ITERATED LOCAL SEARCH APPROACHES FOR MAXIMIN LATIN HYPERCUBE DESIGNS

BY

Aparna Dey

Roll No – 0851551

A Thesis submitted of the partial fulfillment of the requirements for the degree of Master of Philosophy in Mathematics

Khulna University of Engineering & Technology
Khulna 9203, Bangladesh

July, 2012

Declaration

This is to certify that the thesis work entitled "Iterated Local Search Approaches for maximin Latin Hypercube Designs " has been carried out by *Aparna Dey* in the Department of Mathematics, Khulna University of Engineering & Technology, Khulna, Bangladesh. The above thesis work or any part of this work has not been submitted anywhere for the award of any degree or diploma.

.....

Signature of Supervisor
Name: A. R. M. Jalal Uddin Jamali
Designation: Professor,
Depertment of Mathematics

.....

Signature of Candidate
Name: Aparna Dey
Roll No.: 0851551

Approval

This is to certify that the thesis work submitted by *Aparna Dey* entitled " *Iterated Local Search Approaches for maximin Latin Hypercube Designs* " has been approved by the board of examiners for the partial fulfillment of the requirements for the degree of *Master of Philosophy* in the Department of *Mathematics*, Khulna University of Engineering & Technology, Khulna, Bangladesh in July, 2012.

BOARD OF EXAMINERS

1. 26/07/12 — Chairman (Supervisor)

Prof. Dr. A. R. M. Jalal Uddin Jamali
Department of Mathematics
KUET, Khulna-9203

2. — Member

Prof. Dr. Abul Kalam Azad
Head
Department of Mathematics
KUET, Khulna-9203

3. 26-7-2012 — Member

Prof. Dr. Fousia Rahaman
Department of Mathematics
KUET, Khulna-9203

4. 26.7.12 — Member

Prof. Dr.Mohammad Arif Hossain
Department of Mathematics
KUET, Khulna-9203

5. 26/7/12 — Member (External)

Prof. Dr. Subrata Majumdar
Department of Mathematics
Rajshahi University
Rajshahi

Dedication

To my respectable and beloved parents

Krishna Dey and

Adhir Ranjan Dey

Acknowledgement

First of all I would like to articulate my gratitude to the Almighty for His blessings upon me for the successful completion of this thesis paper. Except His special kindness, this thesis paper would not be completed.

I would like to give sincere gratitude and intense appreciation to my thesis supervisor, Dr. A. R. M. Jalal Uddin Jamali, Professor, Department of Mathematics, Khulna University of Engineering & Technology, for his continuous supervision, guidance, support and regular advice to prepare this thesis paper. Without his kind supervision and high encouragement, I could not complete this thesis.

I wish to express my profound gratitude to authority of the Khulna University of Engineering & Technology for accepting me in M. Phil. Course. Its outstanding learning environment helped me lay a firm academic foundation.

I express my deep gratefulness to Prof. Dr. Fouzia Rahman, Prof. Dr. Bazlar Rahman Prof. Dr. Mohammad Arif Hossain, Prof. Dr. Abul Kalam Azad, Dr. M. M. Touhid Hossain and all other members of the Mathematics Department who helped me in different phases of my M. Phil Course. They helped me whenever I needed assistance regarding administrative as well as other academic necessities. Without their continuous encouragement I would still be in my M.Phil dream.

Next with deep gratitude I remember the co-operation and inspiration of my parents that helped me to move forward in my life. I am thankful to them, because they are a source of inspiration for me. Without their co-operation I would not be able to terminate this thesis. I would specially be grateful to my mother who always inspired me for my studies and gave me mental strength in my frustrations. Also thanks to my beloved younger brother Anik Dey and sister Suparna Dey for always being with me.

Finally I do express my thanks to my husband, Mr. Susanta Kumar Das, who helped me time to time in my thesis course and always inspired me for my study and gave me mental strength.

Publication

1. A.R.M. J. U. Jamali, **Aparna Dey**, A. Grosso, M. Locatelli, Correlation Analysis of the Latin Hypercube Designs Obtaining by ILS Approach, International Conference on Mechanical, Industrial an Energy Engineering 2010, 23-24 December 2010, Khulna, BANGLADESH, pp. MIE10-078-1-6

Abstract

Many scientific phenomena are now investigated by experimental design. In the field of experimental design problems we consider maximin Latin Hypercube Designs (LHDs). Problems are formulated as optimization ones. Optimal Latin Hypercube designs has good space-filling as well as non-collapsing properties, is widely used experimental design. Here, we consider Iterated Local search (ILS) heuristic approach to obtain optimized (maximin) LHD in Euclidian distance measure. There is another important property required in experimental design is the *multicollinearity*. Since if two factors are correlated then it will not be possible to distinguish between the effects of the two factors based on this experiment. There are procedures to find good LHDs by minimizing the pair-wise correlation or maximizing the inter-site distances. Several experiments are performed to compare our results with available ones in the literature. Many improvements, regarding maximin LHD, are obtained. Extensive experiments have been also performed to analyze the *multicollinearity* of the optimal LHDs. Some other structural information of the optimal (maximin) LHDs obtained by the ILS approach are analyzed. Finally experimentally it has been shown that the maximin LHD obtained by ILS is the state of arts regarding space-*filling* as well as *non-collapsing* properties and comparable regarding *multicollinearity* property. Moreover the dependency among the factors of the experimental design obtained by ILS approach is negligible for large value of design points.

Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Background

The design of computer experiments has much recent interest and this is likely to grow as more and more simulation models are used to carry out research. Computer simulation experiments are used in a wide range of application to learn about the effect of input variables x on a response of interest y [Butler (2001)]. In computer experiments, instead of physically doing an experiment on the product, mathematical models describing the performance of the product are developed using laws of engineering/physics and solved on computers through numerical methods. As simulation programs are usually deterministic so the output of a computer experiment is not subject to random variations, which makes the design of computer experiments different from that of physical experiments [Fang et al. (2006)]. Many simulation models involved several hundred factors or even more. It is desirable to avoid replicates when projecting the design on to a subset of factors. This is because a few, out of the numerous factors in the system, usually dominate the performance of the product. Thus a good model can be fitted using only these few important factors. Therefore, when projecting the design on to these factors, replication is not required.

As is recognized by several authors, the choice of the design points for computer experiments should at least fulfill two requirements (details can be found in [Johnson et al. (1990), Morris and Mitchell (1995)]. First of all, the design should be *space-filling* in some sense. When no details on the functional behavior of the response parameters are available, it is important to be able to obtain information from the entire design space. Therefore, design points should be *evenly spread* over the entire region. Secondly, the design should be *non-collapsing*. When one of the design parameters has (almost) no influence on the function value, two design points that differ only in this parameter will *collapse*, i.e., they can be considered as the same point that is evaluated twice. For deterministic functions this is not a desirable situation. Therefore, two design points should not share any coordinate value when it is not known a priori which parameters are important. There is another important property for the design of

experiments – the *multicollinearity* property. That is the factors /coordinates should be uncorrelated or they are mutually (approximately) orthogonal. The *multicollinearity* property is also important, because if two factors are correlated then it will not possible to distinguish between the effects of the two factors based on this experiment.

For the design of computer experiments **Latin Hypercube Designs (LHDs)**, first introduced by McKay and his co-authors in 1979, fulfill the *non-collapsing* property. LHDs are important in the design of computer-simulated experiments [Fang et al. (2006)]. Here LHD is defined a bit different than McKay et al. (1979) but similar to Johnson et al. (1990), Husslage et al. (2006), Morris and Mitchell (1995), Grosso et al. (2008). It is assumed that there are N design points have to be placed and each point has k distinct parameters. The points are placed such a way that they are uniformly spread when projected along each single parameter axis. It will be assumed that each parameter range is normalized to the interval [0, N-1]; Then, a LHD is made up by N points, each of which has k integer coordinates with values in 0,1, . . . , N-1 and such that there do not exist two points with one common coordinate value. This allows a non-collapsing design because points are evenly spreaded when projected along a single parameter axis. Note that the number of possible LHDs is huge: there are $(N!)^k$ possible LHDs (where N is number of design point and k is number of factors). A configuration

$$\mathbf{X} = \begin{pmatrix} x_0 \\ \vdots \\ x_{N-1} \end{pmatrix} = \begin{pmatrix} x_{01} & \cdots & x_{0k} \\ \vdots & \cdots & \vdots \\ x_{(N-1)1} & \cdots & x_{(N-1)k} \end{pmatrix}$$

with all $x_{ij} = \{0, 1, \cdots N-1\}$ is a LHD if each column has no duplicate entries. This one-dimensional projective property ensures that there is little redundancy of design points when some of the factors have a relatively negligible effect (sparsity principle).
Unfortunately, randomly generated LHDs almost always show poor *space-filling* properties or / and the factors are highly correlated. On the other hand, *maximin distance* objective based designs have very good *space-filling* properties but often no good projection properties under the Euclidean (L^2) or the Rectangular (L^1) distance. To overcome this shortcoming, Morris and Mitchell (1995) suggested for searching **maximin LHDs** which has both the important properties when looking for "optimal"

designs. The definition of optimal LHDs through the maximin criterion has been proposed in Johnson et al. (1990): given a point-to-point distance metric $d(x_i, x_j); i, j \in I$ (I is the index set), then the maximin LHD problem is to find a LHD such that the minimum point-to-point distance occurring in such configuration is maximized (as large as possible). In the literature the optimal criterion for maximin LHD are defined in several ways [Grosso et al. (2009)] but the main objective is identical i.e. searching the LHD with the maximizing the minimum pair-wise distance. Also different definitions for the distance $d(x_i, x_j)$ are considered in literature; in this work d will be considered in the Euclidean distance, which is one of the most frequently used distance-measure in the applications.

For the presence of combinatorial nature, the number of possible LHDs is very high - $(N!)^k$. For example, to optimize the location of 20 samples in two dimensions, the algorithm has to select the best design from more than 10^{36} possible designs. If the number of variables is increased to 3, the number of possible designs is more than 10^{55}. Consequently, when number of factors and/ or number of design points are large then it requires hundreds of hours by the *brute-force* approach to find out the optimal design. So researchers choose heuristic approaches to find out optimal designs. Here, Iterated Local search (ILS) heuristic approach will be considered to find the optimal (maximin) LHD [Grosso et al. (2009)]. For the optimal criterion the following maximin optimal criteria in Euclidean distance measure will be considered which is similar to Johnson et al. (1990) but a quite different regarding computational effort:

$$\Phi_p(X) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left[\frac{1}{d_{ij}^p} \right]^{\frac{1}{p}} \tag{1.1}$$

Where $d_{ij} = d(x_i, x_j)$ the Euclidian distance between points x_i and x_j and p is a positive integer parameter and which can be computed without the need of detecting and ordering all D_i (pair-wise inter side distance) values which is required in Johnson et al. (1990). This optimal criterion will be denoted as Opt(φ). Under this criterion, LHD Y is better than X if

$$\Phi_p(Y) < \Phi_p(X) \tag{1.2}$$

In Johnson et al. (1990) the definition of maximin optimal criterion is as follows

$$\Phi_p(X) = \sum_{r=1}^{R} \left[\frac{J_r(X)}{D_r(X)^p} \right]^{\frac{1}{p}}, \tag{1.3}$$

where $D_1(X) < D_2(X), \cdots\cdots, D_R(X)$ (pair-wise inter side distances), R is the number of different distances in X. When as authors in Grosso et al. (2008), Grosso et al. (2009) considered another maximin optimal criterion denoted as Opt (D1), which is also considered in Johnson et al. (1990), is given below.

$$\begin{aligned} &max\ D_1(X) \text{ such that} \\ &D_1 = D_1(X) = \min d(x_i, x_j) \quad i \neq j;\ X \in LHD \\ &with\ \min \quad J = \left|(i,j)\right| : d(x_i, x_j) = D_1(X) \end{aligned} \tag{1.4}$$

Under this criteria, LHD Y is better than X if

$$\begin{aligned} &D_1(Y) > D_1(X) \qquad \text{or} \\ &D_1(Y) > D_1(X) \text{ and } J_1(Y) < J_1(X) \end{aligned} \tag{1.5}$$

and so on.

An apparent drawback of the Opt(φ) criterion, for maximin values (maximum D_1 value), is that LHDs with smaller (better) φ_p may have a worse (smaller) D_1 value, i.e. for X and Y, though $\varphi_p(X) < \varphi_p(Y)$ but $D_1(X) < D_1(Y)$. This phenomenon has been frequently observed in the computational experiments [Grosso et al. (2009)]. Nevertheless, a profitable choice is to work in order to minimize the φ_p function, but at the same time, keep track of the best (D_1, J_1) values observed during such minimization. This way the search in the solution space is guided by a kind of heuristic function. Such a mixed approach might appear strange but, as it will be demonstrated experimentally, it can be extremely effective. Such objective will be denoted as Opt(φ, D1).

Anyway in multiple regression problems, there is an expectation to find dependencies between the response variable y, and the regressor variables x_i (factors of the design points). In the most regression problems, however, it is found that there are also dependencies among the factors of the designed points x_i. In the situations where these dependencies are strong, it is said that *multicollinearity* exists. *Multicollinearity* can have serious effects on the estimates of the regression coefficients and on the general

applicability of the model. As *multicollinearity* measures the linear dependency among the factors of the design points, so *multicollinearity* can be measured by the partial pair-wise correlations among the factors. There are several ways available in the literature to measure the pair-wise correlations i.e. *multicollinearity.* Here, the following measure of average pair-wise correlations is considered

$$\rho^2 = \frac{\sum_{i=2}^{k}\sum_{j=1}^{i-1}\rho_{ij}^2}{k(k-1)/2}. \tag{1.6}$$

Here ρ_{ij} denote the partial-correlation between the factors i and j and k denotes the number of factors in the design considered. Note that this definition is frequently used in literature like Fang et al. (2000b). Also definition of maximum pair-wise correlation is given below:

$$\rho_{\max} = \max_{1\le i,j\le k} \rho_{ij} \tag{1.7}$$

1.2 Literature Review

1.2.1 Experimental Designs

Since physical experiments are inevitably very expensive and time consuming, computer experiments are widely used for simulating physical characteristics and for the design and development of products (for examples, [Fang et al. (2006)]). A computer experiment is modeled as a realization of a stochastic process, often in the presence of nonlinearity and high dimensional inputs. In order to perform efficient data analysis and prediction and in order to determine the best settings for a number of design parameters that have an impact on the response variable(s) of interest and which influence the critical quality characteristics of the product or process, it is often necessary to set a good design as well as to optimize the product or process design. In computer experiments, instead of physically doing an experiment on the product, mathematical models describing the performance of the product are developed using laws of engineering/physics. Then the mathematical models are solved on computers through numerical methods such as the finite element method. A computer simulation of the mathematical models is usually time-consuming and there is a great variety of possible input combinations. For these reasons meta-models, Barthelemy and Haftka (1993), Sobieski and Haftka (1997) model with the quality characteristics as explicit

functions of the design parameters, are constructed. Such a meta-model, also called a (global) approximation model or surrogate model, is obtained by simulating a number of design points. Since a meta-model evaluation is much faster than a simulation run, in practice such a meta-model is used, instead of the simulation model, to gain insight into the characteristics of the product or process and to optimize it. Therefore, a careful choice of the design points at which performing simulations in order to build the meta-model is of primary importance.

As it is recognized by several authors, the choice of the design points for computer experiments should at least fulfill two requirements (details can be found in Johnson et al. (1990) and Morris and Mitchell (1995)). First of all, the design should be space-filling in some sense. When no details on the functional behavior of the response parameters are available, it is important to be able to obtain information from the entire design space. Therefore, design points should be evenly spread over the entire region. Secondly, the design should be non-collapsing. When one of the design parameters has (almost) no influence on the function value, two design points that differ only in this parameter will collapse, i.e., they can be considered as the same point that is evaluated twice. For deterministic functions this is not a desirable situation. Therefore, two design points should not share any coordinate value when it is not known a priori which parameters are important.

The latter requirement is fulfilled by employing Latin Hypercube Designs (LHDs). Such designs, proposed by McKay and his co-authors (1979), are evenly distributed in each one-dimensional projection and are thus non-collapsing. Unfortunately, randomly generated LHDs almost always show poor space-filling properties. On the other hand, maximin distance designs, proposed by Johnson, Moore and Ylvisaker (1990), have very good space-filling properties but often no good projection properties under the Euclidean or the Rectangular distance. To overcome this shortcoming, Morris and Mitchell (1995) suggested searching for maximin LHDs when looking for "optimal" designs. Although the search for maximin LHDs will be one of the problems discussed in this thesis, it will be important to point out that also other definitions of "optimality" for designs exist in the literature. These are not discussed in detail throughout the thesis (detail can be found in Santner et al. (2003)), but, for the sake of completeness, in the following literature review some of them will be mentioned, together with a short

discussion of the methods employed to return "optimal" (according to the selected definition) designs.

Fang et al. (2000a); Fang et al. (2000b) defined a uniform design as a design that allocates experimental points uniformly scattered on the domain. Uniform designs do not require being orthogonal. They considered projection uniformity over all sub dimensions. In Fang et al. (2000b) they classified uniform designs as space-filling designs.

Lee and Jung (2000) proposed maximin eigen value sampling, that maximizes minimum eigen value, for Kriging model where maximin eigen value sampling uses eigen values of the correlation matrix. The Kriging model is obtained from sampled points generated by the proposed method. The Kriging model is used to compare the characteristics of proposed sampling design with those of maximum entropy sampling.

The maximin design problem has also been studied in location theory. In this area of research, the problem is usually referred to as the max-min facility dispersion problem; facilities are placed such that the minimal distance to any other facility is maximal. Again, the resulting solution is certainly space-filling, but not necessarily non-collapsing.

In statistical environments Latin Hypercube sampling is often used. In such an approach, points on the grid are sampled without replacement, thereby deriving a random permutation for each dimension (detail can be found in McKay et al. (1979)).

Giunta and his co-authors (2003) gave an overview of pseudo- and quasi-Monte Carlo sampling, Latin hypercube sampling, orthogonal array sampling, and Hammersley (1960) sequence sampling.

McKay and his co-authors (1979), Stein (1987) and Owen (1994) had shown that LHDs perform much better than completely randomized designs. More recently, algorithms have been used to construct systematic LHDs under various optimality criteria. A LHD always has non-collapsing properties but not necessarily good space-filling property. In particular, as already remarked, randomly generated LHDs often show poor space-

filling properties. Therefore, the search for "optimal" LHDs has attracted attention (detail can be found in Morris and Mitchell (1995), Park (1994), Ye (1998), Ye et al. (2000)). Different optimality criteria for LHDs have been proposed, including maximum entropy designs Shewry and Wynn (1987), Currin et al. (1991)], Integrated Mean Squared Error (IMSE) of prediction and minimax and maximin distance designs [Johnson et al. (1990)].

Lin and Steinberg (2006) proposed several methods for extending the uniform sampling to higher dimensions. The method has also been used to construct LHDs with low correlation of first-order and second-order terms. It generates orthogonal LHDs that can include many more factors than those proposed by Ye (1998).

Cioppa (2002) in his dissertation developed a set of experimental designs by considering orthogonal Latin hypercube and uniform designs to create designs having near orthogonality and excellent space-filling properties. Multiple measures were used to assess the quality of candidate designs and to identify the best one.

Morris (1991) made it clear that many simulation models involve several hundred factors or even more. Consequently, factor screening is useful in computer experiments for reducing the dimension of the factor space before carrying out more detailed experiments. Butler (2001) proposed optimal and orthogonal LHDs which is suitable for factor screening.

Olsson (2003) suggested Latin Hypercube sampling as a tool to improve the efficiency of different importance sampling methods for structural reliability analysis. Liefvendahl and Stocki (2006) proposed probabilistic search algorithm, namely Column-wise Pair-wise (CP) search algorithms and Genetic algorithms to construct optimal LHDs. For the optimal criterion they considered energy function (the sum of the norms of the repulsive forces if the samples are considered as electrically charged particles) as proposed by Audze and Eglais (1977). To improve the reliability, Stocki considered the pairwise correlation. Liefvendahl and Stocki (2006) also compared the performance of the CP and genetic algorithms for optimal LHDs.

By using the Latin Hypercube sampling method, Hwan Yang performed the uncertainty and sensitivity analysis for the time-dependent effects in concrete structure. The results of the Latin Hypercube simulations were used to determine which of the model parameters are most significant in affecting the uncertainty of the design. For each sample, a time-dependent structural analysis was performed to produce response data, which were then analyzed statistically.

Wang (2003) used the Latin Hypercube Design (LHD) instead of the Central Composite Designs (CCD), for improvement of Adaptive Response Surface Method (ARSM). Note that ARSM was developed to search for the global design optimum for computation-intensive design problems. Also note that Response Surface Method (RSM) plans a group of design alternatives and performs the design analysis and simulation simultaneously on these design alternatives. Then an approximation model, called a response surface, is constructed.

1.2.2. Optimal Criteria and Approaches

Some literature reviews will be presented here regarding optimal criteria as well as the solution approaches regarding experimental design. As the complexity of the problem is, to the authors' knowledge, open (but suspected to be NP-complete [Grosso et al. (2008)]. So several heuristics approaches (rather than exact optimization methods) have been proposed in the literature to detect optimal experimental designs.

Fang and his co-authors (2000a) considered Simulated Annealing approach to detect maximin LHD. Li and Wu (1997) proposed a class of algorithms based on column pair-wise exchange to build supersaturated designs. In Ye et al. (2000) an exchange algorithm for finding approximate maximin LHDs has been proposed with the further restriction to Symmetric LHDs (SLHDs). A general formulae for maximin LHDs with $k = 2$ are given by Dam and his co-authors (2007a) with the 1-norm (L^1) and infinite norm (L^∞) distances. Morover, for the Euclidean distance maximin LHDs up to $N = 1000$ design points are obtained by (adapted) periodic designs, while, using a branch-and-bound algorithm, exact solutions have been obtained for N up to 70. Inspired by Dam et al. (2007a), Husslage et al. (2006) proposed (adapted) periodic designs and

simulated annealing to extend the known results and construct approximate maximin latin hypercube designs for k up to 10 and N up to 100. All these designs are available in the website http:// www.spacefillingdesigns.nl. In Husslage et al. (2006), it has been shown that the periodic heuristic tends to work when the number N of design points gets above some threshold which depends on the dimension k of the design (more precisely), such threshold increases with k. Viana and his co-authors (2010) proposed the translational propagation algorithm, a new method for obtaining optimal or near optimal Latin hypercube designs (LHDs) without using formal optimization. For the optimal criterion they also considered Opt(ϕ) to maximin LHD. Monte Carlo simulations were used to evaluate the performance of the algorithm for different design configurations where both the dimensionality and the point density were studied. Gross and his co-authors (2008) successfully implemented Iterated local search (ILS) approach for finding maximin LHDs for k = 3, 4, . .10, and N = 3, …, 100. For the optimal criterion they considered maximin LHDs with Opt(D1, J1) and Opt(ϕ) optimal criteria with Euclidian distance measure (Eq. (1.1) to . Eq. (1.4)).

Dam and his co-authors (2007b) proposed some bounds, for the separation distance of certain classes of maximin LHDs, which are useful for assessing the quality of approximate maximin LHDs. By using some of the special properties of LHDs, they were able to found new and tighter bounds for maximin LHDs. Besides these bounds, they presented a method to obtain a bound for three-dimensional LHDs that is better than Baer's bound for many values of N. They also constructed maximin LHDs attaining Baer's bound for infinitely many values of N in all dimensions.

Morris and Mitchell (1995) proposed the maximin distance criterion which maximizes the minimum distance between design points. Morris and Mitchell (1995) adopted a simulated annealing to find approximate maximin LHDs for up to five dimensions and up to 12 design points, and a few larger values, with respect to the ℓ^1- and ℓ^2-distance measure. In Morris and Mitchell's algorithm, a search begins with a randomly chosen LHD, and proceeds through examination of a sequence of designs, each generated as a perturbation of the preceding one. A perturbation D_{try} of a design D is generated by interchanging two randomly chosen elements within a randomly chosen column in D. The perturbation D_{try} will replace D if it leads to an improvement. Otherwise, it will replace D with probability $\pi = \exp[-\{\phi(D_{try}) - \phi(D)\}/t]$, where t is the preset

parameter known as the "temperature "and ϕ is some measure of the quality of the design. Li and Wu (1997) considered a class of Column-wise Pair-wise (CP) algorithms in the context of the construction of optimal supersaturated designs. A CP algorithm makes exchanges on the columns in a design and can be particularly useful for designs that have structure requirements on the columns. Note that each column in a LHD is a permutation of $\{0, \ldots, N-1\}$. At each step, another permutation of $\{0, \ldots, N-1\}$ is chosen to replace a column so that the LHD structure is retained.

Husslage et al. (2005) constructed nested maximin designs in two dimensions. They showed that different types of grids should be considered when constructing nested designs and discussed how to determine which grid is the best for a specific computer experiment.

Using (adapted) periodic designs and simulated annealing, Husslage et al. (2006) extended the known results and construct approximate maximin Latin hypercube designs for up to ten dimensions and for up to 100 design points. All these designs can be downloaded from http://www.spacefillingdesigns.nl. Inspired by the paper of Morris and Mitchell (1995), in which authors show that LHDs often have a nice periodic structure, Husslage et al. (2006) developed adapted periodic designs. By considering periodic and adapted periodic designs, approximate maximin LHDs for up to seven dimensions and for up to 100 design points are constructed. They have shown that the periodic heuristic tends to work well even for a small number N of design points at low values of the dimension k, but as k increases the periodic heuristic tends to get better than other approaches like simulated annealing only at large N values.

In the simulated annealing algorithm, Husslage et al. (2006) considered four different neighborhoods. In all four neighborhoods the main idea is to change two points of the current LHD by exchanging one or more coordinate values. In three of the four neighborhoods, one point is required to be a critical point (a critical point is a point which is at separation distance, i.e., at a distance equal to the minimal one, from one of the other points). In the first neighborhood, one point j_1 is selected randomly from all critical points and the other point j_2 randomly from all remaining points. This implies that the second point can either be a critical or noncritical point. Once the points are selected, the number of coordinates to change is randomly selected. Due to symmetry,

at most $\lfloor k/2 \rfloor$ coordinates are changed. Subsequently, the coordinates to change are randomly selected. The values of the two points in these coordinates are then exchanged, which results in a new LHD. The second neighborhood is very similar to the first. The only difference is that always one coordinate is selected instead of a random number of coordinates. Note that for $k = 3$ the two neighborhoods are the same. In the third neighborhood, also one coordinate is changed, however, now the coordinate is not randomly selected. Instead, all coordinates are tried and the one which results in the neighbor with the largest separation distance is selected. If more coordinates result in the same separation distance, the one with the lowest index is selected. The fourth neighborhood is again very similar to the second neighborhood. The difference is that the first point is randomly selected from all points, instead of only the critical points. Although simulated annealing algorithms have been used before to deal with this type of problem, this adapted neighborhood structure, which is based on critical points, and the use of a different objective function, turned out to work well.

Jin [Jin et al. (2005)] proposed an enhanced stochastic evolutionary algorithm for finding maximin LHDs. They also applied their method to other space-filling criteria, namely the optimal entropy and centered L_2 discrepancy criteria.

Dam et al. (2007a) derived general formulas for two-dimensional maximin LHDs, when the distance measure is ℓ^∞ or ℓ^1, while for the ℓ^2-distance measure (approximate) maximin LHDs up to 1000 design points are obtained by using a branch-and-bound algorithm and constructing (adapted) periodic designs.

Stinstra and his coauthhors (2003) proposed sequential heuristic algorithms for constrained maximin designs by considering high number of design sites with small volume of feasible design space and other constraints. They also used their methods in many practical situations.

It is remarked that the maximin criterion is not the only one used in the literature. Other criteria are the maximum entropy [Shewry and Wynn (1987)], the integrated mean squared error, the minimum correlation between components [Owen (1994)] and a mixed criterion involving both maximin distance and correlation [Joseph and Hung

(2008)]. For more details the book [Santner et al. (2003)] will be referred but for the completeness, in the following literature review, some articles will be mentioned in which criteria related to correlation are considered.

Dam van (2005) derived interesting results for two-dimensional minimax LHDs. Bates Bates and his co-authors (2004) proposed a permutation genetic algorithm to find optimal Audze-Eglais LHDs. Crary and his co-authors (2000) developed I-OPTTMto generate LHDs with minimal IMSE.

Iman and Conover (1982a) proposed a design by minimizing a linear correlation criterion for pairwise factors. This is modified into a polynomial canonical correlation criterion by Tang (1998). Tang (1998) proposed a LHD by the extension of the concept of Iman and Conover (1982a), namely minimizing a polynomial canonical correlation criterion for pair-wise factors.

Park (1994) constructed optimal LHDs in which IMSE and entropy optimization criteria were considered. To construct optimal LHDs, Park presented an approach based on the exchanges of several pairs of elements in two rows. His algorithm first selects some active pairs which minimize the objective criterion value by excluding that pair from the design. Then, for each chosen pair of two points i_1 and i_2, the algorithm considers all possible exchanges $\chi_{i1j1} \leftrightarrow \chi_{i2j1}, \ldots, \chi_{ikj1} \leftrightarrow \chi_{ikj1}$ for $k \leq l$ and find the best exchange among them.

Ye (1998) constructed orthogonal LHDs in order to enhance the utility of LHDs for regression analysis. Ye defined an Orthogonal Latin Hypercube (OLHC) as a Latin Hypercube for which every pair of columns has zero correlation. Furthermore, in Ye's OLHC construction, the element-wise square of each column has zero correlation with all other columns, and the element-wise product of every two columns has zero correlation with all other columns. These properties ensure the independence of estimates of linear effects of each variable and the estimates of the quadratic effects and bilinear interaction effects are uncorrelated with the estimates of the linear effects.

Joseph and Hung (2008) proposed a multi-objective optimization approach to find good LHDs by combining correlation and distance performance measure. They proposed a

modified simulated annealing algorithm with respect to Morris and Mitchell (1995). Instead of randomly choosing a column and two elements within that column, Morris and Mitchell (1995) chose them judiciously in order to achieve improvement in their multi-objective function.

Ye and his co-authors (2000) and Li and Kenny (2009) proposed an exchange algorithm for finding approximate optimal LHDs, but they consider symmetric Latin hypercube designs (SLHDs). The symmetry property is used as a compromise between computing effort and design optimality. However, one important change had made to accommodate the special structure of SLHD. For a SLHD two simultaneous pair exchanges were made in each column to retain the symmetry. Ye and his co-authors (2000) considered maximin as an optimal criterion, whereas Li and Kenny (2009) considered both the maximin and the entropy optimal criterion.

Fang and his co-authors (2000a) proposed threshold accepting heuristic approaches for optimal LHDs to produce low discrepancy designs compared to theoretic expectation and variance. They considered centered L_2-discrepancy for optimizing the designs.

Sebastiani and Wynn (2000) considered maximum entropy sampling criterion for the optimal Bayesian experimental design. The main contribution of this paper is the extension of the MES principle for the estimation of the problems. Currin and his co-authors (1991) also considered an entropy-based design criterion for Bayesian prediction of deterministic functions. Crombecq and his co-authors (2011) considered space-filling and non-collapsing sequential design strategies for simulation based modeling.

Xu Hongquan (1999) introduced the concept of universal optimality from optimum design theory into computer experiments, and then exhibited some universally optimal designs with respect to different distance measures. He showed that Latin Hypercube and saturated orthogonal arrays are universally optimal with respect to Hamming distance, and that universally optimal designs with respect to Lee distance are also derived from Latin Hypercubes and saturated orthogonal arrays.

Recently Jourdan and Franco (2010) proposed a space-filling LHD design, where they considered a new optimal criterion called Kullback–Leibler criterion. This Kullback–Leibler criterion is relatively very new proposed by Jourdan and Franco (2009). The new designs are compared with several traditional optimal Latin hypercube designs.

Leary et al. proposed orthogonal-array-based LHDs for obtaining better space-filing property. As an optimal criterion, they considered the sum of (square of) reverse inter-site distances. Ye (1998) constructed orthogonal LHDs in order to enhance the utility of LHDs for regression analysis. Author defined an Orthogonal Latin Hypercube (OLHC) as a Latin Hypercube for which every pair of columns has zero correlation. Furthermore, in Ye's OLHC construction, the element-wise square of each column has zero correlation with all other columns, and the element-wise product of every two columns has zero correlation with all other columns. These properties ensure the independence of estimates of linear effects of each variable and the estimates of the quadratic effects and bilinear interaction effects are uncorrelated with the estimates of the linear effects.

Steinberg and Dennis (2006) constructed LHDs in which all main effects are orthogonal. Their method can also be used to construct LHDs with low correlation of first-order and second-order terms. It also generates orthogonal LHDs that can include many more factors than those proposed by Ye (1998). Butler (2001) proposed optimal and orthogonal LHDs which are suitable for factor screening. Fang and his co-authors (2000a) proposed threshold accepting heuristic approaches for optimal LHDs to produce low discrepancy designs compared to theoretic expectation and variance. They considered centered L^2-discrepancy for optimizing the designs.

On the other hand Joseph and Hung (2008) shew that maximization of inter-site distances criteria and minimizing the pair-wise correlation criteria need not necessarily agree with each other. In fact, maximization of inter-site distances can result in LHDs where the variables are highly correlated and vice versa. But it has been already discussed above that for the present of highly correlation, the design has failed to analysis individual effect of the factors.

1.3 Goals of the Thesis

After the invention of hi-speed computer the design of computer experiments is likely to grow as more and more simulation models to carry out research. Many simulation models involve several hundred factors or even more. It is desirable to avoid replicates when projecting the design on to a subset of factors. This is because a few, out of the numerous factors in the system, usually dominate the performance of the product. Thus a good model can be fitted using only these few important factors. Therefore, when projecting the design on to these factors, replication is not required. The experimental design should fulfill three important properties – Non-collapsing, Space-filling, and non-multicollinearity. Latin Hypercube Design (LHD) has good non-collapsing property. But randomly generated LHD often has poor space-filling. So researchers seek LHD with good space-filling property. Many researchers have shown that optimal LHD mainly maximin LHD has good space-filling including non-collapsing property. But recently some researchers have shown that maximin LHD are highly correlated among the factors i.e. there exist *multicollinearity*. As it is mentioned earlier that the *multicollinearity* property is also important, because if two factors are correlated then it will not possible to distinguish between the effects of the two factors based on this experiment. Several approaches existed in literature to find out the maximin LHD such as Simulated annealing, Tabu search, Iterated Local Search (ILS) etc. In the paper of Grosso et al. (2008), authors have shown that ILS approach able to find out a remarkable improved optimal experimental design (maximin LHD) regarding available one in the literature. As m*ulticollinearity* can have serious effects on the estimates of the regression coefficients and on the general applicability of the model. In this study, the *multicollinearity* among the factors of the design obtained by the ILS approach is investigated. The main goal of the study is pointed out as follow:

(1) Implements the ILS approach in windows environments. Note that the approaches are successfully implemented in Sun Operating environment.
(2) Compare experimental results regarding maximin distance with available one in the literature in Euclidian distance measure (L^2- measure).
(3) Compare experimental results regarding maximin distance with available one in the literature in Rectangular distance measure (L^1- measure).
(4) Analysis the average correlation among the factors of the design regarding Euclidian distance measure (L^2- measure).

(5) Analysis the maximum correlation among the factors of the design regarding Euclidian distance measure (L^2- measure).

(6) Analysis the average correlation among the factors of the design regarding Rectangular distance measure (L^1- measure).

(7) Analysis the maximum correlation among the factors of the design regarding Rectangular distance measure (L^1- measure).

(8) Compare the several properties of the design regarding good experimental design with available one in the literature.

1.4 Structure of the Thesis

After the introduction which is in this **Chapter**, the remaining thesis is organized as follows:

Chapter 2 will present the overview of experimental design. Mainly this chapter will point out the roll of statistics, experimental error, basic principles of experimental design, sample design versus experimental design etc. Moreover this chapter will mention briefly the requirements of a good experiment as well as discusses several types of the experimental design issues.

Chapter 3 will also present an overview of correlation. This chapter will point out several types of correlations and how they measure. Here the effects of multicollinearity on the model as how the occur has mentioned here briefly.

In **Chapter 4** the heuristic approach mainly Iterated Local Search (ILS) approach will be discussed. ILS approach for optimizing LHD is elaborately presented here as well as the several optimal criteria for maximizes the minimum inter-site distance of the design point.

In **Chapter 5,** the experimental analyses will be performed extensively regarding Euclidian distance measure. At first the performance of the algorithm is compared with available one in the literature regarding inter-site Euclidian distance

measure. From the experimental design it is shown that the algorithm is state-of-arts regarding maximin LHD. Then these optimal designs are considered for the multicollinearity analysis of the factors of each design.

In **Chapter 6** there several experiments will be performed regarding Rectangular distance measure. Then several characteristics of the designs are compared with available one in the literature. The coefficient of correlations are compared multicollinearity analysis are discussed in this chapter.

Finally **Chapter 7** will contain the detail discussions, concluding remarks and recommendations for possible future extensions of the present works.

References will be included last of the thesis as well and publications are mentions before the index of the thesis.

CHAPTER 2

OVERVIEW OF EXPERIMENTAL DESIGNS

2.1 Introduction

An experiment is a well defined act or an investigation conducted to discover the underlying facts about a phenomenon which are utilized to test some hypotheses of interest, to verify the results of pervious investigation or to study the effect of new conditions on the system. Mainly an experiment is the process of data collection from a non-existent population to get answer to certain problems under investigation. Experimental methods are widely used in research as well as in industrial settings, however, sometimes for very different purposes. The primary goal in scientific research is usually to show the statistical significance of an effect that a particular factor exerts on the dependent variable of interest. In general, every machine used in a production process allows its operators to adjust various settings, affecting the resultant quality of the product manufactured by the machine. Experimentation allows the production engineer to adjust the settings of the machine in a *systematic* manner and to learn which factors have the greatest impact on the resultant quality. Using this information, the settings can be constantly improved until optimum quality is obtained.

Computer modeling is having a profound effect on scientific research. Many processes are so complex that physical experimentation is too time consuming or too expensive; or, as in the case of weather modeling, physical experiments may simply be impossible. As a result, experimenters have increasingly turned to mathematical models to simulate these complex systems. Advances in computational power have allowed both greater complexity and more extensive use of such models. Computer models (or codes) often have high- dimensional inputs, which can be scalars or functions. The output may also be multivariate. In particular, it is common for the output to be a time-dependent function from which a number of summary responses are extracted. For simplicity here, it will be assumed that interest is focused on a relatively small set of scalar inputs, x, and on a single scalar response, y. Making a number of runs at various input configurations, which is called a computer experiment. The design problem is the choice of inputs for efficient analysis of the data.

In the design of complex systems, computer experiments are frequently the only practical approach to obtaining a solution. Typically, a simulation model of system performance is constructed based on knowledge of how the system operates. Performance measures are specified to be incorporated into optimization criteria and constraints, and the design parameters which affect performance are identified. The design solution method depends on the computational demands of the simulation model. In the simplest case, the simulation model may be used directly to calculate performance measures and optimize the system. If a performance measure is not straightforward to calculate, such as one that involves an integral, then sampling via computer experiments may be employed to estimate the measure. If the simulation model is computationally expensive, then the optimization may instead rely on a meta-model, i.e., a mathematical model surrogate of system performance, to approximate the relationship between system performance and the design parameters. In meta-modeling, there are two basic tasks that must be conducted: (1) select a set of sample points in the design parameter space (i.e., an experimental design); (2) fit statistical model(s) of performance to the sample points. Methods for the first task may be used to conduct sampling in general.

Typically, a meta-model is constructed based on data generated from a complex deterministic simulation of the system in which the random variation that exists in the real system is not represented. The uncertainty in the system is modeled by simulating external noise and internal variation in the input variables. Design decisions, then, are based on system analysis and evaluation by approximating the system performance using the constructed meta-model. The primary objectives of meta-modeling are to obtain an accurate estimate of the response and to minimize the required computational effort. This includes minimizing the necessary number of sample points and utilizing a computationally efficient modeling method. In addition, an important item underlying both tasks is the issue of performance evaluation and optimization of the system.

When the simulation of the system is stochastic, then performance measures may be considered that involve expected values (means) and/or variances. If the stochastically is over a finite set of elements, then exact calculations of these types of measures is possible (as in the case of a deterministic simulation). However, if the stochastic

components are continuous (or infinite), then a measure requires evaluation of an integral and is, thus, difficult to compute. In this case, a sampling approach, such as via Monte Carlo simulation is warranted.

2.2 The Roll of Statistics

These deterministic computer experiments differ substantially from the physical experiments performed by agricultural and biological scientists of the early 20th century. Their experiments had substantial random error due to variability in the experimental units. Relatively simple models were often successful. The remarkable methodology for design of experiments introduced by Fisher (1935) and the associated analysis of variance is a systematic way of separating important treatment effects from the background noise (as well as from each other). Fisher's stress on blocking, replication and randomization in these experiments reduced the effect of random error, provided valid estimates of uncertainty, and preserved the simplicity of the models.

The above deterministic examples also differ from codes in the simulation literature, which incorporate substantial random error through random number generators. It has been natural, therefore, to design and analyze such stochastic simulation experiments using standard techniques for physical experiments.

Apparently, McKay, Beckman and Conover (1979) were the first to explicitly consider experimental design for deterministic computer codes. They introduced Latin hypercube sampling, an extension of stratified sampling which ensures that each of the input variables has all portions of its range represented. Latin hypercubes are computationally cheap to generate and can cope with many input variables.

Despite some similarities to physical experiments, then, the lack of random (or replication) error leads to important distinctions. In deterministic computer experiments:

- The adequacy of a response-surface model fitted to the observed data is determined solely by systematic bias.
- The absence of random error allows the complexity of the computer model to emerge.

- Usual measures of uncertainty derived from least squares residuals have no obvious statistical meaning. Though deterministic measures of uncertainty are available (e.g., max | $\hat{y}$ (x) - y (x) | over x and a class of y's), they may be very difficult to compute.
- Classical notions of experimental unit, blocking, replication and randomization are irrelevant.

While the pioneering work of Box and Draper (1959) has relevance to the first of these points, it is unclear that current methodologies for the design and analysis of physical experiments are ideal for complex, deterministic computer models. Lest the reader wonder whether statistics has any role here, it is asserted that:

- The selection of inputs at which to run a computer code is still an experimental design problem.
- Statistical principles and attitudes to data analysis are helpful however the data are generated.
- There is uncertainty associated with predictions from fitted models, and the quantification of uncertainty is a statistical problem.
- Modeling a computer code as if it were a realization of a stochastic process, the approach taken below, gives a basis for the quantification of uncertainty and a statistical framework for design and analysis.

2.3 Experimental Error

Experimental error refers to the variation among the observations of the experimental units treated alike. Since the results of an experiment are affected both by the action of treatments under study and by various extraneous variation, so all out efforts are made to control or remove the effects of all external and extraneous factors. In spite of best precaution, there is some amount of uncontrolled variation in experimental results, which are caused by factors which have not been or cannot be controlled by the experimenter. This variation which is unexplained is called experimental error and the unexplained factors are called uncontrolled nuisance factors. In fact, experimental error represents the variation due to uncontrolled factors and random chance factors. In other words, the variation which is not controlled in an experiment is called experimental error. Experimental error arises from two main sources:

i) Inherent variation exists among the experimental units to which the treatments are applied such as soil fertility varies from one part of land to another, raw materials vary in quality, merit and intellectual capacity differs from person to person of same age, observers and measuring devices vary, change occurs in day to day in a week or other environmental condition. Hence some inevitable variation exists among the results of an experiment.

ii) Variation also arises from any lack of uniformity in the physical conduct of experiment.

However experimental error is composed of (a) observation errors, (b) measurement errors, (c) errors of experimentation (d) Inherent variation among the experimental units (e) joint effects of all other influencing factors which have been ignored in the investigation. Experimental error is used as yardstick in test of hypothesis whether observed treatment differences are real or apparent. In an agricultural experiment variation in yield arises from the difference in soil fertility and amount of irrigation, the type of manure, method of cultivation can be kept constant, the difference in fertility cannot be controlled beyond certain extent.

It is essential to consider the manner in which experimental data are collected as the choice of proper technique of data analysis considerably depends on the method of data collection. Thus it is highly desirable to use proper experimental design which will enable us to separate the effects of interest from the uncontrolled variation. An experiment without proper planning often provides useless and irrelevant data. So emphasis is given on proper planning and design of experiment in order to ensure valid inferences.

2.4 Experimental Design

Experimental design is the plane used in experimentation. More specifically, experimental design is the formulation of a set of rules and principles according to which an experiment is to be conducted to collect appropriate data whose analysis will lead to valid inferences for the problem under investigation. In fact, experimental design consists of the following steps:

i) Choosing a set of treatments for comparison

ii) Selection of experimental units to which chosen treatments will be applied

iii) Specification of the number of experimental units for inclusion in the experiment.

iv) Specification of the method of allocating the treatments to the experimental unit.

v) Specification of the grouping of experimental units to control extraneous sources of variation.

The design of deterministic computer experiments has been partly addressed in the literature. For example, Welch (1983) and references mentioned therein have considered nonparametric systematic departures from regression models. Random error is also included, but the resulting sampling-variance contribution to mean squared error can be set to zero, and these approaches have helped shape our formulation. For the most part, however, the designs used for fitting predictors have been those developed for physical experiments. Such designs typically have appealing features of symmetry and are often optimal in one or more senses in settings which include random noise. Their appropriateness for computer experiments, however, is by no means clear. Latin hypercube sampling is aimed at objectives different from those we have in mind.

There has also been some work in design for numerical integration, where function evaluations can be viewed as a computationally cheap computer experiment. Much is known about design for one dimensional quadrature. In particular, Sacks and Ylvisaker (1970) constructed good designs (finite n) from asymptotically ($n \rightarrow \infty$) optimal designs. These methods, however, do not carry over to $d > 1$ dimensions [Ylvisaker (1975)]. Similarly in the numerical analysis literature, results for $d = 1$offer little guide to $d > 1$.

2.5 Basic Principles of Experimental Design

According to Prof. R. A. Fisher basic principles of experimental design are:

i) Replication ii) Randomization and iii) Local control

(i) Replication:

Replication means repetition of basic treatment under investigation. Thus replication is the repetition of same treatment on several experimental units. In fact, an experiment in which each treatment is allocated that even if same treatment is assigned to all experimental units. So it is essential to replicate the treatment to study the variation in the yields of each variety. Moreover, the average of differences in yields of two treatments is a reliable measure of difference in performance of the two treatments.

(ii) Randomization:

Professor R. A. Fisher introduced the principle of randomization in modern experimental design. Randomization is the process of distributing the treatments to the experimental units purely by chance mechanism is such a way that any experimental unit is equally likely to receive any treatment. Then randomization ensures that no treatment is unduly favored or handicapped in the experiment. The purpose of randomization is to reduce as far as possible any systematic effect of uncontrolled factors in the experiment and to give increased justification for the application of statistical theory so that bias is avoided from treatment comparisons. Separate randomization is necessary in separate replication.

(iii) Local control:

Local control is the procedure of reducing and controlling error variation by arranging the experimental units in blocks. By blocking variation among the blocks is eliminated from the experimental error. Thus error variation is reduced to a considerable extent and precision of experiment is increased. Thus local control makes the experimental design more efficient and test procedure more powerful by the reduction of error variation. Further such arrangement in blocks provides information about the difference between the blocks. It is important to note that while experimental error can only be estimated because of replication. This is sufficiently controlled by the principle of local control. In fact local control aims at eliminating all extraneous sources of variation from treatment comparisons. Thus local control leads to adequate reduction of error variation without unduly increasing the number of replication. So that real differences among the treatments can be detected significant.

2.6 Requirements of a Good Experiment

The object of a comparative experiment dealing with a set of alternating treatments is to separate the treatment effects from the uncontrolled variations. Once the treatments, the experimental units and the nature of observations have been decided upon a good experiment should satisfy the following conditions.

i) Absence of bias or free from systematic errors: It is essential to plan an experiment so that unbiased estimates of treatment differences and treatment effects can be obtained from the data of the experiment. This requires that experimental units receiving different treatments differ in no systematic way from one another. The absence of bias is achieved by the principle of randomization.

ii) Measure of experimental error: Since the treatments under comparison apparently produce differences. Hence an experiment should furnish a measure of experimental error, which is employed as a measuring stick in test of significance about treatment differences. Replication provides an estimate of experimental error and thus makes a test of significance possible.

iii) Precision: Precision refers to the closeness with which different effects are estimated. Estimates of treatment effects and of treatment differences should be precise and precision is measured by the reciprocal of variance which in turn depends on experimental error. In fact, a small value of standard error indicates increased precision of the experiment. Important methods of increasing precision are increased replication, refinement of experimental technique, blocking in various efficient designs, use of confounding in factorial experiments, use of auxiliary observations in analysis of covariance.

iv) Clearly defined objective: Every experiment should have clearly defined objective on which the design and analysis of data considerably depend.

v) Simplicity: Experimental design should be very simple and consistent with the objective on which the design and analysis of data.

vi) Scope or range of validity: The conclusions drawn from experimental data should have a wide range of validity.

2.7 Sample Design versus Experimental Design

i) Sample design deals with collection of samples from existent populations while experimental design deals with the collection of data from nonexistent population.

ii) Sample design is usually concerned with collection of samples from finite population. But design of experiments is concerned with data collection from infinite population.

iii) In sampling theory, information is obtained from sample units under the assumption that the sample can adequately represent the population. But in experimental study, data have to be manufactured by proper experimentation where the experimenter controls and modifies certain factors of interest and then he observes the effect of the modification.

iv) Sample design is concerned with absolute experiment whereas design of experiment is concerned with comparative experiments often called controlled experiments.

2.8 Experimental Design Issues

2.8.1 Classical Designs and Computer Experiments

Below some classical designs are reviewed, including factorial designs, RSM designs, and optimal designs, and their design criteria. These designs were originally developed for planning physical experiments where the experimental data are subject to experimental (replication) error. Nevertheless, these designs can be used to plan computer experiments and can sometimes have good performance [Sacks et al. (1989)].

2.8.2 Full and Fractional Designs

Fisher observed that in (full) factorial designs each experimental run contains information on many factors (input variables) and thus allows great flexibility and efficiency for studying marginal effects (main effects and interaction effects) in ANOVA decomposition. Full factorial designs include all possible combinations of factors with multiple levels. For example, two factors with an I-level and J-level full factorial include IJ combinations. As a result, full factorial designs of multiple factors and multiple levels can result in experiments with many combinations (design points), which may be too expensive in practice.

Finney (1945) developed fractional factorial designs to improve design economy by considering selected fractions (subsets) of the full factorial. These designs require far

fewer runs than the full factorial designs to study the main effects and low-order interactions of some factors by sacrificing the ability to estimate higher-order interactions. For example, a saturated seven factors 2-level fractional factorial design only requires eight combinations while a seven factors 2-level full factorial design requires 128 combinations. Plackett and Burman (1946) developed additional fractional factorial designs to achieve further savings. Some further work in factorial designs considered blocking, asymmetric factorials, incomplete replications, and different design criteria [Wang and Wu (1991)].

Fractional factorial designs are very efficient for fitting first order models or for screening out important factors, but may have complex confounding problems for fitting second or higher order polynomial models. For fitting second or third order polynomial models, experimental designs developed for response surface methodology (RSM) have been found to be quite efficient.

2.8.3 RSM Designs

Box and Wilson (1951) developed RSM to explore relationships between chemical process yield and process variables, such as ambient temperature and pressure. Box and Draper (1969) applied RSM for process improvement which is referred to as evolutionary operation (EVOP). EVOP involves two stages of empirical optimizations. In the first stage, a first-order empirical model approximation and the steepest ascent method are used to move toward the (local) optimum. In the second stage, the optimal conditions are obtained via statistical and graphical analyses of a second-order model approximation over the region which is suspected to contain the (local) optimum. Both the RSM and EVOP methods have been shown to be very effective for process characterization and optimization in industry. Excellent reviews of the methods and applications can be found in the papers of Hill and Hunter (1966) and Myers and Montgomery (1995).

Common first-order RSM designs include fractional factorial designs (discussed above) and simplex designs [Myers and Montgomery (1995)] which refer to the vertices of a simplex. Common second-order designs include the central composite design [CCD, Box and Wilson (1951)]. CCDs are basically fractional factorial designs augmented by center points and facial points so that the quadratic effects can be estimated. The Box-Behnken designs are combinations of incomplete block designs that require fewer levels than CCDs.

Both CCDs and Box-Behnken designs maintain a good balance between design efficiency (small prediction variation) and design properties (orthogonality, rotabilitiy, and robustness). Some alternative RSM designs focus on design efficiency (e.g., optimal design) by sacrificing design properties while other research focuses on design properties solely.

2.8.4 Optimal Designs

Early work in optimal designs can be found in the papers of Kiefer and Wolforitz (1960). Optimal designs are motivated by optimizing the statistical inference possibility, such as estimation or prediction capability, which are critically dependent on the underlying model. For example, consider a linear model:

$$Y_i = \beta_1 b_1(x_i) + \cdots + \beta_M b_M(x_i) + \in_i,\ i = 1, \ldots\ldots, n, \tag{2.1}$$

where Y_i is the i-th response variable, x_i is the i-th vector of input variables, $b_1(.), \ldots, b_M(.)$ are known functions, $\beta_1, \ldots, \beta_M$ are unknown coefficients, and $\in_i$ represents the experimental error and is identically and independently distributed with mean zero and variance σ^2.

In terms of estimation, it is important to minimize the variance of the estimates of the β_j's. It follows that the variance-covariance matrix of the least squares estimates of $\beta = (\beta_1, \ldots, \beta_M)^T$, $\hat{\beta}$, is

$$\sum(\hat{\beta}) = \sigma^2 (X^T X)^{-1}, \tag{2.2}$$

where $\mathbf{X^T} = (b(x_1)^T, \ldots, b(x_n)^T)$, with $b(x_i) = (b_1(x_i), \ldots, b_M(x_i))^T$.

In terms of prediction, it is important to have high accuracy in the predicted response at any given value of the input variable (x_0), $Y|x_0$. It follows that the prediction variance is

$$v(x_0) = VAR(Y|x_0) = \sigma^2\, b(x_0)^T (\mathbf{X^T X})^{-1} b(x_0). \tag{2.3}$$

As the variance-covariance is a matrix which depends on $\mathbf{X^T X}$ only and the prediction variance is a function of the value x_0 and $\mathbf{X^T X}$, the following criteria have been proposed for optimal designs:

(i) D-Optimal: minimizing determinant of $\mathbf{X}^T\mathbf{X}$,

(ii) A-Optimal: minimizing trace of $\mathbf{X}^T\mathbf{X}$,

(iii) G-Optimal: minimizing maximal value of $v(x_0)$ over all possible x_0,

(iv) Q- or IV-Optimal: minimizing the average value of $v(x_0)$ over all possible x_0.

Reviews on some optimal designs include St. John and Draper (1975) and Steinberg and Hunter (1984). Note that criteria (i) and (ii) focus on estimation error and criteria (iii) and (iv) focus on prediction error. Some of these criteria can be shown to be asymptotically equivalent [Kiefer and Wolforitz (1960)]. Obviously, many other alternatives can be considered. The primary disadvantage of optimal designs is their dependence on the underlying model, an issue which has created controversy. Nevertheless, optimal designs have been very popular in applications, including RSM optimal designs, due to their clear and simple objectives. Similar optimal criteria have been extended to model-independent designs, which are less sensitive to underlying modes.

2.8.5 Orthogonal Array Designs

The most well known orthogonal array (OA) design is the Latin square design, and both are special forms of fractional factorial designs. A complete factorial design for p variables, each at q levels, contains all q^p level combinations. A Latin square is an experimental design for three variables, each at q levels, with q^2 design points chosen so that all possible level combinations for every pair of variables appears in the design exactly once. Spatially, if the q^2 design points are projected onto any two variables, the two-dimensional complete factorial is fully represented. By overlaying two independent Latin squares in a manner that preserves orthogonality, a Graeco-Latin square for four variables is constructed. The extension to higher dimensions is called a hyper-Graeco-Latin square and is equivalent to an OA of strength two.

2.8.6 Latin Hypercube Designs

Latin hypercube designs (LHDs) are a particular class of non-collapsing designs. For LHDs on the $[0,n-1]^k$ hypercube, the values of the input parameters are chosen from the set $\{0,1,\ldots,n-1\}$ and for each input parameter each value in this set is chosen exactly once. More formally, a k-dimensional LHD of n design points can be described as a set of n points $x_i = (x_{i1}, x_{i2}, \ldots, x_{ik}) : i = 0, 1, \ldots, n-1$, is given by a n×k matrix (i.e. a matrix with n rows and k columns) X, where each column of X consists of a permutation of integers $0,1, \ldots, n-1$. Each row of X will be referred as a design point and each column of X as a factor of the design points. X can be represented as follows

$$\mathbf{X} = \begin{pmatrix} x_0 \\ \vdots \\ x_{N-1} \end{pmatrix} = \begin{pmatrix} x_{01} & \cdots & x_{0k} \\ \vdots & \cdots & \vdots \\ x_{(N-1)1} & \cdots & x_{(N-1)k} \end{pmatrix}$$

such that for each $j \in \{1,2, \ldots, k\}$ and for all $p,q \in \{0,1, \ldots, n-1\}$ with $p \neq q$, $x_{pj} \neq x_{qj}$ holds.

An important issue in generating a LH design is that the LH balance does not guarantee low correlations between input variables. In fact, perfect correlation between two input variables is possible. In practice, several LH designs are generated and input variable correlations are used to select the "best" one. Owen (1994) addressed the controlling of correlations in LH designs and the rank correlation method of Iman and Conover (1982b) is also applicable. Ye (1998) presented a method to construct orthogonal LH designs, however, only for certain restrictive samples sizes.

McKay and his coauthors (1979) introduced LH sampling, in which the continuous range of each variable is partitioned into n intervals (strata), each interval for each variable is sampled exactly once, and the univariate sample values are randomly matched across all the variables to form the n sample points.

If the n distinct levels are assumed for a LH design are the midpoints of the n intervals utilized by the LH sampling scheme, then LH sampling may be viewed as a random perturbation of the LH design points, using a uniform distribution over each interval.

A hybrid OA-LH design, for which a LH design is created from an OA structure, was introduced by Tang (1993). Starting with, say, a q-level OA of strength two and index

unity, there will be n = q2 points. The set of integer values {1, . . . , n} is partitioned into the q groups:

$$\{1, \ldots, q\}; \{q + 1, \ldots, 2q\}; \ldots; \{q(q - 1) + 1, \ldots, n\}.$$

Within the OA, the levels labeled {0, 1, . . . , q −1} correspond (in order) to the q groups. For a given dimension, each OA-level appears q times in the OA design. To create the OA-LH design, the q appearances of each OA-level are replaced by distinct integer values that are sampled without replacement from the corresponding group. Finally, the OA-LH sampling scheme utilizes random perturbations (within intervals). Both designs attain only the LH properties, except that balance is maintained in bivariate margins (with a strength two OA or higher).

2.8.7 Integrated Mean Squared Error Based Design

For a fixed number of runs, n, and for specified correlation structure R, a criterion will be necessary for choosing a design that predicts the response well at untried inputs in the experimental region *X*. Here functional of the MSE matrix or kernel is considered

$$M = \{E[Y(\omega) - \hat{y}(\omega)][Y(x) - \hat{y}(x)]\}$$

for all w and x in *X*.The diagonal elements are the MSE [$\hat{y}(x)$]. In the Bayes case when the β's in (1) are known constants, M is just the posterior covariance matrix of the process. When the β's have prior variances that tend to infinity, M is the limiting posterior covariance matrix of Y(ω). Now Integrated Mean Squared Error (IMSE) based on M is discussed. The IMSE criterion chooses the design S to minimize

$$\int_x MSE[\hat{y}(x)]\phi(x)dx$$

for **a** given weight function φ(x). The IMSE can be written as

$$\sigma^2\left\{1 - trace\left[\begin{pmatrix} 0 & F' \\ F & R \end{pmatrix}^{-1} \int \begin{pmatrix} f(x)f'(x) & f(x)r'(x) \\ r(x)f'(x) & r(x)r'(x) \end{pmatrix}\phi(x)dx\right]\right\}$$

These integrals simplify to products of one-dimensional integrals if X is rectangular and the elements of f(x) and r(x) are products of functions of a single input factor. Thus, polynomial regression models and product correlations can be numerically convenient.

The IMSE criterion is essentially the trace of M (suitably normalized). It is assumed that $\phi(x)$ is uniform, though other weights cause no real difficulty. This criterion has proved to be effective in terms of actual squared error of prediction in test examples reported by et al. (1989).

CHAPTER 3

OVERVIEW OF CORRELATION

3.1 Introduction

There are many situations in which the objective in studying the joint behavior of two set of variables is to see whether they are related, rather than to use one to predict the value of the other. Correlation is a widely used statistical technique. Correlation coefficients are the index of the measurement of the relationship among the sets of variables. There are several kinds of correlation coefficients regarding the number of variables considered. They are (i) Simple correlation (ii) Multiple correlation and (iii) Partial correlation. The correlation may be linear or nonlinear. Below they will be discussed briefly.

(i) Simple correlation: If there are only two variables, then the measure of the relation is called simple correlation. In order to compute simple correlation, two variables must be needed, with values of one variable (X) paired in some logical way with values of the second variable (Y). Such an organization of data is referred to as a bivariate (two-variable) distribution. Two variables may be positively correlated, be negatively correlated, or have no relationship to each other (zero correlation).

In the case of a positive correlation between two variables, high measurements on one variable tend to be associated with high measurements on the other and low measurements on one with low measurements on the other. With negative correlation, high scores of one variable are associated with low scores of the other. The two variables thus tend to vary together but in opposite directions. A zero correlation means that there is no relationship between the two variables. High and low scores on the two variables are not associated in any predictable manner.

A simple correlation coefficient is a measure of the relationship between two variables. It describes the tendency of two variables to vary together (co-vary); that is , it describes the tendency of high or low values of one variable to be regularly associated with either high or low values of the other variable. The linear correlation coefficient of Pearson product moment (Blanched formula) formula [Ross (2005)]:

$$r = \frac{S_{XY}}{\sqrt{(S_{XX})(S_{YY})}} \tag{3.1}$$

where X and Y are paired observations,

$$S_{XY} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) \quad ,$$

$$S_{XX} = \sum (x_i - \bar{x})^2$$

$$S_{YY} = \Sigma(y_i - \bar{y})^2$$

(ii) Multiple Correlation:

The degree of relationship existing among three or more variables is called multiple correlation. To allow for generalizations to large numbers of variables, it is convenient to adopt a notation involving subscripts.

Llet $X_1, X_2, X_3, \ldots X_k$ denote the variables (factors of the design) under consideration. Then the partial correlation between the factors i and j are given by ρ_{ij}, where each ρ_{ij} are computed as Eq. (3.1).

Then the measure the multicollinearity among the factors can be defined by the following measure of average pair-wise correlations

$$\rho^2 = \frac{\sum_{i=2}^{k}\sum_{j=1}^{i-1}\rho_{ij}^2}{k(k-1)/2} \tag{3.2}$$

Note that this definition is frequently used in literature [Fang et al. (2000b), Joseph and Hung (2008)]. Whereas the definition of maximum pair-wise correlation is given below:

$$\rho_{\max} = \max_{1\le i,j\le k} \rho_{ij} \tag{3.3}$$

(iii) Partial Correlation:

It is often important to measure the correlation between a dependent variable and one particular independent variable when all other variables involved are kept constant; that is, when the effects of all other variables are removed (often indicated by the phrase "other things being equal"). This can be obtained by defining a *coefficient of partial correlation*, except that the explained and unexplained variations must be considered that arise both with and without the particular independent variable.

If $r_{12,3}$ is denoted the coefficient of partial correlation between X_1 and X_2 , keeping X_3 constant, it is found that

$$r_{12,3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}} \tag{3.4}$$

Now some important issues regarding the analysis among the observed variables will be discussed.

3.2 Coefficient of Determination:

A statement that specifies corresponding values on the two variables is a common way to express a relationship between two variables. To predict about one variable, a value on the second variable must be needed. A more complicated method of expressing the relationship between two variables will be learnt. It is widely used because it gives, with one number, an overall index that specifies the proportion of variance the two variables have in common. The name of this index is the coefficient of determination. It is easy to calculate

At first the linear regression model has been defined by the relationship

$$Y = \alpha + \beta X + \varepsilon \tag{3.5}$$

Then the quantities $Y_i - \hat{\alpha} - \hat{\beta}X_i$, $i = 1,2,\cdots,N$ are called residuals, where $\hat{\alpha}$ and $\hat{\beta}$ are the estimators of α and β. They represent the differences between actual and predicted responses. SS_R will be considered the sum of squares of these residuals. Then SS_R can be represented as [Ross (2005)]

$$SS_R = \frac{S_{XX}S_{YY} - S_{XY}^2}{S_{XX}} \tag{3.6}$$

Then the coefficient of determination denotes by R^2 is defined by

$$R^2 = \frac{S_{YY} - SS_R}{S_{YY}} \tag{3.7}$$

It can be shown that the absolute value of the sample correlation r can be expressed as

$$|r| = \sqrt{\frac{S_{YY} - SS_R}{S_{YY}}} = \sqrt{R^2} \qquad (3.8)$$

It is worthwhile not mention here that this relation is hold for the case of simple linear correlation.

3.3 Multicollinearity :

In multiple regression problems, it is expected finding dependencies between the response variable y, and the regressor variables (factors of the designed points x_i). In the most regression problems, however, it is found that there are also dependencies among the factors of the designed points x_i. In the situations where these dependencies are strong, it is said that *multicollinearity* exists. *Multicollinearity* can have serious effects on the estimates of the regression coefficients and on the general applicability of the model.

The effects of *multicollinearity* may be easily demonstrated. The diagonal elements of the matrix $C = (X'X)^{-1}$ can be written as

$$C_{ii} = 1/(1- R^2_{ii}) \qquad i = 1, 2, \ldots, k \qquad (3.9)$$

where R^2_i is the coefficient of multiple determination resulting from regressing x_i on the other k-1 regressor variables. Clearly, the stronger the linear dependency of x_i on the remaining regressor variables, and hence the stronger the *multicollinearity,* the larger the value of R^2_i will be.

3.3.1 The effect of Multicollinearity

Multicollinearity can result in several problems. These problems are as follows:

- The partial regression coefficient due to multicollinearity may not be estimated precisely. Due to multicollinearity, the standard errors are likely to be high.
- Multicollinearity results in a change in the signs as well as in the magnitudes of the partial regression coefficients from one sample to another sample.

- Multicollinearity makes it tedious to assess the relative importance of the independent variables in explaining the variation caused by the dependent variable.

3.3.2 The Cause of Occurrence of Muliticollinearity

There are some reasons the occurrence of multicollinearity are point out below:

- Multicollinearity is caused by an inaccurate use of dummy variables.
- Multicollinearity is caused by the inclusion of a variable which is computed from other variables in the equation.
- Multicollinearity can also result from the repetition of the same kind of variable. *Practical examples of this include a Nokia N Series user and Nokia 1101 user; the height of the person in feet and the height of the person in inches, etc.* In other words, multicollinearity is caused by the inclusion of an almost identical variable twice.
- Multicollinearity generally occurs when the variables are highly and truly correlated to each other.

In the presence of high multicollinearity, the confidence intervals of the coefficients tend to become very wide and the statistics tend to be very small. It becomes difficult to reject the null hypothesis of any study when multicollinearity is present in the data under study.

- Multicollinearity is not something that can be counted.
- Multicollinearity is not discrete in nature; rather, it is continuous.
- Multicollinearity is nothing but a matter of degree.

The presence of multicollinearity can be detected in several ways.

3.3.3 Detection of Multicollinearity

(i) Multicollinearity can be detected with the help of tolerance and its reciprocal, called variance inflation factor (VIF). Some authors have suggested a formal detection-tolerance or the **variance inflation factor** (VIF) for multicollinearity:

$$\text{tolerance} = 1 - R_i^2 \ , \qquad \text{VIF} = \frac{1}{tolerance}$$

where R_i^2 is the coefficient of determination of a regression of explanator i on all the other explanators. A tolerance of less than 0.20 or 0.10 and/or a VIF of 5 or 10 and above indicates a multicollinearity problem.

(ii) Multicollinearity can also be examined with the help of a condition number. The standard measure of **ill-conditioning** in a matrix is the condition index. It will indicate that the inversion of the matrix is numerically unstable with finite-precision numbers (standard computer floats and doubles). This indicates the potential sensitivity of the computed inverse to small changes in the original matrix. The Condition Number is computed by finding the square root of (the maximum eigenvalue divided by the minimum eigenvalue). If there is no multicollinearity, then the condition number will give the value of one. If the multicollinearity increases, then the eigenvalues will be greater and smaller than one. If the eigenvalue becomes close to zero, then there is a serious multicollinearity problem. Basically, if the condition number is 15, then multicollinearity is a concern. If it is greater than 30, then multicollinearity is a very serious concern for the researcher performing the study. If the Condition Number is above 30, the regression is said to have significant multicollinearity.

(iii) Insignificant regression coefficients for the affected variables in the multiple regression, but a rejection of the joint hypothesis that those coefficients are all zero (using an F-test).

(iv) Large changes in the estimated regression coefficients when a predictor variable is added or deleted.

(v) Farrar-Glauber Test: [Farrar and Glauber (1967)] If the variables are found to be orthogonal, there is no multicollinearity; if the variables are not orthogonal, then multicollinearity is present.

(vi) As *multicollinearity* measures the linear dependency among the factors of the design points, so *multicollinearity* can be measured by the partial pair-wise correlations among the factors. There are several ways are available in the literature to measure the

pair-wise correlations i.e. *multicollinearity.* Here, the following measure of average pair-wise correlations is considered which described above

$$\rho^2 = \frac{\sum_{i=2}^{k}\sum_{j=1}^{i-1}\rho_{ij}^2}{k(k-1)/2}.$$

Or maximum pair-wise correlation:

$$\rho_{\max} = \max_{1\le i,j\le k}\rho_{ij}$$

CHAPTER 4

ITERATED LOCAL SEARCH APPROACH FOR MAXIMIN LATIN HYPERCUBE DESIGNS

4.1 Introduction

The Latin hypercube design is a popular choice of experimental design when computer simulation is used to study a physical process. These designs guarantee uniform samples for the marginal distribution of each single input. A number of methods have been proposed for extending the uniform sampling to higher dimensions. How to construct Latin hypercube designs in which all main effects are orthogonal is shown here, that ILS method can also be used to construct Latin hypercube designs with low correlation of first-order and second-order terms. ILS method generates orthogonal Latin hypercube designs that can include many more factors than those proposed by Ye (1998).

4.2 Iterated Local Search

The importance of high performance algorithms for tackling difficult optimization problems cannot be understated, and in many cases the only available methods are metaheuristics. The word metaheuristics contains all heuristics methods that show evidence of achieving good quality solutions for the problem of interest within an acceptable time. Metaheuristic techniques have become more and more competitive. When designing a metaheuristic, it is preferable that it be simple, both conceptually and in practice. Naturally, it also must be effective, and if possible, general purpose. The main advantage of this approach is the ease of implementation and the quickness.

As metaheuristics have become more and more sophisticated, this ideal case has been pushed aside in the quest for greater performance. As a consequence, problem-specific knowledge (in addition to that built into the heuristic being guided) must now be incorporated into metaheuristics in order to reach the state of the art level. Unfortunately, this makes the boundary between heuristics and metaheuristics fuzzy, and also it makes the risk of loosing both simplicity and generality.

Here a well known metaheuristics approache, namely general Iterated Local Search (ILS) has been discussed. Iterated Local Search is a metaheuristic designed to embed

another, problem specific, local search as if it were a black box. This allows Iterated Local Search to keep a more general structure than other metaheuristics currently in practice.

The essence of metaheuristic - the iterated local search - can be given in a nut-shell: one iteratively builds a sequence of solutions generated by the embedded heuristic, leading to far better solutions than if one were to use repeated random trials of that heuristic. This simple idea has a long history, and its rediscovery by many authors has lead to many different names for iterated local search like *iterated descent* [Baum (1986)], *large-step Markov chains, iterated Lin-Kernighan, chained local optimization* [Martin and Otto (1996)], or *combinations of these*. There are two main points that make an algorithm an iterated local search: (i) there must be a single chain that is being followed (this then excludes population-based algorithms); (ii) the search for better solutions occurs in a reduced space defined by the output of a black box heuristic. In practice, local search has been the most frequently used embedded heuristic, but in fact any optimizer can be used, be-it deterministic or not.

The purpose of this review is to give a detailed description of iterated local search and to show where it stands in terms of performance. So far, in spite of its conceptual simplicity, it has lead to a number of state-of-the art results without the use of too much problem-specific knowledge; perhaps this is because iterated local search is very malleable, many implementation choices being left to the developer. In what follows a formal description of ILS and comment on its main components will be given.

Procedure *Iterated Local Search*

s_0 = Generate Initial Solution

s^* = Local Search(s_0)

repeat

s' = Perturbation(s^*, history)

s'' = Local Search(s')

s^* = Acceptance Criterion (s^*; s'', history)

until termination condition met

end

ILS involves four main components:

1. creating an initial solution;
2. a black-box heuristic that acts as a local search on the set S;
3. the perturbation operator, which modifies a local solution;
4. the acceptance criterion, which determines whether or not a perturbed solution will become the starting point of the next iteration.

Local search applied to the initial solution s_0 gives the starting point s^* of the walk in the set S^*. Starting with a good s^* can be important if high-quality solutions are to be reached as fast as possible. The initial solution s_0 used in the ILS is typically found one of two ways: a random starting solution is generated or a greedy construction heuristic is applied. A "random restart" approach with independent samplings is sometimes a useful strategy (in particular when all other options fail), it breaks down as the instance size grows because in that time the tail of the distribution of costs collapses. A greedy initial solution s_0 has two main advantages over random starting solutions: (i) when combined with local search, greedy initial solutions often result in better quality solutions s^*; (ii) a local search from greedy solutions takes, on average, less improvement steps and therefore the local search requires less CPU time.

The current s^*, first a change or perturbation is applied that leads to an intermediate state s' (which belongs to S). Then Local Search is applied to s' and reaching a solution $s^{*\prime}$ in S^*. If $s^{*\prime}$ passes an acceptance test, it becomes the next element of the walk in S^*; otherwise, one returns to s^*. The resulting walk is a case of a stochastic search in S^*, but where neighborhoods are never explicitly introduced. This iterated local search procedure should lead to good biased sampling as long as the perturbations are neither too small nor too large. If they are too small, one will often fall back to s^* and few new solutions of S^* will be explored. If on the contrary the perturbations are too large, s' will be random, there will be no bias in the sampling, and a random restart type algorithm will be recovered.

In practice, much of the potential complexity of ILS is hidden in the history dependence. If there happens to be no such dependence, the walk has no memory: the

perturbation and acceptance criterion do not depend on any of the solutions visited previously during the walk, and one accepts or not $s^{*\prime}$ with a fixed rule. This leads to random walk dynamics on S^* that are "Markovian", the probability of making a particular step from s_1^* to s_2^* depending only on s_1^* and s_2^*. Most of the work using ILS has been of this type, though recent studies show unambiguously that incorporating memory enhances performance [Stutzle (1998)].

The main drawback of any local search algorithm is that, by definition, it gets trapped in local optima that might be significantly worse than the global optimum. The strategy employed by ILS to escape from local optima is represented by perturbations to the current local minimum. The perturbation scheme takes a locally optimal solution, s^*, and produces another solution from which a local search is started at the next iteration. Hopefully, the perturbation will return a solution outside the basins of attraction of previously visited local minima. That is, it will be "near" a previously unvisited local optimum. Choosing the correct perturbation scheme is of primary importance, because it has a great influence on the intensification/diversification characteristics of the overall algorithm. Generally, the local search should not be able to undo the perturbation; otherwise one will fall back into the local optimum just visited. Perturbation schemes are commonly referred to as "strong" and "weak", depending on how much they affect the solution that they change. A perturbation scheme that is too strong has too much diversity and will reduce the ILS to an iterated random restart heuristic. A perturbation scheme that is too weak has too little diversity and will result in the ILS not searching enough of the search space. The perturbation scheme should be chosen in such a way that it is as weak as possible while still maintaining the following condition: the likelihood of revisiting the perturbed solution on the next execution of Local Search should be low [Lourenco et al. (2002)]. The strength should remain as low as possible to speed up execution time. The desired perturbation scheme will return a solution near a locally optimal value. If this is the case, the local search algorithm should take less time to reach the next locally optimal value. Components from other meta-heuristics can sometimes be incorporated into the perturbation phase. Battiti and Protasi (1997) use memory structures to control the perturbation. In doing so, one can force intensification when globally good values are reached and force diversification when the search stagnates in an area of the search space. Borrowing from Simulated

Annealing [Kirkpatrick et al. (1983)], temperature controlled techniques have been used to force the perturbation to change in a deterministic manner. Basic variable neighborhood search employs a deterministic perturbation scheme. Just as perturbation can range from too much intensification (no perturbations) to too much diversification (perturb all elements of the solution), acceptance criterion choices affect the search in a similar way. The most dramatic acceptance criterion on the side of diversification is to accept all perturbed solutions. This type of practice can undermine the foundations of ILS, since it encourages a "random-walk" type search. Contrasting with this, the algorithm accepts only solutions that are improvements to the globally optimal value (a sort of greedy strategy). Many implementations of ILS employ this type of acceptance strategy. This type of criterion, especially with a weak perturbation scheme, can restrict the search from escaping the current basin of attraction. Moreover, with this type of scheme the probability of reaching the same locally optimal value increases a trait that reduces the algorithm's overall effectiveness. In this case random restart when the search stagnates is a good way to ensure some diversification and to counterbalance the (possible) negative effects of too greedy a search. Large perturbations are only useful if they can be accepted. This only occurs if the acceptance criterion is not too biased toward better solutions [Lourenco et al. (2001)]. In the paper of St¨utzle (1998) author show that acceptance criteria that accept some worse solutions outperform their best-only counterparts.

For what concerns the stopping rule, generally the algorithm executes until one of the following conditions is met:

- a predetermined number of cycles have occurred;
- the best solution has not changed for a predetermined number of cycles;
- a solution has been found that is beyond some predetermined threshold.

ILS has many of the desirable features of a metaheuristic: it is simple, easy to implement, robust, and highly effective. The essential idea of ILS lies in focusing the search not on the full space of solutions but on a smaller subspace defined by the solutions that are locally optimal for a given optimization engine. The success of ILS

lies in the biased sampling of this set of local optima. How effective this approach turns out to be depends mainly on the choice of the local search, the perturbations, and the acceptance criterion. Interestingly, even when using the most naive implementations of these parts, ILS can do much better than random restart. But with further work so that the different modules are well adapted to the problem at hand, ILS can often become a competitive or even state of the art algorithm. This dichotomy is important because the optimization of the algorithm can be done progressively, and so ILS can be kept at any desired level of simplicity. This, plus the modular nature of iterated local search, leads to short development times and gives ILS an edge over more complex metaheuristics in the world of industrial applications. As an example of this, recall that ILS essentially treats the embedded heuristic as a black box; then upgrading an ILS to take advantage of a new and better local search algorithm is nearly immediate. Because of all these features, it is believable that ILS is a promising and powerful algorithm to solve real complex problems in industry and services, in areas ranging from finance to production management and logistics. Finally, notice that although all of the present review was given in the context of tackling combinatorial optimization problems, in reality much of what is covered can be extended in a straight-forward manner to continuous optimization problems.

4.3 Maximin Latin Hypercube Designs

The s-norm distance between two points x_i and $x_j, \forall\ i, j = 1, 2, \cdots, N$ will be denoted as follows:

$$d_{ij} = \| x_i - x_j \|_s , \tag{4.1}$$

Unless otherwise mentioned, we will only consider the Euclidean distance measure ($s = 2$). In fact, the squared value of d_{ij} will be usually considered (in brief d), i.e. d^2 (saving the computation of the square root). This has a noticeable effect on the execution speed since the distances d will be evaluated many times.

4.4 Definition of LHD

A Latin Hypercube Design (LHD) is a statistical design of experiments, which was first defined by McKay and his coauthors in (1979). An LHD of k-factors (dimensions) with N design points, $x_i = (x_{i1}, x_{i2} \cdots x_{ik}) : i = 0, 1, \ldots, N-1$, is given by a $N \times k$- matrix (i.e. a matrix with N rows and k columns) X, where each column of X consists of a permutation of integers $0, 1, \cdots, N-1$ (note that each factor range is normalized to the interval $[0, N-1]$ so that for each dimension j all x_{ij} , $i = 0, 1, \cdots, N-1$ are distinct. Each row of X will be referred as a (discrete) design point and each column of X as a factor (parameter) of the design points.

X can be represented as follows

$$X = \begin{pmatrix} x_0 \\ \vdots \\ x_{N-1} \end{pmatrix} = \begin{pmatrix} x_{01} & \cdots & x_{0k} \\ \vdots & \cdots & \vdots \\ x_{(N-1)1} & \cdots & x_{(N-1)k} \end{pmatrix} \tag{4.2}$$

such that for each $j \in \{1, 2 \cdots, k\}$ and for all $p, q \in \{0, 1, \cdots, N-1\}$ with $p \neq q$; $x_{pj} \neq x_{qj}$ holds.

Given a LHD X and a distance d, let

$$D = \{d(x_i, x_j) : 1 \leq i < j \leq N\}.$$

Note that $|D| \leq \binom{n}{2}$. $D_r(X)$ is defined as the r-th minimum distance in D, and $J_r(X)$ as the number of pairs $\{x_i, x_j\}$ having $d(x_i, x_j) = D_r(X)$ in X.

The maximin LHD problem aims at finding a LHD X^* such that $D_1(X)$ is as large as possible. However, a search which only takes into account the D_1 values is certainly not efficient. Indeed, the landscape defined by the D_1 values is "too flat". For this reason the search should be driven by other optimality criteria, which take into account also other values besides D_1.

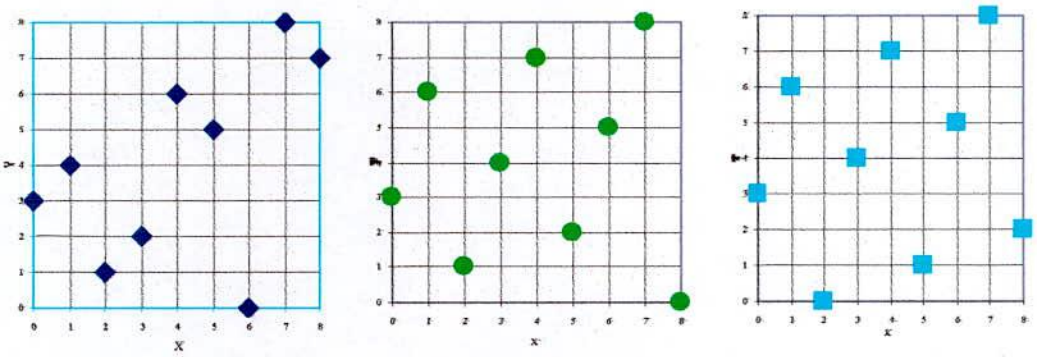

Fig: (a) $D_1(X_r)=2$, $J_1(X_r)=4$ Fig: (b) $D_1(X_{sm})=8$, $J_1(X_{sm})=4$ Fig: (c) $D_1(X_M)=8$, $J_1(X_M)=2$

Fig. 4.1: Some LHDs and their corresponding (D_1, J_1) values.

4.5 Optimality Criteria

A criterion is necessary in order to drive the search through LHDs. Some of the criteria employed in the literature are given below.

Opt(D_1, J_1) Optimality Criterion : Under this criterion a LHD Y can be considered better than another one X if a lexicographic ordering holds:

$$\begin{aligned} &D_1(Y) > D_1(X) \quad \text{or} \\ &D_1(Y) = D_1(X) \quad \text{and} \quad J_1(Y) < J_1(X). \end{aligned} \tag{4.3}$$

This optimality criterion is illustrated as follows. In Figure 4.1(a) X_r is a randomly generated LHD with $(N, k,) = (9,2)$ where $D_1(X_r) = 2$ and $J(X_r) = 4$; Figure 4.1 (b) presents an improved configuration X_{sm} where $D_1(X_{sm}) = 8$ with $J(X_{sm}) = 4$. A third LHD X_M is given in Figure 4.1 (c) where $D_1(X_M) = 8$ and $J_1(X_M) = 2$; by the Opt(D_1, J_1) criterion this is the best configuration among the three.

By generalizing this approach, the problem like a multi-objective problem can be considered with priorities: maximize the objective with highest priority D_1; within the set of optimal solutions with respect to D_1, minimize the objective with second highest priority J_1. Note that Johnson and his coauthors (1990) first proposed this optimality criterion.

Opt(φ) Optimality Criterion : As previously remarked, if there exist different LHDs with equal D_1 and J_1 values, i.e. in case there exist at least two LHDs X, Y such that $D_1(X) = D_1(Y) = D_1$ and $J_1(X) = J_1(Y) = J_1$, further the objective D_2 could be considered and maximized $D_2(X)$, the second smallest distance in X, and, if equality still holds, minimize $J_2(X)$, the number of occurrence of $D_2(X)$, and so on. Then an optimal design X sequentially maximizes D_{is} and minimizes J_{is} in the following order: $D_1, J_1; D_2, J_2, \cdots, D_m, J_m$. Morris and Mitchell (1995) have used all the above measures to define a family of scalar-valued functions (to be minimized), which can be used to rank competing designs in such a way that a maximin design receives the highest ranking. This family of functions, indexed by p, is given by

$$\phi_p(x) = \sum_{r=1}^{m}\left[\frac{J_r(X)}{(D_r(X))^p}\right]^{1/p} \tag{4.4}$$

where p is a positive integer parameter. Under this criterion, LHD Y is better than X if

$$\phi_P(Y) < \phi_R(X).$$

Note that for large enough p, each term in the sum in (4.4) dominates all subsequent terms. Through p the impact of the different D_r distances can be controled: as p increases, the impact of distance D_1 becomes more and more relevant. In the form (4.4), the evaluation of ϕ_p would be computationally costly. However, it has a computationally cheaper form Jin et al. (2005). Indeed, (4.4) can be simplified as

$$\phi_P(X) = \left[\sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{1}{d_{ij}^p}\right]^{\frac{1}{p}} \tag{4.5}$$

which can be computed without the need of detecting and ordering all the D_i values.

An apparent drawback of the Opt(ϕ) criterion, if there is an interest in maximin values (maximum D_1 value), is that LHDs with smaller (better) ϕ_p can have a worse(smaller) D_1, i.e. there are X and Y such that $\phi_P(X) < \phi_P(Y)$ and $D_1(X) < D_1(Y)$. This phenomenon has been frequently observed in our computational experiments. Nevertheless, a profitable choice is to work in order to minimize the ϕ_p function but, at the same time, keep track of the best (D_1, J_1) values observed during such

minimization. This way the search in the solution space is guided by a kind of heuristic function. Such mixed approach might appear strange but, as it will be demonstrated experimentally, it can be extremely effective.

While the two criteria above are strictly related to maximin values and they will be widely employed in the definition of approaches for detecting maximin solutions, for the sake of completeness, it is also mentioned that other optimality criteria, not necessarily related with maximin values, are available in the literature. A couple of them is presented below.

Correlation Optimality Criterion: Iman and Conover (1982a), Owen (1994), and Tang (1998) proposed to choose designs by minimizing correlations among factors within the class of LHDs. Owen (1994) used the following performance measure for evaluating the goodness of the LHD with respect to pairwise correlations. It is defined as follows

$$\rho^2 = \frac{\sum_{i=1}^{k}\sum_{j=i+1}^{k}\rho_{ij}^2}{k(k-1)/2} \tag{4.6}$$

For calculating the correlation between each single pair of factors (say column q and column r), here Pearson's formula is used:

$$\rho_{qr} = \frac{N\sum x_{iq}x_{ir} - \sum x_{iq}\sum x_{ir}}{\sqrt{N\sum x_{iq}^2 - (\sum x_{iq})^2}\sqrt{N\sum x_{ir}^2 - (\sum x_{ir})^2}} \tag{4.7}$$

where the sums are over $i = 1, \cdots, N$. Here Y is better than X if $\rho^2(Y) < \rho^2(X)$.

4.6 ILS heuristic for maximin LHD

A general scheme for ILS-based algorithms has been discussed in Section 4.1. Now the ILS based procedure will be presented here for maximin Latin hypercube design. It is mentioned earlier that, the main components of ILS heuristic approaches are Initialization (I_S), LocalSearch (L_M), Perturbation Move (P_M), and the Stopping Rule (S_R). Now the pseudo-code of the proposed ILS heuristic for maximin is given bellow

LHD problems is the following:

Step 1. **Initialization :** $X = I_S(\{0, 1, \ldots, N-1\})$

Step 2. **Local Search :** $X^* = L_M(X)$

while S_R not satisfied do

Step 3. **Perturbation Move :** $X' = P_M(X)$

Step 4. **Local Search :** $X^* = L_M(X')$

Step 5. **Improvement test :** if X^* is better than X,

set $X = X^*$

end while

Return X

In order to fully specify the algorithm the components are detailed below.

4.6.1 Initialization (I_S)

The initialization (I_S) procedure embedded in the algorithm is extremely simple: the first initial solution is randomly generated. In particular, the first initial solution generation is built as follows. For each component $h \in \{1, \ldots, k\}$ a random permutation $v_0, \ldots, v_{N-1}$ of the integers $0, 1, \ldots, N-1$ is generated and set

$$x_{rh} = v_r \quad \text{for all } r \in \{0, \ldots, N-1\}.$$

Although more aggressive procedures could be designed, we chose random generation because it is fast and unbiased.

4.6.2 Local Search Procedure (L_S)

In order to define a local search procedure (L_S), it is necessary to define a concept of neighborhood of a solution. Given a LHD $X = (x_1, \ldots, x_N)$, its neighborhood is made of all other LHDs obtained by applying local moves to X. Before introducing some local moves, first the notion of critical point is introduced.

Critical point: It is said that x_i is a critical point for X, if

$$\min_{j \neq i} d(x_i, x_j) = D_1(X),$$

i.e., the minimum distance from $\mathbf{x}_i$ to all other points is also the minimum one among all the distances in X. $I(X)$ $\{1, \ldots, N\}$ the set of indices of the critical points in X is denoted.

Local moves (L_M): A local move is an operator that applies some form of slight perturbation to a solution X, in order to obtain a different solution. Different local moves define different neighborhoods for local search. In the literature two different local moves are available: Rowwise-Pairwise (RP) exchange [Park (1994)] and Columnwise-Pairwise (CP) exchange [Morris and Mitchell (1995)]. In Park's algorithm [Park (1994)] some active pairs (pairs of critical points, in the terminology) are selected. Then, for each chosen pair of two active rows, say i_1 and i_2, the RP exchange algorithm considered all the possible exchanges of corresponding elements as follows:

$$x_{i1,p} \leftrightarrow x_{i2,q} \quad \forall \quad p, q = 1, 2, \ldots, k : p \neq q,$$

and found the best exchange among them. The CP algorithm proposed by Morris and Mithchell (1995) exchanges two randomly selected elements within a randomly chosen column. But Li and Wu (1997) defined the CP algorithm in a bit different way: they randomly chose a column and replaced it by its random permutations if a better LHD is obtained.

It is observed that the effect of CP based local search and RP based local search is not significance [Jamali (2009)]. So, here, RP based local move is considered as defined in Jamali (2009) which is a bit different than that of Park (1994). For optimal criteria, Opt(ϕ) optimal criteria is considered.

The definition of Rowwise-Pairwise Critical Local Moves (it is called LM_{RpD1}) as follows. The algorithm sequentially chooses two points (rows) such that at least one of them is a critical point, then exchanges two corresponding elements (factors) of the selected pair. If $i \in I(X)$, $r, j \in \{1, \ldots, N\}$, $h, \ell \in \{1, \ldots, k\}$, swapping the ℓ-th component gives the neighbor Y defined by

$$y_{rh} = \begin{cases} x_{rh} & \text{if } r \neq i \text{ or } h \neq \ell \\ x_{ih} & \text{if } r = j \text{ and } h = \ell \\ x_{jh} & \text{if } r = i \text{ and } h = \ell \end{cases} \tag{4.8}$$

It is remarked that, if Opt(D_1, J_1) be the optimality criterion, it perfectly makes sense to avoid considering pairs x_i and x_j such that $I(X) \cap \{x_i, x_j\} = \emptyset$ since any swap involving two non-critical points cannot improve the D_1 value of the current LHD.

When Opt(ϕ) is adopted as optimality criterion, any exchange can, in general, lead to an improved value of ϕ. The RP local move for Opt(ϕ) optimality criterion is denoted by $LM_{RP}\phi$ and is also defined as Eq. (4.8), the only difference being that the requirement is dropped that at least one point must be critical.

Now the RP based local moves are illustrated by considering a randomly generated initial design A : $(N,k) = (7,2)$ (see Figure 4.2(a)). Then a neighborhood solution of A, by considering points (0,2), (4,4) (here both are critical points), is LHD B, obtained after swapping the second coordinate of the points (0, 2) and (4,4) (Figure 4.2 (b)).

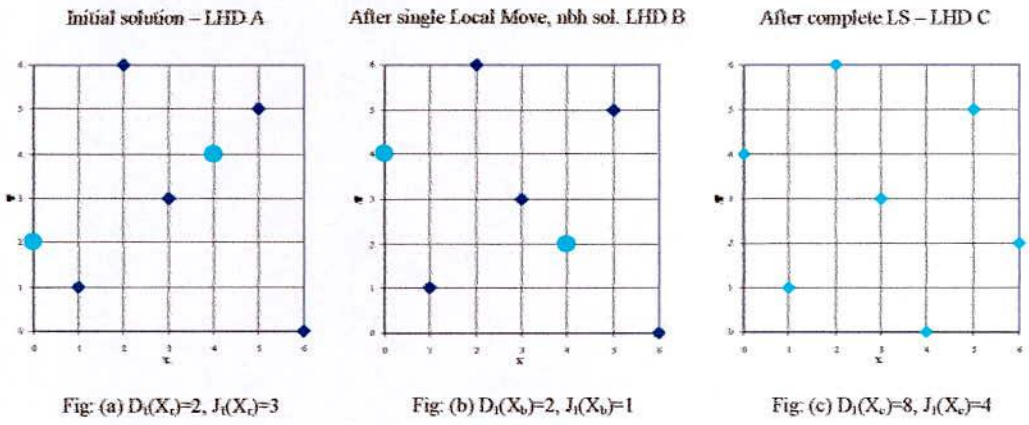

Fig. 4.2: Illustration of Neighborhood solutions for $LM_{RP_{D1}}$ based local search (LS) procedure

Also note that LHD B is an improving neighbor of LHD A, since $(D_1, J_1)(B) = (2,1)$ whereas $(D_1, J_1)(A) = (2,3)$. Finally Figure 4.2 (c) shows the maximin LHD produced by the Local search procedure.

Acceptance Rule: Between the two type of local moves [Jamali (2009)], Best Improve (BI) acceptance rule is considered as there are no significant difference regarding output [Jamali (2009)]. For the BI acceptance rule, the whole neighborhood of the

current solution is searched for the best improving neighbor. The reader was warned again that the meaning of "Y is better than X" can be defined accordingly with the Opt(D_1, J_1) or Opt(ϕ) optimality criterion. So for the Opt(D_1, J_1) optimality criterion: "Y is better than X" if

$$D_1(Y) > D_1(X) \text{ or } (D_1(X) = D_1(Y) \text{ and } J_1(X) > J_1(Y)).$$

On the other hand for Opt(ϕ) optimality criterion : "Y is better than X" if

$$\phi_p(Y) < \phi_p(X),$$

where ϕ_p is defined by (5).

4.6.3 Perturbation Move (P_M)

Perturbation is the key operator in ILS, allowing the algorithm to explore the search space by jumping from one local optimum to another. Basically, a perturbation is similar to a local move, but it must be somehow less local, or, more precisely, it is a move within a neighborhood larger than the one employed in the local search. Actually the perturbation operator produces the initial solutions for all the local searches after the first one. Between the two types of perturbation operators, say, (i) Cyclic Order Exchange (COE) and (ii) Pairwise Crossover (PC) proposed in Jamali (2009), COE is considered.

1. Cyclic Order Exchange (COE): First perturbation move procedure is Cyclic Order Exchange (COE). The operator COE produce a cyclic order exchange upon a randomly selected single component (column) of a randomly selected portion of the design points (rows). Among the three variant of COE perturbation move techniques: Single Cyclic Order Exchange (SCOE) perturbation operation, Multiple Components Cyclic Order Exchange (MCCOE), and Multiple Single Cyclic Order Exchange (MSCOE) [Jamali (2009)], only SCOE technique is considered.

Single Cyclic Order Exchange (SCOE): Two different rows (points) are randomly chosen, say x_i and x_j , such that $i < j$ and $j - i \geq 2$. Then, a column (component), say ℓ is randomly choosen. Finally, the value of component ℓ is swapped in cyclic order from point x_i to point x_j. The pseudo-code structure for SCOE is the following.

Step 1: randomly select two different points x_i and x_j
such that $i < j$ and $j - i \geq 2$

Step 2: Randomly choose a component ℓ

Step 3a: set temporarily $x'_{j\ell} = x_{j\ell}$
for $t = j, j-1, \ldots, i+1$ **do**

Step 3b: Replace the component $x(t)\ell$ by $x(t-1)\ell$
end for

Step 3c: and replace $x_{i\ell}$ by $x'_{j\ell}$

Note that $j - i \geq 2$ is required otherwise the perturbation would be a special case of the local move employed in the local search procedure. The SCOE perturbation is illustrated by an example. Let the current LHD X^* with $N = 6$ and $k = 8$ is

$$X^* = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 0 & 5 & 3 & 1 & 5 & 2 & 4 & 4 \\ 1 & 0 & 4 & 2 & 4 & 3 & 3 & 5 \\ 2 & 1 & 5 & 3 & 3 & 4 & 2 & 0 \\ 3 & 2 & 0 & 4 & 2 & 5 & 1 & 1 \\ 4 & 3 & 1 & 5 & 1 & 0 & 0 & 2 \\ 5 & 4 & 2 & 0 & 0 & 1 & 5 & 3 \end{pmatrix} \tag{4.9}$$

Now two rows (points), say $\boldsymbol{x_2}$ and $\boldsymbol{x_5}$ and the column (component), say $\ell = 4$ are randomly chosen. Then, after the SCOE perturbation the following LHD X' (bold faces denote the values modified with respect to X^*) is obtained,

$$X' = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 0 & 5 & 3 & 1 & 5 & 2 & 4 & 4 \\ 1 & 0 & 4 & \mathbf{5} & 4 & 3 & 3 & 5 \\ 2 & 1 & 5 & \mathbf{2} & 3 & 4 & 2 & 0 \\ 3 & 2 & 0 & 3 & 2 & 5 & 1 & 1 \\ 4 & 3 & 1 & 4 & 1 & 0 & 0 & 2 \\ 5 & 4 & 2 & 0 & 0 & 1 & 5 & 3 \end{pmatrix} \tag{4.10}$$

Note that SCOE only slightly modifies the current LHD X^* but this exactly follows the spirit of ILS, where the perturbation should keep unchanged large portions of the current solution and should not completely disrupt it structure.

4.6.4 Stopping Rule (S_R)

A very simple stopping Rule (S_R) is used here. An integer parameter called MaxNonImp (MNI) is introduced and the algorithm will stop if the currently best local optimizer X^* cannot be improved for MaxNonImp consecutive perturbations.

CHAPTER 5

COMPUTATIONAL EXPERIMENTS AND DISCUSSION REGARDING EUCLIDEAN DISTANCE

5.1 Introduction

In the study two kinds of distance measure have been considered, one is Euclidean distance measure (L^2) another one is Rectangular distance measure (L^1). If $x = (x_1, x_2, x_3, \ldots\ldots, x_n)$ and $y = (y_1, y_2, y_3, \ldots\ldots y_n)$ be two points, then the distance of the two points in Rectangular distance measure is defined by

$$d(x,y:R) = \sum_{i=1}^{n} |x_i - y_i| \tag{5.1}$$

Whereas the distance in Euclidian distance measure is gives as follows

$$d(x,y:E) = \left| \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \right| \tag{5.2}$$

For simplicity square of the distance is considered i.e.

$$d^2(x,y:E) = \sum_{i=1}^{n} (x_i - y_i)^2 \tag{5.3}$$

In what follow, in this chapter, only Euclidian distance measure is considered but square. D_1 is also denoted as a minimum inter-site square distance of the design points in Euclidian distance measure.

5.2 Comparison of ILS with the existing literature

At first several experiments have been performed to find out the maximin LHDs. For the experiments following parameter setting is considered:

Table 5.1: Parameter setting for the experiments of ILS approach

Experimental design	LHD
Method	ILS
Optimal Criteria	Opt(ϕ)
Local Move	RP
Acceptance Rule	BI
Perturbation Technique	SCOE
Stopping Rule	MaxNonImp parameter value
MaxNonImp setting	100
Parameter , p	20

Table 5.2: The setting of number of runs for the ILS approach

k	N	R
3-10	2-25	500
3-10	26-50	100
3,4,5	51-100	50
6-10	51-100	10

Now the results are compared with available ones in the web site **www.spacefillingdesigns.nl/** and the literatures. It is worthwhile to mention here that this website is frequently updated. The best results also uploaded here. The comparisons of the experimental results are reputed in the table 5.3 and table 5.4.

Several computational experiments are performed to test the ILS heuristic in Euclidian measure (L^2) and compared it with available ones in the literature. The summary of the experimental results are given in **Table 5.3** and **Table 5.4**. Note that D_1 values are given in square of the actual values in Euclidian distance measure in order to avoid the fractional terms. In the table the results for ***PD*** (Periodic Design) and ***SA*** (Simulated Annealing) algorithms are taken from the paper of Husslage et al. (2006). The results of ***SA_M*** (Simulated Annealing proposed by Morris and Mitchell) are taken from Morris and Mitchell (1995).

Table 5.3: Comparison among several approaches of finding maximin LHDs for N=2 to 100

	Number of best solutions (maximin LHD)						Identical solutions	Worse Solution
k	***PD***	***SA***	***SA_M***	***Web***	***MS***	***ILS***	***ILS***	***ILS***
3	**61**	0	0	*65*	0	**14**	20	65
4	**02**	0	0	*47*	0	**34**	18	47
5	00	0	0	*11*	0	**78**	10	11
6	00	0	0	00	0	**90**	09	00
7	00	0	0	00	0	**92**	07	00
8		0	0	00	0	**93**	06	00
9		0	0	00	0	**93**	06	00
10		0	0	00	0	**92**	07	00

Notice that in the paper of Morris and Mitchell (1995), few points were considered with Euclidian distance measure, whereas results for the column ***Web*** are taken from the "Archived results for 3-10 dimensional L^2 -maximin Latin hypercube designs – 14/3/2006" of the web www.spacefillingdesigns.nl. Note that this web portal is updated later by Grosso et al. (2009) for maximin LHDs, where updated values are obtained by the proposed ILS approach. The results of ***MS*** (MultiStart) are taken form Grosso et al. (2008). The results in column **ILS** are obtained by the proposed Iterated Local search approach with Opt (ϕ) criterion. Note that in the column "Identical solution" means the solution of ILS is identical of the Web value.

It is observed in the **Table 5.3** that the approach is able to obtain 14 improved maximin LHDs for k=3; 34 improved solutions for k = 4; 78 improved solutions for k = 5; 90 improved solutions for k = 6; 92 improved solutions for k = 7; 93 improved solutions for k = 8; 93 improved solutions for k = 9; 92 improved solutions for k = 10. It is observed that the performance of the ILS approach outperform the SA, SA_M and MS approaches for all dimension considered here. It is also observed (the detail results are not reputed here) that ILS approach is also better compare to PD approach accept k = 3 and 4 when N is large. It is remarked that the increasing of k or/and N, the performance of the proposed ILS approaches is rather better. Note that, for k = 8, 9, 10, there are no value available in literature for the PD approach and very few values are available in SA_M [Morris and Mitchell (1995), Grosso et al. (2008)].

Table 5.4: Comparison of computational cost

Total Elapsed Time (hrs)			
k	PD	SA	ILS
3	145	500	164
4	61	181	507
5	267	152	767
6	108	520	1235
7	232	246	698
8	--	460	846
9	--	470	1087
10	--	470	1166

The computational cost of the approaches is reported in the table 5.4.
It is still needed to comment about the computation times. As already remarked there is no information regarding times to obtain the Web's results. It is however quite clear that ILS is more computationally demanding with respect to PD and SA. Such higher costs are clearly rewarded in terms of quality of the results but the quality of the results might be wondered if the time restrictions are imposed on ILS. According to some further experiments that were performed, it would be realized that, especially at large k values, equivalent or better results with respect to the PD and SA ones, could quickly be reached by ILS. Therefore, it seems that at large k values even few and short runs of ILS are able to deliver results better than those reached by PD and SA.

Now several experiments are performed for analysis the multicollinearity among the design points obtained by the ILS approach. The correlation will be investigated among the design points obtained by ILS approach. For the analysis, both the pair-wise correlation and maximum correlation are considered. The experimental results are compared with literature namely SA approach regarding D_1 values. It is worthwhile to mention here that both the approaches (ILS as well as SA) optimized D_1 values rather than ρ (average correlation) value. Here it is considered that the number of factors k = 3, … 10 and for each k factor the number of points $N = 5i$; i = 1, 2,…, 20.

Table 5.5: Multicollinearity analysis of the design obtained by ILS approach and comparison of maximin LHD for $k = 3, 4, 5$

N	$k=3$			$k=4$			$k=5$		
	ρ	D_1		ρ	D_1		ρ	D_1	
	ILS	ILS	SA	ILS	ILS	SA	ILS	ILS	SA
5	***0.200***	11(6)	11	**0.261**	15(4)	15	**0.255**	24(5)	24
10	**0.127**	27(3)	27	**0.163**	50(12)	50	0.006	82(20)	82
15	0.043	48(8)	48	0.046	**89(2)**	83	0.073	**131(3)**	124
20	0.077	**66(12)**	62	0.023	**137(8)**	123	0.036	**210(3)**	184
25	**0.280**	**91(48)**	81	0.06	**181(1)**	162	0.080	**286(1)**	255
30	0.073	**109(1)**	102	0.051	**234(1)**	209	0.020	**403(2)**	335
35	0.040	**129(1)**	122	0.035	**289(2)**	255	0.022	**482(1)**	418
40	0.025	**161(3)**	146	0.029	**345(1)**	301	0.019	**590(1)**	505
45	**0.189**	**186(47)**	166	0.039	**412(1)**	362	0.024	**706(1)**	615
50	**0.141**	**213(33)**	185	0.031	**480(1)**	414	0.03	**834(1)**	699
55	0.068	**243(35)**	214	0.041	**550(1)**	477	0.026	**966(1)**	805
60	**0.108**	**273(41)**	237	0.033	**624(1)**	530	0.019	**1101(1)**	928
65	0.085	**314(43)**	260	0.021	**705(1)**	582	0.023	**1239(2)**	1035
70	0.006	**321(1)**	285	0.032	**779(2)**	658	0.013	**1439(1)**	1135
75	0.018	**353(1)**	310	0.039	**867(3)**	714	0.015	**1571(1)**	1282
80	**0.142**	**403(52)**	344	0.026	**949(5)**	786	0.015	**1702(1)**	1430
85	0.037	**426(29)**	369	0.035	**1043(6)**	877	0.017	**1871(1)**	1566
90	0.028	**481(180)**	384	0.022	**1134(7)**	940	0.016	**2032(1)**	1696
95	0.011	**482(4)**	413	0.030	**1223(1)**	1010	0.031	**2206(2)**	1846
100	0.054	**554(4)**	451	0.016	**1313(1)**	1074	0.019	**2401(3)**	1975

It is observed in the table 5.5 that for $k = 3, 4, 5$, except few designs (in the table see the ρ values in bold face), the average pair-wise correlations are less than 0.1. Similarly it is also noticed in the table 5.6 and 5.7 that for $k = 6,7, 8, 9, 10$, except few designs (the ρ values are shown in bold face in these tables) the average pair-wise correlations are less than 0.1. Moreover it is observed in the tables that accept the designs $(k, N) = \{(3,5); (3,25); (4,5); (5,5); (6,5); (7,5); (8,5); (9,5); (10,5),\}$ the average correlation are less than 0.2. It is remark that accept the design (3,5), the average correlations are greater than 0.2 when the number of design point $N = 5$. Perhaps the design points $N = 5$ has inherent muliticollinearity property. From the above experiments it may be concluded that the average pair-wise correlation among the factor of design points are tolerable and the average pair-wise correlations among the factors of each design are decreasing (though not strictly monotonic) as increasing the number of design points N

and/or factors k. On the other hand it is also observed that regarding D_1 values (maximin LHD value) the ILS approach able to obtain, accept N = 3,4, better solutions for all dimension as well as design points. It is also observed that the ILS approach performed better regarding D_1 values. For $N = 3, 4$, the ILS approach able to identical value compare to SA approach.

Table 5.6: Multicollinearity analysis of the design obtained by ILS approach and comparison of maximin LHD for $k = 6, 7, 8$

N	$k=6$			$k=7$			$k=8$		
	ρ	D_1		ρ	D_1		ρ	D_1	
	ILS	ILS	SA	ILS	ILS	SA	ILS	ILS	SA
5	**0.342**	27(3)	27	**0.357**	32(2)	32	**0.395**	40(10)	40
10	**0.110**	**93(3)**	91	**0.146**	**112(4)**	110	**0.127**	**133(4)**	130
15	**0.102**	**175(2)**	167	0.054	**224(1)**	211	0.075	**280(1)**	257
20	0.042	**285(1)**	260	**0.113**	**327(3)**	332	0.061	**434(1)**	403
25	0.037	**408(1)**	368	0.023	**531(2)**	467	0.051	**637(1)**	583
30	0.044	**545(1)**	473	0.045	**725(1)**	620	0.014	**897(1)**	787
35	0.059	**697(3)**	601	0.016	**936(1)**	811	0.048	**1151(2)**	1002
40	0.029	**886(1)**	739	0.035	**1162(1)**	970	0.021	**1459(1)**	1224
45	0.030	**1065(1)**	891	0.032	**1408(1)**	1192	0.026	**1755(1)**	1480
50	0.024	**1218(1)**	1042	0.033	**1707(1)**	1397	0.026	**2089(1)**	1772
55	0.018	**1432(1)**	1198	0.012	**2043(1)**	1639	0.036	**2462(1)**	2084
60	0.019	**1647(1)**	1381	0.018	**2284(2)**	1899	0.032	**2888(1)**	2393
65	0.020	**1884(1)**	1565	0.025	**2579(1)**	2132	0.025	**3321(1)**	2723
70	0.018	**2116(1)**	1759	0.032	**2895(1)**	2417	0.013	**3779(1)**	3130
75	0.018	**2365(1)**	1963	0.027	**3230(1)**	2703	0.013	**4228(1)**	3513
80	0.015	**2597(1)**	2152	0.029	**3641(1)**	2979	0.018	**4695(1)**	3877
85	0.017	**2877(1)**	2399	0.016	**4023(1)**	3299	0.016	**5110(1)**	4324
90	0.013	**3134(1)**	2633	0.021	**4400(2)**	3661	0.016	**5608(1)**	4699
95	0.015	**3455(1)**	2817	0.017	**4844(1)**	3940	0.022	**6148(1)**	5154
100	0.015	**3710(1)**	3117	0.017	**5206(1)**	4335	0.013	**6692(1)**	5597

Table 5.7: Multicollinearity analysis of the design obtained by ILS approach and comparison of maximin LHD for $k = 9, 10$

N	$k = 9$			$k = 10$		
	ρ	D_1		ρ	D_1	
	ILS	ILS	SA		ILS	SA
5	**0.404**	43(2)	43	**0.411**	50(10)	50
10	0.043	**157(4)**	154	**0.121**	**174(5)**	172
15	0.098	**318(1)**	296	**0.115**	**358(1)**	337
20	0.062	**517(1)**	472	0.016	**645(6)**	542
25	0.066	**752(1)**	688	0.040	**875(2)**	792
30	0.035	**1041(1)**	925	0.0410	**1210(1)**	1086
35	0.012	**1398(1)**	1190	0.036	**1595(1)**	1398
40	0.043	**2102(1)**	1489	0.008	**2102(1)**	1742
45	0.028	**2126(1)**	1820	0.028	**2466(1)**	2130
50	0.015	**2569(1)**	2179	0.028	**2991(1)**	2556
55	0.016	**2996(1)**	2570	0.018	**3530(1)**	3054
60	0.021	**3446(1)**	2939	0.017	**4109(1)**	3500
65	0.019	**3991(1)**	3357	0.018	**4695(1)**	4034
70	0.019	**4516(1)**	3841	0.016	**5366(1)**	4539
75	0.027	**5141(1)**	4298	0.018	**6015(1)**	5171
80	0.025	**5792(1)**	4807	0.018	**6733(1)**	5773
85	0.013	**6479(1)**	5340	0.026	**7508(1)**	6397
90	0.008	**7152(1)**	5832	0.023	**8325(1)**	7040
95	0.010	**7765(1)**	6396	0.020	**9252(1)**	7741
100	0.011	**8520(1)**	6983	0.014	**10233(1)**	8450

Now experiments are performed for analysis the maximum pair-wise correlation, ρ_{max}, as well as average pair-wise correlation, ρ, of the maximin LHD designs. For this experiments the design points $N = 3, \ldots, 100$, and the dimensions $k = 4, 6, 8, 10$ are considered. All the other settings remain unchanged. The experimental results are reputed in the Figure 5.1 and 5.2. It is observed that, accept for low value of N the average pair-wise correlations are less than 0.2 among the factors of each experimental design. Figure 5.2 shows that the maximum pair-wise correlations decreasing but not strictly monotonic as increasing the number of design points.

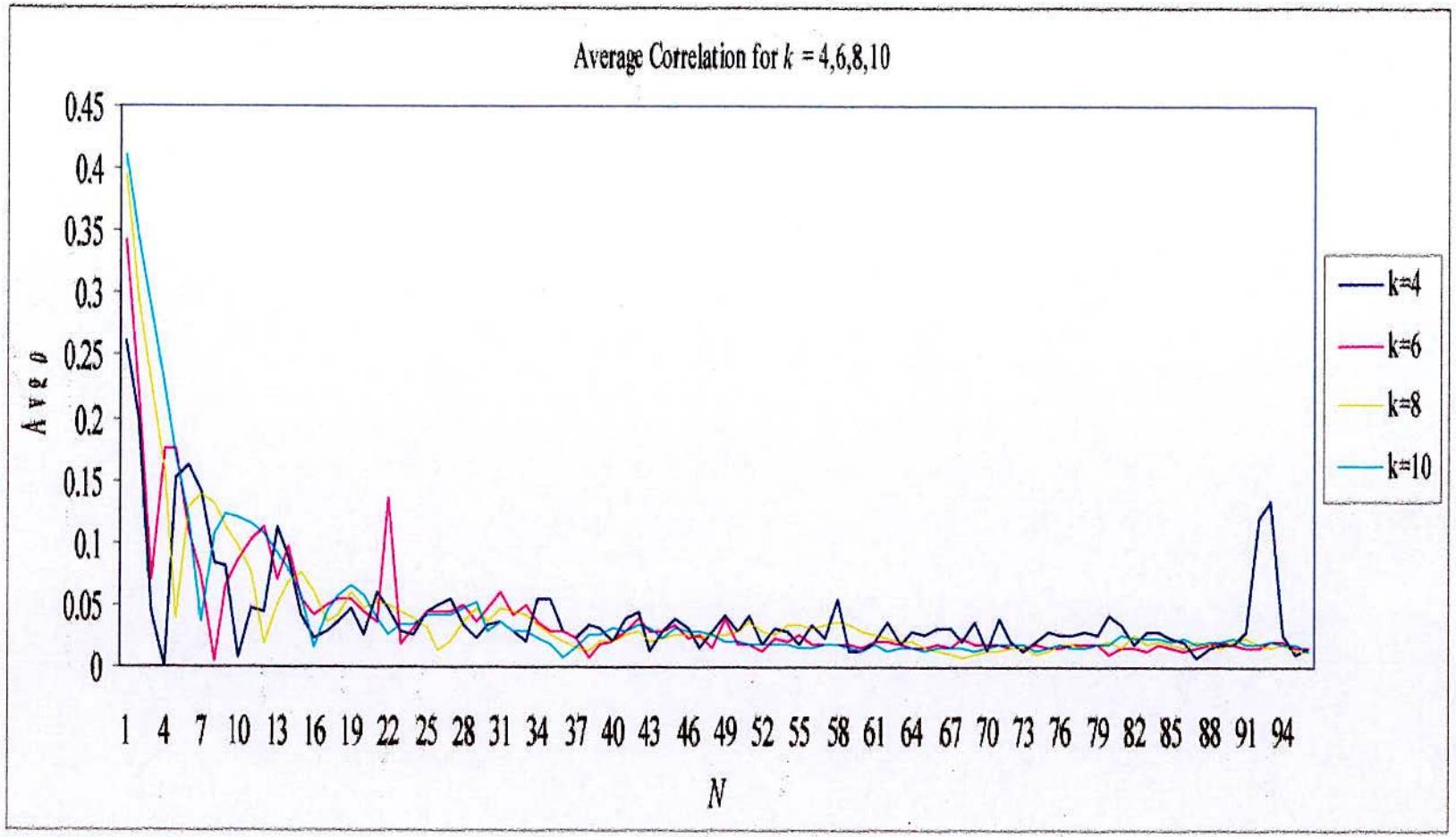

Fig. 5.1: Trends of the average pair-wise correlations ρ for points $N = 3, 4, \ldots, 100$

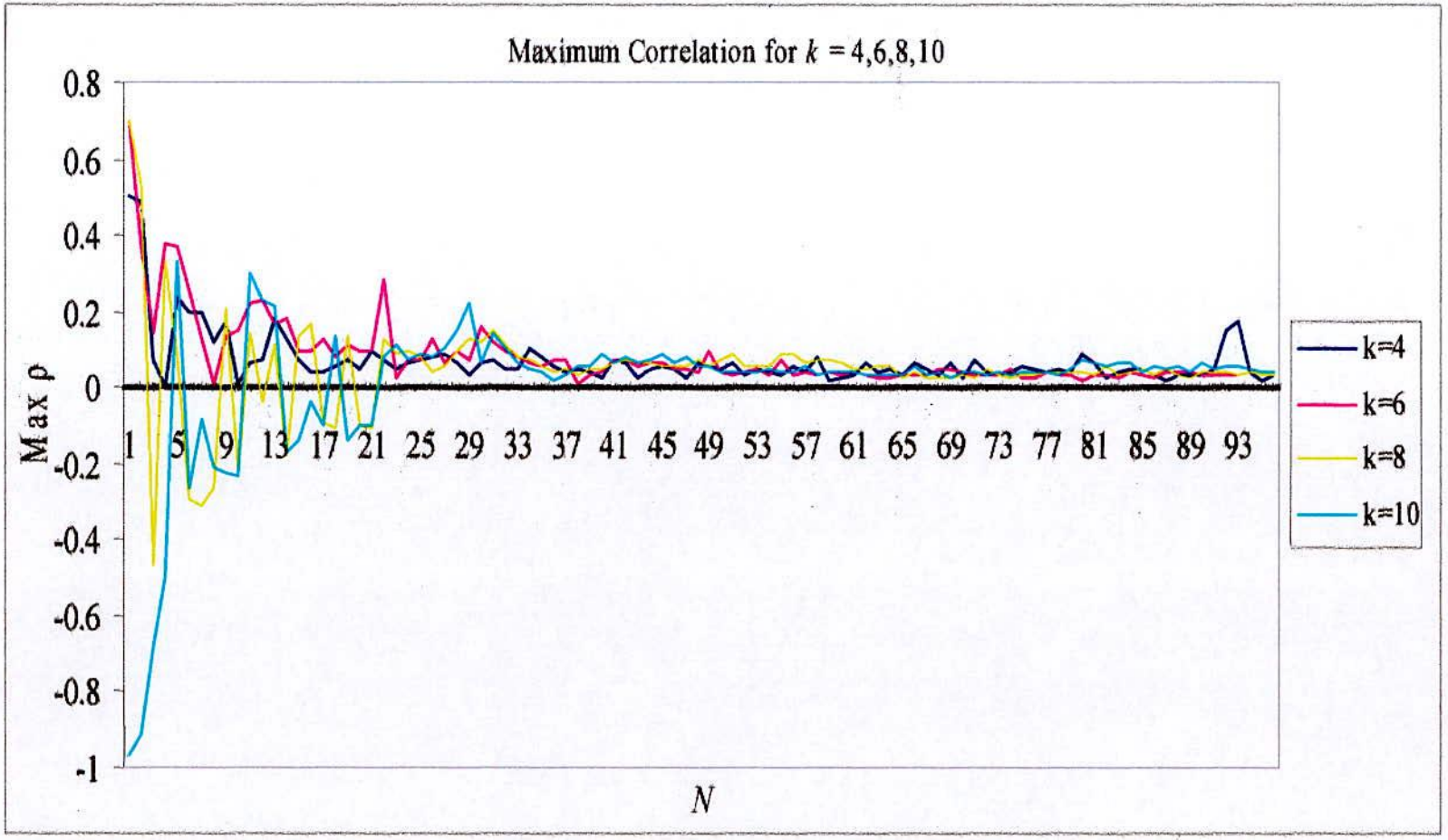

Fig. 5.2: Trends of the maximum pair-wise correlations ρ_{max} for points N = 3, 4, … , 100

From these above primary experiments it can be concluded that the proposed ILS approach able to find out the optimal (maximin) LHDs which are state-of-the-arts regarding *space-filling* as well as *non-collapsing* properties and have tolerable *Multicollinearity* property. Particularly it can be remarked that when numbers of design points are large, the linear dependency among the factors of the proposed maximin LHDs are insignificant.

CHAPTER 6

COMPUTATIONAL EXPERIMENTS AND DISCUSSION REGARDING RECTANGULAR DISTANCE

6.1 Introduction

In the Previous chapter the experiments in which distances are measured in Euclidian distance measure have been performed. In this chapter Rectangular distance measure will be considered. It is known that for any two points $x = (x_1, x_2, x_3, \ldots\ldots, x_n)$ and $y = (y_1, y_2, y_3, \ldots\ldots, y_n)$, the distance between this points, in Rectangular measure (L^1), be

$$d(x, y: R) = \sum_{i=1}^{s} |x_i - y_i| \tag{6.1}$$

It is noted that for the L^1-distance measure, to find optimal LHD is more complicated. Anyway the purpose is not to perform experiments to optimize LHD by ILS in Rectangular distance measure; rather the maximin LHD in Rectangular distance measure will be studied where the designs are optimized by ILS approach regarding Euclidian distance measure.

6.2 Experimental Result and Discussion

Actually in this study, the optimal LHD namely maximin LHD obtained by the proposed ILS approach (MLH-ILS) will be considered in which distance is measured in Euclidian distance measure (L^2). Then the minimum inter-site distance will be measured among the design points of the MLH-ILS design by Rectangular distance measure, (L^1). In what follow $D_1^{(L1)}$ denotes the minimum inter-site distance measured by the Rectangular distance measure. At first two experimental design namely $(N, k) = (5, 3)$; $(9, 4)$ are considered. Two experiments are performed to optimize (maximin LHD) by ILS approach in L^2 measure and find out several importance characteristics. The outputs are reputed in the Table 6.1 and 6.2 at columns 4 and 5 respectively. Now these characteristics of the designs are compared with that of literature. The comparisons are mentioned in the Tables 6.1 and 6.2 .

Table 6.1: The comparison of MLH-ILS vs MLH-SA and OMLH – MSA for $(N, k) = (5, 3)$

Method →	**MLH-SA**	**OMLH-MSA**	**MLH-ILS**
Optimal Latin Hypercube Design Matrix →	1 1 2 2 5 3 3 2 5 4 3 1 5 4 4	1 2 3 2 4 5 3 5 1 4 1 2 5 3 4	1 3 5 2 2 2 3 5 1 4 4 4 5 1 3
Optimal Criteria →	Φ_p	Φ_p , ρ^2	**(Φ, D_1)**
Distance measure →	L^1	L^1	L^2
PROPERTIES ↓			
ρ →	0.265	**0.0816**	0.200
ρ_{max} →	0.4	**0.1**	0.200
$D_1(J_1)^{(L1)}$ →	**5(3)**	5(4)	5(6)
$\Phi^{(L1)}$ →	**0.2170**	.2201	0.21879
$D_1(J_1)^{(L2)}$ →	9(1)	9(2)	**11(6)**
$\Phi^{(L2)}$ →	0.1113	0.1151	**0.09956**

In the tables MLH-SA denotes maximin LHD obtained by Simulated annealing approach proposed by Morris and Mitchell (1995), OMLH-MSA denotes Orthogonal-maximin LHD obtained by the Modified Simulated annealing approach proposed by Joseph and Hung (2008) in which paper multi-objective criterion is considered and MLH-ILS denotes maximin LHD obtained by ILS approach as mentioned earlier which is proposed by Grosso et al. (2008) and also proposed here. In the Table 6.2 another design denoted by OLH- Y is also mentioned. Here OLH-Y denotes Orthogonal LHD obtained by the approach proposed by Ye (1998). In the tables ρ denotes average pair-wise correlation, ρ_{max} denotes maximum pair-wise correlation (Eq. 3.2 and Eq. 3.3). The optimal criterion, Φ_p is given by the Eq. (4.5), the optimal criterion Φ_p , ρ^2 denotes the multi-objective function where the algorithm optimized $\omega_1 \rho^2 + \omega_2 \Phi_p$ criterion [Joseph and Hung (2008)] where ρ is given by Eq. (3.2) whereas Φ_p is also given by Eq. (4.5) and ω_1, ω_2 are weight factors (optimized both minimum distance criterion as well as correlation criterion). On the other hand the optimal criterion (Φ, D1) is also given by Eq. (4.5) but the algorithm track the best D_1 value. Maximized minimum inter-site distance is indicated by $D_1(J_1)$. In the tables, Rectangular distance

measure is indicated by the superscript (L1) whereas superscript (L2) indicates Euclidian distance measure. The J_1, in $D_1(J_1)$, indicates number of site separated by the distance D_1.

In the Table 6.1, it is noted that the designs MLH-SA and OMLH-MSA are optimized regarding Rectangular distance measure (L^1) whereas the proposed design of this thesis – MLH-ILS is optimized regarding Euclidian distance measure (L^2). It is observed in the table that though in MLH-ILS, considered L^2 distance measure, the $D_1(J_1)^{(L1)}$ and $\Phi^{(L1)}$ values of MLH-ILS design is comparable of the other two designs. On the other hand the $D_1(J_1)^{(L2)}$ and $\Phi^{(L2)}$ values of MLH-ILS design is significantly better than those the other two designs. Now it is observed that the design OMLH-MSA is best than the other two regarding multicollinearity, since the designs are optimized regarding average correlation ρ value. But MLH-ILS design is better than the design MLH-SA regarding both the correlation measure ρ and ρ_{max}.

Table 6.2: The comparison of MLH-ILS vs MLH-SA, OMLH – MSA and OLH-Y for $(N, k) = (9, 4)$

Method →	MLH-SA	OMLH - MSA	OLH- Y	MLH- ILS
Optimal Latin Hypercube Design Matrix →	1 3 3 4 2 5 8 8 3 8 6 2 4 7 1 6 5 2 9 3 6 9 5 9 7 1 4 7 8 4 2 1 9 6 7 5	1 5 3 3 2 2 5 8 3 9 7 5 4 3 8 1 5 7 1 7 6 6 9 9 7 1 2 4 8 8 4 2 9 4 6 6	1 2 6 3 2 9 7 6 3 4 2 9 4 7 1 2 5 5 5 5 6 3 9 8 7 6 8 1 8 1 3 4 9 8 4 7	1 5 8 4 2 7 4 9 3 2 1 6 4 8 3 3 5 1 5 1 6 3 7 8 7 6 9 2 8 9 6 7 9 4 2 5
Optimal Criteria →	Φ_p	Φ_p , ρ^2	$\rho = 0$	Φ_p, D_1
Distance measure →	L^1	L^1	L^1	L^2
PROPERTIES ↓				
ρ →	0.108	0.063	0.000	0.151
ρ_{max} →	0.217	0.117	0.000	0.233
$D_1(J_1)^{(L1)}$ →	**11(3)**	11(4)	10(8)	10(4)
$\Phi^{(L1)}$ →	**0.105**	**0.105**	0.115	0.108
$D_1(J_1)^{(L2)}$ →	33(2)	31(1)	30(8)	**42(6)**
$\Phi^{(L2)}$ →	0.031	0.033	0.037	**0.026**

Again in the Table 6.2, it is observed that the designs MLH-SA, OMLH-MSA and OLH-Y are optimized regarding Rectangular distance measure (L^1) whereas the proposed design – MLH-ILS is optimized regarding Euclidian distance measure (L^2). It is observed in the table that though L^2 distance measure is considered in MLH-ILS design, the $D_1(J_1)^{(L1)}$ and $\Phi^{(L1)}$ values of MLH-ILS design are comparable with respect to the other three designs. On the other hand $D_1(J_1)^{(L2)}$ and $\Phi^{(L2)}$ values of MLH-ILS design are significantly better than the other three designs considered. It is also noticed that regarding correlation parameters ρ and ρ_{max}, OLH-Y design is better comparing with the other three designs but regarding $D_1(J_1)^{(L2)}$, OLH-Y design is worst one. It is noted that in the design OMLH-MSA, correlation criterion ρ is partially minimized and in the OLH-Y design, designs are chosen so that ρ be zero. On the other hand MLH-ILS design is comparable with both the designs MLH-SA and OMLH-MSA with respective to correlation parameter ρ and ρ_{max}.

Now some experiments will be performed to compare D_1 values (minimum inter-site distance) measured in Rectangular distance measure of the designs available in the web www.spacefillingdesigns.nl. The L^1 based maximin LHDs, available in the web, are denoted by MLH-Web. Note that the maximin designs, considered here from the web, are optimized regarding Rectangular distance measure. On the other hand it is mentioned here again that the proposed designs MLH- ILS is optimized regarding Euclidian distance measure. So in this experiments the $D_1^{(L1)}$ values are just calculated from the MLH-ILS designs which is optimized regarding L^2 distance measure. Note that in the Table 6.3 the symbol $D_1^{(L1)}$ denotes minimum inter-site distance among the points of a design in which distance is measured in Rectangular distance measure. For this experiments factors $k = 3, 4, \ldots, 6$ and number of points $N = 4, 5, \ldots, 25$ are considered. It is noted that there are few values available in the literature for L^1 measure. It is observed in the Table 6.3 that, though the designs MLH-ILS are optimized regarding L^2 measure but the $D_1^{(L1)}$ values of MLH-ILS designs are comparable with that of MLH-Web designs in which designs are optimized regarding L^1 distance measure.

Again some experiments will be performed to find out the $D_1^{(L2)}$ (minimum inter-site distance in Euclidian distance measure) values of the MLH-Web designs considered in the previous experiment. Note that, the designs MLH-Web are optimized in Rectangular distance measure

L^1 rather than L^2 distance measure which are available in the web www.spacefillingdesigns.nl. Now regarding $D_1^{(L2)}$ values of MLH-ILS which is optimized in L^2 measure are compared with MLH-Web, which is optimized in L^1 measure. The experimental results are reputed in the Table 6.4. Note that there are few designs are available in the web regarding MLH-Web as mention earlier. It is noticed that the designs MLH- ILS outperform the designs MLH-web regarding L^2. It is remarked that when number of design points N and/ or k are large, the performance of ILS approach is much better.

Table 6.3: The comparison of MLH-ILS vs MLH-Web regarding rectangular distance measure (L^1) for $k = 3, 4, 5, 6$

N	$k=3$		$k=4$		$K=5$		$K=6$	
	MLH-ILS ($D_1^{(L1)}$)	MLH-Web ($D_1^{(L1)}$)	MLH-ILS ($D_1^{(L1)}$)	MLH-Web ($D_1^{(L1)}$)	MLH-ILS ($D_1^{(L1)}$)	MLH-Web ($D_1^{(L1)}$)	MLH-ILS $D_1^{(L1)}$	MLH-Web ($D_1^{(L1)}$)
4	4	4	6	6	8	8	10	10
5	5	5	7	7	10	10	11	12
6	6	6	8	8	10	11	14	14
7	6	6	8	10	12	12	14	16
8	7	7	10		13		16	
9	8	8	10		13		17	
10	7	8	12		15		19	
11	8	8	11		15		19	
12	8	9	13		17		23	
13	9	10	12		17		21.	
14	9	10	14		19		24	
15	10	11	14		17		22	
16	9	11	14		19		24	
17	10		14		19		26	
18	10		16		19		27	
19	10		16		21		26	
20	10		18		21		29	
21	11		20		25		29	
22	11		17		23		31	
23	11		18		26		32	
24	11		19		26		33	
25	13		19		27		34	

Table 6.4: The comparison of MLH-ILS vs MLH-Web regarding Euclidian distance measure (L^2) for $k = 3, 4, 5, 6$

N	*k*=3		*k*=4		*K*=5		*K*=6	
	MLH-ILS ($D_1^{(L2)}$)	MLH-Web ($D_1^{(L2)}$)	MLH-ILS ($D_1^{(L2)}$)	MLH-Web ($D_1^{(L2)}$)	MLH-ILS ($D_1^{(L2)}$)	MLH-Web ($D_1^{(L2)}$)	MLH-ILS ($D_1^{(L2)}$)	MLH-Web ($D_1^{(L2)}$)
3	6	6	7	7	8	8	**12**	8
4	6	6	12	12	14	14	**20**	18
5	11	9	**15**	13	**24**	22	**27**	24
6	14	14	**22**	18	**32**	27	**40**	36
7	**17**	12	**28**	26	**40**	32	52	52
8	21	21						
9	22	22						
10	**27**	22						
11	**30**	22						
12	**36**	27						
13	**41**	36						
14	**42**	34						
15	**48**	41						
16	**50**	41						

CHAPTER 7
DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

7.1 Introduction

The designs of experiments are used in a wide range of application to learn about the effect of input variables on a response of interest and also finding out the redundancy of the factors. As it is recognized, the choice of the design points for computer experiments should fulfill at least two properties - *space-filling and non-collapsing*. The *multicollinearty,* another property, is also important, because if two factors are correlated then it will be not possible to distinguish between the effects of the two factors based on this experiment. Latin Hypercube Designs (LHDs), a widely used experimental design, fulfilled the *non-collapsing* property. The experimental design namely Maximin LHD obtained by ILS approach has nice *space-filling* as well as by default *non-collapsing* property. The main goal of this thesis was to analyze the *multicollinearty* of the designs obtained by the ILS approach. Another goal was to observe the evenly distribution of the design points regarding rectangular distance measure.

7.2 Discussions

The importance of high performance algorithms for tackling difficult optimization problems cannot be understated. In many cases the only available methods are *metaheuristics*. The word *metaheuristics* contains all heuristics methods that show evidence of achieving good quality solutions for the problem of interest within an acceptable time. *Metaheuristic* techniques have become more and more competitive. When designing a *metaheuristic*, it is preferable that it be simple, both conceptually and in practice. Naturally, it also must be effective and if possible for general purpose. The main advantage of this approach is the ease of implementation and the quickness.

In this thesis Iterated Local Search (ILS) approach, a well known *metaheuristics* method has been considered, for finding optimal solution in the discrete space. For the initial solution (design), randomly generated Latin Hypercube Design has been considered which has very good non-overlapping property. For optimal criteria, it is

considered that maximize the minimum inter-site distance of the design points in Euclidian distance measure.

In **Chapter 2,** an attempt has been made to overview the basic constituents and properties of experimental design. As in the introduction the necessity of the experimental design has been discussed, so chapter 2 has argued about basic principles of experimental design, experimental error etc. Moreover in this chapter the requirements of a good experiment have been mentioned in brief. Also the several type of the experimental designs have been discussed.

In **Chapter 3,** another attempt has been made to overview the basic constituents and properties of correlation and multicollinearity. Mainly an overview of correlation is discussed in this chapter. This chapter is pointed out several types of correlations. How they are measured is also discussed here. Here the effect of multicollinearity on the model is also briefly discussed.

In **Chapter 4,** the heuristic approach mainly Iterated Local Search (ILS) approach has been discussed and presented in brief. ILS approach for optimizing LHD is elaborately presented here as well as the various optimal criteria for maximizing the minimum inter-site distance of the design point. The definition of maximin LHD is also given here for the proposed ILS approach.

In **Chapter 5,** the experimental analysis is performed extensively regarding Euclidian distance measure. At first the performance of the algorithm is compared with available one in the literature regarding inter-site Euclidian distance measure. From the experimental design it is shown that the algorithm is state-of-arts regarding maximin LHD. Then these optimal designs are considered for the multicollinearity analysis of the factors of each design.

In **Chapter 6** several experiments performed regarding Rectangular distance measure have been presented. Then several characteristics of the designs are compared

with available ones in the literature. Mainly average correlation coefficients as well as maximum correlation coefficients are considered for the comparison.

7.3 Concluding Remarks

Essentially, the credibility of Iterated Local Search (ILS) algorithms relies on their ability to solve difficult, real-world problems with the minimal amount of human effort. If it cannot make the successful transition from academic exercise to physical application it will be abandoned in favor of other techniques.

One of the goals of this thesis was to implement the ILS approach in PC window environment. It is worthwhile to mention here again that the ILS approach is successfully implemented in a *Sun-Cluster* with *Solar* OS for maximin LHD. Some time it is very difficult to implement algorithms from one environment to another environment. Here the algorithms have been implemented in windows PC environment (Windows visual C++) by considering some modification.

In this thesis two kinds of distance measure namely Euclidian distance measure (L^2) and Rectangular distance measure (L^1) to calculate the inter-site distance of each two design points are considered. But in the experiments for optimization (maximization) the minimum inter-site distance, only the Euclidian distance measure was considered. Then the L^1 distance of the optimized points was also calculated. It is shown that ILS approach able to obtain a large amount of best experimental design compared to available one in the literature.

It is mentioned earlier that a good experimental design should have *space-filling*, *non-collapsing* and non-correlation (poor *multicollinearity*) property. It is known that LHD, by default, has good *non-collapsing* property. On the other hand for the achievement of *space-filling* property, one should optimize the experimental design like LHD in some criteria so that design points spread over the design space evenly. Maximin (maximization of the minimum inter-site distance) is one of the criteria which is frequently used to optimize experimental design regarding fulfilling the space-filling properties. It is also mentioned earlier that maximin LHD problem is discrete

optimization problem. That is the search space is discrete and number of search points is huge namely $(N!)^{k-1}$. It is also known that the optimal experimental design does not need necessarily exact optimal one regarding optimal criteria; rather it should cover three above mentioned criteria. Therefore m*eta-heuristic* approaches are frequently used in the real world problems like finding experimental design. In this thesis Iterated Local search (ILS) approach, which is a meta-heuristic approach, considered for the maximin LHD. For the distance measure, Euclidian distance measure is considered.

ILS approach has the ability to obtain a huge number of improved optimal designs comparable to the literature, regarding *space-filling* criteria. That is ILS approach is able to find out an experimental design which is both space-filling and non-collapsing, when it is applied on LHDs.

Another main interest of this thesis was to analyze the experimental design regarding another important property namely *multicollinearity* property. The final goal was to find out the tolerance of the designs regarding *non-correlation* property. For the measure of the multicollinearity property, average pair-wise correlation as well maximum pair-wise correlation was considered. For the calculation of average pair-wise correlation the following well known formula (see section 4.5) is used

$$\rho^2 = \frac{\sum_{i=1}^{k}\sum_{j=i+1}^{k}\rho_{ij}^2}{k(k-1)/2}$$

Where ρ_{qr} denotes the single pair simple correlation between the factors p and r. For calculating the correlation between each single pair of factors (say column q and column r), the well known Kearl Pearson's product moment formula is used:

$$\rho_{qr} = \frac{N\sum x_{iq}x_{ir} - \sum x_{iq}\sum x_{ir}}{\sqrt{N\sum x_{iq}^2 - (\sum x_{iq})^2}\sqrt{N\sum x_{ir}^2 - (\sum x_{ir})^2}}$$

where the sums over $i = 1, \cdots, N$. Here Y is better than X if $\rho^2(Y) < \rho^2(X)$

and the maximum correlation is defined as follow

$$\rho_{\max} = \max_{1 \le i, j \le k} \rho_{ij}$$

At first several experiments were performed to analyze the average pair-wise correlation among the factors of each optimal design obtained by ILS approach. For the experiments number of factors (dimensions) k= 3,..., 10 and number of design points (rows) $N = 5 \times i; i = 1, \cdots, 20$ are considered. For the measure of the inter-site distance is considered in Euclidian distance measure (L^2). Experimentally it is shown that most of the designs have tolerable average pair-wise correlation as well as maximum correlation. It is remarkable that when number of design points is large then average pair-wise correlation is less than 0.1. It is also worthwhile to mention here again that the maximin LHDs, obtained by ILS approach, are compared with available ones in the literature. The maximin LHDs obtained by the ILS approach are comparable with those available in the literature even in term of non-collinearity they perform much better when number of design points and /or are large.

Several experiments regarding Rectangular distance measure (L^1) also have been performed. Several designs with fix number of factors as well as points are considered, and then the designs are analyzed in several aspects. It is noted that the designs are optimized regarding Euclidian distance measure but also measurement of inter-site pair of points in rectangular distance measure along with Euclidian distance measure is denoted. Though the algorithms maximized the minimum inter-site distance of the design in Euclidian distance measure, the design also evenly spread over the design space with respect to Rectangular distance measure. It is also noted that all the properties considered are comparable with that of literature.

Finally it may be concluded that the proposed ILS algorithm is state-of arts regarding maximin LHD in Euclidian distance measure. Though the algorithm considered Euclidian distance measure and optimized maximin LHD, the designs have tolerable correlations among the factors. Moreover though the algorithm optimized the design space regarding *space-filling* criterion in Euclidian measure, the design space is almost evenly spread over the space in Rectangular distance measure too.

REFERENCE

1. Audze P., and Eglais V., New approach to planning out of experiments, problems of dynamics and strength, vol. 35, pp. 104107(1977). (in Russian).
2. Barthelemy J. F. M., and Haftka R. T., Approximation concepts for optimum structural design – A review, Structural Optimization, vol. 5(3), pp.129-144(1993).
3. Bates, R. A., Riccomagno E., Schwabe R., and Wynn H. P., Lattices and dual lattices in optimal experimental design for fourier models. Computational Statistics and Data Analysis 28, 283–296, (1998).
4. Bates S. J., Sienz J., and Toropov V. V., Formulation of the optimal Latin hypercube design of experiments using a permutation genetic algorithm, AIAA 2004-2011, pp. 1-7(2004).
5. Baum, E. B., Iterated descent: A better algorithm for local search in combinatorial optimization problems. Technical report, Caltech, Pasadena, CA, (1986). manuscript.
6. Box, G. E. P. and Draper N. R., A basis for the selection of a response surface design. J. Amer. Statist. Assoc. 54 622-654, (1959).
7. Box, G. E. P. and Draper N. R., Evolutionary Operation. John Wiley, New York, NY, (1969).
8. Box, G. E. P. and Wilson K. B., On the experimental attainment of optimum condition Journal of the Royal Statistical Society, Series B 13, 1–45, (1951).
9. Butler N. A., Optimal and orthogonal Latin hypercube designs for computer experiments, Biometrika,vol. 88(3), pp. 847-857(2001).
10. Cioppa, T. M., Efficient nearly orthogonal and space-filling experimental designs for high-dimensional complex models, PhD thesis, naval postgraduate school Monterey, California,USA,(2002).
11. Crary S. B., Cousseau P., Armstrong D., Woodcock D. M., Mok E. H., Dubochet O., Lerch P., and Renaud P., Optimal design of computer experiments for metamodel generation using I-OPTTM, Computer Modeling in Engineering & Sciences, vol. 1(1), pp. 127-139(2000).

12. Crombecq K. , Laermans E., Dhaene T.; Efficient space-filling and non-collapsing sequential design strategies for simulation-based modeling, European Journal of Operational Research, Elsevier, vol.-214, pp. 683-696 (2011).
13. Currin C., Mitchell, T., Morris, M. D., and Ylvisaker D., Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments, Journal of the American Statistical Association,vol. 86, pp. 953-963 (1991).
14. Dam E. R. van, Two-dimensional minimax Latin hypercube designs, CentER Discussion Paper 2005-105. Tilburg University, (2005).
15. Dam E. R. van, Husslage B. G. M. , Hertog D. den and Melissen J. B. M., *Maximin latin hypercube designs in two dimensions*, Operations Research vol. 55(1), pp. 158–169(2007a).
16. Dam E. R. van, Rennen G., and Husslage B., *Bounds for maximin Latin hypercube designs,* Department of Econometrics and Operations Research, Diss. Paper, Tilburg University, The Netherlands,(2007b).
17. Fang K. T., Ma C., and Winker P., Centered L2-Discrepancy of Random Sampling and Latin Hypercube Design, and Construction of Uniform Designs, Mathematics of Computation, vol. 71(237), pp. 275-296 (2000a).
18. Fang, K. T., Lin D. K. J. , Winkler P., and Zhang Y., Uniform design: theory and application. Technometrics 42, 237–248, (2000b).
19. Fang K. T., Li R., and Sudjianto A., Design and Modeling for Computer Experiments, CRC Press, New York, (2006).
20. Farrar Donald E. and Glauber, Robert R., "Multicollinearity in Regression Analysis: The Problem Revisited, " *The Review of Economics and Statistics* 49(1):92-107, (1967).
21. Finney, D. J., Fractional replication of factorial arrangments. Annals of Eugenics 12, 291–301, (1945).
22. Fisher, R. A., The Design of Experiments. Oliver and Boyd, Edinburgh, (1935).
23. Grosso A., Jamali A. R. M. J. U. and Locatelli M., Iterated Local Search Approaches to Maximin Latin Hypercube Designs, Innovations and Advanced Techniques in Systems, *Computing Sciences and Software Engineering*, *Springer Netherlands,* pp. 52-56, (2008).

24. Grosso A., Jamali A. R. J. U., and Locatelli M., Finding Maximin Latin Hypercube Designs by Iterated Local Search Heuristics (2008), *European Journal of Operations Research, Elsevier,* Vol. 197, pp. 541-547 (2009).
25. Giunta A. A., Wojtkiewicz S. F., and Eldred M. S., Overview of modern design of experiments methods for computational simulations, AIAA 2003-0649, pp. 1-17(2003).
26. Hammersley, J. M., Monte Carlo methods for solving multivariable problems. Annals of the New York Academy of Sciences 86, 844–874, (1960)..
27. Hill, W. J. and Hunter W. G., A review of response surface methodology: a literature survey. Technometrics 8, 571–590, (1966).
28. Husslage B., van Dam E. R., and Hertog D. den, Nested maximin latin hypercube designs in two dimensions, CentER Discussion Paper No. 200579(2005).
29. Husslage B., Rennen G., van Dam E. R., and Hertog D. den, Space-Filling Latin Hypercube Designs for Computer Experiments, CentER Discussion Paper No. 2006-18 (2006).
30. Iman R. L., and Conover W. J., A distribution-free approach to inducing rank correlation among input variables, Comm. Stat. Part B – Simulation Comput., vol. 11, pp. 311-334 (1982a).
31. Iman, R. L. and Conover W. J., Small-sample sensitivity analysis techniques for computer models, with an application to risk assessment. Communications in Statistics – Part A, Theory and Methods 17, 1749–1842, (1982b).
32. Jamali, A. R. M. J. U., Heuristic Approaches for Maximin Distance and Packing Problems. Ph.D. dissertation, Dipartimento di Informatica, Universitá degli Studi di Torino, Turin, Italy, (2009).
33. Jin R., Chen W., and Sudjianto A., An efficient algorithm for constructing optimal design of computer experiments, Journal of Statistical Planning and Inference, vol. 134(1), pp. 268-287(2005).
34. Johnson, M. E., Moore L. M., and Ylvisaker D., Minimax and maximin distance designs. Journal of Statistical Planning and Inference, vol. 26, pp.131–148, (1990).
35. Joseph V. R. , and Hung Y., Orthogonal-Maximin Latin Hypercube Designs, *Statistica Sinica*, vol. 18, pp 171-186 (2008).

36. Jourdan A. and Franco J.: Optimal Latin hypercube designs for the Kullback–Leibler criterion, AStA Advances in Statistical Analysis, Springer-Verlag , Volume 94, Number 4, pp- 341-351, DOI: 10.1007/s10182-010-0145-y, (2010).

37. Jourdan, A., Franco, J. : Plans d'experiences numeriques d'information de Kullback–Leibler minimale. J. Soc. Fr. Stat., **150**(2), 52–64 (2009).

38. Kiefer, J. C., Jack Carl Kiefer Collected Papers 3: Design of Experiments (L. D. Brown, I. Olkin, J. Sacks, and H. P. Wynn, eds.). Springer, New York, (1985).

39. Kiefer, J. and Wolforitz, J., The equivalence of two extremum problems. The Canadian Journal of Mathematics 12, 363–366, (1960).

40. Kirkpatrick S., Gelatt Jr. C. D., and Vecchi M. P., Optimization by Simulated Annealing, Science, vol. 220, pp. 671-680 (1983).

41. Leary S., Bhaskar A., and Keane A., Optimal orthogonal-array-based Latin hypercubes, Journal of Applied Statistics, vol. 30(5), pp. 585-598(2003).

42. Lee T. H., and Jung J. J., Maximin Eigenvalue Sampling of Kriging model, 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 30 August - 1 September 2004, Albany, New York, (2000).

43. Li W., and Kenny Q. Y., Optimal Symmetric Latin Hypercube Designs, (2009). Available in http://www.csom.umn.edu/WWWPages/Faculty/WLi/research/super.ps.

44. Li W. W., and Wu C. F. J., Columnwise-Pairwise Algorithms With Applications to the Construction of Supersaturated Designs, Technometrics, American Society for Quality Control and American Statistical Association, Alexandria, Va, USA, vol. 39(2), pp. 171-179(1997).

45. Liefvendahl M., and Stocki R., A study on algorithms for optimization of Latin hypercubes, Journal of Statistical Planning and Inference, vol. 136 (9), pp. 3231-3247(1 September 2006).

46. Lin D. K. J., and Steinberg D. M., A Construction Method for Orthogonal Latin Hypercube Designs, Biometrika, Oxford University Press, vol. 93(2), pp. 279 - 288(2006).

47. Lourenco H. R. , Martin O. C., and Stutzle T., In Iterated Local Search Handbook of Metaheuristics, ISORMS 57(Eds.: Glover F., and G. Kochenberger,), Kluwer, pp. 321-353 (2002).

48. Lourenco H. R., Marting O., and Stutzl T., A beginner's introduction to Iterated Local Search, In Proceedings of MIC'2001-Meta-heuristics International Conference, Porto-Portugal, vol. 1, pp. 1-6(2001).
49. Martin O. and Otto S. W., Combining simulated annealing with local search heuristics. Annals of Operations Research, 63:57–75, (1996).
50. McKay M. D., Beckman, R. J., and Conover W. J., A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, vol. 21, pp. 239-245(1979).
51. Morris M. D., Factorial plans for preliminary computational experiments, Technometrics, vol. 33, pp. 161-74 (1991).
52. Morris M. D., and Mitchell T. J., Exploratory designs for computer experiments, Journal of Statistical Planning and Inference, vol. 43, pp. 381 402(1995).
53. Myers, R. H. and Montgomery D. C., Response Surface Methodology: Process and Product Optimization Using Designed Experiments. John Wiley & Sons, New York, NY, (1995).
54. Olsson A., Sandberg G., and Dahlblom O., On Latin hypercube sampling for structural reliability analysis , Structural Safety, Elsevier, vol. 25(1), pp. 47-68(January 2003).
55. Owen, A. B., Controlling correlations in Latin hypercube samples. Journal of the American Statistical Association, vol. 89, pp. 1517-1522 (1994).
56. Park J. S., Optimal Latin hypercube designs for computer experiments, Journal of Statistical Planning and Inference, vol. 39, pp. 95-111(1994).
57. Plackett, R. L. and Burman J. P., The design of multifactorial experiments. Biometrika 33, 305–325, (1946).
58. Ross S. M. ; Introductory Statistics, Second edition, Elservier, (2005).
59. Sacks, J. and Ylvisaker, D., Statistical designs and integral approximation. In Proc. 12th Bien. Sem. Canad. Math. Congress (R. Pyke, ed.) 115-136. Canadian Mathematical Congress, Montreal, (1970).
60. Sacks, J., Schiller, S. B. and WELCH, W. J., Designs for computer experiments. Technometrics 3141-47, (1989).
61. Santner T. J. , Williams B. J., and Notz W. I., The design and analysis of computer experiments, Springer Series in Statistics, Springer-Verlag, New York, (2003).

62. Schmid, W. Ch., (t,m, s)-Nets: Digital Constructions and Combinatorial Aspects. Ph.D. dissertation, Institute of Mathematics, University of Salzburg, Salzburg, Austria, (1995).

63. Sebastiani P., and Wynn H. P., Maximum entropy sampling and optimal Bayesian experimental design, J. R. Statist. Soc., vol. 62(1), pp. 145-157(2000).

64. Shewry M. , Wynn H., Maximum entropy design, Journal of Applied Statistics vol. 14, pp. 165–170 (1987).

65. Sobieski -S. J., and Haftka R. T., Multidisciplinary aerospace design optimization: Survey of recent developments, Structural and Multidisciplinary Optimization, vol. 14(1), pp. 1-23(1997).

66. St. John, R. C. and Draper N. R., D-optimality for regression designs: a review. Technometrics 17, 15–23, (1975).

67. Stein M., Large sample properties of simulation using Latin hypercube sampling, Technometrics, vol. 29, pp. 143-51 (1987).

68. Steinberg, D. M. and Hunter W. G., Experimental design: review and comment (with discussion). Technometrics 26, 71–130, (1984).

69. Steinberg, G. D. M. and Dennis K. J. N., A construction method for orthogonal Latin hypercube designs, Biometrika, vol. 93 (2), pp. 279-288 (2006).

70. Stinstra E. D., Hertog D. den, Stehouwer H. P., and Vestjens A., Constrained maximin designs for computer experiments. Technometrics, vol.45(4), pp. 340-346(2003).

71. St¨utzle, T., Local Search Algorithms for Combinatorial Problems — Analysis, Improvements, and New Applications. PhD thesis, Darmstadt University of Technology, Department of Computer Science, (1998).

72. Tang, B., Orthogonal array-based Latin hypercubes. Journal of the American Statistical Association 88, 1392–1397, (1993).

73. Tang B. X., Selection Latin hypercube designs using correlation criteria, Statist. Sinica, vol. 8, pp. 965 -77(1998).

74. Viana, F. A. C. , Venter G., and Balabanov V., An algorithm for fast optimal Latin hypercube Design of experiments, International Journal For Numerical Methods In Engineering, Wiley inter science (DOI: 10.1002/nme.2750), vol. 82, pp. 135–156 (2010).

75. Wang G. G., Adaptive Response Surface Method Using Inherited Latin Hypercube Design Points, Transactions of the ASME, Journal of Mechanical Design, vol. 125, pp. 210-220(June 2003).
76. Wang, J. C. and Wu C. F. J., An approach to the construction of asymmetrical orthogonal arrays. Journal of the American Statistical Association 86, 450–456, (1991).
77. Welch, W. J., A mean squared error criterion for the design of experiments. Biometrika 70 205 -213, (1983).
78. Xu H., Universally optimal designs for computer experiments, Statistica Sinica, vol. 9, pp. 1083-1088 (1999).
79. Ye, K. Q., Orthogonal column Latin hypercubes and their application in computer experiments. Journal of the American Statistical Association 93, 1430-1439, (1998).
80. Ye, K. Q,, Li W., and Sudjainto A., Algorithmic construction of optimal symmetric Latin hypercube designs, Journal of Statistical Planning and Inference, vol. 90, pp. 145-159(2000).
81. Ylvisaker, D., Designs on random fields. In A Survey of Statistical Design and Linear Models (J. N. Srivastava, ed.) 593-607, (1975). North-Holland, Amsterdam.