η

σύγκριση των οποίων παρουσιάζεται στην ενότητα 7.4.2, Μέρος I. Ένας εναλλακτικός τρόπος για την ελαχιστοποίηση του προβλήματος των ποινικοποιημένων ελαχίστων τετραγώνων της εξίσωσης (9.3) είναι η τοπική γραμμική προσέγγιση που προτάθηκε από τους Zou και Li [160]. Αυτή η προσέγγιση οδηγεί σε ένα κυρτό πρόβλημα που μπορεί να λυθεί με τη χρήση του αλγορίθμου LARS [39].

Πολλά κριτήρια επιλογής μεταβλητών μπορούν να εξαχθούν από την προσέγγιση των ποινικοποιημένων ελαχίστων τετραγώνων, με διαφορετικές συναρτήσεις ποινών. Οι πιο κοινές συναρτήσεις ποινής περιλαμβάνουν την ποινή LASSO [109, 135]  $p_\lambda(|\beta|) = l|\beta|$ , την ποινή SCAD, και την minimax μη κοίλη ποινή [156]. Οι Fan και Li [42] εστιάζουν στις μη κοίλες συναρτήσεις ποινής προκειμένου να διασφαλίσουν ότι η συνάρτηση ποινής  $p_{\lambda_n}(\cdot)$  ικανοποιεί κάποιες συγκεκριμένες συνθήκες, έτσι ώστε οι προκύπτοντες εκτιμητές να κατέχουν αρκετές επιθυμητές ιδιότητες [42], όπως αμεροληψία, σποραδικότητα (να αποφέρουν αραιές λύσεις στις οποίες κάποιοι από τους συντελεστές μπορεί να τεθούν ίσοι με το μηδέν) και συνέχεια. Έδειξαν λοιπόν ότι η ποινή  $L_1$ , η οποία οδηγεί στη μέθοδο του απόλυτου τελεστή συρρίκνωσης και επιλογής (Least Absolut Shrinkage and Selection Operator, LASSO) δημιουργεί μεροληπτικές εκτιμήσεις. Ωστόσο, ικανοποιεί τη συνθήκη της σποραδικότητας και της συνέχειας του παραγόμενου εκτιμητή. Από την άλλη πλευρά, η ποινή  $L_p$ , με  $0 \leq p < 1$  δεν ικανοποιεί τη συνθήκη της συνέχειας και η ποινή  $L_p$ , με  $p > 1$  δεν ικανοποιεί τη συνθήκη της σποραδικότητας.

Η ποινή ομαλά περικομμένης απόλυτης απόκλισης (smoothly clipped absolute deviation penalty, SCAD) αποτελεί μια συνεχή διαφορίσιμη ποινή, η πρώτη παράγωγος της οποίας ορίζεται ως

$$p'_\lambda(\beta) = \lambda \left\{ I(\beta \leq \lambda) + \frac{(\alpha\lambda - \beta)_+}{(\alpha - 1)\lambda} I(\beta > \lambda) \right\}, \quad (7.3)$$

όπου  $p_\lambda(0) = 0$ , για  $\alpha > 2$ ,  $\beta > 0$  και  $I(\cdot)$  μία δείκτρια συνάρτηση. Οι Fan και Li (2001) πρότειναν την επιλογή του  $\alpha = 3.7$  η οποία φαίνεται να αποδίδει αρκετά ικανοποιητικά σε πολλά προβλήματα επιλογής μεταβλητών.

Για να βρεθεί η λύση των ποινικοποιημένων ελαχίστων τετραγώνων και για να προσεγγιστεί τοπικά η ποινή SCAD, οι Fan και Li [42] χρησιμοποίησαν την επαναληπτική παλινδρόμηση κορυφογραμμής επιλέγοντας μια αρχική τιμή για τους

άγνωστους συντελεστές. Δοθείσας μιας αρχικής τιμής  $\beta^{(0)}$  κοντά στην πραγματική  $\beta$ , όταν  $\beta_j^{(0)}$  δεν είναι πολύ κοντά στο 0, η ποινή  $p_\lambda(|\beta_j|)$  μπορεί να προσεγγιστεί τοπικά από την ακόλουθη τετραγωνική συνάρτηση

$$[p_\lambda(|\beta_j|)]' = p'_\lambda(\beta_j) \text{sgn}(\beta_j) \approx \{p'_\lambda(|\beta_j^{(0)}|)/|\beta_j^{(0)}|\} \beta_j, \quad (7.4)$$

αλλιώς, όταν  $\beta_j^{(0)}$  είναι πολύ κοντά στο 0, θέτουμε  $\hat{\beta}_j = 0$ . Με την τοπική τετραγωνική προσέγγιση, χρησιμοποιώντας την αρχική τιμή  $\beta^{(0)}$ , παίρνουμε τη λύση για τα ποινικοποιημένα ελάχιστα τετράγωνα με επαναληπτικό υπολογισμό της ακόλουθης παλινδρόμησης κορυφογραμμής

$$\beta^{(1)} = \{\mathbf{X}^T \mathbf{X} + n \sum_{\lambda} (\beta^{(0)})\}^{-1} \mathbf{X}^T \mathbf{y}, \quad (7.5)$$

όπου  $\sum_{\lambda} (\beta^{(0)}) = \text{diag}\{p'_\lambda|\beta_1^{(0)}|/|\beta_1^{(0)}|, \dots, p'_\lambda|\beta_d^{(0)}|/|\beta_d^{(0)}|\}$ . Ο τύπος του τυπικού σφάλματος από την επαναληπτική παλινδρόμηση κορυφογραμμής λαμβάνεται ως

$$\text{cov}(\hat{\beta}) = \{\mathbf{X}^T \mathbf{X} + n \sum_{\lambda} (\hat{\beta})\}^{-1} \mathbf{X}^T \mathbf{X} \{\mathbf{X}^T \mathbf{X} + n \sum_{\lambda} (\hat{\beta})\}^{-1}. \quad (7.6)$$

Ο τύπος (7.7) αναφέρεται μόνο στα μη-μηδενικά συστατικά του  $\hat{\beta}$ . Η διαδικασία για την εξεύρεση της λύσης των ποινικοποιημένων ελαχίστων τετραγώνων και η τοπική προσέγγιση της ποινής SCAD, είναι η χρήση μιας επαναληπτικής παλινδρόμησης κορυφογραμμής επιλέγοντας μία αρχική τιμή για τους άγνωστους συντελεστές. Η συνάρτηση της ποινής μπορεί να προσεγγιστεί τοπικά από την τετραγωνική συνάρτηση της εξίσωσης (7.4). Εμείς αποκτήσαμε τη λύση για τα ποινικοποιημένα ελάχιστα τετράγωνα υπολογίζοντας με επαναληπτικό τρόπο την παλινδρόμηση κορυφογραμμής της εξίσωσης (7.5) και στη συνέχεια υπολογίζοντας τον τύπο του τυπικού σφάλματος της εξίσωσης (7.7). Η διαδικασία που περιγράφηκε προηγουμένως μπορεί να εφαρμοστεί εύκολα μέσω πολλών στατιστικών πακέτων, όπως η R ή η MATLAB.

Πρόσφατα, ο Zhang [156] πρότεινε την ποινή  $MC_+$  που αποτελεί μια γρήγορη, συνεχή, σχεδόν αμερόληπτη ποινικοποιημένη μέθοδο επιλογής μεταβλητών στη γραμμική παλινδρόμηση για δεδομένα υψηλής διάστασης. Η παράγωγος της ελά-

χιστης κοίλης ποινής (minimum concavity penalty, MCP) ορίζεται ως

$$p'_\lambda(\beta) = (\lambda - \beta/a)_+. \quad (7.7)$$

Οι συναρτήσεις ποινής SCAD και MC, σε αντίθεση με LASSO, έχουν επίπεδες ουρές κάτι που είναι θεμελιώδους σημασίας για τη μείωση της μεροληψίας λόγω της ποινικοποίησης [3, 42]. Η ιδιότητα oracle έχει δημιουργηθεί για μια κατηγορία συναρτήσεων ποινής συμπεριλαμβανομένης της SCAD (Fan και Li [42]) και της LASSO που προτάθηκε από τον [159], όταν το  $d$  είναι πεπερασμένο. Για μερικές περαιτέρω θεωρητικές μελέτες παραπέμπουμε τον ενδιαφερόμενο αναγνώστη στους Zhang και Huang [157] και Zhang [156]. Στη μελέτη μας χρησιμοποιήσαμε τις ποινές: LASSO, SCAD και μία ποινή από τα στοιχεία της  $MC_+$ , την ελάχιστη-μέγιστη (minimax) κοίλη ποινή (minimax concave penalty, MCP).

### 7.2.2 Μέθοδος Σίγουρου Κρησαρίσματος

Οι Fan και Lv [44] πρότειναν μία διαδικασία κρησαρίσματος μεταβλητών χρησιμοποιώντας την παλινδρόμηση συνιστωσών προκειμένου να εντοπιστούν οι σημαντικοί παράγοντες, βασιζόμενοι στις μεμονωμένες συσχετίσεις τους με τη μεταβλητή απόκρισης. Δημιούργησαν ένα  $p \times 1$  διάνυσμα ως  $\omega$  από τις περιθωριακές συσχετίσεις μεταξύ των κύριων επιδράσεων και της μεταβλητής απόκρισης. Ο  $n \times p$  πίνακας σχεδιασμού  $\mathbf{X}$  είναι κανονικοποιημένος, δηλαδή τέτοιος ώστε κάθε στήλη του να έχει μηδενική μέση τιμή και τυπική απόκλιση ίση με ένα. Αυτό το διάνυσμα  $\omega = \omega_1, \dots, \omega_p$  δίνεται ως

$$\omega = \mathbf{X}'\mathbf{y}, \quad (7.8)$$

όπου  $\mathbf{y}$  είναι ένα  $n \times 1$  διάνυσμα απόκρισης,  $\mathbf{X}$  είναι ο  $n \times p$  κανονικοποιημένος πίνακας του μοντέλου και  $p$  είναι ο αριθμός των παραγόντων.

Για ένα δεδομένο  $\gamma \in (0, 1)$ , οι  $p$  παραγοντικές συνιστώσες του  $\omega$ , ταξινομούνται κατά απόλυτη τιμή μεγέθους σε φθίνουσα σειρά και σχηματίζουν το ακόλουθο σύνολο παραγόντων

$$M_\gamma = \{1 \leq i \leq p : |\omega_i| \text{ is among the first } [\gamma n] \text{ largest of all}\} \quad (7.9)$$

όπου  $[\gamma n]$  δηλώνει το ακέραιο μέρος του  $\gamma n$ . Αυτή η τιμή ενεργεί ως ένας τρόπος συρρίκνωσης του αρχικού μοντέλου  $1, \dots, p$  σε ένα υπομοντέλο με μέγεθος  $d = [\gamma n] < n$ . Μια τέτοια διαδικασία μάθησης συσχετίσεων κατατάσσει τα σημαντικά χαρακτηριστικά σύμφωνα με τις περιθωριακές συσχετίσεις τους με τη μεταβλητή απόκρισης και αγνοεί εκείνα που έχουν αδύναμες οριακές συσχετίσεις με τη μεταβλητή απόκρισης. Οι Fan και Lv [44] ονόμασαν αυτή τη μέθοδο που βασίζεται στις συσχετίσεις, ως Μέθοδο του Σίγουρου Κρησαρίσματος (Sure Independence Screening, SIS), δεδομένου ότι κάθε χαρακτηριστικό χρησιμοποιείται ανεξάρτητα ως ανεξάρτητη μεταβλητή προκειμένου να αποφασίσουμε πώς επηρεάζει την πρόβλεψη της μεταβλητής απόκρισης. Η έννοια αυτή εφαρμόζεται σε γενικευμένα γραμμικά μοντέλα, σε προβλήματα ταξινόμησης καθώς και στη μη παραμετρική μάθηση όπως μελετήθηκε από τους Ravikumar et al. [113].

Τα γραμμικά μοντέλα με περισσότερους από  $n$  παράγοντες δεν μπορούν να ανιχνεύσουν όλες τις παραμέτρους μόνο με  $n$  σημεία δεδομένων στο δείγμα. Επομένως, εφαρμόσαμε τη μέθοδο SIS επιλέγοντας  $d = n - 1$ , λαμβάνοντας υπόψη το γεγονός ότι η SIS προτάθηκε για να μειώσει τη διάσταση ενός προβλήματος από το υψηλό  $p$ , κάτω από το μέγεθος του δείγματος,  $n$ . Οι συνθήκες υπό τις οποίες η ιδιότητα του σίγουρου κρησαρίσματος

$$P(M_* \subset M_\gamma) \rightarrow 1 \quad \text{καθώς} \quad n \rightarrow \infty, \quad (7.10)$$

ισχύει για τη SIS, όπου  $M_*$  είναι το πραγματικό μοντέλο και  $M_\gamma$  το τελικό μοντέλο, απα-ντώνται στο Θεώρημα 1 των Fan και Lv (2008).

### **Επαναληπτική Μέθοδος Σίγουρου Κρησαρίσματος**

Ωστόσο, σε μερικές περιπτώσεις η μεθοδολογία SIS ενδέχεται να αποτυγχάνει. Για παράδειγμα, αν ένας παράγοντας είναι οριακά ασυσχέτιστος, αλλά συσχετίζεται από κοινού με τη μεταβλητή απόκρισης, δεν μπορεί να επιλεγεί από τη SIS και συνεπώς αυτός ο σημαντικός παράγοντας θα αφαιρεθεί από το πείραμά μας από το πρώτο στάδιο. Ένα άλλο μειονέκτημα της SIS εμφανίζεται εάν ένας παράγοντας είναι από κοινού ασυσχέτιστος με την απόκριση αλλά έχει υψηλότερη οριακή συσχέτιση με τη μεταβλητή απόκρισης σε σύγκριση με ορισμένους σημαντικούς παράγοντες.



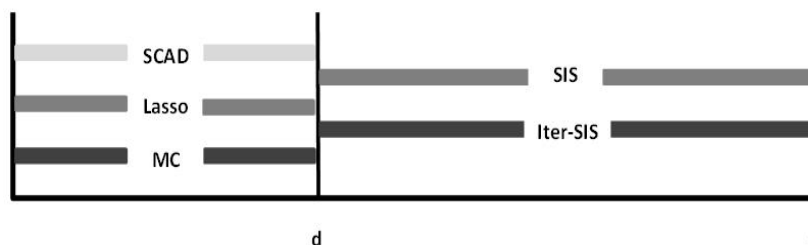
Τότε, ο μη σημαντικός παράγοντας κατατάσσεται υψηλότερα από τη διαδικασία SIS. Υπάρχει επίσης το ενδεχόμενο συγγραμμικότητας μεταξύ των παραγόντων στο τελικό επιλεγμένο μοντέλο, δεδομένου ότι ελήφθησαν υπόψη ένας προς έναν κατά τη διάρκεια της διαδικασίας κρησαρίσματος. Μετά την εκτέλεση της SIS, προκειμένου να μειωθεί το αρχικό σύνολο όλων των πιθανών παραγόντων σ' ένα σύνολο με μέγεθος μικρότερο από  $n$ , μπορεί να εφαρμοστεί η SCAD για την εκτίμηση των συντελεστών και να εξαλείψει εκείνους που δεν είναι σημαντικοί, με την παρουσία των άλλων παραγόντων, δηλαδή τις εκτιμήσεις των συντελεστών που είναι κοντά στο μηδέν. Αυτό μπορεί να μειώσει περαιτέρω τον αριθμό των σημαντικών παραγόντων παρέχοντας τις εκτιμήσεις που είναι απαραίτητο για να γραφεί το προσαρμοσμένο μοντέλο. Αυτός είναι ο λόγος για τον οποίο η SCAD μπορεί να οδηγήσει στην εξάλειψη κάποιου προβλήματος συγγραμμικότητας που παράγεται μόνο με τη μέθοδο SIS. Προσπαθώντας να ξεπεράσουν αυτές τις δυσκολίες, οι Fan και Lv [44] εισήγαγαν μία επαναληπτική διαδικασία της αρχικής τεχνικής SIS που ονομάζεται επαναληπτική μέθοδος σίγουρου κρησαρίσματος (Iter-SIS). Η Iter-SIS λειτουργεί ως εξής:

- Πρώτον, επιλέγεται ένα υποσύνολο  $A_1 = \{X_{i_1}, \dots, X_{i_{m_1}}\}$  από  $m_1$  μεταβλητές χρησιμοποιώντας τη μέθοδο SIS σε συνδυασμό με μια μέθοδο ποινικοποιημένης πιθανοφάνειας, για παράδειγμα τη SIS-SCAD ή τη SIS-LASSO, έτσι ώστε να επιλέξουμε ένα υποσύνολο  $\hat{M}_1$  αυτών των δεικτών. Στη συνέχεια δημιουργείται ένα  $n$  διαστάσεων διάνυσμα των υπολοίπων με την παλινδρόμηση της απόκρισης  $Y$  πάνω στις μεταβλητές  $X_{i_1}, \dots, X_{i_{m_1}}$ .
- Στην επόμενη επανάληψη, εφαρμόζεται αυτή η μέθοδος σε  $m - m_1$  παράγοντες που χρησιμοποιούν τα υπόλοιπα ως νέες μεταβλητές απόκρισης. Αυτό έχει ως αποτέλεσμα ένα υποσύνολο  $m_2$  μεταβλητών  $A_2 = \{X_{j_1}, \dots, X_{j_{m_2}}\}$ .
- Στα επόμενα βήματα, η SIS σε συνδυασμό με μια συνάρτηση ποινής εφαρμόζεται στους υπόλοιπους  $m - m_1 - m_2$  παράγοντες χρησιμοποιώντας το νέο σύνολο των υπολοίπων ως νέα μεταβλητή απόκρισης.
- Αφού εκτελέσουμε αρκετές επαναλήψεις παίρνουμε τα  $A_1, \dots, A_l$  τα οποία είναι διαφορετικά υποσύνολα, η ένωση των οποίων έχει μέγεθος  $d$  μικρότερο από  $n$ . Εφαρμόζοντας μια ποινικοποιημένη μέθοδο στην ένωση των  $A_i$  προκύπτει ένα μέτριο μοντέλο το οποίο είναι πολύ κοντά στο πραγματικό.

Σημειώνουμε ότι χρησιμοποιώντας τα υπόλοιπα από την πρώτη επανάληψη στο  $\{X_1, \dots, X_m\} \setminus A_1$  μπορεί πιθανώς να βοηθήσει στην αναγνώριση αυτών των παραγόντων που έχουν υψηλή συσχέτιση με τη μεταβλητή απόκρισης μόνο από την παρουσία ορισμένων παραγόντων από το σύνολο των  $X_{i_1}, \dots, X_{i_{m_1}}$ . Αυτό συμβαίνει επειδή αυτοί οι παράγοντες είναι πιθανόν να είναι συσχετισμένοι με τα υπόλοιπα από την πρώτη παλινδρόμηση. Έτσι, θα ενσωματωθούν στο σύνολο  $A_2$ .

### 7.3 Νέα μέθοδος επιλογής μεταβλητών που βασίζεται στη SIS

Για το πρόβλημα της επιλογής μεταβλητών σε ΥΣ προτείνουμε, στο πρώτο στάδιο, την εφαρμογή της μεθόδου του σίγουρου κρησαρίσματος, είτε της SIS είτε της Iter-SIS, για να μειώσουμε τη διάσταση του προβλήματος από τις αρχικές  $p$  επιδράσεις σε μία μέτρια κλίμακα, έστω  $d$ , κάτω από το μέγεθος του δείγματος.



**Σχήμα 7.1:** Μέθοδοι επιλογής μοντέλων σε ΥΣ. Η SIS ή η Iter-SIS χρησιμοποιού-νται για να μειώσουν τη διάσταση του προβλήματος από τις αρχικές  $p$  επιδράσεις σε μία μέτρια κλίμακα, έστω  $d$ , κάτω από το μέγεθος του δείγματος. Στο δεύτερο στάδιο για να υπολογίσουμε τους συντελεστές των επιδράσεων χρησιμοποιήσαμε μία μέθοδο ποινής όπως είναι η SCAD, η LASSO και η MC.

Στο δεύτερο στάδιο χρησιμοποιήσαμε μια ποινικοποιημένη μέθοδο όπως είναι η SCAD, η LASSO και η MC για την εκτίμηση των συντελεστών των επιδράσεων, οδηγώντας τους μη σημαντικούς συντελεστές του μοντέλου στο μηδέν. Αυτό μπορεί να μειώσει περαιτέρω τον αριθμό των σημαντικών παραγόντων, και με αυτό τον τρόπο οδηγούμαστε μόνο στις εκτιμήσεις που είναι απαραίτητες για τη διαμόρφωση του τελικού μοντέλου. Μία γραφική απεικόνιση αυτών των μεθόδων παρουσιάζεται στο Σχήμα 7.1. Αυτή η μέθοδος η οποία ουσιαστικά είναι ο συνδυασμός της SIS ακολουθούμενη από τις ποινές SCAD, LASSO ή MC, καλείται ως SIS-SCAD,

SIS-Lasso και SIS-MC, αντίστοιχα. Ομοίως, η επαναληπτική SIS ακολουθούμενη από τις ποινές SCAD, τη LASSO και την MC καλούνται ως Iter-SIS-SCAD, Iter-SIS-Lasso και Iter-SIS-MC. Σε κάθε ενδιάμεσο στάδιο χρησιμοποιήσαμε είτε το κριτήριο AIC είτε το κριτήριο BIC για το συντονισμό της παραμέτρου κανονικοποίησης των ποινικοποιημένων υποπροβλημάτων πιθανοφάνειας και του τελικού μοντέλου που επιλέχθηκε από τις διαδικασίες που βασίζονται στη SIS, ενώ ρυθμίσαμε την παράμετρο  $\alpha$  να είναι ίση με 3.7 [42] στην περίπτωση της ποινής SCAD. Το πακέτο "SIS" της R χρησιμοποιήθηκε για την εφαρμογή των μεθόδων SIS και Iter-SIS. Αυτό το πακέτο απαιτεί ως είσοδο τον αριθμό σημαντικών παραγόντων που πρέπει να επιστρέψουν από τις διαδικασίες SIS και Iter-SIS. Τελικά επιλέξαμε  $n - 1$ , αλλά ένας χρήστης θα μπορούσε ενδεχομένως να επιλέξει το αρχικό υποσύνολο των παραγόντων να είναι κάτι μεγαλύτερο από  $n$ , εάν θεωρείται ένας μεγάλος αριθμός αρχικών παραγόντων. Αυτός ο αριθμός μπορεί να μειωθεί περαιτέρω από τη μέθοδο της ποινής που χρησιμοποιήθηκε στο δεύτερο στάδιο. Οι μέθοδοι συρρίκνωσης επιτυγχάνουν την επιλογή των μεταβλητών με τη μεροληψία, ή τη συρρίκνωση, των εκτιμώμενων συντελεστών παλινδρόμησης στο μηδέν. Γενικά, οι μεροληπτικοί εκτιμητές έχουν συνήθως χαμηλότερη διασπορά από τους εκτιμητές που λαμβάνονται από τα συνήθη ελάχιστα τετράγωνα. Αυτή είναι η πρώτη φορά που η SIS και η επέκτασή της, εφαρμόστηκε σε ΥΣ. Αν και μπορεί να φαίνεται μια απλή ιδέα, η εφαρμογή της για την ανάλυση ΥΣ οδηγεί σε εξαιρετικά ισχυρά αποτελέσματα, ιδίως όσον αφορά τη στατιστική ισχύ του ελέγχου.

## 7.4 Αξιολόγηση της προτεινόμενης μεθόδου

Για να δούμε την εφαρμογή των μεθόδων επιλογής χαρακτηριστικών με τις μεθόδους που βασίζονται στη SIS στους ΥΣ, μελετήσαμε τις επιδόσεις τους τόσο σε προσομοιωμένα όσο και σε πραγματικά δεδομένα. Στο πρώτο μέρος αυτής της ενότητας παρέχουμε πολλές συγκρίσεις μεταξύ των μεθόδων SIS με τη χρήση διαφορετικών ποινών για να αποφασίσουμε ποιος συνδυασμός είναι ο καλύτερος, ενώ στο δεύτερο μέρος παρουσιάζουμε διάφορες συγκρίσεις με κάποιες άλλες πολύ γνωστές μεθόδους. Στους πίνακες που παρέχονται στο Παράρτημα Γ, παρουσιάζουμε διάφορα μέτρα απόδοσης, τα οποία είναι όλα με βάση 1000 προσομοιώσεις.

### 7.4.1 Μελέτη προσομοιώσεων

Η απόδοση της προτεινόμενης μεθόδου που βασίζεται στη SIS εξετάζεται στο παρόν κεφάλαιο μέσω πειραμάτων προσομοίωσης. Οι επιδόσεις των διαδικασιών SIS συγκρίθηκαν με τέσσερις διαφορετικές υπάρχουσες προσεγγίσεις που προτείνονται στη βιβλιογραφία:

- (i) την *stepwise-SCAD*, την προσέγγιση των ποινικοποιημένων ελαχίστων τετραγώνων που οφείλεται στους Li και Lin [82].
- (ii) τον επιλογέα Dantzig (DS), μία μέθοδο επιλογής χαρακτηριστικών για ΥΣ που πρότειναν οι Phoa et al. [112] (όπως εφαρμόζεται στους [101, 139]). Σημειώνουμε ότι επιλέξαμε την τιμή της παραμέτρου κατωφλίου βασιζόμενοι στην τροποποίηση του AIC που υιοθετήθηκε από τους Phoa et al. [112].
- (iii) την ποινή  $L_q$  LASSO. Η  $L_q$  LASSO ισοδυναμεί με τη LAD LASSO [138] όταν  $q = 1$  και με την SQR LASSO [7] όταν το  $q = 2$ , αντίστοιχα. Η παράμετρος κανονικοποίησης της SQR LASSO είναι  $C\sqrt{\log(d)/n}$ , όπου  $C$  είναι κάποια σταθερά και  $d$  είναι η διάσταση του σχεδιασμού. Επιλέξαμε χειροκίνητα την ελάχιστη παράμετρο κανονικοποίησης ως  $\sqrt{\log(d)/n}$  και για την LAD και για την SQR LASSO.

### Σχέδιο προσομοιώσεων

Για να συγκρίνουμε την απόδοση της προτεινόμενης μεθοδολογίας με αυτή των τεσσάρων μεθόδων προσομοίωσης, δημιουργήσαμε δεδομένα από το γραμμικό μοντέλο που περιλαμβάνει μόνο τις κύριες επιδράσεις  $y = \mathbf{X}\beta + \epsilon$ , όπου  $\mathbf{X}$  είναι ένας  $n \times p$  πίνακας ΥΣ και  $\epsilon_i \sim N(0, 1)$ ,  $i = 1, 2, \dots, n$  το τυχαίο σφάλμα για κάθε παρατήρηση  $y_i$ . Χρησιμοποιήσαμε ως πίνακα σχεδιασμού, οκτώ ελάχιστους-βέλτιστους κυκλικούς (minimax-optimal cyclic) ΥΣ που κατασκευάστηκαν από τους Koukouvinos et al. [76] και έξι ΥΣ, τρεις  $E(s^2)$ -βέλτιστους ΥΣ και τρεις  $D$ -βέλτιστους ΥΣ, που κατασκευάστηκαν από τους Marley και Woods [101] (Πίνακας 7.1).

Θεωρούμε τα ακόλουθα δύο μοντέλα για τους συντελεστές  $\beta$ :

- οι ενεργοί παράγοντες παράγονται από την κατανομή  $N(\mu, 1)$ , οι απόλυτες τιμές της  $\mu$  ισούνται με 2 και τα άλλα, μη ενεργά συστατικά του  $\beta$  ισούνται με

**Πίνακας 7.1:** Σχεδιασμοί που εξετάζονται στη μελέτη προσομοίωσης

ΥΣ	$d$ Παράγοντες	$n$ Πειραματικές εκτελέσεις	
Koukouninos <i>et al.</i>	14	8	$E(s^2)$ -βέλτιστος
Koukouninos <i>et al.</i>	18	10	$E(s^2)$ -βέλτιστος
Koukouninos <i>et al.</i>	22	12	$E(s^2)$ -βέλτιστος
Koukouninos <i>et al.</i>	26	14	$E(s^2)$ -βέλτιστος
Koukouninos <i>et al.</i>	30	16	$E(s^2)$ -βέλτιστος
Koukouninos <i>et al.</i>	34	18	$E(s^2)$ -βέλτιστος
Koukouninos <i>et al.</i>	38	20	$E(s^2)$ -βέλτιστος
Koukouninos <i>et al.</i>	42	22	$E(s^2)$ -βέλτιστος
Marley and Woods	26	12	$E(s^2)$ -βέλτιστος
Marley and Woods	24	14	$E(s^2)$ -βέλτιστος
Marley and Woods	22	18	$E(s^2)$ -βέλτιστος
Marley and Woods	26	12	$D$ -βέλτιστος
Marley and Woods	24	14	$D$ -βέλτιστος
Marley and Woods	22	18	$D$ -βέλτιστος

0 (Μοντέλο 1, **Model 1**).

- οι ενεργοί παράγοντες παράγονται από την κατανομή  $N(\mu, 1)$ , οι απόλυτες τιμές της  $\mu$  ισούνται 5 και τα άλλα, μη ενεργά συστατικά του  $\beta$  ισούνται με 0 (Μοντέλο 2, **Model 2**).

Κάθε προσομοίωση εκτελέστηκε 1000 φορές και οι ενεργοί παράγοντες  $p$  επιλέχθηκαν τυχαία από το σύνολο των  $\{1, \dots, p\}$  στηλών του  $X$  σύμφωνα με τα ακόλουθα σενάρια:

- Σενάριο I (σποραδικότητα των επιδράσεων): για ενεργούς παράγοντες που κυμαίνονται από  $p = 1, \dots, [n/6]$ .
- Σενάριο II (ενδιάμεση πολυπλοκότητα): για ενεργούς παράγοντες  $p = [n/6] + 1, \dots, [n/4]$ .
- Σενάριο III (μεγάλος αριθμός επιδράσεων): για ενεργούς παράγοντες  $p = [n/4] + 1, \dots, [n/2]$ .

Στους Πίνακες Γ.1-Γ.30 που παρέχονται στο παράρτημα Γ, συγκρίνονται οι επιδόσεις των μεθόδων. Σε αυτούς τους πίνακες ο όρος «Type I» αντιπροσωπεύει το σφάλμα Τύπου I (το κόστος δήλωσης μιας αδρανούς επίδρασης ως ενεργής), το «Type II» αντιπροσωπεύει το σφάλμα Τύπου II (το κόστος δήλωσης μιας ενεργής επίδρασης ως αδρανούς), η «Power» αντιπροσωπεύει τη στατιστική ισχύ και το «MS» τον

μέσο αριθμό παραγόντων που περιέχονται στα τελικά μοντέλα. Οι προσομοιώσεις που πραγματοποιήσαμε στην R πραγματοποιήθηκαν χρησιμοποιώντας τα πακέτα «flare» [84] και «SIS» [41], η οποία απαιτεί ως είσοδο την επιστροφή του αριθμού των σημαντικών παραγόντων όπου εμείς επιλέξαμε να είναι  $d = n - 1$ .

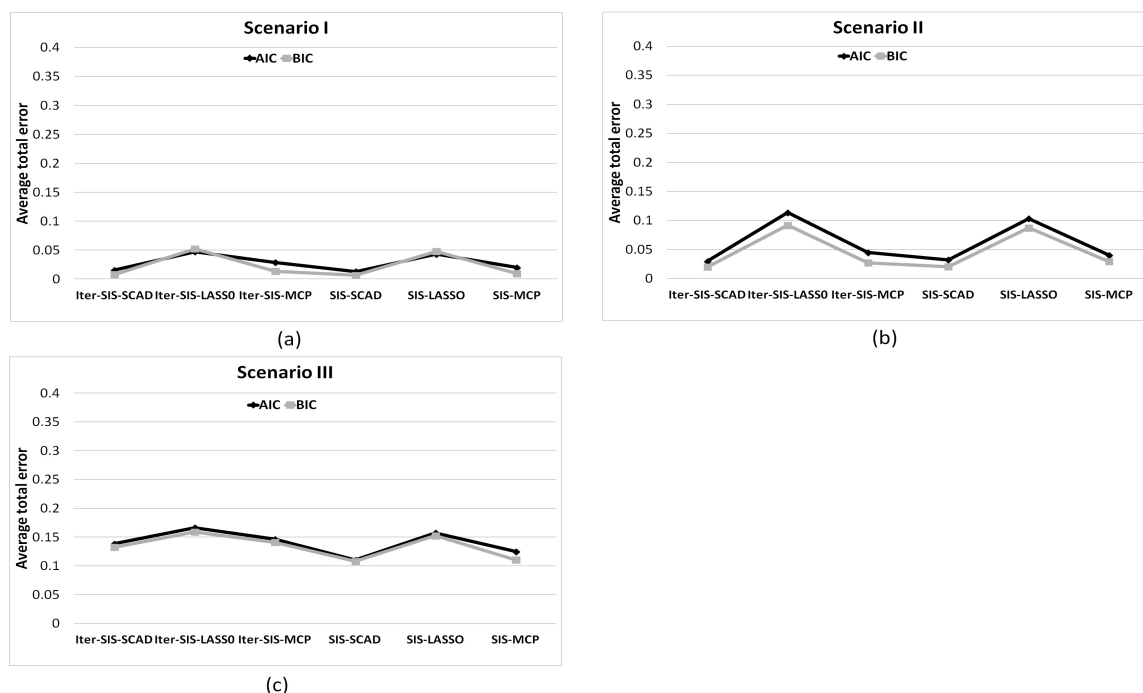
## 7.4.2 Αποτελέσματα προσομοιώσεων

### ΜΕΡΟΣ Ι: Συγκρίσεις μεταξύ των Ποινών

Πριν προχωρήσουμε στη μελέτη προσομοίωσης μεταξύ των διαδικασιών που βασίζονται στη SIS και των τεσσάρων προαναφερθέντων μεθόδων προσθέτουμε κάποια συγκριτικά αποτελέσματα μεταξύ όλων των μεθόδων SIS όπου εξετάστηκαν όλοι οι πιθανοί συνδυασμοί μεταξύ της SIS και της Iter-SIS χρησιμοποιώντας τρεις διαφορετικές ποινές. Η σύγκριση της SIS με την Iter-SIS είναι πραγματικά χρήσιμη για τον προσδιορισμό του καλύτερου συνδυασμού που οδηγεί σε χαμηλά επίπεδα του σφάλματος Τύπου I ή Τύπου II και σε υψηλά επίπεδα της ισχύος. Για να βοηθήσουμε τα συμπεράσματα, λάβαμε τον μέσο όρο των σφαλμάτων Τύπου I και Τύπου II για κάθε μία από τις 1000 προσομοιώσεις και στη συνέχεια προστέθηκε το συνολικό σφάλμα, δηλαδή το μέσο συνολικό σφάλμα για όλες τις προσομοιώσεις. Υπολογίζοντας το μέσο συνολικό σφάλμα, μπορούμε να αποφασίσουμε σχετικά με τη μέθοδο που διατηρεί και τα δύο σφάλματα στο χαμηλότερο επίπεδο, ταυτόχρονα. Τα αποτελέσματα συνοψίζονται στους Πίνακες Γ.1-Γ.35 του Παραρτήματος Γ.

Το Σχήμα 7.2 δείχνει ορισμένες συγκρίσεις των μεθόδων που βασίζονται στις SIS χρησιμοποιώντας είτε το κριτήριο AIC (μαύρες ράβδοι) είτε το κριτήριο BIC (γκρι ράβδοι) για τον συντονισμό της παραμέτρου κανονικοποίησης του ποινικοποιημένου υποπροβλήματος πιθανοφάνειας και αυτού του τελικού μοντέλου που επιλέχθηκε από την SIS ή την Iter-SIS. Για να εκτελέσουμε τις παραπάνω συγκρίσεις, οι συντελεστές  $\beta$  επιλέγονται σύμφωνα με το Μοντέλο 2. Παρατηρώντας το Σχήμα 7.2 καταλήγουμε στο συμπέρασμα ότι το κριτήριο BIC είναι προτιμότερο δεδομένου ότι δίνει χαμηλότερες τιμές για το μέσο συνολικό σφάλμα από εκείνες του κριτηρίου AIC. Επιπλέον, όσον αφορά τις συναρτήσεις των ποινών, η ποινή SCAD υπερέρχει όλων των άλλων και στα τρία σενάρια που εξετάστηκαν, δεδομένου ότι τόσο η SIS όσο και η Iter-SIS ακολουθούμενες από την ποινή SCAD δίνουν ταυτόχρονα τα χαμη-

## Penalty comparisons



**Σχήμα 7.2:** Μέσο συνολικό σφάλμα για τις μεθόδους που βασίζονται στη SIS χρησιμοποιώντας διαφορετικές ποινές. Υπάρχουν συγκρίσεις χρησιμοποιώντας είτε το κριτήριο AIC (μαύρες ράβδοι) είτε το κριτήριο BIC (γκρι ράβδοι). Το πάνω αριστερό πλαίσιο (a) αναφέρεται στο Σενάριο I, το πάνω δεξί πλαίσιο (b) αναφέρεται στο Σενάριο II, και το κάτω αριστερό πλαίσιο (c) αναφέρεται στο Σενάριο III.

λότερα σφάλματα. Καλά αποτελέσματα συγκεντρώνονται χρησιμοποιώντας και την ποινή MC. Ωστόσο, φαίνεται ότι η ποινή SCAD είναι ελαφρώς καλύτερη. Η SIS-SCAD ξεπερνά την Iter-SIS-SCAD μόνο στο Σενάριο III ως προς το σφάλμα Τύπου II (Πίνακες Γ.2-Γ.15). Συνοψίζοντας, οι μέθοδοι που βασίζονται στη SIS αποδίδουν περισσότερο σε συνδυασμό με τη συνάρτηση ποινής SCAD. Έτσι θα χρησιμοποιήσουμε μόνο την SIS-SCAD και την Iter-SIS-SCAD χρησιμοποιώντας το κριτήριο BIC για τις επικείμενες συγκρίσεις.

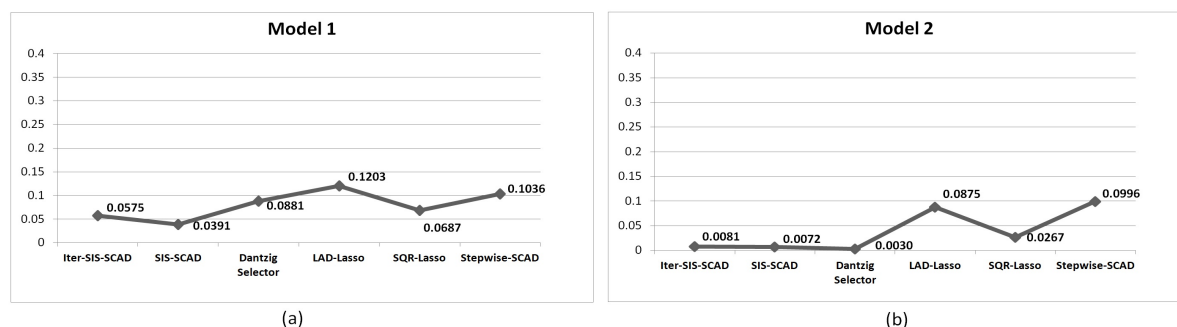
## ΜΕΡΟΣ II : Συγκρίσεις με άλλες μεθόδους

Οι Πίνακες Γ.16-Γ.30 στο Παράρτημα Γ συνοψίζουν τα αποτελέσματα για μερικά μέτρα απόδοσης. Τα αποτελέσματα των προσομοιώσεων παρουσιάζονται και μέσω των σχημάτων 7.2, 7.3, 7.4 και 7.5 όσον αφορά το μέσο συνολικό σφάλμα καθώς και μέσω του Σχήματος 7.6 όσον αφορά τη στατιστική ισχύ. Στόχος μας είναι να αποφασίσουμε ποια μέθοδος διατηρεί και τα δύο σφάλματα, Τύπου I και Τύπου II,

ταυτόχρονα στο χαμηλότερο επίπεδο, δεδομένου ότι όσο χαμηλότερο είναι το μέσο σφάλμα, τόσο χαμηλότερες είναι οι τιμές και των δύο σφαλμάτων. Το Μοντέλο 1 είναι ένα μοντέλο με τυχαίους συντελεστές οι οποίοι έχουν μικρό μέγεθος, ίσο με 2 και διακύμανση ίση με 1. Το γεγονός αυτό οδηγεί σε πιο πολύπλοκα μοντέλα αφού οι συντελεστές του είναι τόσο μικροί όσο και οι συντελεστές για τον όρο του σφάλματος. Το Μοντέλο 2 είναι ένα μοντέλο με τυχαίους συντελεστές οι οποίοι όμως έχουν σχετικά μεγάλο μέγεθος που ισούται με 5 και με διακύμανση ίση με 1.

**Σενάριο I:** Στο Σχήμα 7.3 είναι προφανές για το Μοντέλο 1 (αριστερό πλαίσιο) ότι υπάρχουν λίγες αξιοσημείωτες διαφορές μεταξύ των μεθόδων που βασίζονται στη SIS. Συνολικά, οι SIS-SCAD, Iter-SIS-SCAD, DS και SQR-Lasso καταφέρνουν να διατηρούν και τα δύο σφάλματα Τύπου I και Τύπου II, σ' ένα χαμηλό επίπεδο ταυτόχρονα. Ωστόσο, το Σχήμα 7.6 (πάνω αριστερό πλαίσιο, μαύρη γραμμή) δείχνει ότι μόνο η SIS-SCAD και η stepwise-SCAD καταφέρνουν να βρουν σχεδόν όλες τις ενεργές επιδράσεις με υψηλή πιθανότητα, επιτυγχάνοντας ποσοστά για την ισχύ κοντά στο 100%. Επιπλέον, η LAD LASSO δεν φαίνεται να είναι μια λογική επιλογή δεδομένου ότι, ακόμη και κάτω από το σενάριο σποραδικότητας των επιδράσεων, δεν εντοπίζονται ορισμένοι ενεργοί παράγοντες και στις περισσότερες περιπτώσεις παρουσιάζεται σφάλμα Τύπου II. Το γεγονός αυτό γίνεται εμφανές και μέσω του σχήματος της Ισχύος (Σχήμα 7.6, επάνω αριστερό πλαίσιο, μαύρη γραμμή), όπου η LAD LASSO μαζί με το DS επιτυγχάνει συγκριτικά χαμηλότερα ποσοστά. Για το Μοντέλο 2 όλες οι

Scenario I



**Σχήμα 7.3:** Μέσο συνολικό σφάλμα για το Σενάριο I. Το αριστερό πλαίσιο (a) αναφέρεται στο Μοντέλο 1, το δεξί πλαίσιο (b) αναφέρεται στο Μοντέλο 2.

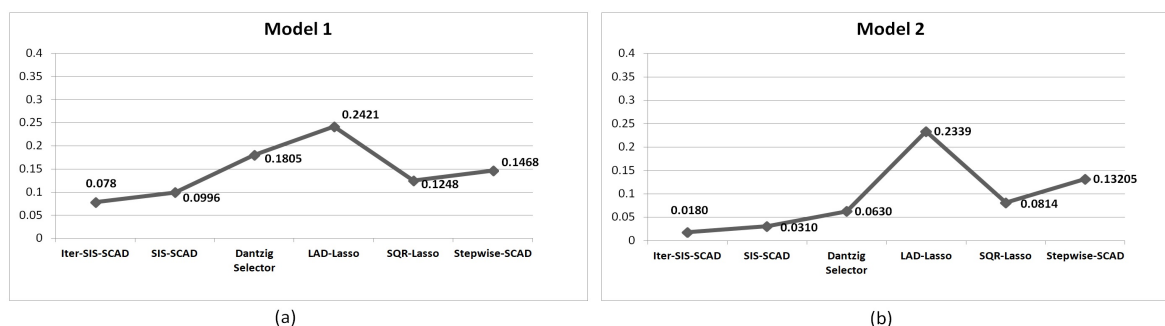
μέθοδοι εκτός από τη LAD LASSO και τη stepwise-SCAD έδωσαν εξαιρετικά



αποτελέσματα κάτω από το σενάριο σποραδικότητας των επιδράσεων, Σενάριο I, επιτυγχάνοντας χαμηλές τιμές και για τα ποσοστά σφάλματος Τύπου I και Τύπου II καθώς και υψηλές τιμές της στατιστικής ισχύος (Σχήμα 7.6, επάνω αριστερά, γκρι γραμμή). Το δεξιό πλαίσιο του Σχήματος 7.3 υποδηλώνει ότι το DS ξεπερνά ελαφρώς σε απόδοση την SIS-SCAD και την Iter-SIS-SCAD όσον αφορά τον μέσο όρο συνολικού σφάλματος αλλά όχι όσον αφορά την ισχύ. Το Σχήμα 7.6 (επάνω αριστερό πλαίσιο, γκρι γραμμή) δείχνει σαφώς ότι δεν υπάρχουν διαφορές στη στατιστική Ισχύ μεταξύ SIS-SCAD, Iter-SIS-SCAD και DS καθώς αυτές οι μέθοδοι προσδιορίζουν το πραγματικό μοντέλο περισσότερο από 99% των φορών επιτυγχάνοντας τις τιμές 99.97%, 99.85% και 99.77%, αντίστοιχα. Είναι ενδιαφέρον ότι οι μέθοδοι που βασίζονται στη SIS παρουσιάζουν πολύ καλύτερες επιδόσεις από αυτές των άλλων τριών μεθόδων, stepwise-SCAD, SQR και LAD LASSO.

**Σενάριο II:** Στο Σχήμα 7.4 (αριστερός πίνακας) είναι προφανές ότι οι μέθοδοι που βασίζονται στη SIS, ειδικά η Iter-SIS-SCAD, κυριαρχούν σε όλες τις άλλες μεθόδους στην περίπτωση του Μοντέλου 1. Αυτό το γεγονός ενισχύεται μέσω του άνω δεξιού πλαισίου του Σχήματος 7.6, όπου η Iter-SIS-SCAD για το Μοντέλο 1 (μαύρη γραμμή) επιτυγχάνει το ποσοστό 89.73% για τη στατιστική ισχύ σε σύγκριση με τα 86.18%, 65.82%, 55.95%, 77.91% και 89.82% των SIS-SCAD, DS, LAD-Lasso SQR-Lasso και stepwise-SCAD, αντίστοιχα. Μόνο η stepwise-SCAD δίνει πολύ καλά αποτελέσματα σε αυτό το σενάριο αν λάβουμε υπόψη την Ισχύ. Για μια ακόμη φορά, οι διαδικασίες που βασίζονται στη

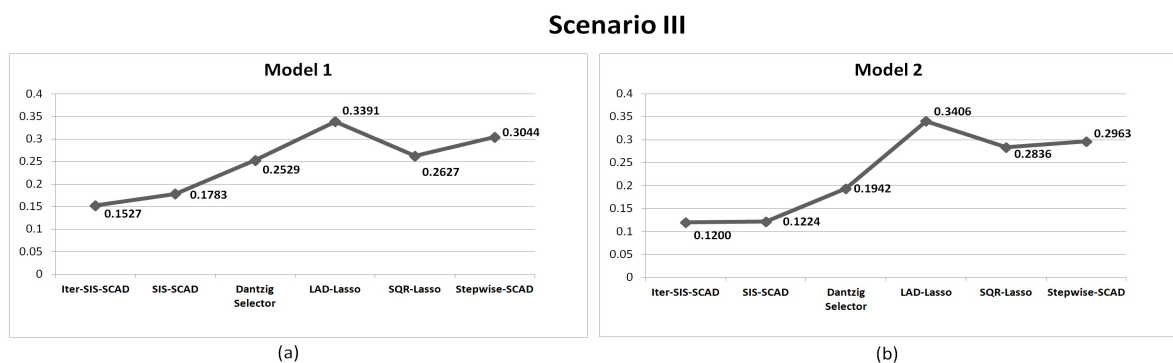
Scenario II



**Σχήμα 7.4:** Μέσο συνολικό σφάλμα για το Σενάριο II. Το αριστερό πλαίσιο (a) αναφέρεται στο Μοντέλο 1, το δεξιό πλαίσιο (b) αναφέρεται στο Μοντέλο 2.

SIS σαφώς έχουν καλύτερες επιδόσεις έναντι όλων των υπολοίπων όσον αφορά όλα τα εξεταζόμενα μέτρα, στην περίπτωση του Μοντέλου 2. Η Iter-SIS-SCAD επιτυγχάνει την καλύτερη απόδοση, ακόμα καλύτερη και από αυτή της SIS-SCAD (δεξί πλαίσιο του Σχήματος 7.4). Πιο συγκεκριμένα, το συνολικό μέσο σφάλμα της ανέρχεται στο 1% ενώ η Ισχύς ισούται με 98.77% (Σχήμα 7.6, επάνω δεξί πλαίσιο, γκρι γραμμή). Η SIS-SCAD επιτυγχάνει εξαιρετικά καλά αποτελέσματα, παρόμοια με αυτά της Iter-SIS-SCAD, ακολουθούμενη από το DS. Έτσι, η προτεινόμενη μέθοδος φαίνεται να αποδίδει ακόμη και κάτω από το Σενάριο II.

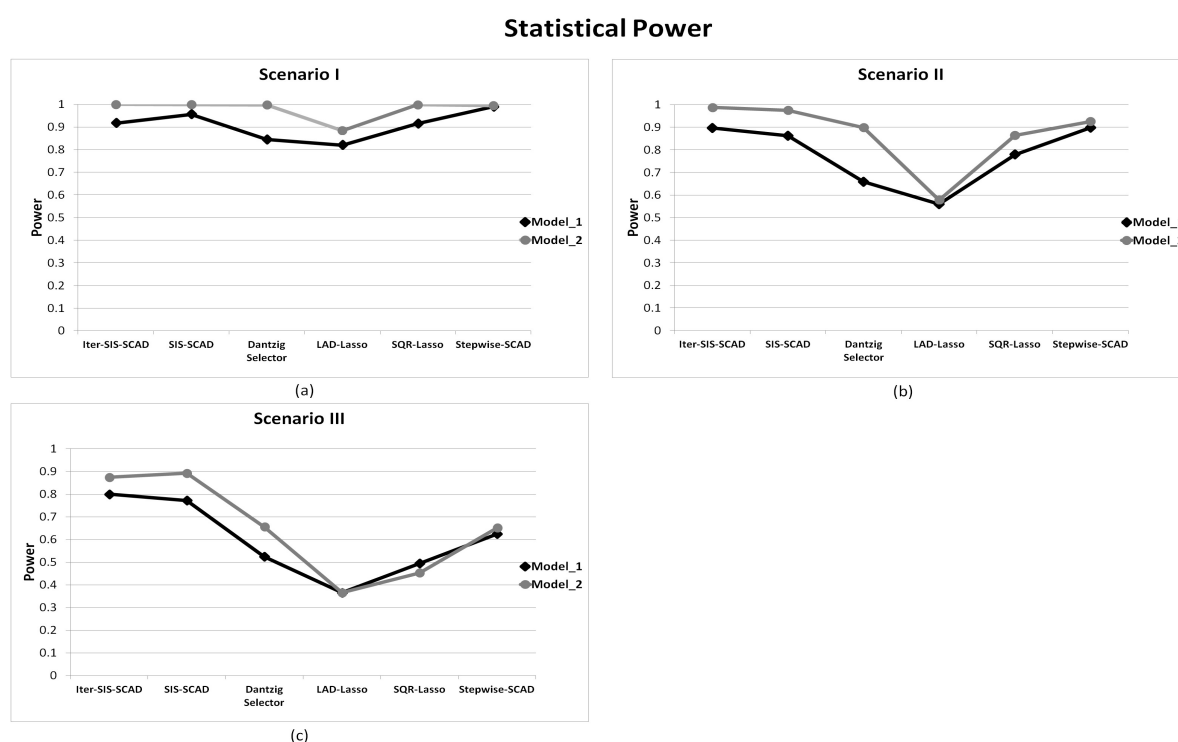
**Σενάριο III:** Το Σχήμα 7.5 παρουσιάζει με σαφήνεια, μέσω του αριστερού πλαισίου του, την αναμενόμενη υπεροχή των διαδικασιών που βασίζονται στη SIS σε σχέση με τις άλλες μεθοδολογίες ιδίως όσον αφορά το μέσο συνολικό σφάλμα στο Μοντέλο 1. Επιπλέον, είναι σαφές το όφελος των SIS-SCAD και Iter-SIS-SCAD σε σχέση με τις άλλες μεθόδους σε σχέση με την Ισχύ, κάτι που απεικονίζεται στο Σχήμα 7.6 (κάτω αριστερό πλαίσιο, μαύρη γραμμή). Συγκεκριμένα, η Iter-SIS-SCAD παρουσιάζει περίπου 53% βελτίωση στην Ισχύ έναντι του DS και κατά περίπου 28% βελτίωση στην Ισχύ έναντι της stepwise-SCAD. Για το Μοντέλο 2 η προτεινόμενη μεθοδολογία επιτυγχάνει πολύ καλύτερες



**Σχήμα 7.5:** Μέσο συνολικό σφάλμα για το Σενάριο III. Το αριστερό πλαίσιο (a) αναφέρεται στο Μοντέλο 1, το δεξί πλαίσιο (b) αναφέρεται στο Μοντέλο 2.

επιδόσεις στον προσδιορισμό των σημαντικών παραγόντων. Φαίνεται ότι μόνο η SIS-SCAD και Iter-SIS-SCAD είναι ανταγωνιστικές, όπου η πρώτη επιτυγχάνει υψηλότερες τιμές της στατιστικής ισχύος (Σχήμα 7.6, κάτω αριστερό πλαίσιο, γκρι γραμμή) ενώ η τελευταία συγκεντρώνει χαμηλότερο μέσο όρο συνολικού

σφάλματος (Σχήμα 7.5, δεξιό πλαίσιο). Το DS επιτυγχάνει σχετικά χαμηλό μέσο συνολικό σφάλμα. Ωστόσο, το Σχήμα 7.6 της Ισχύος δείχνει την κατωτερότητά της έναντι των προτεινόμενων μεθόδων. Το Σχήμα 7.6 ενισχύει την αιτιολόγησή μας αφού σαφώς απεικονίζει την υπεροχή των διαδικασιών που βασίζονται στη SIS. Οι πίνακες στο Παράρτημα παρουσιάζουν τον πλήρη πίνακα των αποτελεσμάτων, συμπεριλαμβανομένου του σφάλματος Τύπου I καθώς και το μέσο μέγεθος (MS) των τελικά επιλεγμένων μοντέλων.



**Σχήμα 7.6:** Στατιστική ισχύς στο Μοντέλο 1 (μαύρες γραμμές) και στο Μοντέλο 2 (γκρι γραμμές). Το επάνω αριστερό πλαίσιο (a) αναφέρεται στο Σενάριο I, το επάνω δεξιό πλαίσιο (b) αναφέρεται το Σενάριο II και το κάτω αριστερό πλαίσιο (c) αναφέρεται στο Σενάριο III.

Συνολικά, η προτεινόμενη μεθοδολογία υπερβαίνει όλες τις άλλες μεθόδους ειδικά όταν δεν ισχύει η υπόθεση της σποραδικότητας των επιδράσεων. Αυτό είναι εξαιρετικά σημαντικό σε περιπτώσεις που υπάρχουν στο πείραμα πολλοί ενεργοί παράγοντες. Έτσι, η υπεροχή της προτεινόμενης μεθόδου έγκειται στην απόδοσή της τόσο στο Σενάριο II (ενδιάμεση πολυπλοκότητα) όσο και το Σενάριο III (μεγάλος αριθμός ενεργών επιδράσεων), αφού η ανίχνευση των ενεργών παραγόντων σε ΥΣ πέραν του ενός ή των δύο ενεργών παραγόντων (Σενάριο I) φαίνεται ότι δεν είναι ικανοποιητική με τη χρήση άλλων μεθόδων.

### 7.4.3 Παραδείγματα

Στο κεφάλαιο αυτό παρουσιάζουμε τον τρόπο με τον οποίο η προτεινόμενη μεθοδολογία αποδίδει σε πραγματικά δεδομένα από ένα πείραμα που έχει συζητηθεί εκτενώς στη βιβλιογραφία. Αυτό το παράδειγμα παρουσιάζει τα δεδομένα Rubber του Williams [141], χρησιμοποιώντας τον ΥΣ του Lin [91].

#### Παράδειγμα 1

**Πίνακας 7.2:** Δύο επιπέδων υπερκορεσμένος σχεδιασμός και τα δεδομένα Rubber του Williams.

Εκτελέσεις	Παράγοντες																							Απόκριση
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
1	1	1	1	-1	-1	-1	1	1	1	1	1	-1	-1	-1	1	1	-1	-1	1	-1	-1	-1	1	133
2	1	-1	-1	-1	-1	-1	1	1	1	-1	-1	1	1	1	-1	1	-1	-1	-1	1	1	-1	-1	62
3	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	1	1	1	1	-1	-1	-1	-1	1	1	-1	45
4	1	1	-1	1	-1	1	-1	-1	-1	1	1	-1	-1	1	1	-1	1	1	1	-1	-1	-1	-1	52
5	-1	-1	1	1	1	1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1	-1	1	1	1	56
6	-1	-1	1	1	1	1	1	-1	1	1	1	-1	-1	1	1	-1	1	1	1	1	1	1	-1	47
7	-1	-1	-1	-1	1	-1	-1	1	-1	1	-1	1	1	-1	1	1	1	1	1	1	-1	-1	1	88
8	-1	1	1	-1	-1	1	-1	1	-1	1	-1	-1	-1	-1	-1	-1	-1	1	-1	1	1	1	-1	193
9	-1	-1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1	-1	1	1	-1	-1	-1	-1	1	1	32
10	1	1	1	1	-1	1	1	1	1	-1	-1	1	1	1	-1	1	-1	1	-1	1	-1	-1	1	53
11	-1	1	-1	1	1	-1	1	1	1	-1	1	-1	1	-1	-1	-1	1	1	-1	-1	-1	-1	1	276
12	1	-1	-1	-1	1	1	1	1	-1	1	1	1	-1	-1	1	-1	-1	-1	1	-1	1	1	1	145
13	1	1	1	1	1	-1	1	-1	1	-1	-1	1	-1	-1	-1	-1	1	-1	1	1	-1	1	-1	130
14	-1	-1	1	-1	-1	-1	-1	-1	-1	-1	1	1	1	-1	-1	-1	-1	-1	1	-1	1	-1	-1	127

**Πίνακας 7.3:** Μοντέλα που αποκομήθηκαν από την ανάλυση των δεδομένων Rubber.

	Μέθοδος	Μέγεθος	Τελικό μοντέλο
Williams [141]	Αρχική Ανάλυση	7	14 16 19 (σημαντικές)
			4 21 13 8 (μέτριας σημαντικότητας)
Williams [141]	Τελική πρόταση	4	14 10 19 4
Lin [91]	Παλινδρόμηση κατά βήματα	5	4 10 12 14 19
Westfall et al. [140]	Προς τα εμπρός παλινδρόμηση	6	14 12 19 4 10 11
Abraham et al. [1]	all-subsets regression	3	1 14 19
Abraham et al. [1]	all-subsets regression	4	1 13 14 19
Abraham et al. [1]	all-subsets regression	5	1 3 13 14 19
Beattie et al. [6]	πολλές μεθόδους	1	14
Li and Lin [82]	Stepwise-SCAD	4	14 12 19 4
Zhang et al. [158]	Μερικά ελάχιστα τετράγωνα (PLSVS)	4	14 12 19 4
Phoa et al. [112]	Dantzig Selector	1	14
Phoa [111]	SRRS	1	14

Λαμβάνοντας υπόψη τον ΥΣ του Lin [91], δηλαδή το μισό κλάσμα του 28-εκτελέσεων και 24-παραγόντων Plackett-Burman σχεδιασμού, εκτελούμε την ανάλυση χρησιμοποιώντας τα αρχικά δεδομένα από τον Williams [141]. Τα αρχικά δεδομένα απόκρισης του μισού κλάσματος παρουσιάζονται στον Πίνακα 7.2. Οι παράγοντες  $X_{13}$  και  $X_{16}$  είναι ίδιοι και, ως εκ τούτου, διαγράψαμε τη στήλη 13 που

**Πίνακας 7.4:** Μοντέλα που αποκομήθηκαν από την ανάλυση των δεδομένων Rubber χρησιμοποιώντας τις μεθόδους που βασίζονται στη SIS.

Δεδομένα	Μέθοδος	Μέγεθος	Τελικό Μοντέλο
Orig.-scale	Iter-SIS-SCAD	12	14 20 18 9 19 15(σημαντικές) 11 12 16 6 7 17 (μέτριας σημαντικότητας)
Orig.-scale	SIS-SCAD	13	1 2 7 12 14 15 22(σημαντικές) 6 8 9 11 16 18 (μέτριας σημαντικότητας)
$\log_{10}$ -scale	Iter-SIS-SCAD	1	14
$\log_{10}$ -scale	SIS-SCAD	1	14

αντιστοιχεί στον παράγοντα  $X_{13}$ . Ακολούθως, όπως οι Beattie et al. [6], οι Phoa et al. [112] και ο Phoa [111] μετά τη διαγραφή της στήλης 13, άλλαξαμε την επισήμανση των παραγόντων από 14 – 24 ως 13 – 23 (Πίνακας 7.2). Ο προκύπτων σχεδιασμός 23 παραγόντων και 14 εκτελέσεων είναι ένας  $E(s^2)$ -βέλτιστος ΥΣ και αποτελεί ένα δημοφιλή σχεδιασμό για τη μελέτη πολλών μεθόδων ανάλυσης σε ΥΣ. Μια σύντομη ματιά στη στήλη της απόκρισης,  $y$  στον Πίνακα 7.2 δείχνει ότι η μεταβολή του μεγέθους των τιμών της  $y$  είναι μεγάλη, με το μικρότερο ( $y = 32$ ) στο μεγαλύτερο ( $y = 300$ ). Ο Williams [141] επεσήμανε ότι η λογαριθμική κλίμακα είναι πιο κατάλληλη από την αρχική για το συγκεκριμένο πείραμα. Με αυτό τον τρόπο, αντίθετα με άλλους συγγραφείς που ανέλυσαν μόνο τα αρχικά δεδομένα, εμείς παρέχουμε αποτελέσματα τόσο για την αρχική απόκριση όσο και για τη μετασχηματισμένη ( $\log(y)$ ).

Η αρχική ανάλυση βρίσκεται στη μελέτη του Williams [141] και βασίζεται στο σχεδιασμό των Plackett-Burman 28 πειραματικών εκτελέσεων και 24 παραγόντων. Τα δεδομένα αυτά έχουν προηγουμένως αναλυθεί από πολλούς συγγραφείς, όπως π.χ. από τον Lin [91] ο οποίος χρησιμοποίησε την κατά βήματα παλινδρόμηση [35], από τους Westfall et al. [140], Abraham et al. [1] που χρησιμοποίησαν την all-subsets παλινδρόμηση περιορίζοντας την αναζήτησή τους για τα καλύτερα υποσύνολα σε μοντέλα με πέντε ή λιγότερους παράγοντες, από τους Beattie et al. [6] που σύγκριναν αρκετές μεθόδους επιλογής μοντέλων, τους Li και Lin [82], τους Zhang et al. [158] καθώς και από δύο σχετικά νέες ερευνητικές εργασίες από τον Phoa, η πρώτη αναφέρεται στον DS με το αντίστοιχο profile plot [112] και η δεύτερη στη μέθοδο SRRS [111]. Ο Πίνακας 7.3 παρουσιάζει μια συνοπτική περιγραφή των μεθόδων που χρησιμοποιήθηκαν από κάθε συγγραφέα καθώς και το αντίστοιχο μέγεθος και τις μεταβλητές του τελικά επιλεγμένου μοντέλου.

Εφαρμόσαμε τη SIS-SCAD και την Iter-SIS-SCAD. Δίνοντας ως προκαθορισμένη

τιμή για το μέγιστο αριθμό των παραγόντων  $n - 1 = 13$ , η Iter-SIS-SCAD αναγνώρισε τον  $X_{14}$ , ακολουθούμενο από τους  $X_{20}$ ,  $X_{18}$ ,  $X_9$ ,  $X_{19}$  και  $X_{15}$ , ως τους πρώτους 6 πιο σημαντικούς παράγοντες και από τους  $X_{11}$ ,  $X_{12}$ ,  $X_{16}$ ,  $X_6$ ,  $X_7$  και  $X_{17}$  ως μέτριας σημαντικότητας παράγοντες, ανάλογα με το μέγεθος των επιδράσεων. Όλες οι μεταβλητές που παρουσιάστηκαν από την αρχική μελέτη των Westfall et al. [140], ανιχνεύτηκαν από τις μεθόδους μας εκτός του παράγοντα  $X_{10}$ . Ο παράγοντας  $X_{14}$  βρίσκεται σε κάθε επιλεγμένο μοντέλο είτε από τη SIS είτε από την Iter-SIS, ενώ μόνος του εξηγεί το 63% της μεταβλητότητας στην παρατηρούμενη απόκριση. Ορισμένες χρήσιμες παρατηρήσεις σχετικά με το εν λόγω σύνολο δεδομένων δίνονται από τον Sundberg [130] ο οποίος έδειξε ότι το σύνολο δεδομένων του Williams θα έπρεπε να θεωρείται σε λογαριθμική κλίμακα αντί για την αρχική και ότι επηρεάζεται από μία έκτροπη παρατήρηση (outlier), η οποία επηρεάζει σημαντικά τα μικρά σύνολα δεδομένων που κατασκευάζονται ως ΥΣ. Η κυριαρχία του παράγοντα  $X_{14}$  ενισχύθηκε στην περίπτωση που θεωρούσαμε το σύνολο δεδομένων του Williams σε λογαριθμική κλίμακα αντί για την αρχική κλίμακα όπου είτε χρησιμοποιώντας τη SIS είτε την Iter-SIS μόνο ο  $X_{14}$  παραμένει ως ενεργός παράγοντας (Πίνακας 7.4). Αν και τα καλύτερα επιλεγμένα μοντέλα, είναι αυτά που περιέχουν τους παράγοντες 12, 11 ή 9, μόνο ο  $X_{14}$  είναι πολύ σημαντικός για την απόκριση. Έτσι, τελικά συμφωνούμε με τους Abraham et al. [1], ότι δεν είναι ακόμα σαφές ποιοι είναι οι ενεργοί παράγοντες.

#### **7.4.4 Γενικές παρατηρήσεις**

Προκειμένου να συμβάλουμε στη διαμόρφωση των συμπερασμάτων, αναπτύσσουμε ορισμένες γενικές παρατηρήσεις.

1. Η απώλεια ενός ενεργού παράγοντα στο πρώτο στάδιο (SIS) έχει επίπτωση στο σφάλμα Τύπου II το οποίο δεν μπορεί να βελτιωθεί στο δεύτερο στάδιο (δηλαδή με την εφαρμογή της SCAD). Έτσι, το σφάλμα Τύπου II της SIS-SCAD δίνει μια αίσθηση της πιθανότητας του πρώτου συνόλου  $d$ . Επομένως, η πιθανότητα του πρώτου συνόλου,  $d$ , είναι ίση ή μεγαλύτερη από την πιθανότητα (1- σφάλμα Τύπου II) της SIS-SCAD. Συμπερασματικά, η πιθανότητα του πρώτου σετ,  $d$ , είναι  $\geq$  από την Ισχύ (Power) της SIS-SCAD.

2. Πειράματα με μεγάλο αριθμό ενεργών επιδράσεων, όπως στο Σενάριο III στη μελέτη προσομοίωσης, αποτελούν μια σπάνια περίπτωση στους ΥΣ, δεδομένου ότι ισχύει η αρχή σποραδικότητας των επιδράσεων. Οι διαδικασίες SIS είναι ιδιαίτερα αποδοτικές όταν ο αριθμός των ενεργών παραγόντων στο μοντέλο είναι ίσος με  $n/6$  ή και περισσότερο (Σενάριο II και III) σε αντίθεση με τις άλλες μεθόδους, οι οποίες απέτυχαν στην εύρεση λύσης για το πρόβλημα πέρα από έναν ή δύο ενεργούς παράγοντες (Σχήμα 7.5). Τα Σχήματα 7.4 και 7.5 δείχνουν πόσο απότομη είναι η πτώση για τις τιμές των σφαλμάτων και το Σχήμα 7.6 δείχνει ότι οι αντίστοιχες τιμές για την Ισχύ είναι στο υψηλότερο επίπεδο έναντι των υπολοίπων μεθόδων. Συγκεκριμένα, εάν κοιτάξουμε και τα τρία πάνελ του Σχήματος 7.6 παρατηρούμε ότι η συγκριτική αποδοτικότητα των μεθόδων που βασίζονται στη SIS είναι ιδιαίτερα εμφανής για το Σενάριο III, λιγότερο εμφανής για το Σενάριο II, και σχεδόν εξαφανίζεται, ειδικά σε σύγκριση με το ΔΣ, για το Σενάριο I, στην περίπτωση του Μοντέλου 2 (γκρίζα γραμμή). Ωστόσο, στο Μοντέλο 1 (μαύρη γραμμή), οι μέθοδοι που βασίζονται στη SIS έχουν υπερβολικά υψηλή Ισχύ σε σύγκριση με τους υπόλοιπους αλγορίθμους. Αναμφίβολα, οι εξαιρετικά καλές επιδόσεις των μεθόδων SIS συνεχίζουν να υφίστανται και στην περίπτωση της σποραδικότητας των επιδράσεων (Σενάριο I) όπου η πλειονότητα των μεθόδων που παρουσιάζονται αποδίδουν σε μεγάλο βαθμό. (Σχήμα 7.3).
3. Φαίνεται ότι όλες οι μέθοδοι επιτυγχάνουν υψηλότερες τιμές της Ισχύος στο Μοντέλο 2, όπου το μέγεθος των επιδράσεων είναι ίσο με 5 με διασπορά 1, παρά στο Μοντέλο 1 όπου το μέγεθος των επιδράσεων είναι ίσο με 2 με διασπορά 1, αποδεικνύοντας ότι η πολυπλοκότητα στο Μοντέλο 1 είναι ακόμη υψηλότερη από ότι στο Μοντέλο 2. Αυτό οφείλεται στο γεγονός ότι οι συντελεστές στο Μοντέλο 1 είναι τόσο μικροί όσο ο όρος σφάλματος, ενώ οι συντελεστές στο Μοντέλο 2 είναι αρκετά μεγάλοι σε σχέση με τον όρο σφάλματος.
4. Στο Μοντέλο 1, η Iter-SIS-SCAD επιτυγχάνει υψηλότερες τιμές της Ισχύος για το Σενάριο II και το Σενάριο III από την SIS-SCAD η οποία έχει καλύτερη απόδοση στο Σενάριο I. Επομένως, προτείνουμε τη χρήση της επαναληπτικής μεθόδου SIS στο Σενάριο II-και το Σενάριο III, αλλά τη χρήση της SIS στην

περίπτωση που ισχύει η σποραδικότητα των επιδράσεων δηλαδή στο Σενάριο I. Στο Μοντέλο 2, τα αποτελέσματα και των δύο διαδικασιών που βασίζονται στη SIS είναι σχεδόν τα ίδια.

5. Για να ενισχύσουμε περαιτέρω την αιτιολόγησή μας για τη χρήση των διαδικασιών που βασίζονται στη SIS πραγματοποιήσαμε ένα παράδειγμα με πραγματικά δεδομένα όπου η προτεινόμενη μεθοδολογία επιτυγχάνει τον εντοπισμό όλων των επιδράσεων, οι οποίες έχουν εντοπιστεί και από άλλες μελέτες που έχουν γίνει στην βιβλιογραφία.

## 7.5 Συμπεράσματα

Οι ΥΣ τυπικά χρησιμοποιούνται για να κρησάρουμε τους λίγους, ενεργούς, από τους πολλούς υποψήφιους παράγοντες. Σε αυτό το κεφάλαιο μελετάμε τη μέθοδο SIS ως μέθοδο επιλογής μοντέλου στους ΥΣ. Από όσο γνωρίζουμε, αυτή είναι η πρώτη φορά, που αυτή η μέθοδος έχει εφαρμοστεί σε ΥΣ. Τα βασικότερα οφέλη αυτής της ιδέας είναι ότι, αν και απλή, επιτυγχάνει εξαιρετική απόδοση όσον αφορά τόσο τα σφάλματα Τύπου I και Τύπου II όσο και την Ισχύ. Μια εκτενής μελέτη προσομοίωσης δείχνει ότι η SIS-SCAD και η Iter-SIS-SCAD αποφέρουν τα καλύτερα αποτελέσματα και μπορούν να ανταγωνιστούν τις υπάρχουσες μεθόδους ανάλυσης που έχουν προταθεί μέχρι τώρα στην βιβλιογραφία.

Σε ένα πείραμα κρησαρίσματος, είναι πιο σημαντικό να προσδιοριστούν οι πραγματικά ενεργές επιδράσεις από το να διασφαλιστεί ότι δεν περιλαμβάνονται μη ενεργές επιδράσεις. Οι μέθοδοι SIS είναι πραγματικά αποτελεσματικές αφού επιτυγχάνουν συγκριτικά τις καλύτερες επιδόσεις ως προς τη στατιστική Ισχύ που αποτελεί και το κρισιμότερο μέτρο για την ανάλυση των ΥΣ. Τα αποτελέσματα που παρέχουν είναι τόσο ικανοποιητικά που ξεπερνούν ακόμη και εκείνα του DS, στις περισσότερες περιπτώσεις. Επιπλέον, αποδίδουν εξαιρετικά καλά όχι μόνο όταν το πραγματικό μοντέλο είναι πολύ αραιό αλλά και όταν υπάρχει μεγάλος αριθμός ενεργών επιδράσεων (Σενάριο III). Μια τέτοια μέθοδος λοιπόν είναι σε θέση να χειριστεί σχεδιασμούς που περιλαμβάνουν πολλούς παράγοντες δύο ή και πολλών επιπέδων.



## **Μέρος ΙΙΙ**

### **Εύρωστοι Παραμετρικοί Σχεδιασμοί**



---

## Εισαγωγικά στοιχεία και βασικές έννοιες για τους εύρωστους παραμετρικούς σχεδιασμούς

---

Cost is more important than quality  
but quality is the best way to reduce cost.

—Genichi Taguchi (1924-2012)

**Σ**το παρόν κεφάλαιο παρουσιάζονται κάποιες βασικές έννοιες για τους εύρωστους παραμετρικούς σχεδιασμούς.

### 8.1 Εισαγωγή

Μετά το τέλος του Δευτέρου Παγκοσμίου Πολέμου, οι ιαπωνικές εταιρίες έχοντας αντιληφθεί τις ανάγκες των καιρών για συνεχή ποιοτική βελτίωση των προϊόντων και των υπηρεσιών, προς όφελος του καταναλωτή υιοθέτησαν προγράμματα, τα οποία αποσκοπούσαν στην βελτίωση της ποιότητας. Σύμφωνα με την ιαπωνική φιλοσοφία, όσο υψηλότερη είναι η ποιότητα, τόσο χαμηλότερο είναι το κόστος παραγωγής και η τιμή του προϊόντος. Οι πιο δημοφιλείς μέθοδοι που αποσκοπούν στην ελάττωση του παραγωγικού κόστους και την ποιοτική βελτίωση των προϊόντων μέσω του στατιστικού ποιοτικού ελέγχου πριν από την έναρξη της μαζικής παραγωγής «Off-line Quality Control» είναι οι γνωστές ως «Μέθοδοι του Taguchi» (Taguchi methods). Ο Taguchi, με τη βοήθεια του παραμετρικού σχεδιασμού, στοχεύει στο να μειωθεί η μεταβλητότητα της παραγωγής ενός προϊόντος ή ενός συστήματος.

Η μεθοδολογία των εύρωστων παραμετρικών σχεδιασμών (Robust Parameter Design Methodology, RPDM) αποτελεί μια στατιστική μεθοδολογία που στοχεύει στη μείωση της διασποράς ενός συστήματος λόγω των μεταβλητών που είναι δύσκολο να ελεγχθούν, δηλαδή των μεταβλητών θορύβου. Σύμφωνα με τον Taguchi, η απόκλιση ενός προϊόντος από τον επιθυμητό στόχο οφείλεται σε δύο παράγοντες, τους παράγοντες που μπορούν να ελεγχθούν και τους παράγοντες που δεν μπορούν να ελεγχθούν. Ο εύρωστος παραμετρικός σχεδιασμός συνεπάγεται το σχεδιασμό του συστήματος όχι με την έννοια του πειραματικού σχεδιασμού αλλά με την έννοια της επιλογής των επιπέδων των παραγόντων ελέγχου με στόχο να επιτευχθεί η μείωση της διασποράς και η διατήρηση της μέσης απόδοσης. Ο στόχος είναι να επιλέξουμε τα επίπεδα των παραγόντων ελέγχου, ώστε να γίνει η διαδικασία λιγότερο ευαίσθητη στη μεταβολή του θορύβου και άρα πιο εύρωστη.

Ελεγχόμενοι παράγοντες (control factors,  $C$ ) είναι αυτοί των οποίων τις τιμές μπορεί να ρυθμίσει εύκολα ο πειραματιστής. Αυτοί οι παράγοντες χωρίζονται σε αυτούς που επηρεάζουν τη διασπορά του αποτελέσματος και ονομάζονται παράγοντες ελέγχου διασποράς (variability control factors), σε αυτούς που επιδρούν στο μέσο του αποτελέσματος και ονομάζονται παράγοντες ελέγχου του στόχου (target control factors), και τέλος, σε αυτούς που δεν επιδρούν ούτε στη μέση απόδοση ούτε στην μεταβλητότητα και επομένως μπορούν να ρυθμιστούν έτσι ώστε να εναρμονίζονται με τις οικονομικές απαιτήσεις και ονομάζονται παράγοντες κόστους. Η επιλογή των επιπέδων των παραγόντων ελέγχου εστιάζεται σε μεγάλο βαθμό στην μεταβλητότητα της μεταβλητής απόκρισης, το μεγαλύτερο ποσοστό της οποίας οφείλεται στην παρουσία ενός δεύτερου συνόλου παραγόντων, οι οποίοι καλούνται παράγοντες θορύβου (noise factors) ή μεταβλητές θορύβου (noise variables).

Οι μη ελεγχόμενοι παράγοντες ή παράγοντες θορύβου (noise factors,  $N$ ) είναι πηγές διασποράς που δε συνδέονται συνήθως με το περιβάλλον της παραγωγής ή της λειτουργίας και δεν μπορούν να ελεγχθούν κατά τη διαδικασία του σχεδιασμού ενός προϊόντος. Οι παράγοντες αυτοί διακρίνονται στους εξωτερικούς παράγοντες θορύβου, όπως η υγρασία, η σκόνη, η θερμοκρασία περιβάλλοντος, κ.ά. και στους εσωτερικούς παράγοντες θορύβου, όπως οι αποκλίσεις των μηχανημάτων από τις προδιαγραφές της παραγωγικής διεργασίας.

## 8.2 Διασταυρωμένοι και Συνδυασμένοι Σχηματισμοί

Ο Taguchi [131] πρότεινε τη χρήση των διασταυρωμένων σχηματισμών, δηλαδή της διασταύρωσης δύο παραγοντικών σχεδιασμών, του ενός που αντιστοιχεί στους παράγοντες ελέγχου και του άλλου που αντιστοιχεί στους παράγοντες θορύβου, με σκοπό την εξεύρεση των βέλτιστων επιπέδων για τους παράγοντες ελέγχου. Ο διασταυρωμένος σχηματισμός (crossed array) περιέχει δύο πειραματικούς σχεδιασμούς, έναν για τους παράγοντες θορύβου (εξωτερικός σχηματισμός, outer array) και έναν για τους παράγοντες ελέγχου (εσωτερικός σχηματισμός, inner array), οι οποίοι σχεδιασμοί διασταυρώνονται μεταξύ τους. Ανεξάρτητα ο ένας από τον άλλον, ο εσωτερικός και ο εξωτερικός σχηματισμός, είναι γενικά αρκετά οικονομικοί σε μέγεθος (πλήθος εκτελέσεων), γιατί είναι κορεσμένοι (saturated) ή σχεδόν κορεσμένοι σχεδιασμοί. Όμως η διασταύρωση αυτών των δυο δεν παράγει, πάντα, έναν οικονομικό σχεδιασμό. Σε μαθηματικούς όρους, αν  $N_1$  και  $N_2$  συμβολίζουν τα μεγέθη των εκτελέσεων για τους σχηματισμούς παραγόντων ελέγχου και θορύβου αντίστοιχα, τότε  $N = N_1 N_2$  είναι το μέγεθος των εκτελέσεων για το διασταυρωμένο σχηματισμό.

Ο διασταυρωμένος σχηματισμός σε πάρα πολλές περιπτώσεις είναι ιδιαίτερα δαπανηρός, όσον αφορά τις πειραματικές εκτελέσεις. Ωστόσο, άλλες στρατηγικές που χρησιμοποιούν τους λεγόμενους συνδυασμένους σχηματισμούς [122] χρειάζονται λιγότερες πειραματικές εκτελέσεις για τον ίδιο αριθμό παραγόντων. Μπορούμε να αναλύσουμε ένα σχεδιασμό που περιέχει τους παράγοντες ελέγχου και τους παράγοντες θορύβου με την βοήθεια της Μεθοδολογίας Αποκριτικών Επιφανειών (Response Surface Methodology, RSM). Για αυτή την προσέγγιση, ο λεγόμενος συνδυασμένος σχηματισμός (combined array, Wu and Hamada [143]) είναι πιο οικονομικός και αρκετά αποδοτικός. Η λέξη συνδυασμένος σημαίνει ότι δεν χρησιμοποιούνται διαφορετικοί σχηματισμοί, αλλά ένας σχηματισμός και για τους παράγοντες ελέγχου και για τους παράγοντες θορύβου. Οι συνδυασμένοι σχηματισμοί είναι οικονομικοί όσον αφορά στο μέγεθος των πειραματικών εκτελέσεων ενώ προσφέρουν και μεγαλύτερη ευχέρεια όσον αφορά στην επιλογή των παραγόντων που θα εκτιμήσουμε.

### 8.3 Αρχή σημαντικότητας των επιδράσεων στον Παραμετρικό Σχεδιασμό

Η ιεραρχία των επιδράσεων στους παραγοντικούς σχεδιασμούς δε λειτουργεί για πειράματα παραμετρικού σχεδιασμού. Στους παραγοντικούς σχεδιασμούς, οι επιδράσεις της ίδιας τάξης αντιμετωπίζονται με τον ίδιο τρόπο. Αυτό σημαίνει ότι οι κύριες επιδράσεις είναι υψηλότερης σημασίας, ακολουθούν οι επιδράσεις δεύτερης τάξης, έπειτα οι επιδράσεις τρίτης τάξης και ούτω καθεξής. Αντίθετα, οι επιδράσεις (κύριες ή αλληλεπιδράσεις) της ίδιας τάξης στους RPDM, δεν είναι απαραίτητα του ίδιου ενδιαφέροντος. Για παράδειγμα, οι  $C \times N$  αλληλεπιδράσεις είναι σημαντικότερες από τις  $C \times C$  αλληλεπιδράσεις. Σύμφωνα με τους Wu και Hamada [143] η σημαντικότητα των επιδράσεων θα πρέπει να κατατάσσεται σε φθίνουσα σειρά σπουδαιότητας όπως αυτή εμφανίζεται στον Πίνακα 8.1.

**Πίνακας 8.1:** Κατάταξη των επιδράσεων στους παραγοντικούς έναντι των εύρωστων παραμετρικών σχεδιασμών μέχρι τρίτης τάξης σημαντικότητας των παραγόντων.

Κατάταξη σημαντικότητας		Τύπος επίδρασης
Παραγοντικά Πειράματα		
1.		$C, N$
2.		$C \times N, C \times C, N \times N$
3.		$C \times C \times C, C \times C \times N, C \times N \times N$ $N \times N \times N$
Εύρωστα Παραμετρικά Πειράματα		
1.		$C, N, C \times N$
2.		$C \times C, C \times C \times N$
3.		$N \times N, C \times C \times C, C \times N \times N$ $N \times N \times N$

- (i) κύριες επιδράσεις ελέγχου (C), κύριες επιδράσεις θορύβου (N), ελέγχου επί θορύβου ( $C \times N$ ) αλληλεπιδράσεις
- (ii) ελέγχου επί ελέγχου ( $C \times C$ ) αλληλεπιδράσεις, ελέγχου επί ελέγχου επί θορύβου ( $C \times C \times N$ ) αλληλεπιδράσεις.
- (iii) θορύβου επί θορύβου ( $N \times N$ ) αλληλεπιδράσεις.

Παρατηρούμε ότι οι ( $C \times N$ ) αλληλεπιδράσεις είναι οι πιο σημαντικές από τις αλληλεπιδράσεις δεύτερης τάξης (όπως τις ( $N \times N$ ) για παράδειγμα) διότι, μαζί με τις

κύριες επιδράσεις, εμπλέκονται στον καθορισμό του μοντέλου θέσης και διασποράς άρα και στη διαδικασία βελτιστοποίησης του παραμετρικού σχεδιασμού. Η ακριβής εκτίμηση των επιδράσεών τους οδηγεί στην βελτίωση της διαδικασίας βελτιστοποίησης.

Οι κύριες επιδράσεις θορύβου επίσης περιλαμβάνονται στην πρώτη κλάση σημαντικότητας

- (i) παρόλο που δεν συμμετέχουν στη διαδικασία βελτιστοποίησης, το μέγεθός τους μπορεί να καθιστά δύσκολη την ερμηνεία της σημαντικότητας των άλλων επιδράσεων που είναι ταυτόσημες με αυτές. Επίσης οι κύριες επιδράσεις θορύβου είναι ιδιαίτερα χρήσιμες για την ευρωστία του μοντέλου διότι δεν μπορούν να υπάρχουν αλληλεπιδράσεις που εμπλέκονται οι αντίστοιχες επιδράσεις δίχως την ύπαρξη αυτών στο υπό εξέταση μοντέλο.
- (ii) Οι  $(C \times C \times N)$  αλληλεπιδράσεις μπορούν να θεωρηθούν ως επίδραση δεύτερης τάξης. Είναι υψηλής σημασίας όχι μόνο μεταξύ των αλληλεπιδράσεων τρίτης τάξης, αλλά και μεταξύ μερικών αλληλεπιδράσεων δεύτερης τάξης, όπως των  $(N \times N)$  για παράδειγμα. Η σημαντικότητά τους υποδεικνύει ότι το ζεύγος των παραγόντων ελέγχου θα πρέπει να προσαρμοστεί από κοινού για να ελαχιστοποιήσει την επίδραση του παράγοντα θορύβου (που περιέχεται στην αλληλεπίδραση  $(C \times C \times N)$  τριών παραγόντων) επί της απόκρισης.

## 8.4 Στρατηγικές μοντελοποίησης

Οι στρατηγικές μοντελοποίησης είναι είτε η λεγόμενη μοντελοποίηση θέσης - διασποράς που χρησιμοποιείται στους διασταυρωμένους σχηματισμούς ή η αποκριτική μοντελοποίηση που χρησιμοποιείται σε συνδυασμένους σχηματισμούς και η οποία υιοθετείται στο Κεφάλαιο 9. Η προσέγγιση της μοντελοποίησης θέσης-διασποράς είναι μία προσέγγιση που γίνεται εύκολα κατανοητή και μπορεί να εφαρμοστεί εξίσου εύκολα. Είναι μια φυσική προσέγγιση για τα πειράματα που βασίζονται στους διασταυρωμένους σχηματισμούς. Στην φόρμα του διασταυρωμένου σχηματισμού, οι  $\bar{y}_i$  και  $s_i^2$  υπολογίζονται πάνω στον ίδιο σχηματισμό παραγόντων θορύβου. Τα μοντέλα που βρίσκουμε καθορίζουν ποιοι παράγοντες είναι οι σημαντικοί παράγοντες

θέσης (αυτοί που εμφανίζονται στο μοντέλο του μέσου) και ποιοι είναι οι σημαντικοί παράγοντες διασποράς (αυτοί που εμφανίζονται στο μοντέλο διασποράς).

### 8.4.1 Μοντελοποίηση θέσης και διασποράς (Location and Dispersion Modeling)

Στην προσέγγιση της μοντελοποίησης θέσης-διασποράς κατασκευάζουμε ξεχωριστά μοντέλα για τα μέτρα θέσης και ξεχωριστά μοντέλα για τα μέτρα διασποράς για τις κύριες επιδράσεις και τις αλληλεπιδράσεις των παραγόντων ελέγχου. Σε κάθε συνδυασμό επιπέδων, υπολογίζονται τα  $\bar{y}_i$  (δειγματικός μέσος, ως μέτρο θέσης) και τα  $lns_i^2$  (λογαριθμοποιημένη δειγματική διασπορά, ως μέτρο διασποράς) επί των επαναλήψεων και υπολογίζονται ως εξής:

$$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}, \quad s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2, \quad (8.1)$$

όπου  $n_i$  είναι ο αριθμός των επαναλήψεων θορύβου για τη στάθμη του  $i$ -οστού παράγοντα ελέγχου. Από τα μοντέλα θέσης και διασποράς που προκύπτουν, οι παράγοντες θέσης (location factors) ορίζονται να είναι εκείνοι οι παράγοντες που εμφανίζονται ως σημαντικοί στο μοντέλο θέσης και οι παράγοντες διασποράς (dispersion factors) ορίζονται να είναι εκείνοι οι παράγοντες που εμφανίζονται ως σημαντικοί στο μοντέλο διασποράς. Ο παράγοντας θέσης που δεν είναι παράγοντας διασποράς καλείται παράγοντας προσαρμογής (adjustment factor). Για τα προβλήματα nominal-the-best οι προτεινόμενες στάθμες για τους παράγοντες ελέγχου μπορούν να ληφθούν χρησιμοποιώντας την ακόλουθη διαδικασία δύο βημάτων:

- 1 Επιλέγουμε τις στάθμες των παραγόντων διασποράς για την ελαχιστοποίηση της διασποράς.
- 2 Επιλέγουμε τις στάθμες του παράγοντα προσαρμογής για να φέρουμε τη θέση/το μέσο στο στόχο.

Αν ο παράγοντας που χρησιμοποιείται στο βήμα 2 ήταν επίσης παράγοντας διασποράς, η αλλαγή της στάθμης του θα μπορούσε επίσης να επηρεάσει και τη θέση και τη διασπορά. Τότε ενδεχομένως η θέση μπορεί να προσαρμοστεί στο στόχο αλλά η διασπορά να αυξηθεί, κάτι το οποίο απαιτεί την προσαρμογή των παραγόντων



διασποράς και στη συνέχεια την επανάληψη μεταξύ των δύο βημάτων. Αν ο ένας παράγοντας προσαρμογής δεν αρκεί για να φέρει την θέση στον επιθυμητό στόχο, μπορεί να χρειαστούν δύο ή περισσότεροι παράγοντες προσαρμογής στο βήμα 2 για να φτάσουμε τον στόχο. Για τα προβλήματα «the larger-the-better» και «the smaller-the-better», συνήθως συνίσταται να εναλλάσσουμε τα δύο βήματα της προηγούμενης διαδικασίας ως εξής:

- 1 Επιλέγουμε τις στάθμες των παραγόντων θέσης για τη μεγιστοποίηση (ή την ελαχιστοποίηση) της θέσης.
- 2 Επιλέγουμε τις στάθμες των παραγόντων διασποράς που δεν είναι παράγοντες θέσης για την ελαχιστοποίηση της διασποράς.

#### 8.4.2 Μοντελοποίηση Taguchi

Η ιδέα του Taguchi για την βελτίωση της ποιότητας δίνει έμφαση στη μείωση της διασποράς. Η προσέγγισή του για τη μείωση της διασποράς αναγνωρίζεται ευρέως ως μια πολύ σημαντική συμβολή στη στατιστική και τη μηχανική. Η μέθοδος ανάλυσης είναι απλή και είναι συνδεδεμένη με την έννοια του διασταυρωμένου σχηματισμού. Υπάρχει μια προσπάθεια να μεγιστοποιηθεί το SNR (Signal to Noise Ratio). Η μοντελοποίηση που γίνεται, ουσιαστικά σχετίζει το SNR με τους παράγοντες ελέγχου στο σενάριο «μόνο με κύριες επιδράσεις». Οι τυπικές τεχνικές του ANOVA χρησιμοποιούνται για να ανιχνεύσουν τους παράγοντες ελέγχου που επηρεάζουν το SNR. Αυτοί οι παράγοντες είναι ουσιαστικά και αυτοί που επηρεάζουν τη μεταβλητότητα του προβλήματος. Επιπλέον, ο ANOVA γίνεται στο μέσο της απόκρισης, όπως παρουσιάστηκε και στην 8.4.1, προκειμένου να διαπιστωθούν πιθανοί παράγοντες προσαρμογής.

Ο Taguchi πρότεινε την ομαδοποίηση των παρατηρήσεων σε κάθε εξωτερικό σχηματισμό με την βοήθεια κάποιων μέτρων απόδοσης, τα οποία παρέχουν πληροφορία για το μέσο και τη διασπορά. Τα μέτρα απόδοσης ονομάζονται μέτρα απόδοσης θορύβου SNRs και η στατιστική ανάλυση γίνεται βασισμένη σε αυτά. Υπάρχουν τέσσερα κύρια μέτρα απόδοσης θορύβου που συνιστά ο Taguchi. Η επιλογή εξαρτάται από το στόχο του πειράματος. Όπως στην περίπτωση της αποκριτικής επιφάνειας, υπάρχουν τρεις κύριοι στόχοι:

1. The Smaller the Better: Ο πειραματιστής επιθυμεί την ελαχιστοποίηση της απόδοσης.
2. The Larger the Better: Ο πειραματιστής επιθυμεί την μεγιστοποίηση της απόδοσης.
3. The Target is Best: Ο πειραματιστής επιθυμεί να επιτύχει μια συγκεκριμένη τιμή στόχο.

### 8.4.3 Αποκριτική Μοντελοποίηση (Response Modeling)

Οι στρατηγικές μοντελοποίησης στην μεθοδολογία των εύρωστων παραμετρικών σχεδιασμών είναι, είτε το λεγόμενο μοντέλο διασποράς θέσης που χρησιμοποιείται στους διασταυρωμένους σχηματισμούς είτε η αποκριτική μοντελοποίηση που χρησιμοποιείται στους συνδυασμένους σχηματισμούς. Η στρατηγική μοντελοποίησης που υιοθετήθηκε στην παρούσα διδακτορική διατριβή είναι η αποκριτική μοντελοποίηση. Εφαρμόσαμε το μοντέλο απόκρισης σύμφωνα με τα παρακάτω βήματα :

- Επιλέγουμε ποιο μοντέλο θέλουμε να εφαρμόσουμε. Για παράδειγμα για το μοντέλο πρώτης σημαντικότητας, θα λάβουμε υπόψη μόνο τους παράγοντες C, N και  $C \times N$ . Το τελικό εκτιμώμενο μοντέλο  $\hat{y}$  θα περιέχει τις σημαντικές επιδράσεις των παραπάνω παραγόντων.
- Σύμφωνα με την RPDM θεωρούμε ότι οι παράγοντες θορύβου είναι τυχαίες μεταβλητές με  $E[\mathbf{z}] = 0$  και  $Var[\mathbf{z}] = \Sigma_{\mathbf{z}}$ . Με βάση αυτές τις υποθέσεις μπορούμε να διαμορφώσουμε το μοντέλο θέσης  $\widehat{E}[y]$  καθώς και το μοντέλο διασποράς  $\widehat{Var}[y]$ , αντίστοιχα.
- Η RPDM μεθοδολογία στοχεύει στη μείωση της διασποράς της διαδικασίας, ενώ βελτιστοποιεί τη μέση απόκριση. Στόχος είναι να επιλέξουμε σωστά τα επίπεδα των παραγόντων ελέγχου έτσι ώστε η διαδικασία να μην είναι ευαίσθητη στις αλλαγές των παραγόντων θορύβου. Με αυτόν τον τρόπο, επιλέγουμε τα επίπεδα των συντελεστών ελέγχου έτσι ώστε να βελτιστοποιήσουμε το μοντέλο του μέσου  $\widehat{E}[y]$  κρατώντας το μοντέλο της διασποράς  $\widehat{Var}[y]$  όσο το δυνατόν χαμηλότερα.

- Αυτό μπορεί να γίνει, για παράδειγμα, εξετάζοντας την αλληλεπίδραση μεταξύ των παραγόντων ελέγχου και θορύβου.



---

## Προσέγγιση με τη μοντελοποίηση της απόκρισης στη μεθοδολογία των εύρωστων παραμετρικών σχεδιασμών με τη χρήση υπερκορεσμένων σχεδιασμών

---

**T**α τελευταία χρόνια, τόσο οι εύρωστοι παραμετρικοί σχεδιασμοί (robust parameter designs, RPDs) όσο και οι ΥΣ έχουν προσελκύσει μία πληθώρα ερευνητών. Στο παρόν κεφάλαιο, θεωρούμε μία «χρυσή τομή» μεταξύ των δύο αυτών πεδίων. Πιο συγκεκριμένα, προτείνουμε μια κατασκευή αποτελεσματικών ΥΣ μαζί με μία μέθοδο ανάλυσης, προκειμένου να αντιμετωπιστεί το ιδιαίτερα σημαντικό πρόβλημα που αφορά στη μεθοδολογία των εύρωστων παραμετρικών σχεδιασμών (RPDM). Συνδυάζοντας την επαναληπτική μέθοδο σίγουρου κρησαρίσματος (iterative sure independence screening, Iter-SIS) για την επιλογή των μεταβλητών με μία ποινικοποιημένη μέθοδο, την αποκαλούμενη SCAD, προχωρήσαμε στην ανάλυση των ΥΣ που κατασκευάστηκαν στο παρόν κεφάλαιο. Η προτεινόμενη μεθοδολογία εφαρμόστηκε σε διαφορετικά μοντέλα με σκοπό να γίνει ορατή η αποτελεσματικότητά της σε πολλά διαφορετικά σενάρια, υποθέτοντας τόσο μοντέλα πρώτης όσο και μοντέλα δεύτερης τάξης υπό την έννοια των σχεδιασμών της επιφάνειας απόκρισης. Στο παρόν κεφάλαιο παρουσιάζουμε δύο επεξηγηματικά παραδείγματα καθώς και πολυάριθμα αριθμητικά πειράματα προσομοίωσης για πολλές διαφορετικές περιπτώσεις. Τα αποτελέσματα υποδηλώνουν ότι η προτεινόμενη μέθοδος είναι ιδιαίτερα αποτελεσματική για τον προσδιορισμό των ενεργών επιδράσεων των κύριων παραγόντων, των αλληλεπιδράσεων δύο παραγόντων, των αλληλεπιδράσεων τριών παραγόντων καθώς επίσης και των καθαρά τετραγωνικών όρων, υπό την προϋπόθεση της σπο-

ραδικότητας των επιδράσεων. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [III].

## **9.1 Ερευνητικό πρόβλημα-Λύση**

Μεταξύ πολλών εφαρμογών του RPDM, μία από τις πιο πρόσφατες και ενδιαφέρουσες αποτελεί η εφαρμογή της μεθοδολογίας των αποκριτικών επιφανειών (RSM). Αυτή η προσέγγιση προτάθηκε για πρώτη φορά από τους Vining και Myers [137] οι οποίοι προσάρμοσαν μοντέλα αποκριτικών επιφανειών ξεχωριστά για το μέσο και τη διασπορά της απόκρισης. Οι Myers et. al. [106] αντιστοίχησαν τους παράγοντες ελέγχου και τους παράγοντες θορύβου σε ένα συνδυασμένο σχεδιασμό και εφάρμοσαν ένα μοντέλο αποκριτικής επιφάνειας.

Οι ΥΣ αποτελούν μια μεγάλη κατηγορία παραγο-νικών σχεδιασμών που μπορούν να χρησιμοποιηθούν για πειράματα κρησαρίσματος σημαντικών παραγόντων από ένα μεγάλο σύνολο πιθανών ενεργών παραγόντων. Τέτοιοι σχεδιασμοί χρησιμοποιούνται στο αρχικό στάδιο βιομηχανικών είτε επιστημονικών πειραμάτων για τον προσδιορισμό των ενεργών επιδράσεων και είναι εξαιρετικά χρήσιμοι υπό την προϋπόθεση ότι πληρείται η υπόθεση της «σποραδικότητας των επιδράσεων» [17]. Αυτή η υπόθεση σημαίνει ότι μόνο λίγοι από τους πολλούς πειραματικούς παράγοντες επηρεάζουν πραγματικά την απόκριση που μας ενδιαφέρει. Για περισσότερες λεπτομέρειες, παραπέμπουμε στο μέρος II της παρούσας διδακτορικής διατριβής.

Η έννοια των ΥΣ στα RPD προβλήματα είναι πρόσφατη και η μόνη ερευνητική εργασία που εξετάζει αυτή την προσέγγιση προτάθηκε από τους Matsuura et al. [102]. Στο παρόν κεφάλαιο προτείνουμε έναν ΥΣ δύο επιπέδων στην περίπτωση ενός προβλήματος RPD υποθέτοντας ότι η απόκριση που μας ενδιαφέρει ακολουθεί την κανονική κατανομή. Εκτός από τις κύριες επιδράσεις, τις αλληλεπιδράσεις δεύτερης τάξης και τους καθαρά τετραγωνικούς όρους, έχουμε επίσης εξετάσει και τις αλληλεπιδράσεις τρίτης τάξης  $C \times C \times N$  που είναι ίδιας σημαντικότητας με εκείνες των αλληλεπιδράσεων δεύτερης τάξης  $C \times C$  σύμφωνα με την αρχή σημαντικότητας των επιδράσεων που παρουσιάζονται στην ενότητα 8. Αναμφίβολα, η χρήση πολλών διαφορετικών σχεδιασμών που προσπαθούν να αντιμετωπίσουν το παρόν πρόβλημα μπορεί να οδηγήσει σ' έναν εξαιρετικά μεγάλο αριθμό πειραματικών εκτελέσεων.

Στο παρόν κεφάλαιο κατασκευάζουμε αποδοτικούς ΥΣ για εφαρμογές στους RPD που μειώνουν δραματικά τις πειραματικές εκτελέσεις. Όλες οι θεωρούμενες επιδράσεις αντιστοιχίζονται στις στήλες του προτεινόμενου σχεδιασμού με τη χρήση ενός  $(8m + 4) \times (8m + 4)$  πίνακα Hadamard ( $m$  είναι ένας φυσικός αριθμός) παίρνοντας τον foldover σχεδιασμό. Με αυτόν τον τρόπο αποκτάμε και ξεχωριστές/διακριτές στήλες που αντιστοιχούν σε κάθε επίδραση και την ιδιότητα κάθε στήλη να είναι ισορροπημένη. Επιπλέον, πραγματοποιήσαμε μία ενδελεχή αναζήτηση για κάποιες τάξεις (order), όπου  $n \neq 8m + 4$ , τα αποτελέσματα των οποίων παρουσιάζονται στην ενότητα 9.3.1. Η ισχύς της προτεινόμενης μεθόδου ενισχύθηκε επιπλέον με την παροχή της πλήρους δομής του πίνακα συσχέτισης για μια σειρά διαφορετικών περιπτώσεων.

Στην παρούσα ενότητα παρουσιάζουμε μία προσέγγιση με τη μοντελοποίηση της απόκρισης στη μεθοδολογία των εύρωστων παραμετρικών σχεδιασμών με τη χρήση ΥΣ. Η ενότητα 9.3 παρουσιάζει τις βασικές έννοιες τόσο του RPDM όσο και της διαδικασίας επιλογής μεταβλητών μέσω της επαναληπτικής μεθόδου σίγουρου κρησαρίσματος σε συνδυασμό με τη μη κοίλη ποινικοποιημένη πιθανοφάνεια. Στην παράγραφο 9.5 περιγράφεται η προτεινόμενη μεθοδολογία μαζί με δύο επεξηγηματικά παραδείγματα και στην παράγραφο 9.4 διεξάγουμε πολλά αριθμητικά πειράματα που αποκαλύπτουν την αποτελεσματικότητα της προτεινόμενης διαδικασίας.

## 9.2 Θεωρητικό υπόβαθρο

Σε αυτή την ενότητα περιγράφονται οι βασικές έννοιες του RPDM, των μοντέλων που εξετάζονται στη μελέτη μας, καθώς επίσης και της μεθοδολογίας της επιλογής μέσω της ποινικοποιημένης μη κοίλης πιθανοφάνειας, ώστε να αντιμετωπιστεί το πρόβλημα της ανάλυσης του μοντέλου απόκρισης.

### 9.2.1 Μεθοδολογία Εύρωστων Παραμετρικών Σχεδιασμών

Η δυσκολία στον αποτελεσματικό έλεγχο ορισμένων μεταβλητών εισόδου όταν μια διαδικασία είναι υπό εξέλιξη οδήγησε στον χαρακτηρισμό των μεταβλητών εισόδου σε δύο κατηγορίες: ελέγχου και θορύβου. Οι μεταβλητές ελέγχου είναι εκείνες που μπορούν εύκολα να ελεγχθούν ή να ρυθμιστούν κατά τη διάρκεια διεξαγωγής

της διαδικασίας, ενώ οι μεταβλητές θορύβου είναι εκείνες που δεν μπορούν ή είναι εξαιρετικά δύσκολο να ελεγχθούν.

Για  $1 \leq u \leq N$ , έστω  $x_{u1}, x_{u2}, \dots, x_{up}$  και  $z_{u1}, z_{u2}, \dots, z_{uq}$  είναι αντίστοιχα τα επίπεδα των  $p$  παραγόντων ελέγχου και των  $q$  παραγόντων θορύβου που αντιστοιχούν στην  $u$ -οστή πειραματική εκτέλεση. Θεωρούμε δύο μοντέλα εύρωστου παραμετρικού σχεδιασμού (RPD): το πρώτο, που αναφέρεται ως μοντέλο 1, είναι το εξής

$$y_u = \mu + \sum_{i=1}^p a_i x_{ui} + \sum_{i_1 < i_2 = 1}^p a_{i_1 i_2} x_{ui_1} x_{ui_2} + \sum_{j=1}^q b_j z_{uj} + \sum_{i=1}^p \sum_{j=1}^q c_{ij} x_{ui} z_{uj} + \sum_{j=1}^q \sum_{i_1 < i_2 = 1}^p d_{i_1 i_2 j} x_{ui_1} x_{ui_2} z_{uj} + \epsilon_u, \quad u = 1, 2, \dots, N, \quad (9.1)$$

όπου  $y_u$  δηλώνει την απόκριση που αντιστοιχεί στην  $u$ -οστή πειραματική εκτέλεση. Το δεύτερο, που αναφέρεται ως μοντέλο 2, είναι ένα μοντέλο επιφάνειας απόκρισης και περιγράφεται ως εξής

$$y_u = \mu + \sum_{i=1}^p a_i x_{ui} + \sum_{i_1 < i_2 = 1}^p a_{i_1 i_2} x_{ui_1} x_{ui_2} + \sum_{i=1}^p a_{ii} x_{ui}^2 + \sum_{j=1}^q b_j z_{uj} + \sum_{i=1}^p \sum_{j=1}^q c_{ij} x_{ui} z_{uj} + \sum_{j=1}^q \sum_{i_1 < i_2 = 1}^p d_{i_1 i_2 j} x_{ui_1} x_{ui_2} z_{uj} + \epsilon_u, \quad u = 1, 2, \dots, N. \quad (9.2)$$

Εκτός από αυτά τα δύο μοντέλα μπορούμε να θεωρήσουμε δύο επιπλέον μοντέλα που δεν περιέχουν τους όρους αλληλεπιδράσεων τριών παραγόντων που περιλαμβάνονται στο μοντέλο (1) και (2). Θα μπορούσαμε να αναφερθούμε σε αυτά τα δύο μοντέλα ως μοντέλο 3 και μοντέλο 4, αντίστοιχα. Με αυτόν τον τρόπο, θα αποδείξουμε στα ακόλουθα κεφάλαια ότι η μεθοδολογία μας λειτουργεί καλά και στα τέσσερα μοντέλα. Υποθέτουμε ότι τα επίπεδα των παραγόντων θορύβου δεν είναι ελεγχόμενα και είναι τυχαίες μεταβλητές στη διαδικασία έτσι ώστε  $E[\mathbf{z}] = \mathbf{0}$  και  $\text{Var}[\mathbf{z}] = \Sigma_{\mathbf{z}}$ .

Επομένως, το μοντέλο του μέσου και το μοντέλο της διασποράς προκύπτουν βάσει αυτής της υπόθεσης. Η RPD μεθοδολογία στοχεύει στη μείωση της διασποράς της διαδικασίας ενώ βελτιστοποιεί τη μέση απόκριση με τη σωστή επιλογή των επιπέδων των παραγόντων ελέγχου έτσι ώστε η διαδικασία να γίνει εύρωστη στις αλλαγές των παραγόντων θορύβου. Αυτό μπορεί να γίνει, για παράδειγμα, εξετάζοντας την αλληλεπίδραση μεταξύ των παραγόντων ελέγχου και θορύβου.



### 9.2.2 Επιλογή μεταβλητών μέσω της μη κοίλης ποινικοποιημένης πιθανοφάνειας

Οι Fan και Li [42] πρότειναν μία κλάση από μεθόδους για την επιλογή μεταβλητών μέσω της μη κοίλης ποινικοποιημένης πιθανοφάνειας. Οι Li και Lin [81] εισήγαγαν μια επέκταση αυτής της μεθόδου στα μη κυρτά ποινικοποιημένα ελάχιστα τετράγωνα και επικεντρώθηκαν στην κατάσταση στην οποία ο πίνακας σχεδιασμού δεν είναι πλήρης. Σύμφωνα με τους Li και Lin [81], ορίζουμε τα ποινικοποιημένα ελάχιστα τετράγωνα ως

$$\mathcal{Q}(\boldsymbol{\beta}) = \frac{1}{2n} \sum_{i=1}^n (y_i - x_i^T \boldsymbol{\beta})^2 + \sum_{j=1}^d p_{\lambda_n}(|\beta_j|), \quad (9.3)$$

όπου  $d$  είναι η διάσταση του  $\boldsymbol{\beta}$ ,  $p_{\lambda_n}(\cdot)$  είναι η συνάρτηση ποινής, και  $\lambda_n$  είναι μία παράμετρος συντονισμού η οποία ελέγχει την πολυπλοκότητα του μοντέλου και μπορεί να επιλεγεί μέσω προσεγγίσεων βάσει των δεδομένων, όπως είναι η μέθοδος της διασταυρωμένης επικύρωσης (cross-validation) (CV) και η γενικευμένη μέθοδος της διασταυρωμένης επικύρωσης (GCV; [29]).

Επιλέγοντας τη συνάρτηση ποινής να είναι η  $p_{\lambda}(|\boldsymbol{\beta}|) = \lambda|\boldsymbol{\beta}|$ , καταλήγουμε στη Lasso [135] ή, γενικότερα, στην  $p_{\lambda}(|\boldsymbol{\beta}|) = \lambda|\boldsymbol{\beta}|^q$  που οδηγεί στην παλινδρόμηση κορυφογραμμής (ridge regression, [50]). Για να επιτευχθεί ο σκοπός της επιλογής μεταβλητών, οι Fan και Li [42] πρότειναν τη χρήση της ποινής SCAD. Επιλέγοντας, λοιπόν, μια αρχική τιμή για τους άγνωστους συντελεστές, οι Li και Lin [81] χρησιμοποίησαν την επαναληπτική παλινδρόμηση κορυφογραμμής έτσι ώστε να βρεθεί η λύση για τα ποινικοποιημένα ελάχιστα τετράγωνα παρέχοντας τόσο θεωρητικές όσο και εμπειρικές τεκμηριώσεις στην προσέγγισή τους.

### 9.2.3 Επαναληπτική μέθοδος σίγουρου κρησαρίσματος στους Υπερκορεσμένους Σχεδιασμούς

Οι Fan και Lv [44] πρότειναν μία διαδικασία επιλογής μεταβλητών χρησιμοποιώντας component-wise παλινδρόμηση έτσι ώστε να εντοπίσουν τους σημαντικούς παράγοντες βάσει μόνο των ατομικών συσχετίσεων με τη μεταβλητή απόκρισης. Κατασκεύασαν ένα  $(m + 1) \times 1$  διάνυσμα  $\boldsymbol{\omega}$  περιθωριακών συσχετίσεων μεταξύ των

επιδράσεων και της μεταβλητής απόκρισης το οποίο δίνεται από τη σχέση  $\omega = \mathbf{X}^T \mathbf{y}$ , όπου  $\mathbf{y}$  είναι ένα  $n \times 1$  διάνυσμα απόκρισης,  $\mathbf{X}$  είναι ο  $n \times (m+1)$  κανονικοποιημένος πίνακας του μοντέλου και το  $m$  αντιπροσωπεύει τον αριθμό των παραγόντων. Έτσι, το  $m+1$  είναι ο αριθμός των παραμέτρων του μοντέλου. Στη συνέχεια, δημιούργησαν ένα υπο-μοντέλο χρησιμοποιώντας μια τιμή κατωφλίου έτσι ώστε να εξαλειφθούν οι μη σημαντικές μεταβλητές. Οι Fan και Lv [44] αποκάλεσαν τη διαδικασία αυτή Μέθοδο Σίγουρου Κρησαρίσματος (SIS). Ο αριθμός των σημαντικών παραγόντων μπορεί να μειωθεί περαιτέρω με τη χρήση μιας ποινικοποιημένης μεθόδου όπως είναι η SCAD ή η Lasso. Αυτή η μέθοδος μπορεί να εφαρμοστεί για την εκτίμηση των σημαντικών επιδράσεων αγνοώντας εκείνες που είναι μη σημαντικές, εκτιμώντας τες κοντά στο μηδέν. Επιπλέον, οι Fan και Lv [44] παρουσίασαν μία επαναληπτική διαδικασία της αρχικής μεθόδου SIS, δηλαδή την Επαναληπτική Μέθοδο Σίγουρου Κρησαρίσματος (iterative SIS), προκειμένου να ξεπεραστούν ορισμένες αδυναμίες της μεθόδου SIS, όπως η πολυσυγγραμμικότητα. Αυτή η μέθοδος ουσιαστικά εκτελεί μια επαναληπτική διαδικασία χρησιμοποιώντας τα υπόλοιπα ως νέες αποκρίσεις.

### 9.3 Προτεινόμενη μέθοδος

Σε αυτή την ενότητα περιγράφουμε την προτεινόμενη μέθοδο που περιλαμβάνει τόσο την κατασκευή του σχεδιασμού όσο και την ανάλυση των εξεταζόμενων μοντέλων 1 και 2. Η επέκταση στο μοντέλο 3 και το μοντέλο 4 είναι απλή. Για την κατασκευή ή χρησιμοποιήσαμε έναν πίνακα Hadamard διαστάσεων  $(8m+4) \times (8m+4)$  από τον οποίο προκύπτει ένας foldover σχεδιασμός αυτού του πίνακα Hadamard. Πιο συγκεκριμένα, προτείνουμε δύο ξεχωριστές διαδικασίες ανάλογα με το μοντέλο ενδιαφέροντος, 1 ή 2. Ο προτεινόμενος σχεδιασμός παρουσιάζεται στην ακόλουθη αλγοριθμική διαδικασία:

1. Θεωρούμε έναν  $n \times n$  πίνακα Hadamard,  $\mathbf{H}_n^*$ , του οποίου όλα τα στοιχεία της πρώτης στήλης είναι μονάδες, και θέτουμε  $n = 8m+4$ , όπου  $m$  είναι ένας φυσικός αριθμός.
2. Ορίζουμε  $\mathbf{H}_N = [\mathbf{H}_n^{*T}, -\mathbf{H}_n^{*T}]^T$ .

3. Οι  $r$ , όπου  $r \leq 8m + 4$ , στήλες του  $\mathbf{H}_N$  αντιστοιχίζονται σε  $r$  παράγοντες, όπου  $N = 16m + 8$ .
4. Κατασκευάζουμε ένα σύνολο από  $2p$  πειραματικές εκτελέσεις όπου κάθε παράγοντας τοποθετείται διαδοχικά σε  $\pm\alpha$  και όλοι οι άλλοι παράγοντες ελέγχου και θορύβου τοποθετούνται στο μηδέν. Σημειώνουμε ότι τα αξονικά σημεία αντιστοιχίζονται μόνο στους παράγοντες ελέγχου.
5. Κατασκευάζουμε ένα σύνολο  $n_0$  κεντρικών σημείων που σημαίνει ότι όλοι οι παράγοντες ελέγχου και θορύβου τοποθετούνται στο μηδέν σε κάθε πειραματική εκτέλεση.

Οι συνδυασμοί επιπέδων των παραγόντων ελέγχου και θορύβου στον προτεινόμενο σχεδιασμό συμβολίζονται με  $x_i, i = 1, 2, \dots, N$  και  $z_i, i = 1, 2, \dots, N$ , αντίστοιχα, όπου  $N = 2(8m + 4)$ . Ο πίνακας σχεδιασμού για το μοντέλο 1 είναι ο  $\mathbf{X} = (f(x_1, z_1), f(x_2, z_2), \dots, f(x_N, z_N))$  ο οποίος αποτελεί έναν ΥΣ εάν  $2(8m + 4) < 1 + p + q + pq + \binom{p}{2} + \binom{p}{2}q$ , θεωρώντας το μοντέλο 1. Για αυτό το μοντέλο θα ακολουθήσουμε μόνο τα βήματα **1** - **3**, έτσι ώστε να δημιουργήσουμε τον πίνακα σχεδιασμού. Η ενσωμάτωση καθαρά τετραγωνικών όρων οδηγεί σε μία προσέγγιση με τη μεθοδολογία της επιφάνειας της απόκρισης. Το μοντέλο 2 αποτελεί ένα μοντέλο δεύτερης τάξης το οποίο πρέπει να προσαρμοστεί υπό τη μεθοδολογία της επιφάνειας απόκρισης. Για το λόγο αυτό, προτείνουμε το σχεδιασμό, σύμφωνα με το μοντέλο 2, το οποίο δημιουργείται αν ακολουθήσουμε τα βήματα **1** - **5** της αλγοριθμικής διαδικασίας. Αν θεωρήσουμε το μοντέλο 3 και το μοντέλο 4 ακολουθούμε τα αντίστοιχα βήματα βάσει των μοντέλων 1 και 2 αντίστοιχα. Ωστόσο, διαγράφουμε τις αλληλεπιδράσεις τριών παραγόντων στο αρχικά υποτιθέμενο RPD μοντέλο. Τα πειραματικά αποτελέσματα που παρέχονται στην ενότητα 9.4 υποδεικνύουν τον τρόπο επιλογής της παραμέτρου  $m$ . Πιο συγκεκριμένα, εάν οι σημαντικές επιδράσεις μπορούν να εκτιμηθούν (ο αριθμός τους θα πρέπει να είναι μικρότερος από το ένα τρίτο του αριθμού των παραμέτρων) τότε συνιστούμε το  $m$  να επιλεγεί έτσι ώστε ο εκτιμώμενος αριθμός των ενεργών επιδράσεων να είναι μικρότερος από περίπου το ένα τρίτο του αριθμού των πειραματικών εκτελέσεων. Διαφορετικά, το  $m$  θα πρέπει να επιλέγεται με τέτοιο τρόπο ώστε να μπορεί να προσαρμοστεί το πλήρες μοντέλο. Προκειμένου να δημιουργηθεί ο πίνακας σχεδιασμού, υπενθυμίζουμε την ιδιότητα

που παρέχεται από το Θεώρημα 1 που δίνεται στην ενότητα 9.3.1.

### 9.3.1 Κατασκευή

**Λήμμα 12** Θεωρούμε έναν  $n \times n$  πίνακα Hadamard  $\mathbf{H}_n^*$  στον οποίο η πρώτη στήλη είναι μία στήλη με μονάδες. Ορίζουμε  $\mathbf{H}_N = [\mathbf{H}_n^{*T}, -\mathbf{H}_n^{*T}]^T$ . Οι  $r$  ( $r \leq n$ ) στήλες του  $\mathbf{H}_N$  αντιστοιχίζονται στους  $r$  παράγοντες. Οι  $r$  στήλες συμβολίζονται ως  $\mathbf{d}_i = (d_{i1}, d_{i2}, \dots, d_{iN})^T, i = 1, 2, \dots, r$ , όπου  $N = 2n$ . Έστω

$$\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r, \mathbf{d}_{12}, \mathbf{d}_{13}, \dots, \mathbf{d}_{(r-1)r}, \mathbf{d}_{123}, \mathbf{d}_{124}, \dots, \mathbf{d}_{(r-2)(r-1)r}],$$

όπου, για  $1 \leq i_1 < i_2 \leq r$ ,  $\mathbf{d}_{i_1 i_2} = (d_{i_1 1} d_{i_2 1}, d_{i_1 2} d_{i_2 2}, \dots, d_{i_1 N} d_{i_2 N})^T$  και, για  $1 \leq i_1 < i_2 < i_3 \leq r$ ,  $\mathbf{d}_{i_1 i_2 i_3} = (d_{i_1 1} d_{i_2 1} d_{i_3 1}, d_{i_1 2} d_{i_2 2} d_{i_3 2}, \dots, d_{i_1 N} d_{i_2 N} d_{i_3 N})^T$ . Τότε ο  $\mathbf{D}$  είναι ένας δύο επιπέδων ισορροπημένος σχεδιασμός και οι ανά δύο στήλες του  $\mathbf{D}$  δεν είναι πλήρως εξαρτημένες δεδομένου ότι ικανοποιούνται οι ακόλουθες συνθήκες

1. Για  $1 \leq i_1 < i_2 < i_3 < i_4 \leq r$ , η αλληλεπίδραση  $\mathbf{d}_{i_1 i_2 i_3 i_4}^*$  δεν είναι πλήρως εξαρτημένη από την επίδραση του μέσου, όπου  $\mathbf{d}_i^* = (d_{i1}, d_{i2}, \dots, d_{iN})^T$ .
2. Για  $1 \leq i_1 < i_2 < i_3 < i_4 < i_5 < i_6 \leq r$ , η αλληλεπίδραση  $\mathbf{d}_{i_1 i_2 i_3 i_4 i_5 i_6}^*$  δεν είναι πλήρως εξαρτημένη από την επίδραση του μέσου.

Στην προτεινόμενη μεθοδολογία χρησιμοποιήσαμε την ιδιότητα που δίνεται στο ακόλουθο θεώρημα:

**Θεώρημα 1** Θεωρούμε έναν  $n \times n$  πίνακα Hadamard  $\mathbf{H}_n^*$  στον οποίο η πρώτη στήλη είναι μία στήλη με μονάδες. Θέτουμε  $n = 8m + 4$ , όπου  $m$  είναι ένας φυσικός αριθμός. Ορίζουμε  $\mathbf{H}_N = [\mathbf{H}_n^{*T}, -\mathbf{H}_n^{*T}]^T$ . Οι  $r$  ( $r \leq 8m + 4$ ) στήλες του  $\mathbf{H}_N$  αντιστοιχίζονται στους  $r$  παράγοντες. Οι  $r$  στήλες συμβολίζονται ως  $\mathbf{d}_i = (d_{i1}, d_{i2}, \dots, d_{iN})^T, i = 1, 2, \dots, r$ , όπου  $N = 16m + 8$ . Έστω

$$\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_r, \mathbf{d}_{12}, \mathbf{d}_{13}, \dots, \mathbf{d}_{(r-1)r}, \mathbf{d}_{123}, \mathbf{d}_{124}, \dots, \mathbf{d}_{(r-2)(r-1)r}],$$

όπου, για  $1 \leq i_1 < i_2 \leq r$ ,  $\mathbf{d}_{i_1 i_2} = (d_{i_1 1} d_{i_2 1}, d_{i_1 2} d_{i_2 2}, \dots, d_{i_1 N} d_{i_2 N})^T$  και, για  $1 \leq i_1 < i_2 < i_3 \leq r$ ,  $\mathbf{d}_{i_1 i_2 i_3} = (d_{i_1 1} d_{i_2 1} d_{i_3 1}, d_{i_1 2} d_{i_2 2} d_{i_3 2}, \dots, d_{i_1 N} d_{i_2 N} d_{i_3 N})^T$ . Τότε ο  $\mathbf{D}$  είναι

ένας ισορροπημένος σχεδιασμός δύο επιπέδων και οι ανά δύο στήλες του  $\mathbf{D}$  δεν είναι πλήρως εξαρτημένες.

**Απόδειξη.** Δίνεται στο παράρτημα.

Το να μην έχουμε πλήρως εξαρτημένες στήλες στον  $\mathbf{D}$  είναι μια σημαντική στατιστική ιδιότητα για τους εκτιμητές του RPD μοντέλου. Αντιστοιχίζοντας σε αυτό το σχεδιασμό όλους τους παράγοντες ελέγχου και θορύβου, μας δίνεται η δυνατότητα να εκτιμήσουμε όλες τις κύριες επιδράσεις (C, N), τις αλληλεπιδράσεις δεύτερης τάξης ( $C \times N$ ,  $C \times C$ ), καθώς και τις αλληλεπιδράσεις τρίτης τάξης ( $C \times C \times N$ ).

**Παρατήρηση 1** Το Θεώρημα 1 αποδεικνύει ότι όλοι οι πίνακες Hadamard τάξεως  $n = 8m + 4$  παρέχουν σχεδιασμούς που μπορούν να εκτιμήσουν το μοντέλο που μας ενδιαφέρει. Αυτό δεν σημαίνει ότι οποιαδήποτε άλλη κλάση είναι ακατάλληλη για αυτή τη μεθοδολογία, αλλά ότι η αναζήτηση τέτοιων σχεδιασμών πρέπει να γίνει για κάθε περίπτωση χωριστά για να προσδιοριστεί αν ένας πίνακας Hadamard τάξεως  $n \neq 8m + 4$  είναι κατάλληλος για τα προτεινόμενα μοντέλα. Πραγματοποιήσαμε αναζήτηση σε υπολογιστή για τις περιπτώσεις πινάκων Hadamard τάξεως 8, 16, 24 και 40 και βρήκαμε τα ακόλουθα: Για την τάξη 8 υπάρχει μόνο ένας inequivalent Hadamard matrix και για τάξη 16 υπάρχουν 5 inequivalent Hadamard matrix αλλά κανένας από αυτούς δεν είναι κατάλληλος για την προσαρμογή του μοντέλου που μας ενδιαφέρει (δηλαδή του μοντέλου όπου εκτιμώνται όλες οι υπό εξέταση παράμετροι). Για την τάξη 24 υπάρχουν 60 inequivalent Hadamard matrix αλλά μόνο 2 από αυτούς είναι κατάλληλοι για την προσαρμογή του μοντέλου που μας ενδιαφέρει (δηλαδή του μοντέλου όπου εκτιμώνται όλες οι υπό εξέταση παράμετροι). Αυτοί οι δύο πίνακες παρέχονται στο Παράρτημα Δ. Ο αριθμός των inequivalent Hadamard matrix τάξεως 40 είναι άγνωστος, αλλά μπορέσαμε να βρούμε μερικούς από αυτούς που είναι κατάλληλοι για την προτεινόμενη μεθοδολογία και παραθέτουμε έναν εξ αυτών στο Παράρτημα Δ.

**Παρατήρηση 2** Η ισχύς των προτεινόμενων σχεδιασμών ενισχύεται επιπλέον, αφού εκτός από το θεώρημά μας παρέχουμε την πλήρη δομή των πινάκων συσχέτισης για έναν αριθμό διαφορετικών περιπτώσεων και τα αποτελέσματα δείχνουν ότι στην πλειονότητα των περιπτώσεων οι παράγοντες στον πίνακα του μοντέλου παραμένουν ασυσχέτιστοι. Πιο συγκεκριμένα, ο Πίνακας 9.1 παρέχει την πλήρη δομή των πινάκων

συσχέτισης ενός αριθμού διαφορετικών περιπτώσεων για τους σχεδιασμούς που κατασκευάστηκαν από τους πίνακες Hadamard τάξεως  $n = 12$  και  $n = 20$ . Συμπεριλάβαμε μόνο τις περιπτώσεις όπου ο προκύπτων πίνακας του μοντέλου είναι υπερκορεσμένος και ο συνολικός αριθμός των κύριων επιδράσεων στο μοντέλο (παράγοντες ελέγχου και θορύβου) είναι μέχρι  $2n/3$ . Οι παράμετροι  $p$  και  $q$  συμβολίζουν τον αριθμό των παραγόντων ελέγχου και θορύβου που περιλαμβάνονται στο σχεδιασμό, αντίστοιχα, ενώ η παράμετρος  $t$  υποδηλώνει το συνολικό αριθμό των στηλών στον πίνακα του μοντέλου (εξαιρουμένου του σταθερού όρου) του μοντέλου 1. Στον Πίνακα 9.1 συμπεριλάβαμε τις απόλυτες τιμές των συσχετίσεων καθώς και τη συχνότητα στον αντίστοιχο πίνακα συσχέτισης. Ο συμβολισμός  $\text{Corr}(\mathbf{M}_1)$  αναφέρεται στον πίνακα συσχέτισης του πίνακα του μοντέλου για το μοντέλο 1, ενώ το  $\text{Corr}(A, B)$  αναφέρεται στον πίνακα συσχέτισης του αντίστοιχου πίνακα του μοντέλου που περιλαμβάνει μόνο τις επιδράσεις του  $A$  και τις επιδράσεις του  $B$ . Σημειώνουμε ότι τα  $C$  και  $N$  αντιπροσωπεύουν τις κύριες επιδράσεις των παραγόντων ελέγχου και θορύβου, αντίστοιχα. Ομοίως, οι  $C \times C$  υποδηλώνουν τις αλληλεπιδράσεις μεταξύ των παραγόντων ελέγχου, οι  $C \times N$  υποδηλώνουν τις αλληλεπιδράσεις μεταξύ των παραγόντων ελέγχου και των παραγόντων θορύβου και οι  $C \times C \times N$  υποδηλώνουν τις αλληλεπιδράσεις τρίτης τάξης μεταξύ των παραγόντων ελέγχου και των παραγόντων θορύβου.

Επιπλέον, χρησιμοποιούμε το  $r$  για να δηλώσουμε την απόλυτη τιμή της συσχέτισης και  $f$  για να δηλώσουμε τη συχνότητα αυτής της συσχέτισης στον πίνακα συσχέτισης. Δεν λαμβάνουμε υπόψη τη διαγώνιο του πίνακα συσχέτισης. Τα αποτελέσματα που παρουσιάζονται στον Πίνακα 9.1 δείχνουν ότι σε όλες τις περιπτώσεις υπάρχει ένας μεγάλος αριθμός μη συσχετισμένων παραγόντων στον πίνακα του μοντέλου και στις περισσότερες περιπτώσεις ο συσχετισμός παραμένει χαμηλός, και ίσος με 0.33. Υπάρχουν υψηλότερες συσχετίσεις (0.66) μόνο αν θεωρήσουμε ένα μεγάλο μοντέλο (πάνω από 50 παραμέτρους που πρέπει να εκτιμηθούν με 24 πειραματικές εκτελέσεις), αλλά αυτές οι συσχετίσεις εμφανίζονται μόνο μεταξύ των κύριων επιδράσεων των παραγόντων ελέγχου και των  $C \times C \times N$  αλληλεπιδράσεων.

Ακόμα και σ' αυτές τις ακραίες περιπτώσεις, έχουμε μόνο μερικές συσχετίσεις που είναι ίσες με 0.66 ενώ οι υπόλοιπες είναι είτε ίσες με μηδέν είτε ίσες με 0.33. Αυτό το γεγονός δεν περιορίζει τα αποτελέσματά μας, αφού, μπορούμε ακόμη να εκτιμήσουμε το πλήρες μοντέλο. Παρόμοια, μπορούμε να παράγουμε τα αποτελέσματα για όλες τις

άλλες περιπτώσεις των σχεδιασμών που κατασκευάσαμε από τους πίνακες Hadamard τάξεως 20, 24, 28, 36 και 40, αλλά τα αποτελέσματα είναι αρκετά εκτενή για να μπορέσουν να συμπεριληφθούν στην παρούσα διδακτορική διατριβή.

### 9.3.2 Ανάλυση του προτεινόμενου σχεδιασμού

Η αντιμετώπιση πειραμάτων που περιλαμβάνουν πολλούς παράγοντες οδηγεί σε έναν πίνακα ο οποίος αποτελεί έναν ΥΣ. Αυτό το γεγονός σημαίνει ότι δεν μπορούν να εκτιμηθούν όλοι οι παράμετροι του μοντέλου. Ως εκ τούτου, εφαρμόζουμε τη μεθοδολογία των ποινικοποιημένων ελάχιστων τετραγώνων σε συνδυασμό με την επαναληπτική SIS, προκειμένου να εντοπιστούν οι σημαντικοί παράγοντες που επηρεάζουν την απόκριση που μας ενδιαφέρει. Πιο συγκεκριμένα, πραγματοποιούμε την ανάλυση σε δύο βήματα: το πρώτο είναι η εφαρμογή της επαναληπτικής SIS, για την επιλογή μιας αρχικής τιμής για τα ποινικοποιημένα ελάχιστα τετράγωνα, το δεύτερο βήμα είναι η εφαρμογή της μεθοδολογίας των ποινικοποιημένων ελαχίστων τετραγώνων χρησιμοποιώντας τη συνάρτηση SCAD. Με αυτό τον τρόπο επιτυγχάνουμε τόσο την επιλογή όσο και την εκτίμηση των σημαντικών επιδράσεων. Η διαδικασία που προτείνουμε περιγράφεται συνοπτικά στα ακόλουθα τέσσερα στάδια:

1. Θεώρησε το γραμμικό μοντέλο παλινδρόμησης  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ , όπου  $\mathbf{X} = [\mathbf{1}_N, \mathbf{D}]$  είναι ο πίνακας σχεδιασμού,  $\mathbf{y}$  είναι ένα  $N$ -διαστάσεων διάνυσμα,  $\boldsymbol{\beta}$  είναι ένα διάνυσμα μήκους  $r$ , όπου  $r$  είναι ο αριθμός των παραμέτρων του μοντέλου που αντιστοιχούν στις στήλες του πίνακα σχεδιασμού και  $\boldsymbol{\epsilon}$  είναι το κανονικά κατανομημένο σφάλμα.
2. Διάλεξε τους  $n - 1$  πιο σημαντικούς παράγοντες χρησιμοποιώντας την επαναληπτική SIS έτσι ώστε να επιλεγεί η αρχική τιμή για την μέθοδο SCAD.
3. Εφάρμοσε τα ποινικοποιημένα ελάχιστα τετράγωνα χρησιμοποιώντας τον πίνακα σχεδιασμού  $\mathbf{X}$  που περιγράφεται στο στάδιο 1, επέλεξε την αρχική τιμή που βρέθηκε στο στάδιο 2, και επέλεξε την παράμετρο ρύθμισης χρησιμοποιώντας το κριτήριο BIC.
4. Επέλεξε την παράμετρο  $\alpha = 3.7$  σύμφωνα με τους Li και Lin [81].

**Πίνακας 9.1:** Δομή συσχετίσεων για τον πίνακα του μοντέλου που περιλαμβάνει 24 πειραματικών εκτελέσεων που κατασκευάζεται από έναν πίνακα Hadamard τάξεως 12.

$p$	$q$	$t$	Πίνακας συσχετίσεων για τον πίνακα του μοντέλου που περιλαμβάνει													
			$M_1$		$C$		$C, C \times C$		$C, C \times N$		$C, C \times C \times N$		$C \times C, N$		$C \times C, C \times N$	
			$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$
3	3	27	0.00	558	0.00	6	0.00	30	0.00	96	0.00	78	0.00	30	0.00	78
			0.33	142			0.33	36		54			0.33	54		54
4	2	32	0.00	782	0.00	12	0.00	84	0.00	108	0.00	132	0.00	50	0.00	104
			0.33	210			0.33	6		24		108		6		78
5	1	31	0.00	750	0.00	20	0.00	180	0.00	90	0.00	120	0.00	80	0.00	120
			0.33	180			0.33	30				90		30		90
3	4	34	0.00	858	0.00	6	0.00	30	0.00	138	0.00	114	0.00	42	0.00	114
			0.33	264			0.33	72		96				42		96
4	3	43	0.00	1350	0.00	12	0.00	84	0.00	168	0.00	228	0.00	66	0.00	156
			0.33	456			0.33	6		72		234		6		150
5	2	47	0.00	1620	0.00	20	0.00	180	0.00	170	0.00	288	0.00	102	0.00	190
			0.33	530			0.33	30		40		300		30		190
			0.66	12								12				
6	1	43	0.00	1386	0.00	30	0.00	330	0.00	132	0.00	210	0.00	150	0.00	210
			0.33	420			0.33	90				210		90		210



Πίνακας 9.1: (συνέχεια)

$p$	$q$	$t$	Πίνακας συσχέτισης για τον πίνακα του μοντέλου που περιλαμβάνει																				
			$M_1$			$C$			$C, C \times C$			$C, C \times N$			$C, C \times C \times N$			$C \times C, N$			$C \times C, C \times N$		
			$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$	$r$	$f$			
2	6	27	0.00	582	0.00	2	0.00	6	0.00	122	0.00	56	0.00	42	0.00	96	0.33	120	0.33	60			
3	5	41	0.00	1220	0.00	6	0.00	30	0.00	186	0.00	156	0.00	56	0.00	156	0.33	420	0.33	150			
4	4	54	0.00	2052	0.00	12	0.00	84	0.00	236	0.00	336	0.00	84	0.00	216	0.33	798	0.33	246			
			0.66	12	0.66		0.33	6	0.33	144	0.33	408	0.33	6	0.33	246	0.66	12	0.66	12			
5	3	63	0.00	2790	0.00	20	0.00	180	0.00	260	0.00	524	0.00	126	0.00	270	0.33	1080	0.33	330			
			0.66	36	0.66		0.33	30	0.33	120	0.33	630	0.33	30	0.33	330	0.66	36	0.66	36			
6	2	65	0.00	3026	0.00	30	0.00	330	0.00	246	0.00	576	0.00	182	0.00	312	0.33	1110	0.33	390			
			0.66	24	0.66		0.33	90	0.33	60	0.33	660	0.33	90	0.33	390	0.66	24	0.66	24			
7	1	57	0.00	2352	0.00	42	0.00	546	0.00	182	0.00	336	0.00	252	0.00	336	0.33	840	0.33	420			
			0.33	840	0.33		0.33	210	0.33		0.33	420	0.33	210	0.33	420	0.66	210	0.66	210			

Η προκύπτουσα εκτίμηση αποτελείται από δύο κομμάτια: τα στοιχεία που θα είναι ακριβώς 0, τα οποία αντιστοιχούν στο συντελεστή των μη σημαντικών επιδράσεων και τα υπόλοιπα μη μηδενικά στοιχεία που αντιστοιχούν στις ενεργές επιδράσεις. Υποθέτουμε ότι οι παρατηρήσεις λαμβάνονται από το μοντέλο  $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}$ , όπου, χωρίς απώλεια της γενικότητας, όλα τα στοιχεία του πρώτου τμήματος  $\mathbf{X}_1$  είναι ενεργά, ενώ του δεύτερου τμήματος  $\mathbf{X}_2$  δεν είναι ενεργά. Ο εκτιμητής oracle  $\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{X}_1^T \mathbf{y}$  αποτελεί έναν ιδανικό εκτιμητή και η συνάρτηση SCAD είναι αυτή η οποία δίνει μια εκτίμηση που κατέχει την ιδιότητα oracle υπό την ασυμπτωτική έννοια. Σημειώνουμε ότι  $\hat{\boldsymbol{\beta}}_2 = 0$ . Συνεπώς, αυτή η διαδικασία ταυτόχρονα εξαιρεί τις μη ενεργές επιδράσεις και υπολογίζει τους συντελεστές παλινδρόμησης των ενεργών. Για την εφαρμογή της μεθόδου χρησιμοποιήσαμε κώδικες της R κάνοντας παράλληλα χρήση του πακέτου "SIS".

### 9.3.3 Παραδείγματα

Στο παρόν κεφάλαιο δείχνουμε, μέσω δύο παραδειγμάτων με προσομοιωμένα δεδομένα, πόσο αποτελεσματικά η προτεινόμενη μέθοδος μπορεί να εκτιμήσει τις παραμέτρους του μοντέλου σε ένα πρόβλημα εύρωστου παραμετρικού σχεδιασμού. Και στα δύο παραδείγματα, ο κύριος στόχος είναι να προσδιοριστούν τα βέλτιστα επίπεδα των παραγόντων ελέγχου, δηλαδή τα επίπεδα που ελαχιστοποιούν τη μέση απόκριση καθώς επίσης και τη διασπορά του μοντέλου.

**Πίνακας 9.2:** Πίνακες Hadamard που χρησιμοποιήθηκαν στη μελέτη.

Παραγόντων C: ελέγχου, N: θορύβου	Στήλες για τα μοντέλα				Hadamard	Εκτελέσεις σχεδιασμού	Χρήση
	1	2	3	4			
C=6,N=3	87				20 × 20	40	Παράδειγμα 1
C=5,N=5	95		45		20 × 20	40	Πειράματα προσομοίωσης Παράδειγμα 2 και Πειράματα προσομοίωσης
		100		50	20 × 20	51	
C=3,N=3	27		18		12 × 12	24	Πειράματα προσομοίωσης Πειράματα προσομοίωσης
		30		22	12 × 12	31	

#### Παράδειγμα 1

Σε αυτό το πείραμα εξετάζουμε έξι παράγοντες ελέγχου, A, B, C, D, E, F και τρεις παράγοντες θορύβου m, n, o, καθένα σε δύο επίπεδα. Ο στόχος είναι να ελαχιστοποιηθεί η μέση απόκριση, διατηρώντας παράλληλα τη διασπορά του μοντέλου

στο μικρότερο δυνατό επίπεδο. Αυτό θα επιτευχθεί με την κατάλληλη επιλογή των στάθμεων των παραγόντων ελέγχου. Στο παρόν παράδειγμα εξετάζουμε τα δεδομένα από την εργασία των Kunert et al. [80], όπου το πραγματικό μοντέλο είναι το ακόλουθο:

$$y = 4.08 - 0.85x_B + 0.63x_E + 0.88x_F + 0.31z_n + 0.20x_C z_n + 0.18x_E z_o - 0.18x_A x_B z_n + \epsilon. \quad (9.4)$$

Υποθέτουμε ότι το πειραματικό σφάλμα  $\epsilon$  ακολουθεί την  $N(0, \sigma^2)$  με το  $\sigma$  να ισούται με 0.043 και πριν προχωρήσουμε στην εκτίμηση του μοντέλου απόκρισης χρησιμοποιούμε τα δεδομένα από το σχεδιασμένο πείραμα. Αυτά τα στοιχεία παρουσιάζονται στον Πίνακα Δ.1 του Παραρτήματος Δ. Η εξίσωση 9.4 αποτελεί το εκτιμώμενο μοντέλο της απόκρισης που δίνεται στους Kunert et al. [80], το οποίο αποκτήθηκε με τη χρήση των δεδομένων από το διασταυρωμένο σχηματισμό. Χρησιμοποιώντας τον προτεινόμενο πειραματικό σχεδιασμό και θέτοντας  $m = 2$ , εκτελούμε το ίδιο πείραμα όπως και οι Kunert et al. [80], στο παράδειγμα προσομοίωσής μας. Πρώτον, όλοι οι παράγοντες ελέγχου και θορύβου κατανέμονται στις  $p + q = 9$  στήλες του  $40 \times 40$  πίνακα. Αυτό επιτεύχθηκε παίρνοντας τον πίνακα  $\mathbf{H}_N = [\mathbf{H}_n^{*T}, -\mathbf{H}_n^{*T}]^T$ , όπου  $\mathbf{H}_n^{*T}$  είναι ένας  $20 \times 20$  πίνακας Hadamard του οποίου η πρώτη στήλη είναι μια στήλη μόνο με μονάδες. Ο πίνακας σχεδιασμού  $\mathbf{X}$ , με διαστάσεις  $40 \times 87$ , που χρησιμοποιήθηκε στη γραμμική παλινδρόμηση, είναι ένας ΥΣ κάτι που σημαίνει ότι δεν μπορούν να είναι εκτιμήσιμες όλες οι παράμετροι του μοντέλου παλινδρόμησης. Ο Πίνακας Α.1 δείχνει το διάνυσμα των δεδομένων που λαμβάνεται από το πραγματικό μοντέλο απόκρισης χρησιμοποιώντας κώδικες της R. Εκτελώντας την ανάλυση σύμφωνα με την προαναφερθείσα διαδικασία τριών σταδίων (στην ενότητα 9.3.2), το εκτιμώμενο μοντέλο έχει ως εξής

$$\hat{y} = 4.079 - 0.84x_B + 0.62x_E + 0.87x_F + 0.30z_n + 0.19x_C z_n + 0.17x_E z_o - 0.18x_A x_B z_n. \quad (9.5)$$

Σημειώνουμε ότι επιλέξαμε την ποινή Lasso αντί της ποινής SCAD. Κάναμε αυτή την επιλογή διότι η ποινή Lasso έδωσε τελικά το επιθυμητό μοντέλο, ενώ η ποινή SCAD έχασε την αλληλεπίδραση τριών παραγόντων η οποία, σύμφωνα με το πραγματικό μοντέλο, φαίνεται να είναι σημαντική. Σύμφωνα με το RPDM θεωρούμε τους παράγοντες θορύβου να είναι τυχαίες μεταβλητές με  $E(\mathbf{z}) = 0$  και  $\text{Var}(\mathbf{z}) = \Sigma_z = \mathbf{I}$ .

Επομένως, το μοντέλο θέσης μπορεί να περιγραφεί ως

$$\widehat{E}[y] = 4.079 - 0.84x_B + 0.62x_E + 0.87x_F,$$

και το αντίστοιχο μοντέλο διασποράς ως

$$\widehat{\text{Var}}[y] = \text{constant} + (0.30 + 0.19x_C - 0.18x_Ax_B)^2 \text{Var}(z_n),$$

$$\widehat{\text{Var}}[y] = \text{constant} + 2(0.30)(0.19)x_C - 2(0.19)(0.18)x_Ax_B.$$

Τέλος, το μοντέλο διασποράς είναι το ακόλουθο

$$\widehat{\text{Var}}[y] = \text{constant} + 0.114x_C - 0.068x_Ax_B.$$

Ο όρος  $x_Ax_Bz_n$  έχει το  $z_n$  ως παράγοντα θορύβου και ο παράγοντας θορύβου  $z_n$  εμφανίζεται στην εξίσωση (9.5). Έτσι, η παρουσία της αλληλεπίδρασης  $x_Ax_Bz_n$  και του παράγοντα  $z_n$  σημαίνει ότι ο  $x_C$  και η αλληλεπίδραση  $x_Ax_B$  θα έπρεπε να συμπεριλαμβάνονται στο μοντέλο της διασποράς  $\widehat{\text{Var}}[y]$ . Ο κύριος στόχος είναι να ελαχιστοποιηθεί η απόκριση. Ωστόσο, την ίδια στιγμή πρέπει να διατηρήσουμε τη διακύμανση στο χαμηλότερο επίπεδο. Έτσι, εάν οι περιοχές ενδιαφέροντος για τις τιμές των μεταβλητών μας είναι  $x_i \leq 1$ ,  $i = 1, 2, \dots, 6$ , τότε υπολογίζονται οι βέλτιστες στάθμες των επιπέδων των παραγόντων ελέγχου μέσω του μοντέλου για το μέσο και του μοντέλου διακύμανσης ως  $(x_A, x_B, x_C, x_D, x_E, x_F) = (1, 1, -1, 0, -1, -1)$ . Οι στάθμες αυτές που ελαχιστοποιούν το μοντέλο διακύμανσης είναι ακριβώς ίδιες με αυτές που επιλέχθηκαν στο μοντέλο του μέσου, δεδομένου ότι οι μόνες μεταβλητές που επηρεάζουν το μοντέλο διακύμανσης είναι η κύρια επίδραση  $x_C$  και η αλληλεπίδραση δεύτερης τάξης  $x_Ax_B$ , όπου και οι δύο ελαχιστοποιούν τη διακύμανση με τις προαναφερθείσες στάθμες.

### **Παράδειγμα 2**

Υποθέτουμε ότι υπάρχουν πέντε παράγοντες ελέγχου, πέντε παράγοντες θορύβου και μία μεταβλητή απόκρισης. Ο κύριος στόχος είναι να επιλέξουμε τις στάθμες για τους παράγοντες ελέγχου που ελαχιστοποιούν ταυτόχρονα τη μέση απόκριση και τη διακύμανση του μοντέλου. Δανειζόμαστε, εν μέρει το παράδειγμα που παρουσιάζεται

στους Matsuura et al. [102]. Στο παράδειγμα τους δεν υπάρχουν αλληλεπιδράσεις τριών παραγόντων. Ως εκ τούτου, θεωρούμε το ακόλουθο μοντέλο αποκριτικής επιφάνειας

$$y = 4.08 - 0.85x_1 + 0.63x_4 + 0.88x_5 - 0.81x_1^2 + 0.6x_3^2 + 1.01x_5^2 + \\ + 0.31z_1 - 0.72x_1z_1 + 0.20x_2z_2 + 0.20x_4z_3 + 0.18x_1x_4z_1 + \epsilon,$$

όπου το πειραματικό σφάλμα  $\epsilon$  κατανέμεται σύμφωνα με την  $N(0, \sigma^2)$  με το  $\sigma$  να ισούται με 0.1. Με βάση αυτό το μοντέλο αποκριτικής επιφάνειας προσομοιώσαμε μία απόκριση, που παρουσιάζεται στον Πίνακα Δ.1 του Παραρτήματος Δ, έτσι ώστε για την εκτίμηση του μοντέλου αποκριτικής επιφάνειας να χρησιμοποιήσουμε τα δεδομένα από το σχεδιασμένο πείραμα. Σημειώνουμε ότι αν χρησιμοποιήσουμε έναν κεντρικό σύνθετο σχεδιασμό (Central Composite Design, CCD), τότε χρειαζόμαστε τουλάχιστον  $128 + 2p + n_0$  πειραματικές εκτελέσεις, έτσι ώστε όλες οι παράμετροι να είναι εκτιμήσιμες χωρίς όμως να περιέχουν τις αλληλεπιδράσεις τριών παραγόντων. Χρησιμοποιούμε τον πειραματικό σχεδιασμό που προτείνεται στο Θεώρημα 1 και εφαρμόζουμε το μοντέλο 2. Πρώτα απ' όλα, θεωρώντας  $m = 2$ , αντιστοιχούμε τους πέντε παράγοντες ελέγχου και τους πέντε παράγοντες θορύβου στις  $p + q = 10$  στήλες ενός  $40 \times 40$  πίνακα. Στη συνέχεια, θεωρούμε ένα σύνολο αξονικών σημείων  $2p = 10$ , όπου ο κάθε παράγοντας ελέγχου τίθεται διαδοχικά στο  $\pm\alpha$ , όπου επιλέξαμε ότι το  $\alpha = 1$ , και όλοι οι υπόλοιποι παράγοντες τοποθετούνται στο 0. Τέλος, επιλέξαμε να τοποθετήσουμε ένα κεντρικό σημείο που σημαίνει ότι το  $n_0 = 1$ . Συνοπτικά, οι πειραματικές εκτελέσεις που χρειάζονται στο σχεδιασμό είναι 51 οι οποίες μειώνουν δραματικά τις πειραματικές εκτελέσεις ενός CCD. Επιπλέον, ο προτεινόμενος σχεδιασμός περιέχει όλες τις αλληλεπιδράσεις  $C \times C \times N$ . Ως εκ τούτου, ο πίνακας σχεδιασμού αποτελείται από 101 παραμέτρους (μαζί με το μέσο όρο του μοντέλου). Ο πίνακας σχεδιασμού αποτελείται από 25 τοις εκατό περισσότερες πειραματικές εκτελέσεις από εκείνες των Matsuura et al. [102]. Ωστόσο, οι παράμετροι που πρέπει να εκτιμηθούν είναι κατά 200 τοις εκατό περισσότερες από των Matsuura et

α). [102]. Το εκτιμώμενο μοντέλο είναι τελικά το

$$\hat{y} = 4.09 - 0.85x_1 + 0.62x_4 + 0.85x_5 - 0.84x_1^2 + 0.73x_3^2 + 0.90x_5^2 + \\ + 0.30z_1 - 0.73x_1z_1 + 0.19x_2z_2 + 0.20x_4z_3 + 0.16x_1x_4z_1.$$

Το μοντέλο για το μέσο (ή μοντέλο θέσης) είναι το ακόλουθο

$$\widehat{E}[y] = 4.09 - 0.85x_1 + 0.62x_4 + 0.85x_5 - 0.84x_1^2 + 0.73x_3^2 + 0.90x_5^2,$$

και το μοντέλο διασποράς είναι το ακόλουθο

$$\widehat{\text{Var}}[y] = \text{Var}(0.30z_1 - 0.73x_1z_1 + 0.19x_2z_2 + 0.20x_4z_3 + 0.16x_1x_4z_1) = \\ = \text{constant} + (0.30 - 0.73x_1 + 0.16x_1x_4)^2 \text{Var}(z_1) + (0.19x_2)^2 \text{Var}(z_2) \\ + (0.20x_4)^2 \text{Var}(z_3) = \\ = \text{constant} - 2(0.30)(0.73)x_1 + 2(0.73)(0.16)x_1x_4 = \text{constant} - 0.44x_1 + 0.23x_1x_4.$$

Ο κύριος στόχος είναι να ελαχιστοποιηθεί η απόκριση διατηρώντας τη διακύμανση στο χαμηλότερο επίπεδο. Έτσι, εάν οι περιοχές ενδιαφέροντος για τις τιμές των μεταβλητών είναι  $x_i \leq 1$ ,  $i = 1, 2, \dots, 6$ , τότε οι βέλτιστες στάθμες των επιπέδων των παραγόντων ελέγχου υπολογίζονται μέσω του μοντέλου για το μέσο και του μοντέλου διακύμανσης. Εάν θεωρήσουμε αρχικά το μοντέλο του μέσου πρέπει να επιλέξουμε τις στάθμες των παραγόντων ως  $(x_1, x_2, x_3, x_4, x_5) = (1, 0, 0, -1, -1)$ . Αν θεωρήσουμε μόνο τους παράγοντες που ελαχιστοποιούν τη διασπορά του μοντέλου απόκρισης θα πρέπει να ορίσουμε τη μεταβλητή  $(x_1)$  να είναι ίση με 1 ενώ η  $(x_4)$  θα πρέπει να οριστεί ίση με  $-1$ . Καταλήγουμε στο συμπέρασμα ότι οι στάθμες είναι συνεπείς και για τα δύο μοντέλα. Έτσι, προτείνουμε τις στάθμες να είναι ως εξής:  $(x_1, x_2, x_3, x_4, x_5) = (1, 0, 0, -1, -1)$ . Σημειώνουμε ότι η στάθμη των τετραγωνικών όρων μπορεί να επιλεγεί έτσι ώστε να επωφελείται όλη η διαδικασία.

## 9.4 Πειράματα προσομοίωσης

Σε αυτή την ενότητα παρέχουμε μια συγκριτική μελέτη με προσομοιωμένα δεδομένα ώστε να εξετάσουμε την απόδοση της προτεινόμενης μεθοδολογίας. Οι πίνακες

που αναφέρονται σε αυτή την ενότητα μπορούν να βρεθούν στο Παράρτημα Δ. Το πρώτο μέρος του παρόντος κεφαλαίου αποτελεί μία μελέτη προσομοίωσης για το μοντέλο 1. Θεωρήσαμε δύο διαφορετικά σενάρια :

1.  $p = 3, q = 3$  και  $m = 1$

2.  $p = 5, q = 5$  και  $m = 2$

προκειμένου να παρουσιάσουμε την αποτελεσματικότητα της μεθόδου μας σε διαφορετικές καταστάσεις. Το δεύτερο μέρος του παρόντος τμήματος αποτελεί μια μελέτη προσομοίωσης για το μοντέλο αποκριτικής επιφάνειας, δηλαδή για το μοντέλο 2, στο οποίο εφαρμόσαμε τα προαναφερθέντα δύο σενάρια. Σημειώνουμε ότι για το μοντέλο 2 επιλέξαμε  $\alpha = 1$  και  $n_0 = 1$ . Τα αριθμητικά πειράματα διεξήχθησαν, όταν ένας με δύο από τους παράγοντες ελέγχου και από τους παράγοντες θορύβου είναι ενεργοί, όπως παρουσιάζεται και στις ακόλουθες περιπτώσεις. Έτσι, τα ποσοστά σφάλματος τύπου I και τύπου II στον ΥΣ εξετάζονται για τρεις διαφορετικές περιπτώσεις όσον αφορά τους ενεργούς παράγοντες που επηρεάζουν την απόκριση, δηλαδή για τις περιπτώσεις 1 έως 3, οι οποίες θεωρούνται ως εξής :

**Περίπτωση I :** ένας παράγοντας ελέγχου και ένας παράγοντας θορύβου ενεργοί.

**Περίπτωση II :** δύο παράγοντες ελέγχου και ένας παράγοντας θορύβου ενεργοί.

**Περίπτωση III :** δύο παράγοντες ελέγχου και δύο παράγοντες θορύβου ενεργοί.

Για απλούστευση υποθέτουμε ότι όλες οι επιδράσεις των ενεργών παραγόντων είναι σταθερές και ίσες με  $\beta$  ενώ οι επιδράσεις των μη ενεργών παραγόντων ίσες με 0. Η τιμή του  $\beta$  ισούται με  $1\sigma, 2\sigma$  και  $3\sigma$  αντίστοιχα, και χωρίς απώλεια της γενικότητας θεωρήσαμε  $\sigma = 1$ . Επιπλέον, θεωρήσαμε τις περιπτώσεις όπου το  $\beta_i$  έχει κατανομηθεί τυχαία ως συν ή πλην δείχνοντας την ευρωστία της μεθοδολογίας μας. Θεωρήσαμε ότι το πειραματικό σφάλμα κατανέμεται σύμφωνα με την  $N(0, \sigma^2)$  και ότι η τιμή του μέσου  $\mu$  των μοντέλων ισούται με 100. Συνολικά πραγματοποιήθηκαν 1000 αριθμητικά πειράματα, σε κάθε περίπτωση, για τα δύο διαφορετικά σενάρια των παραγόντων ελέγχου και θορύβου.

Οι προτεινόμενοι πίνακες σχεδιασμού που χρησιμοποιήθηκαν στην παρούσα μελέτη προσομοίωσης και των τεσσάρων μοντέλων παρέχονται στον Πίνακα Α.1. Πρέπει να αναφερθεί ότι εκτελούμε πρόσθετες προσομοιώσεις λαμβάνοντας υπόψη τα μοντέλα 1 και 2 χωρίς να συμπεριλαμβάνονται αλληλεπιδράσεις τριών παραγόντων.

Αναφέρουμε αυτά τα μοντέλα ως μοντέλα 3 και 4 και παρουσιάζουμε τα αποτελέσματα που αποκομήθηκαν στους Πίνακες Δ.10-Δ.13 του Παραρτήματος Δ. Έχουμε επίσης διεξαγάγει αριθμητικά πειράματα για ένα CCD με  $n = 128 + 2p + n_0 = 128 + 10 + 1 = 139$  πειραματικές εκτελέσεις, τα αποτελέσματα των οποίων παρουσιάζονται στον Πίνακα Δ.14 του Παραρτήματος Δ.

Οι Πίνακες Δ.2-Δ.14 του Παραρτήματος Δ δείχνουν τα αποτελέσματα των ποσοστών επιλογής καθώς και τις εκτιμήσεις των συντελεστών κάθε μοντέλου σε κάθε μία από τις περιπτώσεις που θεωρήσαμε. Πιο συγκεκριμένα, σε κάθε στήλη  $\beta$  υπάρχουν 1000 πειραματικές εκτελέσεις. Οι αριθμοί στις στήλες “ $E(\beta_i|\beta_i \neq 0)$ ” και “ $V(\beta_i|\beta_i \neq 0)$ ” υποδεικνύουν τις εκτιμήσεις για το μέσο και τη διακύμανση αντίστοιχα, για την παράμετρο  $\beta_i$  με δεδομένο ότι  $\beta_i \neq 0$ . Όπως μπορεί να παρατηρηθεί οι επιδράσεις των ενεργών παραγόντων μπορεί να βρεθούν, με μεγάλη πιθανότητα, όταν ο αριθμός των επιδράσεων των ενεργών παραγόντων είναι περίπου το ένα τρίτο του αριθμού των πειραματικών εκτελέσεων για το μέγεθος των επιδράσεων που κυμαίνονται από  $1\sigma$  έως  $3\sigma$  και που ορίζονται τυχαία ως συν ή μείον. Επιπροσθέτως, αυτοί οι πίνακες παρουσιάζουν τα αποτελέσματα των σφαλμάτων Τύπου I και Τύπου II που βρέθηκαν για τις περιπτώσεις 1 έως 3. Για παράδειγμα, στον Πίνακα Δ.4 του Παραρτήματος Δ, το ποσοστό επιλογής όλων των ενεργών παραγόντων είναι περίπου 100% (δηλαδή, σφάλμα Τύπου II 0 %) για όλες τις εξεταζόμενες περιπτώσεις.

Συνοπτικά, για τα μοντέλα 1 και 3 η χρήση των προτεινόμενων σχεδιασμών παρέχει υψηλή πιθανότητα για την εύρεση των επιδράσεων των ενεργών παραγόντων στην πλειοψηφία των περιπτώσεων που θεωρήσαμε και για τα δύο σενάρια. Σημειώνουμε ότι υπάρχουν ορισμένες περιπτώσεις όπου εμφανίζεται ποσοστό σφάλματος Τύπου I. Αυτό σημαίνει ότι ορισμένοι παράγοντες προσδιορίζονται ως ενεργοί, ενώ είναι ανενεργοί. Ωστόσο, το ποσοστό αυτού του σφάλματος είναι αμελητέο κάτι που φαίνεται και από τους Πίνακες Δ.2-Δ.5 και τους Πίνακες Δ.10-Δ.11 του Παραρτήματος Δ για τα μοντέλα 1 και 3, αντίστοιχα.

Όσον αφορά τα μοντέλα αποκριτικής επιφάνειας, όπως είναι τα μοντέλα 2 και 4, οι πειραματικές εκτελέσεις για τον προτεινόμενο σχεδιασμό αυξήθηκαν αφού υπάρχουν επιπλέον  $2p = 10$  αξονικά σημεία και ένα κεντρικό σημείο. Τα αριθμητικά αποτελέσματα είναι αξιοσημείωτα, αφού παρατηρήσαμε ότι σε ορισμένες περιπτώσεις, οι τετραγωνικές επιδράσεις δεν μπορούν να εκτιμηθούν. Αυτό συμβαίνει όταν το



μέγεθος του συντελεστή  $\beta$  είναι αρκετά μικρό, για παράδειγμα στην περίπτωση όπου το  $\beta = 1$ . Το γεγονός αυτό παρατηρήθηκε και για το σχεδιασμό CCD ο οποίος όμως έχει και το μειονέκτημα του αρκετά μεγάλου αριθμού πειραματικών εκτελέσεων σε σύγκριση με τον προτεινόμενο σχεδιασμό. Για να μην αναφέρουμε το γεγονός ότι η εκτίμηση των επιδράσεων χρησιμοποιώντας τη μέθοδό μας είναι εξαιρετικά καλή (βλ. Πίνακες Δ.6- Δ.9 και τους Πίνακες Δ.12- Δ.14 του Παραρτήματος Δ για το μοντέλο 2 και 4 αντίστοιχα).

Εν κατακλείδι, η εφαρμογή τόσο του σχεδιασμού όσο και της προτεινόμενης διαδικασίας στους εύρωστους παραμετρικούς σχεδιασμούς είναι πραγματικά αποτελεσματική για τον προσδιορισμό των ενεργών επιδράσεων, ειδικά όταν οι ενεργοί παράγοντες είναι περίπου το ένα τρίτο των πειραματικών εκτελέσεων.

## 9.5 Συμπεράσματα

Στο παρόν κεφάλαιο παρουσιάζεται μία συγκριτική μελέτη της εφαρμογής των ΥΣ σε προβλήματα εύρωστων παραμετρικών σχεδιασμών. Οι ΥΣ που κατασκευάστηκαν χρησιμοποιώντας το fold-over ενός  $(8m + 4) \times (8m + 4)$  πίνακα Hadamard ( $m$  είναι ένας φυσικός αριθμός) αντί για έναν ορθογώνιο σχεδιασμό δύο επιπέδων, επιτυγχάνουν μια αξιοσημείωτη μείωση των πειραματικών εκτελέσεων. Επιπρόσθετα, μελετήθηκαν και άλλες διαστάσεις πινάκων Hadamard καθώς και μια πλήρης δομή συσχετίσεων για διαφορετικές περιπτώσεις. Τα αποτελέσματα δείχνουν ότι στην πλειονότητα των περιπτώσεων οι παράγοντες στον πίνακα του μοντέλου παραμένουν ασυσχέτιστοι.

Θεωρήσαμε πολλά διαφορετικά μοντέλα. Το πρώτο αποτελείται από τις κύριες επιδράσεις  $(C, N)$  και τις αλληλεπιδράσεις  $C \times N$ ,  $C \times C$  και  $C \times C \times N$ , ενώ το δεύτερο αποτελεί ένα μοντέλο αποκριτικής επιφάνειας αφού περιέχει όλες τις επιδράσεις του προηγούμενου μοντέλου μαζί με τις τετραγωνικές επιδράσεις ( $C^2$ ) των παραγόντων ελέγχου. Εκτός από αυτούς τους σχεδιασμούς εκτελέσαμε τον προτεινόμενο σχεδιασμό και την προτεινόμενη διαδικασία ανάλυσης σε δύο ακόμα σχεδιασμούς σύμφωνα με τον εύρωστο παραμετρικό σχεδιασμό, δηλαδή αυτούς τους σχεδιασμούς με όλες τις παραπάνω επιδράσεις χωρίς όμως να ενσωματώσουμε αλληλεπιδράσεις τριών παραγόντων. Ωστόσο, πρέπει να σημειώσουμε ότι οι  $C \times C \times N$  αλληλεπιδράσεις

στο RPDM είναι το ίδιο σημαντικές με τις  $C \times C$  αλληλεπιδράσεις και η σημασία τους μπορεί να υποδηλώνει ότι το ζεύγος των παραγόντων ελέγχου θα πρέπει να προσαρμοστεί από κοινού έτσι ώστε να μειωθεί ελαφρώς η επίδραση του παράγοντα θορύβου (που περιέχεται στην αλληλεπίδραση τριών παραγόντων) στην απόκριση. Έτσι, όλες οι επιδράσεις μπορούν να αντιστοιχηθούν στις στήλες του ΥΣ χωρίς την ύπαρξη δύο στηλών που είναι πλήρως ταυτόσημες όπως αποδείχθηκε στο Θεώρημα 1.

Μέσα από τα αριθμητικά πειράματα αποκαλύπτεται ότι αν ο εκτιμώμενος αριθμός των ενεργών παραγόντων είναι μικρότερος από το ένα τρίτο περίπου του αριθμού των ενεργών παραμέτρων, τότε ο προτεινόμενος ΥΣ, με την τιμή του  $m$  να είναι τέτοια ώστε ο εκτιμώμενος αριθμός ενεργών επιδράσεων να είναι μικρότερος από περίπου το ένα τρίτο του αριθμού των πειραματικών εκτελέσεων, είναι ιδιαίτερα αποτελεσματικός για την εξεύρεση όλων των ενεργών επιδράσεων. Για να αναλύσουμε τέτοιους σχεδιασμούς σύμφωνα με τον RPDM, εφαρμόζουμε την επαναληπτική διαδικασία SIS σε συνδυασμό με την SCAD, προκειμένου να μπορέσουμε ταυτόχρονα να ανιχνεύσουμε και να εκτιμήσουμε τις ενεργές επιδράσεις. Εξετάσαμε τα ποσοστά σφάλματος Τύπου I και Τύπου II και παρατηρήσαμε ότι και τα δύο διατηρούνται στο χαμηλότερο επίπεδο. Πιο συγκεκριμένα, η μέθοδός μας φαίνεται να αποδίδει αρκετά καλά, ειδικά όσον αφορά το ποσοστό επιλογής των ενεργών παραγόντων (δηλαδή το ποσοστό σφάλματος Τύπου II) στις περισσότερες περιπτώσεις.

Πρέπει να επισημανθεί ότι υπάρχουν πολλά ζητήματα που απαιτούν μελλοντική έρευνα. Για παράδειγμα, το μέγεθος των ενεργών επιδράσεων θα μπορούσε να επιλεγεί σύμφωνα με τη σημαντικότητα των επιδράσεων.

---

## Υπερκορεσμένοι Split-Plot Σχεδιασμοί για Εύρωστα Παραμετρικά Πειράματα

---

**Σ**τόχος των εύρωστων παραμετρικών σχεδιασμών, όπως έχουμε ήδη εκτενώς αναφερθεί στα προηγούμενα κεφάλαια, είναι ο προδιορισμός της επίδρασης των παραγόντων ελέγχου στη διασπορά που οφείλεται στους παράγοντες που δεν μπορούν να ελεγχθούν-παράγοντες θορύβου. Οι split-plot σχεδιασμοί μπορούν να εφαρμοστούν αποτελεσματικά για να μειώσουμε το κόστος στους εύρωστους παραμετρικούς σχεδιασμούς (robust parameter designs, RPDs), δηλαδή, όταν υπάρχουν παράγοντες που είναι δύσκολο να αλλάξουν ή δύσκολο να ελεγχθούν. Υπό την παρουσία πολλών παραγόντων, είτε οι παραδοσιακοί επαναλαμβανόμενοι split-plot σχεδιασμοί, είτε οι σχεδιασμοί των καρτεσιανών γινομένων που προτάθηκαν από τον Taguchi, δεν είναι τόσο πρακτικοί δεδομένου ότι απαιτούν μεγάλο αριθμό πειραματικών εκτελέσεων. Έτσι, είναι απαραίτητο να χρησιμοποιηθούν κλασματικοί παραγοντικοί σχεδιασμοί, όπως είναι οι ΥΣ. Στο παρόν κεφάλαιο προτείνουμε μια κατασκευή για υπερκορεσμένους split-plot σχεδιασμούς (SSSPD) κατάλληλους για RPD πειράματα. Εφαρμόσαμε τη μεθοδολογία των σύνθετων ορθογώνιων σχεδιασμών χρησιμοποιώντας ορθογώνιους σχεδιασμούς ισχύος μεγαλύτερης από 1, για τις split-plot μονάδες και ισορροπημένους ΥΣ για όλες τις whole-plot μονάδες. Η προτεινόμενη μεθοδολογία υποδεικνύει ότι τα πειράματα που διεξάγονται υπό το πρίσμα των split-plot σχεδιασμών μπορούν να έχουν τεράστια αξία στη μεθοδολογία των RPD, αφού επιτρέπουν την εκτίμηση των επιδράσεων που μας ενδιαφέρουν με πολύ πιο εύκολο τρόπο από τους σχεδιασμούς που προτάθηκαν από τον Taguchi. Τα ερευνητικά αποτελέσματα που παρουσιάζονται στο παρόν κεφάλαιο δημοσιεύτηκαν στην επιστημονική εργασία [VII].

## 10.1 Ερευνητικό πρόβλημα-Λύση

Σε βιομηχανικά πειράματα, οι σχεδιασμοί διηρημένων τεμαχίων ή αλλιώς οι split-plot σχεδιασμοί έχουν κερδίσει σε σημαντικό βαθμό την προσοχή των πειραματιστών. Υπό την παρουσία split-plot δομής, ορισμένοι παράγοντες είναι πιο δαπανηροί να αλλάξουν από άλλους παράγοντες [9, 69, 120] λόγω του γεγονότος ότι είτε μια συγκεκριμένη θεραπεία είναι δαπανηρή είτε χρονοβόρα για να εκτελεστεί. Οι παράγοντες που η αλλαγή τους είναι δαπανηρή εφαρμόζονται σε μεγάλες παρτίδες (whole-plot μονάδες), ενώ οι παράγοντες που η αλλαγή τους είναι εύκολη εφαρμόζονται σε μικρές παρτίδες (split-plot μονάδες) μέσα στις μεγαλύτερες.

Όπως έχουμε ήδη αναφέρει εκτενώς στα προηγούμενα κεφάλαια η RPD μεθοδολογία στοχεύει στην επιλογή των στάθμεων για τους παράγοντες που μπορούν να ελεγχθούν (παράγοντες ελέγχου  $C$ ) που οδηγούν σε προϊόντα ή διαδικασίες που είναι λιγότερο ευαίσθητες στις επιδράσεις των μη ελεγχόμενων παραγόντων (των παραγόντων θορύβου  $N$ ). Η βέλτιστη επιλογή των παραγόντων ελέγχου είναι αυτή που οδηγεί τη μέση απόκριση πλησιέστερα στο στόχο οδηγώντας παράλληλα και στη μικρότερη δυνατή διασπορά καθώς οι παράγοντες θορύβου αλλάζουν. Η χρήση των διασταυρωμένων σχηματισμών που προτάθηκαν από τον Taguchi [131] οδηγεί σε πολύ κοστοβόρα πειράματα όσον αφορά στο μέγεθος των εκτελέσεων. Πρόσφατα, διάφοροι συγγραφείς, ξεκινώντας από τους Box και Jones [16], έδειξαν πώς μπορούν να χρησιμοποιηθούν επαρκώς οι split-plot σχεδιασμοί για τη μείωση του κόστους των πειραμάτων RPD καθώς και για την αύξηση της αποτελεσματικότητας παρουσία παραγόντων θορύβου.

Οι Box και Jones [16] συζήτησαν τρεις τρόπους με τους οποίους οι split-plot σχεδιασμοί μπορούν να υπάρξουν σε εύρωστα παραμετρικά πειράματα. Ο πρώτος θεωρεί ότι οι παράγοντες θορύβου (noise factors) είναι δύσκολο να αλλάξουν (whole-plot), ενώ οι παράγοντες ελέγχου (control factors) είναι εύκολο να αλλάξουν (split-plot). Ο δεύτερος θεωρεί τους παράγοντες ελέγχου (control factors) δύσκολο να αλλάξουν (whole-plot), ενώ τους παράγοντες θορύβου (noise factors) εύκολο να αλλάξουν (split-plot), και στον τελευταίο, οι Box και Jones [16] εξέτασαν τη χρήση strip-plots, τα οποία μπορούν να μειώσουν περαιτέρω το κόστος του πειράματος.

Γενικά, στα RPD πειράματα, το ενδιαφέρον επικεντρώνεται στις επιδράσεις των παραγόντων ελέγχου καθώς και των αλληλεπιδράσεων μεταξύ των παραγόντων ελέγ-

χου επί των παραγόντων θορύβου. Οι κύριες επιπτώσεις των παραγόντων θορύβου είναι δευτερεύοντος ενδιαφέροντος, καθώς δεν επηρεάζουν στον προσδιορισμό των βέλτιστων στάθμεων. Δεδομένου αυτού, οι Box και Jones [16] καταλήγουν στο συμπέρασμα ότι οι παράγοντες θορύβου είναι καταλληλότερο να χρησιμοποιηθούν ως whole-plots. Αυτός είναι και ο λόγος για τον οποίο θεωρήσαμε τους παράγοντες ελέγχου ως split-plots και τους παράγοντες θορύβου ως whole-plots.

Οι ερευνητές έχουν διερευνήσει το ερώτημα του πώς να επιλεγούν οι καλύτεροι split-plot σχεδιασμοί για ένα δεδομένο σύνολο παραγόντων ελέγχου και θορύβου [10, 12]. Μέχρι στιγμής έχει δημοσιευθεί ένας σημαντικός αριθμός ερευνητικών εργασιών, η πλειοψηφία των οποίων επικεντρώνεται σε κλασματικούς split-plot σχεδιασμούς [8, 9, 12, 79]. Σε περιπτώσεις όπου είναι σημαντικό να εξετάσουμε ένα μεγάλο αριθμό παραγόντων με περιορισμένο αριθμό πειραματικών εκτελέσεων με μία split-plot δομή, η χρήση υπερκορεσμένων split-plot σχεδιασμών (SSSPDs) μπορεί να αποδειχθεί ένας οικονομικός και ταυτόχρονα αποδοτικός τρόπος κρησαρίσματος. Ωστόσο, παρόλο που η ενσωμάτωση ΥΣ στην split-plot δομή είναι ένας ιδιαίτερα αποδοτικός τρόπος πειραματισμού, η συγκεκριμένη περιοχή παραμένει ανεξερεύνητη. Οι πιο πρόσφατες δημοσιευμένες εργασίες περιλαμβάνουν αυτή της Koh et al. [73] η οποία κατασκεύασε SSSPDs χρησιμοποιώντας έναν  $E(s^2)$ -βέλτιστο ΥΣ για τους whole-plot παράγοντες και μισά κλάσματα Plackett-Burman σχεδιασμών για τους split-plot παράγοντες. Η δική μας κατασκευή, που παρουσιάζεται στο Θεώρημα 1, είναι ένα γενικό θεώρημα κατασκευής που θεωρεί αυτό των Koh et al. [73] ως ειδική περίπτωση.

Στην παρούσα ενότητα παρουσιάζουμε μία μέθοδο κατασκευής SSSPDs κατάλληλων για RPD πειράματα. Εφαρμόσαμε τη μεθοδολογία των σύνθετων ορθογώνιων σχεδιασμών χρησιμοποιώντας ορθογώνιους σχεδιασμούς ισχύος μεγαλύτερης από 1, για τις split-plot μονάδες και ισορροπημένους ΥΣ για όλες τις whole-plot μονάδες. Επιπλέον παρουσιάζουμε τον πίνακα των ταυτώσεων επιδράσεων των προτεινόμενων σχεδιασμών καθώς επίσης παρουσιάζουμε και έναν αποδοτικό τρόπο ανάλυσης.

## 10.2 Θεωρητικό υπόβαθρο

Σε αυτή την ενότητα περιγράφουμε τις βασικές έννοιες του RPDM, τα υπό εξέταση μοντέλα καθώς επίσης και τη μεθοδολογία των σύνθετων ορθογώνιων σχεδιασμών.

### 10.2.1 Μεθοδολογία Εύρωστων Παραμετρικών Σχεδιασμών

Για  $1 \leq u \leq N$ , έστω  $x_{u1}, x_{u2}, \dots, x_{up}$  τα επίπεδα των  $p$  παραγόντων ελέγχου και  $z_{u1}, z_{u2}, \dots, z_{uq}$  τα επίπεδα των  $q$  παραγόντων θορύβου που αντιστοιχούν στη  $u$ -οστή πειραματική εκτέλεση. Έστω  $y_u$  η απόκριση που μας ενδιαφέρει και αντιστοιχεί στη  $u$ -οστή πειραματική εκτέλεση. Θεωρούμε δύο μοντέλα RPD.

Το πρώτο, που αναφέρεται ως μοντέλο 1, περιγράφεται ως εξής:

$$y_k = \mu + \sum_{i=1}^p a_i x_{ki} + \sum_{j=1}^q b_j z_{kj} + \sum_{i=1}^p \sum_{j=1}^q c_{ij} x_{ki} z_{kj} + \epsilon_k, \quad k = 1, 2, \dots, n \quad (10.1)$$

Το δεύτερο μοντέλο, που αναφέρεται ως μοντέλο 2, περιγράφεται ως εξής

$$y_k = \mu + \sum_{i=1}^p a_i x_{ki} + \sum_{i_1 < i_2 = 1}^p a_{i_1 i_2} x_{ki_1} x_{ki_2} + \sum_{j=1}^q b_j z_{kj} + \sum_{i=1}^p \sum_{j=1}^q c_{ij} x_{ki} z_{kj} + \epsilon_k, \quad k = 1, 2, \dots, n. \quad (10.2)$$

Υποθέτουμε ότι τα επίπεδα των παραγόντων θορύβου δεν είναι ελεγχόμενα και είναι τυχαίες μεταβλητές στη διαδικασία έτσι ώστε  $E[\mathbf{z}] = \mathbf{0}$  και  $Var[\mathbf{z}] = \Sigma_z$ . Σύμφωνα με αυτή την υπόθεση λαμβάνουμε το μοντέλο του μέσου και το μοντέλο της διακύμανσης. Η μεθοδολογία RPDM στοχεύει στην βελτιστοποίηση της μέσης απόκρισης διατηρώντας τη διασπορά της διαδικασίας στο χαμηλότερο επίπεδο. Αυτό μπορεί να επιτευχθεί με σωστή επιλογή των επιπέδων των παραγόντων ελέγχου, έτσι ώστε η διαδικασία να μην επηρεάζεται από τις αλλαγές των παραγόντων θορύβου. Αυτό μπορεί να επιτευχθεί με την εξέταση της αλληλεπίδρασης μεταξύ παραγόντων ελέγχου και θορύβου.



και  $D$  είναι ένας ορθογώνιος πίνακας με  $n_1 n_2$  πειραματικές εκτελέσεις και  $m_1 + m_2$  παράγοντες ισχύος  $t_3 = \min\{t_1, t_2\}$ . Ο πίνακας  $D$  συμβολίζεται ως

$$COA((n_1, n_2), (m_1, m_2), 2, (t_1, t_2, t_3)).$$

Οι πρώτοι  $m_1$  παράγοντες πρόκειται να χρησιμοποιηθούν για τους παράγοντες θορύβου, και οι τελευταίοι  $m_2$  παράγοντες πρόκειται να χρησιμοποιηθούν για τους παράγοντες ελέγχου. Οι σχεδιασμοί που χρησιμοποιήθηκαν στη μεθοδολογία του Taguchi αποτελούν μια ειδική περίπτωση των σύνθετων ορθογώνιων σχεδιασμών που αντιστοιχούν στην ίδια επιλογή για καθένα από τα  $b_i$ . Ο Rosenbaum [115] έδειξε πώς μπορεί κανείς να αποκτήσει σχεδιασμούς που έχουν υψηλότερη αναλυτική τάξη από τους διασταυρωμένους σχηματισμούς για την ανάλυση των  $y_i$  επιτρέποντας ακόμα και την ανάλυση  $s_i^2$ .

### 10.3 Κατασκευή σχεδιασμών και δομή ταυτόσημων επιδράσεων

Σε αυτή την ενότητα αναπτύσσουμε τους προτεινόμενους σχεδιασμούς προτείνοντας ένα γενικό θεώρημα κατασκευής καθώς και δύο ειδικές περιπτώσεις κατάλληλες για το μοντέλο 1 και το μοντέλο 2, αντίστοιχα. Συγκεκριμένα προτείνουμε την κατασκευή ενός split-plot σχεδιασμού για RPD πειράματα χρησιμοποιώντας ορθογώνιους σχηματισμούς ισχύος μεγαλύτερης από 1, δηλαδή έναν  $OA(n_2, m_2, 2, t_2)$ , για  $t_2 \geq 1$  για τα split-plots και οποιονδήποτε ισορροπημένο ΥΣ, δηλαδή έναν  $OA(n_1, m_1, 2, t_1)$ ,  $t_1 = 1$  για τις whole-plot μονάδες.

Αν και οι προτεινόμενοι σχεδιασμοί μπορούν να εφαρμοστούν αποτελεσματικά σε οποιουδήποτε τύπου split-plot πειράματα, εμείς εξετάζουμε την περίπτωση των πειραμάτων κρησαρίσματος μέσα σε μια δομή RPD. Τα πειράματα περιλαμβάνουν παράγοντες ελέγχου (C) και θορύβου (N), όλοι σε δύο επίπεδα, με τους παράγοντες ελέγχου να έχουν αντιστοιχηθεί στα split-plots και οι παράγοντες θορύβου στα whole-plots. Στόχος αυτών των σχεδιασμών είναι να εκτιμηθεί πλήρως το μοντέλο πρώτης σημαντικότητας το οποίο αποτελείται κυρίως από τις κύριες επιδράσεις των whole-plots και των split-plots καθώς επίσης και από μερικές αλληλεπιδράσεις δύο



παραγόντων μεταξύ των whole-plots και των split-plots (Πίνακας 8.1).

### 10.3.1 Κατασκευή

Ας θεωρήσουμε μια κλάση  $n$ -εκτελέσεων σχεδιασμών  $D(n, 2^m, t)$  που αφορούν  $m$  παράγοντες κάθε έναν σε δύο επίπεδα, έτσι ώστε κάθε σχεδιασμός που ανήκει σε αυτή την κλάση να είναι ένας  $OA(n, m, 2, t)$  ισχύος  $t \geq 1$ . Είναι προφανές ότι οποιοσδήποτε σχεδιασμός που ανήκει στο  $D(n, 2^m, t)$  με την προϋπόθεση ότι  $t = 1$  και  $n \leq m$ , είναι ένας ισορροπημένος ΥΣ.

**Θεώρημα 1** Θεωρούμε έναν σχεδιασμό  $D_1 \in D(n_1, 2^{m_1}, t_1)$  με πίνακα σχεδιασμού  $X_W$ . Για την περίπτωση  $t_1 = 1$ , υποθέτουμε ότι οι στήλες του πίνακα σχεδιασμού  $X_W$  είναι ανά δύο ανεξάρτητες (δεν ταυτίζονται). Για  $1 \leq i \leq n_1$ , κάθε  $b_i \in D(n_2, 2^{m_2}, t_2)$ ,  $t_2 \geq 1$ , με πίνακα σχεδιασμού  $X_{S_i}$ . Ορίζουμε το

$$D_2 = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n_1} \end{bmatrix}$$

έτσι ώστε ο  $D_2$  να αποτελεί έναν  $OA$  ισχύος μεγαλύτερης ή ίσης με 1. Έστω  $D^{(COA)}(n_1 n_2, 2^{m_1+m_2})$  μία κλάση από  $n_1 n_2$ -εκτελέσεων σχεδιασμούς που περιλαμβάνει  $(m_1 + m_2)$  παράγοντες, κάθε ένας σε δύο επίπεδα, έτσι ώστε κάθε σχεδιασμός  $D \in D^{(COA)}(n_1 n_2, 2^{m_1+m_2})$  να είναι ένας  $COA$  με μία κατάλληλη επιλογή παραμέτρων. Ακολουθώντας τον ορισμό 2, ορίζουμε ( $D \in D^{(COA)}(n_1 n_2, 2^{m_1+m_2})$ ) ως

$$D = \begin{bmatrix} d_1^T \\ \vdots \\ d_1^T \\ \vdots \\ d_{n_1}^T \\ \vdots \\ d_{n_1}^T \end{bmatrix} \begin{bmatrix} b_1 \\ b_{n_1} \end{bmatrix}. \quad (10.3)$$

Τότε ο  $D$  είναι ένας ισορροπημένος δύο επιπέδων split-plot σχεδιασμός ο οποίος επι-

τρέπει την εκτίμηση του RPD μοντέλου πρώτης σημαντικότητας.

### Απόδειξη Θεωρήματος 1.

Έστω  $X_W$  είναι ο πίνακας σχεδιασμού οποιουδήποτε  $n_1 \times m_1$  δύο επιπέδων ισορροπημένου ΥΣ, είναι ένας  $OA(n_1, m_1, 2, t_1)$ ,  $t_1 = 1$  και  $X_{S_i}, i = 1, \dots, n_1$  είναι ο πίνακας σχεδιασμού για κάθε  $n_2 \times m_2$ ,  $OA(n_2, m_2, 2, t_2)$ ,  $t_2 \geq 1$  έτσι ώστε όλα μαζί να αποτελούν έναν  $OA$  ισχύος μεγαλύτερης ή ίσης με 2. Επιπλέον,  $1_{n_2}$  έστω ένα  $n_2 \times 1$  διάνυσμα στήλη με 1's και  $\otimes$  το δεξιό γινόμενο Kronecker. Ο whole-plot πίνακας σχεδιασμού για τον ΥΣ στο περιβάλλον των split-plot σχεδιασμών είναι

$$X_{WP} = X_W \otimes 1_{n_2} \quad (10.4)$$

όπου ο  $X_{WP}$  είναι ένας  $n_1 n_2 \times m_1$  πίνακας. Επιπρόσθετα, ο split-plot πίνακας σχεδιασμού στο περιβάλλον των split-plot σχεδιασμών είναι

$$X_{SP} = 1_{n_1} \otimes X_{S_i}, \quad (10.5)$$

όπου  $X_{SP}$  είναι ένας  $n_1 n_2 \times m_2$  πίνακας. Μπορεί να αποδειχθεί ότι για  $n_1 OA(n_2, m_2, 2, t_2)$ ,  $t_2 \geq 2$  στα split-plots τέτοια ώστε το  $D_2$  να αποτελεί ένα  $OA(n_1 n_2, m_2, 2, t)$  ισχύος  $t \geq 2$ , καταλήγουν σε ορθογώνιες στήλες όπου

$$X'_{SP} X_{SP} = n_1 n_2 \mathbf{I}_{m_2}. \quad (10.6)$$

Αυτό υποδεικνύει ότι οι split-plot παράγοντες είναι ορθογώνιοι μεταξύ τους, κάτι το οποίο οδηγεί σε ορθογωνιότητα των παραγόντων ελέγχου στο υπό εξέταση RPD μοντέλο. Αυτό είναι εξαιρετικά σημαντικό για την ακριβή εκτίμηση των παραγόντων ελέγχου.

Οποιοσδήποτε δύο στήλες του πίνακα σχεδιασμού  $D$  (9.5), η μία να αντιστοιχεί σε έναν whole-plot παράγοντα και η άλλη να αντιστοιχεί σε έναν split-plot παράγοντα, σχηματίζουν έναν  $OA(n_1 n_2, 2, 2, 2)$ ,  $n_1 n_2 = n$ . Έτσι για κάθε στήλη  $c_{j_1 k}$ ,  $1 \leq j_1 \leq m_1$ ,  $1 \leq k \leq n$  που ανήκει σε ένα whole-plot πίνακα σχεδιασμού και για οποιαδήποτε στήλη  $c_{j_2 k}$ ,  $1 \leq j_2 \leq m_2$ ,  $1 \leq k \leq n$  που ανήκει σε ένα split-plot πίνακα σχεδιασμού, αυτό αυτόματα υποδηλώνει ότι  $c_{j_1 k} c_{j_2 k} = 0$ . Έτσι, προκύπτει

ότι οι whole-plot παράγοντες είναι ορθογώνιοι στους split-plot παράγοντες καθώς

$$\begin{pmatrix} X'_{WP} \\ X'_{SP} \end{pmatrix} \begin{pmatrix} X_{WP} & X_{SP} \end{pmatrix} = \begin{pmatrix} X'_{WP}X_{WP} & \mathbf{0} \\ \mathbf{0} & n_1n_2\mathbf{I}_{m_2} \end{pmatrix}. \quad (10.7)$$

Αυτό υποδεικνύει ότι οι split-plot παράγοντες είναι ορθογώνιοι στους whole-plot παράγοντες, δηλαδή οι παράγοντες ελέγχου είναι ορθογώνιοι με τους παράγοντες θορύβου. Αυτό έχει ως αποτέλεσμα την ορθογωνικότητα μεταξύ των whole-plot παραγόντων και των αλληλεπιδράσεων των whole-plot με τους split-plot (το  $C \times N$  στο πλαίσιο RPD). **Τέλος απόδειξης**

**Πίνακας 10.1:** Εύρωστος Παραμετρικός Split-plot σχεδιασμός.

Whole Plot εκτέλεση	Whole-plots					Split-plots	
$W_1$	$W_2$	...	...	...	$W_{n_1}$		
1	$d_{11}$	$d_{12}$	...	...	...	$d_{1m_1}$	$X_{S_1}$
2	$d_{21}$	$d_{22}$	...	...	...	$d_{2m_1}$	$X_{S_2}$
⋮	⋮	⋮	⋱			⋮	⋮
⋮	⋮	⋮		⋱		⋮	⋮
⋮	⋮	⋮			⋱	⋮	⋮
$n_1$	$d_{n_11}$	$d_{n_12}$	...	...	...	$d_{n_1m_1}$	$X_{S_{n_1}}$

Στην εξίσωση (10.7), υποτίθεται ότι οι αλληλεπιδράσεις δύο παραγόντων μεταξύ των whole-plot και των split-plot παραγόντων είναι αμελητέες, επομένως αποκλείστηκαν από τον πίνακα σχεδιασμού. Αυτή η κατασκευή επιτρέπει τη χρήση ισορροπημένων ΥΣ τόσο στους whole-plot όσο και στους split-plot παράγοντες. Ανάλογα με την κατασκευή αυτών των σχεδιασμών μπορούμε εύκολα να καταλήξουμε ως προς τους περιορισμούς στις πειραματικές μονάδες. Με αυτόν τον τρόπο, δεν είναι γενικά δυνατή η τοποθέτηση οποιασδήποτε whole-plot επί whole-plot ή split-plot επί split-plot αλληλεπίδρασης. Εντούτοις, μπορούμε να εξάγουμε κάποια χρήσιμα συμπεράσματα σχετικά με κάποιες από τις whole-plot επί split-plot αλληλεπιδράσεις (τις  $C \times N$  στο RPD πλαίσιο) οι οποίες είναι και οι πιο σημαντικές αλληλεπιδράσεις στην RPD μεθοδολογία (Πίνακας 8.1).

Θεωρούμε ότι  $X_{WP_i}$  είναι η  $i$ -οστή στήλη του  $X_{WP}$ ,  $i = 1, \dots, m_1$ ,  $X_{SP_j}$  είναι η

$j$ -οστή στήλη του  $X_{SP}$ ,  $j = 1, \dots, m_2$  και θέτουμε

$$X_{WPSP_i} = [X_{WP_i} * X_{SP_1} \quad X_{WP_i} * X_{SP_2} \quad \dots \quad X_{WP_i} * X_{SP_{m_2}}], \quad (10.8)$$

όπου  $*$  είναι το γινόμενο ανά στοιχείο. Τότε ο πίνακας  $X_{WPSP}$  για τις whole-plot επί split-plot αλληλεπιδράσεις μπορεί να γραφεί ως

$$X_{WPSP} = [X_{WPSP_1} \quad X_{WPSP_2} \quad \dots \quad X_{WPSP_{m_1}}], \quad (10.9)$$

όπου ο  $X_{WPSP}$  είναι  $n_1 n_2 \times m_1 m_2$  διαστάσεων. Ο Πίνακας 10.1 δίνει μία σχηματική εξήγηση του split-plot σχεδιασμού.

### 10.3.2 Δομή ταυτόσημων επιδράσεων

Λόγω του γεγονότος ότι οι ΥΣ είναι ένας ειδικός τύπος κλασματικών παραγοντικών σχεδιασμών, μερικές από τις επιδράσεις είναι ταυτόσημες με άλλες επιδράσεις. Αυτό σημαίνει ότι δεν μπορούμε να εκτιμήσουμε όλες τις επιδράσεις ανεξάρτητα από άλλες επιδράσεις (Wu και Hamada, [143]). Για να εξετάσουμε τη δομή των ταυτόσημων σχέσεων μεταξύ των επιδράσεων θα πρέπει να δημιουργήσουμε τον αντίστοιχο πίνακα των ταυτόσημων επιδράσεων (alias matrix). Ας εξετάσουμε το προσαρμοσμένο μοντέλο

$$y = X_1 \beta_1 + \epsilon_1, \quad (10.10)$$

αλλά το πραγματικό μοντέλο είναι

$$y = X_1 \beta_1 + X_2 \beta_2 + \epsilon. \quad (10.11)$$

Μπορούμε εύκολα να καταλήξουμε ότι ο πίνακας των ταυτόσημων επιδράσεων μεταξύ των δύο αυτών μοντέλων ισούται με  $X_{alias} = (X_1' X_1)^{-1} (X_1')' X_2$ .

Για να εκτιμήσουμε τις ταυτόσημες σχέσεις στο SSSPDs θεωρούμε ότι

$$X_1 = [X_{WP} \quad X_{SP}]$$

και θέτουμε  $X_2$  να είναι ο πίνακας σχεδιασμού των whole-plot επί split-plot αλλη-

λεπιδράσεων

$$X_2 = X_{WPSP} = [X_{WPSP_1} \quad X_{WPSP_2} \quad \dots \quad X_{WPSP_{m_1}}]$$

Επομένως, ο πίνακας των ταυτώσεων επιδράσεων μεταξύ των κύριων επιδράσεων (split-plot  $C$  και whole-plot  $N$ ) και των αλληλεπιδράσεων δύο παραγόντων μεταξύ των split-plot και των whole-plot ( $C \times N$ ) παραγόντων δίνεται ως

$$\begin{aligned} X_{alias} &= (X_1' X_1)^{-1} (X_1')' X_2 \\ &= \begin{bmatrix} X'_{WP} X_{WP} & X'_{WP} X_{SP} \\ X'_{SP} X_{WP} & X'_{SP} X_{SP} \end{bmatrix}^{-1} \times \begin{bmatrix} X'_{WP} X_{WPSP_1} & \dots & X'_{WP} X_{WPSP_{m_1}} \\ X'_{SP} X_{WPSP_1} & \dots & X'_{SP} X_{WPSP_{m_1}} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \dots & 0 \\ \frac{1}{n_1 n_2} X'_{SP} X_{WPSP_1} & \dots & \frac{1}{n_1 n_2} X'_{SP} X_{WPSP_{m_1}} \end{bmatrix}. \end{aligned}$$

Ο πίνακας  $X_{alias}$  δείχνει ότι οι split-plot παράγοντες είναι ορθογώνιοι μεταξύ τους ( $X'_{SP} X_{SP} = \frac{1}{n_1 n_2} I_{m_2}$ ), οι split-plot παράγοντες είναι ορθογώνιοι στους whole-plot παράγοντες ( $X'_{WP} X_{SP} = 0$ ) και οι whole-plot παράγοντες είναι ορθογώνιοι στις whole-plot επί split-plot αλληλεπιδράσεις ( $X'_{WP} X_{WPSP_i} = 0$ ). Είναι προφανές ότι οι whole-plot παράγοντες είναι συσχετισμένοι μεταξύ τους αφού βρίσκονται σ' έναν ΥΣ καθώς επίσης και οι αλληλεπιδράσεις μεταξύ των whole-plots και των split-plots συσχετίζονται με τους split-plot παράγοντες.

**Παρατήρηση 1** Οι ΥΣ,  $E(s^2)$ -βέλτιστοι ή  $D$ -βέλτιστοι, κατασκευάζονται χρησιμοποιώντας είτε αναζήτηση στον υπολογιστή είτε διάφορες άλλες μεθόδους. Έτσι, δεν υπάρχει ένα σχέδιο πάνω στο πως οι whole-plot επί split-plot αλληλεπιδράσεις είναι μερικώς ταυτώσιμες με τους split-plot παράγοντες. Πρέπει να σημειώσουμε ότι οι ταυτώσιμες δομές πρέπει να εξεταστούν πριν από τη χρήση των προαναφερθέντων ΥΣ σε πρακτικά προβλήματα.

### 10.3.3 Γενικές παρατηρήσεις

- Σε γενικές γραμμές, οι προτεινόμενοι σχεδιασμοί αποτελούν έναν SSSPD τουλάχιστον σε σχέση με τους whole-plot παράγοντες. Μια κλάση SSSPD σχεδιασμών έχει παρουσιαστεί από τους Koh et al. [73]. Η κατασκευή τους είναι μια ειδική περίπτωση κατασκευής του Θεωρήματος 1 που προτείνουμε στο παρόν κεφάλαιο. Πιο συγκεκριμένα, οι Koh et al. [73] χρησιμοποίησαν έναν  $E(s^2)$ -βέλτιστο ΥΣ στους whole-plot παράγοντες και μισά κλάσματα Plackett-Burman σχεδιασμών στους split-plot παράγοντες. Η κατασκευή μας επιτρέπει τη χρήση ισορροπημένων ΥΣ τόσο στους whole-plot όσο και στους split-plot παράγοντες. Αυτό έχει ως αποτέλεσμα σχεδιασμούς που μπορούν να φιλοξενήσουν πολυάριθμους παράγοντες με περιορισμένες πειραματικές εκτελέσεις κάτι που είναι κατάλληλο και εξαιρετικά χρήσιμο για πειράματα κρησαρίσματος.
- Στο θεώρημα κατασκευής που προτείνουμε χρησιμοποιήσαμε COA με κατάλληλη επιλογή παραμέτρων χωρίς τη χρήση οποιουδήποτε συγκεκριμένου τύπου βέλτιστου σχεδιασμού είτε whole-plot είτε split-plot σχεδιασμούς. Αυτοί οι σχεδιασμοί μπορούν να εφαρμοστούν αποτελεσματικά για RPD πειράματα και είναι σε θέση να φιλοξενήσουν πολλούς παράγοντες με περιορισμένες πειραματικές εκτελέσεις. Παρουσιάζουμε επιπλέον κάποιες χρήσιμες ειδικές περιπτώσεις που επιτρέπουν στους ερευνητές να επιλέξουν έναν σχεδιασμό ανάλογα με το μοντέλο που τους ενδιαφέρει.
- Τα RPD πειράματα επιτρέπουν τη χαλάρωση μεταξύ των ταυτόσημων επιδράσεων μεταξύ των παραγόντων θορύβου. Έτσι, οι σχεδιασμοί που παρουσιάζουμε σε αυτή τη διδακτορική διατριβή όπου προσαρμόζουμε έναν ΥΣ σε όλους τους whole-plot παράγοντες λειτουργούν πολύ καλά σε RPD πειράματα.
- Οι σχεδιασμοί που παρουσιάζονται εξυπηρετούν το μοντέλο πρώτης σημαντικότητας (Πίνακας 8.1) καθώς και μερικούς από τους όρους αλληλεπίδρασης  $C \times N$ . Με αυτόν τον τρόπο είμαστε σε θέση να εξετάσουμε πολλούς παράγοντες του μοντέλου πρώτης σημαντικότητας στο πρώτο στάδιο του πειράματος ώστε να εντοπίσουμε τους πιο σημαντικούς. Οι σημαντικοί παράγοντες που προέκυψαν μπορούν να χρησιμοποιηθούν στα επακόλουθα πειράματα στα ο-



2 στα split-plot οδηγούμεστε σε ορθογώνιες στήλες όπου

$$X'_{SP}X_{SP} = n_1n_2\mathbf{I}_{m_2}. \quad (10.13)$$

Αυτό υποδεικνύει ότι οι split-plot παράγοντες είναι ορθογώνιοι ο ένας στον άλλο. Επιπλέον, από το Θεώρημα 1, όλοι οι whole-plot παράγοντες είναι ορθογώνιοι στους split-plot παράγοντες ως

$$\begin{pmatrix} X'_{WP} \\ X'_{SP} \end{pmatrix} \begin{pmatrix} X_{WP} & X_{SP} \end{pmatrix} = \begin{pmatrix} X'_{WP}X_{WP} & \mathbf{0} \\ \mathbf{0} & n_1n_2\mathbf{I}_{m_2} \end{pmatrix}. \quad (10.14)$$

Αυτό υποδεικνύει ότι οι split-plot παράγοντες είναι ορθογώνιοι στους whole-plot παράγοντες, δηλαδή οι παράγοντες ελέγχου είναι ορθογώνιοι στους παράγοντες θορύβου. Επιπλέον, ο πίνακας των ταυτόσημων επιδράσεων μεταξύ των κύριων επιδράσεων (split-plot  $C$  και whole-plot  $N$ ) και των αλληλεπιδράσεων δύο παραγόντων μεταξύ των split-plot και των whole-plot ( $C \times N$ ) παραγόντων δίνεται ως

$$\begin{aligned} X_{alias} &= (X_1'X_1)^{-1}(X_1)'X_2 \\ &= \begin{bmatrix} X'_{WP}X_{WP} & X'_{WP}X_{SP} \\ X'_{SP}X_{WP} & X'_{SP}X_{SP} \end{bmatrix}^{-1} \times \begin{bmatrix} X'_{WP}X_{WPSP_1} & \dots & X'_{WP}X_{WPSP_{m_1}} \\ X'_{SP}X_{WPSP_1} & \dots & X'_{SP}X_{WPSP_{m_1}} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \end{bmatrix}. \end{aligned} \quad (10.15)$$

Ο πίνακας  $X_{alias}$  καταδεικνύει ότι οι split-plot παράγοντες είναι ορθογώνιοι ο ένας στον άλλο ( $X'_{SP}X_{SP} = \frac{1}{n_1n_2}I_{m_2}$ ), οι split-plot παράγοντες είναι ορθογώνιοι στους whole-plot παράγοντες ( $X'_{WP}X_{SP} = 0$ ) και στις αλληλεπιδράσεις μεταξύ των whole-plot και των split-plot παραγόντων ( $X'_{SP}X_{WPSP_i} = 0$ ) και οι whole-plot παράγοντες είναι ορθογώνιοι στις αλληλεπιδράσεις μεταξύ των whole-plot και των split-plot παραγόντων ( $X'_{WP}X_{WPSP_i} = 0$ ). Αυτό έχει ως αποτέλεσμα την καλή προσαρμογή του μοντέλου (10.1). **Τέλος απόδειξη του Λήμματος 13.**





Επιπλέον, όλοι οι whole-plot παράγοντες είναι ορθογώνιοι στους split-plot παράγοντες λόγω του Θεωρήματος 1. Επιπροσθέτως, ο πίνακας των ταυτόσημων επιδράσεων μεταξύ των κύριων επιδράσεων (split-plot  $C$  και whole-plot  $N$ ) και των αλληλεπιδράσεων δύο παραγόντων μεταξύ των split-plot και whole-plot ( $C \times N$ ) παραγόντων δίδονται ως

$$X_{alias} = (X_1' X_1)^{-1} (X_1)' X_2 = \begin{bmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \end{bmatrix},$$

όπως ακριβώς στην εξίσωση 10.15.

Θεωρούμε ότι  $X_{SPSP_i} = X_{SP_k} * X_{SP_l}$  είναι το  $i$ -οστό γινόμενο ανά στοιχείο (elementwise product) μεταξύ της  $k$ -οστής και της  $l$ -οστής στήλης του  $X_{SP}$ ,  $\forall k \neq l$ , για  $i = 1, \dots, \binom{m_2}{2}$ , τότε

$$X_{SPSP} = \begin{bmatrix} X_{SPSP_1} & X_{SPSP_2} & \dots & X_{SPSP_{\binom{m_2}{2}}} \end{bmatrix}, \quad (10.18)$$

όπου  $X_{SPSP}$  είναι  $n_1 n_2 \times \binom{m_2}{2}$  διαστάσεων.

Ας θεωρήσουμε το προσαρμοσμένο μοντέλο

$$y = X_1 \beta_1 + \epsilon_1, \quad (10.19)$$

αλλά το πραγματικό μοντέλο είναι

$$y = X_1 \beta_1 + X_3 \beta_3 + \epsilon. \quad (10.20)$$

Μπορούμε εύκολα να καταλήξουμε στο συμπέρασμα ότι ο πίνακας των ταυτόσημων επιδράσεων μεταξύ των δύο μοντέλων είναι ίσος με  $X_{alias} = (X_1' X_1)^{-1} (X_1)' X_3$ . Προκειμένου να εκτιμηθούν οι ταυτόσημες σχέσεις στον SSSPD, θεωρούμε ότι

$$X_1 = [X_{WP} \quad X_{SP}]$$

και έστω  $X_3$  ένας πίνακας σχεδιασμού των αλληλεπιδράσεων split-plot επί split-plot

$$X_3 = X_{SPSP} = \begin{bmatrix} X_{SPSP_1} & X_{SPSP_2} & \dots & X_{SPSP_{\binom{m_2}{2}}} \end{bmatrix}.$$

Επομένως, ο πίνακας των ταυτόσημων επιδράσεων μεταξύ των κύριων επιδράσεων (split-plot  $C$  και whole-plot  $N$ ) και των αλληλεπιδράσεων δύο παραγόντων μεταξύ των split-plot ( $C \times C$ ) παραγόντων δίνεται ως

$$\begin{aligned} X_{alias} &= (X_1' X_1)^{-1} (X_1)' X_3 \\ &= \begin{bmatrix} X'_{WP} X_{WP} & X'_{WP} X_{SP} \\ X'_{SP} X_{WP} & X'_{SP} X_{SP} \end{bmatrix}^{-1} \times \begin{bmatrix} X'_{WP} X_{SPSP_1} & \dots & X'_{WP} X_{SPSP_{\binom{m_2}{2}}} \\ X'_{SP} X_{WPSP_1} & \dots & X'_{SP} X_{WPSP_{\binom{m_2}{2}}} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \end{bmatrix}. \end{aligned}$$

Ο πίνακας  $X_{alias}$  όχι μόνο καταδεικνύει ότι οι split-plot παράγοντες είναι ορθογώνιοι ο ένας στον άλλο ( $X'_{SP} X_{SP} = \frac{1}{n_1 n_2} I_{m_2}$ ) και στους whole-plot παράγοντες ( $X'_{WP} X_{SP} = 0$ ) αλλά επίσης δείχνει ότι οι split-plot παράγοντες είναι ορθογώνιοι στις αλληλεπιδράσεις μεταξύ των split-plot επί split-plot παραγόντων ( $X'_{SP} X_{SPSP_i} = 0$ ) και ότι οι whole-plot παράγοντες είναι ορθογώνιοι στις αλληλεπιδράσεις μεταξύ των split-plot επί split-plot παραγόντων ( $X'_{WP} X_{SPSP_i} = 0$ ). Αυτό έχει ως αποτέλεσμα την καλή προσαρμογή του μοντέλου (10.2). **Τέλος απόδειξης του Λήμματος 14.**

### 10.3.6 Αλγόριθμος δημιουργίας σχεδιασμών

Σε αυτή την ενότητα παρουσιάζουμε έναν αλγόριθμο για την εφαρμογή του προαναφερθέντος θεωρήματος. Αυτό είναι πολύ σημαντικό για τους ενδιαφερόμενους, ώστε να δημιουργήσουν εύκολα τους προτεινόμενους σχεδιασμούς. Ένα σκίτσο του αλγορίθμου μας αποτελείται από τα ακόλουθα βήματα:

1. Θεωρούμε έναν σχεδιασμό  $D_1 \in D(n_1, 2^{m_1}, t_1)$  με πίνακα σχεδιασμού  $X_W$ .

Θεωρούμε έναν ισορροπημένο ΥΣ έτσι ώστε να μπορούμε να φιλοξενήσουμε

πολλούς whole-plot παράγοντες.

2. Για  $1 \leq i \leq n_1$ , θεωρούμε ένα  $b_i \in D(n_2, 2^{m_2}, t_2)$ , με  $t_2 \geq 1$ . Επιλέγουμε είτε έναν πλήρη είτε έναν κλασματικό παραγοντικό σχεδιασμό ισχύος  $t_2 \geq 1$  (π.χ. έναν ισορροπημένο ΥΣ) για κάθε ένα από τα  $b_i$ .
3. Ορίζουμε τον πίνακα

$$D_2 = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n_1} \end{bmatrix},$$

έτσι ώστε ο  $D_2$  να αποτελεί έναν ΟΑ ισχύος μεγαλύτερης ή ίσης με 1.

4. Ορίζουμε  $D \in D^{(COA)}(n_1 n_2, 2^{m_1+m_2})$  ως

$$D = \begin{bmatrix} d_1^T \\ \vdots \\ d_1^T \\ \vdots \\ d_{n_1}^T \\ \vdots \\ d_{n_1}^T \end{bmatrix} \begin{matrix} \\ \\ b_1 \\ \\ \\ b_{n_1} \end{matrix}. \quad (10.21)$$

5. Ο κατασκευασμένος σχεδιασμός  $D$  είναι ένας ισορροπημένος split-plot σχεδιασμός δύο επιπέδων που επιτρέπει την εκτίμηση του RPD μοντέλου πρώτης σημαντικότητας.

**Παρατήρηση 1:** Στο βήμα 2, επιλέγουμε ένα  $b_k \in D(n_2, 2^{m_2}, t_2)$ , με  $t_2 \geq 1$ . Μπορούμε να χρησιμοποιήσουμε την ίδια επιλογή για κάθε ένα από τα  $b_k$ . Τότε οδηγούμαστε σε διασταυρωμένους σχηματισμούς που έχουν χρησιμοποιηθεί μέχρι τώρα στη μεθοδολογία του Taguchi. Έστω  $X_W$  ένας ισορροπημένος ΥΣ. Αυτός ο

σχεδιασμός μπορεί εύκολα να δημιουργηθεί ως

$$D = \begin{bmatrix} d_1^T \\ \vdots \\ d_1^T \\ \vdots \\ d_{n_1}^T \\ \vdots \\ d_{n_1}^T \end{bmatrix} \cdot b_k \quad (10.22)$$

Τότε ο  $D$  είναι ένας ισορροπημένος δύο επιπέδων split-plot σχεδιασμός ο οποίος επιτρέπει την εκτίμηση του RPD μοντέλου 10.1.

**Παρατήρηση 2:** Στο βήμα 2, επιλέγουμε ένα  $b_k \in D(n_2, 2^{m_2}, t_2)$ , με  $t_2 \geq 1$  και  $\bar{b}_k$  ο foldover σχεδιασμός του  $b_k$ . Έστω  $X_W$  ένας ισορροπημένος ΥΣ. Αυτός ο σχεδιασμός μπορεί εύκολα να δημιουργηθεί ως

$$D = \begin{bmatrix} d_1^T \\ \vdots \\ d_1^T \\ \vdots \\ d_2^T \\ \vdots \\ d_2^T \\ \vdots \\ d_{n_1-1}^T \\ \vdots \\ d_{n_1-1}^T \\ \vdots \\ d_{n_1}^T \\ \vdots \\ d_{n_1}^T \end{bmatrix} \cdot \begin{matrix} b_k \\ \\ \bar{b}_k \\ \\ b_k \\ \\ \bar{b}_k \end{matrix} \quad (10.23)$$

Τότε, ο  $D$  είναι ένας ισορροπημένος δύο επιπέδων split-plot σχεδιασμός ο οποίος επιτρέπει την εκτίμηση του RPD μοντέλου 10.2.

**Παράδειγμα κατασκευής, εφαρμόζοντας τον αλγόριθμο 10.3.6**

Ας θεωρήσουμε τον σχεδιασμό  $D_1 \in D(6, 2^{10}, 1)$  για τους whole-plot παράγοντες που δίνεται ως

**Πίνακας 10.2:** Σχεδιασμός  $D_1$  ο οποίος αποτελείται από  $m_1 = 10$  whole-plot παράγοντες και  $n_1 = 6$  πειραματικές εκτελέσεις.

		Whole-plots										
		$W_1$	$W_2$	$W_3$	$W_4$	$W_5$	$W_6$	$W_7$	$W_8$	$W_9$	$W_{10}$	
$D_1 =$		1	1	1	1	1	1	1	1	1	1	$d_1^T$
		1	1	1	1	-1	-1	-1	-1	-1	-1	$d_2^T$
		1	-1	-1	-1	1	1	1	-1	-1	-1	$d_3^T$
		-1	1	-1	-1	1	-1	-1	1	1	-1	$d_4^T$
		-1	-1	1	-1	-1	1	-1	1	-1	1	$d_5^T$
		-1	-1	-1	1	-1	-1	1	-1	1	1	$d_6^T$

Θεωρούμε τους σχεδιασμούς  $b_k \in D(4, 2^{11}, 1)$  ο καθένας αντιστοιχεί στους sub-plot παράγοντες ως ακολούθως

$$b_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 \end{pmatrix}$$

$$b_2 = \begin{pmatrix} 1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \end{pmatrix}$$

$$b_3 = \begin{pmatrix} 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 \end{pmatrix}$$

$$b_4 = \begin{pmatrix} 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \end{pmatrix}$$

$$b_5 = \begin{pmatrix} -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix}$$

$$b_6 = \begin{pmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 \\ -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 \end{pmatrix}$$

**Πίνακας 10.3:** Σχεδιασμός  $D$  ο οποίος αποτελείται από  $m_1 + m_2 = 21$  παράγοντες και  $n_1 n_2 = 24$  πειραματικές εκτελέσεις.

		Whole-plots								Split-plots		
		$W_1$	$W_2$	$W_3$	$W_4$	$W_5$	$W_6$	$W_7$	$W_8$	$W_9$	$W_{10}$	
$D =$		1	1	1	1	1	1	1	1	1	1	$b_1$
		1	1	1	1	-1	-1	-1	-1	-1	-1	$b_2$
		1	-1	-1	-1	1	1	1	-1	-1	-1	$b_3$
		-1	1	-1	-1	1	-1	-1	1	1	-1	$b_4$
		-1	-1	1	-1	-1	1	-1	1	-1	1	$b_5$
		-1	-1	-1	1	-1	-1	1	-1	1	1	$b_6$

Ακολουθώντας αυτή τη διαδικασία, ο πίνακας σχεδιασμού  $D$  είναι ο ακόλουθος

$w_1$	$w_2$	$w_3$	$w_4$	$w_5$	$w_6$	$w_7$	$w_8$	$w_9$	$w_{10}$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$	$s_7$	$s_8$	$s_9$	$s_{10}$	$s_{11}$
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	-1	-1	1	-1	-1	-1	1	1	1	-1	1
1	1	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
1	1	1	1	1	1	1	1	1	1	1	1	-1	1	1	1	-1	-1	-1	1	-1
1	1	1	1	-1	-1	-1	-1	-1	-1	1	-1	-1	1	-1	-1	1	1	1	1	-1
1	1	1	1	-1	-1	-1	-1	-1	-1	-1	1	-1	-1	1	-1	-1	-1	1	1	1
1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1
1	-1	-1	-1	1	1	1	-1	-1	-1	1	-1	1	-1	-1	1	-1	-1	-1	1	1
1	-1	-1	-1	1	1	1	-1	-1	-1	1	1	-1	1	-1	1	-1	-1	-1	-1	1
1	-1	-1	-1	1	1	1	-1	-1	-1	-1	1	-1	1	1	-1	1	1	1	-1	-1
1	-1	-1	-1	1	1	1	-1	-1	-1	-1	-1	1	-1	1	1	-1	1	1	1	-1
-1	1	-1	-1	1	-1	-1	1	1	-1	1	1	1	-1	1	-1	-1	1	-1	-1	-1
-1	1	-1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	1	-1	-1	1	1	-1	-1	-1	-1	1	-1	1	1	-1	1	1	1
-1	1	-1	-1	1	-1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	1
-1	-1	1	-1	-1	1	-1	1	-1	1	-1	-1	1	1	-1	1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	1	-1	1	-1	1	-1	-1	-1	1	1	-1	1	-1	-1	-1	1
-1	-1	1	-1	-1	1	-1	1	-1	1	1	1	-1	-1	-1	1	-1	1	1	-1	1
-1	-1	1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	-1	1	-1	1	1	-1
-1	-1	-1	1	-1	-1	1	-1	1	1	1	-1	-1	-1	1	1	-1	1	-1	1	-1
-1	-1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	-1	1	1	-1	1	1	-1
-1	-1	-1	1	-1	-1	1	-1	1	1	-1	1	1	1	-1	-1	-1	1	-1	1	1
-1	-1	-1	1	-1	-1	1	-1	1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1

Ο σχεδιασμός  $D \in D^{(COA)}(24, 2^{21}, 1)$  είναι ένας ισορροπημένος δύο επιπέδων split-plot σχεδιασμός ο οποίος επιτρέπει την εκτίμηση του RPD μοντέλου πρώτης σημαντικότητας.

## 10.4 Ανάλυση των προτεινόμενων σχεδιασμών

Η δυσκολία της ανάλυσης των SSSPD έγκειται στο γεγονός ότι υπάρχουν πολλοί παράγοντες που εξετάζονται με εξαιρετικά περιορισμένους πόρους. Αυτό οδηγεί σε έναν πίνακα ΥΣ που σημαίνει ότι δεν μπορούν να εκτιμηθούν όλες οι παράμετροι. Η ακριβής ανάλυση αυτών των πειραμάτων εξαρτάται τόσο από τη σποραδικότητα των επιδράσεων όσο και από την ιεραρχία τους (περισσότερες πληροφορίες θα βρείτε στο κεφάλαιο 4) [104]. Με βάση αυτές τις βασικές αρχές παρουσιάζουμε μια μεθοδολογία για την ανάλυση των προτεινόμενων σχεδιασμών. Πραγματοποιούμε την ανάλυση σε τρία στάδια: το πρώτο συνιστά την ανάλυση των whole-plot, το δεύτερο



αποτελεί την ανάλυση των split-plot και των αλληλεπιδράσεων μεταξύ των whole-plot και των split-plot, και το τρίτο αφορά στην ενσωμάτωση της ανάλυσης στο RPD πλαίσιο. Στην ανάλυση λαμβάνουμε υπόψη ότι οι παράγοντες ελέγχου αντιστοιχούνται στα split-plot και οι παράγοντες θορύβου στα whole-plot.

### 10.4.1 Στάδιο I: Whole-plot (N) ανάλυση

Η ανάλυση των whole-plot παρατηρήσεων γίνεται με διαδικασίες κατά βήματα επιλογής [91]. Οι παρατηρήσεις για τους whole-plots παράγοντες υπολογίζονται λαμβάνοντας με τη μέση τιμή των αποκρίσεων πάνω στις split-plot μονάδες μέσα σε κάθε whole-plot μονάδα. Λόγω του γεγονότος ότι προσαρμόσαμε έναν ΥΣ στον whole-plot πίνακα  $X_W$  υπάρχει ένας μικρός αριθμός whole-plot μονάδων και κατά συνέπεια, ελάχιστοι βαθμοί ελευθερίας για το whole-plot σφάλμα, κάτι που περιορίζει τη δύναμη της ανάλυσης των whole-plot. Σε τέτοιες περιπτώσεις, θα μπορούσαν να χρησιμοποιηθούν μεγαλύτερες τιμές του  $\alpha$  (π.χ.,  $\alpha = 0.2$ ) έτσι ώστε να αυξηθούν οι πιθανότητες ανίχνευσης σημαντικών whole-plot επιδράσεων. Ωστόσο, η επίπτωση του κρησαρίσματος πολλών παραγόντων με λίγους βαθμούς ελευθερίας για τις whole-plot μονάδες είναι ότι ακόμη και για υψηλές τιμές των τιμών του  $\alpha$ , ακόμη και σημαντικές whole-plot επιδράσεις εξακολουθούν να είναι μη σημαντικές. Στις περιπτώσεις αυτές, μπορεί να είναι αναπόφευκτο να χρησιμοποιηθεί η κρίση εμπειρογνομόνων για τον προσδιορισμό των σημαντικών παραγόντων. Χρησιμοποιήσαμε τελικά  $\alpha = 0.2$  καθώς επίσης επιλέξαμε τους ενεργούς παράγοντες θορύβου σε 5% επίπεδο σημαντικότητας. Για να συνοψίσουμε, τα βήματα που αποτελούν το Στάδιο I είναι τα ακόλουθα:

1. Υπολογίζουμε το μέσο όρο των αποκρίσεων πάνω στις split-plot μονάδες μέσα σε κάθε whole-plot μονάδα έτσι ώστε να λάβουμε τις παρατηρήσεις για τα whole-plots .
2. Αναλύουμε τα whole-plots χρησιμοποιώντας κατά βήματα επιλογή με  $\alpha = 0.2$ .
3. Επιλέγουμε τις σημαντικές whole-plot επιδράσεις σε 5% επίπεδο σημαντικότητας.

### 10.4.2 Στάδιο II : Ανάλυση split-plot (C) και whole-plot επί split-plot (C × N) αλληλεπιδράσεων.

Μόλις οι ενεργοί whole-plot παράγοντες έχουν ανιχνευτεί, θεωρούμε τις αλληλεπιδράσεις μεταξύ των whole-plot και των split-plot παραγόντων. Άλλες προτεινόμενες μέθοδοι εξετάζουν μια αυστηρή προσέγγιση κληρονομικότητας των επιδράσεων, σύμφωνα με την οποία αλληλεπιδράσεις θεωρούνται μόνο εάν και οι δύο whole-plot και split-plot παράγοντες ανιχνεύονται από τα προηγούμενα βήματα. Αντιθέτως, εμείς θεωρήσαμε όλους τους split-plot παράγοντες καθώς επίσης και τις αλληλεπιδράσεις μεταξύ των ενεργών whole-plot και όλων των split-plot παραγόντων. Σε περίπτωση που εξετάζουμε το μοντέλο 2, σε αυτό το στάδιο ενσωματώνουμε τις split-plot επί split-plot αλληλεπιδράσεις. Επιπλέον θεωρήσαμε ένα μοντέλο που περιλαμβάνει την whole-plot επίδραση η οποία μετράει για όλη τη διασπορά των whole-plot μονάδων έτσι ώστε να διασφαλιστεί ότι το whole-plot σφάλμα δεν συγχωνεύεται με το split-plot σφάλμα. Πραγματοποιήσαμε την ανάλυση για την εύρεση των σημαντικών επιδράσεων των split-plot παραγόντων και των αλληλεπιδράσεων μεταξύ των whole-plot και των split-plot παραγόντων χρησιμοποιώντας κατά βήματα επιλογή με  $\alpha = 0.2$ . Επιλέξαμε τους ενεργούς παράγοντες σε 1% επίπεδο σημαντικότητας. Για να συνοψίσουμε, τα βήματα που αποτελούν το Στάδιο II είναι τα ακόλουθα :

1. Θεωρούμε το μοντέλο με (α) όλους τους split-plot παράγοντες (β) τις αλληλεπιδράσεις μεταξύ των ενεργών whole-plot (που ανιχνεύτηκαν στο στάδιο I) και όλων των split-plot παραγόντων (γ) μία whole-plot επίδραση η οποία μετράει για όλη τη διασπορά των whole-plot μονάδων.
2. Αναλύουμε το μοντέλο του βήματος 1 χρησιμοποιώντας κατά βήματα επιλογή με  $\alpha = 0.2$ .
3. Επιλέγουμε τους σημαντικούς split-plot (C, παράγοντες ελέγχου) και τις αλληλεπιδράσεις μεταξύ των whole-plot και των split-plot (C × N) επιδράσεων σε επίπεδο σημαντικότητας 1%.

### 10.4.3 Στάδιο III : Αποκριτική Μοντελοποίηση

Οι στρατηγικές μοντελοποίησης στη μεθοδολογία των εύρωστων παραμετρικών σχεδιασμών είναι, είτε το λεγόμενο μοντέλο διασποράς θέσης που χρησιμοποιείται στους διασταυρωμένους σχηματισμούς είτε η αποκριτική μοντελοποίηση που χρησιμοποιείται στους συνδυασμένους σχηματισμούς. Η στρατηγική μοντελοποίησης που υιοθετήθηκε στο παρόν κεφάλαιο είναι η αποκριτική μοντελοποίηση. Εφαρμόσαμε λοιπόν το μοντέλο απόκρισης σύμφωνα με τα παρακάτω βήματα :

- Το εκτιμώμενο μοντέλο διατυπώνεται με βάση τόσο το στάδιο I όσο και το στάδιο II. Ακριβέστερα, μόλις επιλεγούν οι ενεργοί whole-plot (N) παράγοντες από το στάδιο I και οι ενεργοί split-plot (C) παράγοντες, καθώς επίσης και οι αλληλεπιδράσεις δύο παραγόντων μεταξύ των whole-plot και των split-plot παραγόντων (C×N) ή μεταξύ των split-plot με άλλους split-plot παράγοντες (C×C) από το στάδιο II, δημιουργούμε το τελικό μοντέλο  $\hat{y}$ .
- Σύμφωνα με την RPDM θεωρούμε ότι οι παράγοντες θορύβου είναι τυχαίες μεταβλητές με  $E[\mathbf{z}] = 0$  και  $Var[\mathbf{z}] = \Sigma_{\mathbf{z}}$ . Με βάση αυτές τις υποθέσεις μπορούμε να διαμορφώσουμε το μοντέλο θέσης  $\widehat{E}[y] = 0$  καθώς και το μοντέλο διασποράς  $\widehat{Var}[y]$ , αντίστοιχα.
- Η RPDM στοχεύει στη μείωση της διασποράς της διαδικασίας, ενώ βελτιστοποιεί τη μέση απόκριση. Στόχος είναι να επιλέξουμε σωστά τα επίπεδα των παραγόντων ελέγχου έτσι ώστε η διαδικασία να μην είναι ευαίσθητη στις αλλαγές των παραγόντων θορύβου. Με αυτόν τον τρόπο, επιλέγουμε τα επίπεδα των συντελεστών ελέγχου έτσι ώστε να βελτιστοποιήσουμε το μοντέλο του μέσου  $\widehat{E}[y] = 0$  κρατώντας το μοντέλο της διασποράς  $\widehat{Var}[y]$  όσο το δυνατόν χαμηλότερα.

## 10.5 Παραδείγματα

Σε αυτή την ενότητα παρουσιάζουμε ορισμένα επεξηγηματικά παραδείγματα. Κατασκευάσαμε τους σχεδιασμούς σύμφωνα με το Θεώρημα 1 καθώς και σύμφωνα με τις ειδικές περιπτώσεις που παρουσιάζονται μέσω του Λήμματος 13 και του Λήμματος 14 που προτάθηκαν στις προηγούμενες ενότητες.

**Παράδειγμα 12** Θεωρούμε τον σχεδιασμό  $D_1 \in D(6, 2^{10}, 1)$  για τους whole-plot παράγοντες και τους σχεδιασμούς  $b_k \in D(4, 2^{11}, 1)$  που αντιστοιχούν στους split-plot παράγοντες που δίνονται στο Παράδειγμα 10.3.6. Ο προκύπτων σχεδιασμός  $D \in D^{(COA)}(24, 2^{21}, 1)$  που κατασκευάστηκε στο παράδειγμα του κεφαλαίου 10.3.6 (Πίνακας 10.3) είναι ένας ισορροπημένος split-plot σχεδιασμός δύο επιπέδων που επιτρέπει την εκτίμηση του RPD μοντέλου πρώτης σημαντικότητας.

Υποθέτουμε ότι το πείραμα περιλαμβάνει μία απόκριση,  $m_1 = 10$  παράγοντες θορύβου ( $N$ ) και  $m_2 = 11$  παράγοντες ελέγχου ( $C$ ) όλοι σε δύο επίπεδα, με τους παράγοντες ελέγχου να λειτουργούν ως split-plot παράγοντες, και με τους παράγοντες θορύβου να λειτουργούν ως whole-plot παράγοντες. Ο κύριος στόχος είναι να επιλέξουμε τις στάθμες για τους παράγοντες ελέγχου που ελαχιστοποιούν τη μέση απόκριση και ταυτόχρονα ελαχιστοποιούν τη διακύμανση του μοντέλου. Θεωρούμε το ακόλουθο μοντέλο απόκρισης

$$y = 25.00 + 15.31x_1 + 10.88x_5 + 11.18x_{11} + 12.85z_2 + 13.63z_5 + 11.18x_2z_2 + \epsilon, \quad (10.24)$$

όπου το πειραματικό σφάλμα  $\epsilon$  κατανέμεται ως  $N(0, 1)$ . Υπάρχουν δύο ενεργοί παράγοντες ελέγχου ( $C$ ) και δύο ενεργοί παράγοντες θορύβου ( $N$ ) καθώς και μια  $C \times N$  αλληλεπίδραση. Χρησιμοποιώντας τον προτεινόμενο πειραματικό σχεδιασμό  $D$  προσομοιώσαμε δεδομένα από την εξίσωση (10.24) πριν προχωρήσουμε στην εκτίμηση του μοντέλου απόκρισης.

**Πίνακας 10.4:** Whole-plot ανάλυση διασποράς (Στάδιο I) για το Παράδειγμα 1. Παλινδρόμηση κατά βήματα με  $\alpha = 0.2$ .

Πηγή μεταβλητότητας	βε	Adj SS	Adj MS	F-Value	P-Value
$z_2$	1	790.49	790.49	131503.94	0.00
$z_5$	1	895.87	895.87	149035.66	0.00
$z_8$	1	0.12	0.12	19.92	0.05
Σφάλμα	2	0.01	0.01		
Συνολικό	5	2680.94			

Εκτελώντας την ανάλυση σύμφωνα με την προαναφερθείσα διαδικασία (Κεφάλαιο 10.4) εντοπίσαμε πρώτα τους σημαντικούς whole-plot παράγοντες (Πίνακας 10.4). Σε επίπεδο σημαντικότητας 5% μόνο ο  $z_2$  και ο  $z_5$  φαίνεται να είναι σημαντικοί. Έτσι χρησιμοποιήθηκαν στο επόμενο στάδιο II.

Μετά τη διεξαγωγή της ανάλυσης για τους split-plot παράγοντες σύμφωνα με το

**Πίνακας 10.5:** Split-plot ανάλυση διασποράς (Στάδιο II) για το Παράδειγμα 1. Παλινδρόμηση κατά βήματα με  $\alpha = 0.2$ .

Πηγή μεταβλητότητας	βε	Adj SS	Adj MS	F-Value	P-Value
$x_1$	1	3730.00	3730.05	13172.55	0.00
$x_5$	1	1498.30	1498.25	5291.02	0.00
$x_{11}$	1	2060.10	2060.10	7275.19	0.00
$x_2z_2$	1	1223.30	1223.30	4320.06	0.00
$x_5z_2$	1	1.90	1.86	6.58	0.03
$x_6z_2$	1	0.80	0.81	2.87	0.12
$x_4z_5$	1	2.90	2.86	10.09	0.01
$x_{10}z_5$	1	0.60	0.60	2.13	0.18
WPE	5	10723.80	2144.75	7574.12	0.00
Σφάλμα	10	2.80	0.28		
Συνολικό	23	27692.00			

στάδιο II αποκτήσαμε τους σημαντικούς παράγοντες που παρουσιάζονται στον Πίνακα 10.5. Είναι προφανές ότι σε επίπεδο σημαντικότητας 1% οι παράγοντες  $x_1$ ,  $x_5$ ,  $x_{11}$  και η αλληλεπίδραση  $x_2z_2$  φαίνεται να είναι σημαντικοί. Ακολουθώντας τα βήματα του σταδίου III, αξιολογήσαμε πρώτα το εκτιμώμενο μοντέλο το οποίο περιγράφεται ως

$$\hat{y} = 24.90 + 15.38x_1 + 10.81x_5 + 11.43x_{11} + 12.57z_2 + 13.39z_5 + 9.76x_2z_2. \quad (10.25)$$

Σύμφωνα με την RPDM θεωρούμε ότι οι παράγοντες θορύβου είναι τυχαίες μεταβλητές με  $E[\mathbf{z}] = 0$  και  $Var[\mathbf{z}] = \Sigma_{\mathbf{z}} = \mathbf{I}$ . Επομένως το μοντέλο για το μέσο μπορεί να περιγραφεί ως

$$\widehat{E}[y] = 24.90 + 15.38x_1 + 10.81x_5 + 11.43x_{11},$$

και το αντίστοιχο μοντέλο της διασποράς ως

$$\widehat{Var}[y] = constant + (12.57 + 9.76x_2)^2 Var(z_2),$$

$$\widehat{Var}[y] = constant + 2(12.57)(9.76)x_2.$$

Τέλος, το μοντέλο της διασποράς λαμβάνεται ως

$$\widehat{Var}[y] = constant + 245.37x_2.$$

Κύριος στόχος είναι να ελαχιστοποιήσουμε την απόκριση κρατώντας, την ίδια στιγμή, τη διασπορά στο χαμηλότερο δυνατό επίπεδο. Έτσι, εάν οι περιοχές ενδιαφέροντος είναι  $x_i \leq 1$ ,  $i = 1, 2, \dots, 6$ , τότε τα βέλτιστα επίπεδα των παραγόντων ελέγχου εκτιμώνται τόσο μέσω του μοντέλου του μέσου όσο και του μοντέλου της διασποράς ως  $(x_1, x_2, x_5, x_{11}) = (-1, -1, -1, -1)$ . Οι στάθμες εκείνες που ελαχιστοποιούν το μοντέλο διασποράς, είναι οι ίδιες με εκείνες που επιλέχθηκαν στο μοντέλο του μέσου αφού η μόνη μεταβλητή που επηρεάζει το μοντέλο διασποράς είναι η κύρια επίδραση  $x_2$  η οποία ελαχιστοποιεί τη διασπορά με τα προαναφερόμενα επίπεδα.

**Παράδειγμα 13** Ας θεωρήσουμε τον σχεδιασμό  $D_1 \in D(6, 2^{10}, 1)$  για τους whole-plot παράγοντες ως

$$D_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 \\ \cdot & & & & & & & & & \end{pmatrix}$$

Θεωρούμε τους σχεδιασμούς  $d_k \in D(4, 2^3, 2)$  κάθε ένας να αντιστοιχεί στους split-plot παράγοντες ως

$$b_k = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

Ακολουθώντας το Λήμμα 14 κατασκευάζουμε τον σχεδιασμό ο οποίος είναι ένας δύο επιπέδων ισορροπημένος split-plot σχεδιασμός ο οποίος επιτρέπει την εκτίμηση του μοντέλου 10.1.

Υποθέτουμε ότι το πείραμα περιλαμβάνει μία απόκριση,  $m_1 = 10$  παράγοντες θορύβου (N) και  $m_2 = 3$  παράγοντες ελέγχου (C) όλοι σε δύο επίπεδα, με τους παράγοντες ελέγχου να λειτουργούν ως split-plot παράγοντες, και με τους παρά-

γοντες θορύβου να λειτουργούν ως whole-plot παράγοντες. Ο κύριος στόχος είναι να επιλέξουμε τις στάθμες για τους παράγοντες ελέγχου που μεγιστοποιούν τη μέση απόκριση και ταυτόχρονα ελαχιστοποιούν τη διακύμανση του μοντέλου. Θεωρούμε το ακόλουθο μοντέλο απόκρισης

$$y = 25.00 + 15.6x_1 + 16x_3 + 13.8z_1 + 10.6z_8 + 10.8x_2z_1 + \epsilon, \quad (10.26)$$

όπου το πειραματικό σφάλμα  $\epsilon$  κατανέμεται ως  $N(0, 1)$ . Υπάρχουν δύο ενεργοί παράγοντες ελέγχου (C) και δύο ενεργοί παράγοντες θορύβου (N) καθώς και μια  $C \times N$  αλληλεπίδραση. Χρησιμοποιώντας το προτεινόμενο πειραματικό σχέδιο  $D$  προσομοιώσαμε δεδομένα από την εξίσωση (10.26) πριν προχωρήσουμε στην εκτίμηση του μοντέλου απόκρισης.

$$D \in D^{(COA)}(24, 2^{13}, 1)$$

$$D = \begin{pmatrix} w_1 & w_2 & w_3 & w_4 & w_5 & w_6 & w_7 & w_8 & w_9 & w_{10} & s_1 & s_2 & s_3 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 \end{pmatrix}$$

Εκτελώντας την ανάλυση σύμφωνα με την προαναφερθείσα διαδικασία (Κεφάλαιο 10.4) εντοπίσαμε πρώτα τους σημαντικούς whole-plot παράγοντες (Πίνακας 10.6).

**Πίνακας 10.6:** Whole-plot ανάλυση διασποράς (Στάδιο I) για το Παράδειγμα 2. Παλινδρόμηση κατά βήματα με  $\alpha = 0.2$ .

Πηγή μεταβλητότητας	βε	Adj SS	Adj MS	F-Value	P-Value
$z_1$	1	23.190	23.190	321.18	0.04
$z_8$	1	34.580	34.580	478.94	0.03
$z_9$	1	323.139	323.139	4475.53	0.01
$z_{10}$	1	326.465	326.465	4521.61	0.01
Σφάλμα	1	0.072	0.072		
Συνολικό	5	857.275			

Σε επίπεδο σημαντικότητας 5%, οι παράγοντες  $z_1$ ,  $z_8$ ,  $z_9$  και ο  $z_{10}$  φαίνεται να είναι σημαντικοί. Έτσι χρησιμοποιήθηκαν στο επόμενο στάδιο II.

**Πίνακας 10.7:** Split-plot ανάλυση διασποράς (Στάδιο II) για το Παράδειγμα 2. Παλινδρόμηση κατά βήματα με  $\alpha = 0.2$ .

Πηγή μεταβλητότητας	βε	Adj SS	Adj MS	F-Value	P-Value
$x_1$	1	2464.3	2464.35	22372.18	0.00
$x_2$	1	0.4	0.40	3.67	0.11
$x_3$	1	3571.9	3571.88	32426.78	0.00
$x_1z_1$	1	2.4	2.40	21.77	0.01
$x_2z_1$	1	1629.7	1629.72	14795.12	0.00
$x_3z_1$	1	4.6	4.64	42.15	0.00
$x_1z_8$	1	0.4	0.39	3.54	0.12
$x_2z_8$	1	0.8	0.85	7.67	0.04
$x_3z_8$	1	3.1	3.07	27.86	0.00
$x_1z_9$	1	12.1	12.06	109.51	0.00
$x_2z_9$	1	1.0	0.96	8.74	0.03
$x_1z_{10}$	1	0.6	0.56	5.10	0.07
$x_3z_{10}$	1	4.1	4.06	36.89	0.00
WPE	5	3429.1	685.82	6226.11	0.000
Σφάλμα	5	0.6	0.11		
Συνολικό	23	18654.0			

Μετά τη διεξαγωγή της ανάλυσης για τους split-plot παράγοντες σύμφωνα με το στάδιο II αποκτήσαμε τους σημαντικούς παράγοντες που παρουσιάζονται στον Πίνακα 10.7. Είναι προφανές ότι σε επίπεδο σημαντικότητας 1% οι παράγοντες  $x_1$ ,  $x_3$  και οι αλληλεπιδράσεις  $x_2z_1$ ,  $x_3z_1$ ,  $x_3z_8$ ,  $x_1z_9$ ,  $x_3z_{10}$  φαίνεται να είναι σημαντικοί. Όπως μπορούμε να παρατηρήσουμε από τον Πίνακα 10.7, τα  $x_1$  και  $x_3$  προέκυψαν ως οι δύο σημαντικοί παράγοντες θέσης. Με αυτόν τον τρόπο θα πρέπει να συμπεριληφθούν στο μοντέλο του μέσου και όχι στο μοντέλο διασποράς. Το γεγονός αυτό ενισχύεται από το γεγονός ότι οι όροι αλληλεπίδρασης μεταξύ των παραγόντων θορύβου και των παραγόντων ελέγχου  $x_1$  και  $x_3$  φαίνεται πως είναι λιγότερο σημαντικοί



από τις κύριες επιδράσεις των  $x_1$  και  $x_3$ , αντίστοιχα. Έτσι, λαμβάνοντας υπόψη μόνο τους όρους αλληλεπίδρασης μεταξύ του  $x_2$  και των ενεργών παραγόντων θορύβου συμπεραίνουμε ότι θα πρέπει να λάβουμε υπόψη μόνο τη  $x_2z_1$  στο μοντέλο διασποράς. Ακολουθώντας τα βήματα του σταδίου III, αξιολογήσαμε πρώτα το εκτιμώμενο μοντέλο το οποίο περιγράφεται ως

$$\hat{y} = 16.96 + 15.20x_1 + 16.70x_3 + 2.20z_1 + 2.68z_8 + 8.21z_9 + 8.25z_{10} + 11.28x_2z_1.$$

Σύμφωνα με την RPDM θεωρούμε ότι οι παράγοντες θορύβου είναι τυχαίες μεταβλητές με  $E[\mathbf{z}] = 0$  και  $Var[\mathbf{z}] = \Sigma_{\mathbf{z}} = \mathbf{I}$ . Επομένως το μοντέλο για το μέσο μπορεί να περιγραφεί ως

$$\widehat{E}[y] = 16.96 + 15.20x_1 + 16.70x_3,$$

και το αντίστοιχο μοντέλο της διασποράς ως

$$\widehat{Var}[y] = constant + (2.20 + 11.28x_2)^2 Var(z_1),$$

$$\widehat{Cov}[y] = constant + 2(2.20)(11.28)x_2.$$

Τέλος, το μοντέλο της διασποράς λαμβάνεται ως

$$\widehat{Var}[y] = constant + 49.63x_2.$$

Κύριος στόχος είναι να μεγιστοποιήσουμε την απόκριση κρατώντας, την ίδια στιγμή, τη διασπορά στο χαμηλότερο δυνατό επίπεδο. Έτσι, εάν οι περιοχές ενδιαφέροντος είναι  $x_i \leq 1$ ,  $i = 1, 2, \dots, 6$ , τότε τα βέλτιστα επίπεδα των παραγόντων ελέγχου εκτιμώνται τόσο μέσω του μοντέλου του μέσου όσο και του μοντέλου της διασποράς ως  $(x_1, x_2, x_3) = (+1, -1, +1)$ . Οι στάθμες εκείνες που ελαχιστοποιούν το μοντέλο διασποράς, είναι οι ίδιες με εκείνες που επιλέχθηκαν στο μοντέλο του μέσου αφού η μόνη μεταβλητή που επηρεάζει το μοντέλο διασποράς είναι η κύρια επίδραση  $x_2$  η οποία ελαχιστοποιεί τη διασπορά με τα προαναφερόμενα επίπεδα.

**Παράδειγμα 14** Ας θεωρήσουμε τον σχεδιασμό  $D_1 \in D(6, 2^{10}, 1)$  για τους *whole plot* παράγοντες ως

$$D_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 \end{pmatrix}$$

Θεωρούμε τους σχεδιασμούς  $b_k \in D(4, 2^3, 2)$  και τους *foldover* σχεδιασμούς τους  $\bar{b}_k$  ο καθένας να αντιστοιχεί στους *split-plot* παράγοντες.

$$b_k = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}, \bar{b}_k = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

Ακολουθώντας το Λήμμα 14 κατασκευάζουμε έναν σχεδιασμό  $D \in D^{(COA)}(24, 2^{13}, 1)$  ο οποίος είναι ένας δύο επιπέδων ισορροπημένος *split-plot* σχεδιασμός και επιτρέπει την εκτίμηση του μοντέλου 10.2.

Υποθέτουμε ότι το πείραμα περιλαμβάνει μία απόκριση,  $m_1 = 10$  παράγοντες θορύβου ( $N$ ) και  $m_2 = 3$  παράγοντες ελέγχου ( $C$ ) όλοι σε δύο επίπεδα, με τους παράγοντες ελέγχου να λειτουργούν ως *split-plot* παράγοντες, και με τους παράγοντες θορύβου να λειτουργούν ως *whole-plot* παράγοντες. Ο κύριος στόχος είναι να επιλέξουμε τις στάθμες για τους παράγοντες ελέγχου που ελαχιστοποιούν τη μέση απόκριση και ταυτόχρονα ελαχιστοποιούν τη διακύμανση του μοντέλου. Θεωρούμε το ακόλουθο μοντέλο απόκρισης

$$y = 20.00 + 14.50x_2 + 13.00x_3 + 9.80z_1 + 8.60z_8 + 10.80x_2x_3 + 11.00x_1z_1 + \epsilon. \quad (10.27)$$

όπου το πειραματικό σφάλμα  $\epsilon$  κατανέμεται ως  $N(0, 1)$ . Υπάρχουν δύο ενεργοί παράγοντες ελέγχου ( $C$ ) και δύο ενεργοί παράγοντες θορύβου ( $N$ ) καθώς και μια  $C \times N$

αλληλεπίδραση. Χρησιμοποιώντας τον προτεινόμενο πειραματικό σχεδιασμό  $D$  προσομοιώσαμε δεδομένα από την εξίσωση (10.27) πριν προχωρήσουμε στην εκτίμηση του μοντέλου απόκρισης.

$$D = \begin{pmatrix} w_1 & w_2 & w_3 & w_4 & w_5 & w_6 & w_7 & w_8 & w_9 & w_{10} & s_1 & s_2 & s_3 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & -1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 \\ -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 \end{pmatrix},$$

**Πίνακας 10.8:** Whole-plot ανάλυση διασποράς (Στάδιο II) για το Παράδειγμα 3. Παλινδρόμηση κατά βήματα με  $\alpha = 0.2$ .

Πηγή μεταβλητότητας	$\beta\epsilon$	Adj SS	Adj MS	F-Value	P-Value
$z_1$	1	461.93	461.93	11584.48	0.00
$z_7$	1	0.38	0.38	9.43	0.09
$z_8$	1	384.92	384.92	9653.21	0.00
Σφάλμα	2	0.08	0.04		
Συνολικό	5	680.34			

Εκτελώντας την ανάλυση σύμφωνα με την προαναφερθείσα διαδικασία (Κεφάλαιο 10.4) εντοπίσαμε πρώτα τους σημαντικούς whole-plot παράγοντες (Πίνακας 10.8). Σε επίπεδο σημαντικότητας 5%, ο παράγοντας  $z_1$  και ο  $z_8$  φαίνεται να είναι σημαντικοί. Έτσι χρησιμοποιήθηκαν στο επόμενο στάδιο II.

**Πίνακας 10.9:** Split-plot ανάλυση διασποράς (Στάδιο II) για το Παράδειγμα 3. Παλινδρόμηση κατά βήματα με  $\alpha = 0.2$ .

Πηγή μεταβλητότητας	βε	Adj SS	Adj MS	F-Value	P-Value
$x_1$	1	2.30	2.27	9.09	0.02
$x_2$	1	4922.80	4922.77	19691.78	0.00
$x_3$	1	3743.40	3743.41	14974.18	0.00
$x_1x_2$	1	1.80	1.78	7.14	0.03
$x_1x_3$	1	1.40	1.38	5.54	0.05
$x_2x_3$	1	2376.80	2376.82	9507.62	0.00
$x_1z_1$	1	2500.20	2500.21	10001.19	0.00
$x_3z_1$	1	2.80	2.81	11.25	0.01
$x_1z_7$	1	2.30	2.33	9.31	0.02
$x_3z_8$	1	4.80	4.78	19.11	0.00
WPE	5	2721.3	544.27	2177.16	0.00
Σφάλμα	8	2.0	0.25		
Συνολικό	23	15301.90			

Μετά τη διεξαγωγή της ανάλυσης για τους split-plot παράγοντες σύμφωνα με το στάδιο II αποκτήσαμε τους σημαντικούς παράγοντες που παρουσιάζονται στον Πίνακα 10.9. Είναι προφανές ότι σε επίπεδο σημαντικότητας 1% οι παράγοντες  $x_2$  και  $x_3$  και οι αλληλεπιδράσεις  $x_2x_3$ ,  $x_1z_1$ ,  $x_3z_8$  φαίνεται να είναι σημαντικοί. Όπως μπορούμε να παρατηρήσουμε από τον Πίνακα 10.9, ο  $x_3$  εμφανίζεται ως σημαντικός παράγοντας θέσης. Με αυτόν τον τρόπο θα πρέπει να συμπεριληφθεί στο μοντέλο του μέσου και όχι στο μοντέλο διασποράς. Το γεγονός αυτό ενισχύεται από το γεγονός ότι ο όρος αλληλεπίδρασης  $x_3z_8$  φαίνεται ότι είναι λιγότερο σημαντικός από την κύρια επίδραση  $x_3$ . Έτσι, λαμβάνουμε υπόψη μόνο τον όρο αλληλεπίδρασης  $x_1z_1$  στο μοντέλο διασποράς. Ακολουθώντας τα βήματα του σταδίου III, αξιολογήσαμε πρώτα το εκτιμώμενο μοντέλο το οποίο περιγράφεται ως

$$\hat{y} = 19.99 + 14.32x_2 + 12.49x_3 + 9.61z_1 + 8.77z_8 + 10.90x_2x_3 + 11.18x_1z_1. \quad (10.28)$$

Σύμφωνα με την RPDM θεωρούμε ότι οι παράγοντες θορύβου είναι τυχαίες μεταβλητές με  $E[\mathbf{z}] = 0$  και  $Var[\mathbf{z}] = \Sigma_{\mathbf{z}} = \mathbf{I}$ . Επομένως το μοντέλο για το μέσο μπορεί να περιγραφεί ως

$$E[\hat{y}] = 19.99 + 14.32x_2 + 12.49x_3 + 10.90x_2x_3,$$

και το αντίστοιχο μοντέλο της διασποράς ως

$$\widehat{\text{Var}}[y] = \text{constant} + (9.61 + 11.18x_1)^2 \text{Var}(z_1),$$

$$\widehat{\text{Var}}[y] = \text{constant} + 2(9.61)(11.18)x_1.$$

Τέλος, το μοντέλο της διασποράς λαμβάνεται ως

$$\widehat{\text{Var}}[y] = \text{constant} + 214.88x_1.$$

Κύριος στόχος είναι να ελαχιστοποιήσουμε την απόκριση κρατώντας, την ίδια στιγμή, τη διασπορά στο χαμηλότερο δυνατό επίπεδο. Έτσι, εάν οι περιοχές ενδιαφέροντος είναι  $x_i \leq 1$ ,  $i = 1, 2, \dots, 6$ , τότε τα βέλτιστα επίπεδα των παραγόντων ελέγχου εκτιμώνται τόσο μέσω του μοντέλου του μέσου όσο και του μοντέλου της διασποράς ως  $(x_1, x_2, x_3) = (-1, -1, -1)$ . Οι στάθμες εκείνες που ελαχιστοποιούν το μοντέλο διασποράς, είναι οι ίδιες με εκείνες που επιλέχθηκαν στο μοντέλο του μέσου αφού η μόνη μεταβλητή που επηρεάζει το μοντέλο διασποράς είναι η κύρια επίδραση  $x_2$  η οποία ελαχιστοποιεί τη διασπορά με τα προαναφερόμενα επίπεδα.

## 10.6 Συμπεράσματα

Στη βιομηχανία, λόγω περιορισμών χρόνου και κόστους, το μέγεθος των πειραμάτων πρέπει να διατηρείται όσο το δυνατόν μικρότερο. Μερικές φορές τα πειράματα αυτά περιλαμβάνουν όχι μόνο παράγοντες που μπορούν να ελεγχθούν αλλά και ορισμένους ανεξέλεγκτους παράγοντες οδηγώντας σε RPD πειράματα. Οι split-plot σχεδιασμοί μπορούν να χρησιμοποιηθούν αποτελεσματικά για να μειώσουν το κόστος στον RPD πειραματισμό υπό την παρουσία παραγόντων που είναι δύσκολο να ελεγχθούν. Σε αυτό το πλαίσιο, όταν είναι απαραίτητο να εξεταστεί ένας μεγάλος αριθμός παραγόντων με περιορισμένες πειραματικές εκτελέσεις και όπου οι παραδοσιακοί split-plot σχεδιασμοί (επαναλαμβανόμενοι ή κλασματικοί παραγοντικοί) δεν είναι εφικτοί, τότε θα πρέπει να συμπεριληφθούν ΥΣ στο πλαίσιο των split-plot. Στο παρόν κεφάλαιο προτείνουμε μια μέθοδο κατασκευής για υπερκορεσμένους split-plot σχε-

διασμούς (SSSPDs) κατάλληλους για RPD πειράματα. Εφαρμόσαμε τη μεθοδολογία των σύνθετων ορθογώνιων σχεδιασμών χρησιμοποιώντας ορθογώνιους σχεδιασμούς ισχύος μεγαλύτερης από 1, για τις split-plot μονάδες και ισορροπημένους ΥΣ για τις whole-plot μονάδες. Οι SSSPDs που προτείνονται σε αυτό το κεφάλαιο μπορούν να θεωρηθούν ως μια χρήσιμη προσέγγιση για τον προσδιορισμό των σημαντικών παραγόντων στην RPDM. Αυτοί οι σχεδιασμοί οδηγούν πάντοτε σε ορθογωνιότητα μεταξύ των whole-plot και των split-plot παραγόντων και, κυρίως, αποτελούν μία γενίκευση άλλων μεθόδων που έχουν ήδη προταθεί στην βιβλιογραφία.

## **Μέρος IV**

### **Παραρτήματα – Βιβλιογραφία**





# Παράρτημα Α

## Παράρτημα Κεφαλαίου 5

**Πίνακας Α.1:**  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ.

$n$	$k$	$q$	$m$	Μέθοδος Κατασκευής	$E(f_{NOD})$	$L_{f_{NOD}}$	Αποδοτικότητα
6	2	3	10		2.6667	2.6667	1.0000
$G = (2111221233)$							
5	2	3	10	Πόρισμα 8	2.6667	2.6667	1.0000
4	2	3	8	Θεώρημα 3	2.3651	2.3651	1.0000
6	3	3	15		2.8571	2.8571	1.0000
$G = (211312123222131)$							
5	3	3	15	Πόρισμα 8	2.7937	2.7937	1.0000
4	3	3	12	Θεώρημα 3	2.4040	2.4040	1.0000
6	4	3	20		2.9474	2.9474	1.0000
$G = (12112222333211212113)$							
5	4	3	20	Πόρισμα 8	2.8538	2.8538	1.0000
4	4	3	16	Θεώρημα 3	2.4556	2.4556	1.0000
6	5	3	25		3.0000	3.0000	1.0000
$G = (1123211123222112312132312)$							
5	5	3	25	Πόρισμα 8	2.8889	2.8889	1.0000
4	5	3	20	Θεώρημα 3	2.4749	2.4749	1.0000
6	6	3	30		3.0345	3.0345	1.0000
$G = (12121232133323311112121211222)$							
5	6	3	30	Πόρισμα 8	2.9119	2.9119	1.0000
4	6	3	24	Θεώρημα 3	2.5266	2.4831	0.9828
6	7	3	35		3.0588	3.0588	1.0000
$G = (13122232211312313121122221321123121)$							
5	7	3	35	Πόρισμα 8	2.9281	2.9281	1.0000
4	7	3	28	Θεώρημα 3	2.4974	2.4974	1.0000
6	8	3	40		3.0769	3.0769	1.0000
$G = (2122211212323123112322113211123223111321)$							
5	8	3	40	Πόρισμα 8	2.9402	2.9402	1.0000
4	8	3	32	Θεώρημα 3	2.5045	2.5045	1.0000
6	9	3	45		3.0909	3.0909	1.0000
$G = (213332122112112213321223321121111211232221132)$							
5	9	3	45	Πόρισμα 8	2.9495	2.9495	1.0000
4	9	3	36	Θεώρημα 3	2.5079	2.5079	1.0000
6	10	3	50		3.1020	3.1020	1.0000
$G = (22132123221111122121333122121322221112121122333131)$							
5	10	3	50	Πόρισμα 8	2.9569	2.9569	1.0000
4	10	3	40	Θεώρημα 3	2.5197	2.5145	0.9980
6	11	3	55		3.1111	3.1111	1.0000
$G = (1212131111321312213211311321222212222113133213213222122)$							
5	11	3	55	Πόρισμα 8	2.9630	2.9630	1.0000
4	11	3	44	Θεώρημα 3	2.5520	2.5182	0.9867
6	12	3	60		3.1186	3.1186	1.0000
$G = (12222212221132311221213313123113122211111332131221232121312)$							
5	12	3	60	Πόρισμα 8	2.9680	2.9680	1.0000
4	12	3	48	Θεώρημα 3	2.5307	2.5201	0.9958

**Πίνακας Α.1:**  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ (συνέχεια) - Α \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.

$n$	$k$	$q$	$m$	Μέθοδος Κατασκευής	$E(f_{NOD})$	$L_{f_{NOD}}$	Αποδοτικότητα
9	3	3	24		3.9130	3.9130	1.0000
$G = (323122321113232112211231)$							
8	3	3	24	Πόρισμα 8	3.9324	3.9324	1.0000
7	3	3	18	Θεώρημα 3	3.6209	3.5163	0.9711
9	4	3	32		4.0645	4.0645	1.0000
$G = (22231121221123321321123231133112)$							
8	4	3	32	Πόρισμα 8	4.0502	4.0502	1.0000
7	4	3	24	Θεώρημα 3	3.7222	3.6208	0.9727
9	5	3	40		4.1539	4.1539	1.0000
$G = (2313232222211111132222323122111321331131)$							
8	5	3	40	Πόρισμα 8	4.1197	4.1197	1.0000
8	5	3	30	Θεώρημα 3*	3.7762	3.6797	0.9744
8	5	3	30	Θεώρημα 3	3.7946	3.6797	0.9697
9	6	3	48		4.1655	4.1655	1.0000
$G = (321311213123132123213212222211331222121332112131)$							
8	6	3	48	Πόρισμα 8	4.1655	4.1655	1.0000
7	6	3	36	Θεώρημα 3*	3.7714	3.7270	0.9882
7	6	3	36	Θεώρημα 3	3.7524	3.7270	0.9932
9	7	3	56		4.2545	4.2545	1.0000
$G = (22222323121223131321121231332112112123232212311113311321)$							
8	7	3	56	Πόρισμα 8	4.1980	4.1980	1.0000
7	7	3	42	Θεώρημα 3*	3.7995	3.7507	0.9872
7	7	3	42	Θεώρημα 3	3.8064	3.7507	0.9854
12	3	3	33		5	5	1.0000
$G = (212233121322223131111112331312223)$							
11	3	3	33	Πόρισμα 8	5.0556	5.0556	1.0000
10	3	3	24	Θεώρημα 3*	4.7512	4.5628	0.9603
10	3	3	24	Θεώρημα 3	4.7150	4.5628	0.9677
12	4	3	44		5.2093	5.2093	1.0000
$G = (1212231221213111111223212212333332212233133)$							
11	4	3	44	Πόρισμα 8	5.2300	5.2300	1.0000
10	4	3	32	Θεώρημα 3*	4.9252	4.7437	0.9632
10	4	3	32	Θεώρημα 3	4.9050	4.7437	0.9671
12	5	3	55		5.3333	5.3333	1.0000
$G = (111232113222231121232311112133312132332123312321123122)$							
11	5	3	55	Πόρισμα 8	5.3333	5.3333	1.0000
10	5	3	40	Θεώρημα 3*	4.9915	4.8504	0.9717
10	5	3	40	Θεώρημα 3	4.9966	4.8504	0.9707
8	2	4	14		4.6154	4.6154	1.0000
$G = (23311222144133)$							
7	2	4	14	Πόρισμα 8	4.3990	4.3990	1.0000
6	2	4	12	Θεώρημα 3	4.0227	4.0227	1.0000
8	3	4	21		4.8000	4.8000	1.0000
$G = (344423111213232121332)$							
7	3	4	21	Πόρισμα 8	4.3990	4.3990	1.0000
6	3	4	18	Θεώρημα 3	4.1291	4.1029	1.0000
8	4	4	28		4.8889	4.8889	1.0000
$G = (313111134434322224213112323)$							
7	4	4	28	Πόρισμα 8	4.6042	4.6042	1.0000
6	4	4	24	Θεώρημα 3*	4.1558	4.1413	0.9965
6	4	4	24	Θεώρημα 3	4.1630	4.1413	0.9948

**Πίνακας Α.1:**  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ (συνέχεια) - Α \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.

$n$	$k$	$q$	$m$	Μέθοδος Κατασκευής	$E(f_{NOD})$	$L_{f_{NOD}}$	Αποδοτικότητα
8	5	4	35		4.9412	4.9412	1.0000
$G = (23321322311431423132421121122331443)$							
7	5	4	35	Πόρισμα 8	4.6434	4.6434	1.0000
6	5	4	30	Θεώρημα 3	4.1822	4.1637	0.9956
8	6	4	42		4.9756	4.9756	1.0000
$G = (14343323112132211123323332224111312414342)$							
7	6	4	42	Πόρισμα 8	4.6692	4.6692	1.0000
6	6	4	30	Θεώρημα 3	4.1976	4.1881	0.9977
8	7	4	49		5	5	1.0000
$G = (3431422213423132223431213112134311143112232122334)$							
7	7	4	49	Πόρισμα 8	4.6875	4.6875	1.0000
6	7	4	42	Θεώρημα 3*	4.2030	4.2030	1.0000
6	7	4	42	Θεώρημα 3	4.2123	4.2030	0.9978
8	8	4	56		5.0182	5.0182	1.0000
$G = (21411341113441331333231234223211321114224312222422233133)$							
7	8	4	56	Πόρισμα 8	4.7011	4.7011	1.0000
6	8	4	48	Θεώρημα 3*	4.2163	4.2128	0.9992
6	8	4	48	Θεώρημα 3	4.2116	4.2128	0.9979
12	3	4	33		6.7500	6.7500	1.0000
$G = (343133442211234114122332211323421)$							
11	3	4	33	Πόρισμα 8	6.5625	6.5625	1.0000
10	3	4	27	Θεώρημα 3*	6.1660	6.1090	0.9908
10	3	4	27	Θεώρημα 3	6.2001	6.1090	0.9853
12	4	4	44		6.9070	6.9070	1.0000
$G = (23243321113332242413133212123143424241312411)$							
11	4	4	44	Πόρισμα 8	6.6933	6.6933	1.0000
10	4	4	36	Θεώρημα 3*	6.2802	6.2357	0.9929
10	4	4	36	Θεώρημα 3	6.2706	6.2357	0.9944
12	5	4	55		7	7	1.0000
$G = (323124111323144312132224133211343224341342311112342432)$							
11	5	4	55	Πόρισμα 8	6.7708	6.7708	1.0000
10	5	4	45	Θεώρημα 3*	6.3641	6.2955	0.9892
10	5	4	45	Θεώρημα 3	6.3621	6.2955	0.9895
16	4	4	60		8.9492	8.9492	1.0000
$G = (14323111124221231324121341242441424134222134231142133333332)$							
15	4	4	60	Πόρισμα 8	8.7680	8.7680	1.0000
14	4	4	48	Θεώρημα 3*	8.4309	8.2766	0.9817
14	4	4	48	Θεώρημα 3	8.3759	8.2766	0.9881
10	2	5	18		6.5882	6.5882	1.0000
$G = (321353451431242241)$							
9	2	5	18	Πόρισμα 8	6.2306	6.2306	1.0000
8	2	5	16	Θεώρημα 3*	5.7733	5.7733	1.0000
8	2	5	16	Θεώρημα 3	5.7933	5.7733	0.9971
10	3	5	27		6.7692	6.7692	1.0000
$G = (124212314151433335422541243)$							
9	3	5	27	Πόρισμα 8	6.3754	6.3754	1.0000
8	3	5	24	Θεώρημα 3*	5.8748	5.8748	1.0000
8	3	5	24	Θεώρημα 3	5.9183	5.8748	0.9927
10	4	5	36		6.8571	6.8571	1.0000
$G = (123423232142431244345451352512133141)$							
9	4	5	36	Πόρισμα 8	6.4457	6.4457	1.0000
8	4	5	32	Θεώρημα 3*	5.9360	5.9239	0.9980
8	4	5	32	Θεώρημα 3	5.9400	5.9239	0.9973

**Πίνακας Α.1:**  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ (συνέχεια) - Α \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.

$n$	$k$	$q$	$m$	Μέθοδος Κατασκευής	$E(f_{NOD})$	$L_{f_{NOD}}$	Αποδοτικότητα
10	5	5	45		6.9091	6.9091	1.0000
$G = (144432334412155332114142134222553322151442133)$							
9	5	5	45	Πόρισμα 8	6.4873	6.4873	1.0000
8	5	5	40	Θεώρημα 3*	5.9605	5.9528	0.9987
8	5	5	40	Θεώρημα 3	5.9579	5.9528	0.9991
10	6	5	54		6.9434	6.9434	1.0000
$G = (113113124424532545223411342332445343351224434232211151)$							
9	6	5	54	Πόρισμα 8	6.5147	6.5147	1.0000
8	6	5	48	Θεώρημα 3*	5.9790	5.9719	
8	6	5	48	Θεώρημα 3	5.9896	5.9719	
15	2	5	28		9.3333	9.3333	1.0000
$G = (5143442334451321243252153211)$							
14	2	5	28	Πόρισμα 8	9.0489	9.0489	1.0000
13	2	5	24	Θεώρημα 3*	8.6530	8.5878	0.9924
13	2	5	24	Θεώρημα 3	8.6168	8.5878	0.9966
15	3	5	42		9.6585	9.6585	1.0000
$G = (125213121415322541431132454214533353342244)$							
14	3	5	42	Πόρισμα 8	9.3307	9.3307	1.0000
13	3	5	36	Θεώρημα 3*	8.8654	8.8114	0.9939
13	3	5	36	Θεώρημα 3	8.8717	8.8114	0.9932
12	2	6	22		8.5714	8.5714	1.0000
$G = (3531425643642352151124)$							
11	2	6	22	Πόρισμα 8	8.1151	8.1151	1.0000
10	2	6	20	Θεώρημα 3*	7.5906	7.5906	1.0000
10	2	6	20	Θεώρημα 3	7.6012	7.5906	0.9986
12	3	6	33		8.7500	8.7500	1.0000
$G = (245112423532535656224411354163341)$							
11	3	6	33	Πόρισμα 8	8.2639	8.2639	1.0000
10	3	6	30	Θεώρημα 3	7.7096	7.7050	0.9994
12	4	6	44		8.8372	8.8372	1.0000
$G = (45455235241643222151145433241212563365433161)$							
11	4	6	44	Πόρισμα 8	8.3366	8.3366	1.0000
10	4	6	40	Θεώρημα 3*	7.7607	7.7607	1.0000
10	4	6	40	Θεώρημα 3	7.7658	7.7607	0.9993
12	5	6	55		8.8889	8.8889	1.0000
$G = (5352642121364123325224444113332123555561643551261445143)$							
11	5	6	55	Πόρισμα 8	8.3796	8.3796	1.0000
10	5	6	50	Θεώρημα 3*	7.8002	7.7937	0.9992
10	5	6	50	Θεώρημα 3	7.7969	7.7937	0.9996
14	2	7	26		10.5600	10.5600	1.0000
$G = (67154621125235544333266471)$							
13	2	7	26	Πόρισμα 8	10.0310	10.0310	1.0000
12	2	7	24	Θεώρημα 3*	9.4525	9.4525	1.0000
12	2	7	24	Θεώρημα 3	9.4598	9.4525	0.9992
14	3	7	39		10.7368	10.7368	1.0000
$G = (23432241315633457324545726657411662161)$							
13	3	7	39	Πόρισμα 8	10.1826	10.1826	1.0000
12	3	7	36	Θεώρημα 3*	9.5850	9.5755	0.9990
12	3	7	36	Θεώρημα 3	9.5850	9.5755	0.9993
14	4	7	52		10.8235	10.8235	1.0000
$G = (6621213465524263424234157564577521161657143133265343)$							
13	4	7	52	Πόρισμα 8	10.2569	10.2569	1.0000
12	4	7	48	Θεώρημα 3*	9.6410	9.6357	0.9994
12	4	7	48	Θεώρημα 3	9.6375	9.6357	0.9998

**Πίνακας Α.1:**  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί πολλαπλών επιπέδων ΥΣ (συνέχεια) - Α \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.

$n$	$k$	$q$	$m$	Μέθοδος Κατασκευής	$E(f_{NOD})$	$L_{f_{NOD}}$	Αποδοτικότητα
16	2	8	30		12.5517	12.5517	1.0000
$G = (515312666435267423481785724137)$							
15	2	8	30	Πόρισμα 8	11.9671	11.9671	1.0000
14	2	8	28	Θεώρημα 3	11.3449	11.3449	1.0000
16	3	8	45		12.7273	12.7273	1.0000
$G = (652578524445413127636354385236712261873161747)$							
15	3	8	45	Πόρισμα 8	12.1207	12.1207	1.0000
14	3	8	42	Θεώρημα 3	11.4741	11.4741	1.0000
16	4	8	60		12.8136	12.8136	1.0000
$G = (132634158872716336651742424651886237647423235554471125517637)$							
15	4	8	60	Πόρισμα 8	12.1962	12.1962	1.0000
14	4	8	56	Θεώρημα 3*	11.5414	11.5375	0.9997
14	4	8	56	Θεώρημα 3	11.5401	11.5375	0.9998
24	2	8	46		18.2000	18.2000	1.0000
$G = (1643374874546551473136522375842262651621187387)$							
23	2	8	46	Πόρισμα 8	17.6677	17.6677	1.0000
22	2	8	42	Θεώρημα 3*	17.1018	17.0716	0.9982
22	2	8	42	Θεώρημα 3	17.0879	17.0716	0.9990
18	2	9	34		14.5455	14.5455	1.0000
$G = (6437261815727188294687925533446351)$							
17	2	9	34	Πόρισμα 8	13.9169	13.9169	1.0000
16	2	9	32	Θεώρημα 3	13.2589	13.2589	1.0000
18	3	9	51		14.7200	14.7200	1.0000
$G = (372568735572815444926897121661783353137646218289454)$							
17	3	9	51	Πόρισμα 8	14.0721	14.0721	1.0000
16	3	9	48	Θεώρημα 3*	13.3927	13.3927	1.0000
16	3	9	48	Θεώρημα 3	13.3945	13.3927	0.9999
20	2	10	38		16.5405	16.5405	1.0000
$G = (9314311042749727265815991028566845133786)$							
19	2	10	38	Πόρισμα 8	15.8765	15.8765	1.0000
18	2	10	36	Θεώρημα 3	15.1886	15.1886	1.0000
20	3	10	57		16.7143	16.7143	1.0000
$G = (928126647448162891101355931578189248796353563427975637102104)$							
19	3	10	57	Πόρισμα 8	16.0329	16.0329	1.0000
18	3	10	54	Θεώρημα 3*	15.3260	15.3260	1.0000
18	3	10	54	Θεώρημα 3	15.3274	15.3274	0.9999

**Πίνακας Α.2:**  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί μικτών επιπέδων ΥΣ - Α \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.

$n$	$k$	$m$	Διάγραμμα Επιπέδων	Τύπος Σχεδιασμού	$E(f_{NOD})$	$L_{f_{NOD}}$	Αποδοτικότητα
8	8	56	(2, 2, 2, 2, 2, 2, 4)		2.3455	2.3455	1.0000
$G = (2111211412212112212212212212221422212231121122312122122)$							
7	8	56		Πόρισμα 9	2.5395	2.5395	1.0000
6	8	34		Θεώρημα 5	2.1003	2.0574	0.9796
8	8	56	(2, 2, 2, 2, 2, 2, 4, 4)		2.6727	2.6727	1.0000
$G = (2222213421121212111212231121213412112241222112322212242)$							
7	8	34		Πόρισμα 9	2.8134	2.8134	1.0000
6	8	34		Θεώρημα 5*	2.5357	2.4881	0.9812
6	8	34		Θεώρημα 5	2.6893	2.6353	0.9799
8	8	56	(2, 2, 2, 2, 2, 4, 4, 4)		3.0182	3.0182	1.0000
$G = (2112233211221213222242211212123221113442221124112122434)$							
7	8	56		Πόρισμα 9	3.0989	3.0989	1.0000
6	8	36		Θεώρημα 5*	2.8645	2.8645	1.0000
6	8	36		Θεώρημα 5	2.8702	2.8702	0.9980
8	8	56	(2, 2, 2, 2, 4, 4, 4, 4)		3.3818	3.3818	1.0000
$G = (2211223411214134221133121212144321224421222332211222243)$							
7	8	56		Πόρισμα 9	3.3960	3.3960	1.0000
6	8	36		Θεώρημα 5*	3.2346	3.1962	0.9881
6	8	36		Θεώρημα 5	3.1093	3.0657	0.9860
8	7	49	(2, 2, 2, 2, 2, 2, 4)		2.3571	2.3571	1.0000
$G = (12222411211122112213112212222221122111232211224)$							
7	7	49		Πόρισμα 9	2.5525	2.5525	1.0000
6	7	31		Θεώρημα 5	2.2070	2.0887	0.9464
8	7	49	(2, 2, 2, 2, 2, 4, 4)		2.7381	2.7381	1.0000
$G = (1212234212123411122122212143122124222211232122121)$							
7	7	49		Πόρισμα 9	2.8702	2.8702	1.0000
6	7	32		Θεώρημα 5*	2.7208	2.5796	0.9481
6	7	32		Θεώρημα 5	2.5559	2.4110	0.9432
8	7	49	(2, 2, 2, 2, 4, 4, 4)		3.1429	3.1429	1.0000
$G = (11212212122242121231311114442224342211332222123)$							
7	7	49		Πόρισμα 9	3.2031	3.2031	1.0000
6	7	32		Θεώρημα 5*	3.0619	2.9906	0.9767
6	7	32		Θεώρημα 5	2.9287	2.8396	0.9696
8	6	42	(2, 2, 2, 2, 2, 4)		2.3740	2.3740	1.0000
$G = (11122422112321212311221222121412221222121)$							
7	6	42		Πόρισμα 9	2.5706	2.5706	1.0000
6	6	26		Θεώρημα 5	2.2592	2.1485	0.9510
8	6	42	(2, 2, 2, 2, 4, 4)		2.8293	2.8293	1.0000
$G = (12224311114422122411213321123122212222212)$							
6	42			Πόρισμα 9	2.8293	2.8293	1.0000
6	6	28		Θεώρημα 5*	2.7923	2.6865	0.9621
6	6	28		Θεώρημα 5	2.6342	2.5073	0.9518
8	5	35	(2, 2, 2, 2, 4)		2.4000	2.4000	1.0000
$G = (11114111222221222223212241221322121)$							
7	5	35		Πόρισμα 9	2.5978	2.5978	1.0000
6	5	22		Θεώρημα 5	2.2694	2.2175	0.9771
8	2	14	(2, 4)		2.7692	2.7692	1.0000
$G = (22131424132221)$							
7	2	14		Πόρισμα 9	2.9375	2.9375	1.0000
6	2	10		Θεώρημα 5	2.9833	2.9833	0.9796

**Πίνακας Α.2:**  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί μικτών επιπέδων ΥΣ (συνέχεια) - Α \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.

$n$	$k$	$m$	Διάγραμμα Επιπέδων	Μέθοδος Κατασκευής	$E(f_{NOD})$	$L_{f_{NOD}}$	Αποδοτικότητα
8	3	21	(2, 4, 4)		3.6000	3.6000	1.0000
$G = (242113231234144223122)$							
7	3	21		Πόρισμα 9	3.5896	3.5896	1.0000
6	3	16		Θεώρημα 5*	3.5125	3.4125	0.9715
6	3	16		Θεώρημα 5	3.1813	3.0813	0.9686
10	3	27	(2, 2, 5)		3.3077	3.3077	1.0000
$G = (213124115211214122225122223)$							
9	3	27		Πόρισμα 9	3.4869	3.4869	1.0000
8	3	18		Θεώρημα 5*	3.6622	3.4008	0.9286
8	3	18		Θεώρημα 5	3.6099	3.4008	0.9421
10	2	18	(2, 5)		3.7647	3.7647	1.0000
$G = (121424252311251322)$							
9	2	18		Πόρισμα 9	3.8906	3.8906	1.0000
8	2	13		Θεώρημα 5*	4.2092	3.9015	0.9269
8	2	13		Θεώρημα 5	3.8390	3.6338	0.9466
12	5	55	(2, 2, 2, 6)		3.6741	3.6741	1.0000
$G = (1212421113111231121612112222262222522214112212112522212)$							
11	5	55		Πόρισμα 9	3.8743	3.8743	1.0000
10	5	22		Θεώρημα 5*	3.9376	3.7166	0.9439
10	5	22		Θεώρημα 5	3.8520	3.7166	0.9648
12	4	44	(2, 2, 2, 6)		3.8140	3.8140	1.0000
$G = (1123211122222231215212522142226111212161124)$							
11	4	44		Πόρισμα 9	4.0052	4.0052	1.0000
10	4	28		Θεώρημα 5	4.2169	3.9418	0.9348
12	3	33	(2, 2, 6)		4.0833	4.0833	1.0000
$G = (123222112226224215211116125123214)$							
11	3	33		Πόρισμα 9	4.2523	4.2523	1.0000
10	3	22		Θεώρημα 5*	4.6190	4.3506	0.9419
10	3	22		Θεώρημα 5	4.6537	4.3506	0.9349
12	2	22	(2, 6)		4.7619	4.7619	1.0000
$G = (2422132612161425252311)$							
11	2	22		Πόρισμα 9	4.8585	4.8585	1.0000
10	2	16		Θεώρημα 5*	5.3333	5.1667	0.9688
10	2	16		Θεώρημα 5	5.4667	5.1667	0.9451
14	4	52	(2, 2, 2, 7)		4.5000	4.5000	1.0000
$G = (1115121722242111222511222126122422132122212312161217)$							
13	4	52		Πόρισμα 9	4.6889	4.6889	1.0000
12	4	33		Θεώρημα 5*	5.0910	4.7917	0.9412
12	4	33		Θεώρημα 5	5.2547	4.9858	0.9488
14	4	52	(2, 2, 7, 7)		6.3627	6.3627	1.0000
$G = (1153227612321275221421472266223221531125224411211167)$							
13	4			Πόρισμα 9	6.3511	6.3511	1.0000
12	4			Θεώρημα 5*	7.0596	6.8007	0.9633
12	4			Θεώρημα 5	7.4449	7.1974	0.9668
14	4	55	(2, 7, 7, 7)		8.4706	8.4706	1.0000
$G = (2672276414362631152313742723231511572552124612452467)$							
13	4	52		Πόρισμα 9	8.2071	8.2071	1.0000
12	4	43		Θεώρημα 5*	8.4901	8.3683	0.9857
12	4	43		Θεώρημα 5	8.2895	8.2031	0.9896
14	4	39	(2, 7, 7)		7.5789	7.5789	1.0000
$G = (164144153232175147266113257221276222235)$							
13	4	39		Πόρισμα 9	7.4284	7.4284	1.0000
12	4	31		Θεώρημα 5*	7.9202	7.7525	0.9788
12	4	31		Θεώρημα 5	8.1581	8.0334	0.9847

**Πίνακας Α.2:**  $E(f_{NOD})$ -βέλτιστοι σχεδόν ισορροπημένοι  $k$ -κυκλικοί μικτών επιπέδων ΥΣ (συνέχεια) - Α \* δείχνει ότι η γραμμή με το ίδιο επίπεδο έχει αφαιρεθεί.

$n$	$k$	$m$	Διάλυση επιπέδων	Μέθοδος Κατασκευής	$E(f_{NOD})$	$L_{f_{NOD}}$	Αποδοτικότητα
14	2	26	(2, 7)		5.7600	5.7600	1.0000
$G = (16241213252511221327272614)$							
13	2	26		Πόρισμα 9	5.8351	5.8351	1.0000
12	2	19		Θεώρημα 5*	6.7371	6.3043	0.9358
12	2	19		Θεώρημα 5	6.2996	5.9487	0.9443
15	4	56	(3, 3, 5, 5)		8.0182	8.0182	1.0000
$G = (11321144213513453222324323542223231233513334111132232255)$							
14	4	56		Πόρισμα 9	7.8781	7.8781	1.0000
13	4	44		Θεώρημα 5*	7.6369	7.5142	0.9839
13	4	44		Θεώρημα 5	7.8521	7.7379	0.9855
15	4	56	(3, 5, 5, 5)		8.9000	8.9000	1.0000
$G = (1215233134233444124425151152244225522233333355111353324)$							
14	4	56		Πόρισμα 9	8.6589	8.6589	1.0000
13	4	46		Θεώρημα 5*	8.3509	8.2620	0.9894
13	4	46		Θεώρημα 5	8.3954	8.2620	0.9841
15	3	42	(3, 3, 5)		7.2195	7.2195	1.0000
$G = (332123315213124324121232122211334313235235)$							
14	3	42		Πόρισμα 9	7.1736	7.1736	1.0000
13	3	32		Θεώρημα 5*	6.9638	6.7945	0.9757
13	3	32		Θεώρημα 5	6.8715	6.7022	0.9754
15	3	42	(3, 5, 5)		8.4065	8.4065	1.0000
$G = (342343134322213223121255335244154232155311)$							
14	3	42		Πόρισμα 9	8.2260	8.2260	1.0000
13	3	34		Θεώρημα 5*	8.0183	7.8614	0.9804
13	3	34		Θεώρημα 5	7.8957	7.7780	0.9851
15	2	28	(3, 5)		7.3704	7.3704	1.0000
$G = (2321152124341235321313352432)$							
14	2	28		Πόρισμα 9	7.3167	7.3167	1.0000
13	2	22		Θεώρημα 5*	7.2287	6.9950	0.9677
13	2	22		Θεώρημα 5	7.1941	6.9950	0.9723



## Παράρτημα Β

### Παράρτημα Κεφαλαίου 6

**Πίνακας Β.1:** Αποτελέσματα συγκρίσεων προσομοίωσης μεταξύ των μεθόδων επιλογής μεταβλητών για τους δύο-επιπέδων « $K$ »,  $E(s^2)$ -βέλτιστους, σχεδιασμούς. Η πρώτη στήλη παρουσιάζει το σχεδιασμό, η δεύτερη στήλη αναφέρεται στον αριθμό των πραγματικά ενεργών παραγόντων  $d$  ενώ οι άλλες στήλες παρουσιάζουν τα σφάλματα που αντιστοιχούν σε κάθε μέθοδο. Οι σειρές που αναφέρονται ως **mean**, παρουσιάζουν το συνολικό μέσο όρο των αντιστοιχών σφαλμάτων για κάθε σχεδιασμό.

Designs	$d$	SVR-RFE		Dantzig selector		$L_q$ -penalty		stepwise regression	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
$K1$	1	0.0995	0.0000	0.0032	0.0000	0.08553	0.0000	0.1404	0.0000
	2	0.0916	0.0440	0.0780	0.0115	0.06566	0.0270	0.1298	0.0410
	3	0.0833	0.1257	0.0947	0.1213	0.0414	0.2913	0.1348	0.2350
	4	0.0882	0.2210	0.1204	0.2328	0.0447	0.5043	0.1968	0.4090
	<b>mean</b>	<b>0.0907</b>	<b>0.0977</b>	<b>0.0741</b>	<b>0.0914</b>	<b>0.05932</b>	<b>0.2057</b>	<b>0.1505</b>	<b>0.1710</b>
$K2$	1	0.1268	0.0000	0.0012	0.0000	0.0795	0.0000	0.1674	0.0000
	2	0.0889	0.0065	0.0303	0.0040	0.0528	0.0001	0.1443	0.0065
	3	0.0641	0.0240	0.0679	0.0590	0.0381	0.1760	0.1658	0.1217
	4	0.0580	0.0850	0.0846	0.1580	0.0316	0.3985	0.2154	0.2983
	5	0.0626	0.1636	0.109	0.2472	0.0325	0.5570	0.2688	0.4240
<b>mean</b>	<b>0.1160</b>	<b>0.0977</b>	<b>0.0743</b>	<b>0.0914</b>	<b>0.0752</b>	<b>0.2057</b>	<b>0.1839</b>	<b>0.1710</b>	
$K3$	1	0.1144	0.0000	0.0005	0.0000	0.0718	0	0.1744	0.0000
	2	0.0913	0.0015	0.0108	0.0005	0.0481	0.0000	0.1697	0.0000
	3	0.0679	0.0176	0.0304	0.0400	0.028	0.0783	0.1646	0.0473
	4	0.0629	0.0400	0.0563	0.1125	0.0198	0.2848	0.1900	0.1460
	5	0.0579	0.1028	0.0698	0.1902	0.0155	0.4722	0.2424	0.2992
	6	0.0615	0.1653	0.0823	0.2528	0.0154	0.5868	0.2843	0.4157
<b>mean</b>	<b>0.0760</b>	<b>0.0545</b>	<b>0.0417</b>	<b>0.0993</b>	<b>0.0331</b>	<b>0.2370</b>	<b>0.2042</b>	<b>0.1514</b>	
$K4$	1	0.12	0.0000	0.0001	0.0000	0.0636	0.0000	0.1902	0.0000
	2	0.098	0.004	0.005	0.0025	0.045	0.0000	0.18	0.009
	3	0.0756	0.0087	0.0217	0.0257	0.0326	0.0427	0.1783	0.0343
	4	0.0548	0.0195	0.0372	0.0793	0.0204	0.2005	0.1807	0.0895
	5	0.05	0.053	0.0553	0.1476	0.0193	0.3952	0.2162	0.2048
	6	0.0526	0.1005	0.0719	0.2108	0.0167	0.5533	0.2589	0.3207
	7	0.0575	0.1564	0.0806	0.2807	0.0139	0.6537	0.2949	0.4100
<b>mean</b>	<b>0.0726</b>	<b>0.0489</b>	<b>0.0388</b>	<b>0.1067</b>	<b>0.0302</b>	<b>0.2636</b>	<b>0.2142</b>	<b>0.1526</b>	
$K5$	1	0.1080	0.0000	0.0001	0.0000	0.062	0.0000	0.1979	0.0000
	2	0.0871	0.0030	0.0007	0.0010	0.0450	0.0000	0.1938	0.0000
	3	0.071	0.0057	0.0059	0.0097	0.0294	0.0190	0.1817	0.0000
	4	0.0561	0.0153	0.0186	0.0488	0.0163	0.1033	0.1827	0.0340
	5	0.0508	0.0432	0.0334	0.0980	0.0116	0.2830	0.1915	0.0934
	6	0.0510	0.0745	0.0456	0.1568	0.0091	0.4407	0.2275	0.2015
	7	0.0553	0.1196	0.0573	0.2114	0.0072	0.5651	0.2697	0.3146
	8	0.0634	0.1751	0.0717	0.2534	0.0069	0.6600	0.3034	0.4058
<b>mean</b>	<b>0.0678</b>	<b>0.0546</b>	<b>0.0292</b>	<b>0.0974</b>	<b>0.0234</b>	<b>0.2589</b>	<b>0.2185</b>	<b>0.1312</b>	

**Πίνακας Β.2:** Αποτελέσματα συγκρίσεων προσομοίωσης μεταξύ των μεθόδων επιλογής μεταβλητών για τους δύο-επιπέδων « $K$ »,  $E(s^2)$ -βέλτιστους, σχεδιασμούς. Η πρώτη στήλη παρουσιάζει το σχεδιασμό, η δεύτερη στήλη αναφέρεται στον αριθμό των πραγματικά ενεργών παραγόντων  $d$  ενώ οι άλλες στήλες παρουσιάζουν τα σφάλματα που αντιστοιχούν σε κάθε μέθοδο. Οι σειρές που αναφέρονται ως **mean**, παρουσιάζουν το συνολικό μέσο όρο των αντίστοιχων σφαλμάτων για κάθε σχεδιασμό.

Designs	$d$	SVR-RFE		Dantzig selector		$L_q$ -penalty		stepwise regression	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
$K6$	1	0.1043	0.0000	0.0001	0.0000	0.0565	0.0000	0.2149	0.0000
	2	0.0850	0.0010	0.0011	0.0010	0.0415	0.0000	0.1998	0.0000
	3	0.0696	0.0047	0.0044	0.0090	0.0293	0.005	0.1917	0.0023
	4	0.0567	0.0108	0.0156	0.0418	0.0193	0.0625	0.1835	0.0238
	5	0.0499	0.0252	0.0232	0.0894	0.01262	0.1944	0.1915	0.0658
	6	0.0482	0.0440	0.0353	0.1388	0.00882	0.3648	0.2128	0.1367
	7	0.0502	0.0776	0.0461	0.1816	0.0078	0.4927	0.2444	0.2261
	8	0.0522	0.1120	0.0570	0.2243	0.00735	0.6046	0.2840	0.3197
	9	0.0551	0.1537	0.0629	0.2699	0.00504	0.6737	0.3088	0.395
<b>mean</b>		<b>0.0635</b>	<b>0.0477</b>	<b>0.0273</b>	<b>0.1062</b>	<b>0.0209</b>	<b>0.2664</b>	<b>0.2257</b>	<b>0.1299</b>
$K7$	1	0.0871	0.0000	0.0001	0.0000	0.0535	0.0000	0.2139	0.0000
	2	0.0787	0.0000	0.0000	0.0020	0.0410	0.0000	0.2054	0.0001
	3	0.0703	0.002	0.0014	0.0067	0.0292	0.0000	0.1974	0.0000
	4	0.0537	0.0085	0.0061	0.0238	0.0191	0.0303	0.1887	0.0043
	5	0.0503	0.0244	0.0162	0.0570	0.0112	0.1302	0.1880	0.0228
	6	0.0467	0.0348	0.0276	0.0990	0.0072	0.2792	0.1983	0.0678
	7	0.0447	0.0534	0.0345	0.1480	0.0056	0.4259	0.2230	0.1571
	8	0.0515	0.0896	0.0434	0.1868	0.0047	0.5401	0.2609	0.2532
	9	0.0574	0.1302	0.0536	0.2340	0.0047	0.6239	0.2993	0.3456
	10	0.0608	0.1707	0.05932	0.2579	0.0036	0.6967	0.3173	0.4065
<b>mean</b>		<b>0.0601</b>	<b>0.0514</b>	<b>0.0776</b>	<b>0.1015</b>	<b>0.0180</b>	<b>0.2726</b>	<b>0.2292</b>	<b>0.1257</b>
$K8$	1	0.0803	0.0000	0.0001	0.0000	0.0493	0.0000	0.2239	0.0000
	2	0.0799	0.0015	0.0001	0.0020	0.0398	0.0000	0.2141	0.0000
	3	0.0625	0.0000	0.0014	0.0053	0.0290	0.0003	0.2096	0.0000
	4	0.0558	0.0053	0.0057	0.0253	0.0204	0.0148	0.1949	0.0018
	5	0.0505	0.0144	0.0124	0.0518	0.0144	0.0752	0.1993	0.0182
	6	0.0497	0.0295	0.0228	0.0892	0.0099	0.2123	0.1960	0.0515
	7	0.0447	0.0390	0.0322	0.1339	0.0073	0.3756	0.2157	0.1114
	8	0.0461	0.0585	0.0396	0.1779	0.0059	0.4916	0.2443	0.1886
	9	0.0470	0.0810	0.0460	0.2061	0.0055	0.5839	0.2820	0.2780
	10	0.0513	0.1167	0.0545	0.2377	0.0038	0.6634	0.3030	0.3380
	11	0.0585	0.1531	0.0607	0.2741	0.0039	0.7209	0.3259	0.3970
<b>mean</b>		<b>0.0569</b>	<b>0.0454</b>	<b>0.0250</b>	<b>0.1094</b>	<b>0.0172</b>	<b>0.2853</b>	<b>0.2372</b>	<b>0.1259</b>

**Πίνακας Β.3:** Αποτελέσματα συγκρίσεων προσομοίωσης μεταξύ των μεθόδων επιλογής μεταβλητών για τους δύο-επιπέδων «*MW*», Bayesian *D*-βέλτιστους σχεδιασμούς. Η πρώτη στήλη παρουσιάζει το σχεδιασμό, η δεύτερη στήλη αναφέρεται στον αριθμό των πραγματικά ενεργών παραγόντων *d* ενώ οι άλλες στήλες παρουσιάζουν τα σφάλματα που αντιστοιχούν σε κάθε μέθοδο. Οι σειρές που αναφέρονται ως **mean**, παρουσιάζουν τον συνολικό μέσο όρο των αντιστοιχών σφαλμάτων για κάθε σχεδιασμό.

Designs	<i>d</i>	SVR-RFE		Dantzig selector		$L_q$ -penalty		stepwise regression	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
<i>MW1</i>	1	0.1012	0.0000	0.0006	0.0000	0.0635	0.0000	0.1743	0.0000
	2	0.0761	0.0060	0.0178	0.0160	0.0451	0.0070	0.1648	0.0145
	3	0.0670	0.0520	0.0497	0.1030	0.0343	0.1687	0.1659	0.0877
	4	0.0770	0.1450	0.0755	0.2125	0.0305	0.4050	0.1865	0.2380
	5	0.0935	0.2594	0.0942	0.3114	0.0308	0.5610	0.2080	0.4096
	6	0.1165	0.3937	0.1125	0.4228	0.0353	0.6608	0.2220	0.5355
<b>mean</b>		<b>0.0886</b>	<b>0.1427</b>	<b>0.0584</b>	<b>0.1776</b>	<b>0.0399</b>	<b>0.3004</b>	<b>0.1869</b>	<b>0.2142</b>
<i>MW2</i>	1	0.1180	0.0020	0.0002	0.0000	0.0683	0.0000	0.2136	0.0000
	2	0.0938	0.0035	0.0049	0.0010	0.0795	0.0000	0.2080	0.0040
	3	0.0756	0.0110	0.0209	0.0450	0.0346	0.0600	0.1930	0.0190
	4	0.0683	0.0433	0.0461	0.1005	0.0241	0.2330	0.2005	0.0808
	5	0.0697	0.0886	0.0688	0.1788	0.0215	0.4098	0.2134	0.1778
	6	0.0830	0.1578	0.0832	0.2665	0.0218	0.5563	0.2339	0.3088
	7	0.0955	0.2337	0.0948	0.3361	0.0220	0.6371	0.2594	0.4154
<b>mean</b>		<b>0.0863</b>	<b>0.0771</b>	<b>0.0456</b>	<b>0.1326</b>	<b>0.0388</b>	<b>0.2709</b>	<b>0.2174</b>	<b>0.1437</b>
<i>MW3</i>	1	0.1041	0.0000	0.0000	0.0000	0.0762	0.0000	0.2465	0.0000
	2	0.0963	0.0000	0.0002	0.0001	0.0534	0.0000	0.2530	0.0000
	3	0.0759	0.0003	0.0011	0.0063	0.0341	0.0006	0.2397	0.0010
	4	0.0669	0.0018	0.0066	0.0220	0.0157	0.0345	0.2429	0.0022
	5	0.0527	0.0034	0.0163	0.0454	0.0079	0.1438	0.2273	0.0136
	6	0.0438	0.0105	0.0314	0.0798	0.0053	0.3503	0.2249	0.0373
	7	0.0405	0.0229	0.0445	0.1283	0.0040	0.5294	0.2273	0.0851
	8	0.0423	0.0419	0.0562	0.1614	0.0037	0.6509	0.2347	0.1761
	9	0.0492	0.0717	0.0769	0.2152	0.0037	0.7192	0.2398	0.2918
<b>mean</b>		<b>0.0635</b>	<b>0.0169</b>	<b>0.0259</b>	<b>0.0732</b>	<b>0.0227</b>	<b>0.2699</b>	<b>0.2373</b>	<b>0.0674</b>
<i>MW4</i>	1	0.1131	0.0010	0.0004	0.0000	0.0636	0.0000	0.1554	0.0000
	2	0.0865	0.0110	0.0213	0.0060	0.0489	0.0015	0.1415	0.0195
	3	0.0735	0.0620	0.0554	0.0963	0.0371	0.1847	0.156	1 0.1317
	4	0.0857	0.1775	0.0953	0.2353	0.0374	0.4328	0.1945	0.2900
	5	0.1095	0.3158	0.1096	0.3524	0.0358	0.5892	0.2406	0.4534
	6	0.1254	0.4223	0.1184	0.4313	0.0361	0.6633	0.2559	0.5283
<b>mean</b>		<b>0.0990</b>	<b>0.1649</b>	<b>0.0667</b>	<b>0.1869</b>	<b>0.0432</b>	<b>0.3119</b>	<b>0.1907</b>	<b>0.2372</b>
<i>MW4</i>	1	0.1212	0.0000	0.0001	0.0000	0.0706	0.0000	0.1917	0.0000
	2	0.0985	0.0010	0.0055	0.0020	0.0473	0.0000	0.1904	0.0000
	3	0.0785	0.0097	0.0229	0.0353	0.0349	0.0600	0.1835	0.0223
	4	0.0621	0.0295	0.0470	0.0865	0.0242	0.2258	0.1914	0.0858
	5	0.0704	0.0858	0.0697	0.1896	0.0253	0.4136	0.2161	0.1876
	6	0.0804	0.1555	0.0878	0.2625	0.0246	0.5570	0.2674	0.3127
	7	0.0899	0.2190	0.0948	0.3343	0.0243	0.6396	0.3019	0.4169
<b>mean</b>		<b>0.0859</b>	<b>0.0715</b>	<b>0.0468</b>	<b>0.1300</b>	<b>0.0359</b>	<b>0.2709</b>	<b>0.2203</b>	<b>0.1465</b>
<i>MW5</i>	1	0.1026	0.0000	0.0001	0.0000	0.0741	0.0000	0.2510	0.0000
	2	0.1007	0.0000	0.0003	0.0025	0.0555	0.0000	0.1808	0.001
	3	0.0767	0.0003	0.0030	0.0043	0.0328	0.0010	0.1785	0.0247
	4	0.0683	0.0050	0.0125	0.0155	0.0183	0.0493	0.1870	0.0685
	5	0.0615	0.0124	0.0359	0.0540	0.0109	0.2672	0.2262	0.1952
	6	0.0626	0.0305	0.0663	0.1087	0.0107	0.4713	0.2709	0.3233
	7	0.0793	0.0703	0.1039	0.1697	0.0123	0.6109	0.3096	0.4071
	8	0.1074	0.1321	0.1409	0.2494	0.0152	0.7019	0.3281	0.4838
	9	0.1516	0.2199	0.1762	0.3240	0.1946	0.7542	0.3368	0.5053
<b>mean</b>		<b>0.0901</b>	<b>0.0523</b>	<b>0.0599</b>	<b>0.1031</b>	<b>0.0472</b>	<b>0.3173</b>	<b>0.2521</b>	<b>0.2232</b>

**Πίνακας Β.4:** Αποτελέσματα συγκρίσεων προσομοίωσης μεταξύ των μεθόδων επιλογής μεταβλητών για τους σχεδιασμούς μικτών επιπέδων. Η πρώτη στήλη παρουσιάζει τον σχεδιασμό, η δεύτερη στήλη αναφέρεται στον αριθμό των πραγματικά ενεργών παραγόντων  $d$  ενώ οι άλλες στήλες παρουσιάζουν τα σφάλματα που αντιστοιχούν σε κάθε μέθοδο. Οι σειρές που αναφέρονται ως **mean**, παρουσιάζουν τον συνολικό μέσο όρο των αντίστοιχων σφαλμάτων για κάθε σχεδιασμό.

Designs	$d$	SVR-RFE		Dantzig selector		$L_q$ -penalty		stepwise regression	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
$n12k2p22$	1	0.0773	0.0100	0.0006	0.0000	0.0521	0.0000	0.1325	0.0000
	2	0.0592	0.0255	0.0155	0.0580	0.0455	0.0145	0.1231	0.0205
	3	0.0559	0.0640	0.0380	0.1657	0.0403	0.1267	0.1260	0.0970
	4	0.0609	0.1385	0.0565	0.2768	0.0384	0.3320	0.1494	0.2550
	5	0.0600	0.1998	0.0697	0.3606	0.0370	0.5072	0.1827	0.4004
	6	0.0659	0.3020	0.0786	0.4283	0.0373	0.6217	0.2028	0.5048
<b>mean</b>		<b>0.0632</b>	<b>0.1233</b>	<b>0.0431</b>	<b>0.2149</b>	<b>0.0418</b>	<b>0.2670</b>	<b>0.1527</b>	<b>0.2130</b>
$n12k3p33$	1	0.0595	0.0010	0.0006	0.0000	0.0461	0.0000	0.0975	0.0000
	2	0.0485	0.0385	0.0117	0.0275	0.0426	0.0195	0.0924	0.0310
	3	0.0503	0.1210	0.0366	0.1410	0.0413	0.1523	0.0992	0.1447
	4	0.0577	0.2355	0.0580	0.2758	0.0396	0.3338	0.1208	0.3270
	5	0.0636	0.3408	0.0725	0.3724	0.0418	0.4910	0.1495	0.5190
	6	0.0692	0.4410	0.0801	0.4528	0.0387	0.5933	0.1603	0.6252
<b>mean</b>		<b>0.0581</b>	<b>0.1963</b>	<b>0.0432</b>	<b>0.2116</b>	<b>0.0417</b>	<b>0.2650</b>	<b>0.1199</b>	<b>0.2745</b>
$n12k4p44$	1	0.0521	0.0090	0.0008	0.0000	0.0359	0.0000	0.0765	0.0000
	2	0.0425	0.0580	0.0182	0.0705	0.0389	0.0390	0.0735	0.0515
	3	0.0451	0.1687	0.0402	0.2147	0.0402	0.2043	0.0847	0.2127
	4	0.0522	0.2923	0.0553	0.3373	0.0396	0.3790	0.1045	0.4400
	5	0.0584	0.4128	0.0666	0.4328	0.0388	0.5334	0.1234	0.6198
	6	0.0635	0.5235	0.0721	0.5122	0.0380	0.6313	0.1295	0.6747
<b>mean</b>		<b>0.0523</b>	<b>0.244</b>	<b>0.0422</b>	<b>0.2612</b>	<b>0.0386</b>	<b>0.2978</b>	<b>0.0987</b>	<b>0.3331</b>
$n18p8$	1	0.0861	0.0000	0.0001	0.0000	0.1036	0.0000	0.1667	0.0000
	2	0.0768	0.0000	0.0002	0.0165	0.0753	0.0000	0.1756	0.0000
	3	0.0702	0.0000	0.0022	0.0397	0.0519	0.0000	0.1653	0.0000
	4	0.0640	0.0000	0.0045	0.0635	0.0336	0.0043	0.1619	0.0000
	5	0.0589	0.0000	0.0086	0.0822	0.0198	0.0216	0.1555	0.0000
	6	0.0442	0.0000	0.0101	0.1127	0.0086	0.0618	0.1511	0.0017
	7	0.0359	0.0000	0.0119	0.1400	0.0056	0.2967	0.1419	0.0023
	8	0.0219	0.0004	0.0190	0.1619	0.0043	0.4663	0.1459	0.0104
	9	0.0007	0.0007	0.0175	0.1889	0.0020	0.5848	0.1418	0.0262
<b>mean</b>		<b>0.0510</b>	<b>0.0001</b>	<b>0.0082</b>	<b>0.0895</b>	<b>0.0339</b>	<b>0.1595</b>	<b>0.1562</b>	<b>0.0045</b>
$n18k2p34$	1	0.0634	0.0020	0.0003	0.0000	0.0383	0.0000	0.1403	0.0000
	2	0.0607	0.0145	0.0055	0.0525	0.0326	0.0060	0.1325	0.0035
	3	0.0618	0.0540	0.0180	0.1487	0.0290	0.0567	0.1280	0.0140
	4	0.0716	0.1325	0.0403	0.2505	0.0261	0.2668	0.1261	0.1758
	5	0.0839	0.2050	0.0563	0.3140	0.0238	0.4524	0.1413	0.1900
	6	0.0956	0.3212	0.0730	0.3968	0.0232	0.6070	0.1712	0.3302
	7	0.1082	0.4244	0.0858	0.4631	0.0234	0.6989	0.1984	0.4740
	8	0.1150	0.5033	0.0969	0.5138	0.0232	0.7623	0.2137	0.5588
	9	0.1192	0.5600	0.1035	0.5657	0.0242	0.8028	0.2227	0.6084
<b>mean</b>		<b>0.0866</b>	<b>0.2463</b>	<b>0.0533</b>	<b>0.3006</b>	<b>0.0271</b>	<b>0.4059</b>	<b>0.1638</b>	<b>0.2616</b>
$n24k2p46$	1	0.0556	0.0000	0.0000	0.0000	0.0375	0.0000	0.1504	0.0000
	2	0.0555	0.0065	0.0008	0.0195	0.0313	0.0000	0.1396	0.0000
	3	0.0519	0.0237	0.0050	0.0443	0.0277	0.0080	0.1363	0.0060
	4	0.0599	0.0775	0.0140	0.1138	0.0226	0.0675	0.1346	0.0200
	5	0.0708	0.1238	0.0288	0.1870	0.0210	0.2268	0.1369	0.0722
	6	0.0829	0.1997	0.0446	0.2535	0.0206	0.3853	0.1474	0.1457
	7	0.0915	0.2593	0.0556	0.3279	0.0211	0.5180	0.1696	0.2790
	8	0.1017	0.3390	0.0673	0.3901	0.0205	0.6274	0.1902	0.3876
	9	0.1073	0.4087	0.0737	0.4366	0.0212	0.6831	0.2099	0.4967
	10	0.1114	0.4680	0.0805	0.4893	0.0221	0.7362	0.2188	0.5497
	11	0.1174	0.5314	0.0887	0.5350	0.0224	0.7756	0.2232	0.5784
	12	0.1178	0.5607	0.0925	0.5570	0.0225	0.8040	0.2300	0.6188
<b>mean</b>		<b>0.1138</b>	<b>0.3331</b>	<b>0.0613</b>	<b>0.3727</b>	<b>0.0323</b>	<b>0.5369</b>	<b>0.2319</b>	<b>0.3504</b>

---

# Παράρτημα Γ

---

## Παράρτημα Κεφαλαίου 7:

### ΜΕΡΟΣ Ι: Συγκρίσεις ποινών

**Πίνακας Γ.1:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους  $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		Iter-SIS-SCAD				Iter-SIS-Lasso				Iter-SIS-MCP				
$n \times p$	$p'$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.05	0.00	1.00	2.35	0.10	0.00	1.00	3.52	0.08	0.00	1.00	3.00
12 × 26	2	I	0.06	0.00	1.00	3.35	0.17	0.00	1.00	5.96	0.07	0.01	0.99	3.75
12 × 26	3	II	0.08	0.08	0.92	4.69	0.20	0.02	0.98	7.63	0.10	0.08	0.98	4.94
12 × 26	4	III	0.14	0.21	0.79	6.18	0.23	0.11	0.89	8.68	0.14	0.21	0.79	6.33
12 × 26	5	III	0.19	0.34	0.66	7.40	0.26	0.22	0.78	9.42	0.19	0.33	0.67	7.40
12 × 26	6	III	0.24	0.43	0.57	8.18	0.28	0.30	0.70	9.82	0.24	0.42	0.58	8.22
14 × 24	1	I	0.02	0.00	1.00	1.39	-	-	-	-	0.07	0.00	1.00	2.66
14 × 24	2	I	0.03	0.00	1.00	2.61	0.17	0.00	1.00	5.67	0.07	0.00	1.00	3.52
14 × 24	3	II	0.04	0.00	1.00	3.91	0.22	0.00	1.00	7.70	0.08	0.00	1.00	4.57
14 × 24	4	II	0.06	0.02	0.98	5.12	0.25	0.00	1.00	8.95	0.08	0.03	0.97	5.50
14 × 24	5	III	0.11	0.07	0.93	6.65	0.27	0.03	0.97	9.95	0.12	0.08	0.92	6.98
14 × 24	6	III	0.18	0.15	0.85	8.26	0.28	0.07	0.93	10.67	0.18	0.16	0.84	8.36
14 × 24	7	III	0.23	0.21	0.79	9.47	0.30	0.12	0.88	11.27	0.23	0.21	0.79	9.47
18 × 22	1	I	0.00	0.00	1.00	1.00	-	-	-	-	0.00	0.00	1.00	1.00
18 × 22	2	I	0.00	0.00	1.00	2.10	0.11	0.00	1.00	4.27	0.01	0.00	1.00	2.26
18 × 22	3	I	0.01	0.00	1.00	3.15	0.19	0.00	1.00	6.59	0.02	0.00	1.00	3.47
18 × 22	4	II	0.01	0.00	1.00	4.22	0.28	0.00	1.00	9.11	0.03	0.00	1.00	4.59
18 × 22	5	II	0.02	0.00	1.00	5.31	0.36	0.00	1.00	11.04	0.04	0.00	1.00	5.73
18 × 22	6	III	0.06	0.02	0.98	6.77	0.42	0.00	1.00	12.60	0.08	0.02	0.98	7.18
18 × 22	7	III	0.11	0.04	0.96	8.37	0.46	0.02	0.98	13.79	0.14	0.05	0.95	8.76
18 × 22	8	III	0.23	0.10	0.90	10.47	0.50	0.04	0.96	14.64	0.25	0.10	0.90	10.67
18 × 22	9	III	0.33	0.14	0.86	12.03	0.54	0.07	0.93	15.36	0.34	0.14	0.86	12.13

**Πίνακας Γ.2:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους σχεδιασμούς των Koukouvinos et al. [76].

SSD		Iter-SIS-SCAD					Iter-SIS-Lasso				Iter-SIS-MCP			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
8 × 14	1	I	0.20	0.00	1.00	3.65	0.29	0.00	1.00	4.76	0.18	0.00	1.00	3.30
8 × 14	2	II	0.17	0.03	0.97	3.97	0.27	0.01	0.99	5.23	0.17	0.04	0.96	3.92
8 × 14	3	III	0.17	0.15	0.85	4.40	0.26	0.05	0.95	5.71	0.17	0.15	0.85	4.42
8 × 14	4	III	0.24	0.33	0.67	5.08	0.28	0.13	0.87	6.27	0.24	0.34	0.66	5.08
10 × 18	1	I	0.14	0.00	1.00	3.40	-	-	-	-	0.11	0.00	1.00	2.92
10 × 18	2	I	0.12	0.00	1.00	3.89	0.24	0.00	1.00	5.85	0.12	0.00	1.00	3.94
10 × 18	3	II	0.13	0.03	0.97	4.80	0.25	0.01	0.99	6.75	0.13	0.04	0.96	4.79
10 × 18	4	III	0.16	0.13	0.87	5.75	0.24	0.03	0.97	7.28	0.16	0.13	0.87	5.71
10 × 18	5	III	0.21	0.23	0.77	6.60	0.25	0.09	0.91	7.87	0.21	0.23	0.77	6.63
12 × 22	1	I	0.03	0.00	1.00	1.71	-	-	-	-	0.07	0.00	1.00	2.45
12 × 22	2	I	0.06	0.00	1.00	3.13	0.19	0.00	1.00	5.89	0.08	0.00	1.00	3.66
12 × 22	3	II	0.06	0.00	1.00	4.09	0.23	0.00	1.00	7.27	0.08	0.01	0.99	4.48
12 × 22	4	III	0.09	0.04	0.96	5.40	0.23	0.01	0.99	8.11	0.10	0.05	0.95	5.64
12 × 22	5	III	0.14	0.11	0.89	6.84	0.24	0.05	0.95	8.82	0.15	0.13	0.87	6.98
12 × 22	6	III	0.20	0.19	0.81	8.02	0.24	0.08	0.92	9.39	0.20	0.19	0.81	8.00
14 × 26	1	I	0.06	0.00	1.00	2.56	-	-	-	-	0.10	0.00	1.00	3.60
14 × 26	2	I	0.06	0.00	1.00	3.40	0.17	0.00	1.00	6.11	0.09	0.00	1.00	4.25
14 × 26	3	II	0.06	0.00	1.00	4.44	0.20	0.00	1.00	7.70	0.09	0.00	1.00	5.11
14 × 26	4	II	0.07	0.03	0.97	5.54	0.22	0.00	1.00	8.80	0.09	0.03	0.97	5.94
14 × 26	5	III	0.11	0.08	0.92	6.96	0.22	0.02	0.98	9.58	0.12	0.09	0.91	7.16
14 × 26	6	III	0.16	0.16	0.84	8.28	0.23	0.04	0.96	10.41	0.17	0.16	0.84	8.36
14 × 26	7	III	0.22	0.24	0.76	9.45	0.24	0.08	0.92	11.01	0.21	0.23	0.77	9.48
16 × 30	1	I	0.00	0.00	1.00	1.20	-	-	-	-	0.05	0.00	1.00	2.35
16 × 30	2	I	0.02	0.00	1.00	2.45	0.12	0.00	1.00	5.40	0.06	0.00	1.00	3.57
16 × 30	3	I	0.02	0.00	1.00	3.56	0.18	0.00	1.00	7.88	0.06	0.00	1.00	4.52
16 × 30	4	II	0.03	0.01	0.99	4.84	0.21	0.00	1.00	9.57	0.07	0.01	0.99	5.66
16 × 30	5	III	0.05	0.02	0.98	6.15	0.23	0.01	0.99	10.68	0.08	0.03	0.97	6.88
16 × 30	6	III	0.09	0.07	0.93	7.70	0.23	0.02	0.98	11.39	0.11	0.07	0.93	8.24
16 × 30	7	III	0.14	0.13	0.87	9.41	0.24	0.04	0.96	12.20	0.15	0.13	0.87	9.61
16 × 30	8	III	0.20	0.19	0.81	10.89	0.25	0.07	0.93	12.90	0.20	0.19	0.81	10.88
18 × 34	1	I	0.00	0.00	1.00	1.00	-	-	-	-	0.01	0.00	1.00	1.48
18 × 34	2	I	0.00	0.00	1.00	2.06	0.07	0.00	1.00	4.22	0.03	0.00	1.00	3.01
18 × 34	3	I	0.01	0.00	1.00	3.17	0.17	0.00	1.00	8.12	0.04	0.00	1.00	4.20
18 × 34	4	II	0.01	0.00	1.00	4.38	0.23	0.00	1.00	10.75	0.05	0.00	1.00	5.44
18 × 34	5	II	0.03	0.01	0.99	5.68	0.24	0.00	1.00	11.97	0.06	0.02	0.98	6.60
18 × 34	6	III	0.05	0.03	0.97	7.11	0.24	0.00	1.00	12.71	0.08	0.03	0.97	7.97
18 × 34	7	III	0.09	0.07	0.93	8.89	0.24	0.02	0.98	13.31	0.11	0.08	0.92	9.51
18 × 34	8	III	0.14	0.12	0.88	10.52	0.25	0.04	0.96	14.09	0.15	0.12	0.88	10.92
18 × 34	9	III	0.19	0.17	0.83	12.11	0.25	0.06	0.94	14.69	0.19	0.17	0.83	12.25
20 × 38	1	I	0.01	0.00	1.00	1.27	-	-	-	-	0.06	0.00	1.00	3.09
20 × 38	2	I	0.01	0.00	1.00	2.31	0.07	0.00	1.00	4.56	0.05	0.00	1.00	3.75
20 × 38	3	I	0.01	0.00	1.00	3.43	0.10	0.00	1.00	6.47	0.04	0.00	1.00	4.53
20 × 38	4	II	0.01	0.00	1.00	4.50	0.12	0.00	1.00	8.22	0.04	0.00	1.00	5.34
20 × 38	5	II	0.02	0.00	1.00	5.59	0.15	0.00	1.00	9.83	0.04	0.01	0.99	6.33
20 × 38	6	III	0.02	0.01	0.99	6.73	0.17	0.00	1.00	11.37	0.04	0.01	0.99	7.35
20 × 38	7	III	0.04	0.03	0.97	8.13	0.18	0.01	0.99	12.62	0.06	0.04	0.96	8.72
20 × 38	8	III	0.08	0.07	0.93	9.76	0.20	0.02	0.98	13.78	0.09	0.08	0.92	10.19
20 × 38	9	III	0.12	0.12	0.88	11.51	0.21	0.03	0.97	14.71	0.13	0.12	0.88	11.70
20 × 38	10	III	0.17	0.17	0.83	13.17	0.23	0.05	0.95	15.83	0.18	0.17	0.83	13.32
22 × 42	1	I	0.00	0.00	1.00	1.00	-	-	-	-	0.00	0.00	1.00	1.02
22 × 42	2	I	0.00	0.00	1.00	2.00	0.03	0.00	1.00	3.15	0.00	0.00	1.00	2.02
22 × 42	3	I	0.00	0.00	1.00	3.01	0.06	0.00	1.00	5.16	0.00	0.00	1.00	3.05
22 × 42	4	I	0.00	0.00	1.00	4.01	0.09	0.00	1.00	7.28	0.00	0.00	1.00	4.09
22 × 42	5	II	0.00	0.00	1.00	5.00	0.12	0.00	1.00	9.32	0.01	0.00	1.00	5.17
22 × 42	6	II	0.01	0.01	0.99	6.17	0.15	0.00	1.00	11.34	0.01	0.01	0.99	6.36
22 × 42	7	III	0.02	0.02	0.98	7.44	0.17	0.00	1.00	12.89	0.02	0.02	0.98	7.70
22 × 42	8	III	0.04	0.04	0.96	8.99	0.19	0.01	0.99	14.24	0.05	0.04	0.96	9.24
22 × 42	9	III	0.08	0.07	0.93	10.85	0.20	0.02	0.98	15.50	0.08	0.08	0.92	11.07
22 × 42	10	III	0.11	0.11	0.89	12.52	0.21	0.03	0.97	16.38	0.12	0.11	0.89	12.72
22 × 42	11	III	0.16	0.16	0.84	14.26	0.23	0.05	0.95	17.40	0.17	0.16	0.84	14.44

**Πίνακας Γ.3:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους  $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		Iter-SIS-SCAD					Iter-SIS-Lasso				Iter-SIS-MCP			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.04	0.00	1.00	1.89	-	-	-	-	0.08	0.00	1.00	2.99
12 × 26	2	I	0.04	0.00	1.00	2.91	0.17	0.00	1.00	6.07	0.07	0.00	1.00	3.75
12 × 26	3	II	0.06	0.05	0.95	4.35	0.20	0.01	0.99	7.52	0.10	0.08	0.92	4.94
12 × 26	4	III	0.12	0.17	0.83	5.98	0.22	0.10	0.90	8.46	0.14	0.21	0.79	6.33
12 × 26	5	III	0.17	0.28	0.72	7.25	0.24	0.18	0.82	9.11	0.19	0.33	0.33	7.40
12 × 26	6	III	0.23	0.39	0.61	8.26	0.26	0.27	0.73	9.60	0.24	0.42	0.58	8.22
14 × 24	1	I	0.00	0.00	1.00	1.00	-	-	-	-	0.03	0.00	1.00	1.59
14 × 24	2	I	0.01	0.00	1.00	2.20	0.16	0.00	1.00	5.58	0.05	0.00	1.00	3.05
14 × 24	3	II	0.02	0.00	1.00	3.41	0.23	0.00	1.00	7.79	0.06	0.00	1.00	4.24
14 × 24	4	II	0.05	0.03	0.97	4.95	0.25	0.01	0.99	9.05	0.10	0.04	0.96	5.64
14 × 24	5	III	0.11	0.09	0.91	6.60	0.27	0.04	0.96	9.97	0.13	0.09	0.91	7.02
14 × 24	6	III	0.19	0.16	0.84	8.40	0.29	0.08	0.92	10.66	0.20	0.17	0.83	8.55
14 × 24	7	III	0.24	0.23	0.77	9.53	0.30	0.13	0.87	11.17	0.24	0.23	0.77	9.55
18 × 22	1	I	0.01	0.00	1.00	1.13	-	-	-	-	0.08	0.00	1.00	2.63
18 × 22	2	I	0.02	0.00	1.00	2.31	0.14	0.00	1.00	4.73	0.07	0.00	1.00	3.49
18 × 22	3	I	0.03	0.00	1.00	3.50	0.20	0.00	1.00	6.88	0.08	0.00	1.00	4.47
18 × 22	4	II	0.03	0.00	1.00	4.63	0.26	0.00	1.00	8.67	0.08	0.00	1.00	5.41
18 × 22	5	II	0.04	0.00	1.00	5.68	0.32	0.00	1.00	10.37	0.08	0.00	1.00	6.33
18 × 22	6	III	0.06	0.01	0.99	6.95	0.36	0.00	1.00	11.77	0.09	0.01	0.99	7.37
18 × 22	7	III	0.09	0.02	0.98	8.27	0.39	0.00	1.00	12.87	0.11	0.02	0.98	8.57
18 × 22	8	III	0.15	0.04	0.96	9.74	0.42	0.01	0.99	13.77	0.16	0.04	0.96	9.95
18 × 22	9	III	0.21	0.06	0.96	11.11	0.44	0.02	0.98	14.52	0.22	0.06	0.94	11.22

**Πίνακας Γ.4:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους σχεδιασμούς των Koukouninos et al. [76].

SSD		SIS-SCAD					SIS-Lasso				SIS-MCP			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
8 × 14	1	I	0.15	0.00	1.00	3.00	0.25	0.00	1.00	4.27	0.12	0.00	1.00	2.54
8 × 14	2	II	0.14	0.03	0.97	3.59	0.23	0.02	0.98	4.66	0.13	0.03	0.97	3.45
8 × 14	3	III	0.17	0.16	0.84	4.42	0.25	0.07	0.93	5.55	0.17	0.15	0.85	4.45
8 × 14	4	III	0.19	0.22	0.78	5.05	0.25	0.10	0.90	6.17	0.19	0.21	0.79	5.02
10 × 18	1	I	0.09	0.00	1.00	2.52	-	-	-	-	0.11	0.00	1.00	2.87
10 × 18	2	I	0.09	0.00	1.00	3.41	0.21	0.00	1.00	5.36	0.09	0.02	0.98	3.48
10 × 18	3	II	0.11	0.03	0.97	4.57	0.22	0.00	1.00	6.22	0.11	0.03	0.97	4.50
10 × 18	4	III	0.13	0.07	0.93	5.50	0.23	0.03	0.97	7.03	0.13	0.07	0.93	5.49
10 × 18	5	III	0.17	0.14	0.86	6.51	0.24	0.07	0.93	7.71	0.17	0.14	0.86	6.51
12 × 22	1	I	0.05	0.00	1.00	2.06	0.13	0.00	1.00	3.72	0.07	0.00	1.00	2.37
12 × 22	2	I	0.03	0.00	1.00	2.66	0.17	0.00	1.00	5.34	0.05	0.00	1.00	2.93
12 × 22	3	II	0.05	0.02	0.98	3.92	0.21	0.01	0.99	6.92	0.07	0.03	0.97	4.19
12 × 22	4	III	0.08	0.04	0.96	5.24	0.22	0.02	0.98	7.81	0.09	0.04	0.96	5.40
12 × 22	5	III	0.12	0.08	0.92	6.70	0.22	0.05	0.95	8.57	0.13	0.08	0.92	6.73
12 × 22	6	III	0.16	0.12	0.88	7.87	0.24	0.08	0.92	9.35	0.17	0.13	0.87	7.88
14 × 26	1	I	0.04	0.00	1.00	2.11	-	-	-	-	0.06	0.00	1.00	2.59
14 × 26	2	I	0.03	0.00	1.00	2.69	0.14	0.00	1.00	5.27	0.04	0.00	1.00	3.05
14 × 26	3	II	0.04	0.00	1.00	4.12	0.17	0.00	1.00	6.90	0.06	0.00	1.00	4.38
14 × 26	4	II	0.05	0.01	0.99	5.15	0.19	0.00	1.00	8.22	0.06	0.02	0.98	5.36
14 × 26	5	III	0.08	0.04	0.96	6.44	0.20	0.02	0.98	9.18	0.09	0.05	0.95	6.58
14 × 26	6	III	0.12	0.07	0.93	7.93	0.21	0.04	0.96	10.05	0.12	0.07	0.93	8.03
14 × 26	7	III	0.15	0.10	0.90	9.14	0.23	0.07	0.93	10.82	0.16	0.10	0.90	9.16
16 × 30	1	I	0.01	0.00	1.00	1.19	-	-	-	-	0.02	0.00	1.00	1.64
16 × 30	2	I	0.02	0.00	1.00	2.45	0.11	0.00	1.00	5.16	0.04	0.00	1.00	3.03
16 × 30	3	I	0.02	0.00	1.00	3.53	0.16	0.00	1.00	7.42	0.04	0.00	1.00	4.05
16 × 30	4	II	0.03	0.01	0.99	4.84	0.19	0.00	1.00	8.83	0.05	0.01	0.99	5.25
16 × 30	5	III	0.06	0.02	0.98	6.30	0.20	0.02	0.98	9.95	0.07	0.03	0.97	6.65
16 × 30	6	III	0.08	0.04	0.96	7.70	0.21	0.03	0.97	10.96	0.09	0.05	0.95	7.91
16 × 30	7	III	0.13	0.08	0.92	9.35	0.23	0.05	0.95	11.93	0.13	0.08	0.92	9.47
16 × 30	8	III	0.16	0.10	0.90	10.66	0.23	0.07	0.93	12.61	0.16	0.10	0.90	10.68

**Πίνακας Γ.5:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους σχεδιασμούς των Koukouninos et al. [76].

SSD		SIS-SCAD				SIS-Lasso				SIS-MCP				
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
18 × 34	1	I	0.00	0.00	1.00	1.12	-	-	-	-	0.01	0.00	1.00	1.35
18 × 34	2	I	0.00	0.00	1.00	2.14	0.09	0.00	1.00	4.92	0.01	0.00	1.00	2.46
18 × 34	3	I	0.01	0.00	1.00	3.28	0.16	0.00	1.00	7.91	0.02	0.00	1.00	3.75
18 × 34	4	II	0.14	0.00	1.00	4.40	0.19	0.00	1.00	9.82	0.03	0.00	1.00	4.96
18 × 34	5	II	0.03	0.01	0.99	5.73	0.21	0.01	0.99	11.13	0.04	0.01	0.99	6.12
18 × 34	6	III	0.06	0.03	0.97	7.44	0.22	0.02	0.98	11.93	0.07	0.03	0.97	7.72
18 × 34	7	III	0.09	0.05	0.95	8.98	0.22	0.03	0.97	12.77	0.10	0.06	0.94	9.25
18 × 34	8	III	0.13	0.07	0.93	10.68	0.23	0.05	0.95	13.70	0.13	0.07	0.93	10.81
18 × 34	9	III	0.16	0.09	0.91	12.09	0.24	0.07	0.93	14.44	0.16	0.09	0.91	12.19
20 × 38	1	I	0.00	0.00	1.00	1.16	-	-	-	-	0.02	0.00	1.00	1.62
20 × 38	2	I	0.01	0.00	1.00	2.33	0.06	0.00	1.00	4.01	0.02	0.00	1.00	2.59
20 × 38	3	I	0.01	0.00	1.00	3.34	0.08	0.00	1.00	5.93	0.02	0.00	1.00	3.64
20 × 38	4	II	0.01	0.00	1.00	4.34	0.11	0.00	1.00	7.63	0.02	0.00	1.00	4.66
20 × 38	5	II	0.01	0.01	0.99	5.45	0.13	0.00	1.00	9.15	0.02	0.01	0.99	5.73
20 × 38	6	III	0.03	0.02	0.98	6.93	0.15	0.01	0.99	10.84	0.04	0.02	0.98	7.18
20 × 38	7	III	0.06	0.03	0.97	8.48	0.17	0.02	0.98	12.28	0.06	0.03	0.97	8.67
20 × 38	8	III	0.09	0.05	0.95	10.23	0.19	0.03	0.97	13.56	0.09	0.05	0.95	10.36
20 × 38	9	III	0.12	0.07	0.93	11.89	0.21	0.04	0.96	14.63	0.13	0.07	0.93	11.99
20 × 38	10	III	0.16	0.10	0.90	13.42	0.22	0.06	0.94	15.64	0.16	0.10	0.90	13.54
22 × 42	1	I	0.00	0.00	1.00	1.00	-	-	-	-	0.00	0.00	1.00	1.02
22 × 42	2	I	0.00	0.00	1.00	2.01	0.03	0.00	1.00	3.32	0.00	0.00	1.00	2.06
22 × 42	3	I	0.00	0.00	1.00	3.01	0.06	0.00	1.00	5.18	0.00	0.00	1.00	3.07
22 × 42	4	I	0.00	0.00	1.00	4.03	0.08	0.00	1.00	7.22	0.00	0.00	1.00	4.10
22 × 42	5	II	0.00	0.00	1.00	5.16	0.11	0.00	1.00	9.04	0.00	0.00	1.00	5.27
22 × 42	6	II	0.02	0.01	0.99	6.53	0.14	0.01	0.99	10.88	0.02	0.01	0.99	6.66
22 × 42	7	III	0.04	0.02	0.98	8.12	0.16	0.02	0.98	12.46	0.04	0.03	0.97	8.29
22 × 42	8	III	0.06	0.04	0.96	9.77	0.18	0.02	0.98	14.06	0.06	0.04	0.96	9.89
22 × 42	9	III	0.09	0.06	0.94	11.59	0.20	0.03	0.97	15.29	0.10	0.06	0.94	11.72
22 × 42	10	III	0.12	0.07	0.93	13.14	0.21	0.05	0.95	16.23	0.13	0.07	0.93	13.26
22 × 42	11	III	0.15	0.09	0.91	14.62	0.22	0.06	0.94	17.26	0.15	0.09	0.91	14.71

**Πίνακας Γ.6:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους  $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		SIS-SCAD				SIS-Lasso				SIS-MCP				
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.03	0.00	1.00	1.80	-	-	-	-	0.05	0.00	1.00	2.22
12 × 26	2	I	0.04	0.01	0.99	2.87	0.16	0.00	1.00	5.84	0.05	0.01	0.99	3.24
12 × 26	3	II	0.08	0.07	0.93	4.53	0.20	0.04	0.96	7.40	0.09	0.08	0.92	4.75
12 × 26	4	III	0.14	0.17	0.83	6.36	0.23	0.13	0.87	8.49	0.14	0.18	0.82	6.44
12 × 26	5	III	0.20	0.26	0.74	7.90	0.26	0.20	0.80	9.40	0.20	0.26	0.74	7.92
12 × 26	6	III	0.23	0.32	0.68	8.67	0.27	0.26	0.74	9.82	0.23	0.32	0.68	8.68
14 × 24	1	I	0.02	0.00	1.00	1.35	-	-	-	-	0.04	0.00	1.00	1.97
14 × 24	2	I	0.03	0.00	1.00	2.58	0.15	0.00	1.00	5.36	0.04	0.00	1.00	2.95
14 × 24	3	II	0.04	0.00	1.00	3.88	0.20	0.00	1.00	7.19	0.06	0.01	0.99	4.24
14 × 24	4	II	0.06	0.03	0.97	5.12	0.22	0.01	0.99	8.44	0.07	0.03	0.97	5.41
14 × 24	5	III	0.10	0.06	0.94	6.57	0.25	0.03	0.97	9.56	0.11	0.06	0.94	6.71
14 × 24	6	III	0.15	0.09	0.91	8.14	0.27	0.06	0.94	10.42	0.15	0.10	0.90	8.20
14 × 24	7	III	0.20	0.13	0.87	9.46	0.28	0.09	0.91	11.12	0.20	0.13	0.87	9.46
18 × 22	1	I	0.00	0.00	1.00	1.00	-	-	-	-	0.03	0.00	1.00	1.67
18 × 22	2	I	0.00	0.00	1.00	2.06	0.12	0.00	1.00	4.43	0.02	0.00	1.00	2.43
18 × 22	3	I	0.01	0.00	1.00	3.14	0.20	0.00	1.00	6.87	0.03	0.00	1.00	3.50
18 × 22	4	II	0.01	0.00	1.00	4.19	0.27	0.00	1.00	8.91	0.03	0.00	1.00	4.56
18 × 22	5	II	0.02	0.00	1.00	5.30	0.33	0.00	1.00	10.67	0.04	0.00	1.00	5.70
18 × 22	6	III	0.06	0.02	0.98	6.83	0.37	0.01	0.99	11.92	0.08	0.02	0.98	7.11
18 × 22	7	III	0.13	0.04	0.96	8.70	0.43	0.02	0.98	13.25	0.15	0.04	0.96	8.93
18 × 22	8	III	0.22	0.07	0.93	10.53	0.47	0.04	0.96	14.30	0.23	0.07	0.93	10.64
18 × 22	9	III	0.31	0.10	0.90	12.14	0.52	0.07	0.93	15.10	0.32	0.10	0.90	12.20



**Πίνακας Γ.7:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο AIC για τους  $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		SIS-SCAD					SIS-Lasso				SIS-MCP			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.02	0.00	1.00	1.61	-	-	-	-	0.04	0.00	1.00	2.01
12 × 26	2	I	0.04	0.00	1.00	2.93	0.14	0.00	1.00	5.46	0.05	0.00	1.00	3.13
12 × 26	3	II	0.06	0.05	0.95	4.20	0.18	0.03	0.97	7.12	0.07	0.06	0.94	4.42
12 × 26	4	III	0.12	0.14	0.86	6.16	0.22	0.10	0.90	8.34	0.13	0.14	0.86	6.25
12 × 26	5	III	0.17	0.21	0.79	7.58	0.24	0.16	0.84	9.24	0.18	0.21	0.79	7.64
12 × 26	6	III	0.21	0.27	0.73	8.60	0.25	0.23	0.77	9.74	0.21	0.27	0.73	8.60
14 × 24	1	I	0.02	0.00	1.00	1.46	-	-	-	-	0.05	0.00	1.00	2.21
14 × 24	2	I	0.02	0.00	1.00	2.55	0.16	0.00	1.00	5.56	0.05	0.00	1.00	3.15
14 × 24	3	II	0.04	0.01	0.99	3.73	0.18	0.00	1.00	6.76	0.06	0.01	0.99	4.13
14 × 24	4	II	0.06	0.03	0.97	5.09	0.22	0.01	0.99	8.29	0.08	0.03	0.97	5.44
14 × 24	5	III	0.11	0.06	0.94	6.74	0.24	0.04	0.96	9.28	0.12	0.07	0.93	6.89
14 × 24	6	III	0.17	0.11	0.89	8.29	0.27	0.08	0.92	10.35	0.17	0.12	0.88	8.36
14 × 24	7	III	0.20	0.14	0.86	9.48	0.27	0.11	0.89	10.90	0.20	0.14	0.86	9.48
18 × 22	1	I	0.03	0.00	1.00	1.58	-	-	-	-	0.06	0.00	1.00	2.30
18 × 22	2	I	0.02	0.00	1.00	2.46	0.08	0.00	1.00	2.63	0.05	0.00	1.00	3.15
18 × 22	3	I	0.04	0.00	1.00	3.73	0.20	0.00	1.00	6.71	0.06	0.01	0.99	4.13
18 × 22	4	II	0.04	0.00	1.00	4.68	0.24	0.00	1.00	8.40	0.08	0.03	0.97	5.44
18 × 22	5	II	0.05	0.00	1.00	5.85	0.28	0.00	1.00	9.69	0.12	0.07	0.93	6.89
18 × 22	6	III	0.06	0.01	0.99	6.92	0.32	0.00	1.00	11.12	0.17	0.12	0.88	8.36
18 × 22	7	III	0.08	0.02	0.98	8.13	0.34	0.01	0.99	12.03	0.20	0.14	0.86	9.48
18 × 22	8	III	0.12	0.02	0.98	9.53	0.37	0.01	0.99	13.11	0.25	0.18	0.82	10.50
18 × 22	9	III	0.17	0.04	0.96	10.88	0.38	0.02	0.98	13.80	0.26	0.21	0.79	11.02

**Πίνακας Γ.8:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους σχεδιασμούς των Koukouninos et al. [76].

SSD		Iter-SIS-SCAD					Iter-SIS-Lasso				Iter-SIS-MCP			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
8 × 14	1	I	0.20	0.00	1.00	3.61	0.30	0.00	1.00	6.50	0.18	0.00	1.00	3.30
8 × 14	2	II	0.17	0.03	0.97	3.93	0.27	0.01	0.99	5.20	0.16	0.04	0.96	3.87
8 × 14	3	III	0.17	0.15	0.85	4.42	0.26	0.05	0.95	5.71	0.17	0.15	0.85	4.38
8 × 14	4	III	0.24	0.33	0.77	5.05	0.28	0.13	0.87	6.27	0.24	0.34	0.66	5.04
10 × 18	1	I	0.12	0.00	1.00	3.06	0.26	0.00	1.00	8.00	0.11	0.00	1.00	2.92
10 × 18	2	I	0.09	0.00	1.00	3.49	0.23	0.00	1.00	5.64	0.10	0.00	1.00	3.66
10 × 18	3	II	0.10	0.04	0.96	4.44	0.24	0.01	0.99	6.61	0.11	0.05	0.95	4.58
10 × 18	4	III	0.15	0.13	0.87	5.56	0.24	0.03	0.97	7.18	0.15	0.13	0.87	5.58
10 × 18	5	III	0.21	0.23	0.77	6.54	0.25	0.09	0.91	7.82	0.20	0.23	0.77	6.53
12 × 22	1	I	0.02	0.00	1.00	1.50	0.16	0.00	1.00	5.10	0.05	0.00	1.00	2.03
12 × 22	2	I	0.02	0.00	1.00	2.44	0.16	0.00	1.00	5.11	0.04	0.00	1.00	2.82
12 × 22	3	II	0.02	0.00	1.00	3.39	0.20	0.00	1.00	6.75	0.04	0.01	0.99	3.81
12 × 22	4	III	0.05	0.05	0.95	4.80	0.22	0.01	0.99	7.85	0.07	0.05	0.95	5.12
12 × 22	5	III	0.12	0.12	0.88	6.42	0.23	0.05	0.95	8.68	0.13	0.13	0.87	6.63
12 × 22	6	III	0.19	0.20	0.80	7.79	0.24	0.08	0.92	9.31	0.19	0.20	0.80	7.78
14 × 26	1	I	0.01	0.00	1.00	1.24	0.10	0.00	1.00	4.40	0.02	0.00	1.00	1.54
14 × 26	2	I	0.01	0.00	1.00	2.34	0.11	0.00	1.00	4.57	0.04	0.00	1.00	2.91
14 × 26	3	II	0.02	0.00	1.00	3.45	0.16	0.00	1.00	6.77	0.04	0.00	1.00	4.04
14 × 26	4	II	0.04	0.03	0.97	4.72	0.20	0.00	1.00	8.29	0.06	0.03	0.97	5.15
14 × 26	5	III	0.08	0.08	0.92	6.20	0.21	0.01	0.99	9.36	0.10	0.10	0.90	6.59
14 × 26	6	III	0.14	0.16	0.84	7.80	0.23	0.04	0.96	10.26	0.14	0.16	0.84	7.89
14 × 26	7	III	0.20	0.24	0.76	9.13	0.24	0.08	0.92	10.90	0.20	0.23	0.77	9.19
16 × 30	1	I	0.00	0.00	1.00	1.00	0.07	0.00	1.00	4.00	0.01	0.00	1.00	1.16
16 × 30	2	I	0.00	0.00	1.00	2.05	0.07	0.00	1.00	3.99	0.01	0.00	1.00	2.41
16 × 30	3	I	0.00	0.00	1.00	3.07	0.12	0.00	1.00	6.34	0.02	0.00	1.00	3.49
16 × 30	4	II	0.01	0.01	0.99	4.22	0.18	0.00	1.00	8.55	0.02	0.01	0.99	4.58
16 × 30	5	III	0.03	0.03	0.97	5.50	0.20	0.01	0.99	10.07	0.04	0.03	0.97	5.88
16 × 30	6	III	0.06	0.07	0.93	7.04	0.22	0.02	0.98	11.07	0.08	0.07	0.93	7.39
16 × 30	7	III	0.12	0.14	0.86	8.79	0.23	0.04	0.96	12.00	0.13	0.14	0.86	9.06
16 × 30	8	III	0.19	0.20	0.80	10.47	0.24	0.08	0.92	12.77	0.18	0.20	0.80	10.43

**Πίνακας Γ.9:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους σχεδιασμούς των Koukouninos et al. [76].

SSD		Iter-SIS-SCAD				Iter-SIS-Lasso				Iter-SIS-MCP				
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
18 × 34	1	I	0.00	0.00	1.00	1.00	0.03	0.00	1.00	2.90	0.00	0.00	1.00	1.00
18 × 34	2	I	0.00	0.00	1.00	2.00	0.03	0.00	1.00	2.90	0.00	0.00	1.00	2.01
18 × 34	3	I	0.00	0.00	1.00	3.00	0.08	0.00	1.00	5.51	0.00	0.00	1.00	3.04
18 × 34	4	II	0.00	0.00	1.00	4.07	0.15	0.00	1.00	8.39	0.01	0.00	1.00	4.16
18 × 34	5	II	0.01	0.01	0.99	5.20	0.19	0.00	1.00	10.56	0.02	0.02	0.98	5.37
18 × 34	6	III	0.03	0.03	0.97	6.58	0.21	0.01	0.99	11.95	0.04	0.04	0.96	6.83
18 × 34	7	III	0.07	0.08	0.92	8.30	0.22	0.02	0.98	12.90	0.08	0.08	0.92	8.50
18 × 34	8	III	0.11	0.12	0.88	9.85	0.23	0.03	0.97	13.83	0.12	0.13	0.87	10.14
18 × 34	9	III	0.17	0.19	0.81	11.59	0.25	0.07	0.93	14.53	0.17	0.18	0.82	11.57
20 × 38	1	I	0.00	0.00	1.00	1.02	0.04	0.00	1.00	3.37	0.01	0.00	1.00	1.41
20 × 38	2	I	0.00	0.00	1.00	2.03	0.04	0.00	1.00	3.37	0.02	0.00	1.00	2.56
20 × 38	3	I	0.00	0.00	1.00	3.07	0.06	0.00	1.00	5.10	0.02	0.00	1.00	3.59
20 × 38	4	II	0.00	0.00	1.00	4.09	0.08	0.00	1.00	6.88	0.02	0.00	1.00	4.58
20 × 38	5	II	0.01	0.00	1.00	5.18	0.11	0.00	1.00	8.67	0.02	0.01	0.99	5.62
20 × 38	6	III	0.01	0.01	0.99	6.33	0.14	0.00	1.00	10.41	0.02	0.01	0.99	6.63
20 × 38	7	III	0.03	0.04	0.96	7.75	0.16	0.01	0.99	12.00	0.04	0.04	0.96	8.04
20 × 38	8	III	0.07	0.08	0.92	9.35	0.18	0.02	0.98	13.26	0.08	0.08	0.92	9.57
20 × 38	9	III	0.11	0.13	0.87	11.05	0.20	0.03	0.97	14.45	0.11	0.13	0.87	11.10
20 × 38	10	III	0.16	0.18	0.82	12.71	0.22	0.05	0.95	15.64	0.17	0.19	0.81	12.81
22 × 42	1	I	0.00	0.00	1.00	1.00	0.01	0.00	1.00	2.50	0.00	0.00	1.00	1.00
22 × 42	2	I	0.00	0.00	1.00	2.00	0.01	0.00	1.00	2.47	0.00	0.00	1.00	2.00
22 × 42	3	I	0.00	0.00	1.00	3.00	0.03	0.00	1.00	4.20	0.00	0.00	1.00	3.01
22 × 42	4	I	0.00	0.00	1.00	4.00	0.05	0.00	1.00	5.98	0.00	0.00	1.00	4.02
22 × 42	5	II	0.00	0.00	1.00	5.01	0.08	0.00	1.00	7.90	0.00	0.00	1.00	5.05
22 × 42	6	II	0.01	0.01	0.99	6.16	0.11	0.00	1.00	10.02	0.01	0.01	0.99	6.19
22 × 42	7	III	0.01	0.02	0.98	7.38	0.14	0.00	1.00	11.80	0.02	0.02	0.98	7.50
22 × 42	8	III	0.04	0.05	0.95	8.88	0.16	0.01	0.99	13.51	0.04	0.05	0.95	8.96
22 × 42	9	III	0.07	0.08	0.92	10.63	0.19	0.02	0.98	14.94	0.08	0.09	0.91	10.70
22 × 42	10	III	0.11	0.12	0.88	12.23	0.20	0.03	0.97	16.05	0.11	0.12	0.88	12.32
22 × 42	11	III	0.15	0.17	0.83	13.82	0.22	0.05	0.95	17.15	0.15	0.18	0.82	13.84

**Πίνακας Γ.10:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους  $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		Iter-SIS-SCAD				Iter-SIS-Lasso				Iter-SIS-MCP				
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.01	0.00	1.00	1.19	0.15	0.00	1.00	5.50	0.05	0.00	1.00	2.22
12 × 26	2	I	0.00	0.00	1.00	2.00	0.14	0.00	1.00	5.43	0.05	0.00	1.00	3.18
12 × 26	3	II	0.05	0.08	0.92	3.92	0.19	0.02	0.98	7.28	0.08	0.09	0.91	4.50
12 × 26	4	III	0.12	0.22	0.78	5.74	0.23	0.11	0.89	8.53	0.13	0.21	0.79	6.01
12 × 26	5	III	0.18	0.35	0.65	7.14	0.26	0.22	0.78	9.29	0.19	0.34	0.66	7.21
12 × 26	6	III	0.23	0.43	0.57	7.99	0.28	0.30	0.70	9.72	0.23	0.43	0.57	8.02
14 × 24	1	I	0.00	0.00	1.00	1.08	0.13	0.00	1.00	5.00	0.04	0.00	1.00	1.84
14 × 24	2	I	0.01	0.00	1.00	2.17	0.12	0.00	1.00	4.57	0.03	0.00	1.00	2.67
14 × 24	3	II	0.01	0.00	1.00	3.26	0.18	0.00	1.00	6.84	0.03	0.00	1.00	3.71
14 × 24	4	II	0.03	0.02	0.98	4.50	0.22	0.01	0.99	8.38	0.06	0.03	0.97	5.00
14 × 24	5	III	0.08	0.07	0.93	6.06	0.25	0.03	0.97	9.64	0.10	0.08	0.92	6.42
14 × 24	6	III	0.15	0.15	0.85	7.79	0.27	0.07	0.93	10.50	0.17	0.16	0.84	8.02
14 × 24	7	III	0.22	0.23	0.77	9.17	0.29	0.12	0.88	11.16	0.22	0.22	0.78	9.18
18 × 22	1	I	0.00	0.00	1.00	1.00	0.05	0.00	1.00	3.00	0.00	0.00	1.00	1.00
18 × 22	2	I	0.00	0.00	1.00	2.00	0.05	0.00	1.00	3.01	0.00	0.00	1.00	2.00
18 × 22	3	I	0.00	0.00	1.00	3.00	0.11	0.00	1.00	5.02	0.00	0.00	1.00	3.01
18 × 22	4	II	0.00	0.00	1.00	4.00	0.18	0.00	1.00	7.23	0.00	0.00	1.00	4.05
18 × 22	5	II	0.00	0.00	1.00	5.05	0.27	0.00	1.00	9.58	0.01	0.00	1.00	5.11
18 × 22	6	III	0.03	0.02	0.98	6.43	0.35	0.01	0.99	11.52	0.04	0.02	0.98	6.54
18 × 22	7	III	0.09	0.05	0.95	7.98	0.42	0.02	0.98	13.21	0.10	0.05	0.95	8.13
18 × 22	8	III	0.21	0.10	0.90	10.06	0.48	0.04	0.96	14.38	0.21	0.10	0.90	10.11
18 × 22	9	III	0.31	0.16	0.84	11.68	0.53	0.07	0.93	15.20	0.32	0.16	0.84	11.68

**Πίνακας Γ.11:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των επαναληπτικών μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους  $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		Iter-SIS-SCAD					Iter-SIS-Lasso				Iter-SIS-MCP			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.01	0.00	1.00	1.13	0.14	0.00	1.00	5.40	0.02	0.00	1.00	1.55
12 × 26	2	I	0.01	0.00	1.00	2.34	0.14	0.00	1.00	5.40	0.03	0.00	1.00	2.80
12 × 26	3	II	0.04	0.04	0.96	3.75	0.18	0.01	0.99	7.18	0.05	0.05	0.95	4.09
12 × 26	4	III	0.10	0.18	0.82	5.60	0.21	0.10	0.90	8.28	0.12	0.18	0.82	5.81
12 × 26	5	III	0.16	0.29	0.71	6.96	0.23	0.18	0.82	8.99	0.17	0.30	0.70	7.01
12 × 26	6	III	0.22	0.40	0.60	8.04	0.26	0.27	0.73	9.51	0.22	0.40	0.60	8.07
14 × 24	1	I	0.00	0.00	1.00	1.00	0.11	0.00	1.00	4.30	0.00	0.00	1.00	1.00
14 × 24	2	I	0.00	0.00	1.00	2.02	0.11	0.00	1.00	4.36	0.01	0.00	1.00	2.11
14 × 24	3	II	0.01	0.00	1.00	3.11	0.18	0.00	1.00	6.81	0.01	0.00	1.00	3.27
14 × 24	4	II	0.03	0.03	0.97	4.46	0.22	0.01	0.99	8.44	0.04	0.04	0.96	4.70
14 × 24	5	III	0.09	0.10	0.90	6.15	0.26	0.04	0.96	9.66	0.10	0.10	0.90	6.34
14 × 24	6	III	0.16	0.17	0.83	7.92	0.28	0.09	0.91	10.50	0.17	0.18	0.82	8.07
14 × 24	7	III	0.23	0.24	0.76	9.19	0.29	0.13	0.87	11.03	0.23	0.24	0.76	9.18
18 × 22	1	I	0.00	0.00	1.00	1.00	0.08	0.00	1.00	3.60	0.00	0.00	1.00	1.04
18 × 22	2	I	0.00	0.00	1.00	2.00	0.08	0.00	1.00	3.61	0.01	0.00	1.00	2.11
18 × 22	3	I	0.00	0.00	1.00	3.01	0.12	0.00	1.00	5.30	0.01	0.00	1.00	3.16
18 × 22	4	II	0.00	0.00	1.00	4.04	0.17	0.00	1.00	7.14	0.01	0.00	1.00	4.21
18 × 22	5	II	0.01	0.00	1.00	5.08	0.23	0.00	1.00	8.94	0.02	0.00	1.00	5.32
18 × 22	6	III	0.02	0.01	0.99	6.24	0.30	0.00	1.00	10.74	0.04	0.01	0.99	6.51
18 × 22	7	III	0.05	0.02	0.98	7.60	0.35	0.00	1.00	12.19	0.07	0.02	0.98	7.90
18 × 22	8	III	0.11	0.04	0.96	9.17	0.39	0.01	0.99	13.31	0.12	0.04	0.96	9.38
18 × 22	9	III	0.17	0.07	0.93	10.58	0.42	0.02	0.98	14.24	0.19	0.07	0.93	10.78

**Πίνακας Γ.12:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους σχεδιασμούς των Koukouninos et al. [76].

ΣΣΔ		ΣΙΣ-Σ'ΑΔ					ΣΙΣ-Λασσο				ΣΙΣ-Μ'Π			
$n \times p$	$p$	SN	Τυπε I	Τυπε II	Πωερ	ΜΣ	Τυπε I	Τυπε II	Πωερ	ΜΣ	Τυπε I	Τυπε II	Πωερ	ΜΣ
8 × 14	1	I	0.15	0.00	1.00	2.97	0.25	0.00	1.00	5.00	0.12	0.00	1.00	2.58
8 × 14	2	II	0.12	0.03	0.97	3.43	0.22	0.02	0.98	4.58	0.12	0.03	0.97	3.37
8 × 14	3	III	0.17	0.16	0.84	4.37	0.25	0.07	0.93	5.52	0.17	0.15	0.85	4.40
8 × 14	4	III	0.19	0.22	0.78	5.01	0.25	0.10	0.90	6.16	0.19	0.22	0.78	4.99
10 × 18	1	I	0.07	0.00	1.00	2.17	0.19	0.00	1.00	5.00	0.09	0.00	1.00	2.51
10 × 18	2	I	0.07	0.00	1.00	3.08	0.19	0.00	1.00	5.05	0.07	0.00	1.00	3.16
10 × 18	3	II	0.09	0.03	0.97	4.30	0.21	0.00	1.00	6.08	0.10	0.03	0.97	4.35
10 × 18	4	III	0.12	0.07	0.93	5.36	0.22	0.03	0.97	6.94	0.12	0.07	0.93	5.36
10 × 18	5	III	0.16	0.15	0.85	6.37	0.23	0.07	0.93	7.68	0.16	0.14	0.86	6.36
12 × 22	1	I	0.03	0.00	1.00	1.58	0.14	0.00	1.00	4.70	0.04	0.00	1.00	1.90
12 × 22	2	I	0.02	0.00	1.00	2.32	0.13	0.00	1.00	4.69	0.03	0.00	1.00	2.52
12 × 22	3	II	0.03	0.02	0.98	3.45	0.18	0.01	0.99	6.45	0.04	0.03	0.97	3.72
12 × 22	4	III	0.06	0.04	0.96	4.86	0.20	0.02	0.98	7.56	0.07	0.04	0.96	5.05
12 × 22	5	III	0.11	0.09	0.91	6.36	0.22	0.05	0.95	8.43	0.11	0.09	0.91	6.48
12 × 22	6	III	0.15	0.13	0.87	7.67	0.24	0.08	0.92	9.28	0.16	0.13	0.87	7.70
14 × 26	1	I	0.01	0.00	1.00	1.32	0.10	0.00	1.00	4.42	0.03	0.00	1.00	1.70
14 × 26	2	I	0.01	0.00	1.00	2.31	0.10	0.00	1.00	4.42	0.02	0.00	1.00	2.46
14 × 26	3	II	0.02	0.01	0.99	3.44	0.14	0.00	1.00	6.19	0.03	0.01	0.99	3.68
14 × 26	4	II	0.03	0.01	0.99	4.58	0.17	0.00	1.00	7.78	0.04	0.02	0.98	4.80
14 × 26	5	III	0.06	0.05	0.95	6.03	0.19	0.02	0.98	8.91	0.07	0.05	0.95	6.19
14 × 26	6	III	0.10	0.08	0.92	7.59	0.20	0.04	0.96	9.85	0.11	0.08	0.92	7.69
14 × 26	7	III	0.14	0.11	0.89	8.88	0.22	0.07	0.93	10.71	0.14	0.11	0.89	8.92
16 × 30	1	I	0.00	0.00	1.00	1.00	0.07	0.00	1.00	4.00	0.01	0.00	1.00	1.15
16 × 30	2	I	0.00	0.00	1.00	2.10	0.07	0.00	1.00	3.97	0.01	0.00	1.00	2.30
16 × 30	3	I	0.00	0.00	1.00	3.10	0.12	0.00	1.00	6.21	0.01	0.00	1.00	3.33
16 × 30	4	II	0.01	0.01	0.99	4.27	0.16	0.00	1.00	8.10	0.02	0.01	0.99	4.48
16 × 30	5	III	0.04	0.03	0.97	5.76	0.18	0.02	0.98	9.38	0.05	0.03	0.97	5.99
16 × 30	6	III	0.06	0.05	0.95	7.23	0.20	0.03	0.97	10.60	0.07	0.05	0.95	7.37
16 × 30	7	III	0.11	0.09	0.91	8.99	0.22	0.05	0.95	11.75	0.12	0.09	0.91	9.08
16 × 30	8	III	0.14	0.11	0.89	10.30	0.22	0.09	0.91	11.33	0.14	0.10	0.90	10.34

**Πίνακας Γ.13:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους σχεδιασμούς των Koukouninos et al. [76].

SSD		SIS-SCAD					SIS-Lasso				SIS-MCP			
$n \times p$	$p'$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
18 × 34	1	I	0.00	0.00	1.00	1.00	0.04	0.00	1.00	3.17	0.00	0.00	1.00	1.00
18 × 34	2	I	0.00	0.00	1.00	2.00	0.04	0.00	1.00	3.17	0.00	0.00	1.00	2.02
18 × 34	3	I	0.00	0.00	1.00	3.01	0.09	0.00	1.00	5.85	0.00	0.00	1.00	3.05
18 × 34	4	II	0.00	0.00	1.00	4.13	0.14	0.00	1.00	8.32	0.01	0.00	1.00	4.19
18 × 34	5	II	0.02	0.01	0.99	5.39	0.18	0.01	0.99	9.38	0.02	0.01	0.99	5.46
18 × 34	6	III	0.04	0.03	0.97	7.01	0.20	0.04	0.96	9.86	0.05	0.03	0.97	7.13
18 × 34	7	III	0.07	0.06	0.94	8.55	0.22	0.07	0.93	10.71	0.08	0.06	0.94	8.70
18 × 34	8	III	0.11	0.08	0.92	10.25	0.23	0.13	0.87	10.45	0.11	0.08	0.92	10.35
18 × 34	9	III	0.15	0.10	0.90	11.77	0.22	0.16	0.84	10.36	0.15	0.10	0.90	11.81
20 × 38	1	I	0.00	0.00	1.00	1.02	0.04	0.00	1.00	3.35	0.00	0.00	1.00	1.02
20 × 38	2	I	0.00	0.00	1.00	2.06	0.04	0.00	1.00	3.35	0.00	0.00	1.00	2.17
20 × 38	3	I	0.00	0.00	1.00	3.09	0.06	0.00	1.00	4.96	0.01	0.00	1.00	3.24
20 × 38	4	II	0.00	0.00	1.00	4.10	0.08	0.00	1.00	6.76	0.01	0.00	1.00	4.22
20 × 38	5	II	0.01	0.01	0.99	5.25	0.10	0.00	1.00	8.37	0.01	0.01	0.99	5.38
20 × 38	6	III	0.03	0.02	0.98	6.73	0.13	0.01	0.99	10.11	0.03	0.02	0.98	6.87
20 × 38	7	III	0.05	0.04	0.96	8.21	0.16	0.02	0.98	11.77	0.05	0.04	0.96	8.32
20 × 38	8	III	0.08	0.06	0.94	9.95	0.18	0.03	0.97	13.19	0.08	0.06	0.94	10.00
20 × 38	9	III	0.11	0.08	0.92	11.62	0.20	0.05	0.95	14.44	0.12	0.08	0.92	11.66
20 × 38	10	III	0.15	0.10	0.90	13.14	0.22	0.06	0.94	15.46	0.15	0.11	0.89	13.23
22 × 42	1	I	0.00	0.00	1.00	1.00	0.02	0.00	1.00	1.00	0.00	0.00	1.00	1.00
22 × 42	2	I	0.00	0.00	1.00	2.00	0.02	0.00	1.00	2.65	0.00	0.00	1.00	2.00
22 × 42	3	I	0.00	0.00	1.00	3.00	0.03	0.00	1.00	4.32	0.00	0.00	1.00	3.01
22 × 42	4	I	0.00	0.00	1.00	4.00	0.06	0.00	1.00	6.16	0.00	0.00	1.00	4.01
22 × 42	5	II	0.00	0.00	1.00	5.12	0.08	0.00	1.00	8.01	0.00	0.00	1.00	5.15
22 × 42	6	II	0.02	0.01	0.99	6.46	0.11	0.01	0.99	9.90	0.02	0.02	0.98	6.51
22 × 42	7	III	0.03	0.03	0.97	8.03	0.14	0.02	0.98	11.77	0.04	0.03	0.97	8.12
22 × 42	8	III	0.06	0.04	0.96	9.64	0.17	0.02	0.98	13.56	0.06	0.04	0.96	9.66
22 × 42	9	III	0.09	0.07	0.93	11.39	0.19	0.03	0.97	14.93	0.09	0.07	0.93	11.44
22 × 42	10	III	0.12	0.08	0.92	12.89	0.20	0.05	0.95	15.97	0.12	0.08	0.92	12.94
22 × 42	11	III	0.14	0.10	0.90	14.35	0.22	0.06	0.94	17.09	0.16	0.11	0.89	14.19

**Πίνακας Γ.14:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους  $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		SIS-SCAD					SIS-Lasso				SIS-MCP			
$n \times p$	$p'$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.00	0.00	1.00	1.11	0.13	0.00	1.00	5.18	0.02	0.00	1.00	1.39
12 × 26	2	I	0.01	0.00	1.00	2.34	0.13	0.00	1.00	5.18	0.03	0.01	0.99	2.66
12 × 26	3	II	0.05	0.07	0.93	4.01	0.18	0.04	0.96	7.12	0.07	0.08	0.92	4.28
12 × 26	4	III	0.12	0.18	0.82	5.99	0.22	0.13	0.87	8.35	0.13	0.19	0.81	6.16
12 × 26	5	III	0.19	0.26	0.74	7.62	0.25	0.20	0.80	9.31	0.19	0.26	0.74	7.67
12 × 26	6	III	0.22	0.33	0.67	8.45	0.26	0.26	0.74	9.73	0.22	0.33	0.67	8.45
14 × 24	1	I	0.00	0.00	1.00	1.00	0.10	0.00	1.00	4.36	0.02	0.00	1.00	1.42
14 × 24	2	I	0.01	0.00	1.00	2.13	0.11	0.00	1.00	4.36	0.01	0.00	1.00	2.33
14 × 24	3	II	0.01	0.00	1.00	3.27	0.17	0.00	1.00	6.52	0.03	0.01	0.99	3.58
14 × 24	4	II	0.04	0.03	0.97	4.61	0.20	0.01	0.99	8.00	0.05	0.03	0.97	4.88
14 × 24	5	III	0.08	0.06	0.94	6.21	0.23	0.03	0.97	9.27	0.09	0.06	0.94	6.34
14 × 24	6	III	0.14	0.10	0.90	7.85	0.26	0.07	0.93	10.24	0.14	0.10	0.90	7.93
14 × 24	7	III	0.19	0.14	0.86	9.20	0.28	0.09	0.91	11.03	0.19	0.14	0.86	9.21
18 × 22	1	I	0.00	0.00	1.00	1.00	0.06	0.00	1.00	2.15	0.00	0.00	1.00	1.04
18 × 22	2	I	0.00	0.00	1.00	2.00	0.06	0.00	1.00	3.15	0.00	0.00	1.00	2.00
18 × 22	3	I	0.00	0.00	1.00	3.00	0.12	0.00	1.00	5.22	0.00	0.00	1.00	3.03
18 × 22	4	II	0.00	0.00	1.00	4.00	0.20	0.01	0.99	8.00	0.00	0.00	1.00	4.06
18 × 22	5	II	0.01	0.00	1.00	5.09	0.26	0.00	1.00	9.35	0.09	0.06	0.94	6.43
18 × 22	6	III	0.04	0.02	0.98	6.56	0.33	0.01	0.99	11.16	0.05	0.02	0.92	6.68
18 × 22	7	III	0.11	0.05	0.95	8.37	0.40	0.02	0.98	12.76	0.12	0.05	0.95	8.47
18 × 22	8	III	0.20	0.08	0.92	10.22	0.29	0.14	0.86	11.60	0.20	0.08	0.92	10.25
18 × 22	9	III	0.29	0.11	0.89	11.82	0.29	0.16	0.84	11.89	0.30	0.12	0.88	11.88

**Πίνακας Γ.15:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων SIS χρησιμοποιώντας διαφορετικές ποινές σε συνδυασμό με το κριτήριο BIC για τους  $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		SIS-SCAD				SIS-Lasso				SIS-MCP				
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.02	0.00	1.00	1.46	0.12	0.00	1.00	4.90	0.02	0.00	1.00	1.56
12 × 26	2	I	0.02	0.00	1.00	2.51	0.12	0.00	1.00	4.90	0.03	0.00	1.00	2.73
12 × 26	3	II	0.04	0.05	0.95	3.86	0.17	0.03	0.97	6.80	0.06	0.06	0.94	4.08
12 × 26	4	III	0.11	0.14	0.86	5.90	0.21	0.10	0.90	8.19	0.12	0.14	0.86	5.99
12 × 26	5	III	0.17	0.22	0.78	7.39	0.23	0.16	0.84	9.12	0.17	0.22	0.78	7.43
12 × 26	6	III	0.21	0.28	0.72	8.44	0.25	0.23	0.77	9.67	0.21	0.28	0.72	8.43
14 × 24	1	I	0.00	0.00	1.00	1.00	0.12	0.00	1.00	4.62	0.00	0.00	1.00	1.09
14 × 24	2	I	0.00	0.00	1.00	2.07	0.12	0.00	1.00	4.62	0.01	0.00	1.00	2.21
14 × 24	3	II	0.01	0.01	0.99	3.23	0.16	0.00	1.00	6.32	0.02	0.01	0.99	3.43
14 × 24	4	II	0.04	0.03	0.97	4.64	0.20	0.01	0.99	7.91	0.05	0.04	0.96	4.85
14 × 24	5	III	0.09	0.07	0.93	6.32	0.23	0.04	0.96	9.07	0.10	0.07	0.93	6.46
14 × 24	6	III	0.15	0.12	0.88	7.96	0.26	0.08	0.92	10.21	0.15	0.12	0.88	8.04
14 × 24	7	III	0.19	0.15	0.85	9.19	0.27	0.11	0.89	10.80	0.19	0.15	0.85	9.20
18 × 22	1	I	0.00	0.00	1.00	1.00	0.07	0.00	1.00	3.44	0.01	0.00	1.00	1.31
18 × 22	2	I	0.00	0.00	1.00	2.01	0.07	0.00	1.00	3.44	0.01	0.00	1.00	2.23
18 × 22	3	I	0.00	0.00	1.00	3.08	0.12	0.00	1.00	5.23	0.02	0.00	1.00	3.31
18 × 22	4	II	0.01	0.00	1.00	4.10	0.17	0.00	1.00	7.01	0.02	0.00	1.00	4.33
18 × 22	5	II	0.01	0.00	1.00	5.18	0.22	0.00	1.00	8.75	0.03	0.00	1.00	5.42
18 × 22	6	III	0.02	0.01	0.99	6.35	0.27	0.00	1.00	10.30	0.04	0.01	0.99	6.59
18 × 22	7	III	0.05	0.02	0.98	7.65	0.31	0.01	0.99	11.53	0.06	0.02	0.98	7.83
18 × 22	8	III	0.09	0.03	0.97	9.10	0.35	0.01	0.99	12.73	0.11	0.03	0.97	9.28
18 × 22	9	III	0.15	0.05	0.95	10.57	0.37	0.02	0.98	13.55	0.16	0.05	0.95	10.68

## ΜΕΡΟΣ II : Συγκρίσεις με άλλες μεθόδους

**Πίνακας Γ.16:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 1, για τους σχεδιασμούς των Koukouvinos et al. [76].

US		Iter-SIS-SCAD					SIS-SCAD					DS			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS	
8 × 14	1	I	0.17	0.30	0.70	3.00	0.17	0.00	1.00	3.16	0.06	0.06	0.94	1.66	
8 × 14	2	II	0.18	0.20	0.80	3.71	0.14	0.23	0.77	3.32	0.06	0.22	0.78	2.23	
8 × 14	3	III	0.18	0.30	0.70	4.02	0.15	0.31	0.69	3.70	0.06	0.35	0.65	2.65	
8 × 14	4	III	0.19	0.39	0.61	4.38	0.17	0.36	0.64	4.20	0.07	0.45	0.55	2.89	
10 × 18	1	I	0.13	0.10	0.90	3.11	0.11	0.05	0.95	2.81	0.04	0.07	0.93	1.69	
10 × 18	2	I	0.13	0.14	0.86	3.77	0.11	0.14	0.86	3.50	0.03	0.20	0.80	2.09	
10 × 18	3	II	0.16	0.27	0.73	5.10	0.12	0.21	0.79	4.24	0.04	0.33	0.67	2.58	
10 × 18	4	III	0.16	0.27	0.73	5.10	0.14	0.26	0.74	4.89	0.04	0.41	0.59	2.99	
10 × 18	5	III	0.18	0.35	0.65	5.53	0.16	0.32	0.68	5.53	0.06	0.47	0.53	3.38	
12 × 22	1	I	0.05	0.09	0.91	1.87	0.04	0.00	1.00	1.80	0.04	0.07	0.93	1.69	
12 × 22	2	I	0.06	0.12	0.88	2.94	0.06	0.00	1.00	3.00	0.02	0.20	0.80	1.93	
12 × 22	3	II	0.07	0.12	0.88	4.05	0.07	0.12	0.88	3.91	0.02	0.29	0.71	2.48	
12 × 22	4	III	0.09	0.15	0.85	4.99	0.09	0.18	0.82	4.87	0.03	0.38	0.62	2.95	
12 × 22	5	III	0.12	0.23	0.77	5.93	0.12	0.24	0.76	5.81	0.03	0.44	0.56	3.30	
12 × 22	6	III	0.15	0.31	0.69	6.58	0.14	0.28	0.72	6.63	0.03	0.48	0.52	3.66	
14 × 26	1	I	0.03	0.06	0.94	1.71	0.03	0.00	1.00	1.70	0.03	0.04	0.96	1.67	
14 × 26	2	I	0.04	0.09	0.91	2.90	0.02	0.00	1.00	2.40	0.01	0.19	0.81	1.92	
14 × 26	3	II	0.06	0.10	0.90	4.07	0.05	0.12	0.88	3.84	0.01	0.28	0.72	2.49	
14 × 26	4	II	0.07	0.11	0.89	5.19	0.06	0.15	0.85	4.86	0.02	0.36	0.64	3.02	
14 × 26	5	III	0.09	0.16	0.84	6.15	0.10	0.20	0.80	6.09	0.02	0.43	0.57	3.39	
14 × 26	6	III	0.12	0.23	0.77	6.98	0.11	0.23	0.77	6.93	0.03	0.48	0.52	3.72	
14 × 26	7	III	0.16	0.31	0.69	7.82	0.14	0.28	0.72	7.69	0.04	0.50	0.50	4.15	
16 × 30	1	I	0.01	0.04	0.96	1.37	0.03	0.00	1.00	2.00	0.02	0.03	0.97	1.64	
16 × 30	2	I	0.02	0.08	0.92	2.32	0.07	0.10	0.90	3.80	0.01	0.20	0.80	1.77	
16 × 30	3	I	0.02	0.08	0.92	3.37	0.03	0.11	0.89	3.55	0.01	0.28	0.72	2.34	
16 × 30	4	II	0.04	0.08	0.92	4.69	0.05	0.12	0.88	4.84	0.01	0.35	0.65	2.86	
16 × 30	5	III	0.05	0.10	0.90	5.84	0.07	0.15	0.85	6.11	0.02	0.40	0.60	3.37	
16 × 30	6	III	0.07	0.13	0.87	6.90	0.10	0.18	0.82	7.37	0.02	0.44	0.56	3.80	
16 × 30	7	III	0.10	0.18	0.82	8.05	0.13	0.21	0.79	8.38	0.03	0.47	0.53	4.21	
16 × 30	8	III	0.14	0.26	0.74	8.96	0.15	0.25	0.75	9.33	0.03	0.51	0.49	4.38	
18 × 34	1	I	0.01	0.04	0.96	1.25	0.01	0.00	1.00	1.40	0.02	0.04	0.96	1.66	
18 × 34	2	I	0.00	0.07	0.93	1.99	0.00	0.07	0.93	3.38	0.02	0.20	0.80	1.77	
18 × 34	3	I	0.01	0.08	0.92	3.05	0.02	0.09	0.91	3.15	0.02	0.28	0.72	2.32	
18 × 34	4	II	0.01	0.07	0.93	4.15	0.03	0.11	0.89	4.45	0.02	0.34	0.66	2.84	
18 × 34	5	II	0.03	0.08	0.92	5.43	0.05	0.14	0.86	5.75	0.02	0.40	0.60	3.32	
18 × 34	6	III	0.04	0.11	0.89	6.61	0.07	0.15	0.85	7.02	0.02	0.44	0.56	3.75	
18 × 34	7	III	0.07	0.13	0.87	7.93	0.10	0.18	0.82	8.40	0.02	0.47	0.53	4.09	
18 × 34	8	III	0.10	0.19	0.81	9.11	0.12	0.20	0.80	9.50	0.02	0.49	0.51	4.53	
18 × 34	9	III	0.13	0.25	0.75	10.1	0.15	0.25	0.75	10.6	0.02	0.52	0.48	4.85	
20 × 38	1	I	0.01	0.03	0.97	1.27	0.02	0.00	1.00	1.60	0.02	0.04	0.96	1.77	
20 × 38	2	I	0.01	0.07	0.93	2.07	0.01	0.00	1.00	2.20	0.01	0.20	0.80	1.76	
20 × 38	3	I	0.01	0.06	0.94	3.13	0.01	0.09	0.91	3.13	0.01	0.27	0.73	2.30	
20 × 38	4	II	0.01	0.06	0.94	3.13	0.02	0.11	0.89	4.38	0.01	0.33	0.67	2.80	
20 × 38	5	II	0.02	0.06	0.94	5.29	0.04	0.12	0.88	5.77	0.01	0.38	0.62	3.30	
20 × 38	6	III	0.02	0.07	0.93	6.34	0.06	0.14	0.86	6.97	0.01	0.41	0.59	3.80	
20 × 38	7	III	0.03	0.08	0.92	7.39	0.08	0.16	0.84	8.35	0.01	0.45	0.55	4.12	
20 × 38	8	III	0.04	0.11	0.89	8.45	0.09	0.18	0.82	9.34	0.01	0.49	0.51	4.52	
20 × 38	9	III	0.07	0.15	0.85	9.60	0.12	0.21	0.79	10.59	0.02	0.50	0.50	4.96	
20 × 38	10	III	0.10	0.21	0.79	10.80	0.15	0.24	0.76	11.70	0.02	0.53	0.47	5.21	
22 × 42	1	I	0.01	0.03	0.97	1.18	0.00	0.00	1.00	1.30	0.02	0.04	0.96	1.66	
22 × 42	2	I	0.00	0.07	0.93	1.89	0.00	0.03	0.97	2.03	0.01	0.19	0.81	1.75	
22 × 42	3	I	0.00	0.07	0.93	2.85	0.00	0.09	0.91	2.89	0.01	0.27	0.73	2.28	
22 × 42	4	I	0.00	0.06	0.94	3.84	0.01	0.09	0.91	3.96	0.01	0.32	0.68	2.86	
22 × 42	5	II	0.00	0.06	0.94	4.87	0.02	0.11	0.89	5.33	0.01	0.37	0.63	3.34	
22 × 42	6	II	0.01	0.07	0.93	5.89	0.03	0.12	0.88	6.48	0.01	0.41	0.59	3.85	
22 × 42	7	III	0.01	0.07	0.93	6.96	0.06	0.15	0.85	8.06	0.01	0.45	0.55	4.22	
22 × 42	8	III	0.02	0.08	0.92	8.07	0.07	0.16	0.84	9.13	0.01	0.47	0.53	4.64	
22 × 42	9	III	0.04	0.11	0.89	9.22	0.09	0.18	0.82	10.59	0.01	0.49	0.51	4.97	
22 × 42	10	III	0.07	0.16	0.84	10.53	0.12	0.21	0.79	11.68	0.01	0.53	0.47	5.17	
22 × 42	11	III	0.10	0.21	0.79	11.80	0.13	0.23	0.77	12.55	0.02	0.54	0.46	5.55	

**Πίνακας Γ.17:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 1, για τους  $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		Iter-SIS-SCAD					SIS-SCAD				DS			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.04	0.08	0.92	1.91	0.03	0.00	1.00	2.20	0.03	0.05	0.95	1.69
12 × 26	2	I	0.05	0.12	0.88	3.07	0.05	0.10	0.90	3.12	0.03	0.23	0.77	0.23
12 × 26	3	II	0.08	0.17	0.83	4.22	0.09	0.28	0.72	4.26	0.03	0.34	0.66	2.74
12 × 26	4	III	0.10	0.27	0.73	5.22	0.12	0.32	0.68	5.38	0.04	0.45	0.55	3.17
12 × 26	5	III	0.13	0.36	0.64	6.02	0.13	0.35	0.65	6.11	0.05	0.52	0.48	3.49
12 × 26	6	III	0.16	0.46	0.54	6.47	0.18	0.46	0.54	6.97	0.06	0.56	0.44	3.80
14 × 24	1	I	0.02	0.05	0.95	1.50	0.01	0.00	1.00	1.20	0.03	0.04	0.96	1.66
14 × 24	2	I	0.03	0.08	0.92	2.56	0.03	0.03	0.97	2.62	0.01	0.20	0.80	1.85
14 × 24	3	II	0.05	0.08	0.92	3.76	0.05	0.12	0.88	3.68	0.02	0.29	0.71	2.47
14 × 24	4	II	0.06	0.10	0.90	4.88	0.08	0.15	0.85	4.98	0.02	0.37	0.63	2.99
14 × 24	5	III	0.10	0.16	0.84	6.03	0.11	0.19	0.81	6.16	0.03	0.44	0.56	3.37
14 × 24	6	III	0.12	0.20	0.80	6.99	0.14	0.23	0.77	7.06	0.03	0.49	0.51	3.71
14 × 24	7	III	0.18	0.31	0.69	7.81	0.18	0.29	0.71	7.98	0.04	0.55	0.45	3.83
18 × 22	1	I	0.02	0.05	0.95	1.31	0.00	0.00	1.00	1.10	0.03	0.04	0.96	1.57
18 × 22	2	I	0.01	0.09	0.91	1.94	0.02	0.05	0.95	2.22	0.01	0.20	0.80	1.70
18 × 22	3	I	0.01	0.08	0.92	2.96	0.01	0.09	0.91	2.99	0.01	0.27	0.73	2.27
18 × 22	4	II	0.02	0.07	0.93	4.05	0.03	0.08	0.92	4.21	0.01	0.33	0.67	2.79
18 × 22	5	II	0.03	0.07	0.93	5.18	0.06	0.11	0.89	5.49	0.01	0.39	0.61	3.27
18 × 22	6	III	0.05	0.08	0.92	6.26	0.10	0.13	0.87	6.82	0.02	0.44	0.56	3.67
18 × 22	7	III	0.07	0.10	0.90	7.45	0.17	0.17	0.83	8.35	0.03	0.49	0.51	4.04
18 × 22	8	III	0.12	0.13	0.87	8.71	0.22	0.19	0.81	9.55	0.04	0.52	0.48	4.38
18 × 22	9	III	0.18	0.16	0.84	9.92	0.28	0.22	0.78	10.64	0.05	0.56	0.44	4.66

**Πίνακας Γ.18:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 1, για τους  $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		Iter-SIS-SCAD					SIS-SCAD				DS			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.04	0.09	0.91	1.84	0.05	0.00	1.00	2.30	0.03	0.07	0.93	1.73
12 × 26	2	I	0.05	0.12	0.88	2.96	0.09	0.20	0.80	3.70	0.02	0.23	0.77	2.08
12 × 26	3	II	0.07	0.15	0.85	4.20	0.08	0.21	0.79	4.15	0.03	0.34	0.66	2.64
12 × 26	4	III	0.10	0.25	0.75	5.13	0.10	0.29	0.71	5.13	0.04	0.44	0.56	3.09
12 × 26	5	III	0.13	0.36	0.64	5.95	0.14	0.37	0.63	6.09	0.05	0.51	0.49	3.48
12 × 26	6	III	0.16	0.44	0.56	6.51	0.16	0.40	0.60	6.75	0.06	0.56	0.44	3.73
14 × 24	1	I	0.02	0.07	0.93	1.36	0.01	0.00	1.00	1.20	0.03	0.04	0.96	1.60
14 × 24	2	I	0.02	0.09	0.91	2.34	0.02	0.05	0.95	2.30	0.02	0.21	0.79	1.86
14 × 24	3	II	0.04	0.09	0.91	3.62	0.07	0.10	0.90	4.08	0.02	0.30	0.70	2.45
14 × 24	4	II	0.07	0.12	0.88	4.89	0.09	0.16	0.84	5.12	0.02	0.38	0.62	2.93
14 × 24	5	III	0.10	0.16	0.84	6.05	0.11	0.19	0.81	6.21	0.03	0.44	0.56	3.30
14 × 24	6	III	0.13	0.23	0.77	7.11	0.15	0.26	0.74	7.15	0.04	0.50	0.50	3.62
14 × 24	7	III	0.18	0.30	0.70	7.91	0.18	0.29	0.71	8.07	0.04	0.52	0.48	3.98
18 × 22	1	I	0.02	0.06	0.94	1.33	0.02	0.00	1.00	1.50	0.03	0.03	0.97	1.63
18 × 22	2	I	0.02	0.08	0.92	2.29	0.05	0.03	0.97	2.41	0.01	0.19	0.81	1.71
18 × 22	3	I	0.03	0.07	0.93	3.39	0.03	0.08	0.92	3.42	0.01	0.28	0.72	2.24
18 × 22	4	II	0.04	0.06	0.94	4.44	0.04	0.08	0.92	4.41	0.01	0.33	0.67	2.74
18 × 22	5	II	0.05	0.07	0.93	5.59	0.07	0.09	0.91	5.71	0.01	0.39	0.61	3.17
18 × 22	6	III	0.06	0.07	0.93	6.60	0.09	0.10	0.90	6.87	0.01	0.42	0.58	3.64
18 × 22	7	III	0.08	0.07	0.93	7.62	0.11	0.11	0.89	7.90	0.02	0.44	0.56	4.06
18 × 22	8	III	0.11	0.09	0.91	8.73	0.14	0.13	0.87	9.03	0.02	0.48	0.52	4.36
18 × 22	9	III	0.14	0.12	0.88	9.80	0.19	0.15	0.85	10.13	0.03	0.51	0.49	4.68

**Πίνακας Γ.19:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους σχεδιασμούς των Koukouνι- nos et al. [76].

SSD		LAD-Lasso					SQR-Lasso				Stepwise-SCAD			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
8 × 14	1	I	0.10	0.07	0.93	2.17	0.09	0.09	0.91	2.13	0.14	0.02	0.98	2.85
8 × 14	2	II	0.09	0.36	0.64	2.39	0.06	0.17	0.83	2.41	0.14	0.15	0.85	3.39
8 × 14	3	III	0.09	0.42	0.58	2.76	0.05	0.34	0.66	2.54	0.16	0.35	0.65	3.72
8 × 14	4	III	0.09	0.51	0.49	2.88	0.04	0.49	0.51	2.49	0.19	0.54	0.46	3.72
10 × 18	1	I	0.15	0.07	0.93	3.44	0.08	0.08	0.92	2.31	0.16	0.00	1.00	3.70
10 × 18	2	I	0.13	0.26	0.74	3.48	0.05	0.13	0.87	2.60	0.15	0.04	0.96	4.30
10 × 18	3	II	0.10	0.35	0.65	3.51	0.04	0.28	0.72	2.78	0.17	0.21	0.79	4.88
10 × 18	4	III	0.10	0.47	0.53	3.47	0.03	0.42	0.58	2.78	0.20	0.39	0.61	5.28
10 × 18	5	III	0.09	0.54	0.46	3.42	0.03	0.55	0.45	2.63	0.23	0.51	0.49	5.45
12 × 22	1	I	0.05	0.08	0.92	2.07	0.07	0.08	0.92	2.49	0.18	0.00	1.00	4.76
12 × 22	2	I	0.04	0.37	0.63	2.11	0.05	0.13	0.87	2.74	0.17	0.03	0.97	5.25
12 × 22	3	II	0.04	0.45	0.55	2.38	0.04	0.20	0.80	3.13	0.17	0.10	0.90	5.94
12 × 22	4	III	0.04	0.56	0.44	2.44	0.03	0.34	0.66	3.09	0.19	0.25	0.75	6.43
12 × 22	5	III	0.04	0.62	0.38	2.57	0.02	0.47	0.53	3.01	0.23	0.40	0.60	6.95
12 × 22	6	III	0.04	0.67	0.33	2.65	0.02	0.56	0.44	2.88	0.24	0.48	0.52	6.96
14 × 26	1	I	0.07	0.07	0.93	2.73	0.06	0.06	0.94	2.56	0.19	0.00	1.00	5.74
14 × 26	2	I	0.06	0.28	0.72	2.90	0.05	0.08	0.92	3.01	0.18	0.03	0.97	6.25
14 × 26	3	II	0.05	0.38	0.62	3.01	0.03	0.16	0.84	3.32	0.18	0.08	0.92	6.94
14 × 26	4	II	0.05	0.49	0.51	3.08	0.03	0.28	0.72	3.45	0.19	0.20	0.80	7.47
14 × 26	5	III	0.05	0.56	0.44	3.23	0.02	0.41	0.59	3.42	0.22	0.30	0.70	8.03
14 × 26	6	III	0.05	0.63	0.37	3.21	0.02	0.51	0.49	3.31	0.24	0.42	0.58	8.32
14 × 26	7	III	0.05	0.66	0.34	3.26	0.02	0.61	0.39	3.06	0.25	0.48	0.52	8.41
16 × 30	1	I	0.01	0.08	0.92	3.72	0.06	0.07	0.93	2.71	0.20	0.00	1.00	6.88
16 × 30	2	I	0.08	0.17	0.83	3.83	0.05	0.10	0.90	3.12	0.18	0.00	1.00	7.26
16 × 30	3	I	0.06	0.25	0.75	3.90	0.03	0.12	0.88	3.50	0.18	0.02	0.98	7.80
16 × 30	4	II	0.06	0.38	0.62	3.91	0.02	0.22	0.78	3.73	0.18	0.10	0.90	8.40
16 × 30	5	III	0.05	0.48	0.52	3.83	0.02	0.34	0.66	3.71	0.20	0.19	0.81	9.02
16 × 30	6	III	0.05	0.56	0.44	3.84	0.01	0.42	0.58	3.75	0.22	0.31	0.69	9.53
16 × 30	7	III	0.05	0.61	0.39	3.81	0.01	0.53	0.47	3.50	0.25	0.41	0.59	9.87
16 × 30	8	III	0.05	0.65	0.35	3.79	0.01	0.61	0.39	3.33	0.27	0.46	0.54	10.2

**Πίνακας Γ.20:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους σχεδιασμούς των Koukouνι- nos et al. [76].

SSD		LAD-Lasso					SQR-Lasso				Stepwise-SCAD			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
18 × 34	1	I	0.04	0.07	0.93	2.25	0.06	0.06	0.94	2.90	0.21	0.00	1.00	7.78
18 × 34	2	I	0.03	0.27	0.73	2.47	0.04	0.07	0.93	3.29	0.19	0.01	0.99	8.21
18 × 34	3	I	0.02	0.37	0.63	2.55	0.03	0.10	0.90	3.76	0.19	0.02	0.98	8.86
18 × 34	4	II	0.02	0.50	0.50	2.58	0.02	0.17	0.83	3.98	0.19	0.05	0.95	9.57
18 × 34	5	II	0.02	0.59	0.41	2.60	0.02	0.26	0.74	4.16	0.19	0.12	0.88	9.87
18 × 34	6	III	0.02	0.65	0.35	2.63	0.01	0.37	0.63	4.14	0.21	0.21	0.79	10.69
18 × 34	7	III	0.02	0.71	0.29	2.67	0.01	0.48	0.52	3.92	0.23	0.30	0.70	11.18
18 × 34	8	III	0.02	0.73	0.27	2.73	0.01	0.55	0.45	3.81	0.25	0.40	0.60	11.55
18 × 34	9	III	0.02	0.77	0.23	2.67	0.01	0.62	0.38	3.61	0.27	0.46	0.54	11.70
20 × 38	1	I	0.05	0.07	0.93	2.68	0.05	0.07	0.93	2.89	0.22	0.00	1.00	8.96
20 × 38	2	I	0.04	0.20	0.80	2.94	0.04	0.06	0.94	3.31	0.20	0.00	1.00	9.30
20 × 38	3	I	0.03	0.29	0.71	3.12	0.03	0.09	0.91	3.86	0.20	0.00	1.00	9.86
20 × 38	4	II	0.03	0.42	0.58	3.15	0.02	0.13	0.87	4.22	0.19	0.02	0.98	10.45
20 × 38	5	II	0.02	0.53	0.47	3.09	0.02	0.22	0.78	4.44	0.19	0.07	0.93	10.94
20 × 38	6	III	0.02	0.61	0.39	3.06	0.01	0.32	0.68	4.41	0.21	0.16	0.84	11.83
20 × 38	7	III	0.02	0.67	0.33	2.99	0.01	0.41	0.59	4.38	0.23	0.26	0.74	12.32
20 × 38	8	III	0.02	0.71	0.29	2.86	0.01	0.51	0.49	4.13	0.26	0.37	0.63	12.83
20 × 38	9	III	0.02	0.75	0.25	2.84	0.01	0.58	0.42	3.96	0.27	0.42	0.58	12.96
20 × 38	10	III	0.02	0.78	0.22	2.77	0.01	0.64	0.36	3.73	0.27	0.46	0.54	13.08
22 × 42	1	I	0.07	0.06	0.94	3.68	0.05	0.05	0.95	3.03	0.22	0.00	1.00	9.85
22 × 42	2	I	0.05	0.09	0.91	3.87	0.04	0.06	0.94	3.48	0.21	0.00	1.00	10.47
22 × 42	3	I	0.04	0.16	0.84	3.95	0.03	0.08	0.92	3.98	0.21	0.01	0.99	11.00
22 × 42	4	I	0.03	0.32	0.68	3.91	0.02	0.11	0.89	4.45	0.20	0.02	0.98	11.34
22 × 42	5	II	0.03	0.43	0.57	3.94	0.02	0.18	0.82	4.75	0.20	0.05	0.95	12.02
22 × 42	6	II	0.03	0.52	0.48	3.94	0.01	0.26	0.74	4.92	0.21	0.12	0.88	12.71
22 × 42	7	III	0.03	0.59	0.41	3.86	0.01	0.37	0.63	4.73	0.22	0.19	0.81	13.37
22 × 42	8	III	0.03	0.65	0.35	3.84	0.01	0.46	0.54	4.59	0.24	0.29	0.71	13.88
22 × 42	9	III	0.03	0.68	0.32	3.80	0.01	0.54	0.46	4.35	0.26	0.35	0.65	14.24
22 × 42	10	III	0.03	0.72	0.28	3.64	0.01	0.60	0.40	4.19	0.27	0.41	0.59	14.57
22 × 42	11	III	0.03	0.74	0.26	3.71	0.01	0.66	0.34	3.92	0.28	0.45	0.55	14.69



**Πίνακας Γ.21:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους  $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		LAD-Lasso					SQR-Lasso				Stepwise-SCAD			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.06	0.09	0.91	2.38	0.07	0.07	0.93	2.60	0.16	0.00	1.00	4.89
12 × 26	2	I	0.06	0.38	0.62	2.47	0.05	0.15	0.85	2.94	0.15	0.07	0.93	5.38
12 × 26	3	II	0.05	0.49	0.51	2.70	0.04	0.32	0.68	3.02	0.17	0.25	0.75	6.14
12 × 26	4	III	0.05	0.59	0.41	2.82	0.04	0.45	0.55	3.01	0.19	0.42	0.58	6.53
12 × 26	5	III	0.06	0.65	0.35	2.95	0.04	0.57	0.43	2.92	0.21	0.52	0.48	6.88
12 × 26	6	III	0.06	0.71	0.29	2.89	0.04	0.64	0.36	2.85	0.22	0.58	0.42	6.97
14 × 24	1	I	0.08	0.06	0.94	2.75	0.07	0.08	0.92	2.51	0.20	0.00	1.00	5.64
14 × 24	2	I	0.06	0.27	0.73	2.84	0.05	0.09	0.91	2.99	0.18	0.01	0.99	5.92
14 × 24	3	II	0.05	0.36	0.64	3.03	0.04	0.16	0.84	3.31	0.18	0.05	0.95	6.66
14 × 24	4	II	0.05	0.49	0.51	3.09	0.03	0.29	0.71	3.46	0.19	0.16	0.84	7.19
14 × 24	5	III	0.05	0.56	0.44	3.12	0.03	0.43	0.57	3.39	0.22	0.30	0.70	7.60
14 × 24	6	III	0.05	0.61	0.39	3.27	0.02	0.51	0.49	3.32	0.25	0.43	0.57	8.02
14 × 24	7	III	0.05	0.65	0.35	3.25	0.02	0.59	0.41	3.27	0.27	0.50	0.50	8.14
18 × 22	1	I	0.12	0.07	0.93	3.49	0.08	0.06	0.94	2.53	0.25	0.00	1.00	6.22
18 × 22	2	I	0.09	0.08	0.92	3.64	0.06	0.07	0.93	3.00	0.25	0.00	1.00	6.96
18 × 22	3	I	0.05	0.31	0.69	3.67	0.03	0.09	0.91	3.45	0.24	0.00	1.00	7.50
18 × 22	4	II	0.05	0.31	0.69	3.67	0.03	0.16	0.84	3.84	0.23	0.01	0.99	8.13
18 × 22	5	II	0.05	0.43	0.57	3.71	0.02	0.29	0.71	3.83	0.22	0.06	0.94	8.40
18 × 22	6	III	0.05	0.52	0.48	3.73	0.01	0.42	0.58	3.69	0.23	0.15	0.85	8.71
18 × 22	7	III	0.06	0.60	0.40	3.73	0.01	0.54	0.46	3.38	0.25	0.29	0.71	8.66
18 × 22	8	III	0.06	0.65	0.35	3.67	0.01	0.63	0.37	3.18	0.26	0.46	0.54	7.99
18 × 22	9	III	0.06	0.68	0.32	3.71	0.02	0.69	0.31	2.98	0.29	0.53	0.47	7.94

**Πίνακας Γ.22:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους  $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		LAD-Lasso					SQR-Lasso				Stepwise-SCAD			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.06	0.08	0.92	2.41	0.07	0.07	0.93	2.58	0.17	0.00	1.00	5.32
12 × 26	2	I	0.05	0.36	0.64	2.50	0.05	0.14	0.86	2.89	0.16	0.06	0.94	5.83
12 × 26	3	II	0.05	0.50	0.50	2.59	0.04	0.29	0.71	3.00	0.16	0.16	0.84	6.24
12 × 26	4	III	0.05	0.58	0.42	2.77	0.03	0.45	0.55	2.93	0.19	0.39	0.61	6.63
12 × 26	5	III	0.05	0.64	0.36	2.80	0.03	0.56	0.44	2.90	0.20	0.51	0.49	6.64
12 × 26	6	III	0.05	0.69	0.31	2.87	0.03	0.64	0.36	2.84	0.21	0.58	0.42	6.80
14 × 24	1	I	0.08	0.07	0.93	2.74	0.07	0.06	0.94	2.55	0.20	0.00	1.00	5.83
14 × 24	2	I	0.05	0.36	0.64	2.88	0.05	0.09	0.91	2.98	0.20	0.01	0.99	6.41
14 × 24	3	II	0.04	0.36	0.64	2.84	0.04	0.17	0.83	3.29	0.20	0.06	0.94	6.91
14 × 24	4	II	0.04	0.49	0.51	2.89	0.03	0.29	0.71	3.45	0.20	0.18	0.82	7.38
14 × 24	5	III	0.05	0.59	0.41	2.93	0.03	0.29	0.71	3.48	0.21	0.29	0.71	7.64
14 × 24	6	III	0.04	0.65	0.35	2.87	0.02	0.52	0.48	3.29	0.24	0.42	0.58	7.81
14 × 24	7	III	0.04	0.69	0.31	2.88	0.02	0.61	0.39	3.13	0.25	0.49	0.51	7.85
18 × 22	1	I	0.08	0.06	0.94	2.56	0.07	0.07	0.93	2.47	0.26	0.00	1.00	6.45
18 × 22	2	I	0.05	0.11	0.89	2.76	0.06	0.06	0.94	3.03	0.25	0.00	1.00	6.93
18 × 22	3	I	0.03	0.21	0.79	2.99	0.04	0.08	0.92	3.47	0.24	0.00	1.00	7.53
18 × 22	4	II	0.02	0.36	0.64	3.02	0.02	0.13	0.87	3.91	0.24	0.01	0.99	8.33
18 × 22	5	II	0.02	0.50	0.50	2.90	0.01	0.23	0.77	4.07	0.23	0.03	0.97	8.82
18 × 22	6	III	0.02	0.59	0.41	2.82	0.01	0.35	0.65	4.04	0.22	0.08	0.92	9.10
18 × 22	7	III	0.02	0.66	0.34	2.76	0.01	0.46	0.54	3.87	0.22	0.18	0.82	9.06
18 × 22	8	III	0.03	0.71	0.29	2.71	0.01	0.58	0.42	3.50	0.23	0.30	0.70	8.84
18 × 22	9	III	0.02	0.74	0.26	2.65	0.01	0.65	0.35	3.20	0.24	0.42	0.58	8.35

**Πίνακας Γ.23:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 2, για τους σχεδιασμούς των Koukouvinos et al. [76].

SSD		Iter-SIS-SCAD					SIS-SCAD				DS			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
8 × 14	1	I	0.20	0.00	1.00	3.59	0.15	0.00	1.00	3.00	0.00	0.00	1.00	1.05
8 × 14	2	II	0.16	0.03	0.97	3.84	0.13	0.04	0.96	3.44	0.08	0.00	1.00	3.00
8 × 14	3	III	0.16	0.12	0.88	4.44	0.16	0.12	0.88	4.35	0.10	0.13	0.87	3.73
8 × 14	4	III	0.23	0.29	0.71	5.10	0.18	0.19	0.81	5.05	0.13	0.24	0.76	4.34
10 × 18	1	I	0.12	0.00	1.00	3.06	0.07	0.00	1.00	2.21	0.00	0.00	1.00	1.00
10 × 18	2	I	0.09	0.00	1.00	3.51	0.07	0.00	1.00	3.10	0.03	0.00	1.00	2.48
10 × 18	3	II	0.10	0.04	0.96	4.44	0.09	0.04	0.96	4.31	0.07	0.04	0.96	4.00
10 × 18	4	III	0.14	0.11	0.89	5.56	0.13	0.09	0.91	5.42	0.10	0.05	0.95	4.80
10 × 18	5	III	0.20	0.22	0.78	6.50	0.17	0.15	0.85	6.45	0.11	0.23	0.77	5.31
12 × 22	1	I	0.02	0.00	1.00	1.43	0.02	0.00	1.00	1.49	0.00	0.00	1.00	1.00
12 × 22	2	I	0.02	0.00	1.00	2.34	0.01	0.00	1.00	2.26	0.01	0.00	1.00	2.21
12 × 22	3	II	0.02	0.00	1.00	3.46	0.03	0.02	0.98	3.53	0.03	0.02	0.98	3.57
12 × 22	4	III	0.05	0.05	0.95	4.80	0.06	0.05	0.95	4.92	0.06	0.08	0.92	4.70
12 × 22	5	III	0.10	0.10	0.90	6.24	0.13	0.10	0.90	6.67	0.08	0.16	0.84	5.49
12 × 22	6	III	0.17	0.20	0.80	7.62	0.17	0.14	0.86	7.84	0.10	0.23	0.77	6.11
14 × 26	1	I	0.01	0.00	1.00	1.28	0.01	0.00	1.00	1.32	0.00	0.00	1.00	1.00
14 × 26	2	I	0.02	0.00	1.00	2.37	0.01	0.00	1.00	2.33	0.00	0.00	1.00	2.12
14 × 26	3	II	0.02	0.00	1.00	3.47	0.02	0.01	0.99	3.46	0.02	0.01	0.99	3.47
14 × 26	4	II	0.03	0.02	0.98	4.69	0.04	0.02	0.98	4.71	0.04	0.05	0.95	4.76
14 × 26	5	III	0.07	0.08	0.92	6.18	0.05	0.03	0.97	5.85	0.06	0.12	0.88	5.69
14 × 26	6	III	0.13	0.14	0.86	7.70	0.12	0.10	0.90	7.76	0.08	0.19	0.81	6.41
14 × 26	7	III	0.18	0.21	0.79	9.01	0.16	0.13	0.87	9.08	0.09	0.26	0.74	6.94
16 × 30	1	I	0.00	0.00	1.00	1.01	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00
16 × 30	2	I	0.00	0.00	1.00	2.05	0.00	0.00	1.00	2.06	0.00	0.00	1.00	2.03
16 × 30	3	I	0.00	0.00	1.00	3.07	0.01	0.01	0.99	3.24	0.00	0.00	1.00	3.17
16 × 30	4	II	0.01	0.01	0.99	4.19	0.03	0.02	0.98	4.59	0.02	0.03	0.97	4.51
16 × 30	5	III	0.02	0.02	0.98	5.45	0.05	0.04	0.96	6.09	0.04	0.07	0.93	5.67
16 × 30	6	III	0.06	0.06	0.94	6.98	0.10	0.07	0.93	7.88	0.05	0.13	0.87	6.50
16 × 30	7	III	0.11	0.12	0.88	8.59	0.13	0.10	0.90	9.30	0.07	0.18	0.82	7.26
16 × 30	8	III	0.16	0.18	0.82	10.1	0.17	0.13	0.87	10.66	0.07	0.22	0.78	7.85

**Πίνακας Γ.24:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 2, για τους σχεδιασμούς των Koukouvinos et al. [76].

SSD		Iter-SIS-SCAD					SIS-SCAD				DS			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
18 × 34	1	I	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00
18 × 34	2	I	0.00	0.00	1.00	2.00	0.00	0.00	1.00	2.00	0.00	0.00	1.00	2.04
18 × 34	3	I	0.00	0.00	1.00	3.01	0.00	0.00	1.00	3.09	0.01	0.00	1.00	3.20
18 × 34	4	II	0.00	0.00	1.00	4.05	0.01	0.01	0.99	4.35	0.02	0.02	0.98	4.38
18 × 34	5	II	0.01	0.01	0.99	5.22	0.04	0.03	0.97	5.98	0.03	0.06	0.94	5.58
18 × 34	6	III	0.02	0.02	0.98	6.40	0.06	0.04	0.96	7.34	0.04	0.10	0.90	6.51
18 × 34	7	III	0.05	0.06	0.94	7.99	0.09	0.07	0.93	9.03	0.05	0.15	0.85	7.39
18 × 34	8	III	0.09	0.11	0.89	9.59	0.12	0.09	0.91	10.49	0.07	0.20	0.80	8.15
18 × 34	9	III	0.15	0.16	0.84	11.31	0.17	0.13	0.87	12.03	0.07	0.23	0.77	8.68
20 × 38	1	I	0.00	0.00	1.00	1.02	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00
20 × 38	2	I	0.00	0.00	1.00	2.05	0.00	0.00	1.00	2.08	0.00	0.00	1.00	2.00
20 × 38	3	I	0.00	0.00	1.00	3.07	0.00	0.00	1.00	3.16	0.00	0.00	1.00	3.00
20 × 38	4	II	0.00	0.00	1.00	4.09	0.02	0.01	0.99	4.52	0.00	0.00	1.00	4.27
20 × 38	5	II	0.01	0.00	1.00	5.18	0.03	0.02	0.98	5.82	0.02	0.02	0.98	5.53
20 × 38	6	III	0.01	0.01	0.99	6.31	0.05	0.04	0.96	7.48	0.06	0.06	0.94	6.64
20 × 38	7	III	0.03	0.03	0.97	7.60	0.09	0.06	0.94	9.25	0.11	0.11	0.89	7.58
20 × 38	8	III	0.05	0.06	0.94	9.15	0.11	0.08	0.92	10.64	0.15	0.15	0.85	8.43
20 × 38	9	III	0.10	0.11	0.89	10.92	0.14	0.10	0.90	12.29	0.20	0.20	0.80	8.97
20 × 38	10	III	0.15	0.18	0.82	12.52	0.18	0.12	0.88	13.73	0.24	0.24	0.76	9.61
22 × 42	1	I	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00
22 × 42	2	I	0.00	0.00	1.00	2.01	0.00	0.00	1.00	2.00	0.00	0.00	1.00	2.00
22 × 42	3	I	0.00	0.00	1.00	3.00	0.00	0.00	1.00	3.06	0.00	0.00	1.00	3.00
22 × 42	4	I	0.00	0.00	1.00	4.02	0.01	0.01	0.99	4.24	0.01	0.00	1.00	4.00
22 × 42	5	II	0.00	0.00	1.00	5.03	0.01	0.01	0.99	5.47	0.02	0.02	0.98	5.00
22 × 42	6	II	0.00	0.00	1.00	6.07	0.03	0.02	0.98	6.96	0.03	0.05	0.95	6.00
22 × 42	7	III	0.01	0.01	0.99	7.30	0.06	0.05	0.95	8.91	0.04	0.10	0.90	7.67
22 × 42	8	III	0.03	0.03	0.97	8.61	0.09	0.06	0.94	10.44	0.05	0.13	0.87	8.54
22 × 42	9	III	0.06	0.08	0.92	10.35	0.12	0.09	0.91	12.20	0.05	0.18	0.82	9.18
22 × 42	10	III	0.11	0.12	0.88	12.23	0.14	0.09	0.91	13.52	0.06	0.21	0.79	9.84
22 × 42	11	III	0.15	0.17	0.83	13.63	0.17	0.11	0.89	15.01	0.07	0.26	0.74	10.39

**Πίνακας Γ.25:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 2, για τους  $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		Iter-SIS-SCAD					SIS-SCAD				DS			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.01	0.00	1.00	1.23	0.01	0.00	1.00	1.13	0.00	0.00	1.00	1.02
12 × 26	2	I	0.02	0.01	0.99	2.41	0.02	0.02	0.98	2.47	0.03	0.00	1.00	2.65
12 × 26	3	II	0.04	0.05	0.95	3.82	0.07	0.08	0.92	4.33	0.07	0.08	0.92	4.46
12 × 26	4	III	0.10	0.17	0.93	5.58	0.14	0.18	0.82	6.28	0.10	0.23	0.77	5.41
12 × 26	5	III	0.17	0.30	0.70	7.01	0.20	0.26	0.74	7.82	0.11	0.33	0.67	5.74
12 × 26	6	III	0.22	0.41	0.59	7.93	0.23	0.33	0.67	8.65	0.13	0.45	0.55	5.96
14 × 24	1	I	0.00	0.00	1.00	1.06	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00
14 × 24	2	I	0.01	0.00	1.00	2.22	0.01	0.00	1.00	2.13	0.01	0.00	1.00	2.10
14 × 24	3	II	0.01	0.00	1.00	3.27	0.02	0.02	0.98	3.42	0.03	0.01	0.99	3.50
14 × 24	4	II	0.03	0.03	0.97	4.53	0.05	0.04	0.96	4.83	0.05	0.08	0.92	4.73
14 × 24	5	III	0.07	0.07	0.93	5.97	0.10	0.08	0.92	6.47	0.07	0.16	0.84	5.58
14 × 24	6	III	0.13	0.13	0.87	7.59	0.15	0.11	0.89	8.08	0.10	0.25	0.75	6.10
14 × 24	7	III	0.22	0.24	0.76	9.14	0.20	0.16	0.84	9.35	0.11	0.23	0.77	6.47
18 × 22	1	I	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00
18 × 22	2	I	0.00	0.00	1.00	2.00	0.00	0.00	1.00	2.00	0.00	0.00	1.00	2.00
18 × 22	3	I	0.00	0.00	1.00	3.00	0.00	0.00	1.00	3.04	0.00	0.00	1.00	3.05
18 × 22	4	II	0.00	0.00	1.00	4.00	0.01	0.00	1.00	4.85	0.01	0.00	1.00	4.23
18 × 22	5	II	0.00	0.00	1.00	5.05	0.02	0.01	0.99	5.31	0.04	0.02	0.98	5.61
18 × 22	6	III	0.03	0.01	0.99	6.33	0.07	0.03	0.97	6.99	0.08	0.07	0.93	6.92
18 × 22	7	III	0.08	0.04	0.96	7.86	0.15	0.06	0.94	8.87	0.13	0.15	0.85	7.87
18 × 22	8	III	0.17	0.08	0.92	9.77	0.23	0.09	0.91	10.58	0.17	0.23	0.77	8.59
18 × 22	9	III	0.28	0.13	0.87	11.45	0.32	0.12	0.88	12.05	0.20	0.30	0.70	8.85

**Πίνακας Γ.26:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων Iter-SIS-SCAD, SIS-SCAD και Dantzig selector στο μοντέλο 2, για τους  $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		Iter-SIS-SCAD					SIS-SCAD				DS			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.01	0.00	1.00	1.14	0.02	0.00	1.00	1.39	0.00	0.00	1.00	1.00
12 × 26	2	I	0.01	0.00	1.00	2.30	0.02	0.01	0.99	2.57	0.02	0.06	0.94	2.30
12 × 26	3	II	0.04	0.04	0.96	3.73	0.06	0.07	0.93	4.12	0.03	0.19	0.81	3.21
12 × 26	4	III	0.08	0.14	0.86	5.29	0.12	0.15	0.85	5.99	0.05	0.34	0.66	3.72
12 × 26	5	III	0.16	0.27	0.73	6.98	0.18	0.23	0.77	7.64	0.06	0.45	0.55	3.92
12 × 26	6	III	0.24	0.44	0.56	8.53	0.21	0.28	0.72	8.58	0.07	0.54	0.46	4.11
14 × 24	1	I	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00
14 × 24	2	I	0.00	0.00	1.00	2.02	0.00	0.00	1.00	2.08	0.00	0.03	0.97	2.05
14 × 24	3	II	0.01	0.01	0.99	3.11	0.02	0.02	0.98	3.42	0.01	0.11	0.89	2.97
14 × 24	4	II	0.02	0.03	0.97	4.35	0.06	0.05	0.95	4.99	0.03	0.22	0.78	3.72
14 × 24	5	III	0.06	0.07	0.93	5.86	0.11	0.09	0.91	6.70	0.04	0.33	0.67	4.05
14 × 24	6	III	0.15	0.16	0.84	7.68	0.17	0.13	0.87	8.25	0.04	0.33	0.67	4.05
14 × 24	7	III	0.22	0.22	0.78	9.18	0.22	0.17	0.83	9.49	0.05	0.47	0.53	4.65
18 × 22	1	I	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	1.00
18 × 22	2	I	0.00	0.00	1.00	2.00	0.00	0.00	1.00	2.03	0.00	0.01	0.99	1.98
18 × 22	3	I	0.00	0.00	1.00	3.01	0.00	0.00	1.00	3.07	0.00	0.05	0.95	2.89
18 × 22	4	II	0.00	0.00	1.00	4.05	0.01	0.00	1.00	4.20	0.01	0.10	0.90	3.68
18 × 22	5	II	0.01	0.00	1.00	5.08	0.02	0.00	1.00	5.30	0.01	0.18	0.82	4.28
18 × 22	6	III	0.02	0.01	0.99	6.22	0.04	0.01	0.99	6.58	0.01	0.24	0.76	4.79
18 × 22	7	III	0.04	0.01	0.99	7.44	0.07	0.02	0.98	7.96	0.02	0.30	0.70	5.16
18 × 22	8	III	0.08	0.03	0.97	8.84	0.12	0.04	0.96	9.39	0.02	0.34	0.66	5.59
18 × 22	9	III	0.16	0.07	0.93	10.44	0.18	0.06	0.94	10.97	0.03	0.38	0.62	5.93

**Πίνακας Γ.27:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 2, για τους σχεδιασμούς των Koukouνι- nos et al. [76].

SSD		LAD-Lasso					SQR-Lasso				Stepwise-SCAD			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
8 × 14	1	I	0.06	0.00	1.00	1.82	0.08	0.00	1.00	2.10	0.14	0.00	1.00	2.84
8 × 14	2	II	0.13	0.29	0.71	3.02	0.08	0.01	0.99	2.95	0.12	0.06	0.94	3.38
8 × 14	3	III	0.10	0.31	0.69	3.15	0.08	0.29	0.71	2.99	0.13	0.23	0.77	3.75
8 × 14	4	III	0.12	0.47	0.53	3.33	0.05	0.54	0.46	2.32	0.19	0.45	0.55	4.12
10 × 18	1	I	0.15	0.00	1.00	3.47	0.08	0.00	1.00	2.35	0.17	0.00	1.00	3.81
10 × 18	2	I	0.12	0.22	0.78	3.40	0.05	0.00	1.00	2.79	0.15	0.01	0.99	4.36
10 × 18	3	II	0.10	0.29	0.71	3.58	0.04	0.15	0.85	3.20	0.16	0.12	0.88	5.02
10 × 18	4	III	0.09	0.50	0.50	3.21	0.04	0.39	0.61	2.95	0.21	0.30	0.70	5.74
10 × 18	5	III	0.08	0.56	0.44	3.29	0.03	0.55	0.45	2.66	0.25	0.45	0.55	6.05
12 × 22	1	I	0.04	0.00	1.00	1.90	0.07	0.00	1.00	2.52	0.18	0.00	1.00	4.78
12 × 22	2	I	0.06	0.32	0.68	2.52	0.05	0.00	1.00	2.96	0.17	0.01	0.99	5.30
12 × 22	3	II	0.04	0.36	0.64	2.67	0.03	0.11	0.89	3.20	0.16	0.06	0.94	5.88
12 × 22	4	III	0.05	0.55	0.45	2.74	0.02	0.27	0.73	3.29	0.18	0.15	0.85	6.68
12 × 22	5	III	0.04	0.57	0.43	2.87	0.02	0.48	0.52	2.91	0.23	0.31	0.69	7.36
12 × 22	6	III	0.05	0.67	0.33	2.82	0.02	0.61	0.39	2.64	0.27	0.44	0.56	7.68
14 × 26	1	I	0.06	0.00	1.00	2.52	0.07	0.00	1.00	2.66	0.19	0.00	1.00	5.83
14 × 26	2	I	0.06	0.21	0.79	3.09	0.05	0.00	1.00	3.12	0.18	0.01	0.99	6.24
14 × 26	3	II	0.04	0.30	0.70	3.05	0.03	0.07	0.93	3.51	0.17	0.03	0.97	6.79
14 × 26	4	II	0.06	0.49	0.51	3.35	0.02	0.23	0.77	3.56	0.18	0.10	0.90	7.55
14 × 26	5	III	0.05	0.52	0.48	3.46	0.02	0.42	0.58	3.32	0.21	0.21	0.79	8.26
14 × 26	6	III	0.05	0.52	0.48	3.46	0.02	0.58	0.42	2.88	0.24	0.34	0.66	8.85
14 × 26	7	III	0.05	0.64	0.36	3.45	0.01	0.66	0.34	2.65	0.27	0.43	0.57	9.15
16 × 30	1	I	0.09	0.00	1.00	3.71	0.06	0.00	1.00	2.75	0.20	0.00	1.00	6.74
16 × 30	2	I	0.08	0.11	0.89	3.90	0.04	0.00	1.00	3.22	0.19	0.00	1.00	7.31
16 × 30	3	I	0.05	0.19	0.81	3.93	0.03	0.02	0.98	3.79	0.17	0.01	0.99	7.68
16 × 30	4	II	0.05	0.40	0.60	3.73	0.02	0.11	0.89	4.06	0.18	0.03	0.97	8.47
16 × 30	5	III	0.05	0.47	0.53	3.90	0.01	0.30	0.70	3.79	0.18	0.10	0.90	9.08
16 × 30	6	III	0.05	0.58	0.42	3.63	0.01	0.46	0.54	3.47	0.22	0.21	0.79	9.94
16 × 30	7	III	0.05	0.61	0.39	3.76	0.01	0.59	0.41	3.07	0.25	0.32	0.68	10.46
16 × 30	8	III	0.05	0.68	0.32	3.57	0.01	0.67	0.33	2.81	0.28	0.43	0.57	10.76

**Πίνακας Γ.28:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 2, για τους σχεδιασμούς των Koukouνι- nos et al. [76].

SSD		LAD-Lasso					SQR-Lasso				Stepwise-SCAD			
$n \times p$	$p$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
18 × 34	1	I	0.04	0.00	1.00	2.19	0.06	0.00	1.00	2.84	0.21	0.00	1.00	7.78
18 × 34	2	I	0.03	0.20	0.80	2.61	0.04	0.00	1.00	3.31	0.20	0.00	1.00	8.48
18 × 34	3	I	0.02	0.29	0.71	2.68	0.03	0.01	0.99	3.84	0.19	0.01	0.99	8.91
18 × 34	4	II	0.03	0.51	0.49	2.75	0.02	0.06	0.94	4.33	0.18	0.02	0.98	9.37
18 × 34	5	II	0.02	0.58	0.42	2.71	0.01	0.21	0.79	4.34	0.19	0.06	0.94	10.06
18 × 34	6	III	0.03	0.67	0.33	2.72	0.01	0.39	0.61	3.96	0.20	0.13	0.87	10.87
18 × 34	7	III	0.03	0.69	0.31	2.83	0.01	0.52	0.48	3.54	0.23	0.22	0.78	11.55
18 × 34	8	III	0.03	0.74	0.26	2.84	0.01	0.62	0.38	3.23	0.26	0.34	0.66	12.04
18 × 34	9	III	0.03	0.76	0.24	2.82	0.04	0.70	0.30	2.79	0.28	0.40	0.60	12.33
20 × 38	1	I	0.04	0.00	1.00	2.54	0.05	0.00	1.00	2.96	0.21	0.00	1.00	8.90
20 × 38	2	I	0.04	0.17	0.83	2.95	0.04	0.00	1.00	3.47	0.21	0.00	1.00	9.43
20 × 38	3	I	0.03	0.21	0.79	3.27	0.03	0.00	1.00	4.03	0.19	0.00	1.00	9.75
20 × 38	4	II	0.02	0.45	0.55	3.02	0.02	0.03	0.97	4.53	0.19	0.01	0.99	10.32
20 × 38	5	II	0.02	0.53	0.47	3.03	0.04	0.13	0.87	4.71	0.18	0.02	0.98	10.83
20 × 38	6	III	0.02	0.64	0.36	2.86	0.01	0.31	0.69	4.38	0.18	0.06	0.94	11.50
20 × 38	7	III	0.02	0.69	0.31	2.75	0.01	0.47	0.53	3.92	0.21	0.15	0.85	12.36
20 × 38	8	III	0.02	0.75	0.25	2.60	0.01	0.57	0.43	3.58	0.24	0.26	0.74	13.13
20 × 38	9	III	0.02	0.78	0.22	2.47	0.00	0.65	0.35	3.26	0.27	0.37	0.63	13.60
20 × 38	10	III	0.02	0.81	0.19	2.42	0.00	0.70	0.30	3.00	0.29	0.42	0.58	13.81
22 × 42	1	I	0.06	0.00	1.00	3.61	0.05	0.00	1.00	3.07	0.22	0.00	1.00	9.96
22 × 42	2	I	0.05	0.05	0.95	3.84	0.04	0.00	1.00	3.56	0.21	0.00	1.00	10.27
22 × 42	3	I	0.03	0.34	0.66	3.85	0.03	0.00	1.00	4.13	0.21	0.00	1.00	11.08
22 × 42	4	I	0.03	0.34	0.66	3.75	0.02	0.02	0.98	4.72	0.20	0.01	0.99	11.46
22 × 42	5	II	0.03	0.34	0.66	3.79	0.01	0.08	0.92	5.06	0.19	0.02	0.98	11.87
22 × 42	6	II	0.03	0.57	0.43	3.70	0.01	0.22	0.78	5.03	0.19	0.05	0.95	12.54
22 × 42	7	III	0.03	0.60	0.40	3.79	0.01	0.38	0.62	4.63	0.21	0.13	0.87	13.54
22 × 42	8	III	0.03	0.68	0.32	3.58	0.01	0.52	0.48	4.05	0.22	0.19	0.81	14.02
22 × 42	9	III	0.03	0.70	0.30	3.60	0.01	0.61	0.39	3.69	0.25	0.28	0.72	14.64
22 × 42	10	III	0.03	0.75	0.25	3.44	0.01	0.68	0.32	3.37	0.27	0.36	0.64	15.05
22 × 42	11	III	0.03	0.77	0.23	3.47	0.00	0.73	0.27	3.07	0.29	0.42	0.58	15.24

**Πίνακας Γ.29:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 2, για τους  $E(s^2)$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		LAD-Lasso				SQR-Lasso				Stepwise-SCAD				
$n \times p$	$p'$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.05	0.00	1.00	2.19	0.06	0.00	1.00	2.60	0.15	0.00	1.00	4.72
12 × 26	2	I	0.07	0.37	0.63	2.96	0.05	0.00	1.00	3.18	0.14	0.03	0.97	5.35
12 × 26	3	II	0.05	0.42	0.58	2.93	0.04	0.24	0.76	3.19	0.15	0.13	0.87	6.13
12 × 26	4	III	0.06	0.59	0.41	3.04	0.04	0.47	0.53	2.95	0.19	0.31	0.69	6.87
12 × 26	5	III	0.06	0.62	0.38	3.22	0.04	0.59	0.41	2.89	0.23	0.48	0.52	7.47
12 × 26	6	III	0.06	0.71	0.29	3.04	0.04	0.68	0.32	2.76	0.25	0.54	0.46	7.69
14 × 24	1	I	0.07	0.00	1.00	2.65	0.07	0.00	1.00	2.57	0.20	0.00	1.00	5.50
14 × 24	2	I	0.07	0.20	0.80	3.12	0.05	0.00	1.00	3.08	0.18	0.00	1.00	5.95
14 × 24	3	II	0.05	0.30	0.70	3.07	0.03	0.08	0.92	3.50	0.18	0.03	0.97	6.70
14 × 24	4	II	0.06	0.47	0.53	3.29	0.03	0.24	0.76	3.57	0.18	0.08	0.92	7.27
14 × 24	5	III	0.05	0.53	0.47	3.31	0.03	0.44	0.56	3.27	0.21	0.19	0.81	8.06
14 × 24	6	III	0.06	0.61	0.39	3.44	0.03	0.57	0.43	3.05	0.25	0.32	0.68	8.64
14 × 24	7	III	0.05	0.63	0.37	3.43	0.03	0.66	0.34	2.80	0.29	0.44	0.56	8.78
18 × 22	1	I	0.11	0.00	1.00	3.44	0.07	0.00	1.00	2.57	0.26	0.00	1.00	6.43
18 × 22	2	I	0.09	0.01	0.99	3.74	0.05	0.00	1.00	3.01	0.25	0.00	1.00	6.93
18 × 22	3	I	0.05	0.03	0.97	3.78	0.04	0.00	1.00	3.70	0.23	0.00	1.00	7.40
18 × 22	4	II	0.05	0.34	0.66	3.60	0.02	0.06	0.94	4.05	0.23	0.00	1.00	8.17
18 × 22	5	II	0.06	0.43	0.57	3.89	0.01	0.32	0.68	3.58	0.22	0.02	0.98	8.73
18 × 22	6	III	0.06	0.57	0.43	3.60	0.01	0.52	0.48	3.04	0.22	0.06	0.94	9.10
18 × 22	7	III	0.06	0.60	0.40	3.74	0.01	0.67	0.33	2.55	0.23	0.16	0.84	9.37
18 × 22	8	III	0.07	0.68	0.32	3.54	0.02	0.74	0.26	2.31	0.25	0.29	0.71	9.13
18 × 22	9	III	0.07	0.70	0.30	3.62	0.02	0.78	0.22	2.22	0.28	0.42	0.58	8.93

**Πίνακας Γ.30:** Αποτελέσματα προσομοιώσεων από τις συγκρίσεις μεταξύ των μεθόδων LAD-Lasso, SQR-Lasso και Stepwise-SCAD στο μοντέλο 1, για τους  $D$ -βέλτιστους σχεδιασμούς των Marley και Woods [101].

SSD		LAD-Lasso				SQR-Lasso				Stepwise-SCAD				
$n \times p$	$p'$	SN	Type I	Type II	Power	MS	Type I	Type II	Power	MS	Type I	Type II	Power	MS
12 × 26	1	I	0.05	0.00	1.00	2.33	0.06	0.00	1.00	2.60	0.17	0.00	1.00	5.20
12 × 26	2	I	0.05	0.33	0.67	2.46	0.05	0.01	0.99	3.08	0.16	0.01	0.99	5.79
12 × 26	3	II	0.05	0.43	0.57	2.84	0.04	0.17	0.83	3.30	0.16	0.10	0.90	6.34
12 × 26	4	III	0.05	0.57	0.43	2.76	0.03	0.41	0.59	3.04	0.17	0.26	0.74	6.77
12 × 26	5	III	0.06	0.63	0.37	3.02	0.03	0.57	0.43	2.78	0.20	0.44	0.56	6.95
12 × 26	6	III	0.05	0.69	0.31	2.85	0.03	0.67	0.33	2.62	0.21	0.54	0.46	6.97
14 × 24	1	I	0.07	0.00	1.00	2.60	0.07	0.00	1.00	2.54	0.21	0.00	1.00	5.88
14 × 24	2	I	0.05	0.20	0.80	2.79	0.05	0.00	1.00	3.02	0.20	0.00	1.00	6.38
14 × 24	3	II	0.04	0.30	0.70	2.95	0.03	0.07	0.93	3.50	0.19	0.02	0.98	6.84
14 × 24	4	II	0.04	0.48	0.52	2.88	0.02	0.22	0.78	3.57	0.18	0.07	0.93	7.40
14 × 24	5	III	0.04	0.57	0.43	2.93	0.02	0.41	0.59	3.35	0.20	0.20	0.80	7.81
14 × 24	6	III	0.04	0.64	0.36	2.89	0.02	0.55	0.45	3.07	0.22	0.32	0.68	8.10
14 × 24	7	III	0.04	0.68	0.32	2.96	0.02	0.64	0.36	2.85	0.25	0.44	0.56	8.11
18 × 22	1	I	0.08	0.00	1.00	2.58	0.08	0.00	1.00	2.58	0.26	0.00	1.00	6.56
18 × 22	2	I	0.05	0.05	0.95	2.87	0.05	0.00	1.00	3.09	0.25	0.00	1.00	7.06
18 × 22	3	I	0.02	0.09	0.91	3.19	0.03	0.00	1.00	3.65	0.23	0.00	1.00	7.36
18 × 22	4	II	0.02	0.50	0.50	2.88	0.02	0.04	0.96	4.18	0.25	0.00	1.00	8.49
18 × 22	5	II	0.02	0.50	0.50	2.89	0.01	0.15	0.85	4.40	0.22	0.01	0.99	8.64
18 × 22	6	III	0.03	0.60	0.40	2.81	0.01	0.36	0.64	3.92	0.22	0.04	0.96	9.30
18 × 22	7	III	0.03	0.68	0.32	2.63	0.01	0.53	0.47	3.34	0.22	0.10	0.90	9.59
18 × 22	8	III	0.02	0.72	0.28	2.58	0.00	0.65	0.35	2.88	0.22	0.19	0.81	9.58
18 × 22	9	III	0.02	0.76	0.24	2.46	0.00	0.72	0.28	2.54	0.22	0.31	0.69	9.07



---

## Παράρτημα Δ

---

### Παράρτημα Κεφαλαίου 9

(Πληροφορίες παρέχονται επίσης στο συμπληρωματικό υλικό το οποίο είναι διαθέσιμο στη διεύθυνση <http://www.asq.org/pub/jqt/>)

#### Απόδειξη του Λήμματος 1.

Οι στήλες  $\mathbf{d}_i, i = 1, 2, \dots, r$  είναι ισορροπημένες και όλες είναι ορθογώνιες μεταξύ τους. Η στήλη  $\mathbf{d}_{i_1 i_2}, 1 \leq i_1 < i_2 \leq r$  είναι ισορροπημένη επειδή

$$\sum_{k=1}^N d_{i_1 k} d_{i_2 k} = 0.$$

Είναι άμεσο από την κατασκευή του πίνακα  $\mathbf{H}_N$  ότι

$$\sum_{k=1}^N d_{i_1 k} d_{i_2 k} \cdots d_{i_{2\ell+1} k} = 0, \quad (.29)$$

για όλα τα  $\ell = 0, 1, \dots$ . Έτσι,  $\mathbf{d}_{i_1 i_2 i_3}, 1 \leq i_1 < i_2 < i_3 \leq r$  είναι επίσης ισορροπημένος.

Ας εξετάσουμε τώρα την δομή των ταυτόσημων επιδράσεων των στηλών του πίνακα  $\mathbf{D}$ . Δύο στήλες  $\mathbf{x}, \mathbf{y}$  του πίνακα  $\mathbf{D}$  θα λέγεται ότι είναι εντελώς ταυτόσημες εάν  $\mathbf{x}^T \mathbf{y} = \pm N$ . Έχει ήδη σημειωθεί ότι οι στήλες  $\mathbf{d}_{i_1}, \mathbf{d}_{i_2}, 1 \leq i_1 < i_2 \leq r$  είναι ορθογώνιες μεταξύ τους. Από την εξίσωση (.29) έχουμε ότι οι ταυτόσημες οποιωνδήποτε δύο παραγόντων του  $\mathbf{D}$  που συνολικά εμπλέκουν ακριβώς  $2\ell + 1$  ( $\ell$  μη αρνητικός ακέραιος) διακριτές στήλες του  $\mathbf{H}_N$ , θα είναι μηδέν και έτσι θα είναι ορθογώνιες. Από τα παραπάνω αποτελέσματα καταλήγουμε στο συμπέρασμα ότι τα ακόλουθα ζεύγη παραγόντων του  $\mathbf{D}$  είναι ορθογώνια  $(\mathbf{d}_{i_1}, \mathbf{d}_{i_1 i_2}), (\mathbf{d}_{i_1}, \mathbf{d}_{i_2 i_3}), (\mathbf{d}_{i_1}, \mathbf{d}_{i_1 i_2 i_3}), (\mathbf{d}_{i_1 i_2}, \mathbf{d}_{i_1 i_2 i_3}), (\mathbf{d}_{i_1 i_2}, \mathbf{d}_{i_1 i_3 i_4}), (\mathbf{d}_{i_1 i_2}, \mathbf{d}_{i_3 i_4 i_5}), (\mathbf{d}_{i_1 i_2 i_3}, \mathbf{d}_{i_1 i_2 i_4})$ . Το υπόλοιπο της απόδειξης ακολουθεί

αμέσως αφού ικανοποιούνται οι (1) και (2).

**Απόδειξη του Θεωρήματος 1** Εδώ αρκεί να αποδειχθεί ότι πληρούνται οι προϋποθέσεις (1) και (2) του Λήμματος 1. Επίσης, πρέπει να σημειωθεί ότι η αλληλεπίδραση  $\mathbf{d}_{i_1 i_2 i_3 i_4}$  θα ταυτίζεται πλήρως με την επίδραση του μέσου εάν και μόνο εάν η αλληλεπίδραση  $\mathbf{d}_{i_1 i_2 i_3 i_4}^*$  ταυτίζεται πλήρως με την επίδραση του μέσου. Η αλληλεπίδραση  $\mathbf{d}_{i_1 i_2 i_3 i_4}^*$  θα είναι πλήρως ταυτόσημη με την επίδραση του μέσου εάν οι στήλες  $i_1, i_2, i_3, i_4$  είναι μία από τις ακόλουθες δύο δυνατότητες.

$i_1$	$i_2$	$i_3$	$i_4$	$i_1$	$i_2$	$i_3$	$i_4$
+	+	+	+	-	+	+	+
+	+	-	-	+	-	+	+
+	-	+	-	+	+	-	+
+	-	-	+	+	+	+	-
-	+	+	-	-	-	-	+
-	+	-	+	-	-	+	-
-	-	+	+	-	+	-	-
-	-	-	-	+	-	-	-

Εάν είναι δυνατόν, ας υποθέσουμε ότι οι στήλες αυτές είναι πλήρως εξαρτημένες και θεωρούμε την πρώτη επιλογή. Η απόδειξη της άλλης επιλογής θα είναι παρόμοια. Έστω  $u_1, u_2, u_3, u_4, u_5, u_6, u_7$  και  $u_8$  είναι ο αριθμός των φορών που οι συνδυασμοί των συμβόλων  $++++, ++--, +-+-, +--+, -++-, -+-+, --++$  και  $----$  εμφανίζονται σε ολόκληρο το σχεδιασμό. Επειδή οι στήλες  $\mathbf{d}_{i_1}^*, \mathbf{d}_{i_2}^*, \mathbf{d}_{i_3}^*$  και  $\mathbf{d}_{i_4}^*$  είναι ισορροπημένες, έχουμε τις ακόλουθες εξισώσεις.

$$u_1 + u_2 = u_7 + u_8$$

$$u_1 + u_3 = u_6 + u_8$$

$$u_1 + u_4 = u_5 + u_8$$

$$u_1 + u_5 = u_4 + u_8$$

$$u_1 + u_6 = u_3 + u_8$$

$$u_1 + u_7 = u_2 + u_8$$



Οι παραπάνω εξισώσεις έχουν ως αποτέλεσμα

$$u_1 = u_8, \quad u_2 = u_7, \quad u_3 = u_6, \quad \text{and} \quad u_4 = u_5.$$

Και πάλι, από τη στήλη  $\mathbf{d}_{i_1}^*$  είναι ορθογώνια στις στήλες  $\mathbf{d}_{i_2}^*$ ,  $\mathbf{d}_{i_3}^*$  και  $\mathbf{d}_{i_4}^*$ , έχουμε περαιτέρω

$$u_5 + u_6 = u_7 + u_8$$

$$u_5 + u_7 = u_6 + u_8$$

$$u_6 + u_7 = u_5 + u_8$$

Αυτές οι εξισώσεις παρέχουν

$$u_5 + u_6 + u_7 = 3u_8 \quad \text{or} \quad u_5 + u_6 + u_7 + u_8 = 4u_8.$$

Αλλά, αφού η στήλη  $\mathbf{d}_{i_1}^*$  είναι ισορροπημένη, έχουμε

$$u_5 + u_6 + u_7 + u_8 = 4m + 2 \quad \text{ή} \quad u_8 = m + 1/2$$

Αυτό δεν είναι δυνατό και έχουμε μια αντίφαση που σημαίνει ότι οι στήλες  $\mathbf{d}_{i_1}$  και  $\mathbf{d}_{i_2 i_3 i_4}$  δεν είναι πλήρως ταυτόσημες.

Τέλος, ας ελέγξουμε εάν η αλληλεπίδραση  $\mathbf{d}_{i_1 i_2 i_3 i_4 i_5 i_6}$ ,  $1 \leq i_1 < i_2 < i_3 < i_4 < i_5 < i_6 \leq r$  είναι πλήρως ταυτόσημη με την επίδραση του μέσου.

$$\sum_{k=1}^N d_{i_1 k} d_{i_2 k} d_{i_3 k} d_{i_4 k} d_{i_5 k} d_{i_6 k} = 2 \sum_{k=1}^n d_{i_1 k} d_{i_2 k} d_{i_3 k} d_{i_4 k} d_{i_5 k} d_{i_6 k}$$

Η αλληλεπίδραση  $\mathbf{d}_{i_1 i_2 i_3 i_4 i_5 i_6}$  θα είναι πλήρως ταυτόσημη με την επίδραση του μέσου εάν και μόνο εάν η αλληλεπίδραση  $\mathbf{d}_{i_1 i_2 i_3 i_4 i_5 i_6}^*$  ταυτίζεται πλήρως με την επίδραση του μέσου. Οι μόνες πιθανές περιπτώσεις για να συμβεί αυτό είναι να

έχουμε μία από τις ακόλουθες δύο συλλογές διακριτών πειραματικών εκτελέσεων:

$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$
+	+	+	+	+	+	+	-	+	+	-	+	-	+	+	+	+	-	-	-	+	+	+	+
+	+	+	+	-	-	+	-	+	-	+	+	-	+	+	+	-	+	-	-	-	-	+	+
+	+	+	-	+	-	+	-	-	+	+	+	-	+	+	-	+	+	-	-	-	+	-	+
+	+	+	-	-	+	+	-	-	-	-	+	-	+	+	-	-	-	-	-	+	-	-	+
+	+	-	+	+	-	+	-	-	-	+	-	-	+	-	+	+	+	-	-	-	+	+	-
+	+	-	-	+	+	+	-	+	-	-	-	-	+	-	-	+	-	-	-	+	+	-	-
+	+	-	-	-	-	+	-	+	+	+	-	-	+	-	-	-	+	-	-	-	-	-	-

ή

$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$
+	+	+	+	+	-	+	+	-	-	+	-	-	+	+	+	-	-	-	-	+	-	+	+
+	+	+	+	-	+	+	+	-	-	-	+	-	+	+	-	+	-	-	-	-	+	+	+
+	+	+	-	+	+	+	-	+	+	-	-	-	+	+	-	-	+	+	-	-	-	-	-
+	+	-	+	+	+	+	-	+	-	+	-	-	+	-	+	+	-	-	+	-	-	-	-
+	-	+	+	+	+	+	-	+	-	-	+	-	+	-	+	-	+	-	-	+	-	-	-
-	+	+	+	+	+	+	-	-	+	+	-	-	+	-	-	+	+	-	-	-	+	-	-
+	+	+	-	-	-	+	-	-	+	-	+	-	-	+	+	+	-	-	-	-	-	+	-
+	+	-	+	-	-	+	-	-	-	+	+	-	-	+	+	-	+	-	-	-	-	-	+

Εάν είναι δυνατόν, ας υποθέσουμε ότι αυτές οι στήλες είναι πλήρως ταυτόσημες και θεωρούμε την πρώτη συλλογή διακριτών πειραματικών εκτελέσεων. Η απόδειξη της άλλης συλλογής είναι παρόμοια. Έστω  $u_i, 1 \leq i \leq 32$  είναι ο αριθμός των φορών που οι συνδυασμοί των συμβόλων που αναφέρονται στην πρώτη επιλογή εμφανίζονται σε ολόκληρο το σχεδιασμό. Πρέπει να σημειωθεί ότι οι στήλες  $\mathbf{d}_{i_1}^*$ ,  $\mathbf{d}_{i_2}^*$ ,  $\mathbf{d}_{i_3}^*$ ,  $\mathbf{d}_{i_4}^*$ ,  $\mathbf{d}_{i_5}^*$  και  $\mathbf{d}_{i_6}^*$  είναι ορθογώνιες μεταξύ τους. Λόγω της ορθογωνικότητας μεταξύ αυτών των στηλών, καθεμία από τις στήλες  $i_2 - i_5$  που αποτελείται από τις πρώτες 16 σειρές πρέπει να είναι τουλάχιστον ισορροπημένη. Αυτό σημαίνει ότι έχουμε τις ακόλουθες

εξισώσεις.

$$\begin{aligned}u_1 + u_2 + u_3 + u_4 + u_5 + u_6 + u_7 + u_8 &= 2m + 1 \\u_9 + u_{10} + u_{11} + u_{12} + u_{13} + u_{14} + u_{15} + u_{16} &= 2m + 1 \\u_1 + u_2 + u_3 + u_4 + u_9 + u_{10} + u_{15} + u_{16} &= 2m + 1 \\u_1 + u_2 + u_5 + u_6 + u_9 + u_{11} + u_{14} + u_{16} &= 2m + 1 \\u_1 + u_3 + u_5 + u_7 + u_{10} + u_{11} + u_{13} + u_{16} &= 2m + 1 \\u_1 + u_4 + u_6 + u_7 + u_9 + u_{10} + u_{11} + u_{12} &= 2m + 1\end{aligned}$$

Από τις παραπάνω εξισώσεις, έχουμε την απλοποίηση,

$$2m + 1 + 2(u_1 - u_8) = 2(u_{12} + u_{13} + u_{14} + u_{15}),$$

που είναι μια αντίφαση.

**Πίνακας Δ.1:** Διανύσματα δεδομένων  $y$  που δημιουργήθηκαν από το πραγματικό αποκρι-  
τικό μοντέλο για το Παράδειγμα 1 και το Παράδειγμα 2, αντίστοιχα.

Υ(Παράδειγμα 1)				Υ(Παράδειγμα 2)			
1.	3.2899	27.	4.4705	1.	2.063	27.	8.055
2.	2.8646	28.	6.7239	2.	4.222	28.	5.309
3.	4.2935	29.	5.3923	3.	4.790	29.	3.814
4.	4.6329	30.	2.2038	4.	4.358	30.	5.214
5.	4.8237	31.	3.5459	5.	5.018	31.	4.545
6.	4.8633	32.	1.2366	6.	4.332	32.	5.661
7.	2.9599	33.	4.1455	7.	4.119	33.	6.685
8.	2.1846	34.	3.3259	8.	3.967	34.	4.191
9.	3.4778	35.	2.0709	9.	4.508	35.	7.227
10.	5.9263	36.	4.3410	10.	5.962	36.	4.538
11.	4.5668	37.	2.3969	11.	6.784	37.	8.307
12.	6.9583	38.	6.6161	12.	4.861	38.	6.704
13.	3.1946	39.	6.6849	13.	5.254	39.	6.474
14.	5.5459	40.	4.8801	14.	3.361	40.	5.768
15.	5.3045			15.	4.838	41.	2.573
16.	4.5311			16.	3.709	42.	4.137
17.	5.6376			17.	2.966	43.	4.050
18.	1.6251			18.	2.651	44.	4.010
19.	2.1511			19.	1.251	45.	4.763
20.	2.5945			20.	2.913	46.	4.611
21.	5.7143			21.	7.181	47.	4.664
22.	5.3713			22.	3.476	48.	3.538
23.	3.1564			23.	6.797	49.	6.054
24.	3.4684			24.	6.162	50.	4.340
25.	2.5655			25.	3.363	51.	4.130
26.	3.3507			26.	6.174		

**Πίνακας Δ.2:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται τρεις παράγοντες ελέγχου και τρεις παράγοντες θορύβου για το Μοντέλο 1. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III(Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	99.880	0.023	99.880	0.023	99.880	0.023
$\alpha_1$	0.912	0.023	1.912	0.023	2.912	0.023
$b_1$	0.992	0.023	1.992	0.024	2.992	0.023
$c_{11}$	1.006	0.023	2.006	0.022	3.006	0.022
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE II</b>						
$\mu$	99.880	0.022	99.880	0.022	99.880	0.022
$\alpha_1$	0.912	0.022	1.912	0.022	2.912	0.022
$\alpha_2$	0.726	0.022	1.726	0.022	2.726	0.022
$b_1$	0.934	0.022	1.934	0.022	2.992	0.022
$\alpha_{12}$	1.141	0.022	2.141	0.022	3.141	0.022
$c_{11}$	0.995	0.022	1.995	0.022	3.006	0.022
$c_{21}$	1.207	0.022	2.207	0.022	2.958	0.022
$d_{121}$	0.974	0.022	1.974	0.022	2.974	0.022
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE III</b>						
$\mu$	99.880	0.023	99.880	0.023	99.880	0.023
$\alpha_1$	0.912	0.023	1.912	0.023	2.912	0.023
$\alpha_2$	0.726	0.023	1.726	0.023	2.726	0.023
$b_1$	0.934	0.026	1.934	0.026	2.934	0.026
$b_2$	0.856	0.026	1.856	0.026	2.856	0.026
$\alpha_{12}$	1.141	0.023	2.141	0.023	3.141	0.023
$c_{11}$	0.995	0.026	1.995	0.026	2.995	0.023
$c_{12}$	1.207	0.026	2.207	0.026	3.207	0.026
$c_{21}$	0.889	0.026	1.889	0.026	2.889	0.026
$c_{22}$	1.031	0.026	2.031	0.026	3.031	0.026
$d_{121}$	1.022	0.026	2.022	0.026	3.022	0.026
$d_{122}$	1.174	0.026	2.174	0.026	3.174	0.026
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	

**Πίνακας Δ.3:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται τρεις παράγοντες ελέγχου και τρεις παράγοντες θορύβου για το Μοντέλο 1. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	99.880	0.023	99.880	0.023	99.880	0.023
$\alpha_1$	-1.088	0.023	-2.088	0.023	-3.088	0.023
$b_1$	0.992	0.023	1.992	0.023	2.992	0.022
$c_{11}$	1.006	0.023	2.006	0.023	-2.994	0.022
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE II</b>						
$\mu$	99.880	0.022	99.880	0.022	100.028	0.022
$\alpha_1$	-1.088	0.022	-2.088	0.022	-3.088	0.022
$\alpha_2$	0.726	0.022	1.726	0.022	2.726	0.022
$b_1$	-1.008	0.022	-2.08	0.022	-3.066	0.022
$\alpha_{12}$	1.141	0.022	2.141	0.022	3.141	0.022
$c_{11}$	-0.994	0.022	-1.994	0.022	-2.994	0.022
$c_{21}$	-1.042	0.022	-2.042	0.022	-3.042	0.022
$d_{121}$	0.974	0.022	1.974	0.022	2.974	0.022
Type I	0.027		0.000		0.001	
Type II	0.000		0.000		0.000	
<b>CASE III</b>						
$\mu$	99.880	0.016	99.880	0.023	99.880	0.023
$\alpha_1$	-1.088	0.016	-2.088	0.023	-3.088	0.023
$\alpha_2$	0.726	0.016	1.726	0.023	2.726	0.023
$b_1$	-1.066	0.018	-2.066	0.026	-3.066	0.026
$b_2$	0.856	0.018	1.856	0.026	2.856	0.026
$\alpha_{12}$	1.141	0.016	2.141	0.023	3.141	0.023
$c_{11}$	-1.049	0.019	-2.005	0.026	-3.004	0.026
$c_{12}$	1.163	0.019	2.207	0.026	3.207	0.026
$c_{21}$	-0.977	0.021	-2.111	0.026	-3.111	0.026
$c_{22}$	-0.834	0.021	-1.969	0.026	-2.969	0.026
$d_{121}$	-0.978	0.018	-1.978	0.026	-2.978	0.026
$d_{122}$	-0.826	0.018	-1.826	0.026	-2.826	0.026
Type I	0.063		0.000		0.000	
Type II	0.000		0.000		0.000	

**Πίνακας Δ.4:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 1. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III), περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$ , αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	99.998	0.024	99.998	0.024	99.998	0.024
$\alpha_1$	1.113	0.024	2.114	0.024	3.114	0.024
$b_1$	1.103	0.024	2.103	0.024	3.103	0.024
$c_{11}$	1.094	0.025	2.094	0.024	3.094	0.024
Type I	0.001		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE II</b>						
$\mu$	99.998	0.024	99.998	0.026	99.998	0.026
$\alpha_1$	1.113	0.024	2.113	0.026	3.113	0.026
$\alpha_2$	0.911	0.024	1.911	0.026	2.911	0.026
$b_1$	1.102	0.024	2.102	0.026	3.102	0.026
$\alpha_{12}$	1.004	0.024	2.004	0.026	3.004	0.026
$c_{11}$	1.093	0.024	2.093	0.026	3.093	0.026
$c_{21}$	0.910	0.025	1.972	0.026	2.910	0.026
$d_{121}$	1.015	0.024	2.015	0.026	3.015	0.026
Type I	0.011		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE III</b>						
$\mu$	99.998	0.013	99.998	0.026	99.983	0.026
$\alpha_1$	1.138	0.014	2.113	0.026	3.113	0.026
$\alpha_2$	0.819	0.013	1.911	0.026	2.911	0.026
$b_1$	1.039	0.014	2.156	0.027	3.156	0.027
$b_2$	0.858	0.014	1.920	0.027	2.920	0.027
$\alpha_{12}$	1.004	0.014	2.004	0.026	3.004	0.026
$c_{11}$	1.123	0.013	2.061	0.027	3.061	0.027
$c_{12}$	0.923	0.014	1.967	0.027	2.967	0.027
$c_{21}$	0.941	0.014	1.979	0.027	2.979	0.027
$c_{22}$	1.106	0.014	2.166	0.027	3.166	0.027
$d_{121}$	1.004	0.014	2.032	0.027	3.032	0.027
$d_{122}$	0.525	0.015	1.735	0.027	2.735	0.027
Type I	0.060		0.000		0.000	
Type II	0.000		0.000		0.000	

**Πίνακας Δ.5:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 1. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III), περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$ , αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	99.998	0.125	100.026	0.131	100.005	0.127
$\alpha_1$	1.114	0.023	2.114	0.023	3.114	0.023
$b_1$	-0.897	0.023	-1.897	0.023	-2.897	0.023
$c_{11}$	-0.906	0.023	-1.906	0.023	-2.906	0.023
Type I	0.071		0.000		0.000	
Type II	0.083		0.000		0.000	
<b>CASE II</b>						
$\mu$	99.998	0.110	99.998	0.130	99.998	0.138
$\alpha_1$	-0.815	0.020	-1.815	0.025	-2.815	0.025
$\alpha_2$	0.982	0.020	1.982	0.025	2.982	0.025
$b_1$	-0.897	0.019	-1.897	0.025	-2.897	0.025
$\alpha_{12}$	1.004	0.019	2.004	0.025	3.004	0.025
$c_{11}$	-0.859	0.020	-1.859	0.025	-2.859	0.025
$c_{21}$	0.920	0.020	1.920	0.025	2.920	0.025
$d_{121}$	-0.913	0.020	-1.913	0.025	-2.913	0.025
Type I	0.011		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE III</b>						
$\mu$	99.997	0.014	99.998	0.026	99.998	0.026
$\alpha_1$	-0.670	0.016	-1.886	0.026	-2.886	0.026
$\alpha_2$	0.912	0.016	1.911	0.026	2.912	0.026
$b_1$	-0.806	0.015	-1.844	0.027	-2.844	0.027
$b_2$	0.849	0.015	1.920	0.027	2.919	0.027
$\alpha_{12}$	-1.004	0.015	-1.996	0.026	-2.995	0.027
$c_{11}$	-0.918	0.014	-1.939	0.027	-2.939	0.027
$c_{12}$	0.954	0.016	1.967	0.027	2.966	0.027
$c_{21}$	-1.030	0.016	-2.021	0.027	-3.021	0.027
$c_{22}$	1.133	0.016	2.166	0.027	3.166	0.027
$d_{121}$	-1.159	0.022	-1.968	0.027	-2.968	0.027
$d_{122}$	0.001	0.001	1.735	0.027	2.735	0.027
Type I	0.071		0.000		0.000	
Type II	0.083		0.000		0.000	



**Πίνακας Δ.6:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται τρεις παράγοντες ελέγχου και τρεις παράγοντες θορύβου για το Μοντέλο 2. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	99.920	0.123	99.920	0.123	99.920	0.123
$\alpha_1$	1.173	0.023	2.173	0.024	3.173	0.024
$\alpha_1^2$	1.155	0.146	2.157	0.147	3.157	0.147
$b_1$	0.923	0.026	1.992	0.026	2.992	0.026
$c_{11}$	1.215	0.026	2.215	0.026	3.215	0.026
Type I	0.000		0.000		0.	
Type II	0.000		0.000		0.	
<b>CASE II</b>						
$\mu$	99.900	0.112	99.900	0.169	99.900	0.169
$\alpha_1$	1.173	0.017	2.173	0.026	3.173	0.026
$\alpha_2$	0.835	0.017	1.835	0.026	2.835	0.026
$\alpha_1^2$	1.129	0.115	2.129	0.220	3.130	0.220
$\alpha_2^2$	1.052	0.115	2.052	0.220	3.052	0.220
$b_1$	0.923	0.018	1.923	0.028	2.923	0.028
$\alpha_{12}$	0.756	0.020	1.874	0.028	3.874	0.028
$c_{11}$	1.216	0.018	2.216	0.028	3.216	0.028
$c_{21}$	0.974	0.018	1.974	0.028	2.974	0.028
$d_{121}$	1.303	0.021	2.184	0.028	2.184	0.028
Type I	0.095		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE III</b>						
$\mu$	99.900	0.141	99.900	0.180	99.930	0.180
$\alpha_1$	1.173	0.021	2.173	0.027	3.173	0.027
$\alpha_2$	0.835	0.021	1.835	0.027	2.835	0.027
$\alpha_1^2$	1.129	0.183	2.129	0.234	3.129	0.234
$\alpha_2^2$	1.052	0.183	2.052	0.234	3.052	0.234
$b_1$	0.891	0.027	1.941	0.033	2.941	0.033
$b_2$	0.903	0.027	1.953	0.033	2.953	0.033
$\alpha_{12}$	0.875	0.023	1.875	0.030	2.875	0.030
$c_{11}$	1.119	0.026	2.119	0.033	3.119	0.033
$c_{12}$	1.044	0.026	2.044	0.033	3.044	0.033
$c_{21}$	0.959	0.026	1.959	0.033	2.959	0.033
$c_{22}$	1.289	0.026	2.289	0.033	3.289	0.033
$d_{121}$	1.349	0.030	1.200	0.033	3.200	0.033
$d_{122}$	1.096	0.030	1.947	0.033	2.947	0.033
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	

**Πίνακας Δ.7:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται τρεις παράγοντες ελέγχου και τρεις παράγοντες θορύβου για το Μοντέλο 2. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	99.921	0.123	99.921	0.131	99.921	0.123
$\alpha_1$	-0.827	0.024	-1.827	0.023	-2.827	0.024
$\alpha_1^2$	1.157	0.147	2.157	0.147	3.157	0.147
$b_1$	-1.077	0.026	-2.077	0.026	-3.077	0.026
$c_{11}$	1.216	0.026	2.216	0.026	3.216	0.026
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE II</b>						
$\mu$	99.900	0.112	99.900	0.169	99.900	0.169
$\alpha_1$	-0.827	0.017	-1.827	0.026	-2.827	0.026
$\alpha_2$	0.835	0.017	1.835	0.026	2.835	0.026
$\alpha_1^2$	-0.871	0.145	-1.871	0.220	-2.871	0.220
$\alpha_2^2$	-0.948	0.145	-1.948	0.220	-2.948	0.220
$b_1$	0.923	0.018	1.923	0.028	2.923	0.028
$\alpha_{12}$	0.756	0.021	1.874	0.028	2.875	0.028
$c_{11}$	-0.784	0.018	-1.784	0.028	-2.784	0.028
$c_{21}$	-1.026	0.018	-2.026	0.028	-3.026	0.028
$d_{121}$	1.302	0.021	2.184	0.028	3.184	0.028
Type I	0.095		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE III</b>						
$\mu$	99.990	0.141	99.990	0.180	99.990	0.180
$\alpha_1$	1.173	0.021	2.173	0.027	3.173	0.027
$\alpha_2$	0.835	0.021	1.835	0.027	2.835	0.027
$\alpha_1^2$	-0.871	0.183	-1.871	0.234	-2.871	0.234
$\alpha_2^2$	-0.948	0.183	-1.948	0.234	-2.948	0.234
$b_1$	0.891	0.027	1.941	0.033	2.941	0.033
$b_2$	0.903	0.027	1.953	0.033	2.953	0.033
$\alpha_{12}$	-1.125	0.023	-2.125	0.030	-3.125	0.030
$c_{11}$	-0.881	0.026	-1.881	0.033	-2.881	0.033
$c_{12}$	-0.956	0.026	-1.956	0.033	-2.956	0.033
$c_{21}$	0.959	0.026	1.959	0.033	2.959	0.033
$c_{22}$	1.289	0.026	2.289	0.033	3.289	0.033
$d_{121}$	1.349	0.030	2.200	0.033	3.200	0.033
$d_{122}$	1.096	0.030	1.947	0.033	2.946	0.033
Type I	0.059		0.000		0.000	
Type II	0.000		0.000		0.000	

**Πίνακας Δ.8:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 2. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	100.000	0.071	100.056	0.059	100.059	0.059
$\alpha_1$	0.920	0.015	1.724	0.018	2.723	0.018
$\alpha_1^2$	0.002	0.312	1.128	0.169	2.117	0.169
$b_1$	0.847	0.017	1.974	0.013	2.970	0.013
$c_{11}$	0.878	0.016	1.942	0.011	2.938	0.011
Type I	0.031		0.008		0.008	
Type II	0.200		0.000		0.000	
<b>CASE II</b>						
$\mu$	100.115	0.089	100.337	0.111	100.068	0.112
$\alpha_1$	0.920	0.019	1.920	0.024	2.920	0.021
$\alpha_2$	1.109	0.019	2.109	0.024	3.109	0.021
$\alpha_1^2$	0.001	0.001	0.001	0.001	2.212	0.259
$\alpha_2^2$	1.985	0.108	3.715	0.134	3.830	0.259
$b_1$	0.834	0.020	1.899	0.025	2.899	0.022
$\alpha_{12}$	1.012	0.020	2.012	0.025	3.012	0.022
$c_{11}$	0.878	0.020	1.878	0.025	2.878	0.022
$c_{21}$	1.076	0.020	2.076	0.025	3.076	0.022
$d_{121}$	0.935	0.020	1.935	0.025	2.935	0.022
Type I	0.011		0.000		0.000	
Type II	0.100		0.099		0.000	
<b>CASE III</b>						
$\mu$	100.115	0.080	100.337	0.118	99.983	0.119
$\alpha_1$	0.975	0.018	1.920	0.025	2.920	0.023
$\alpha_2$	1.165	0.018	2.109	0.025	3.109	0.023
$\alpha_1^2$	0.002	0.000	0.001	0.001	2.212	0.275
$\alpha_2^2$	1.985	0.097	3.715	0.143	3.829	0.275
$b_1$	0.790	0.020	1.890	0.028	2.890	0.025
$b_2$	0.834	0.020	1.958	0.028	2.958	0.025
$\alpha_{12}$	1.012	0.018	2.011	0.027	3.012	0.024
$c_{11}$	0.896	0.019	1.896	0.028	2.896	0.024
$c_{12}$	1.162	0.019	2.163	0.028	3.163	0.025
$c_{21}$	1.044	0.019	2.044	0.028	3.043	0.025
$c_{22}$	0.910	0.019	1.909	0.028	2.909	0.025
$d_{121}$	1.017	0.020	1.943	0.028	2.944	0.025
$d_{122}$	0.994	0.025	2.042	0.028	3.042	0.025
Type I	0.034		0.000		0.000	
Type II	0.071		0.077		0.000	

**Πίνακας Δ.9:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 2. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	100.065	0.125	100.056	0.056	100.060	0.056
$\alpha_1$	-0.775	0.022	-1.859	0.012	-2.859	0.012
$\alpha_1^2$	0.001	0.122	1.127	0.160	2.115	0.160
$b_1$	-0.881	0.022	-2.006	0.024	-3.004	0.024
$c_{11}$	0.944	0.022	1.941	0.011	2.938	0.011
Type I	0.094		0.083		0.083	
Type II	0.200		0.000		0.000	
<b>CASE II</b>						
$\mu$	99.950	0.087	99.950	0.087	99.950	0.100
$\alpha_1$	0.982	0.016	1.977	0.016	2.920	0.019
$\alpha_2$	-0.820	0.016	-1.816	0.016	-2.891	0.019
$\alpha_1^2$	-2.138	0.232	-3.085	0.233	-3.788	0.231
$\alpha_2^2$	1.404	0.232	2.375	0.233	3.349	0.231
$b_1$	0.825	0.232	1.821	0.017	2.899	0.020
$\alpha_{12}$	0.935	0.017	1.931	0.017	3.012	0.020
$c_{11}$	-1.046	0.017	-2.041	0.017	-3.122	0.020
$c_{21}$	0.999	0.017	1.995	0.017	3.077	0.020
$d_{121}$	-1.061	0.017	-2.056	0.017	-2.995	0.021
Type I	0.032		0.044		0.011	
Type II	0.000		0.000		0.000	
<b>CASE III</b>						
$\mu$	99.997	0.104	100.337	0.118	100.068	0.120
$\alpha_1$	-1.080	0.022	-2.080	0.025	-3.080	0.023
$\alpha_2$	1.109	0.022	2.109	0.025	3.109	0.023
$\alpha_1^2$	0.500	0.001	0.500	0.001	2.212	0.275
$\alpha_2^2$	1.985	0.126	3.715	0.143	3.830	0.275
$b_1$	-1.110	0.024	-2.110	0.028	-2.110	0.025
$b_2$	-1.042	0.024	-2.042	0.027	-3.042	0.025
$\alpha_{12}$	-0.988	0.023	-1.988	0.028	-2.988	0.024
$c_{11}$	0.896	0.024	1.896	0.028	2.896	0.025
$c_{12}$	1.163	0.024	2.163	0.028	3.163	0.025
$c_{21}$	1.044	0.024	2.044	0.028	3.044	0.025
$c_{22}$	-1.091	0.024	-2.091	0.028	-3.091	0.025
$d_{121}$	0.944	0.024	1.944	0.028	2.944	0.025
$d_{122}$	1.042	0.024	2.042	0.028	3.042	0.025
Type I	0.000		0.000		0.000	
Type II	0.071		0.071		0.000	

**Πίνακας Δ.10:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 3. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	99.998	0.024	100.026	0.024	99.998	0.024
$\alpha_1$	1.113	0.024	2.114	0.024	3.114	0.024
$b_1$	1.103	0.024	2.103	0.024	3.103	0.024
$c_{11}$	1.093	0.024	2.094	0.024	3.094	0.024
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE II</b>						
$\mu$	99.998	0.023	99.998	0.025	99.998	0.025
$\alpha_1$	1.114	0.023	2.114	0.025	3.114	0.025
$\alpha_2$	0.911	0.023	1.911	0.025	2.911	0.025
$b_1$	1.103	0.023	2.103	0.025	3.103	0.025
$\alpha_{12}$	1.004	0.023	2.004	0.025	3.004	0.025
$c_{11}$	1.094	0.023	2.094	0.025	3.094	0.025
$c_{21}$	0.910	0.024	1.972	0.025	2.972	0.025
Type I	0.026		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE III</b>						
$\mu$	99.998	0.027	99.998	0.027	99.983	0.027
$\alpha_1$	1.113	0.027	2.114	0.027	3.114	0.027
$\alpha_2$	0.911	0.027	1.910	0.027	2.910	0.027
$b_1$	1.103	0.027	2.103	0.027	3.103	0.027
$b_2$	0.926	0.027	1.926	0.027	2.926	0.027
$\alpha_{12}$	1.004	0.027	2.004	0.027	3.004	0.027
$c_{11}$	1.061	0.028	2.061	0.028	3.061	0.028
$c_{12}$	0.967	0.028	1.967	0.028	2.967	0.028
$c_{21}$	0.979	0.028	1.979	0.028	2.979	0.028
$c_{22}$	1.166	0.028	2.165	0.028	3.166	0.028
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	

**Πίνακας Δ.11:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 3. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	99.998	0.024	99.998	0.024	99.998	0.024
$\alpha_1$	-0.886	0.024	-1.886	0.024	-2.886	0.024
$b_1$	1.103	0.024	2.103	0.024	3.103	0.024
$c_{11}$	-0.906	0.024	-1.906	0.024	-2.906	0.024
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE II</b>						
$\mu$	99.998	0.025	99.998	0.139	100.028	0.025
$\alpha_1$	-0.886	0.025	-1.886	0.025	-2.886	0.025
$\alpha_2$	0.911	0.025	1.911	0.025	2.911	0.025
$b_1$	1.102	0.025	2.102	0.025	3.102	0.025
$\alpha_{12}$	-0.996	0.025	-1.996	0.025	-2.996	0.025
$c_{11}$	1.094	0.025	2.094	0.025	3.094	0.025
$c_{21}$	-1.028	0.025	-2.028	0.025	-3.028	0.025
Type I	0.001		0.000		0.000	
Type II	0.000		0.000		0.000	
<b>CASE III</b>						
$\mu$	99.997	0.027	99.997	0.027	99.997	0.027
$\alpha_1$	-0.886	0.027	-1.886	0.027	-2.886	0.027
$\alpha_2$	0.911	0.027	1.911	0.027	2.911	0.027
$b_1$	-0.897	0.027	-1.897	0.027	-2.897	0.027
$b_2$	-1.074	0.027	-2.074	0.027	-3.074	0.027
$\alpha_{12}$	1.004	0.027	2.004	0.027	3.004	0.027
$c_{11}$	1.061	0.028	2.061	0.028	3.061	0.028
$c_{12}$	-1.033	0.028	-2.033	0.028	-3.033	0.028
$c_{21}$	-1.021	0.028	-2.021	0.028	-3.021	0.028
$c_{22}$	-0.834	0.028	-1.834	0.028	-2.834	0.028
Type I	0.000		0.000		0.000	
Type II	0.000		0.000		0.000	

**Πίνακας Δ.12:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 4. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	100.001	0.059	100.223	0.068	100.252	0.097
$\alpha_1$	0.920	0.013	1.920	0.014	2.919	0.021
$\alpha_1^2$	0.001	0.001	0.001	0.001	2.818	0.117
$b_1$	0.899	0.013	1.899	0.015	2.899	0.022
$c_{11}$	0.948	0.014	1.948	0.016	2.878	0.022
Type I	0.087		0.087		0.000	
Type II	0.200		0.200		0.000	
<b>CASE II</b>						
$\mu$	100.115	0.068	100.337	0.108	100.068	0.078
$\alpha_1$	0.920	0.015	1.920	0.023	2.920	0.015
$\alpha_2$	1.109	0.015	2.109	0.023	3.109	0.015
$\alpha_1^2$	0.001	0.001	0.001	0.001	2.212	0.180
$\alpha_2^2$	1.985	0.083	3.715	0.132	3.830	0.180
$b_1$	0.899	0.015	1.899	0.024	2.899	0.016
$\alpha_{12}$	0.939	0.016	2.012	0.024	2.939	0.016
$c_{11}$	0.951	0.016	1.878	0.024	2.951	0.016
$c_{21}$	1.003	0.016	2.076	0.024	3.003	0.016
Type I	0.007		0.000		0.071	
Type II	0.111		0.111		0.000	
<b>CASE III</b>						
$\mu$	100.115	0.087	100.337	0.113	100.068	0.093
$\alpha_1$	0.920	0.019	1.920	0.024	2.920	0.018
$\alpha_2$	1.109	0.019	2.109	0.024	3.109	0.018
$\alpha_1^2$	0.001	0.001	0.001	0.001	2.212	0.215
$\alpha_2^2$	1.985	0.106	3.715	0.137	3.830	0.215
$b_1$	0.899	0.020	1.899	0.025	2.899	0.019
$b_2$	0.947	0.020	1.947	0.025	2.947	0.019
$\alpha_{12}$	1.012	0.020	2.012	0.025	3.012	0.019
$c_{11}$	0.896	0.021	1.896	0.026	2.896	0.019
$c_{12}$	1.163	0.021	2.163	0.026	3.163	0.019
$c_{21}$	1.044	0.021	2.044	0.026	3.044	0.019
$c_{22}$	0.909	0.021	1.909	0.026	2.909	0.019
Type I	0.026		0.000		0.051	
Type II	0.083		0.083		0.000	

**Πίνακας Δ.13:** Αποτελέσματα 1000 αριθμητικών πειραμάτων όπου θεωρούνται πέντε παράγοντες ελέγχου και πέντε παράγοντες θορύβου για το Μοντέλο 4. Η περίπτωση I (Case I), η περίπτωση II (Case II), και η περίπτωση III (Case III) περιγράφονται στην ενότητα 4 και το  $\beta = 1, 2, 3$  υποδεικνύει το μέγεθος των συντελεστών με το πρόσημο του  $\beta_i$  να θεωρείται τυχαία ως θετικό ή αρνητικό. Οι αριθμοί στις στήλες  $E(\beta_i|\beta_i \neq 0)$  και  $V(\beta_i|\beta_i \neq 0)$  υποδεικνύουν τον μέσο όρο και τη διακύμανση των εκτιμήσεων των παραμέτρων  $\beta_i$  αντίστοιχα, δεδομένου ότι  $\beta_i \neq 0$ . Τα σφάλματα Τύπου I και Τύπου II παρουσιάζονται στις δύο τελευταίες γραμμές κάθε περίπτωσης.

	$\beta = 1$		$\beta = 2$		$\beta = 3$	
	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$	$E(\beta_i \beta_i \neq 0)$	$V(\beta_i \beta_i \neq 0)$
<b>CASE I</b>						
$\mu$	100.001	0.059	100.222	0.068	100.252	0.097
$\alpha_1$	-1.080	0.013	-2.080	0.015	-3.080	0.021
$\alpha_1^2$	0.001	0.001	0.001	0.001	2.818	0.117
$b_1$	-1.101	0.013	-2.101	0.015	-3.101	0.022
$c_{11}$	0.948	0.014	-1.948	0.016	-2.878	0.022
Type I	0.087		0.087		0.000	
Type II	0.200		0.200		0.000	
<b>CASE II</b>						
$\mu$	100.030	0.068	99.807	0.108	100.068	0.078
$\alpha_1$	-1.080	0.015	-2.080	0.023	-3.080	0.015
$\alpha_2$	1.109	0.015	2.109	0.023	3.109	0.015
$\alpha_1^2$	-1.912	0.082	-3.642	0.130	-3.788	0.180
$\alpha_2^2$	0.001	0.001	0.001	0.001	-2.170	0.180
$b_1$	0.899	0.015	1.899	0.024	2.899	0.016
$\alpha_{12}$	-1.061	0.016	-1.988	0.024	-3.061	0.016
$c_{11}$	-1.049	0.016	-2.122	0.024	-3.049	0.016
$c_{21}$	1.003	0.016	2.076	0.024	3.003	0.016
Type I	0.071		0.000		0.071	
Type II	0.111		0.111		0.000	
<b>CASE III</b>						
$\mu$	100.115	0.081	100.337	0.113	99.997	0.101
$\alpha_1$	-1.080	0.017	-2.080	0.024	-3.080	0.019
$\alpha_2$	-0.891	0.017	-1.891	0.024	-2.891	0.019
$\alpha_1^2$	0.001	0.001	0.001	0.001	2.212	0.232
$\alpha_2^2$	1.985	0.098	3.715	0.137	3.830	0.232
$b_1$	-1.101	0.018	-2.101	0.025	-3.101	0.020
$b_2$	-1.053	0.018	-2.053	0.025	-3.053	0.020
$\alpha_{12}$	1.012	0.018	2.012	0.025	3.012	0.020
$c_{11}$	0.896	0.019	1.896	0.026	2.896	0.021
$c_{12}$	-0.837	0.019	-1.837	0.026	-2.837	0.021
$c_{21}$	1.043	0.019	2.044	0.026	3.044	0.021
$c_{22}$	-1.091	0.019	-2.091	0.026	-3.091	0.021
Type I	0.051		0.000		0.026	
Type II	0.083		0.083		0.000	











---

## Βιβλιογραφία

---

- [1] ABRAHAM, B., CHIPMAN, H. & VIJAYAN, K. (1999). Some Risks in the Construction and Analysis of Supersaturated Designs. *Technometrics*, **41** (2), 135-141.
- [2] AGGARWAL, M.L. & GUPTA, S. (2004). A new method of construction of multi-level supersaturated designs. *J. Statist. Plann. Inference*, **121**, 127-134.
- [3] ANTONIADIS, A. & FAN, J. (2001). Regularized wavelet approximations (with discussion). *Journal American Statistical Association* 2001; **96** (455), 939-967.
- [4] ATKINSON, A. & DONEV, A. (1992). *Optimum Experimental Designs*. Oxford Science Publications, Oxford.
- [5] BATES, R. A., BUCK, R. J., RICCOMAGNO, E. & WYNN, H. P. (1996). Experimental design and observation for large systems (with discussion). *J. Royal Statist. Soc. Ser. B*, **58**, 77-94.
- [6] BEATTIE, S.D., FONG, D.K.H. & LIN, D.K.J. (2002). A two-stage Bayesian model selection strategy for supersaturated designs. *Technometrics*, **44**, 55-63.
- [7] BELLONI, A., CHERNOZHUKOV, V. & WANG, L. (2011). Square-root lasso: pivotal recovery of sparse signals via conic programming. *Biometrika*, **98**, 791-806.
- [8] BINGHAM, D. R. & SITTER, R. R. (1999). Minimum- Aberration Two-Level Fractional Factorial Split-Plot Designs. *Technometrics*, **41**, 62-70.
- [9] BINGHAM, D. R. & SITTER, R. R. (2001). Design Issues in Fractional Factorial Split-Plot Experiments. *Journal of Quality Technology*, **33**, 2-15.

- [10] BINGHAM, D. & SITTER, R. R. (2003). Fractional Factorial Split-Plot Designs for Robust Parameter Experiments. *Technometrics*, **45**, 80-89.
- [11] BINGHAM, D., SITTER, R.R & TANG, B. (2009). Orthogonal and nearly orthogonal designs for computer experiments. *Biometrika*, **96**, 51-65.
- [12] BISGAARD, S. (2000). The Design and Analysis of  $2^{k-p} \times 2^{q-r}$  Split Plot Experiments. *Journal of Quality Technology*, **32**, 39-56.
- [13] BOOTH, K.H.V. & COX, D.R. (1962). Some systematic supersaturated designs. *Technometrics*, **4**, 489-495.
- [14] BORWEIN, P. B. & FERGUSON, R. A. (2003). A complete description of Golay pairs for lengths up to 100. *Math. Comp.*, **73**, 967-985.
- [15] BOX, G. E. P., HUNTER, W. G. & HUNTER, J. S. (1978). *Statistics for Experimenters, An Introduction to Design, Data Analysis, and Model Building*, Wiley, New York.
- [16] BOX, G. & JONES, S. (1992). Split-Plot Designs for Robust Product Experimentation. *Journal of Applied Statistics*, **19**, 3-26.
- [17] BOX, G. E. P. & MEYER, R. D. (1986). Analysis for unreplicated fractional factorials. *Technometrics*, **28**, 11-18.
- [18] BULUTOGLU, D.A. & CHENG, C.S. (2004). Construction of  $E(s^2)$ -optimal supersaturated designs. *Annals of Statistics*, **32**, 1662-1678.
- [19] BULUTOGLU, D.A. & RYAN, K.J. (2008).  $E(s^2)$ -optimal supersaturated designs with good minimax properties when N is odd. *J. Statist. Plann. Inference*, **138**, 1754-1762.
- [20] BUTLER, N.A., MEAD, R., ESKRIDGE, K.M. & GILMOUR, S.G. (2001). A general method of constructing  $E(s^2)$ -optimal supersaturated designs. *Journal Royal Statistical Society B*, **63**, 621-632.
- [21] CANDÈS, E.J. & TAO, T. (2007). The Dantzig selector: statistical estimation when  $p$  is much larger than  $n$ . *Annals of Statistics*, **35** (6), 2313-2351.

- [22] CHAI, F.S., CHATTERJEE, K., DAS, A. & MIDHA, C. (2012). Optimal supersaturated designs for  $s^m$  factorials in  $N \not\equiv 0 \pmod{s}$  runs. *J. Stat. Theory Pract.*, **6**, 1170-7177.
- [23] CHANG, C.-C. & LIN, C.-J. (2001). LibSVM: a library for support vector machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [24] CHENG, C.S. (1997).  $E(s^2)$ -optimal supersaturated designs. *Statistica Sinica*, **7**, 929-939.
- [25] CHEN, J. & LIU, M.-Q. (2008a). Optimal mixed-level supersaturated designs with general number of runs. *Statistics and Probability Letters*, **78**, 2496-2502.
- [26] CHEN, J. & LIU, M.-Q. (2008b). Optimal mixed-level k-circulant supersaturated designs. *J. Statist. Plann. Inference*, **138**, 4151-4157.
- [27] CHEN, R.B., WENG, J.Z. & CHU, C.H. (2013). Screening procedure for supersaturated designs using a Bayesian variable selection method. *Quality and Reliability Engineering International*, **29**, 89-101.
- [28] CORTES, C. & VAPNIK V. (1995). Support vector networks. *Machine Learning*, **20** (3), 273-297.
- [29] CRAVEN, P. & WAHBA G. (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerical Mathematics*, **31**, 377-403.
- [30] DAS, U., GUPTA, S. & GUPTA, S. (2014). Screening active factors in supersaturated designs. *Computational Statistics and Data Analysis*, **77**, 223-232.
- [31] DEJAEGHER, B. & VANDER HEYDEN, Y. (2008). Supersaturated Designs: Setups, Data Interpretation, and Analytical Applications. *Annals of Bioanalytical Chemistry*, **390** (5), 1227-1240.
- [32] DEY, A. & MUKERJEE, R. (1999). *Fractional Factorial Plans*. John Wiley, New York.

- [33] DJOKOVIC, D. Z. & KOTSIREAS, I. S. (2013). Some new periodic Golay pairs. Arxiv:1310.5773.
- [34] DONOHO, DL.(2000). High-dimensional data analysis: the curses and blessings of dimensionality. In *Proceedings of the American Mathematical Society Conference on Mathematical Challenges of the 21st Century* Los Angeles, CA.
- [35] DRAPER, N.R. & SMITH H.(1981). *Applied Regression Analysis*. New York: John Wiley.
- [36] DROU, K., KOUKOUVINOS, C. & LAPP, A. (2017). Screening Active Effects in Supersaturated Designs with Binary Response via Control Charts. *Quality and Reliability Engineering International*, **33** (7) DOI: 10.1002/qre.2119.
- [37] DRUCKER, H., BURGESS, C.J.C., KAUFMAN, L., SMOLA, A. & VAPNIK, V. (1997). Support vector regression machines. *Advances in Neural Information Processing Systems*, **9**, 155-161.
- [38] DUAN, K. & RAJAPAKSE, J.C. (2005). SVM-RFE peak selection for cancer classification with mass spectrometry data. *Proceedings of the third Asia-Pacific Bioinformatics-Conference (APBC)*, Eds., Phoebe Chen, Y.-P., Wong, L., 191-200, World Scientific, Singapore.
- [39] EFRON, B., HASTIE, T., JOHNSTONE, I. & TIBSHIRANI, R. (2004). Least angle regression (with discussion). *The Annals of Statistics*, **32**, 407-499.
- [40] ERRORE, A., JONES, B., LI, W. & NACHTSHEIM, C. J. (2017). Using Definitive Screening Designs to Identify Active First- and Second-Order Factor Effects. *Journal of Quality Technology*, **49** (3), 244-264.
- [41] FAN, J., FENG, Y., SALDANA, D.F. & SAMWORTH, R., WU Y. (2016). SIS: (Iterative) Sure Independence Screening for Generalized Linear Models and Cox's Proportional NHazards Models. R package version 0.7-4 2016; URL <http://CRAN.R-project.org/package=SIS>.
- [42] FAN, J. & LI, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, **96**, 1348-1360.



- [43] FAN, J. & LI, R. (2006). Statistical challenges with high dimensionality: feature selection in knowledge discovery. In *Proceedings of the International Congress of Mathematicians*, Sanz-Sole M, Soria J, Varona JL, Verdera J (eds). Freiburg: European Mathematical Society, Madrid (Spain), **III**, 595–622.
- [44] FAN, J. & LV, J.(2008). Sure Independence Screening for Ultrahigh Dimensional Feature Space. *Journal of the Royal Statistical Society B*, **70** (5), 849–911.
- [45] FANG, K.T., GE, G., LIU, M.Q. & QIN, H. (2004). Combinatorial constructions for optimal supersaturated designs. *Discrete Math.*, **279**, 191–202.
- [46] FANG, K.T., LI, R., & SUDJIANTO, A. (2006). *Design and Modeling for Computer Experiments*. New York: Chapman and Hall/CRC.
- [47] FANG, K.T., LIN, D.K.J. & LIU, M.Q. (2003). Optimal mixed-level supersaturated design. *Metrika*, **58** (3), 279–291.
- [48] FANG, K. T., LIN, D.K.J. & MA, C. X. (2000). On the construction of multi-level supersaturated designs. *Journal of Statistical Planning and Inference*, **86**, 239–252.
- [49] FANG, J. & TAI, D. (2011). Evaluation of Mutual Information, Genetic Algorithm and SVR for Feature Selection in QSAR Regression. *Current Drug Discovery Technologies*, **8**, 107–111.
- [50] FRANK, I.E. & FRIEDMAN, J. H. (1993). A statistical view of some chemometrics regression tools. *Technometrics*, **35**, 109–148.
- [51] GEORGIU, S.D. (2009). Orthogonal Latin hypercube designs from generalized orthogonal designs. *Journal of Statistical Planning and Inference*, **139**, 1530–1540.
- [52] GEORGIU, S.D. (2014). Supersaturated designs: A review of their construction and analysis. *Journal of Statistical Planning and Inference*, **144**, 92–109.
- [53] GEORGIU, S.D. & EFTHIMIOU, I. (2014). Some classes of orthogonal Latin hypercube designs. *Statist. Sinica*, **24**, 101–120.

- [54] GEORGIU, S.D., STYLIANOU, S., DROU, K. & KOUKOUVINOS, C. (2014), Construction of orthogonal and nearly orthogonal designs for computer experiments. *Biometrika*, **101**, 741-747
- [55] GEORGIU, S. & KOUKOUVINOS, C. (2006). Multi-level k-circulant supersaturated designs. *Metrika*, **64**, 209-220.
- [56] GEORGIU, S., KOUKOUVINOS, C., & LIU, M.-Q. (2014). U-type and column orthogonal designs for computer experiments. *Metrika*, **77**, 1057-1073.
- [57] GEORGIU, S., KOUKOUVINOS, C., & SEBERRY, J. (2004). Generalized orthogonal designs. *Ars Combinatoria*, **71**, 33-47.
- [58] GERAMITA, A.V. & SEBERRY, J. (1979). *Orthogonal Designs: Quadratic Forms and Hadamard Matrices*. Marcel Dekker, New York-Basel.
- [59] GUPTA, S. & KOHLI, P. (2008). Analysis of supersaturated designs: a review. *Journal of Indian Society of Agricultural Statistics*, **62**, 156-168.
- [60] GUPTA, V.K., SINGH, P., KOLE, B. & PARSAD, R. (2009). Construction of efficient unbalanced mixed-level supersaturated designs. *Statistics and Probability Letters*, **79**, 2359-2366.
- [61] GUYON, I., WESTON, J., BARNHILL, S. & VAPNIK, V. (2002). Gene Selection for Cancer Classification Using Support Vector Machines. *Machine Learning*, **46** (1), 389-422.
- [62] HEDAYAT, A. S., SLOANE, N. J. A. & STUFKEN, J. (1999). *Orthogonal Arrays: Theory and Applications*. Springer, New York.
- [63] HERNANDEZ, A.S, LUCAS, T.W., & CARLYLE, M. (2012). Constructing nearly orthogonal Latin hypercubes for any nonsaturated run-variable combination. *ACM Trans. Model. Comput. Simul.* **22**, 1-17.
- [64] HOLCOMB, D.R., MONTGOMERY, D.C. & CARLYLE W.M. (2003). Analysis of Supersaturated Designs. *Journal of Quality Technology*, **35**, 13-27.

- [65] HUANG, M.-L., HUNG, Y.-H., LEE, W. M., LI, R. K. & JIANG B.-R. (2014). SVM-RFE Based Feature Selection and Taguchi Parameters Optimization for Multiclass SVM Classifier. *Scientific World Journal*, Article ID 795624, 10 pages.
- [66] HUANG, T.-M., KECCMAN, V. & KOPRIVA, I. (2006). *Kernel Based Algorithms for Mining Huge Data Sets*. Springer-Verlag, Berlin, Heidelberg.
- [67] HUANG, H., YANG, J. & LIU, M.-Q. (2014). Functionally induced priors for componentwise Gibbs sampler in the analysis of supersaturated designs. *Computational Statistics and Data Analysis*, **72**, 1-12.
- [68] JONES, B. & MAJUMDAR, D. (2014). Optimal Supersaturated Designs. *Journal of the American Statistical Association*, **109** (508), 1592-1600.
- [69] JONES, B. & NACHTSHEIM, C. (2009). Split-Plot Designs: What, Why, and How. *Journal of Quality Technology*, **41**, 340-361.
- [70] JONES, B. & NACHTSHEIM, C.J. (2011). A class of three-level designs for definitive screening in the presence of second-order effects. *Journal of Quality Technology*, **43** (1), 1-14.
- [71] KHARAGHANI, H. (2000). Arrays for orthogonal designs. *J. Combin. Designs*, **8**, 166-173.
- [72] KOEHLER, J. R. & OWEN, A. B. (1996). Computer experiments. in S. Ghosh and C. R. Rao, eds, *Handbook of Statistics*, **13**, 261-308, Elsevier Science B.V., Amsterdam.
- [73] KOH, W. Y., ESKRIDGE, K. M. & HANNA, M.A. (2013). Supersaturated Split-Plot Designs. *Journal of Quality Technology*, **45**, 61-73.
- [74] KOUKOUVINOS, C. & MYLONA, K. (2009). Group screening method for the statistical analysis of  $E - f_{NOD}$ -optimal mixed-level supersaturated designs. *Statistical Methodology*, **6**, 380-388.
- [75] KOUKOUVINOS, C., MYLONA, K. & SKOUNTZOU, A. (2011). A Variable Selection Method for Analyzing Supersaturated Designs. *Communications in Statistics - Simulation and Computation*, **40** (4), 484-496.

- [76] KOUKOUVINOS, C., MYLONA, K. & SIMOS, D.E. (2008).  $E(s^2)$ -optimal and minimax-optimal cyclic supersaturated designs via multi-objective simulated annealing. *Journal of Statistical Planning and Inference*, **138** (6), 1639-1646.
- [77] KOUKOUVINOS, C., MYLONA, K. & SIMOS, D. (2009). A hybrid SAGA algorithm for the construction of  $E(s^2)$ -optimal cyclic supersaturated designs. *Journal of Statistical Planning and Inference* 2008; **139**, 478-485.
- [78] KOUKOUVINOS, C. & SEBERRY, J. (1999). New weighing matrices and orthogonal designs constructed using two sequences with zero autocorrelation function - A review. *Journal of Statistical Planning and Inference*, **81**, 153-182.
- [79] KULAHCI, M. & BISGAARD, S. (2005). The Use of Plackett Burman Designs to Construct Split-Plot Designs. *Technometrics*, **47**, 495-501.
- [80] KUNERT, J., AUER, C., ERDBRUGGE, M. & EWERS, R. (2007). An experiment to compare Taguchi's product array and the combined array. *Journal of Quality Technology*, **39** (1), 17-34.
- [81] LI, R. & LIN, D.K.J. (2002). Data Analysis of Supersaturated Design. *Statistics and Probability Letters*, **59**, 135-144.
- [82] LI, R. & LIN, D.K.J. (2003). Analysis Methods for Supersaturated Design: Some Comparisons. *Journal of Data Science*, **1**, 249-260.
- [83] LI, P.F., LIU, M.Q. & ZHANG, R.C. (2004). Some theory and the construction of mixed-level supersaturated designs. *Statistics & Probability Letters*, **69** (1), 105-116.
- [84] LI, X., ZHAO, T., YUAN, X. & LIU, H. (2015). "The flare Package for High Dimensional Linear Regression and Precision Matrix Estimation in R. *Journal of Machine Learning Research*, **16**, 553-557.
- [85] LI, P., ZHAO, S. & ZHANG, R. (2010). A cluster analysis selection strategy for supersaturated designs. *Computational Statistics and Data Analysis*, **54**, 1605-1612.

- [86] LIU, Y. & DEAN, A. (2004).  $k$ -Circulant supersaturated designs. *Technometrics*, **46**, 32-43.
- [87] LIU, M.Q., FANG, K.T. & HICKERNELL, F.J. (2006). Connections among different criteria for asymmetrical fractional factorial designs. *Statistica Sinica*, **16** (4), 1285-1297.
- [88] LIU, Y. & LIU, M.Q. (2011). Construction of optimal supersaturated design with large number of levels. *Journal of Statistical Planning and Inference*, **141**, 2035-2043.
- [89] LIU, Y. & LIU, M.Q. (2012). Construction of equidistant and weak equidistant supersaturated designs. *Metrika*, **75**, 33-53.
- [90] LIU, Y. & LIU, M.Q. (2013). Construction of supersaturated design with large number of factors by the complementary design method. *Acta Mathematicae Applicatae Sinica*, (English Series), **29**, 253-262.
- [91] LIN, D.K.J. (1993). A New Class of Supersaturated Designs. *Technometrics*, **35** (1), 28-31.
- [92] LIN, D.K.J. (1995). Generating Systematic Supersaturated Designs. *Technometrics*, **37**, 213-225.
- [93] LIN, C.D., BINGHAM, D., SITTER, R.R., & TANG, B. (2010). A new and flexible method for constructing designs for computer experiments. *Ann. Statist.*, **38**, 1460-1477.
- [94] LIN, C.D., MUKERJEE, R., & TANG, B. (2009). Construction of orthogonal and nearly orthogonal Latin hypercubes. *Biometrika*, **96**, 243-247.
- [95] LIU, Y.F. & DEAN, A.M. (2004).  $k$ -circulant supersaturated designs. *Technometrics*, **46**, 32-43.
- [96] LIU, X. & TANG J. (2014). Mass Classification in Mammograms Using Selected Geometry and Texture Features, and a New SVM-Based Feature Selection Method. *Systems Journal IEEE*, **8** (3), 910-920.

- [97] LIU, M. & ZHANG, R. (2000). Construction of  $E(s^2)$ -Optimal Supersaturated Designs Using Cyclic BIBDs. *Journal of Statistical Planning and Inference*, **91**, 139-150.
- [98] LU, X., HU, W. & ZHENG, Y. (2003). A systematic procedure in the construction of multi-level supersaturated design. *J. Statist. Plann. Inference*, **115**, 287-310.
- [99] LU, X. & SUN, Y. (2001). Supersaturated design with more than two levels. *Chinese Ann. Math.*, **22B** (2), 183-194.
- [100] LU, X. & WU, X. (2004). A Strategy of Searching Active Factors in Supersaturated Screening Experiments. *Journal of Quality Technology*, **36**, 392-399.
- [101] MARLEY, C.J. & WOODS, D.C. (2010). A comparison of design and model selection methods for supersaturated experiments. *Computational Statistics and Data Analysis*, **54**, 3158-3167.
- [102] MATSUURA, S., SUZUKI, H., IIDA, T., KURE H. & MORI, H. (2011). Robust Parameter Design Using a Supersaturated Design for a Response Surface Model. *Quality and Reliability Engineering International*, **27**, 541-554.
- [103] MCKAY, M.D., BECKMAN, R.J. & CONOVER, W.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, **21**, 239-245.
- [104] MILLER, A. & SITTER, R.R. (2005). Using Folded-Over Nonorthogonal Designs. *Technometrics*, **47**, 502-513.
- [105] MUNDRA, P.A. & RAJAPAKSE, J.C. (2007). SVM-RFE with relevancy and redundancy criteria for gene selection. Chapter in *Pattern Recognition in Bioinformatics*, Eds. Rajapakse, J.C., Schmidt, B., and Volkert, L.G., Vol. **4774**, 242-252, Springer, Heidelberg.
- [106] MYERS, R.H., KHURI, A.I. & VINING, G.G.(1992). Response surface alternatives to the Taguchi robust parameter design approach. *American Statistician*, **46**(2), 131-139.

- [107] NGUYEN, N.K. (1996). An Algorithmic Approach to Constructing Supersaturated Designs. *Technometrics*, **38**, 69-73.
- [108] NGUYEN, N.K., & CHENG, C.S. (2008). New  $E(s^2)$ -Optimal Supersaturated Designs Constructed from Incomplete Block Designs. *Technometrics*, **50**, 26-31.
- [109] PARK, M.Y. & HASTIE, T. (2007).  $L_1$ -regularization path algorithm for generalized linear models. *Journal of the Royal Statistical Society Series B*, **69** (4), 659-677.
- [110] PARKER, P.A., KOWALSKI, S.M. & VINING, G.G. (2007). Unbalanced and minimal point equivalent estimation second-order split-plot designs. *Journal of Quality Technology*, **39**, 376-388.
- [111] PHOA, F.K.H. (2014). The Stepwise Response Refinement Screener(SRRS). *Statistica Sinica*, **24**, 197-210.
- [112] PHOA, F.K.H., PAN, Y.H. & XU, H. (2009). Analysis of supersaturated designs via the Dantzig selector. *Journal of Statistical Planning and Inference*, **139**, 2362-2372.
- [113] RAVIKUMAR, P., LAFFERTY, J., LIU, H. & WASSERMAN L. (2009). Sparse additive models. *JRSS, Series B*, **71** (5), 1009-1030.
- [114] ROSENBAUM, P. R. (1994). Dispersion effects from fractional factorials in Taguchi's method of quality design. *Journal of the Royal Statistical Society, Series B*, **56**, 641-652.
- [115] ROSENBAUM, P. R. (1996). Some useful compound dispersion experiments in quality design. *Technometrics*, **57**, 488-502.
- [116] SACKS, J., WELCH, W. J., MITCHELL, T. J. & WYNN, H. P. (1989). Design and analysis of computer experiments (with discussion). *Statistica Sinica*, **4**, 409-435.

- [117] SAMB, M.L., CAMARA, F., NDIAYE, S., SLIMANI, Y. & ESSEGHIR, M.A. (2012). A Novel RFE-SVM-based Feature Selection Approach for Classification. *International Journal of Advanced Science and Technology*, **43**, 27-36.
- [118] SANTNER, T. J., WILLIAMS, B. J. & NOTZ, W. I. (2003). *The Design and Analysis of Computer Experiments*. Springer, New York.
- [119] SATTERTHWAITTE, F. (1959). Random balance experimentation (with discussion). *Technometrics*, **1**, 111-137.
- [120] SCHOEN, E., JONES, B., & GOOS, P. (2011). A Split-Plot Experiment with Factor Dependent Whole-Plot Sizes. *Journal of Quality Technology*, **43**, 66-79.
- [121] SCHOLKOPF, B., BURGESS, C.J.C. & SMOLA A.J. (1999). *Advances in Kernel Methods: Support Vector Learning*, MIT Press, Massachusetts.
- [122] SHOEMAKER, A.C., TSUI, K.L & WU, C.F.J. (1991). Economical experimentation methods for robust design. *Technometrics*, **33**(4), 415–427.
- [123] SMOLA, A.J. & SCHOLKOPF, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, **14** (3), 199-222.
- [124] STEINBERG, M. & LIN, D.K.J. (2006). A construction method for orthogonal Latin hypercube designs. *Biometrika*, **93**, 279-288.
- [125] SUEN, C.S. & DAS, A. (2010).  $E(s^2)$ -optimal supersaturated designs with odd number of runs. *J. Statist. Plann. Inference*, **140**, 1398-1409.
- [126] SUN, F.S., LIU, M.Q., & LIN, D.K.J. (2010). Construction of orthogonal Latin hypercube designs with flexible run sizes. *Journal of Statistical Planning and Inference*, **140**, 3236-3242.
- [127] SUN, F.S., LIU, M.Q., & LIN, D.K.J. (2009). Construction of orthogonal Latin hypercube designs. *Biometrika*, **96**, 971-974.
- [128] SUN, F.S., PANG, F., & LIU, M.Q. (2011). Construction of column orthogonal designs for computer experiments. *Sci. China Math.*, **54**, 2683-2692.



- [129] SUN, F. S., LIN, D. K. J. & LIU, M. Q. (2011). On construction of optimal mixed-level supersaturated designs. *Annals of Statistics*, **39**, 1310-1333.
- [130] SUNDBERG, R.(2008) A classical dataset from Williams, and its role in the study of supersaturated designs. *Journal of Chemometrics* **22**, 436-440.
- [131] TAGUCHI, G. (1987). *System of Experimental Design*, Vol. I and II, UNIPUB, New York.
- [132] TANG, B. AND DENG, L.Y. (1999). Minimum  $G_2$ -aberration for nonregular fractional factorial designs. *Ann. Statist.*, **27**, 1914-1926.
- [133] TANG, B. & WU, C.F.J. (1997). A method for constructing supersaturated designs and its  $E(s^2)$  optimality. *Canadian Journal of Statistics*, **25**, 191-201.
- [134] TANG, Y., ZHANG, Y.Q. & HUANG, Z. (2007). Development of two-stage SVM-RFE gene selection strategy for microarray expression data analysis. *IEEE ACM Trans. Comput. Biology Bioinform*, **4** (3), 365-381.
- [135] TIBSHIRANI, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society Series B-Methodological*, **58**, 267-288.
- [136] VAPNIK, V. (2000). *The Nature of Statistical Learning Theory*, 2nd ed. John Wiley and Sons, Inc., New York.
- [137] VINING, G.G. & MYERS R.H. (1990). Combining Taguchi and response surface philosophies: A dual response approach. *Journal of Quality Technology*, **22**(1), 38-45.
- [138] WANG, L. (2013).  $L_1$  penalized lad estimator for high dimensional linear regression. *Journal of Multivariate Analysis*, **120**, 135-151.
- [139] WEESE, M.L., SMUCKER, B.J. & EDWARDS, D.J. (2015). Searching for powerful supersaturated designs. *Journal of Quality Technology*, **47** (1), 66-84.
- [140] WESTFALL, P.H., YOUNG, S.S., & LIN, D.K.J. (1998). Forward selection error control in the analysis of supersaturated designs. *Statistica Sinica*, **8**, 101-117.

- [141] WILLIAMS, K.R. (1968). Designed experiments. *Rubber Age*, **100**, 65-71.
- [142] WU, C. (1993). Construction of Supersaturated Designs Through Partially Aliased Intercations. *Biometrika*, **80**, 661-669.
- [143] WU, C.F.J. & HAMADA, M. (2009). *Experiments: Planning, Analysis, and Parameter Design Optimization, 2nd Ed.*, John Wiley & Sons, Inc., New York.
- [144] XU, H. (2003). Minimum moment aberration for nonregular designs and supersaturated designs. *Statist. Sinica*, **13**, 691-708.
- [145] XU, H. & WU, C.F.J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *Annals of Statistics*, **29**, 1066-1077.
- [146] XU, H. & WU, C.F.J. (2005). Construction of optimal multi-level supersaturated designs. *Annals of Statistics*, **33**, 2811-2836.
- [147] YAMADA, S., IKEBE, Y.T., HASHIGUCHI, H. & NIKI N. (1999). Construction of three-level supersaturated design. *Journal of Statistical Planning and Inference*, **81**, 183-193.
- [148] YAMADA, S. & LIN, D.K.J. (1999). Three-level supersaturated designs. *Statistics and Probability Letters*, **45**, 31-39.
- [149] YAMADA, S. & LIN, D.K.J. (2002). Construction of mixed-level supersaturated design. *Metrika*, **56** (3), 205-214.
- [150] YAMADA, S. & MATSUI, T. (2002). Optimality of mixed-level supersaturated designs. *Journal of Statistical Planning and Inference*, **104** (2), 459-468.
- [151] YAMADA, S., MATSUI, M., MATSUI, T., LIN, D.K.J. & TAKAHASHI, T. (2006). A general construction method for mixed-level supersaturated design. *Computational Statistics & Data Analysis*, **50** (1), 254-265.
- [152] YAN, K. & ZHANG, D. (2015). Feature selection and analysis on correlated gas sensor data with recursive feature elimination. *Sensors and Actuators B: Chemical*, **212**, 353-363.

- [153] YANG, J.Y. & LIU, M.Q. (2012). Construction of orthogonal and nearly orthogonal Latin hypercube designs from orthogonal designs. *Statist. Sinica*, **22**, 433–442.
- [154] YIN, Y.H. & LIU, M.Q. (2013). Orthogonal Latin hypercube designs for Fourier polynomial models. *Journal of Statistical Planning and Inference*, **143**, 307–314.
- [155] YIN, Y.H., ZHANG, Q.Z. & LIU, M.Q. (2013). A two-stage variable selection strategy for supersaturated designs with multiple responses. *Frontiers of Mathematics in China*, **8**, 717–730.
- [156] ZHANG, C.H. (2010). Nearly Unbiased Variable Selection Under Minimax Concave Penalty. *The Annals of Statistics*, **38**, 894–942.
- [157] ZHANG, C.H. & HUANG, J. (2008). The sparsity and bias of the LASSO selection in high-dimensional linear regression. *Ann. Statist.* 2008; **36** (4), 1567–1594.
- [158] ZHANG, Q.Z., ZHANG, R.C. & LIU, M.Q. (2007). A method for screening active effects in supersaturated designs. *Journal of Statistical Planning and Inference*, **137** (9), 235–248.
- [159] ZOU, H. (2006). The adaptive Lasso and its oracle properties. *Journal American Statistical Association*, **101** (476), 1418–1429.
- [160] ZOU, H. & LI, R. (2008). One-step sparse estimates in nonconcave penalized likelihood models. *The Annals of Statistics*, **36** (4), 1509–1533.