

On Bayes and Nash experimental designs for hypothesis testing problems

Satya Prakash Singh

*Department of Mathematics
Indian Institute of Technology Hyderabad
Telangana 502285, India
e-mail: spsingh@math.iith.ac.in*

Ori Davidov

*Department of Statistics
University of Haifa, Mount Carmel
Haifa 3498838, Israel
e-mail: davidov@stat.haifa.ac.il*

Abstract: In this communication we examine the relationship between maxi–min, Bayes and Nash designs for some hypothesis testing problems. In particular we consider the problem of sample allocation in the standard analysis of variance framework and show that the maxi–min design is also a Bayes solution with respect to the least favourable prior, as well as a solution to a game theoretic problem, which we refer to as a Nash design. In addition, an extension to tests for order is provided.

MSC2020 subject classifications: 62K05.

Keywords and phrases: ANOVA, Bayesian design, maxi-min design, Nash equilibrium, power.

Received October 2019.

Contents

1	Introduction	3977
2	The maxi–min design for the ANOVA problem	3977
3	Bayes designs	3981
4	Nash designs	3983
	4.1 Standard ANOVA: The game theoretic view	3984
	4.2 Comparison of multiple treatments with a control: The game theoretic view	3989
5	Summary and discussion	3991
	Appendix	3992
	Proof of Theorem 3.1	3993
	Proof of Theorem 3.2	3994
	Proof of Theorem 3.3	3996
	Proof of Theorem 4.1	3996
	Proof of Theorem 4.2	3999
	Proof of Theorem 4.3	4001

Acknowledgments 4002
 References 4002

1. Introduction

The main focus of the literature on optimal design is on the theory of optimal estimation, usually in the context of the classical linear model. Within this framework an optimal design is a design which either minimizes (or maximizes) some functional of the variance matrix of the estimated parameter, cf., [17] and [1]. For example, an A-optimal design is a design which minimizes the sum of the eigenvalues of the variance matrix whereas a D-optimal design minimizes their product, which equals of course, to the determinant of the variance matrix. There are many other optimality criteria such as E, and MV optimality. Collectively these are known as alphabetic optimality.

Optimal estimation, however, may not be the objective of every scientific study. In fact, there are many situations in which researchers are interested in designing an experiment tailored for optimal testing of hypotheses rather than optimal estimation. In this context [20] recently proposed a maxi-min design (MM-design) for optimal testing under order restrictions. In this communication we study the relationship between their MM-design and Bayesian designs for unordered as well as ordered hypothesis testing problems. In addition we show that the MM-design can be viewed as a Nash Equilibrium when the experimental design is viewed as game theoretic problem. For a broad perspective on the Nash equilibrium see [11].

2. The maxi-min design for the ANOVA problem

We focus on the one way layout in which

$$Y_{ij} = \mu_i + \epsilon_{ij}, \tag{2.1}$$

where Y_{ij} is the response of the j^{th} subject in the i^{th} treatment group $i = 1, \dots, K$ and $j = 1, \dots, n_i$. We further assume that the errors ϵ_{ij} are independent $\mathcal{N}(0, \sigma^2)$ random variables (RVs) and without any loss of generality we fix $\sigma^2 = 1$. A brief discussion of the situation where σ^2 is unknown differed to Remark 2.2 appearing near the end of this Section.

Consider testing the standard ANOVA hypotheses

$$H_0 : \boldsymbol{\mu} \in \mathcal{M}_0 \quad \text{versus} \quad H_1 : \boldsymbol{\mu} \notin \mathcal{M}_0, \tag{2.2}$$

where $\mathcal{M}_0 = \{\boldsymbol{\mu} \in \mathbb{R}^K : \mu_1 = \mu_2 = \dots = \mu_K\}$, i.e., under the null all means are equal, whereas under the alternative at least one pair of means is different, i.e.,

$\mu_i \neq \mu_j$ for some $i \neq j$. The standard test for (2.2) is the likelihood ratio test (LRT) given by

$$T_N = \sum_{i=1}^K n_i (\bar{Y}_i - \bar{Y})^2, \quad (2.3)$$

where $\bar{Y}_i = n_i^{-1} \sum_{j=1}^{n_i} Y_{ij}$ is the average in the i^{th} group when $n_i > 0$ and 0 otherwise, $N = \sum_{i=1}^K n_i$, and $\bar{Y} = N^{-1} \sum_{i=1}^K \bar{Y}_i$ is the grand mean. It is well known that (2.3) follows a non-central chi-square distribution, denote $\chi_\nu^2((N/2)\lambda)$ with $\nu = \sum_{i=1}^K \mathbb{I}_{\{n_i > 0\}} - 1$ degrees of freedom and non-centrality parameter (NCP) $(N/2)\lambda$. Further note that

$$\lambda = \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) = \sum_{i=1}^K \xi_i (\mu_i - \bar{\mu})^2, \quad (2.4)$$

where $\xi_i = n_i/N$ and $\bar{\mu} = \sum_{i=1}^K \xi_i \mu_i$. Clearly under the null $\lambda = 0$ so the null distribution is the usual central chi-square. As a shorthand, and for convenience, we will often refer to λ as the NCP.

The power of the LRT (2.3), as a function of $(\boldsymbol{\mu}, \boldsymbol{\xi})$, is:

$$\pi(\boldsymbol{\mu}; \boldsymbol{\xi}) = \mathbb{P}_{\boldsymbol{\mu}; \boldsymbol{\xi}}(T_N \geq c_{\nu, \alpha}) \quad (2.5)$$

where the critical value $c_{\nu, \alpha}$, which solves $\sup_{\boldsymbol{\mu} \in \mathcal{M}_0} \mathbb{P}_{\boldsymbol{\mu}; \boldsymbol{\xi}}(T_N \geq c_{\nu, \alpha}) = \alpha$, is design dependent since $\nu = \nu(\boldsymbol{\xi})$. In this communication we focus on what are known in the literature (cf., [1]) as *approximate designs*. An approximate design $\boldsymbol{\xi} \in \Xi$, where $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_K)^T$ is a vector of weights, i.e., $\xi_i \geq 0$ and $\sum_{i=1}^K \xi_i = 1$, in which each ξ_i represents the proportion of observations assigned to the i^{th} treatment. Naturally the design space Ξ is the unit simplex. It is obvious that once an optimal approximate design is found, an *exact design*, i.e., a vector $\mathbf{n} = (n_1, n_2, \dots, n_K)^T$ such that $\sum_{i=1}^K n_i = N$ is fixed can be obtained by efficient rounding, see [12].

A design $\boldsymbol{\xi}$ which maximizes the power function (2.5) at a given value of $\boldsymbol{\mu}$ is said to be locally optimal. A pair of treatments, i and j say, is *maximally separated* if $|\mu_i - \mu_j| = \max\{|\mu_s - \mu_t| : 1 \leq s, t \leq K\}$. Theorem 1 in [20] states that if (i, j) is a maximally separated pair then the optimal design is $\boldsymbol{\xi}_{ij} = (\mathbf{e}_i + \mathbf{e}_j)/2$, where \mathbf{e}_l is the l^{th} standard basis of \mathbb{R}^K . In other words, given a maximally separated pair (i, j) the power of (2.3) for testing (2.2) is maximized when half of the observations are assigned to group i and the other half to group j . We refer to such a design as a two-point design. Incidentally, this design maximizes the NCP while simultaneously minimizing the degrees of freedom of the test statistic. Unfortunately, the value of $\boldsymbol{\mu}$ or the identity of the maximally separated pair are rarely known in advance. Moreover, as illustrated below, a two-point design which is optimal for some vectors of means may be grossly deficient for others.

Example 2.1. Consider an experiment with $K = 4$ treatment groups. First let us suppose that the true population means are $\boldsymbol{\mu}_1^T = (-1, 1, 0, 0)$. By Theorem 1

of [20] the design $\xi_1 = (1/2, 1/2, 0, 0)$ maximizes the power when $\mu = \mu_1$. If, however, the true vector of population means is $\mu_2^T = (0, 0, -1, 1)$, then the power maximizing design is $\xi_2 = (0, 0, 1/2, 1/2)$. Note that,

$$\pi(\mu_1; \xi_1) = \pi(\mu_2; \xi_2) = \mathbb{P}(\chi_1^2(N/2) \geq c_\alpha) > \alpha$$

for all positive N and, as expected, in both cases the power increases to unity when $N \rightarrow \infty$. However,

$$\pi(\mu_1; \xi_2) = \pi(\mu_2; \xi_1) = \mathbb{P}(\chi_1^2(0) \geq c_\alpha) = \alpha,$$

regardless of the total sample size N . This means that design ξ_2 has no power to detect a departure from the null if $\mu = \mu_1$ and likewise the design ξ_1 has no power to detect a departure from the null if $\mu = \mu_2$. We conclude that locally optimal designs may perform poorly for values of μ for which they were not designed.

It is clear from Example 2.1 that there does not exist a design ξ which is globally optimal for all μ in the alternative. Moreover, locally optimal designs may perform poorly globally. Thus, it is reasonable to seek a design which is guaranteed to perform well for all values of μ ; it is obvious that such a design will trade local optimality for a fair overall performance. In other words we seek a robust design which guarantees a minimal power, or equivalently, a design which maximize the power in the worst-case scenario.

Formally, we define ξ_{MM} and μ_{LFC} as the pair of values satisfying:

$$\pi(\mu_{LFC}; \xi_{MM}) = \max_{\xi \in \Xi} \min_{\mu \in \mathcal{M}_\delta} \pi(\mu; \xi). \tag{2.6}$$

The maximization in (2.6) is over all designs $\xi \in \Xi$ and the minimization is over the set

$$\mathcal{M}_\delta = \{\mu \in \mathbb{R}^K : \max |\mu_i - \mu_j| \geq \delta\} \tag{2.7}$$

for some $\delta > 0$. The restriction that $\mu \in \mathcal{M}_\delta$ implies that the distance within the maximally separated pair is at least δ . Heuristically, one may view δ as measuring the distance from the null; for a precise explanation see Remark 2.1 below. Further note that the restriction that $\mu \in \mathcal{M}_\delta$ is required for the existence of $\min_{\mu \notin \mathcal{M}_0} \pi(\mu; \xi)$. For more details, as well as for the form of \mathcal{M}_δ in other testing problems, the reader is referred to [20]. The quantities ξ_{MM} and μ_{LFC} are referred to as the *maxi-min design* (MM-design) and the *least favourable configuration* (LFC), respectively. Theorem 2 in [20] states that:

Theorem 2.1. *The balanced design is the MM-design in the standard ANOVA setting.*

By Theorem 2.1 the MM-design is the balanced design, i.e., $\xi_{MM} = (1/K, \dots, 1/K)$. Thus the MM-design allocates an equal proportion of observations to each of the treatment groups. It is well known that a balanced design is an A-optimal design for pairwise multiple comparison and also a maxi-min

design for estimating treatment differences under squared error loss ([26]). Further note that $\boldsymbol{\xi}_{\text{MM}}$ is independent of δ . In the proof of Theorem 2.1 it is shown that any permutation of the vector $(-\delta/2, \delta/2, 0, \dots, 0)$ is a LFC with respect to the balanced design hence the LFCs are functions of δ .

Remark 2.1 below provides a deeper understanding of the role of δ as a distance measure.

Remark 2.1. For any $\delta > 0$ set $\boldsymbol{\mu}_1 = (-(\delta-\epsilon)/2, -(\delta-\epsilon)/2, (\delta-\epsilon)/2, (\delta-\epsilon)/2)$ and $\boldsymbol{\mu}_2 = (-\delta/2, \delta/2, -\epsilon/2, \epsilon/2)$. Observe that $\boldsymbol{\mu}_1 \in \mathcal{M}_{\delta_1}$, where $\delta_1 = \delta - \epsilon$ whereas $\boldsymbol{\mu}_2 \in \mathcal{M}_{\delta_2}$ where $\delta_2 = \delta$. The NCPs associated with the MM-design and the means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are:

$$\lambda(\boldsymbol{\mu}_1; \boldsymbol{\xi}_{\text{MM}}) = \frac{(\delta - \epsilon)^2}{K} \quad \text{and} \quad \lambda(\boldsymbol{\mu}_2; \boldsymbol{\xi}_{\text{MM}}) = \frac{\delta^2}{2K} + \frac{\epsilon^2}{2K}.$$

A little algebra shows that $\lambda(\boldsymbol{\mu}_1; \boldsymbol{\xi}_{\text{MM}}) > \lambda(\boldsymbol{\mu}_2; \boldsymbol{\xi}_{\text{MM}})$ when $0 < \epsilon < \delta(2 - \sqrt{3})$. Since the power function is monotonically increasing in the NCP when the degrees of freedom are fixed it also follows that

$$\pi(\boldsymbol{\mu}_1; \boldsymbol{\xi}_{\text{MM}}) > \pi(\boldsymbol{\mu}_2; \boldsymbol{\xi}_{\text{MM}}) \quad \text{when} \quad 0 < \epsilon < \delta(2 - \sqrt{3}). \quad (2.8)$$

Equation (2.8) shows that it is not necessarily true that for any given $\boldsymbol{\xi}$ the power at $\boldsymbol{\mu} \in \mathcal{M}_{\delta_1}$ is smaller than at $\boldsymbol{\mu} \in \mathcal{M}_{\delta_2}$ even though $\delta_1 < \delta_2$. However, it is not difficult to see that

$$\lambda(\boldsymbol{\mu}_{\text{LFC}}(\delta_1); \boldsymbol{\xi}_{\text{MM}}) = \frac{(\delta - \epsilon)^2}{2K} < \frac{\delta^2}{2K} = \lambda(\boldsymbol{\mu}_{\text{LFC}}(\delta_2); \boldsymbol{\xi}_{\text{MM}}) \quad (2.9)$$

where $\boldsymbol{\mu}_{\text{LFC}}(\delta)$ is any permutation of $(-\delta/2, \delta/2, 0, \dots, 0)^T$. We conclude that $\pi(\boldsymbol{\mu}_{\text{LFC}}(\delta); \boldsymbol{\xi}_{\text{MM}})$ is monotonically increasing in δ . Armed with this perspective we can interpret δ as a “distance” from the null when the design is balanced and $\boldsymbol{\mu} = \boldsymbol{\mu}_{\text{LFC}}(\delta)$.

Remark 2.2. Finally, we note that we have assumed that (2.1) holds and σ^2 is known and for convenience we set its value to unity. If, however, σ^2 is unknown then instead of using the statistic T_n given in (2.3) the hypotheses (2.2) will be tested using the statistic

$$S_n = \frac{\sum_{i=1}^K n_i (\bar{Y}_i - \bar{Y})^2 / (K - 1)}{\sum_{i=1}^K \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 / (N - K)}, \quad (2.10)$$

whose denominator is an unbiased estimator of σ^2 . The statistic S_n follows an $F_{K-1, N-K}(\phi)$ distribution where $K - 1$ and $N - K$ are the degrees of freedom associated with numerator and denominator respectively and ϕ is the non-centrality parameter. It is well known that $\phi = (N/2)(\lambda/\sigma^2)$ where λ is as in (2.4). It immediately follows that the design which maximizes ϕ for any fixed but possibly unknown value of σ^2 will also maximize λ and visa versa. Hence our results apply in full also to situations where σ^2 is unknown and the proposed solution will also maximize the power of the F -test.

Next, in Sections 3 and 4 we examine the relationship between MM–designs with: (i) Bayes–designs; and (ii) a game–theoretic formulation of the design problem which we refer to as a Nash–designs.

3. Bayes designs

In this section we discuss the connection between MM–designs and Bayes designs in the context of the ANOVA testing problem (2.2). As noted earlier when the design criterion depends on an unknown parameter, e.g., $\boldsymbol{\mu}$ in (2.5), it is generally impossible to find a design that is globally optimal, i.e., a design which is optimal for all values of $\boldsymbol{\mu}$. One way of addressing this issue is by using MM–designs, as discussed in Section 2, while another is to adopt a Bayesian approach. In the fully Bayesian framework a design which optimizes a functional of the posterior distribution is chosen, cf., [6]. Our approach, which is not fully Bayesian, is rooted in decision theory and referred to in the literature as the *pseudo–Bayesian* approach ([25]). It is also known as *average optimal design*, e.g., [10]. The idea is simple; choose the design which optimizes the expected value of a design criterion with respect to some prior distribution on the unknown parameter. Formally, we define the Bayes design as

$$\boldsymbol{\xi}_Q^\Psi = \arg \max_{\boldsymbol{\xi} \in \Xi} \int_{\mathbf{M}} \Psi(\boldsymbol{\mu}; \boldsymbol{\xi}) dQ(\boldsymbol{\mu}), \tag{3.1}$$

where $\boldsymbol{\mu} \in \mathbf{M} \subset \mathbb{R}^K$, $\Psi(\boldsymbol{\mu}; \boldsymbol{\xi})$ is the design criterion, and Q is a given prior for $\boldsymbol{\mu}$ such that $\int_{\mathbf{M}} dQ(\boldsymbol{\mu}) = 1$. For more examples, in the spirit of (3.1), see [16], [25], [21], and [22]. As shown by [20] maximizing the power of the unconstrained LRT is equivalent to maximizing the NCP. However, it is not clear that the Bayes design with respect to the NCP will coincide with the Bayes design with respect to the power. Therefore we shall explore both.

Let $\boldsymbol{\xi}_Q^\lambda$ and $\boldsymbol{\xi}_Q^\pi$ denote the Bayes Q -optimal designs when $\Psi(\boldsymbol{\mu}; \boldsymbol{\xi})$ in (3.1) replaced by $\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})$ and $\pi(\boldsymbol{\mu}; \boldsymbol{\xi})$ respectively, i.e.,

$$\boldsymbol{\xi}_Q^\lambda = \arg \max_{\boldsymbol{\xi} \in \Xi} \Lambda(Q; \boldsymbol{\xi}) = \arg \max_{\boldsymbol{\xi} \in \Xi} \int \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) dQ(\boldsymbol{\mu}) \tag{3.2}$$

and

$$\boldsymbol{\xi}_Q^\pi = \arg \max_{\boldsymbol{\xi} \in \Xi} \Pi(Q; \boldsymbol{\xi}) = \arg \max_{\boldsymbol{\xi} \in \Xi} \int \pi(\boldsymbol{\mu}; \boldsymbol{\xi}) dQ(\boldsymbol{\mu}). \tag{3.3}$$

We start with the most natural prior.

Theorem 3.1. *Consider a prior Q of the form $\mathcal{N}(\mu_0 \mathbf{1}, \boldsymbol{\Sigma})$ where $\mu_0 \in \mathbb{R}$ and*

$$\boldsymbol{\Sigma} = (\sigma_{ij}) = \begin{cases} \beta, & \text{if } i = j \\ \gamma, & \text{otherwise.} \end{cases} \tag{3.4}$$

Then,

$$\boldsymbol{\xi}_Q^\lambda = \boldsymbol{\xi}_Q^\pi = \left(\frac{1}{K}, \dots, \frac{1}{K} \right).$$

Theorem 3.1 shows that the MM–design coincides with the NCP and power based Bayes designs when the prior Q is the distribution function of a multivariate normal exchangeable random vector. This family of priors, which when $\gamma = 0$, includes the independent components prior, is often used in applications. As shown in the proof of Theorem 3.1 the value of μ_0 is inconsequential so henceforth we set it equal to 0.

The assumption of normality can be relaxed:

Theorem 3.2. *If Q is any distribution of an exchangeable random vector with a finite second moment, then*

$$\xi_Q^\lambda = \xi_Q^\pi = \left(\frac{1}{K}, \dots, \frac{1}{K}\right).$$

Finally, a prior Q is called a *least favourable prior* (LFP), denoted Q_{LFP} , if

$$\max_{\xi \in \Xi} \Lambda(Q_{\text{LFP}}; \xi) = \min_{Q \in \mathcal{Q}} \max_{\xi \in \Xi} \Lambda(Q; \xi),$$

where \mathcal{Q} is the family of all possible priors.

Theorem 3.3. *Fix $\delta > 0$ and consider a prior Q which assigns equal mass to each permutation of the vector $(-\delta/2, \delta/2, 0, \dots, 0)$. Then we have:*

$$\xi_Q^\lambda = \xi_Q^\pi = \left(\frac{1}{K}, \dots, \frac{1}{K}\right),$$

and Q is the LFP on (2.7).

Thus the MM–design is also the NCP and power based Bayes design with respect to the LFP. Note that the LFP in Theorem 3.3 is the prior which assigns the same probability to each one of the least favourable configurations as discussed after the statement of Theorem 2.1.

As requested by a referee we briefly explore Bayes designs for two non–exchangeable priors; Example 3.1 deals with unequal variances and Example 3.2 with unequal means. In both cases we consider $K = 3$ treatment groups. We focus on NCP–based designs, i.e., ξ_Q^λ defined in (3.2), as these are much easier to calculate.

Example 3.1. *Let Q_1 be the prior $Q_1 \equiv \mathcal{N}(\mathbf{1}_3, \Sigma)$ for μ where $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)$. Using equation (A.2) in the proof of Theorem 3.1 we find that*

$$\Lambda(Q_1; \xi) = \sum_{i=1}^3 \xi_i(1 - \xi_i)\sigma_i^2.$$

Maximizing $\sum_{i=1}^3 \xi_i(1 - \xi_i)\sigma_i^2$ with respect to $\xi \in \Xi$ subject to the constraint $\sum_{i=1}^3 \xi_i = 1$ yields

$$\xi_{Q_1}^\lambda = ((\sigma_1^2\sigma_2^2 + \sigma_1^2\sigma_3^2)/\sigma^2, (\sigma_2^2\sigma_3^2 + \sigma_2^2\sigma_1^2)/\sigma^2, (\sigma_1^2\sigma_3^2 + \sigma_2^2\sigma_3^2)/\sigma^2)$$

where $\sigma^2 = 2(\sigma_1^2\sigma_2^2 + \sigma_2^2\sigma_3^2 + \sigma_3^2\sigma_1^2)$. In particular, when $\sigma_i^2 = \sigma_j^2$, i.e., Q_1 is exchangeable, then $\xi_{Q_1}^\lambda = (1/3, 1/3, 1/3)$ reduces to the balanced design. If however $\sigma_i^2 \neq \sigma_j^2$ for some $i \neq j$ then the resulting Bayes design will not be balanced. Designs for some specific values of $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ are reported in Table 1. It is clear from Table 1 that the optimal Bayes design assigns more observations to groups with higher variances and vice versa. For example, when $(\sigma_1^2, \sigma_2^2, \sigma_3^2) = (0.25, 0.50, 1.00)$ then $\xi_{Q_1}^\lambda = (0.214, 0.357, 0.429)$.

Example 3.2. Let Q_2 be the prior $Q_2 \equiv \mathcal{N}(\boldsymbol{\eta}, \mathbf{I}_3)$ for $\boldsymbol{\mu}$ where $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3)^T$ and \mathbf{I}_3 is an identity matrix. Using equation (A.2) in the proof of Theorem 3.1 we find that

$$\Lambda(Q_2; \boldsymbol{\xi}) = \sum_{i=1}^3 \xi_i \eta_i^2 - \left(\sum_{i=1}^3 \xi_i \eta_i \right)^2 + \sigma^2.$$

which we can maximize with respect to $\boldsymbol{\xi}$. Designs for some specific values of the mean parameter $\boldsymbol{\eta}$ are reported in Table 1. Note that the optimal Bayes design reduces to a balanced two-point design when the maximally separated pair is unique. When $\boldsymbol{\eta} = (-1, 1, 1)$ there are two maximally separated pairs (η_1, η_2) and (η_1, η_3) and the optimal Bayes design is not unique.

TABLE 1
Bayesian designs $\xi_{Q_1}^\lambda$ and $\xi_{Q_2}^\lambda$ based on non-exchangeable priors Q_1 and Q_2 .

$(\sigma_1^2, \sigma_2^2, \sigma_3^2)$	$\xi_{Q_1}^\lambda$	$\boldsymbol{\eta}^T$	$\xi_{Q_2}^\lambda$
$(0.25, 0.50, 1)$	$(0.214, 0.357, 0.429)$	$(-1, 0, 1)$	$(0.5, 0, 0.5)$
$(3, 2, 1)$	$(0.409, 0.364, 0.2273)$	$(-1, 1, 1)$	$(0.5, 0.25, 0.25)/(0.5, 0, 0.5)/(0.5, 0.5, 0)$
$(0.25, 0.75, 0.25)$	$(0.286, 0.429, 0.286)$	$(-0.25, 1, 0.5)$	$(0.5, 0.5, 0)$

4. Nash designs

As noted by [11] game theoretic ideas and, specifically, the notion of the *Nash Equilibrium* is widely and insightfully used in numerous disciplines. Applications of game-theory in statistics have a long, but far from voluminous, history. One exception is statistical decision theory which is strongly rooted in the theory of zero sum games, cf. [23] or [3] for a more modern treatment. An additional important reference is the classic book by [4] which discussed a wide array of statistical problems from a game theoretic perspective. We also mention the paper by [18] which touches on the relationship between game theory and Bayesian statistics. Unfortunately, a modern and vigorous applications of the ideas and tools developed in game theory to bear on current statistical problems is lacking. We believe that game theoretic ideas are well suited to address statistical design problems. This section provides the first steps in that direction. In particular, we explore the game theoretic formulation of the testing problem posed in (2.2). Later on we also consider the hypothesis testing problem (4.5) comparing a control with multiple treatments.

4.1. Standard ANOVA: The game theoretic view

The main result of this Section is Theorem 4.3 which is the game-theoretic equivalent of Theorem 2.1. The new formulation and proof are both mathematically simpler and more illuminating than the original ones, and moreover, may serve as a method for reasoning about complex design problems and as a blueprint for establishing future, even more demanding, results. For completeness, as well as the integrity of the exposition, we will first introduce the relevant terminology and notation with which we establish a sequence of results leading to Theorem 4.3.

We begin with some game theoretic terminology ([14]). Consider a two person game in which Player I is the statistician and Player II is nature. Player I chooses a design ξ while Player II chooses a value μ . In the parlance of game theory the choices available to the Players are called strategies. Initially we shall assume that both Player I and II have at their disposal only a finite set of distinct choices, called *pure strategies*, denoted by \mathfrak{X} and \mathfrak{M} , respectively. We start by setting $\mathfrak{X} = \{\xi_{ij} : 1 \leq i < j \leq K\}$ with $\xi_{ij} = (\mathbf{e}_i + \mathbf{e}_j)/2$ and $\mathfrak{M} = \{\mu_{ij} : 1 \leq i < j \leq K\}$, with $\mu_{ij} = (\delta/2)(\mathbf{e}_j - \mathbf{e}_i)$. The set \mathfrak{X} is the collection of all balanced two-point designs as discussed in Section 2 whereas the set \mathfrak{M} is the set of all permutations of the vector $(-\delta/2, \delta/2, 0, \dots, 0)$, where we restrict the first non-zero element to be negative and the second to be positive. Note that \mathfrak{M} is a subset of the LFCs on \mathcal{M}_δ , i.e., if $\mu \in \mathfrak{M}$ then $-\mu \notin \mathfrak{M}$. We need not consider all LFCs since $\lambda(\mu; \xi) = \lambda(-\mu; \xi)$. See Example 4.1 for further clarification.

Let \mathcal{I}_0 denote the set of pairs of indices $\{(i, j), (k, l)\}$ where there are no matches, i.e., $i \notin \{k, l\}$ and $j \notin \{k, l\}$, \mathcal{I}_1 is the set of pairs with one match and \mathcal{I}_2 is the set of pairs that are fully matched, i.e., $(i, j) = (k, l)$. When Player I chooses ξ_{kl} and Player II chooses μ_{ij} we say that (μ_{ij}, ξ_{kl}) is *played*. Suppose now that when (μ_{ij}, ξ_{kl}) is played the payoff to Player I is the value of the NCP, i.e.,

$$\lambda(\mu_{ij}; \xi_{kl}) = \begin{cases} \delta^2/4 & \text{if } \{(i, j), (k, l)\} \in \mathcal{I}_2 \\ \delta^2/16 & \text{if } \{(i, j), (k, l)\} \in \mathcal{I}_1 \\ 0 & \text{if } \{(i, j), (k, l)\} \in \mathcal{I}_0 \end{cases} \quad (4.1)$$

and to Player II it is $-\lambda(\mu_{ij}; \xi_{kl})$. Formally, we denote this game by $G_A(\lambda, \mathfrak{X}, \mathfrak{M})$, where the subscript A indicates that we are playing the ANOVA game, λ is the payoff function and $\mathfrak{X}, \mathfrak{M}$ describe the strategies available to Players I and II, respectively. Since the payoffs to Players I and II sum to zero the game is referred to as a *zero-sum* game. Similarly, $G_A(\pi, \mathfrak{X}, \mathfrak{M})$ is the corresponding game associated with the power function. Since the power function $\pi(\mu; \xi)$ depends on $(\mu; \xi)$ only through the NCP we can write $\pi(\mu; \xi) \equiv \pi(N, \lambda(\mu; \xi))$ and following (4.1) we have

$$\pi(N, \lambda(\mu_{ij}; \xi_{kl})) = \begin{cases} \pi(N, \delta^2/4) & \text{if } \{(i, j), (k, l)\} \in \mathcal{I}_2 \\ \pi(N, \delta^2/16) & \text{if } \{(i, j), (k, l)\} \in \mathcal{I}_1 \\ \pi(N, 0) & \text{if } \{(i, j), (k, l)\} \in \mathcal{I}_0. \end{cases}$$

Note that the games $G_A(\lambda, \mathfrak{X}, \mathfrak{M})$ and $G_A(\pi, \mathfrak{X}, \mathfrak{M})$ are motivated by our earlier observations on maximally separated means and their relationship with two-point designs.

By definition $\lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi}_{kl}) = 0$ whenever $\{(i, j), (k, l)\} \in \mathcal{I}_0$ so the payoff matrix will contain $\binom{K}{2} \binom{K-2}{2}$ zeros when $K > 3$. When $K \leq 3$, the payoff matrix contains no zeros. For example when $K = 3$ the payoff matrix, or game matrix, is given in Table 2.

TABLE 2
The NCP-based game matrix for ANOVA with $K = 3$.

I \ II	$\boldsymbol{\mu}_{12}$	$\boldsymbol{\mu}_{13}$	$\boldsymbol{\mu}_{23}$	max
$\boldsymbol{\xi}_{12}$	$\delta^2/4$	$\delta^2/16$	$\delta^2/16$	$\delta^2/4$
$\boldsymbol{\xi}_{13}$	$\delta^2/16$	$\delta^2/4$	$\delta^2/16$	$\delta^2/4$
$\boldsymbol{\xi}_{23}$	$\delta^2/16$	$\delta^2/16$	$\delta^2/4$	$\delta^2/4$
min	$\delta^2/16$	$\delta^2/16$	$\delta^2/16$	

The objective of Player I is to maximize his gain, or payoff. Similarly, Player II would like to minimize his loss. Observe that

$$\max_{\boldsymbol{\xi} \in \mathfrak{X}} \min_{\boldsymbol{\mu} \in \mathfrak{M}} \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) = \delta^2/16 \neq \delta^2/4 = \min_{\boldsymbol{\mu} \in \mathfrak{M}} \max_{\boldsymbol{\xi} \in \mathfrak{X}} \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}),$$

i.e., the max–min and min–max values do not agree. This means that the game $G_A(\lambda, \mathfrak{X}, \mathfrak{M})$ does not admit an *equilibrium* in pure strategies. Hence strategies $\boldsymbol{\mu}_{ij}^*$ and $\boldsymbol{\xi}_{ij}^*$ such that

$$\lambda(\boldsymbol{\mu}_{ij}^*, \boldsymbol{\xi}_{ij}^*) \geq \lambda(\boldsymbol{\mu}_{ij}, \boldsymbol{\xi}_{ij}^*) \quad \text{and} \quad \lambda(\boldsymbol{\mu}_{ij}^*, \boldsymbol{\xi}_{ij}^*) \geq \lambda(\boldsymbol{\mu}_{ij}^*, \boldsymbol{\xi}_{ij})$$

for all $\boldsymbol{\mu}_{ij} \in \mathfrak{M}$ and $\boldsymbol{\xi}_{ij} \in \mathfrak{X}$ do not exist; see the proof of Theorem 4.1 for more details. A comprehensive discussion on the existence of equilibria can be found in Chapters 4 and 5 of [14]. Suppose now that Player I may randomly select a pure strategy from \mathfrak{X} using a probability law $\mathbf{p} = (p_{12}, \dots, p_{K-1, K})$ where p_{ij} is the probability of selecting the pure strategy $\boldsymbol{\xi}_{ij}$. Similarly, suppose that Player II may independently select a pure strategy from \mathfrak{M} using a probability distribution \mathbf{q} . It is further assumed that Player I may choose $\mathbf{p} \in \mathbf{P}$ where $\mathbf{P} = \mathbf{P}(\mathfrak{X})$ are all discrete distributions supported on \mathfrak{X} and Player II may choose $\mathbf{q} \in \mathbf{Q}$ where $\mathbf{Q} = \mathbf{Q}(\mathfrak{M})$ are all discrete distributions supported on \mathfrak{M} . The expected payoff when Players I and II choose the, so called, *mixed strategies* (\mathbf{p}, \mathbf{q}) is

$$\Gamma(\mathbf{p}, \boldsymbol{\Lambda}, \mathbf{q}) = \mathbf{p}^T \boldsymbol{\Lambda} \mathbf{q} \tag{4.2}$$

where $\boldsymbol{\Lambda}$ is the game matrix. The game in mixed strategies will be denoted by $G_A(\lambda, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$. It is well known that finite zero–sum games have an equilibrium in mixed strategies. Thus, there are strategies \mathbf{p}_0 and \mathbf{q}_0 for Players I and II such that

$$\Gamma(\mathbf{p}_0, \boldsymbol{\Lambda}, \mathbf{q}_0) = \min_{\mathbf{q} \in \mathbf{Q}} \max_{\mathbf{p} \in \mathbf{P}} \Gamma(\mathbf{p}, \boldsymbol{\Lambda}, \mathbf{q}) = \max_{\mathbf{p} \in \mathbf{P}} \min_{\mathbf{q} \in \mathbf{Q}} \Gamma(\mathbf{p}, \boldsymbol{\Lambda}, \mathbf{q}),$$

i.e., the min–max and max–min adversarial optimization problems result in the same outcome. The strategies \mathbf{p}_0 and \mathbf{q}_0 are said to be *optimal mixed strategies*

and

$$v(\lambda, \mathfrak{X}, \mathfrak{M}) = \Gamma(\mathbf{p}_0, \mathbf{\Lambda}, \mathbf{q}_0) \quad (4.3)$$

is called the *value* of the game. Further note that an optimal mixed strategy for Player I is a probability distribution \mathbf{p} which guarantees (i.e., maximizes) the minimal gain; whereas an optimal mixed strategy for Player II is a probability distribution \mathbf{q} which guarantees (i.e., minimizes) the maximal loss. The strategies \mathbf{p}_0 and \mathbf{q}_0 are said to be a Nash equilibrium if deviating from them is detrimental, i.e., if $\Gamma(\mathbf{p}_0, \mathbf{\Lambda}, \mathbf{q}_0) \leq \Gamma(\mathbf{p}_0, \mathbf{\Lambda}, \mathbf{q})$ and $\Gamma(\mathbf{p}_0, \mathbf{\Lambda}, \mathbf{q}_0) \geq \Gamma(\mathbf{p}, \mathbf{\Lambda}, \mathbf{q}_0)$ for all $\mathbf{p} \in \mathbf{P}$ and $\mathbf{q} \in \mathbf{Q}$. The max–min and Nash equilibrium coincide in zero sum games. The value of the game $G_A(\lambda, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$ and its optimal mixed strategies are similarly defined by simply replacing λ by π in (4.3). Now,

Theorem 4.1. *For any $K \in \mathbb{N}$ the pair*

$$\mathbf{p} = \mathbf{q} = \mathbf{1}_{K(K-1)/2} \binom{K}{2}^{-1}$$

is the unique Nash Equilibrium of the games $G_A(\lambda, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$ and $G_A(\pi, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$.

Examples 4.1–4.3 show that adding any $\boldsymbol{\mu} \in \mathcal{M}_\delta \setminus \mathfrak{M}$ to the strategies of Player II, i.e., to the set \mathfrak{M} , will result in a game that *reduces* to $G_A(\lambda, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}'))$, by which we mean that

$$v(\lambda, \mathfrak{X}, \mathfrak{M}) = v(\lambda, \mathfrak{X}, \mathfrak{M}')$$

whenever $\mathfrak{M} \subset \mathfrak{M}' \subset \mathcal{M}_\delta$. For simplicity we focus on the case where $K = 3$ and denote the resulting game by G_3 .

Example 4.1. *Let $\mathfrak{M}' = \{(-\delta/2, \delta/2, 0), (-\delta/2, 0, \delta/2), (0, -\delta/2, \delta/2), (0, \delta/2, -\delta/2)\}$. The game matrix is displayed in Table 3 below where $\boldsymbol{\mu}'_{23} = (0, \delta/2, -\delta/2)$. It is easy to see that the strategies $\boldsymbol{\mu}_{23}$ and $\boldsymbol{\mu}'_{23}$ satisfy $\boldsymbol{\mu}'_{23} = -\boldsymbol{\mu}_{23}$ and are associated with the same payoffs. Thus they are equivalent from a game–theoretic perspective. It is not hard to show that the pair $\mathbf{p} = (1/3, 1/3, 1/3)$,*

TABLE 3
Game matrix of Example 4.1.

$\Gamma \backslash \Pi$	$\boldsymbol{\mu}_{12}$	$\boldsymbol{\mu}_{13}$	$\boldsymbol{\mu}_{23}$	$\boldsymbol{\mu}'_{23}$
$\boldsymbol{\xi}_{12}$	$\delta^2/4$	$\delta^2/16$	$\delta^2/16$	$\delta^2/16$
$\boldsymbol{\xi}_{13}$	$\delta^2/16$	$\delta^2/4$	$\delta^2/16$	$\delta^2/16$
$\boldsymbol{\xi}_{23}$	$\delta^2/16$	$\delta^2/16$	$\delta^2/4$	$\delta^2/4$

$\mathbf{q} = (1/3, 1/3, \gamma/3, (1-\gamma)/3)$, for any $\gamma \in [0, 1]$ is a Nash Equilibrium. Thus the Nash Equilibrium is not unique; however, the value of the game is independent of γ and equal to the value of G_3 . So, without any loss, we may choose $\gamma = 1$ essentially eliminating the strategy $(0, \delta/2, -\delta/2)$ from the game. This shows that the game associated with \mathfrak{M}' reduces and is equivalent to G_3 .

Example 4.2. Let $\mathfrak{M}' = \{(-\delta/2, \delta/2, 0), (-\delta/2, 0, \delta/2), (0, -\delta/2, \delta/2), (-\delta/2, \delta/2, \delta/2)\}$. The corresponding game matrix, $\mathbf{\Lambda}$, is displayed in Table 4 below where $\boldsymbol{\mu}_{123} = (-\delta/2, \delta/2, \delta/2)$.

TABLE 4
Game matrix of Example 4.2.

I\II	$\boldsymbol{\mu}_{12}$	$\boldsymbol{\mu}_{13}$	$\boldsymbol{\mu}_{23}$	$\boldsymbol{\mu}_{123}$
$\boldsymbol{\xi}_{12}$	$\delta^2/4$	$\delta^2/16$	$\delta^2/16$	$\delta^2/4$
$\boldsymbol{\xi}_{13}$	$\delta^2/16$	$\delta^2/4$	$\delta^2/16$	$\delta^2/4$
$\boldsymbol{\xi}_{23}$	$\delta^2/16$	$\delta^2/16$	$\delta^2/4$	0

Suppose that Player I chooses the strategy $\mathbf{p} = (1/3, 1/3, 1/3)$. The best response of Player II to this strategy is a mixed strategy $\mathbf{q} = (q_{12}, q_{13}, q_{23}, q_{123})$ which minimizes (4.2) which is nothing but $(1/3)\mathbf{1}^T \mathbf{\Lambda} \mathbf{q}$. A bit of algebra shows that in this case

$$\begin{aligned} \Gamma(\mathbf{p}, \mathbf{\Lambda}, \mathbf{q}) &= \frac{\delta^2}{8}(q_{12} + q_{13} + q_{23}) + \frac{2\delta^2 q_{123}}{12} \\ &= \frac{\delta^2}{8}(1 - q_{123}) + \frac{2\delta^2 q_{123}}{12} = \frac{\delta^2}{8} + \frac{\delta^2 q_{123}}{24}. \end{aligned} \tag{4.4}$$

The right hand side of (4.4) is minimized when $q_{123} = 0$. Thus, the strategy $\boldsymbol{\mu}_{123}$ can be eliminated and the game reduces to G_3 .

Example 4.3. Let $\mathfrak{M}' = \{(-\delta/2, \delta/2, 0), (-\delta/2, 0, \delta/2), (0, -\delta/2, \delta/2), (-\delta, 0, \delta)\}$. The corresponding game matrix, $\mathbf{\Lambda}$, is displayed in Table 5 below where $\boldsymbol{\mu}_4 = (-\delta, 0, \delta)$.

TABLE 5
Game matrix of Example 4.3.

I\II	$\boldsymbol{\mu}_{12}$	$\boldsymbol{\mu}_{13}$	$\boldsymbol{\mu}_{23}$	$\boldsymbol{\mu}_4$
$\boldsymbol{\xi}_{12}$	$\delta^2/4$	$\delta^2/16$	$\delta^2/16$	$\delta^2/4$
$\boldsymbol{\xi}_{13}$	$\delta^2/16$	$\delta^2/4$	$\delta^2/16$	δ^2
$\boldsymbol{\xi}_{23}$	$\delta^2/16$	$\delta^2/16$	$\delta^2/4$	$\delta^2/4$

It is not hard to see that $\lambda(\boldsymbol{\mu}_4, \boldsymbol{\xi}_{ij}) \geq \lambda(\boldsymbol{\mu}_{ij}, \boldsymbol{\xi}_{ij})$ where $i \neq j \in \{1, 2, 3\}$ so $\boldsymbol{\mu}_4$ is an inferior, or dominated, strategy for Player II. It follows (cf. Theorem 5.20 in [14]) that it will be assigned zero probability, i.e., the game reduces to G_3 .

Examples 4.1 and 4.2 show that adding any strategy to \mathfrak{M} which satisfies $\max |\nu_i - \nu_j| = \delta$, i.e., $\boldsymbol{\nu}$ lies on the boundary of the set \mathcal{M}_δ , will result in a game that is equivalent to G_3 . Example 4.3 shows that strategies in the interior of \mathcal{M}_δ are dominated and therefore will not be chosen by Player II.

Example 4.4. Let $\mathfrak{M}' = \{(-\delta/2, \delta/2, 0), (-\delta/2, 0, \delta/2)\}$. The corresponding game matrix is displayed in Table 6. It is easily observed that $\boldsymbol{\xi}_{23}$ is dominated by a balanced mixture of the strategies $\boldsymbol{\xi}_{12}$ and $\boldsymbol{\xi}_{13}$. Thus, by Theorem 5.20 in [14], $\boldsymbol{\xi}_{23}$ can be eliminated. The Nash Equilibrium of the resulting game (whose game matrix consists of the first two rows of Table 6) is

TABLE 6
Game matrix of Example 4.4.

I \ II	μ_{12}	μ_{13}
ξ_{12}	$\delta^2/4$	$\delta^2/16$
ξ_{13}	$\delta^2/16$	$\delta^2/4$
ξ_{23}	$\delta^2/16$	$\delta^2/16$

$(\mathbf{p}, \mathbf{q}) = ((1/2, 1/2), (1/2, 1/2))$ and the payoff of the game is $5\delta^2/32$ which is larger than the value of the Game G_3 ($\delta^2/8$).

Example 4.4 shows that removing any strategy from \mathfrak{M} results in a game which has higher value than the original game. To summarize, the preceding examples show that:

Theorem 4.2. *If $\mathfrak{M}_l \subset \mathfrak{M}$, then $v(\lambda, \mathfrak{X}, \mathfrak{M}) < v(\lambda, \mathfrak{X}, \mathfrak{M}_l)$ and if $\mathfrak{M}_b \supset \mathfrak{M}$ where $\mathfrak{M}_b \subset \mathcal{M}_\delta$ then $v(\lambda, \mathfrak{X}, \mathfrak{M}_b) = v(\lambda, \mathfrak{X}, \mathfrak{M})$.*

Theorem 4.2 shows that the set of strategies \mathfrak{M} is complete with respect to \mathcal{M}_δ when Player I's strategies are \mathfrak{X} . By this we mean that Player II can not find any strategies in \mathcal{M}_δ which will reduce the value of the game and, in addition, the omission of any strategy from \mathfrak{M} will increase the value of the game.

Now, after fixing the strategies of Player II we examine the strategies of Player I. First note that in expectation, or under repeated play, Player I allocates $1/K$ of the experimental subject to each treatment group. However, in any specific game Player I allocates $1/2$ of the observations to group i and $1/2$ to groups j for some pair (i, j) . Thus, the payoff is 0 with probability $\binom{K}{2}^{-1} \binom{K-2}{2}$, $\delta^2/16$ with probability $\binom{K}{2}^{-1} \{ \binom{K}{2} - \binom{K-2}{2} - 1 \}$, and equal to $\delta^2/4$ with probability $\binom{K}{2}^{-1}$. This set of possible outcomes is not satisfying and it suggest that the strategy space for Player I should be expanded.

Therefore, we next consider the situation where Player I can choose any $\xi \in \Xi$. The strategy space for Player I is no longer finite and games of this type are called *infinite games*. As noted by a referee the action space of Player I, the set Ξ , is convex and therefore Player I need not consider any mixed strategies. We shall denote such games by $G_A(\cdot, \Xi, \mathcal{Q}(\mathfrak{M}))$. We have:

Theorem 4.3. *The pair $\xi = K^{-1} \mathbf{1}_K$ and $\mathbf{q} = \binom{K}{2}^{-1} \mathbf{1}_{K(K-1)/2}$ is the unique Nash Equilibrium of the games $G_A(\lambda, \Xi, \mathcal{Q}(\mathfrak{M}))$ and $G_A(\pi, \Xi, \mathcal{Q}(\mathfrak{M}))$.*

Theorem 4.3 is the game theoretic equivalent of Theorem 2.1. It provides a new perspective on a classical design problem. In light of the Examples 4.1–4.3, it is also clear that Theorem 4.3 applies to the infinite games $G_A(\lambda, \Xi, \mathcal{M}_\delta)$ and $G_A(\pi, \Xi, \mathcal{M}_\delta)$ where the optimal strategy for Player II is to randomly select strategies from \mathfrak{M} using the probability mass function \mathbf{q} .

4.2. Comparison of multiple treatments with a control: The game theoretic view

There are many experiments in which an inherent ordering among the experimental groups exists. For example, in dose–response studies a large response is often expected with a high dose. This ordering is referred to as the simple order. When a control is compared to multiple treatments, each of which is expected to outperform it, the resulting order is known as a tree order. There are many other orders and a rich literature describing efficient statistical analysis under order restriction. For more details see the books by [2], and [19].

In this section, we find the Nash–Design when testing under a tree order. Formally, we would like to compare a control group to $K - 1$ treatment groups. The hypothesis of interest in this case is:

$$H_0 : \boldsymbol{\mu} \in \mathcal{M}_0 \quad \text{versus} \quad H_1 : \boldsymbol{\mu} \in \mathcal{M}_1 \setminus \mathcal{M}_0, \tag{4.5}$$

where \mathcal{M}_0 was defined immediately after (2.2) and $\mathcal{M}_1 = \{\boldsymbol{\mu} \in \mathbb{R}^K : \mu_1 \leq \mu_i, \forall i \in \{2, \dots, K\}\}$. In this context it is natural to define

$$\mathcal{M}_\delta = \{\boldsymbol{\mu} \in \mathcal{M}_1 : \sum_{i=2}^K (\mu_i - \mu_1) \geq \delta\}. \tag{4.6}$$

Note that this \mathcal{M}_δ is different than considered in (2.7). As earlier δ can be viewed, in the sense described in Remark 2.1, as a distance from the null. Since this testing problem is location invariant we may, without any loss of any generality, assume that $\mu_1 = 0$. One obvious choice for the set of strategies of Player II is the set $\mathfrak{M} = \{\boldsymbol{\mu}_i : \mu_i = \delta \mathbf{e}_i; i = 2, \dots, K\}$. Note that elements of \mathfrak{M} are on the boundary of \mathcal{M}_δ and lie on the extreme rays of the \mathcal{M}_1 . A natural set of strategies for Player I are $\mathfrak{X} = \{\boldsymbol{\xi}_{1i} : \xi_{1i} = 1/2(\mathbf{e}_1 - \mathbf{e}_i); i = 2, \dots, K\}$, i.e., the two point strategies comparing the control to treatment i . The payoff to Player I is assumed to be the NCP, λ , or the power, π , of the test (2.3). As noted by a referee the test (2.3) is not the classic constrained test for testing for the tree order [19]. However, as shown by [20] designs for constrained testing problems and unconstrained testing problems coincide asymptotically. Therefore in this manuscript we shall consider only payoffs derived from the unconstrained test. Let \mathbf{p} and \mathbf{q} be the probability laws by which Player I and II of choose their strategies from \mathfrak{X} and \mathfrak{M} respectively. It can be shown, by repeating the arguments made in the proof of Theorem 4.1, that when $K \geq 2$ the Nash Equilibrium for the games $G_T(\lambda, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$ and $G_T(\pi, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$, where the subscript T on $G_T(\cdot, \cdot, \cdot)$ indicates that these games are associated with the tree order, is

$$\mathbf{p} = \mathbf{q} = (K - 1)^{-1} \mathbf{1}_{K-1}.$$

The value of the game $G_T(\lambda, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$ is $\delta^2/4(K - 1)$.

We now extend the strategy space of Player II by adding the strategy $\boldsymbol{\mu}_{\text{central}} = (0, \delta/(K - 1), \dots, \delta/(K - 1))$ to \mathfrak{M} . Note that $\boldsymbol{\mu}_{\text{central}}$ is the equal weights convex

combination of the strategies \mathfrak{X} . Further assume that Player I chooses the strategies from \mathfrak{X} using the probability law $\mathbf{p} = (K-1)^{-1}\mathbf{1}$ whereas Player II chooses from the strategies $(\boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_{\text{central}})$ with probability law $\mathbf{q} = (q_2, \dots, q_K, q_c)$. The game matrix for this augmented game is presented in Table 7.

TABLE 7
Game matrix for the augmented tree order.

I \ II	$\boldsymbol{\mu}_2$	$\boldsymbol{\mu}_3$	\dots	$\boldsymbol{\mu}_K$	$\boldsymbol{\mu}_{\text{central}}$
$\boldsymbol{\xi}_{12}$	$\delta^2/4$	0	\dots	0	$\delta^2/4(K-1)^2$
$\boldsymbol{\xi}_{13}$	0	$\delta^2/4$	\dots	0	$\delta^2/4(K-1)^2$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
$\boldsymbol{\xi}_{1K}$	0	0	\dots	$\delta^2/4$	$\delta^2/4(K-1)^2$

The payoff of this game is

$$\Gamma(\mathbf{p}, \boldsymbol{\Lambda}, \mathbf{q}) = (K-1)^{-1}\mathbf{1}^T \boldsymbol{\Lambda} \mathbf{q} = \frac{\delta^2}{4(K-1)} \left\{ 1 + \frac{(2-K)q_c}{K-1} \right\}, \quad (4.7)$$

which is minimized at $\delta^2/4(K-1)^2$ when $q_c = 1$. In fact, it is not difficult to show that $\boldsymbol{\mu}_{\text{central}}$ minimizes the loss of Player II over all $\boldsymbol{\mu} \in \mathcal{M}_\delta$ when Player I chooses among the strategies in Table 7 with equal probability. Thus, $\mathbf{p} = (K-1)^{-1}\mathbf{1}$ and $q_c = 1$ is a Nash Equilibrium. In fact, it can be verified that any convex combination of Player I's strategies, including all pure strategies result in a Nash Equilibrium.

However, most of these solutions are not satisfactory when considering repeated experiments. To illustrate, consider a design $\boldsymbol{\xi}_1$ obtained using some convex combination of $\{\boldsymbol{\xi}_{12}, \dots, \boldsymbol{\xi}_{1K}\}$ satisfying $\xi_{1i} > \xi_{1j}$ for some $i, j \neq 1$. Let $\boldsymbol{\xi}_2$ be the strategy obtained from $\boldsymbol{\xi}_1$ by switching its i^{th} and j^{th} components, i.e., $\xi_{2i} = \xi_{1j}$, $\xi_{2j} = \xi_{1i}$ and $\xi_{2k} = \xi_{1k}$ for all $k \notin \{i, j\}$. Further set $\boldsymbol{\xi} = (\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2)/2$. Let $\boldsymbol{\mu}_i = \delta \mathbf{e}_i$ and $\boldsymbol{\mu}_j = \delta \mathbf{e}_j$. Observe that $\lambda(\boldsymbol{\mu}_i; \boldsymbol{\xi}_1) > \lambda(\boldsymbol{\mu}_i; \boldsymbol{\xi}) > \lambda(\boldsymbol{\mu}_i; \boldsymbol{\xi}_2)$ and $\lambda(\boldsymbol{\mu}_j; \boldsymbol{\xi}_2) > \lambda(\boldsymbol{\mu}_j; \boldsymbol{\xi}) > \lambda(\boldsymbol{\mu}_j; \boldsymbol{\xi}_1)$ from which we conclude that

$$\min_{\boldsymbol{\mu} \in \{\boldsymbol{\mu}_i, \boldsymbol{\mu}_j\}} \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) > \min_{\boldsymbol{\mu} \in \{\boldsymbol{\mu}_i, \boldsymbol{\mu}_j\}} \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}_k) \quad \text{for } k = 1, 2.$$

Hence $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ can not be Nash designs for this game in which Player II chooses $\{\boldsymbol{\mu}_i, \boldsymbol{\mu}_j\}$. Since the exercise above can be carried out for any two indices it follows that if $\boldsymbol{\xi}$ is Nash then $\xi_i = \xi_j$ for all pairs i and j . Consequently the Nash design must be of the form $\boldsymbol{\xi}_\gamma = (\gamma, (1-\gamma)/(K-1), \dots, (1-\gamma)/(K-1))$ for some $\gamma \in [0, 1]$. Clearly, to find such a design, we should extend the strategy space of Player I. It is not hard to show that $\boldsymbol{\xi}_{1/2} = (1/2, 1/2(K-1), \dots, 1/2(K-1))$ is the only design which satisfies the above conditions and attains the mini-max value of the game. It is also interesting to note that, $\boldsymbol{\xi}_{1/2}$ is also the design obtained by averaging the designs $\boldsymbol{\xi}_{12}, \dots, \boldsymbol{\xi}_{1K}$. Using arguments similar to those above, we get the same Nash Equilibrium for this game when $\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})$ is replaced by $\pi(\boldsymbol{\mu}; \boldsymbol{\xi})$. We summarize the results in the following theorem:

Theorem 4.4. *The pair $\boldsymbol{\xi} = (1/2, 1/2(K-1), \dots, 1/2(K-1))$ and $\boldsymbol{\mu} = (0, \delta/(K-1), \dots, \delta/(K-1))$ is the unique Nash Equilibrium of the games $G_T(\lambda, \boldsymbol{\Xi}, \mathcal{M}_\delta)$ and $G_T(\pi, \boldsymbol{\Xi}, \mathcal{M}_\delta)$.*

5. Summary and discussion

In this article we considered optimal designs for two simple, but widely used, hypotheses testing problems, i.e., (2.2) and (4.5). Initially MM–designs for the one–way ANOVA problem were discussed. Then the corresponding Bayes and Nash designs were introduced and explored. A pseudo–Bayesian approach, in which an integrated version of the design criteria was optimized, was proposed in Section 3. It was shown that the Bayes design coincides with the maxi–min design for exchangeable priors with a finite second moment, which is a large class of reasonable priors. We also show that if the prior is not exchangeable then a balanced design is not obtained. The Bayesian approach can be extended to deal with many other testing problems. For example, consider the experiment in which multiple treatments are compared to a control, as discussed in Section 4.2. As noted this type of comparison is referred to as the simple tree order and the associated hypothesis testing problem is defined in (4.5). In this setting, it is reasonable to consider a prior Q for $\boldsymbol{\mu}$ which is equally supported on $\{\mathbf{e}_j - \delta \mathbf{e}_1\}$ for $j = 2, \dots, K$. A simple calculation shows that

$$\Lambda(Q; \boldsymbol{\xi}) = \xi_1 \delta^2 + \frac{1}{K-1} \left(\sum_{j=2}^K \xi_j - \sum_{j=2}^K (\xi_j - \delta \xi_i)^2 \right). \tag{5.1}$$

Maximizing (5.1) with respect to $\boldsymbol{\xi}$ leads to

$$\xi_1 = \frac{(K-1)^2 \delta^2 + 2\delta(K-1) + 3 - K}{2(1 + \delta(K-1))^2}, \quad \text{and} \quad \xi_i = \xi_j \quad \text{for all } i, j \geq 2.$$

Further observe that $\xi_1 \rightarrow 1/2$ as $\delta \rightarrow \infty$. Thus, there is a sequence of priors, and Bayes optimal designs which converge to the MM–design for the tree order. Finding such sequences of priors for more complicated testing problems, is of yet, an open problem. Within the Bayesian framework it is also interesting to consider the design criteria

$$\bar{\boldsymbol{\xi}}_Q^\lambda = \int \arg \max_{\boldsymbol{\xi} \in \Xi} \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) dQ(\boldsymbol{\mu})$$

and

$$\bar{\boldsymbol{\xi}}_Q^\pi = \int \arg \max_{\boldsymbol{\xi} \in \Xi} \pi(\boldsymbol{\mu}; \boldsymbol{\xi}) dQ(\boldsymbol{\mu})$$

which are, nothing but, Q –weighted locally optimal designs. For obvious reasons we shall refer to such designs as *weighted designs*. It is not hard to see that if $q_{ij} = \mathbb{P}_Q(\mathcal{E}_{ij})$, where \mathcal{E}_{ij} is the event that the means μ_i and μ_j are maximally separated, then any exchangeable prior would lead in the ANOVA setting to:

$$\begin{aligned} \bar{\boldsymbol{\xi}}_Q^\lambda &= \int \left\{ \arg \max_{\boldsymbol{\xi} \in \Xi} \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) \right\} dQ(\boldsymbol{\mu}) = \sum_{i < j} \int_{\mathcal{E}_{ij}} \left\{ \arg \max_{\boldsymbol{\xi} \in \Xi} \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) \right\} dQ(\boldsymbol{\mu}) \\ &= \sum_{i < j} \int_{\mathcal{E}_{ij}} \boldsymbol{\xi}_{ij} dQ(\boldsymbol{\mu}) = \binom{K}{2}^{-1} \sum_{i < j} \boldsymbol{\xi}_{ij} = \left(\frac{1}{K}, \dots, \frac{1}{K} \right), \end{aligned}$$

with a similar result for $\bar{\xi}_Q^\lambda$. Thus the weighted design coincides with the Bayes design under exchangeable priors. A similar equivalence can be shown to hold also under the tree order and even more general settings. Finally, as noted by a referee under a large class of priors Bayes factors in ANOVA settings are either exactly or asymptotically functions of the F -test statistic ([8] and [5]). Therefore, large values of the F -test are associated with large probabilities for the alternative. Hence, maximising the power of frequentist F -test maximizes the power of these Bayesian tests.

In Section 4 the experimental design problem was formulated, and solved, as a game theoretic problem in which the Nash Equilibrium is shown to coincide with the MM-design as well as the Bayes and Weighted designs. Apart from the simplicity and elegance of the solution the proposed approach suggests that game theoretic ideas and methods should play a more prominent role in statistics, especially within the broad field of experimental design. We are not familiar with any other papers in the statistical literature which apply game theoretic ideas to experimental design problems. There are however some papers in the engineering literature cf. [9] and [7] dealing with design allocation problems which are addressed by finding the Nash equilibrium of a “design game”. The game theoretic view provides a set of tools and method for reasoning in complex situations, which, we believe, will prove useful in solving variety of other problems in statistics.

It is well know that the Nash equilibrium coincides with the maxi–min solution in zero sum games. In some situations it may be difficult to compute the maxi–min solution but relatively easy, using symmetry say, to show that a particular set of strategies is a Nash equilibrium. Moreover it is possible to imagine design problems which are not maxi–min. For example, these could occur if sampling cost vary among units receiving different treatments. In such situations Nash designs are relevant whereas MM–designs are not.

Although the focus of this paper has been on the classical one–way ANOVA we believe that the proposed approach is applicable to two–way designs, as well as problems involving covariates, heteroscedastic errors and so forth. In each case the relevant payoff function needs to be defined and then a game pitting the statistician versus nature set–up, see [4] for a collection of problems viewed through this prism. We have already begun exploring various problems using this approach and believe that it can be extended to other areas of statistics. We realize that viewing nature as a strategizing agent may seem odd at first. However, it is common practice in a various disciplines including statistics ([4]) and in particular statistical decision theory ([3]). One benefit of the game theoretic approach is that it transforms a complicated optimization problems into a game, in which reasoning about outcomes is often easier and more natural.

Appendix

We begin with a remark which will be useful.

Remark 1. In the proofs of Theorem 3.1 and Theorem 3.3 we make use of an elegant result of [24] known as *The Purkiss Principle*. The principle, adapted to our setting, states that if f is a function which is symmetric in its arguments x_1, \dots, x_n then $\max\{f(x_1, \dots, x_n) : \sum_{i=1}^n x_i = 1\}$ is obtained when $x_i = x_j$ for all i and j provided that the Hessian of f does not vanish at the optimum.

Proof of Theorem 3.1

Following lemma is required to prove Theorem 3.1.

Lemma 1. *Let $\mathbf{A} = \mathbf{M}(\boldsymbol{\xi}) - \mathbf{M}(\boldsymbol{\xi})\mathbf{J}\mathbf{M}(\boldsymbol{\xi})$, where $\mathbf{M}(\boldsymbol{\xi}) = \text{diag}(\xi_1, \dots, \xi_K)$, $\mathbf{J} = \mathbf{1}_K\mathbf{1}_K^T$ and $\mathbf{1}_K$ is K dimensional vector of one's. Then $\text{trace}(\mathbf{A}^m)$ is permutation symmetric in (ξ_1, \dots, ξ_K) for any non-negative integer m .*

Proof. Let \mathbf{X} follow a multinomial distribution with parameters K and $\boldsymbol{\xi} = (\xi_1, \dots, \xi_K)^T$. It is clear that \mathbf{A} is the variance matrix of \mathbf{X} . Let \mathbf{S} be any permutation matrix and define $\mathbf{Y} = \mathbf{S}\mathbf{X}$. Observe that \mathbf{Y} follows a multinomial distribution with parameters K and $\mathbf{p} = \mathbf{S}\boldsymbol{\xi}$. The variance matrix of \mathbf{Y} is $\mathbf{B} = \mathbf{S}\mathbf{A}\mathbf{S}^T$. Now for any non-negative integer m and using the fact that $\mathbf{S}^T\mathbf{S} = \mathbf{I}$, where \mathbf{I} is the identity matrix, we have

$$\begin{aligned} \text{trace}(\mathbf{B}^m) &= \text{trace}((\mathbf{S}\mathbf{A}\mathbf{S}^T)^m) = \text{trace}((\mathbf{S}\mathbf{A}\mathbf{S}^T) \cdots (\mathbf{S}\mathbf{A}\mathbf{S}^T)) \\ &= \text{trace}(\mathbf{S}\mathbf{A}^m\mathbf{S}^T) = \text{trace}(\mathbf{S}^T\mathbf{S}\mathbf{A}^m) = \text{trace}(\mathbf{A}^m). \end{aligned}$$

Since the latter holds for all permutation matrices \mathbf{S} the function $\text{trace}(\mathbf{A}^m)$ is permutation symmetric in (ξ_1, \dots, ξ_K) as required. \square

We now continue with the proof of Theorem 3.1.

Proof. Observe that the NCP given in (2.4) can be written in matrix form as

$$\lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) = \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}, \tag{A.1}$$

where \mathbf{A} is defined in the statement of Lemma 1. It follows from (A.1) and the definition of $\Lambda(Q; \boldsymbol{\xi})$ that

$$\Lambda(Q; \boldsymbol{\xi}) = \mathbb{E}_Q(\boldsymbol{\mu})^T \mathbf{A} \mathbb{E}_Q(\boldsymbol{\mu}) + \text{trace}(\mathbb{V}_Q(\boldsymbol{\mu})\mathbf{A}), \tag{A.2}$$

where $\mathbb{E}_Q(\boldsymbol{\mu})$ and $\mathbb{V}_Q(\boldsymbol{\mu})$ are the mean and variance of the random vector $\boldsymbol{\mu}$ with respect to the distribution function Q . By assumption $\boldsymbol{\mu}$ is distributed as $\mathcal{N}(\boldsymbol{\mu}_0\mathbf{1}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ has a compound symmetry structure as described in the statement of the Theorem. It follows from (3.4) that (A.2) reduces to

$$\Lambda(Q; \boldsymbol{\xi}) = \beta \sum_i \xi_i(1 - \xi_i) - 2\gamma \sum_{i < j} \xi_i \xi_j. \tag{A.3}$$

A simple calculation, using the method of Lagrange multipliers, shows that the maximizer of (A.3) under the constraint $\sum_{i=1}^K \xi_i = 1$ is

$$\boldsymbol{\xi}_Q^\lambda = (1/K, \dots, 1/K),$$

as stated.

The average power with respect to the prior Q can be expressed as

$$\begin{aligned} \Pi(Q; \boldsymbol{\xi}) &= \int \mathbb{P}(T_n > c_{K-1, \alpha}) dQ(\boldsymbol{\mu}) = \int \mathbb{P}(\chi_{K-1}^2(\lambda) > c_{K-1, \alpha}) dQ(\boldsymbol{\mu}) \\ &= \int \sum_{k=0}^{\infty} \frac{e^{-\lambda/2} (\lambda/2)^k}{k!} \mathbb{P}(\chi_{K-1+2k}^2 > c_{K-1, \alpha}) dQ(\boldsymbol{\mu}) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(\chi_{K-1+2k}^2 > c_{K-1, \alpha}) \int \frac{e^{-\lambda/2} (\lambda/2)^k}{k!} dQ(\boldsymbol{\mu}) \quad (\text{A.4}) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(\chi_{K-1+2k}^2 > c_{K-1, \alpha}) \int I_k(\lambda) dQ(\boldsymbol{\mu}), \end{aligned}$$

where

$$I_k(\lambda) = \frac{e^{-\lambda/2} (\lambda/2)^k}{k!}.$$

Under the assumption of normality

$$\begin{aligned} \int I_k(\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})) dQ(\boldsymbol{\mu}) &= \int \frac{1}{k!} \sum_{i=0}^{\infty} (-1)^i (\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})/2)^{k+i} dQ(\boldsymbol{\mu}) \\ &= \frac{1}{k!} \sum_{i=0}^{\infty} (-1)^i (1/2)^{k+i} \mathbb{E}(\boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu})^{k+i}, \end{aligned}$$

where $\mathbf{A} = \mathbf{M}(\boldsymbol{\xi}) - \mathbf{M}(\boldsymbol{\xi})\mathbf{J}\mathbf{M}(\boldsymbol{\xi})$. Following [13] the moment $\mathbb{E}(\boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu})^k$ are of the form

$$\sum_{i=1}^k \nu_i \prod_{j=1}^k (\text{trace}(\mathbf{A}\boldsymbol{\Sigma})^{s_j})^{r_j},$$

where ν_i are integers and the product is over all integers s_j and r_j such that $\sum_j s_j r_j = k$. It is easy to verify that if $\boldsymbol{\Sigma}$ is defined as in the statement of the Theorem then $\mathbf{A}\boldsymbol{\Sigma} = (\beta - \gamma)\mathbf{A}$. Furthermore by Lemma 1, the quantity $\text{trace}(\mathbf{A}^s)$ is symmetric in (ξ_1, \dots, ξ_K) for any positive integer s . Consequently so are (5), and $\Pi(Q; \boldsymbol{\xi})$. Thus, by Remark 1, $\max\{\Pi(Q; \boldsymbol{\xi}) : \sum_i \xi_i = 1\}$ is attained when $\xi_i = \xi_j$ for all i, j which gives a balanced design. This concludes the proof. \square

Proof of Theorem 3.2

Proof. It is obvious that Equation (A.3) holds for all distribution functions Q which are exchangeable with respect to the random variable (μ_1, \dots, μ_K) from which the conclusion, $\boldsymbol{\xi}_Q^\lambda = (1/K, \dots, 1/K)$, follows immediately. Now, observe that

$$\Pi(Q; \boldsymbol{\xi}) = \int \Upsilon(\boldsymbol{\mu}, \boldsymbol{\xi}) dQ(\boldsymbol{\mu})$$

where $\Upsilon(\boldsymbol{\mu}, \boldsymbol{\xi}) = \mathbb{P}(\chi_{K-1}^2(N\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})/2) > c_{K-1,\alpha})$ and

$$\lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) = \sum_{i=1}^K \xi_i(\mu_i - \bar{\mu})^2 = \sum_{i < j} \xi_i \xi_j (\mu_i - \mu_j)^2.$$

Observe that $\boldsymbol{\mu}$ is a realization from an exchangeable prior Q so $\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})$ is a RV. Consider a design $\boldsymbol{\xi}'$ obtained from $\boldsymbol{\xi}$ by switching its i th and j th elements. For example, if we exchange the first and second elements we get

$$\xi'_k = \begin{cases} \xi_1 & \text{if } k = 2 \\ \xi_2 & \text{if } k = 1 \\ \xi_k & \text{if } k \neq 1, 2 \end{cases} \tag{A.5}$$

It follows that

$$\begin{aligned} \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}') &= \sum_{i < j} \xi'_i \xi'_j (\mu_i - \mu_j)^2 \\ &= \xi'_1 \xi'_2 (\mu_1 - \mu_2)^2 + \sum_{j \neq 1, 2} \xi'_1 \xi'_j (\mu_1 - \mu_j)^2 + \sum_{j \neq 1, 2} \xi'_2 \xi'_j (\mu_2 - \mu_j)^2 \\ &\quad + \sum_{2 < i < j} \xi'_i \xi'_j (\mu_i - \mu_j)^2 \\ &= \xi_1 \xi_2 (\mu_1 - \mu_2)^2 + \sum_{j \neq 1, 2} \xi_2 \xi_j (\mu_1 - \mu_j)^2 + \sum_{j \neq 1, 2} \xi_1 \xi_j (\mu_2 - \mu_j)^2 \\ &\quad + \sum_{2 < i < j} \xi_i \xi_j (\mu_i - \mu_j)^2 \\ &\stackrel{d}{=} \xi_1 \xi_2 (\mu_1 - \mu_2)^2 + \sum_{j \neq 1, 2} \xi_2 \xi_j (\mu_2 - \mu_j)^2 + \sum_{j \neq 1, 2} \xi_1 \xi_j (\mu_1 - \mu_j)^2 \\ &\quad + \sum_{2 < i < j} \xi_i \xi_j (\mu_i - \mu_j)^2 = \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}), \end{aligned}$$

where the equality in distribution follows from exchangeability of Q . Since the indices we choose in (A.5) are arbitrary it follows that

$$\lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) \stackrel{d}{=} \lambda(\boldsymbol{\mu}; \sigma(\boldsymbol{\xi})),$$

where $\sigma(\boldsymbol{\xi})$ is any permutation of $\boldsymbol{\xi}$. Consequently $\Upsilon(\boldsymbol{\mu}, \boldsymbol{\xi}) \stackrel{d}{=} \Upsilon(\boldsymbol{\mu}, \sigma(\boldsymbol{\xi}))$ from which it follows that

$$\Pi(Q; \boldsymbol{\xi}) = \Pi(Q; \sigma(\boldsymbol{\xi})), \tag{A.6}$$

thus, by (A.6), $\Pi(Q; \boldsymbol{\xi})$ is permutation symmetric. Therefore, by Remark 1, $\max\{\Pi(Q; \boldsymbol{\xi}); \sum_{i=1}^K \xi_i = 1\}$ is attained when $\xi_i = \xi_j$ for all i, j . Hence $\boldsymbol{\xi}_Q^\pi = (1/K, \dots, 1/K)$. \square

Proof of Theorem 3.3

Proof. For $i < j$ let $\boldsymbol{\mu}_{ij} = (\delta/2)(\mathbf{e}_j - \mathbf{e}_i)$ denote a permutation of $(-\delta/2, \delta/2, 0, \dots, 0)$. Given Q in the statement of the Theorem we have

$$\begin{aligned}\Lambda(Q; \boldsymbol{\xi}) &= \int \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}) dQ(\boldsymbol{\mu}) = \frac{1}{2} \binom{K}{2}^{-1} \sum_{i \neq j} \lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi}) \\ &= \binom{K}{2}^{-1} \sum_{i < j} \lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi}) = \frac{\delta^2}{4} \binom{K}{2}^{-1} \sum_{i < j} \{(\xi_i + \xi_j) - (\xi_j - \xi_i)^2\}.\end{aligned}$$

Now maximizing $\sum_{i < j} \{(\xi_i + \xi_j) - (\xi_j - \xi_i)^2\}$ is equivalent to minimizing $\sum_{i < j} (\xi_j - \xi_i)^2$ which subject to $\sum_i \xi_i = 1$ and $\xi_i \geq 0$ is minimized at

$$(1/K, \dots, 1/K).$$

It is easy to verify that $\lambda(\boldsymbol{\mu}; \boldsymbol{\xi}_{\text{MM}}) \geq \Lambda(Q_{\text{LFP}}; \boldsymbol{\xi})$ for all $\boldsymbol{\mu} \in \mathcal{M}_\delta$ and $\boldsymbol{\xi} \in \Xi$. Thus, by Theorem 1 of [15] we get

$$\boldsymbol{\xi}_Q^\lambda = (1/K, \dots, 1/K)$$

and Q is the least favourable prior on \mathcal{M}_δ .

A general expression for average power with respect to the prior Q has been already given in equation (A.5) where $I_k(\lambda)$ is defined in (5). Integrating (5) with respect to the specified prior Q , yields,

$$\int I_k(\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})) dQ(\boldsymbol{\mu}) = \frac{1}{K(K-1)} \sum_{i < j} \frac{\exp(-\lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi})) (\lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi}))^k}{k!}, \quad (\text{A.7})$$

where

$$\lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi}) = (\delta^2/4) \{(\xi_i + \xi_j) - (\xi_j - \xi_i)^2\},$$

and $\boldsymbol{\mu}_{ij} = (\delta/2)(-\mathbf{e}_i + \mathbf{e}_j)$. Observe that the expression given in the right hand side of (A.7) is a symmetric function of $\{\xi_1, \dots, \xi_K\}$. Thus, by Remark 1, $\max\{\Pi(Q; \boldsymbol{\xi}) : \sum_i \xi_i = 1\}$ is attained when $\xi_i = \xi_j$ for all i, j . Again using Theorem 1 of [15], we get the desired result. \square

Proof of Theorem 4.1

Proof. First consider the game $G_A(\lambda, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$. Suppose that $\boldsymbol{\mu}_{ij}$ is Player's II strategy in a Nash Equilibrium. Clearly Player I best response to $\boldsymbol{\mu}_{ij}$ is $\boldsymbol{\xi}_{ij}$. However, Player II best response to $\boldsymbol{\xi}_{ij}$ is any $\boldsymbol{\mu}_{lk}$ where $(l, k) \neq (i, j)$ and it follows that $\boldsymbol{\mu}_{ij}$ can not be a strategy in a Nash Equilibrium. Since the latter argument applies to all indices (i, j) , it follows that Player I does not have a Nash Equilibrium in pure strategies. A similar argument shows that neither does Player II. It is straightforward to check that if

$$\mathbf{p}_0 = \mathbf{1}_{K(K-1)/2} \binom{K}{2}^{-1}$$

then for any two distinct strategies of Player II, e.g., $\boldsymbol{\mu}_{ij}$ and $\boldsymbol{\mu}_{kl}$, the expected payoff to Player I, denoted by $E_{\mathbf{p}}(\cdot)$, satisfies

$$\begin{aligned} E_{\mathbf{p}_0}(\lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi})) &= \binom{K}{2}^{-1} \sum_{s < t} \lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi}_{st}) = \binom{K}{2}^{-1} \sum_{s < t} \lambda(\boldsymbol{\mu}_{kl}; \boldsymbol{\xi}_{st}) \\ &= E_{\mathbf{p}_0}(\lambda(\boldsymbol{\mu}_{kl}; \boldsymbol{\xi})). \end{aligned} \tag{A.8}$$

Thus by Theorem 5.18 in [14], \mathbf{p}_0 is an optimal mixed strategy for Player I. Similarly, if

$$\mathbf{q}_0 = \mathbf{1}_{K(K-1)/2} \binom{K}{2}^{-1}$$

then for any two distinct strategies of Player I, e.g., $\boldsymbol{\xi}_{ij}$ and $\boldsymbol{\xi}_{kl}$ the expected payoff of Player II satisfies

$$\begin{aligned} E_{\mathbf{q}_0}(-\lambda(\boldsymbol{\mu}; \boldsymbol{\xi}_{ij})) &= -\binom{K}{2}^{-1} \sum_{s < t} \lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi}_{st}) = -\binom{K}{2}^{-1} \sum_{s < t} \lambda(\boldsymbol{\mu}_{st}; \boldsymbol{\xi}_{kl}) \\ &= E_{\mathbf{q}_0}(-\lambda(\boldsymbol{\mu}; \boldsymbol{\xi}_{kl})), \end{aligned}$$

which shows that \mathbf{q}_0 is an optimal mixed strategy for Player II. It follows that $(\mathbf{p}_0, \mathbf{q}_0)$ is a Nash Equilibrium.

Next we prove the solution $(\mathbf{p}_0, \mathbf{q}_0)$ is unique. As in Equation (4.1) let \mathcal{I}_1 denote the set of indices $\{(i, j), (k, l)\}$ for which $\lambda(\boldsymbol{\mu}_{ij}, \boldsymbol{\xi}_{kl}) = \delta^2/16$ and let $|\mathcal{I}_1|$ denote its cardinality. It can be shown that $|\mathcal{I}_1| = K(K-1)(K-2)$. The payoff at (\mathbf{p}, \mathbf{q}) is

$$\Gamma(\mathbf{p}, \mathbf{A}, \mathbf{q}) = \frac{\delta^2}{4} \left(\sum_{i < j} p_{ij} q_{ij} + \frac{1}{4} \sum_{\{(i,j), (k,l)\} \in \mathcal{I}_1} p_{ij} q_{kl} \right). \tag{A.9}$$

Therefore, the payoff at $(\mathbf{p}_0, \mathbf{q}_0)$ is

$$\Gamma(\mathbf{p}_0, \mathbf{A}, \mathbf{q}_0) = \frac{\delta^2}{4} \left\{ \frac{2}{K(K-1)} + \frac{|\mathcal{I}_1|}{4} \left(\frac{2}{K(K-1)} \right)^2 \right\} = \frac{\delta^2}{4(K-1)}. \tag{A.10}$$

Suppose that $(\mathbf{p}', \mathbf{q}') \neq (\mathbf{p}_0, \mathbf{q}_0)$ is a different Nash Equilibrium. Let

$$\begin{aligned} \mathcal{S}_0 &= \{(i, j) : p'_{ij} = \binom{K}{2}^{-1}\} \\ \mathcal{S}_- &= \{(i, j) : p'_{ij} = \binom{K}{2}^{-1} - a_{ij}, \text{ where } a_{ij} > 0\} \\ \mathcal{S}_+ &= \{(i, j) : p'_{ij} = \binom{K}{2}^{-1} + b_{ij}, \text{ where } b_{ij} > 0\} \end{aligned}$$

The sets \mathcal{S}_0 , \mathcal{S}_- and \mathcal{S}_+ are mutually exclusive and their union is the set of all pairs (i, j) . Since $\mathbf{p}' \neq \binom{K}{2}^{-1} \mathbf{1}_{K(K-1)/2}$, the sets \mathcal{S}_- and \mathcal{S}_+ are not empty.

In response to the mixed strategy \mathbf{p}' Player II may choose a mixed strategy \mathbf{q}' satisfying

$$q'_{ij} = \begin{cases} \binom{K}{2}^{-1} + a_{ij} & \text{if } (i, j) \in \mathcal{S}_- \\ \binom{K}{2}^{-1} & \text{if } (i, j) \in \mathcal{S}_0 \\ \binom{K}{2}^{-1} - b_{ij} & \text{if } (i, j) \in \mathcal{S}_+ \end{cases}$$

Note that

$$\begin{aligned} \sum_{i < j} p'_{ij} q'_{ij} &= \sum_{(i,j) \in \mathcal{S}_-} p'_{ij} q'_{ij} + \sum_{(i,j) \in \mathcal{S}_0} p'_{ij} q'_{ij} + \sum_{(i,j) \in \mathcal{S}_+} p'_{ij} q'_{ij} \\ &= |\mathcal{S}_-| \binom{K}{2}^{-2} - \sum_{(i,j) \in \mathcal{S}_-} a_{ij}^2 + |\mathcal{S}_0| \binom{K}{2}^{-2} + |\mathcal{S}_+| \binom{K}{2}^{-2} \\ &\quad - \sum_{(i,j) \in \mathcal{S}_+} b_{ij}^2 \\ &= \binom{K}{2}^{-1} - \left(\sum_{(i,j) \in \mathcal{S}_-} a_{ij}^2 + \sum_{(i,j) \in \mathcal{S}_+} b_{ij}^2 \right) \end{aligned} \tag{A.11}$$

and

$$\begin{aligned} \sum_{\{(i,j),(k,l)\} \in \mathcal{I}_1} p'_{ij} q'_{kl} &= \sum_{\{(i,j),(k,l)\} \in \mathcal{S}_- \cap \mathcal{I}_1} p'_{ij} q'_{kl} + \sum_{\{(i,j),(k,l)\} \in \mathcal{S}_0 \cap \mathcal{I}_1} p'_{ij} q'_{kl} \\ &\quad + \sum_{\{(i,j),(k,l)\} \in \mathcal{S}_+ \cap \mathcal{I}_1} p'_{ij} q'_{kl} \\ &= |\mathcal{S}_- \cap \mathcal{I}_1| \binom{K}{2}^{-2} \\ &\quad - \sum_{\{(i,j),(k,l)\} \in \mathcal{S}_- \cap \mathcal{I}_1} a_{ij} a_{kl} + |\mathcal{S}_0 \cap \mathcal{I}_1| \binom{K}{2}^{-2} \\ &\quad + |\mathcal{S}_+ \cap \mathcal{I}_1| \binom{K}{2}^{-2} - \sum_{\{(i,j),(k,l)\} \in \mathcal{S}_+ \cap \mathcal{I}_1} b_{ij} b_{kl} \\ &= |\mathcal{R}| \binom{K}{2}^{-2} - \sum_{\{(i,j),(k,l)\} \in \mathcal{S}_- \cap \mathcal{I}_1} a_{ij} a_{kl} \\ &\quad - \sum_{\{(i,j),(k,l)\} \in \mathcal{S}_+ \cap \mathcal{I}_1} b_{ij} b_{kl}, \end{aligned} \tag{A.12}$$

where (A.12) holds since $|\mathcal{S}_- \cap \mathcal{I}_1| + |\mathcal{S}_0 \cap \mathcal{I}_1| + |\mathcal{S}_+ \cap \mathcal{I}_1| = |\mathcal{I}_1|$. So, it now follows from (A.9), (A.10), (A.11), and (A.12) that $\Gamma(\mathbf{p}', \mathbf{\Lambda}, \mathbf{q}') < \Gamma(\mathbf{p}_0, \mathbf{\Lambda}, \mathbf{q}_0)$ so \mathbf{p}' can not be a min-max optimal strategy for Player I. A similar argument holds for Player II. Since the game is a zero sum game the maxi-min solutions coincide with the Nash Equilibria we conclude that $(\mathbf{p}_0, \mathbf{q}_0)$ is the unique Nash Equilibrium in $G_A(\lambda, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$. Repeating the arguments above shows that

$(\mathbf{p}_0, \mathbf{q}_0)$ is also the unique Nash Equilibrium in $G_A(\pi, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$. This completes the proof. \square

Proof of Theorem 4.2

Proof. We start by showing that if $\mathfrak{M}_l \subset \mathfrak{M}$ then $v(\lambda, \mathfrak{X}, \mathfrak{M}_l) > v(\lambda, \mathfrak{X}, \mathfrak{M})$.

Recall first that by Theorem 4.1

$$\mathbf{p} = \mathbf{q} = \binom{K}{2}^{-1} \mathbf{1}_{\binom{K}{2}}$$

is the Nash equilibrium of the game $G_A(\lambda, \mathbf{P}(\mathfrak{X}), \mathbf{Q}(\mathfrak{M}))$ whose value is

$$v(\lambda, \mathfrak{X}, \mathfrak{M}) = \frac{\delta^2}{4} \frac{1}{(K-1)}. \tag{A.13}$$

Let $\mathfrak{M}_l \subset \mathfrak{M}$ be the strategies available to Player II after removing the T pure strategies $\boldsymbol{\mu}_{i_1 j_1}, \dots, \boldsymbol{\mu}_{i_T j_T}$ from \mathfrak{M} . Since the value of any zero sum game can not decrease when the strategy space for Player II is restricted we have

$$v(\lambda, \mathfrak{X}, \mathfrak{M}_l) \geq v(\lambda, \mathfrak{X}, \mathfrak{M}). \tag{A.14}$$

It follows that if (A.14) holds with a strict inequality for $T = 1$, i.e., when only one strategy is removed, then it will hold with a strict inequality for all $T \in \{2, \dots, K-1\}$. Symmetry considerations imply that the identity of the removed strategy is immaterial. Hence, without any loss of generality, we may restrict our attention to the case where the strategy $\boldsymbol{\mu}_{K-1, K}$ is removed from \mathfrak{M} . Furthermore, it is clear that

$$v(\lambda, \mathfrak{X}, \mathfrak{M} \setminus \boldsymbol{\mu}_{K-1, K}) \geq v(\lambda, \mathfrak{X} \setminus \boldsymbol{\xi}_{K-1, K}, \mathfrak{M} \setminus \boldsymbol{\mu}_{K-1, K}). \tag{A.15}$$

We now continue with an analysis of the game $G_A(\lambda, \mathbf{P}(\mathfrak{X} \setminus \boldsymbol{\xi}_{K-1, K}), \mathbf{Q}(\mathfrak{M} \setminus \boldsymbol{\mu}_{K-1, K}))$. By Theorem 5.22 in [14] the value of this game is equal to the solution of the linear program

$$v = \max \{v_{12}, \dots, v_{K-2, K}\}$$

subject to the inequalities

$$\sum_{i < j} p_{ij} \lambda(\boldsymbol{\mu}_{st}; \boldsymbol{\xi}_{ij}) \geq v_{st} \tag{A.16}$$

for all pairs $1 \leq s < t \leq K$ excluding the pair $(s, t) = (K-1, K)$. As usual $\sum_{i < j} p_{ij} = 1$ with $p_{ij} \geq 0$ for all (i, j) and recall that in the game $G_A(\lambda, \mathbf{P}(\mathfrak{X} \setminus \boldsymbol{\xi}_{K-1, K}), \mathbf{Q}(\mathfrak{M} \setminus \boldsymbol{\mu}_{K-1, K}))$ we have $p_{K-1, K} = 0$. Further note that by (14) we may rewrite (A.16) as

$$\frac{\delta^2}{4} \left\{ p_{st} + \frac{1}{4} \sum_{(i, j) \in G(s, t)} p_{ij} \right\} \geq v_{st} \tag{A.17}$$

where $G(s, t) = \{(i, j) : i \notin \{s, t\} \vee j \notin \{s, t\}\}$. Now, let

$$\begin{aligned}\mathcal{A} &= \{\xi_{ij} : i \notin \{K-1, K\} \wedge j \notin \{K-1, K\}\}, \\ \mathcal{B} &= \{\xi_{ij} : i \in \{K-1, K\} \vee j \in \{K-1, K\}\}\end{aligned}$$

i.e., \mathcal{A} is the collections of strategies indexed by pairs (i, j) in which both i and j are distinct from $\{K-1, K\}$, whereas \mathcal{B} is the collections of strategies in which either i or j but not both are in $\{K-1, K\}$. It is clear, by symmetry considerations, that at equilibrium

$$p_{st} = p_{uv}$$

when (s, t) and (u, v) are either both in \mathcal{A} or both in \mathcal{B} . Therefore at equilibrium we have

$$p_{ij} = \begin{cases} \alpha \binom{K}{2}^{-1} & \text{if } (i, j) \in \mathcal{A} \\ \beta \binom{K}{2}^{-1} & \text{if } (i, j) \in \mathcal{B} \end{cases} \quad (\text{A.18})$$

for some α and β to be determined. Equation (A.18) shows that the mass associated with $\xi_{K-1, K}$ is redistributed among the strategies in \mathcal{A} and \mathcal{B} . Since $\sum_{i < j} p_{ij} = 1$ we have

$$A\alpha + B\beta = \binom{K}{2}, \quad (\text{A.19})$$

where $A = \binom{K}{2} - B - 1$ and $B = 2(K-2)$ are the cardinalities of the sets \mathcal{A} and \mathcal{B} respectively. Carefully substituting (A.18) into (A.17) we find that

$$v_{ij} = \begin{cases} \frac{\delta^2}{4}(\alpha \binom{K}{2}^{-1} + \frac{1}{4}\{B\beta \binom{K}{2}^{-1}\}) & \text{if } (i, j) \in \mathcal{A} \\ \frac{\delta^2}{4}(\beta \binom{K}{2}^{-1} + \frac{1}{4}\{\alpha \binom{K}{2}^{-1} + (B-2)\beta \binom{K}{2}^{-1}\}) & \text{if } (i, j) \in \mathcal{B} \end{cases}. \quad (\text{A.20})$$

At the optimum $v = v_{12} = v_{13} = \dots = v_{K-2, K}$. In particular $v_{ij} = v_{i'j'}$ when $(i, j) \in \mathcal{A}$ and $(i', j') \in \mathcal{B}$ which implies that

$$\alpha + \frac{1}{4}\{B\beta\} = \beta + \frac{1}{4}\{\alpha + (B-2)\beta\} \quad (\text{A.21})$$

Solving (A.19) and (A.21) for α and β we find that

$$\alpha = \frac{2}{3}\beta \quad \text{and} \quad \beta = \frac{3}{2} \frac{K(K-1)}{K^2 + K - 6}. \quad (\text{A.22})$$

Plugging (A.22) into (A.20) and with a bit of algebra we find that

$$v(\lambda, \mathfrak{X} \setminus \xi_{K-1, K}, \mathfrak{M} \setminus \mu_{K-1, K}) = \frac{\delta^2}{4} \frac{3K-2}{2(K^2 + K - 6)}. \quad (\text{A.23})$$

It is straightforward to verify that (A.23) is larger than (A.13) for all $K \geq 3$ which together with (A.15) completes the first part of the proof.

Next, we show that if $\mathfrak{M} \subset \mathfrak{M}_b \subset \mathcal{M}_\delta$ then $v(\lambda, \mathfrak{X}, \mathfrak{M}) = v(\lambda, \mathfrak{X}, \mathfrak{M}_b)$. The reasoning here is quite simple. Suppose that Player I uses the mixed strategy $\mathbf{p} = \binom{K}{2}^{-1} \mathbf{1}_{K(K-1)/2}$. Then for any strategy $\boldsymbol{\mu}$ of Player II the payoff to Player I

$$\sum_{i < j} \binom{K}{2}^{-1} \lambda(\boldsymbol{\mu}; \boldsymbol{\xi}_{ij}) = \frac{1}{2K(K-1)} \sum_{i < j} (\mu_i - \mu_j)^2. \quad (\text{A.24})$$

Now, it is not hard to see that $\min\{\sum_{i < j} (\mu_i - \mu_j)^2 : \boldsymbol{\mu} \in \mathcal{M}_\delta\}$ is attained at any permutation of the LFC $(-\delta/2, \delta/2, 0, \dots, 0)$ in which case the value of $\sum_{i < j} (\mu_i - \mu_j)^2$ is $\delta^2 K/2$. It follows that the right hand side of (A.24) is larger or equal to $\delta^2/4(K-1)$ which is nothing but $v(\lambda, \mathfrak{X}, \mathfrak{M})$. Moreover if $\boldsymbol{\mu} \in \mathcal{M}_\delta \setminus \mathfrak{M}$ then the right hand side of (A.24) is strictly larger than $v(\lambda, \mathfrak{X}, \mathfrak{M})$. However, when (A.24) holds with a strict inequality then Player II will assign 0 probability to this strategy, otherwise his loss is expected to increase. Thus $v(\lambda, \mathfrak{X}, \mathfrak{M}) = v(\lambda, \mathfrak{X}, \mathfrak{M}_b)$. If (A.24) holds with an equality then assigning a positive probability to it will not change the value of the game so again $v(\lambda, \mathfrak{X}, \mathfrak{M}) = v(\lambda, \mathfrak{X}, \mathfrak{M}_b)$. This completes the proof. \square

Proof of Theorem 4.3

Proof. First we consider the game $G_A(\lambda, \Xi, \mathfrak{M})$. The Nash Equilibrium in this case is the the pair $(\boldsymbol{\xi}_0, \mathbf{q}_0)$ which solves

$$\max_{\boldsymbol{\xi} \in \Xi} \min_{\mathbf{q} \in \mathcal{Q}} E_{\mathbf{q}}(\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})), \quad (\text{A.25})$$

where $E_{\mathbf{q}}(\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})) = \sum_{i < j} q_{ij} \lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi})$. Let $\boldsymbol{\xi} \in \Xi$ and $\mathbf{q}_0 = \mathbf{1}_{K(K-1)/2} \binom{K}{2}^{-1}$ and note that

$$E_{\mathbf{q}_0}(\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})) = \binom{K}{2}^{-1} \sum_{i < j} \lambda(\boldsymbol{\mu}_{ij}; \boldsymbol{\xi}) = \frac{\delta^2}{4} \binom{K}{2}^{-1} \sum_{i < j} ((\xi_i + \xi_j) - (\xi_i - \xi_j)^2)$$

is maximized when $\xi_i = \xi_j$ for all pairs (i, j) as in the proof of Theorem 3.3). Thus,

$$\max_{\boldsymbol{\xi} \in \Xi} E_{\mathbf{q}_0}(\lambda(\boldsymbol{\mu}; \boldsymbol{\xi})) = E_{\mathbf{q}_0}(\lambda(\boldsymbol{\mu}; K^{-1} \mathbf{1}_K)),$$

so $\boldsymbol{\xi}_0 = K^{-1} \mathbf{1}_K$ is the unique best response to \mathbf{q}_0 . The uniqueness of Nash Equilibrium $(\boldsymbol{\xi}_0, \mathbf{q}_0)$ can be proved by adopting and appropriately modifying the method used in the proof of Theorem 4.1.

Next consider the game $G_A(\pi, \Xi, \mathfrak{M})$. Let $\boldsymbol{\xi} \in \Xi$ and $\mathbf{q}_0 = \mathbf{1}_{K(K-1)/2} \binom{K}{2}^{-1}$. From (A.4), it follows that

$$\begin{aligned} \pi(\boldsymbol{\xi}; \mathbf{q}_0) &= E_{\mathbf{q}_0}(\pi(\boldsymbol{\mu}; \boldsymbol{\xi})) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(\chi_{K-1+2k}^2 > c_{K-1, \alpha}) \left\{ \frac{1}{\binom{K}{2}} \sum_{i < j} \frac{e^{-\lambda_{ij}/2} (\lambda_{ij}/2)^k}{k!} \right\} \end{aligned} \quad (\text{A.26})$$

where λ_{ij} is the NCP given by

$$\lambda_{ij} = \frac{\delta^2}{4}(\xi_i + \xi_j) - (\xi_i - \xi_j)^2. \quad (\text{A.27})$$

From equations (A.27) and (A.26), it is observed that $E_{\mathbf{q}_0}(\pi(\boldsymbol{\mu}; \boldsymbol{\xi}))$ is symmetric in (ξ_1, \dots, ξ_K) . Thus, by Remark 1, $\max\{E_{\mathbf{q}_0}(\pi(\boldsymbol{\mu}; \boldsymbol{\xi})) : \sum_i \xi_i = 1\}$ is attained when $\xi_i = \xi_j$ for all i, j , i.e., $\boldsymbol{\xi} = K^{-1}\mathbf{1}_K$ is an optimal solution. The uniqueness of the solution can be proved similarly as we proved for the game $G_A(\pi, \Xi, \mathfrak{M})$. \square

Acknowledgments

The research leading to this paper was conducted (mostly) while the first author was a postdoctoral fellow at the University of Haifa, Israel. The work of the second author is supported by the Israeli Science Foundation Grant No. 456/17 and gratefully acknowledged.

References

- [1] ATKINSON, A. C., DONEV, A. N. and TOBIAS, R. D. (2007). *Optimum Experimental Designs, With SAS*. Oxford University Press, Oxford. [MR2323647](#)
- [2] BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference under Order Restrictions*. Wiley, New York.
- [3] BERGER, J. O. (1980). *Statistical Decisions Theory, Foundations, Concepts and Methods*. Springer-Verlag, New York. [MR0580664](#)
- [4] BLACKWELL, D. and GIRSHICK, M. (1954). *Theory of Games and Statistical Decisions*. Wiley, New York. [MR1570712](#)
- [5] BOX, G. E. P. and TIAO, G. C. (1992). *Bayesian Inference in Statistical Analysis*. Wiley, New York. [MR0418321](#)
- [6] CHALONER, K. and VERDINELLI, I. (1995). Bayesian experimental design: A review. *Statistical Science* **10** 237–302. [MR1390519](#)
- [7] DANIELE, E., DE PAOLIS, P., GRECO, G. L. and D'ARGENIO, A. (2017). Location methods and Nash equilibria for experimental design in astrophysics and aerospace engineering. In *Spatial Interaction Models. Springer Optimization and Its Applications*, (L. Mallozzi, E. D'Amato and P. Pardalos, eds.) **118** 53–72. Springer, Cham. [MR3674762](#)
- [8] DUNCAN, D. B. (1965). A Bayesian approach to multiple comparisons. *Technometrics* **7** 171–222.
- [9] D'AMATO, E., ELIA, D. and MALLOZZI, L. (2014). Experimental design problems and Nash equilibrium solutions. In *Network Models in Economics and Finance. Springer Optimization and Its Applications*, (V. A. Kalaygin, P. M. Pardalos and T. M. Rassias, eds.) **100** 1–12. Springer, Cham. [MR3329506](#)

- [10] FEDOROV, V. V. and HACKL, P. (1997). *Model-Oriented Design of Experiments*. Springer, New York. [MR1454123](#)
- [11] HOLT, C. A. and ROTH, A. E. (2004). The Nash equilibrium: A perspective. *Proceedings of the National Academy of Sciences* **101** 3999–4002. [MR2044851](#)
- [12] KIEFER, J. C. (1971). The role of symmetry and approximation in exact design optimality. In *S. S. Gupta and J. Yackel (ed.), Statistical Decision Theory and Related Topics, Proceedings of a Symposium*, Purdue University, New York, 109–118. [MR0350985](#)
- [13] MAGNUS, J. R. (1978). The moments of products of quadratic forms in normal variables. *Statistica Neerlandica* **32** 201–210. [MR0528403](#)
- [14] MASCHLER, M., SOLAN, E. and ZAMIR, S. (2013). *Game Theory*. Cambridge University Press, Cambridge. [MR3154588](#)
- [15] NYQUIST, H. (2013). Convergence of an algorithm for constructing minimax designs. In *mODa 10 – Advances in Model Oriented Design and Analysis* (D. Uciński and A. C. Atkinson, eds.) 187–194. Springer, Heidelberg.
- [16] PRONZATO, L. and WALTER, E. (1985). Robust experiment design via stochastic approximation. *Mathematical Biosciences* **75** 103–120. [MR0800967](#)
- [17] PUKELSHEIM, F. (1993). *Optimal Design of Experiments*. Wiley, New York. [MR1211416](#)
- [18] SCHWARZ, G. (1994). Game theory and statistics. In *Handbook of Game Theory*, (R. J. Aumann and S. Hart, eds.) **2** 21, 769–779. Elsevier, New York. [MR1313217](#)
- [19] SILVAPULLE, M. and SEN, P. (2005). *Constrained Statistical Inference: Order, Inequality, and Shape Constraints*. Hoboken, NJ, New York. [MR2099529](#)
- [20] SINGH, S. P. and DAVIDOV, O. (2019). On the design of experiments with ordered treatments. *Journal of Royal Statistical Society: Series B* **81** 881–900. [MR4025401](#)
- [21] SINGH, S. P. and MUKHOPADHYAY, S. (2016a). Bayesian crossover designs for generalized linear models. *Computational Statistics and Data Analysis* **104** 35–50. [MR3540985](#)
- [22] SINGH, S. P. and MUKHOPADHYAY, S. (2016b). Bayesian optimal cluster designs. *Statistical Methodology* **32** 36–52.
- [23] WALD, A. (1949). Statistical decision functions. *Annals of Mathematical Statistics* **20** 165–205. [MR0044802](#)
- [24] WATERHOUSE, W. C. (1983). Do symmetric problems have symmetric solutions? *The American Mathematical Monthly* **90** 378–388. [MR0707152](#)
- [25] WOODS, D. C. and VAN DE VEN, P. M. (2011). Blocked designs for experiments with correlated non-normal response. *Technometrics* **53** 173–182. [MR2808343](#)
- [26] WU, C. F. (1981). On the robustness and efficiency of some randomized designs. *The Annals of Statistics* **9** 1168–1177. [MR0630100](#)