The Multivariate Tukey-Kramer Multiple Comparison Procedure Among Four Correlated Mean Vectors

Takashi Seo[a] *, Takahiro Nishiyama[b]

[a] Department of Mathematical Information Science , Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo, 162-8601 Japan

[b] Graduate School of Science, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku-ku, Tokyo, 162-8601 Japan

Abstract

In this paper, conservative simultaneous confidence intervals for pairwise comparisons among mean vectors in multivariate normal distributions are considered. The affirmative proof of the multivariate generalized Tukey conjecture in the case of four mean vectors can be completed. Further, the upper bound for the conservativeness of the multivariate Tukey-Kramer procedure is also given. Finally, numerical results by Monte Carlo simulations are given.

Key Words: Conservativeness; Coverage probability; Multivariate Tukey-Kramer procedure; Monte Carlo simulation; Pairwise comparisons

*Corresponding author. Tel.: +81-3-5228-8199; fax: +81-3-3260-4293.
E-mail address: seo@rs.kagu.tus.ac.jp.

1. Introduction

Consider the simultaneous confidence intervals for pairwise multiple comparisons among mean vectors from the multivariate normal populations. Let $\boldsymbol{M} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k]$ be the unknown $p \times k$ matrix of k mean vectors corresponding to the k treatments where $\boldsymbol{\mu}_i$ is the mean vector from ith population. Also, let $\widehat{\boldsymbol{M}} = [\widehat{\boldsymbol{\mu}}_1, \ldots, \widehat{\boldsymbol{\mu}}_k]$ be the estimator of $\boldsymbol{M}$ such that $\mathrm{vec}(\boldsymbol{X})$ is distributed as $N_{kp}(\mathbf{0}, \boldsymbol{V} \otimes \boldsymbol{\Sigma})$, where $\boldsymbol{X} = \widehat{\boldsymbol{M}} - \boldsymbol{M}$, $\boldsymbol{V} = [v_{ij}]$ is a known $k \times k$ positive definite matrix and $\boldsymbol{\Sigma}$ is an unknown $p \times p$ positive definite matrix, and $\mathrm{vec}(\cdot)$ denotes the column vector formed by stacking the columns of the matrix under each other. Further, we assume that $\boldsymbol{S}$ is an unbiased estimator of $\boldsymbol{\Sigma}$ such that $\nu\boldsymbol{S}$ is independent of $\widehat{\boldsymbol{M}}$ and is distributed as a Wishart distribution $\mathrm{W}_p(\boldsymbol{\Sigma}, \nu)$. Then, we have the simultaneous confidence intervals for pairwise comparisons among mean vectors given by

$$\boldsymbol{a}'\boldsymbol{M}\boldsymbol{b} \in \left[\ \boldsymbol{a}'\widehat{\boldsymbol{M}}\boldsymbol{b} \pm t(\boldsymbol{b}'\boldsymbol{V}\boldsymbol{b})^{1/2}(\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a})^{1/2}\right], \quad \forall \boldsymbol{a} \in \mathbb{R}^p,\ \forall \boldsymbol{b} \in \mathbb{B}, \tag{1}$$

where $\mathbb{R}^p$ is the set of any nonzero real p-dimensional vectors and $\mathbb{B}$ is a subset in the k-dimensional space such that

$$\mathbb{B} = \{\boldsymbol{b} \in \mathbb{R}^k : \boldsymbol{b} = \boldsymbol{e}_i - \boldsymbol{e}_j,\ 1 \le i < j \le k\},$$

$\boldsymbol{e}_i$ is a unit vector of the k-dimensional space having 1 at i-th component and 0 at others. We note that (1) can be expressed as

$$\boldsymbol{a}'(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \in \left[\ \boldsymbol{a}'(\widehat{\boldsymbol{\mu}}_i - \widehat{\boldsymbol{\mu}}_j) \pm t\ (d_{ij}\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a})^{1/2}\right], \forall \boldsymbol{a} \in \mathbb{R}^p,\ 1 \le i < j \le k,$$

where $d_{ij} = v_{ii} - 2v_{ij} + v_{jj}$.

In order to give the simultaneous confidence intervals, however, we have to decide the unknown value of $t(> 0)$ which is the upper 100α percentile of the $T^2_{\max}$ statistic,

$$T^2_{\max\cdot\mathrm{p}} = \max_{\boldsymbol{b}\in\mathbb{B}} \left\{ \frac{(\boldsymbol{X}\boldsymbol{b})'\boldsymbol{S}^{-1}\boldsymbol{X}\boldsymbol{b}}{\boldsymbol{b}'\boldsymbol{V}\boldsymbol{b}} \right\} \tag{2}$$

$$= \max_{1\le i<j\le k} \{(\boldsymbol{x}_i - \boldsymbol{x}_j)'(d_{ij}\boldsymbol{S})^{-1}(\boldsymbol{x}_i - \boldsymbol{x}_j)\}. \tag{3}$$

Unfortunately, it is difficult to find the exact value of t. Then, some asymptotic approximations for the upper percentiles of $T^2_{\max\cdot\mathrm{p}}$ statistic have been discussed by Siotani (1959a, 1959b, 1960), Krishnaiah (1979), Siotani, Hayakawa and Fujikoshi (1985), Seo and Siotani (1992), Seo (1995), and so on.

In the case that $\boldsymbol{V} = \boldsymbol{I}$, $T^2_{\max\cdot\mathrm{p}}$ statistic is reduced as the same as half of the multivariate Studentized range statistic $R^2_{\max}$ (see, e.g., Seo and Siotani (1992)). So, as an approximation procedure, Seo, Mano and Fujikoshi (1994) proposed the multivariate Tukey-Kramer procedure which is a simple procedure by replacing with the upper percentile of the $R^2_{\max}$ statistic as an approximation to the one of $T^2_{\max\cdot\mathrm{p}}$ statistic for any positive definite matrix $\boldsymbol{V}$. This procedure is the multivariate version of Tukey-Kramer procedure (Tukey 1953; Kramer 1956, 1957). The Tukey-Kramer procedure is an attractive and simple procedure for pairwise multiple comparisons. (see, e.g., Hochberg and Tamhane(1987)).

As for the univariate Tukey-Kramer procedure, the generalized Tukey conjecture is known as the statement that the TK procedure yields the conservative simultaneous confidence intervals for all pairwise comparisons among means (see, e.g., Benjamini and Braun (2002)). Even for the univariate case, there has been no analytical proof of the generalized Tukey conjecture except the special cases. Theoretical discussions related to this conjecture are referred to Hayter(1984, 1989), Brown(1984), Uusipaikka (1985) and Spurrier and Isham (1985). Further, Lin, Seppänen and Uusipaikka(1990) have discussed the generalized Tukey conjecture for pairwise comparisons among the components of the mean vector.

As for the multivariate Tukey-Kramer procedure, the multivariate version of the generalized Tukey conjecture has been affirmatively proved in the case of three correlated mean vectors by Seo, Mano and Fujikoshi (1994). Further, relating to multivariate Tukey-Kramer conjecture, Seo (1996) gave the upper bound for conservativeness of the procedure for the pairwise comparisons among mean vectors. The related discussion for the univariate case is referred to Somerville (1993). The purpose of this paper is to give the affirmative proof of the multivariate generalized Tukey conjecture in the case of four mean vectors.

Although we do not discuss for comparisons with a control, Seo(1995) proposed the conservative procedure for the case of comparisons with a control, which is similar to the multivariate Tukey-Kramer procedure. Also, Seo and Nishiyama(2006) discuss the bound for these conservative procedures for pairwise comparisons and comparisons with a control in the case of three correlated mean vectors.

The organization of the paper is as follows. In Section 2, the conservativeness of the multivariate Tukey-Kramer procedure for four mean vectors and its upper bound for the conservativeness are discussed. Finally, we also give some numerical results by Monte Carlo simulations.

2. The multivariate Tukey-Kramer procedure

The simultaneous confidence intervals for all pairwise comparisons by the multivariate Tukey-Kramer procedure are given by

$$\boldsymbol{a}'(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \in \Big[\, \boldsymbol{a}'(\widehat{\boldsymbol{\mu}}_i - \widehat{\boldsymbol{\mu}}_j) \pm t_{\mathrm{p}\cdot \boldsymbol{I}}\sqrt{d_{ij}\boldsymbol{a}'\boldsymbol{S}\boldsymbol{a}}\Big], \forall \boldsymbol{a} \in \mathbb{R}^p,\ 1 \le i < j \le k, \qquad (4)$$

where $t^2_{\mathrm{p}\cdot \boldsymbol{I}}$ is the upper α percentile of $T^2_{\max \cdot \mathrm{p}}$ statistic with $\boldsymbol{V} = \boldsymbol{I}$, that is, $t^2_{\mathrm{p}\cdot \boldsymbol{I}} = q^2/2$ and $q^2 \equiv q^2_{p,k,\nu}(\alpha)$ is the upper α percentile of the p-variate Studentized range statistic with parameters k and ν. By a reduction of relating to the coverage probability of (4), Seo, Mano and Fujikoshi (1994) proved that the coverage probability in the case $k = 3$ is equal or greater than $1 - \alpha$ for any positive definite matrix $\boldsymbol{V}$. Using the similar reduction, Seo (1996) discussed the bound of conservative simultaneous confidence levels.

Consider the probability

$$Q(t, \boldsymbol{V}, \mathbb{B}) = \Pr\{\, (\boldsymbol{X}\boldsymbol{b})'(\nu \boldsymbol{S})^{-1}(\boldsymbol{X}\boldsymbol{b}) \le t(\boldsymbol{b}'\boldsymbol{V}\boldsymbol{b}),\ \forall \boldsymbol{b} \in \mathbb{B}\,\}, \tag{5}$$

where t is any fixed constant. Without loss of generality, we may assume $\boldsymbol{\Sigma} = \boldsymbol{I}_p$ when we consider the probability (5).

When $t = t_{\mathrm{p}}^*(\equiv t_{\mathrm{p}\cdot I}^2/\nu)$ and $\mathbb{B} = \mathbb{C}$, the coverage probability (5) is the same as the coverage probability of (4). The conservativeness of the simultaneous confidence intervals (4) means that $Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C}) \ge Q(t_{\mathrm{p}}^*, \boldsymbol{I}, \mathbb{C}) = 1 - \alpha$. The inequality is known as the multivariate generalized Tukey conjecture. Then we have the following theorem for the case $k = 3$ by using same line of the proof of Theorem 3.2 in Seo, Mano and Fujikoshi (1994).

Theorem 1. *Let $Q(t, \boldsymbol{V}, \mathbb{B})$ be the coverage probability (5) with a known matrix $\boldsymbol{V}$ for the case $k = 3$. Then, for any positive definite matrix $\boldsymbol{V}$, it holds that*

$$1 - \alpha = Q(t_{\mathrm{p}}^*, \boldsymbol{I}, \mathbb{C}) \le Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C}) < Q(t_{\mathrm{p}}^*, \boldsymbol{V}_0, \mathbb{C}),$$

where $t_{\mathrm{p}}^* = t_{\mathrm{p}\cdot I}^2/\nu$, $\mathbb{C} = \{\boldsymbol{c} \in \mathbb{R}^k : \boldsymbol{c} = \boldsymbol{e}_i - \boldsymbol{e}_j,\ 1 \le i < j \le k\}$ and $\boldsymbol{V}_0$ has the condition such that $\sqrt{d_{12}} = \sqrt{d_{13}} + \sqrt{d_{23}}$ or $\sqrt{d_{13}} = \sqrt{d_{12}} + \sqrt{d_{23}}$ or $\sqrt{d_{23}} = \sqrt{d_{12}} + \sqrt{d_{13}}$.

In connection with Theorem 1, we have the following conjecture in the case of $k \ge 4$.

Conjecture 2. *Let $Q(t, \boldsymbol{V}, \mathbb{B})$ be the coverage probability for (5) with a known matrix $\boldsymbol{V}$. Then, for any positive definite matrix $\boldsymbol{V}$, it holds that*

$$1 - \alpha = Q(t_{\mathrm{p}}^*, \boldsymbol{I}, \mathbb{C}) \le Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C}) < Q(t_{\mathrm{p}}^*, \boldsymbol{V}_1, \mathbb{C}),$$

where $t_{\mathrm{p}}^* = t_{\mathrm{p}\cdot I}^2/\nu$, $\mathbb{C} = \{\boldsymbol{c} \in \boldsymbol{R}^k : \boldsymbol{c} = \boldsymbol{e}_i - \boldsymbol{e}_j,\ 1 \le i < j \le k\}$, $\boldsymbol{V}_1$ satisfies with one of the conditions "$\sqrt{d_{ij}} = \sqrt{d_{il_1}} + \sqrt{d_{jl_1}}$ and $\sqrt{d_{ij}} = \sqrt{d_{il_2}} + \sqrt{d_{jl_2}}$ and ... and $\sqrt{d_{ij}} = \sqrt{d_{il_{k-2}}} + \sqrt{d_{jl_{k-2}}}$", i, j, l_1, $l_2, \ldots, l_{k-3}$ and l_{k-2} take another value each other.

In this paper, we consider the proof of the case of $k = 4$ in Conjecture 2. That is, $\boldsymbol{V}_1$ is a matrix with one of the following six conditions;

(i) $\sqrt{d_{12}} = \sqrt{d_{13}} + \sqrt{d_{23}}$ and $\sqrt{d_{12}} = \sqrt{d_{14}} + \sqrt{d_{24}}$

(ii) $\sqrt{d_{13}} = \sqrt{d_{12}} + \sqrt{d_{23}}$ and $\sqrt{d_{13}} = \sqrt{d_{14}} + \sqrt{d_{34}}$

(iii) $\sqrt{d_{14}} = \sqrt{d_{12}} + \sqrt{d_{24}}$ and $\sqrt{d_{14}} = \sqrt{d_{13}} + \sqrt{d_{34}}$

(iv) $\sqrt{d_{23}} = \sqrt{d_{12}} + \sqrt{d_{13}}$ and $\sqrt{d_{23}} = \sqrt{d_{24}} + \sqrt{d_{34}}$

(v) $\sqrt{d_{24}} = \sqrt{d_{12}} + \sqrt{d_{14}}$ and $\sqrt{d_{24}} = \sqrt{d_{23}} + \sqrt{d_{34}}$

(vi) $\sqrt{d_{34}} = \sqrt{d_{13}} + \sqrt{d_{14}}$ and $\sqrt{d_{34}} = \sqrt{d_{23}} + \sqrt{d_{24}}$

The proof is as follows. Let $\boldsymbol{A}$ be $k \times k$ nonsingular matrix such that $\boldsymbol{V} = \boldsymbol{A}'\boldsymbol{A}$. Then, by the transformation $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{A}^{-1}$, $\mathrm{vec}(\boldsymbol{Y}) \sim \mathrm{N}_{kp}(\boldsymbol{0}, \boldsymbol{I}_k \otimes \boldsymbol{I}_p)$. Further, let that

$$\Gamma = \{\boldsymbol{\gamma}\ ;\ \boldsymbol{\gamma} = (\boldsymbol{c}'\boldsymbol{V}\boldsymbol{c})^{-1/2}\boldsymbol{A}\boldsymbol{c},\ \forall \boldsymbol{c} \in \mathbb{C}\}.$$

Then, we can write the coverage probability $Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C})$ as

$$\begin{aligned} Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C}) &= \Pr\{(\boldsymbol{Y}\boldsymbol{A}\boldsymbol{c})'(\nu\boldsymbol{S})^{-1}(\boldsymbol{Y}\boldsymbol{A}\boldsymbol{c}) \le t_{\mathrm{p}}^*(\boldsymbol{c}'\boldsymbol{V}\boldsymbol{c}),\ \forall \boldsymbol{c} \in \mathbb{C}\} \\ &= \Pr\{(\boldsymbol{Y}\boldsymbol{\gamma})'(\nu\boldsymbol{S})^{-1}(\boldsymbol{Y}\boldsymbol{\gamma}) \le t_{\mathrm{p}}^*,\ \boldsymbol{\gamma} \in \boldsymbol{\Gamma}\}. \end{aligned}$$

Further, we can write $\nu\boldsymbol{S} = \boldsymbol{H}_1\boldsymbol{L}\boldsymbol{H}_1'$ such that $\boldsymbol{L} = \mathrm{diag}(l_1, \ldots, l_{\mathrm{p}}),\ l_1 \ge \ldots \ge l_{\mathrm{p}}$. Then $\boldsymbol{H}_1$ is a $p \times p$ orthogonal matrix and $\boldsymbol{L}$ and $\boldsymbol{H}_1$ are independent. Then

$$Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C}) = \mathrm{E}_{\boldsymbol{L}}[\Pr\{(\boldsymbol{Y}\boldsymbol{\gamma})'\boldsymbol{L}^{-1}(\boldsymbol{Y}\boldsymbol{\gamma}) \le t_{\mathrm{p}}^*,\ \boldsymbol{\gamma} \in \boldsymbol{\Gamma}\}].$$

Since the dimension of the space spanned by $\mathbb{C}$ equals three when $k = 4$, there exists a $k \times k$ orthogonal matrix $\boldsymbol{H}_2$ such that

$$\boldsymbol{\gamma}_m'\boldsymbol{H}_2 = [\boldsymbol{\delta}_m', 0],\ m = 1, \ldots, 6,$$

where $\boldsymbol{\delta}_m (= (\delta_{m1}, \delta_{m2}, \delta_{m3})')$ is a 3-demensinal vector. Here $\boldsymbol{\delta}_m$'s satisfy $\boldsymbol{\delta}_m'\boldsymbol{\delta}_m = 1$. Then we can write

$$\boldsymbol{\delta}_m = \begin{pmatrix} \sin\beta_{m1}\sin\beta_{m2} \\ \sin\beta_{m1}\cos\beta_{m2} \\ \cos\beta_{m1} \end{pmatrix},\ m = 1, \ldots, 6,$$

where $0 \le \beta_{m1} < \pi$ and $0 \le \beta_{m2} < 2\pi$.

Further, we can write $\boldsymbol{Y}\boldsymbol{H}_2 = [\boldsymbol{U}, \widetilde{\boldsymbol{U}}]$, where $\boldsymbol{U}$ is $p \times 3$. Letting $\boldsymbol{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p]'$ where

$$\boldsymbol{u}_s = ||\boldsymbol{u}_s|| \begin{pmatrix} \sin\theta_{s1}\sin\theta_{s2} \\ \sin\theta_{s1}\cos\theta_{s2} \\ \cos\theta_{s1} \end{pmatrix} = r_s \begin{pmatrix} \sin\theta_{s1}\sin\theta_{s2} \\ \sin\theta_{s1}\cos\theta_{s2} \\ \cos\theta_{s1} \end{pmatrix},\ s = 1, \ldots, p,$$

and r_s^2, θ_{s1} and θ_{s2} are independently distributed as χ^2 distribution with three degrees of freedom, uniform distribution on U[0, π) and uniform distribution on U[0, 2π), respectively. Then the coverage probability can be written as

$$\begin{aligned} Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C}) = \ & \mathrm{E}_{\boldsymbol{L},\boldsymbol{R}}\Bigg[\Pr\Bigg\{\sum_{s=1}^{p}\frac{r_s^2}{l_s}(\sin\theta_{s1}\sin\theta_{s2}\sin\beta_{m1}\sin\beta_{m2} \\ & +\sin\theta_{s1}\cos\theta_{s2}\sin\beta_{m1}\cos\beta_{m2} + \cos\theta_{s1}\cos\beta_{m1})^2 \le t_{\mathrm{p}}^* \\ & \text{for } m = 1, \ldots, 6\Bigg\}\Bigg], \end{aligned}$$

where $\boldsymbol{R} = \mathrm{diag}(r_1, \ldots, r_p)$ is independent of $\boldsymbol{L} = \mathrm{diag}(l_1, \ldots, l_p)$.

Relating the coverage probability $Q(t^*_{\mathrm{p}}, \boldsymbol{V}, \mathbb{C})$, we consider the probability

$$\begin{aligned} G(\boldsymbol{\beta}) = \ &\Pr \Big[\sum_{s=1}^{p} \frac{r_s^2}{l_s} (\sin\theta_{s1} \sin\theta_{s2} \sin\beta_{m1} \sin\beta_{m2} + \sin\theta_{s1} \cos\theta_{s2} \sin\beta_{m1} \cos\beta_{m2} \\ &\qquad + \cos\theta_{s1} \cos\beta_{m1})^2 \le t^*_{\mathrm{p}} \text{ for } m = 1, \ldots, 6 \Big], \end{aligned} \tag{6}$$

where $\boldsymbol{\beta} = (\beta_{11}, \beta_{21}, \beta_{31}, \beta_{41}, \beta_{51}, \beta_{61}, \beta_{12}, \beta_{22}, \beta_{32}, \beta_{42}, \beta_{52}, \beta_{62})'$.

Also, we define the volumes Ω and D_m, $m = 1, \ldots, 6$, as follows.

$$\begin{aligned} \Omega = \ &\{(\theta_{s1}, \theta_{s2})^p : 0 < \theta_{s1} < \pi, 0 < \theta_{s2} < 2\pi, 1 \le s \le p\}, \\ D_m = \ &\Big\{ (\theta_{s1}, \theta_{s2})^p \in \Omega : \sum_{s=1}^{p} \frac{r_s^2}{l_s} (\sin\theta_{s1} \sin\theta_{s2} \sin\beta_{m1} \sin\beta_{m2} \\ &\qquad + \sin\theta_{s1} \cos\theta_{s2} \sin\beta_{m1} \cos\beta_{m2} + \cos\theta_{s1} \cos\beta_{m1})^2 > t^*_{\mathrm{p}} \Big\}. \end{aligned}$$

Then, we note that the probability (6) is equal to $1 - \mathrm{volume}[\cup_{m=1}^{6} D_m]/(2\pi^2)^p$.

Therefore, to minimize $G(\boldsymbol{\beta})$ is equivalent to maximizing the value for volume of the union of D_m's. Similarly, to maximize $G(\boldsymbol{\beta})$ is equivalent to minimizing the value for volume of the union of D_m's.

Next, in the case of $k = 4$ for pairwise comparisons, we can assume that subset $\boldsymbol{c}$'s of the set $\mathbb{C}$ are as follows.

$$\boldsymbol{c_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \boldsymbol{c_2} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \boldsymbol{c_3} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix},$$

$$\boldsymbol{c_4} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \boldsymbol{c_5} = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \boldsymbol{c_6} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Letting $\boldsymbol{\gamma}$ corresponding with $\boldsymbol{c}_a$ as $\boldsymbol{\gamma}_a$ $(a = 1, \ldots, {}_kC_2)$, then we have

$$\boldsymbol{\gamma}_a' \boldsymbol{\gamma}_b = \frac{\boldsymbol{c}_a' \boldsymbol{V} \boldsymbol{c}_b}{\sqrt{\boldsymbol{c}_a' \boldsymbol{V} \boldsymbol{c}_a} \sqrt{\boldsymbol{c}_b' \boldsymbol{V} \boldsymbol{c}_b}}, \tag{7}$$

and $\boldsymbol{\gamma}_a' \boldsymbol{\gamma}_b = \boldsymbol{\delta}_a' \boldsymbol{\delta}_b$.

Here we note that $G(\boldsymbol{\beta})$ and D_m, $m = 1, \ldots, 6$, can be written as

$$\begin{aligned} &G(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\delta}_3, \boldsymbol{\delta}_4, \boldsymbol{\delta}_5, \boldsymbol{\delta}_6) \\ &\quad = \Pr \Big[\sum_{s=1}^{p} \frac{r_s^2}{l_s} (\delta_{m1} \sin\theta_{s1} \sin\theta_{s2} + \delta_{m2} \sin\theta_{s1} \cos\theta_{s2} + \delta_{m3} \cos\theta_{s1})^2 \le t^*_{\mathrm{p}} \\ &\qquad\qquad \text{for } m = 1, \ldots, 6 \Big], \end{aligned}$$

and

$$D_m = \left\{ (\theta_{s1}, \theta_{s2})^p \in \Omega : \sum_{s=1}^{p} \frac{r_s^2}{l_s} (\delta_{m1} \sin\theta_{s1} \sin\theta_{s2} + \delta_{m2} \sin\theta_{s1} \cos\theta_{s2} + \delta_{m3} \cos\theta_{s1})^2 > t_{\mathrm{p}}^* \right\}.$$

Assuming that $d_{ij} = d$, we can put $\boldsymbol{\delta}_m$'s as

$$\boldsymbol{\delta_1} = \begin{pmatrix} \frac{1}{2\sqrt{3}} \\ -\frac{1}{2} \\ -\frac{2}{\sqrt{6}} \end{pmatrix}, \boldsymbol{\delta_2} = \begin{pmatrix} \frac{1}{2\sqrt{3}} \\ \frac{1}{2} \\ -\frac{2}{\sqrt{6}} \end{pmatrix}, \boldsymbol{\delta_3} = \begin{pmatrix} -\frac{1}{\sqrt{3}} \\ 0 \\ -\frac{2}{\sqrt{6}} \end{pmatrix},$$

$$\boldsymbol{\delta_4} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix}, \boldsymbol{\delta_5} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}, \boldsymbol{\delta_6} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix},$$

where $\boldsymbol{\delta}_m$'s satisfy $\boldsymbol{\delta}_1'\boldsymbol{\delta}_2 = \frac{1}{2}$, $\boldsymbol{\delta}_1'\boldsymbol{\delta}_3 = \frac{1}{2}$, $\boldsymbol{\delta}_1'\boldsymbol{\delta}_4 = \frac{1}{2}$, $\boldsymbol{\delta}_1'\boldsymbol{\delta}_5 = 0$, $\boldsymbol{\delta}_1'\boldsymbol{\delta}_6 = \frac{1}{2}$, $\boldsymbol{\delta}_2'\boldsymbol{\delta}_3 = \frac{1}{2}$, $\boldsymbol{\delta}_2'\boldsymbol{\delta}_4 = 0$, $\boldsymbol{\delta}_2'\boldsymbol{\delta}_5 = \frac{1}{2}$, $\boldsymbol{\delta}_2'\boldsymbol{\delta}_6 = -\frac{1}{2}$, $\boldsymbol{\delta}_3'\boldsymbol{\delta}_4 = -\frac{1}{2}$, $\boldsymbol{\delta}_3'\boldsymbol{\delta}_5 = -\frac{1}{2}$, $\boldsymbol{\delta}_3'\boldsymbol{\delta}_6 = 0$, $\boldsymbol{\delta}_4'\boldsymbol{\delta}_5 = \frac{1}{2}$, $\boldsymbol{\delta}_4'\boldsymbol{\delta}_6 = \frac{1}{2}$ and $\boldsymbol{\delta}_5'\boldsymbol{\delta}_6 = -\frac{1}{2}$.

For example, putting $r_1^2/l_1 = 1$ and $t_{\mathrm{p}}^* = 0.5$ in the case of $p = 1$, we have

$$\begin{aligned} &G(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\delta}_3, \boldsymbol{\delta}_4, \boldsymbol{\delta}_5, \boldsymbol{\delta}_6) \\ &= \Pr\left[(\sin\theta_{11} \sin\theta_{12}\delta_{m1} + \sin\theta_{11} \cos\theta_{12}\delta_{m2} + \cos\theta_{11}\delta_{m3})^2 \le 0.5 \right. \\ &\qquad \left. \text{for } m = 1, \ldots, 6 \right], \end{aligned}$$

and

$$D_m = \left\{ (\theta_{11}, \theta_{12}) \in \Omega : (\sin\theta_{11} \sin\theta_{12}\delta_{m1} + \sin\theta_{11} \cos\theta_{12}\delta_{m2} + \cos\theta_{11}\delta_{m3})^2 > 0.5 \right\}.$$

It is noted from Figures 1 $\sim$ 6 that the area of $\cup_{m=1}^{6} D_m$ is equal to Ω when $d_{ij} = d$. Hence, the area of $\cup_{m=1}^{6} D_m$ is maximum when $d_{ij} = d$. It is may be noted that the volume$[\cup_{m=1}^{6} D_m]$ for $p \ge 2$ is maximum when $d_{ij} = d$. Therefore, we follow that $Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C})$ is minimum when $d_{ij} = d$. Further, by using Theorem 2.1 and Corollary 2.2 in Seo, Mano and Fujikoshi (1994), we have that $Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C})$ is minimum when $\boldsymbol{V} = \boldsymbol{I}$.

Secondly, we consider the case which volume$[\cup_{m=1}^{6} D_m]$ is minimum. By the same way, we note that $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$, $\boldsymbol{\delta}_3$, $\boldsymbol{\delta}_4$, $\boldsymbol{\delta}_5$ and $\boldsymbol{\delta}_6$ are all same when volume$[\cup_{m=1}^{6} D_m]$ is minimum. Therefore, $\boldsymbol{\delta}_a'\boldsymbol{\delta}_b = \boldsymbol{\delta}_a'\boldsymbol{\delta}_a = 1 (a \ne b)$, that is, $\boldsymbol{\gamma}_a'\boldsymbol{\gamma}_b = 1$. Hence we can get the condition of $\boldsymbol{V}_1$ as "$\sqrt{d_{14}} = \sqrt{d_{13}} + \sqrt{d_{34}}$ and $\sqrt{d_{14}} = \sqrt{d_{12}} + \sqrt{d_{24}}$". We note that there exists the positive semi-definite matrix $\boldsymbol{V}_1$ such that "$\sqrt{d_{14}} = \sqrt{d_{13}} + \sqrt{d_{34}}$ and $\sqrt{d_{14}} = \sqrt{d_{12}} + \sqrt{d_{24}}$".

We have the following theorem.

Theorem 3. *Let $Q(t, \boldsymbol{V}, \mathbb{B})$ be the coverage probability for (5) with a known matrix $\boldsymbol{V}$. Then, for any positive definite matrix $\boldsymbol{V}$ in the case of $k = 4$, it holds that*

$$1 - \alpha = Q(t_{\mathrm{p}}^*, \boldsymbol{I}, \mathbb{C}) \leq Q(t_{\mathrm{p}}^*, \boldsymbol{V}, \mathbb{C}) < Q(t_{\mathrm{p}}^*, \boldsymbol{V}_1, \mathbb{C}),$$

where $t_{\mathrm{p}}^* = t_{\mathrm{p}}^2/\nu$, $\boldsymbol{V}_1$ satisfies with one of the conditions "$\sqrt{d_{ij}} = \sqrt{d_{il}} + \sqrt{d_{jl}}$ and $\sqrt{d_{ij}} = \sqrt{d_{im}} + \sqrt{d_{jm}}$" and i, j, l, m take another value each other.

We note that there dose not exist a positive definite matrix such that "$\sqrt{d_{12}} = \sqrt{d_{13}} + \sqrt{d_{23}}$ and $\sqrt{d_{12}} = \sqrt{d_{14}} + \sqrt{d_{24}}$" or "$\sqrt{d_{13}} = \sqrt{d_{12}} + \sqrt{d_{23}}$ and $\sqrt{d_{13}} = \sqrt{d_{14}} + \sqrt{d_{34}}$" or "$\sqrt{d_{14}} = \sqrt{d_{12}} + \sqrt{d_{24}}$ and $\sqrt{d_{14}} = \sqrt{d_{13}} + \sqrt{d_{34}}$" or "$\sqrt{d_{23}} = \sqrt{d_{12}} + \sqrt{d_{13}}$ and $\sqrt{d_{23}} = \sqrt{d_{24}} + \sqrt{d_{34}}$" or "$\sqrt{d_{24}} = \sqrt{d_{12}} + \sqrt{d_{14}}$ and $\sqrt{d_{24}} = \sqrt{d_{23}} + \sqrt{d_{34}}$" or "$\sqrt{d_{34}} = \sqrt{d_{13}} + \sqrt{d_{14}}$ and $\sqrt{d_{34}} = \sqrt{d_{23}} + \sqrt{d_{24}}$". However, there exists $\boldsymbol{V}_1$ as a positive semi-definite matrix. For example, in the case $k = 4$, the one of such matrix $\boldsymbol{V}_1$ is given by

$$\boldsymbol{V}_1 = \begin{bmatrix} 3 & 0 & 1 & 2 \\ 0 & 6 & 4 & 2 \\ 1 & 4 & 3 & 2 \\ 2 & 2 & 2 & 2 \end{bmatrix}.$$

3. Numerical examinations

This section gives some numerical results of the coverage probability for $T_{\max\cdot\mathrm{p}}^2$ statistic and the upper percentiles of the statistic by Monte Carlo simulation. The Monte Carlo simulations are made from 10 replications of 1,000,000 simulations for each of parameters based on normal random vectors based on $N_{kp}(\boldsymbol{0},\ \boldsymbol{V} \otimes \boldsymbol{I}_p)$. The sample covariance matrix $\boldsymbol{S}$ is computed on the basis of random vectors from $N_p(\boldsymbol{0},\ \boldsymbol{I}_p)$. Also, we note that $\boldsymbol{S}$ is formed independently in each time with ν degrees of freedom. The average of 10 replications based on 1,000,000 simulations is used as the simulated value of the statistic.

Table 1 gives the upper percentiles $t_{\mathrm{p}\cdot V}$ of $T_{\max\cdot\mathrm{p}}$ $(= \sqrt{T_{\max\cdot\mathrm{p}}^2})$ and the upper bounds of the coverage probability for the following parameters: $\alpha = 0.1, 0.05, 0.01$, $p = 1, 2, 5$, $k = 4$, $\nu = 20, 40, 60$, and $\boldsymbol{V} = \boldsymbol{I}$, $\boldsymbol{V}_1$, $\boldsymbol{V}_2$ and $\boldsymbol{V}_3$, that is,

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{V}_1 = \begin{pmatrix} 3 & 0 & 1 & 2 \\ 0 & 6 & 4 & 2 \\ 1 & 4 & 3 & 2 \\ 2 & 2 & 2 & 2 \end{pmatrix},$$

$$\boldsymbol{V}_2 = \begin{pmatrix} 3 & 2 & 2 & 1 \\ 2 & 3 & 2 & 1 \\ 2 & 2 & 3 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \quad \boldsymbol{V}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.05 \end{pmatrix}.$$

Here we note that $\boldsymbol{I}$ and $\boldsymbol{V}_2$ are positive definite matrices such that $d_{ij} = d$, $\boldsymbol{V}_1$ is a positive semi-definite matrix such that $\sqrt{d_{12}} = \sqrt{d_{13}} + \sqrt{d_{23}}$ and $\sqrt{d_{12}} = \sqrt{d_{14}} + \sqrt{d_{24}}$, and $\boldsymbol{V}_3$ is a positive definite matrix such that $d_{12} = 1.5$, $d_{13} = 1.1$, $d_{14} = 1.05$, $d_{23} = 0.6$, $d_{24} = 0.55$ and $d_{34} = 0.15$ $(d_{ij} \neq d)$.

It can be seen from simulation results in Table 1 that the upper percentiles with $\boldsymbol{V} = \boldsymbol{I}$ are always largest values and those with $\boldsymbol{V} = \boldsymbol{V}_1$ are the lower limit values in any positive definite matrix $\boldsymbol{V}$ for each parameter. Also, the upper percentiles with $\boldsymbol{V} = \boldsymbol{V}_3$ are always between those with $\boldsymbol{V} = \boldsymbol{I}$ and those with $\boldsymbol{V} = \boldsymbol{V}_1$. Since $\boldsymbol{I}$ and $\boldsymbol{V}_2$ are positive definite matrices such that $d_{ij} = d$, it may be confirmed from simulation results that the upper percentiles with $\boldsymbol{V} = \boldsymbol{V}_2$ are same as those with $\boldsymbol{V} = \boldsymbol{I}$.

It is noted from Table 1 that the upper bounds for the conservativeness of multiple pairwise comparisons can be obtained. For example, when $p = 2$, $\nu = 20$ and $\alpha = 0.1$, we note that $0.900 \leq Q(t^*_{\mathrm{p}}, \boldsymbol{V}, \mathbb{C}) < 0.977$ for any positive definite $\boldsymbol{V}$. It may be noted that the coverage probabilities do not depend on p. Also, it may be noted that the coverage probabilities are large as ν is large.

Acknowledgements

The authors wish to thank Ms. Yoshimi Iibuchi for the Monte Carlo simulation in numerical examinations. The research of the first author was supported in part by Grant-in-Aid for Scientific Research (C) under Contract Number 18500218.

References

Benjamini, Y. and Braun, H., 2002. John W. Tukey's contributions to multiple comparisons. *Ann. Statist.* 30, 1576–1594.

Brown, L. D., 1984. A note on the Tukey-Kramer procedure for pairwise comparisons of correlated means, in Design of Experiments: Ranking and Selection (Essays in Honor of Robert E. Bechhofer), eds. T. J. Santner and A. C. Tamhane. New York: Marcel Dekker, 1-6.

Cramér, H., 1946. Mathematical methods of statistics. Prinston University Press. Prinston, New Jersey, pp. 297.

Hayter, A. J., 1984. A proof of the conjecture that the Tukey-Kramer multiple comparisons procedure is conservative. *Ann. Statist.* 12, 61–75.

Hayter, A. J., 1989. Pairwise comparisons of generally correlated means. *J. Amer. Statist. Assoc.* 84, 208–213.

Hochberg, Y. and Tamhane, A. C., 1987. Multiple comparison Procedures. Wiley, New York.

Kramer, C. Y., 1956. Extension of multiple range tests to group means with unequal number of replications. *Biometrics.* 12, 307–310.

Kramer, C. Y., 1957. Extension of multiple range tests to group correlated adjusted means. *Biometrics.* 13, 13–18.

Krishnaiah, P. R., 1979. Some developments on simultaneous test procedures. Developments in Statistics, Vol. 2 (Krishnaiah, P. R. ed.) North-Holland, Amsterdam 157–201.

Lin, F. J., Seppänen, E. and Uusipaikka, E., 1990. On pairwise comparisons of the components of the mean vector in multivariate normal distribution. *Comm. Statist. Theo. Meth.* 19, 395–411.

Seo, T., 1995. Simultaneous confidence procedures for multiple comparisons of mean vectors in multivariate normal populations. *Hiroshima Math. J.* 25, 387–422.

Seo, T., 1996. A note on the conservative multivariate Tukey-Kramer multiple comparison procedure. *Amer. J. Math. Manage. Sci.* 16, 251–266.

Seo, T., Mano, S. and Fujikoshi, Y., 1994. A generalized Tukey conjecture for multiple comparisons among mean vectors. *J. Amer. Statist. Assoc.* 89, 676–679.

Seo, T. and Nishiyama, T., 2006. On the conservative multivariate Tukey-Kramer type procedures for multiple comparisons among mean vectors. *submitted for publication.*

Seo, T. and Siotani, M., 1992. The multivariate Studentized range and its upper percentiles. *J. Japan Statist. Soc.* 22, 123–137.

Siotani, M. 1959a. On the extreme value of the generalized distances of the individual points in the multivariate normal sample. *Ann. Inst. Statist. Math.* 10, 183–208.

Siotani, M. 1959b. On the range in the multivariate case. *Proc. Inst. Statist. Math.* 6, 155–165.

Siotani, M. 1960. Note on multivariate confidence bounds. *Ann. Inst. Statist. Math.* 11, 167–182.

Siotani, M., Hayakawa, T. and Fujikoshi, Y., 1985. Modern Multivariate Analysis : A Graduate Course and Handbook. American Sciences Press, Ohio.

Somerville, P. N., 1993. On the conservatism of the Tukey-Kramer multiple comparison procedure. *Statist. Prob. Lett.* 16, 343–345.

Spurrier, J. D. and Isham, S. P., 1985. Exact simultaneous confidence intervals for pairwise comparisons of three normal means. *J. Amer. Statist. Assoc.* 80, 438–442.

Tukey, J. W., 1953. The problem of multiple comparisons. Unpublished manuscript. Princeton University.

Uusipaikka, E., 1985. Exact simultaneous confidence intervals for multiple comparisons among three or four mean values. *J. Amer. Statist. Assoc.* 80, 196–201.

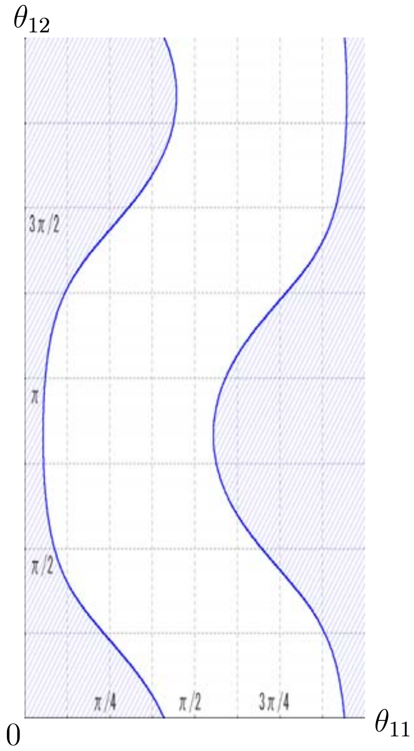

Figure 1. D_1 when $d_{ij} = d$

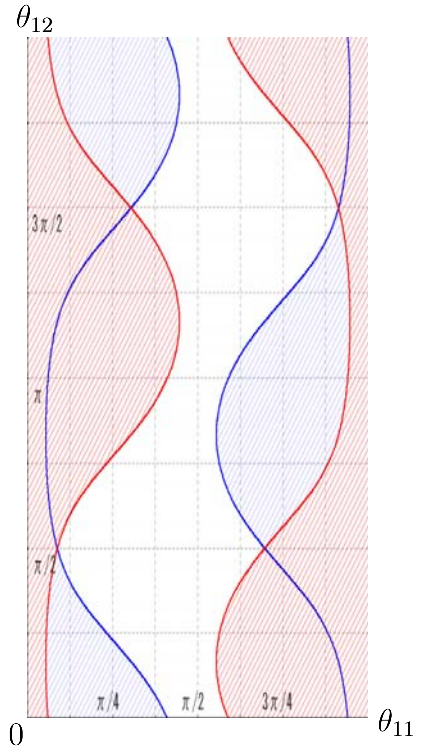

Figure 2. $D_1 \cup D_2$ when $d_{ij} = d$

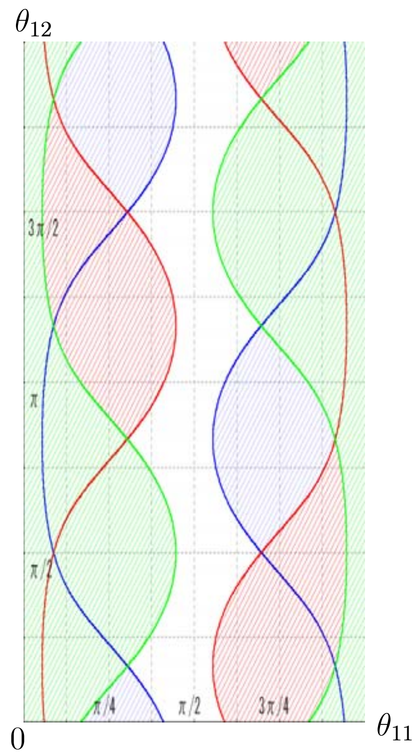

Figure 3. $D_1 \cup D_2 \cup D_3$ when $d_{ij} = d$

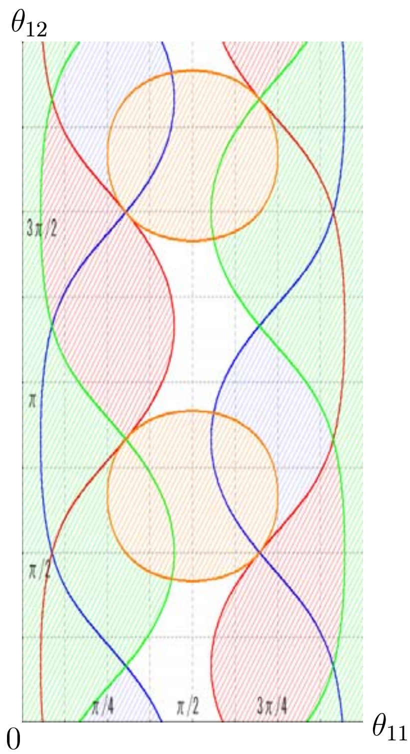

Figure 4. $D_1 \cup D_2 \cup D_3 \cup D_4$ when $d_{ij} = d$

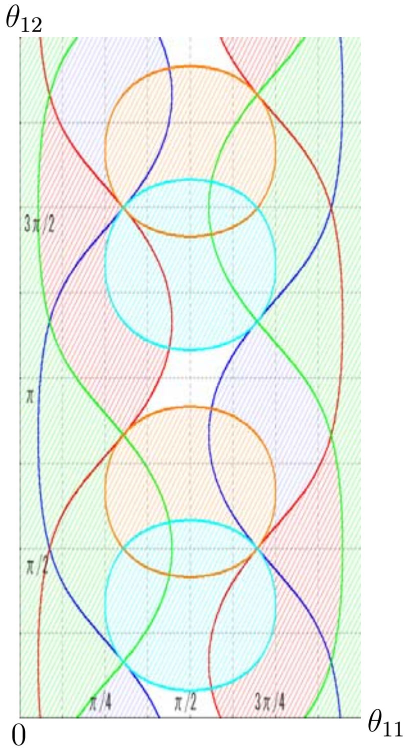

Figure 5. $D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5$ when $d_{ij} = d$

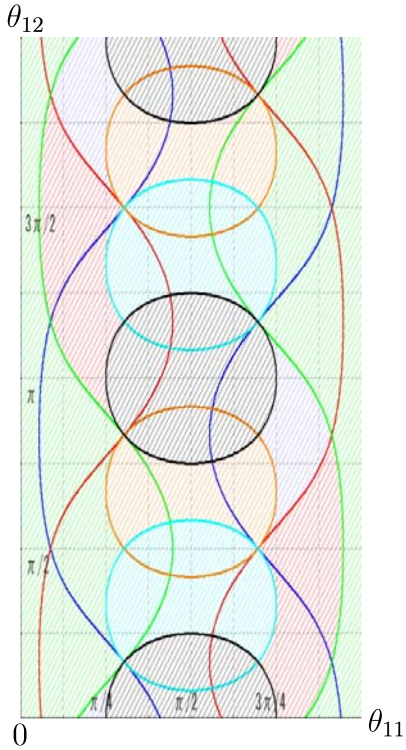

Figure 6. $D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5 \cup D_6$ when $d_{ij} = d$

Table 1: Simulation results for pairwise comparison

p	ν	α	$t_{\mathrm{p}\cdot I}$	$t_{\mathrm{p}\cdot V_2}$	$t_{\mathrm{p}\cdot V_3}$	$t_{\mathrm{p}\cdot V_1}$	$Q(t_{\mathrm{p}}^*, \boldsymbol{V}_1, \mathbb{C})$
1	20	0.01	3.548	3.547	3.509	2.844	0.998
		0.05	2.799	2.798	2.762	2.085	0.989
		0.1	2.448	2.448	2.412	1.724	0.976
	40	0.01	3.320	3.320	3.285	2.706	0.998
		0.05	2.681	2.681	2.646	2.022	0.989
		0.1	2.368	2.367	2.333	1.683	0.977
	60	0.01	3.250	3.250	3.218	2.661	0.998
		0.05	2.642	2.642	2.608	2.000	0.989
		0.1	2.342	2.342	2.308	1.670	0.977
2	20	0.01	4.301	4.302	4.259	3.534	0.998
		0.05	3.491	3.491	3.451	2.723	0.989
		0.1	3.118	3.118	3.079	2.343	0.977
	40	0.01	3.896	3.896	3.861	3.262	0.998
		0.05	3.251	3.250	3.214	2.576	0.990
		0.1	2.938	2.938	2.902	2.238	0.978
	60	0.01	3.774	3.773	3.740	3.182	0.998
		0.05	3.177	3.177	3.142	2.532	0.990
		0.1	2.883	2.883	2.848	2.207	0.978
5	20	0.01	6.293	6.293	6.232	5.266	0.998
		0.05	5.214	5.214	5.160	4.223	0.989
		0.1	4.731	4.731	4.679	3.746	0.977
	40	0.01	5.155	5.154	5.115	4.456	0.998
		0.05	4.450	4.449	4.409	3.709	0.990
		0.1	4.112	4.112	4.071	3.345	0.978
	60	0.01	4.866	4.866	4.832	4.244	0.998
		0.05	4.246	4.247	4.209	3.569	0.990
		0.1	3.944	3.944	3.906	3.234	0.979