**Noname manuscript No.**
(will be inserted by the editor)

# Multivariate small sample tests for two-way designs with applications to industrial statistics

**Rosa Arboretti · Riccardo Ceccato · Livio Corain · Fabrizio Ronchi · Luigi Salmaso**

Received: date / Accepted: date

**Abstract** In this paper, we present a novel nonparametric approach for multivariate analysis of two-way crossed factorial design based on NonParametric Combination applied to Synchronized Permutation tests. This nonparametric hypothesis testing procedure not only allows to overcome the shortcomings of MANOVA test like violation of assumptions such as multivariate normality or covariance homogeneity, but, in an extensive simulation study, reveals to be a powerful instrument both in case of small sample size and many response variables. We contextualize its application in the field of industrial experiments and we assume a linear additive model for the data set analysis. Indeed, the linear additive model interpretation well adapts to the industrial production environment because of the way control of production machineries is implemented. The case of small sample size reflects the frequent needs of practitioners in the industrial environment where there are constraints or limited resources for the experimental design. Furthermore, an increase in rejection rate can be observed under alternative hypothesis when the number of response variables increases with fixed number of observed units. This could lead to a strategical benefit considering that in many real problems it could be easier to collect more information on a single experimental unit than adding a new unit to the experimental design. An application to industrial thermoforming processes is useful to illustrate and highlight the benefits of the adoption of the herein presented nonparametric approach.

**Keywords** Synchronized Permutation · Non-parametric tests · Combining function · NPC tests · Multivariate tests · MANOVA

Arboretti Rosa
Civil, Environmental and Architectural Engineering
University of Padova - Padova - Italy
E-mail: rosa.arboretti@unipd.it

Ceccato Riccardo · Corain Livio · Ronchi Fabrizio · Salmaso Luigi
Department of Management and Engineering
University of Padova - Vicenza - Italy

## 1 Introduction

Industrial experiments are commonly based on factorial designs. Design of Experiment (DOE) is popular in different fields of engineering as for instance for bio-fuel production [36], for industrial production practices [17,35], for machines' production process [19], for alternative raw materials experimentation [18], or for the use of Coordinate Measuring Machines for quality control [34]. The two-way crossed factorial design is a common design used in the exploratory phase. Thanks to two-way multi-level design, practitioners can assess the impact on response variables of two factors they can control during the experiment and of their interaction according to the assumed model.

In many industrial applications (and applied research fields) it is common the need to compare multivariate population obtained in advanced factorial designs. There are manufacturing processes where treatments or control factors in production processes impact on several relevant variables simultaneously [32,24,7,25]. In these cases an overall test is useful to determine for instance whether there is a significant difference on final product or not. Observed data are usually analized using the multivariate analysis of variance (MANOVA) methods. Unfortunately parametric methods rely on assumptions such as multivariate normality and covariance homogeneity, but these prerequisites may be not realistic for several real problems.

How to overcome the violation of MANOVA assumptions has been investigated and nonparametric methods for multivariate inferential tests have been developed. One of the approaches is based on the generalization to the multivariate case [26,6] of the univariate comparison between the group-wise distribution functions $F_i$ and the reference distribution function $H = \frac{1}{N}\sum_i n_i F_i$ that is the pooled distribution [20,21]. In these cases the null hypotheses are formulated in terms of distribution functions. Other approaches are rank-based as for instance in [15,4,5,13,14] but they are for large number of factor levels or for large sample size.

We propose a novel nonparametric approach based on NonParametric Combination (NPC) [31] applied to Synchronized Permutation (SP) tests [3,2] for two-way crossed factorial design assuming a linear additive model. Indeed, the linear additive model interpretation well adapts to the industrial production environment because of the way control of production machineries is implemented. This approach overcomes the shortcomings of MANOVA with the only mild condition of the data set to be analyzed taking values on a multi-dimensional distribution belonging to a nonparametric family of non-degenerate probability distributions. It well works with even only two levels per factor and a small sample size. The case of small sample size reflects the frequent needs of practitioners in the industrial environment where there are constraints or limited resources for the experimental design. Furthermore it allows to formulate test hypotheses in more familiar terms for practitioners such as factor effect size. Indeed, we agree with Lakens [22] that "Effect sizes are the most important outcome of empirical studies. Most articles on effect sizes highlight their importance to communicate the practical significance of results".

A simulation design with fixed factor effects $\delta$ and fixed variance $\sigma$ of data set distributions have been performed in order to evaluate the rejection rate of the NPC applied to SP tests under alternative Hypothesis $H_1$ in the range of interest of significance levels $0 \leq \alpha \leq 0.1$, and in order to compare it with the classical MANOVA test.

A real case study is useful to highlight the benefits of the adoption of the herein presented nonparametric approach in industrial experiments with a small sample size and non-normal data distribution. A two-way two-level design is used to understand whether two control factors and their interaction were significant or not in a project of innovation of the production system of a thermoformed packaging.

The remainder of the paper is organized as follows. We describe Synchronized Permutation methods (Section 2) and NonParametric Combination (Section 3) in a two-way factorial design and the two steps algorithm to apply NPC to SP. Next, the simulations design is described (Section 4) and results are presented (Section 5). Then, a real case study on industrial experiment is presented (Section 6) and, finally, we discuss the results of simulation and make further comments (Section 7).

## 2 Synchronized Permutation methods

The synchronized permutation methods is herein illustrated. We assume the linear additive model of a balanced two-way factorial crossed design:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}, \quad i = 1, \dots, I; \quad j = 1, \dots, J; \quad k = 1, \dots, n, \quad (1)$$

where $\mu$ is the overall mean, $\alpha_i$ is the effect of level $i$ of factor $A$, $\beta_j$ is the effect of level $j$ of factor $B$, $(\alpha\beta)_{ij}$ is the effect of interaction between factor $A$ at level $i$ and factor $B$ at level $j$, $I$ and $J$ are the number of levels of factor $A$ and $B$ respectively, and $\epsilon_{ijk}$ is the error term. The number of replicates of the balanced design is $n$ and the mean of error term is $E(\epsilon_{ijk}) = 0$ for each factor level combination. The total number of observations is $N = \sum_i \sum_j n = I \cdot J \cdot n$

The side conditions are:

$$\sum_i \alpha_i = 0; \quad \sum_j \beta_j = 0; \quad \sum_i (\alpha\beta)_{ij} = 0 \; \forall j; \quad \sum_j (\alpha\beta)_{ij} = 0 \; \forall i. \quad (2)$$

The null hypotheses of no-main effect of factor $A$, no-main effect of factor $B$ and no-interaction effect between factors $A$ and $B$ are:

$$\begin{aligned} H_0^{(A)} &: \quad \alpha_i = 0 \quad \forall i, \\ H_0^{(B)} &: \quad \beta_j = 0 \quad \forall j \quad \text{and} \\ H_0^{(AB)} &: \quad (\alpha\beta)_{ij} = 0 \quad \forall i, j, \end{aligned} \quad (3)$$

respectively. In vector notation, the data can be written as $\boldsymbol{Y} = (Y_{ijk})' = (Y_{111}, \dots, Y_{IJn})'$. And the null hypotheses can be written in terms of contrasts as:

$$\begin{aligned} H_0^{(A)} &: \quad \boldsymbol{C}_A \boldsymbol{\mu} = \boldsymbol{0}, \\ H_0^{(B)} &: \quad \boldsymbol{C}_B \boldsymbol{\mu} = \boldsymbol{0} \quad \text{and} \\ H_0^{(AB)} &: \quad \boldsymbol{C}_{AB} \boldsymbol{\mu} = \boldsymbol{0}, \end{aligned} \quad (4)$$

where $\boldsymbol{\mu} = (\mu_{11}, \ldots, \mu_{IJ})^{'}$, $\mu_{ij} = E(Y_{ijk}) = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ and $\boldsymbol{C}_M$ is a contrast matrix for $M \in \{A, B, AB\}$

Synchronized permutation is developed along two basic concepts. The first is that permutations of observations between two levels of a factor can be made only holding the level of remaining factors in the model constant. For instance, consider the case of a two-way design with factor level combinations $A_1B_1$, $A_1B_2$, $A_2B_1$ and $A_2B_2$. To test the significance of main effect $A$, observations will be exchanged between groups $A_1B_1$ and $A_2B_1$, and between $A_1B_2$ and $A_2B_2$. That is, the level of $B$ is kept constant when performing the test on factor $A$, in the former case such level is 1, in the latter it is 2. The second basic concept in synchronized permutation tests is exchanging the same number of units within each pair of the considered groups [3].

According to Basso et al. [3,2] the test statistics for the main factor $A$ in the two-way design is:

$$\begin{aligned} T_A &= \sum_{i<s} \left[ \sum_j T_{is|j} \right]^2, \quad \text{where} \\ T_{is|j} &= \sum_k Y_{ijk} - \sum_k Y_{sjk}, \quad i, s \in \{1, \ldots, I\}; \quad j \in \{1, \ldots, J\} \end{aligned} \tag{5}$$

The outer sum is made over all possible pairs of levels $1 \leq i \leq s \leq I$ and the inner sum is squared to avoid the cancellation of any of the contributions of effects of factor $A$.

Similarly for factor $B$:

$$\begin{aligned} T_B &= \sum_{j<h} \left[ \sum_i T_{jh|i} \right]^2, \quad \text{where} \\ T_{jh|i} &= \sum_k Y_{ijk} - \sum_k Y_{ihk}, \quad i \in \{1, \ldots, I\}; \quad j, h \in \{1, \ldots, J\} \end{aligned} \tag{6}$$

The statistics for interaction between factor $A$ and factor $B$ is given by the summation of two contributions along the two factors:

$$\begin{aligned} T_{AB} &=^a T_{AB} +^b T_{AB}, \quad \text{where} \\ {}^aT_{AB} &= \sum_{i<s} \sum_{j<h} \left[ T_{is|j} - T_{is|h} \right]^2, \quad \text{and} \\ {}^bT_{AB} &= \sum_{j<h} \sum_{i<s} \left[ T_{jh|i} - T_{jh|s} \right]^2, \quad i, s \in \{1, \ldots, I\}; \quad j, h \in \{1, \ldots, J\} \end{aligned} \tag{7}$$

The statistics for main factors and interaction are uncorrelated.

Then $p$-value is calculated as the proportion of permutations for which test statistics of permuted data set are greater or equal to the test statistic of the original data set.

There are two ways to obtain a synchronized permutation, namely Constrained Synchronized Permutation (CSP) [33] and Unconstrained Synchronized Permutation (USP) [33].

### 2.1 Constrained Synchronized Permutation (CSP)

In the CSP the same permutation is applied in all couples of groups given the initial order of observations. For instance, in the two way design if a permutation consists in exchanging the second observation of group $A_iB_j$ with the first observation of group $A_sB_j$ when testing main effect $A$, then same permutation has to be applied to groups $A_iB_h$ and $A_sB_h$. It is recommended to randomize observations within each group at the beginning before performing the permutation test.

As a result of the application of the same permutation between all possible pairs of groups, the number of possible ways to exchange units depends only on number of replicates n in the balanced design. The total number of possible permutations of CSPs is:

$$C_{csp} = \binom{2n}{n} \tag{8}$$

Thus, according to the way the $p$-value is calculated, the minimum achievable significance error is $\alpha_{min} = 2 \times (C_{csp})^{-1}$. If $n$ is too small, CSP could give a minimum achieved significance level higher than the desired type I error.

### 2.2 Unconstrained Synchronized Permutation (USP)

USP, unlike CSP, can apply different permutations in the various pairs of groups. However, the basic principle of synchronized permutations of exchanging the same number of observations has to be respected. The algorithm provided by Basso et al. [3] guarantees the values of the test statistic to be equally likely. This procedure allows to overcome those cases in which the test statistic is not uniformly distributed.

The total number of possible permutations of USPs depends on a larger number of parameters of dataset respect to CSP. The formula is more complex and there are two cases:

$$\begin{aligned} C^{o}_{USP} &= \sum_{\nu=0}^{(n-1)/2} \binom{n}{\nu}^{J\times I(I-1)} && \text{when n is odd,} \\ C^{e}_{USP} &= \sum_{\nu=0}^{n/2-1} \binom{n}{\nu}^{J\times I(I-1)} + \left[\frac{1}{2}\binom{n}{n/2}^{2J}\right]^{I(I-1)/2} && \text{when n is even,} \end{aligned} \tag{9}$$

where $\nu$ is the number of units exchanged between two groups. The cardinality of the permuted statistics rapidly increases with $n$, $I$ and $J$. The minimum significance level that can be achieved is proportional to the inverse of the cardinality in Equation 9. Thus USP is not expected to suffer of a minimum significance level higher than the desired type I error. However, USP is computationally more intensive compared to CSP, and it is recommended in the case of small number of replicates.

## 3 NonParametric Combination (NPC)

The NonParametric Combination is a natural extension of permutation testing to a variety of multivariate problems. Permutation tests are, in general, distribution-free and non-parametric [11], and have good properties such as exactness, unbiasedness and consistency [31,16,9,10].

To illustrate the NPC we assume the same linear additive model and use the same notation as in Section 2. The notation has to be extended to the multivariate case introducing $V$ observed variables that can be independent or dependent. Because of the objective of this study, we focus on continuous variables. Let us denote a $V$-dimensional data set by $\boldsymbol{Y} = \{\boldsymbol{Y}_{i,j,k}, \quad i = 1, \dots, I, j = 1, \dots, J, k = 1, \dots, n\} = \{Y_{i,j,k,v}, \quad i = 1, \dots, I, j = 1, \dots, J, k = 1, \dots, n, v = 1, \dots, V\}$. According to the extended notation, the multivariate linear additive model of a balanced two-way factorial crossed design is:

$$\boldsymbol{Y}_{i,j,k} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\beta}_j + (\boldsymbol{\alpha\beta})_{ij} + \boldsymbol{\epsilon}_{ijk},\ i = 1, \dots, I;\ j = 1, \dots, J;\ k = 1, \dots, n, \quad (10)$$

where, in the multivariate case, $\boldsymbol{\mu}$ is the vector of overall means, $\boldsymbol{\alpha}_i$ is the vector of effects of level $i$ of factor $A$, $\boldsymbol{\beta}_j$ is the vector of effects of level $j$ of factor $B$, $(\boldsymbol{\alpha\beta})_{ij}$ is the vector of effects of interaction between factor $A$ at level $i$ and factor $B$ at level $j$, $I$ and $J$ are the number of levels of factor $A$ and $B$ respectively, and $\boldsymbol{\epsilon}_{ijk}$ is the vector of error terms. The vector of the means of error terms is $E(\boldsymbol{\epsilon}_{ijk}) = \mathbf{0}$ for each factor level combination, and $\boldsymbol{\Sigma}$ is the variance/covariance matrix of the $V$ observed variables.

Adapted from Pesarin and Salmaso [31] to the design of interest in this study, main assumptions regarding the data structure, hypotheses being tested in NPC contexts, and set of partial tests are:

- (**i**) The response $\boldsymbol{Y}$ takes its values on a $V$-dimensional distribution, $D_{i,j} \in \mathrm{D}, i = 1, \dots, I, j = 1, \dots, J$, belonging to a (possibly not specified) nonparametric family D of non-degenerate probability distributions.
- (**ii**) The null hypothesis refers to equality of effect vectors of the $V$ variables in the $I$ groups for factor $A$, the $J$ groups for factor $B$ and the $I \cdot J$ groups for interaction between factor $A$ and factor $B$:

$$\begin{aligned} H_0^{(A)} &: \quad \boldsymbol{\alpha}_i = \mathbf{0} \quad \forall i, \\ H_0^{(B)} &: \quad \boldsymbol{\beta_j} = \mathbf{0} \quad \forall j \quad \text{and} \\ H_0^{(AB)} &: \quad (\boldsymbol{\alpha\beta})_{ij} = \mathbf{0} \quad \forall i, j, \end{aligned} \quad (11)$$

The null hypotheses $H_0^{(A)}$ and $H_0^{(B)}$ imply that the $V$-dimensional data vectors in $\boldsymbol{Y}$ are exchangeable with respect to the $I$ and $J$ groups respectively.
Considering factor $A$, $H_0^{(A)}$ is supposed to be properly and equivalently broken down into $V$ sub-hypotheses $H_{0v}^{(A)}$, $v = 1, \dots, V$, each appropriate for a partial (univariate) aspect of interest. Therefore, $H_0^{(A)}$ (multivariate) is true if all the $H_{0v}^{(A)}$ are jointly true; and so it may be written as:

$$H_0^{(A)} : \left\{ \bigcap_{v=1}^{V} H_{0v}^{(A)} \right\} \quad (12)$$

$H_0^{(A)}$ is also called *global* or *overall null hypothesis* for factor $A$, and $H_{0v}^{(A)}$, $v = 1, \ldots, V$ are called the *partial null hypotheses.* Similarly $H_0^{(B)}$ and $H_0^{(AB)}$ are supposed to be properly and equivalently broken down into $V$ sub-hypotheses, giving:

$$H_0^{(B)} : \left\{ \bigcap_{v=1}^{V} H_{0v}^{(B)} \right\}, \qquad H_0^{(AB)} : \left\{ \bigcap_{v=1}^{V} H_{0v}^{(AB)} \right\} \tag{13}$$

– **(iii)** Considering factor $A$, the alternative hypothesis states that at least one of the partial null hypotheses $H_{0v}^{(A)}$ is not true. Hence, the alternative may be represented by the union of $V$ partial alternatives hypotheses,

$$H_1^{(A)} : \left\{ \bigcup_{v=1}^{V} H_{1v}^{(A)} \right\}, \tag{14}$$

stating that $H_1^{(A)}$ is true when at least one partial alternative hypotheses $H_{1v}^{(A)}$ is true. In this context, $H_1^{(A)}$ is called the *global* or *overall alternative hypothesis.* Based on the same rationale, we have:

$$H_1^{(B)} : \left\{ \bigcup_{v=1}^{V} H_{1v}^{(B)} \right\}, \qquad H_1^{(AB)} : \left\{ \bigcup_{v=1}^{V} H_{1v}^{(AB)} \right\} \tag{15}$$

– **(iv)** $\boldsymbol{T} = \boldsymbol{T}(\boldsymbol{Y})$ represents a $V$-dimensional vector of test statistics, $V \geq 2$, in which the $v$-th component $T_v = T_v(\boldsymbol{Y})$, $v = 1, \ldots, V$, represents the non-degenerate $v$-th *partial univariate test* appropriate for testing sub-hypothesis $H_{0v}$ against $H_{1v}$. In the NPC context, without loss of generality, all partial tests are assumed to be marginally unbiased, consistent and significant for large values.

The above set of mild conditions should be jointly satisfied. Concerning the partial univariate test, we note that Synchronized Permutation test respects requirements of point (iv) (for more details see [3]). Without loss of generality, from here on we intend the partial univariate test and related statistics to be the Synchronized Permutation test and its statistics.

When developing a multivariate hypothesis testing procedure, a global answer including several response variables is required, and the main point is how to combine the information related to the $V$ variables into one global test. The key idea in the NPC to test the global null hypoteses $H_0^{(A)}$, $H_0^{(B)}$ and $H_0^{(AB)}$ is to combine through an appropriate combining function the partial (univariate) tests which are focused on the $v$-th component variable. Basically, NPC approach corresponds to a method of analysis made up of two phases. In the first phase the univariate permutation tests are performed. In the second phase the $p$-values obtained in the first phase are combined in one second-order global (multivariate) test:

$$T'' = \phi\ (\lambda_1, \ldots, \lambda_V) \tag{16}$$

where $T''$ is the multivariate statistic, $\phi$ is the combining function and $\lambda_v$, $v = 1, \ldots, V$ is the $p$-value of the $v$-th partial univariate test. The test is performed by a continuous, non-increasing and univariate real function $\phi : (0,1)^V \to \mathcal{R}^1$.

Various combining functions can be suitable for this purpose, but according to Pesarin and Salmaso [31], in order to be suitable for test combination all combining functions $\phi$ must satisfy the following properties (see also [27–29] and [12]):

- (**i**) The function $\phi$ must be non-increasing in each argument: $\phi\ (\ldots, \lambda_v, \ldots) \geq \phi\ (\ldots, \lambda'_v, \ldots)$ if $\lambda_v < \lambda'_v$, $v \in \{1, \ldots, V\}$.
- (**ii**) The function $\phi$ must attain its supremum value $\bar{\phi}$, possibly not finite, even when only one argument attains zero: $\phi\ (\ldots, \lambda_v, \ldots) \to \bar{\phi}$ if $\lambda_v \to 0$, $v \in 1, \ldots, V$.
- (**iii**) $\forall \alpha > 0$, the critical value $T''_\alpha$ of every $\phi$ is assumed to be finite and strictly smaller than $\bar{\phi} : T''_\alpha < \bar{\phi}$.

In the simulation study we present in Sections 4 and 5, we selected and compared performances of three combining functions that satisfy the required properties, namely:

- (**i**) The Fisher *omnibus* combining function is based on the statistic:

$$T''_F = -2 \cdot \sum\nolimits_v \log(\lambda_v) \tag{17}$$

  If the $V$ partial test statistics are independent and continuous, then in the null hypothesis $T''_F$ follows a central $\chi^2$ distribution with $2V$ degrees of freedom.
- (**ii**) The Liptak combining function is based on the statistic:

$$T''_L = \sum\nolimits_v \Phi^{-1}(1 - \lambda_v) \tag{18}$$

  where $\Phi$ is the standard normal cumulative distribution function (CDF). If the $V$ partial tests are independent and continuous, then in the null hypothesis $T''_L$ is normally distributed with mean 0 and variance $V$ (see [23]).
- (**iii**) The Tippett combining function is based on the statistic:

$$T''_T = \max\nolimits_{1 \leq v \leq V}(1 - \lambda_v) \tag{19}$$

  significant for large values. Its null distribution, if the $V$ tests are independent and continuous, behaves according to the largest of $V$ random values from the uniform distribution in the open interval $(0, 1)$.

At this stage, once defined data set structure, null and alternative hypotheses, univariate test statistic, combining functions and various assumptions and properties required in the NPC context, we herein illustrate the two phases algorithm to perform the NPC test in the framework of the Conditional Monte Carlo Procedure (CMCP). We resort to CMCP because in most real problems computational difficulties arise in calculating the conditional permutation space when the sample size is large enough, therefore it could be not possible to calculate the exact $p$-value $\lambda_v$ of the observed statistic $T_v^{obs}$ in a reasonable amount of time. It is worth noting that in the multivariate data set CMCP apply permutations of individual data vectors, so that all underlying dependence relations which are present in the component variables are preserved.

For the sake of clearness and simplicity, the algorithm is presented referring to a general test of hypothesis. The reader should take in mind that it has to be repeated three times in a two-way factorial design to test the three global null hypotheses $H_0^{(A)}$, $H_0^{(B)}$ and $H_0^{(AB)}$.

The first phase of the algorithm to perform NPC test is devoted to the estimation of $V$-variate distribution of $\boldsymbol{T}$ and the $p$-values of the univariate tests:

- (i) Calculate the $V$-dimensional vector of the observed values of test statistics $\boldsymbol{T} : \boldsymbol{T}^{obs} = \boldsymbol{T}(\boldsymbol{Y}) = \left[T_v^{obs} = T_v(\boldsymbol{Y}),\ v = 1, \ldots, V\right]$.
- (ii) Consider a random permutation $\boldsymbol{Y}^*$ of $\boldsymbol{Y}$ and calculate the vector of statistics $\boldsymbol{T}^* = \boldsymbol{T}(\boldsymbol{Y}^*) = [T_v^* = T_v(\boldsymbol{Y}^*),\ v = 1, \ldots, V]$.
- (iii) Repeat the previous step $C$ times independently. The set of CMC results $\{\boldsymbol{T}_c^*, c = 1, \ldots, C\}$ is thus a random sampling from the permutation $V$-variate distribution of vector of test statistics $\boldsymbol{T}$.
- (iv) According to Synchronized Permutation test, the $p$-values of the observed values are calculated in each univariate test as the proportion of permutations for which test statistics of permuted data set are greater or equal to the test statistic of the original data set:

$$\hat{\lambda}_v^{obs} = \sum_{c=1}^{C} \mathbb{I}(T_{cv}^* \geq T_v^{obs})/C, \quad v = 1, \ldots, V, \tag{20}$$

  where $\mathbb{I}(\cdot)$ is the indicator function. The result is $\{\hat{\lambda}_1^{obs}, \ldots, \hat{\lambda}_V^{obs}\}$.
- (v) The $p$-values of each of the $C$ elements of the set of permutations carried out at point (iii) are calculated in each univariate test in similar way as point (iv). The result is $\{\hat{\lambda}_{c1}^*, \ldots, \hat{\lambda}_{cV}^*\}$, $c = 1, \ldots, C$.

The second phase of the algorithm to perform NPC test is devoted to the combination of results of the first phase to compute a second-order global (multivariate) test for the overall null hypothesis:

- (i) The combined observed value of the second-order test is calculated by applying a combining function to the $p$-values of the observed values:

$$T''^{obs} = \phi(\hat{\lambda}_1^{obs}, \ldots, \hat{\lambda}_V^{obs}). \tag{21}$$

- (ii) The $c$-th combined value of the $V$-dimensional vector of $p$-values of the $c$-th element of the set of permutations is then calculated by:

$$T_c''^* = \phi(\hat{\lambda}_{c1}^*, \ldots, \hat{\lambda}_{cV}^*), \quad c = 1, \ldots, C. \tag{22}$$

- (iii) Hence, the $p$-value of the combined test $T''$ is estimated as

$$\hat{\lambda}_\phi'' = \sum_{c=1}^{C} \mathbb{I}(T_c''^* \geq T''^{obs})/C. \tag{23}$$

- (iv) If $\hat{\lambda}_\phi'' \leq \alpha$, the global null hypothesis $H_0$ is rejected at significance level $\alpha$.

## 4 Simulations Design

A Monte Carlo simulation study is performed to evaluate the performances of the application of NPC methods to the SP tests described in Sections 3 and 2 respectively.

Data set for simulation are generated according to the cell mean model of a multivariate two-way balanced crossed factorial design (factor $A$ and factor $B$). Consistently with notation in Section 3:

$$\boldsymbol{Y}_{i,j,k} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\beta}_j + (\boldsymbol{\alpha\beta})_{ij} + \boldsymbol{\epsilon}_{ijk}, \tag{24}$$

$i = 1, \ldots, I$, $j = 1, \ldots, J$, $k = 1, \ldots, n$, $v = 1, \ldots, V$, where $I$ and $J$ are the number of levels of factor $A$ and $B$ respectively, $n$ is the number of replicates and $V$ is the number of response variables. As the factorial design is balanced, number of replicates refers to each factor level combination. In this setup, the vector of overall means $\boldsymbol{\mu} = \mathbf{0}$ and the interaction is given by the product of effect of the two factors.

Four distributions are used to generate the error term $\epsilon$. Three are symmetric: normal, Laplace and student's t with 2 degrees of freedom (d.o.f.). One is skewed: lognormal. We consider only homoscedastic case.

Some parameters in the model are fixed:

- The factor effect is $\delta = 1$ for both factors. According to the adopted simulation design, the maximum difference between the means of two levels due to a single factor effect is $\delta$. In the case of two levels of factor $A$ we have $\alpha_1 = 0.5$ and $\alpha_2 = -0.5$, while in the case of three levels we set $\alpha_1 = 0.5$, $\alpha_2 = 0$ and $\alpha_3 = -0.5$. The same for factor $B$: $\beta_1 = 0.5$, $\beta_2 = -0.5$, and $\beta_1 = 0.5$, $\beta_2 = 0$, $\beta_3 = -0.5$ in the case of two and three levels respectively.
- The variance of distributions is fixed as well: $\sigma^2 = 1$

Some other parameters in the model are varied:

- The number of levels of factors: $I, J = 2, 3$. We consider two possible settings: $(I, J) \in \{(2, 2), (3, 3)\}$, so the two factors have the same number of levels in both settings.
- The number of response variables: $V$ = 2,4,8, where the number of active variables (under the alternative hypotesis) is 2 when $V$=2, is 2 when $V$=4 and is 4 when $V$=8.
- The dependence and independence among response variables. In case of independence, the variance/covariance matrix is the identity matrix $\boldsymbol{I}_V$ where $\sigma_{rc} = 0$, $\forall r, c = 1, \ldots, V$, $r \neq c$. In case of dependence, the variance/covariance matrix is $\boldsymbol{\Sigma}_V$ where $\sigma_{rc} = 0.5$, $\forall r, c = 1, \ldots, V$, $r \neq c$.
- The number of replicates: $n = 3, 5$.

It is well known that the number of replicates affects positively the power of the tests as it increases. Studying the performance in case of low number of replicates reflects the frequent needs of practitioners in the industrial environment where there are constraints or limited resources for the experimental design. The SP tests (CSP, USP) combined with the three combining functions (Fisher, Liptak and Tippett) of NPC methods will be investigated in the 24 settings defined as combination of the varying parameters, and will be compared with the MANOVA test along the four distribution functions (normal, Laplace, lognormal and Student t). Furthermore, some simulations with 100 and 50 response variables (50 and 25 active variables respectively) are run with covariance = 0.5, number of levels = 2 and number of replicates = 5, to investigate the behavior of NPC applied to Synchronized Permutation tests with an high number of response variables.

All simulations are performed in R (version 3.4.0; R Development Core Team (2017)). The number of simulations is $n_{sim} = 10000$, and the number of permutations for CSP and USP is $n_{perm} = 2000$.

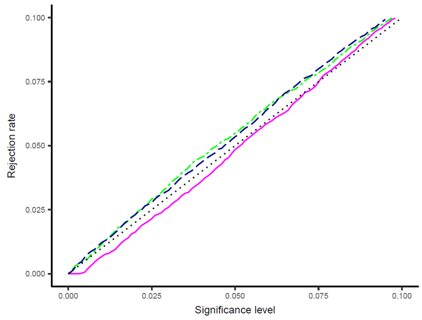
## 5 Simulations Results

In this section main results of the simulation study are presented. The graphs have been obtained plotting the rejection rate of the test (y axis) versus the significance level (x axis). The objective is to compare the performance of the NPC combining functions applied to the permutation tests in the range of interest of significance level $0 \leq \alpha \leq 0.1$. Because of the simmetry of the simulation design, we have same results for factor $A$ and factor $B$.

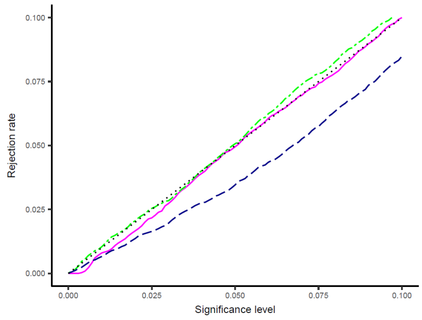
In Figure 1 it is clear that in case of non normal distribution of errors, the MANOVA test does not respect the $\alpha$ level under the null hypothesis unlike the NPC tests. There is a discrepancy between the MANOVA curve and the line of no-discrimination when $H_0^{(A)}$ is true, while in general NPC tests' curves are very close to the hypothetical continuous uniform distribution with every combining function for the main factor (factor $A$).

The respect of the $\alpha$ level under the null hypothesis in the analysis of interaction effect is more challenging for all the considered tests. Under non normality MANOVA shows a clear departure from no-discrimination line. The multivariate combination of USP test reveals to be unreliable as well. The most performing test for the interaction effect in the model assumed is the Liptak combination of CSP test (Figure 2).

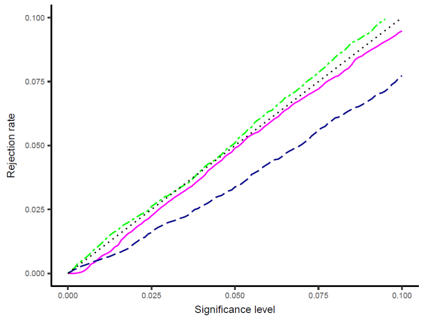
In Section 2 the issue of the minimum achievable significance level related to the cardinality of the univariate permutation tests is shown both for CSP and USP. The curves of NPC of Synchronized Permutation tests reflect the same issue with an initial plateau with rejection rate = 0. A clear example is given by the case of CSP and a data set with $(I, J) = (2, 2)$ levels for the factors $A$ and $B$ respectively and $n = 3$ replicates. The cardinality of the univariate test is $C_{csp} = \binom{2n}{n} = \binom{6}{3} = 20$ and the univariate minimum achievable significance error is $\alpha_{min} = 2 \times (C_{csp})^{-1} = 0.1$. The set of exact $p$-values that can result under such conditions is a set of 10 values at step of 0.1: $S_{p-values} = \{0.1, \ldots, 1\}$ for each univariate test. The application of a combining function gives a set of 10 values for the statistic $T''$ used to compute the $p$-value of the second-order test $\hat{\lambda}''_{\phi}$. The cumulative distribution function of the $p$-values is shown in Figure 3 (a) where we can recognize 10 steps. Recall that we are using a CMC procedure with a number of permutations large enough so that the length of the steps is quite regular thanks to the fact that the likelihood of the values of the statistic $T''$ is the same. A zoom in Figure 3 (b) shows clearly the initial plateau with rejection rate = 0. The NPC tests are affected by the minimum achievable significance level of the permutation test they are applied to. All the graphs referred to CSP in this section show a small initial plateau. This fact reveals to be a shortcoming only in the case of CSP and 3 replicates considering the usual significance levels $\alpha = 0.01$ and $\alpha = 0.05$.



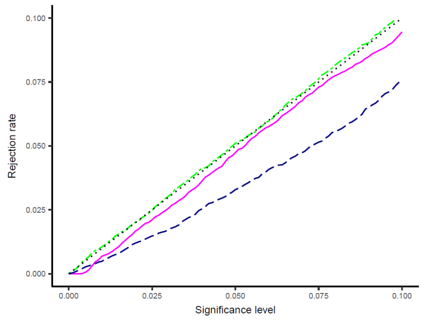

(a) Normal distribution of errors.




(b) Laplace distribution of errors.




(c) Lognormal distribution of errors.




(d) T (2 d.o.f) distribution of errors.

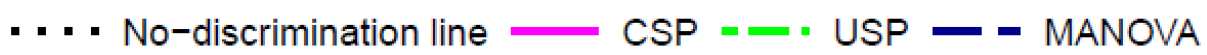
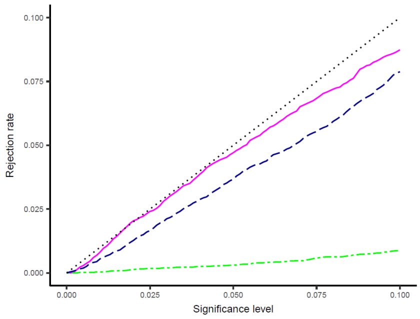


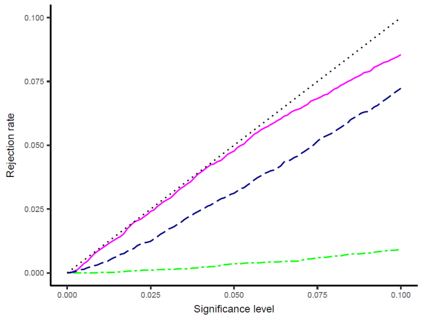

Fig. 1: Factor $A$, behaviour of tests under null hypothesis. Rejection rate at different values of significance level $\alpha$ (x axis). NPC function = Fisher; number of responses = 4; number of levels = 2; covariance = 0; number of replicates = 5.

The capability of NPC applied to SP tests to detect the effect of the main factor under $H_1$ in the assumed model is in general good. Simulation results show that NPC applied to USP and CSP partial tests gives high values of power (rejection rate) both with independent and dependent response variables [1], and both with low number and high number of response variables compared to MANOVA (Figures 4, 5). In general NPC applied to USP partial tests performs better with all distribution of errors, while NPC of CSP partial tests and MANOVA have in some cases similar performances with Laplace and student's t distribution of errors. NPC tests show a better performance compared to MANOVA even in case of normal distribution of errors with the power that can be even more than double at $\alpha < 0.05$ significance level (Figure 4).

In the interaction analysis, the observed rejection rate under $H_1$ is lower than in the main factor analysis (Figure 6). This result is consistent with the model used to generate data set for simulation study where interaction is given by the product of the levels of the factors. In general, the NPC of CSP partial tests performs better than MANOVA and USP, both with independent and dependent response




(a) Lognormal distribution of errors.




(b) T (2 d.o.f) distribution of errors.

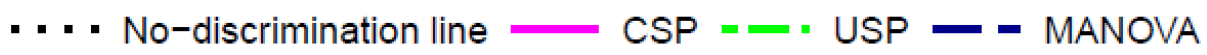
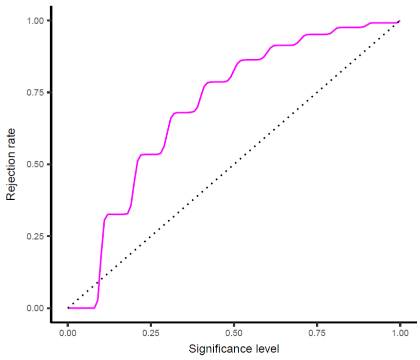
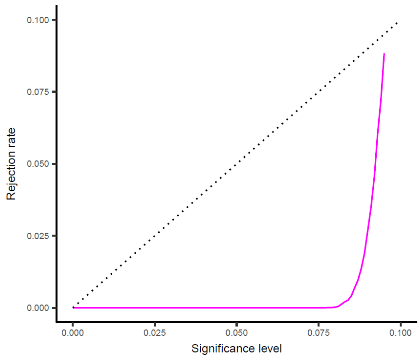



Fig. 2: Interaction $AB$, behaviour of tests under null hypothesis. Rejection rate at different values of significance level $\alpha$ (x axis). NPC function = Liptak; number of responses = 4; number of levels = 2; covariance = 0; number of replicates = 5.



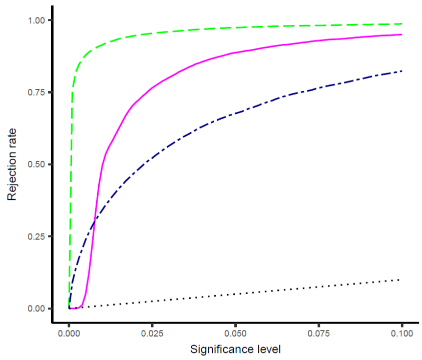

(a) Rejection rate $\in$(0,1), Signif. level $\in$(0,1).



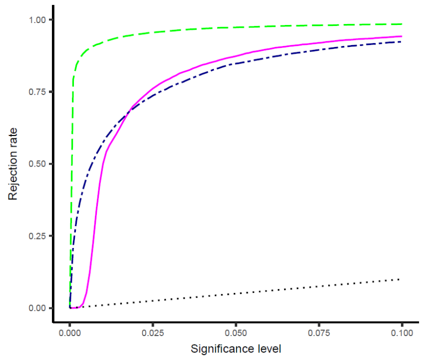

(b) Zoom of graph in (a).



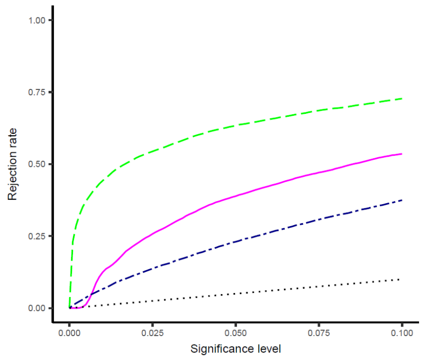

Fig. 3: Factor $A$, effect of cardinality of Synchronized Permutations on minimum significance level of NPC. Rejection rate at different values of significance level $\alpha$ (x axis). NPC function = Fisher; number of responses = 4; distribution of errors = T (2 d.o.f) ; number of levels = 2; covariance = 0.5; number of replicates = 3.



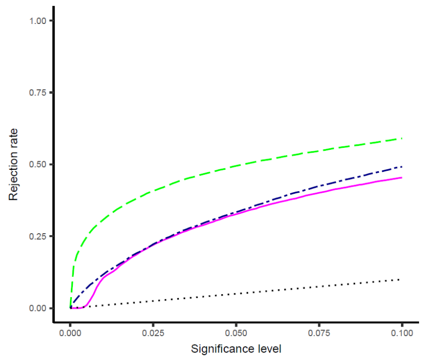

(a) Normal distribution of errors.




(b) Laplace distribution of errors.




(c) Lognormal distribution of errors.




(d) T (2 d.o.f) distribution of errors.

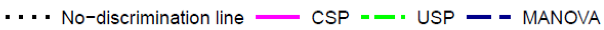


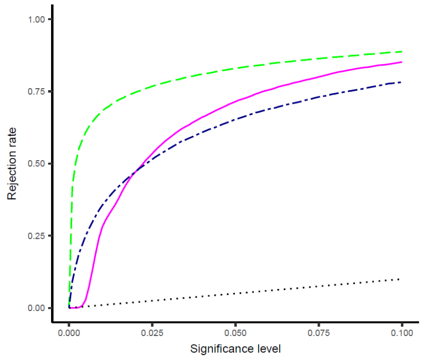

Fig. 4: Factor $A$, comparison of Non-Parametric methods and MANOVA. Rejection rate at different values of significance level $\alpha$ (x axis). NPC function = Fisher; number of responses = 8; number of levels = 3; covariance = 0; number of replicates = 5.

variables, and both with low number and high number of response variables. The NPC of USP partial tests shows a lower power. In some cases with a factor with two levels, MANOVA and NPC applied to CSP have similar rejection rate, with MANOVA performing better with Laplace distribution of errors.

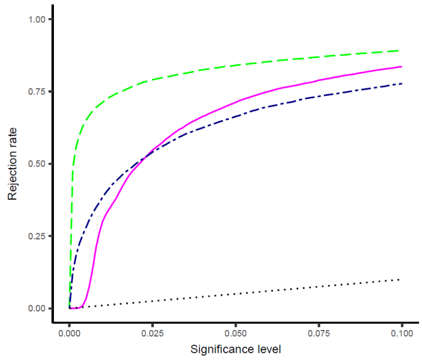
The increase of number of levels of factors $A$ and $B$ from $(I, J) = (2, 2)$ to $(I, J) = (3, 3)$ has a positive effect on the rejection rate under $H_1$ (Figure 7) even if the maximum $\delta = 1$ between factors is constant.

The three combining functions Fisher, Liptak and Tippett show in general good performances. Nevertheless, Fisher and Tippett functions give higher rejection rate compared to Liptak, with Tippett function's curve showing small steps (Figure 8).

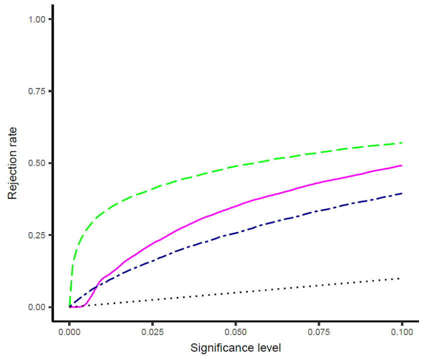
An increase in rejection rate can be observed when the number of response variables increases with fixed number of observed units, meaning an higher power of the test in detecting a factor effect under $H_1$ (Figure 9, 10). This phenomenon is



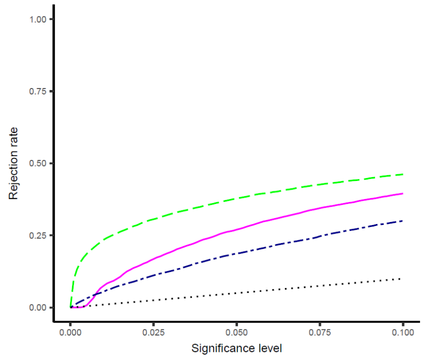

(a) Normal distribution of errors.




(b) Laplace distribution of errors.




(c) Lognormal distribution of errors.




(d) T (2 d.o.f) distribution of errors.

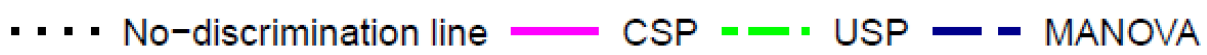
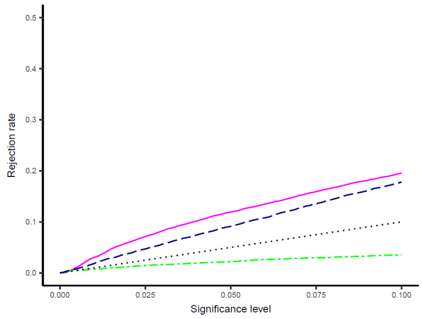
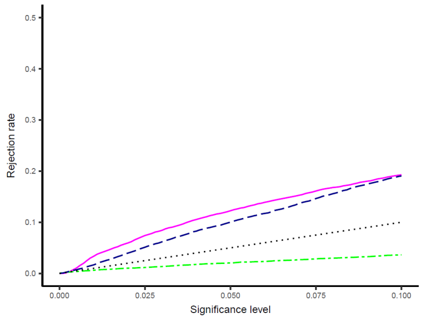



Fig. 5: Factor $A$, comparison of Non-Parametric methods and MANOVA. Rejection rate at different values of significance level $\alpha$ (x axis). NPC function = Fisher; number of responses = 2; number of levels = 3; covariance = 0.5; number of replicates = 5.

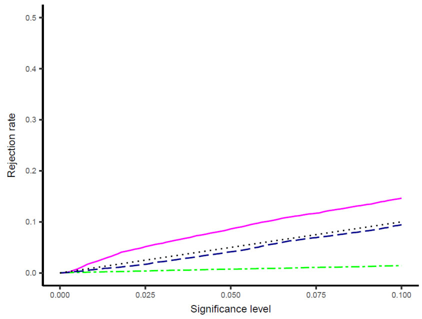
known as *finite sample consistency* and refers to a peculiar property of multivariate combination-based inferences: the power of NPC tests for any added variable monotonically increases if the variable makes larger noncentrality parameter of the underlying population distribution [31,30]. The positive effect on the power of the test that can be obtained adding response variables can be strategically exploited considering that in many real problems it could be easier to collect more information on a single experimental unit than adding a new unit to the experimental design [8]. The effect of the increase of response variables while keeping constant the number of observed units couldn't be investigated for MANOVA test because of the problem of the loss of degrees of freedom that does not allow to apply MANOVA test when the number of response variables is larger than the sample size.



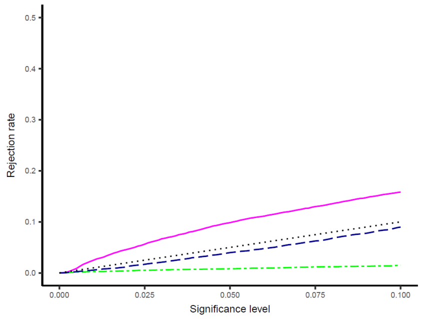

(a) Normal distribution of errors.




(b) Laplace distribution of errors.




(c) Lognormal distribution of errors.




(d) T (2 d.o.f) distribution of errors.

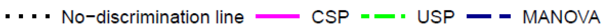



Fig. 6: Interaction $AB$, comparison of Non-Parametric methods and MANOVA. Rejection rate at different values of significance level $\alpha$ (x axis). NPC function = Liptak; number of responses = 4; number of levels = 3; covariance = 0.5; number of replicates = 5.

## 6 A Real Case study on an Industrial Experiment

An industrial experiment according to a two-way two-levels design in the engineering field provides a useful example of the analysis performed using NPC combined with Synchronized Permutations on a dataset with two responses.

The production system of plastic thermoformed packaging is complex, and it is controlled by several factors [32]. In order to innovate the system, the impact of two factors and of their interaction has to be assessed for values of levels outside their usual range. One factor is the temperature of the process (factor $A$) and the other factor is the pace of production line (factor $B$). The packaging is composed by two separate and different-in-shape chambers. The evaluation of the strength of the packaging is done observing the pressure needed to break the packaging by a burst test [bar]. Each chamber is tested separately, so there are two response variables. Five replicates for each factor level combination and for each response variable have been tested. Data collected violate assumption of normality (Table 1), and

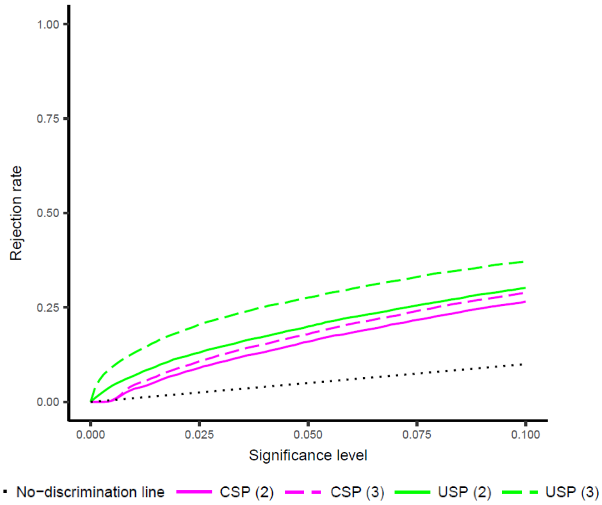
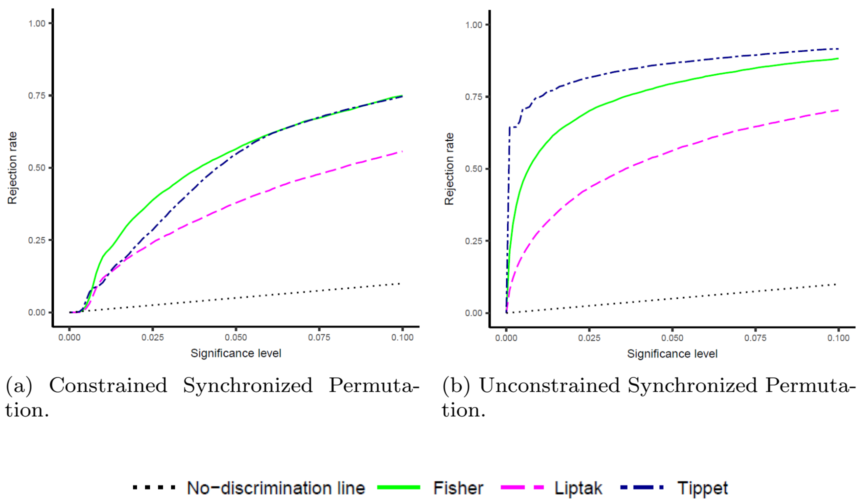


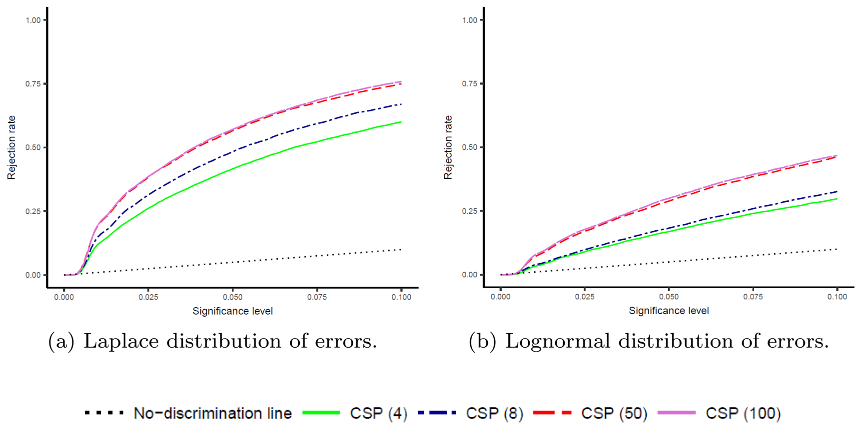

Fig. 7: Factor $A$, performance at different number of levels, 2 and 3. Rejection rate at different values of significance level $\alpha$ (x axis). NPC function = Fisher; distribution of errors = T (2 d.o.f); number of responses = 4; covariance = 0.5; number of replicates = 5.




(a) Constrained Synchronized Permutation.

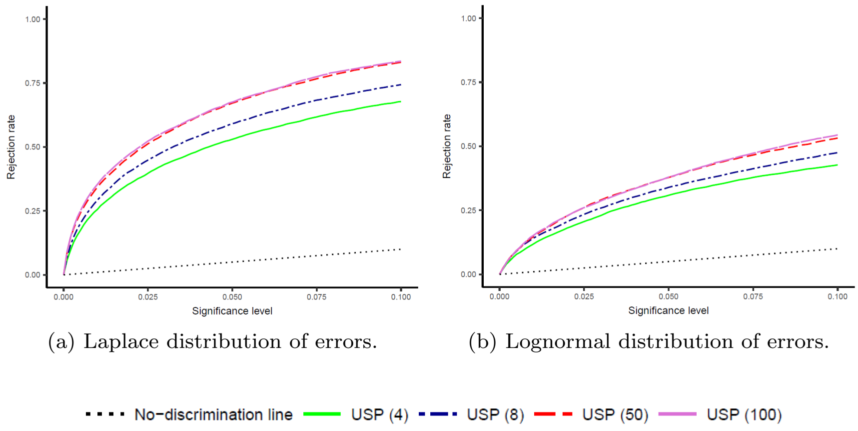
(b) Unconstrained Synchronized Permutation.

Fig. 8: Factor $A$, comparison of NPC functions. Rejection rate at different values of significance level $\alpha$ (x axis). Distribution of errors = Laplace; number of responses = 8; number of levels = 3; covariance = 0.5; number of replicates = 5.




(a) Laplace distribution of errors. (b) Lognormal distribution of errors.

Fig. 9: Factor $A$, effect of increasing number of responses with CSP. Rejection rate at different values of significance level $\alpha$ (x axis). NPC function = Fisher, number of levels = 2; covariance = 0.5; number of replicates = 5.




(a) Laplace distribution of errors. (b) Lognormal distribution of errors.

Fig. 10: Factor $A$, effect of increasing number of responses with USP. Rejection rate at different values of significance level $\alpha$ (x axis). NPC function = Fisher, number of levels = 2; covariance = 0.5; number of replicates = 5.

they reveal heteroscedasticity based on the different values of control factors (Table 2).

MANOVA test is not reliable for the analysis of such data set. The NPC combined with CSP and USP overcome the violation of Manova assumptions. The results of the test are in Table 3 and 4. Note that interaction effect has been analized only with NPC applied to CSP because USP doesn't respect the $\alpha$ level under the null hypothesis. According to simulation study results, main factor effect should preferably be assessed using NPC applied to USP. The null hypothesis is rejected at a significance level $\alpha$ = 0.05. Both the factors and their interaction

Table 1: Multivariate normality tests

| | Shapiro-Wilk | Henze-Zirkler | Royston |
|---|---|---|---|
| Test statistic | 0.908 | 1.374 | 11.626 |
| $p$-value | 0.058 | 0.001 | 0.003 |

Table 2: Box's M-test for Homogeneity of Covariance Matrices

| | Chi-Sq (approx.) |
|---|---|
| Test statistic | 24.576 |
| DF | 9 |
| $p$-value | 0.003 |

Table 3: Constrained Synchronized Permutation (CSP): $p$-values of the NPC tests

| | Fisher | Liptak | Tippett |
|---|---|---|---|
| Temperature | 2.60e-02 | 3.50e-02 | 2.70e-02 |
| Cycles per minute | 1.50e-02 | 3.30e-02 | 7.00e-03 |
| Interaction | 7.00e-03 | 7.00e-03 | 7.00e-03 |

Table 4: Unconstrained Synchronized Permutation (USP): $p$-values of the NPC tests

| | Fisher | Liptak | Tippett |
|---|---|---|---|
| Temperature | 6.00e-03 | 1.75e-02 | 4.00e-03 |
| Cycles per minute | 5.00e-04 | 7.00e-03 | 5.00e-04 |

have a significant impact on the final product. A further investigation allowed to find the setting for the optimal strength of the packaging.

## 7 Conclusions

The application of NonParametric Combination to Synchronized Permutation tests to analyze a multivariate two-way factorial design reveals to be a good instrument for inferential statistics when assumptions of MANOVA are violated. Simulation results show that NPC applied to USP and CSP partial tests gives high values of power (rejection rate) under alternative hypothesis $H_1$ both with independent and dependent response variables, and both with low number and high number of response variables compared to MANOVA. In general NPC applied to USP partial tests performs better for the main factor analysis with all distribution of errors compared to NPC of CSP partial tests and MANOVA. Its power varies under the conditions it has been tested in the simulation study, and it has been observed to be higher than 75% at a significance level $\alpha = 0.05$ in many cases. For the interaction analysis we recommend the adoption of NPC of CSP partial tests with the Liptak combining function because of the higher adherence of the test to the nominal $\alpha$ level. The Fisher combining function, also referred to as *omnibus*, is in general preferable to the Tippett and the Liptak ones in the main factor analysis.

A great advantage given by the adoption of these tests is that they well perform with small sample size. This reflects the frequent needs of practitioners in the industrial environment where there are constraints or limited resources for the experimental design. In case of $n = 3$ replicates we recommend the use of NPC of USP partial tests for main factor analysis because of the shortcomings of the minimum achievable significance level related to the cardinality of the univariate CSP test. The increase of sample size has in general an evident positive effect on the power of NPC of SP tests. Futhermore the power of the test is improved by the increase of number of factor levels with the factor effect fixed.

At last, there is an important property of NPC of SP tests that can be exploited to increase their power: the finite sample consistency. Indeed, an increase in rejection rate can be observed under alternative hypothesis $H_1$ when the number of response variables increases with fixed number of observed units. This could lead to a strategical benefit considering that in many real problems it could be easier to collect more information on a single experimental unit than adding a new unit to the experimental design.

**Acknowledgements** Authors would like to thank the Editor and Referees for useful comments and suggestions which helped improving the manuscript. The research was supported by UNIPD BIRD185315/18.

## References

1. Arboretti Giancristofaro, R., Bonnini, S., Corain, L., Salmaso, L.: Dependency and truncated forms of combinations in multivariate combination-based permutation tests and ordered categorical variables. Journal of Statistical Computation and Simulation **86**(18), 3608–3619 (2016)
2. Basso, D., Chiarandini, M., Salmaso, L.: Synchronized permutation tests in replicated i× j designs. Journal of Statistical Planning and Inference **137**(8), 2564–2578 (2007)
3. Basso, D., Pesarin, F., Salmaso, L., Solari, A.: Permutation Tests for Stochastic Ordering and ANOVA: theory and applications with R. Springer, New York, NY (2009)
4. Bathke, A.C., Harrar, S.W.: Nonparametric methods in multivariate factorial designs for large number of factor levels. Journal of Statistical planning and Inference **138**(3), 588–610 (2008)
5. Bathke, A.C., Harrar, S.W., Madden, L.V.: How to compare small multivariate samples using nonparametric tests. Computational Statistics & Data Analysis **52**(11), 4951–4965 (2008)
6. Brunner, E., Munzel, U., Puri, M.L.: The multivariate nonparametric behrens–fisher problem. Journal of Statistical Planning and Inference **108**(1-2), 37–53 (2002)
7. Corain, L., Salmaso, L.: Nonparametric permutation and combination-based multivariate control charts with applications in microelectronics. Applied Stochastic Models in Business and Industry **29**(4), 334–349 (2013)
8. Corain, L., Salmaso, L.: Improving power of multivariate combination-based permutation tests. Statistics and Computing **25**(2), 203–214 (2015)
9. Giancristofaro, R.A., Corain, L., Ragazzi, S.: A comparison among combination-based permutation statistics for randomized complete block design. Communications in Statistics-Simulation and Computation **41**(7), 964–979 (2012)
10. Giancristofaro, R.A., Corain, L., Ragazzi, S.: The multivariate randomized complete block design: A novel permutation solution in case of ordered categorical variables. Communications in Statistics-Theory and Methods **41**(16-17), 3094–3109 (2012)
11. Good, P.I.: Permutation, Parametric, and Bootstrap Tests of Hypotheses. Springer Series in Statistics. Springer (2005)

12. Goutis, C., Casella, G., Wells, M.T.: Assessing evidence in multiple hypotheses. Journal of the American Statistical Association **91**(435), 1268–1277 (1996)
13. Harrar, S.W., Bathke, A.C.: Nonparametric methods for unbalanced multivariate data and many factor levels. Journal of Multivariate Analysis **99**(8), 1635–1664 (2008)
14. Harrar, S.W., Bathke, A.C.: A nonparametric version of the bartlett-nanda-pillai multivariate test. asymptotics, approximations, and applications. American Journal of Mathematical and Management Sciences **28**(3-4), 309–335 (2008)
15. Harrar, S.W., Bathke, A.C.: A modified two-factor multivariate analysis of variance: asymptotics and small sample approximations. Annals of the Institute of Statistical Mathematics **64**(1), 135–165 (2012)
16. Hoeffding, W.: The large-sample power of tests based on permutations of observations. The Annals of Mathematical Statistics pp. 169–192 (1952)
17. Jarrett, R.: Does theory work in practice? two case studies. Quality Engineering **29**(1), 141–159 (2017)
18. Khaleel, K., Mousa, A., Qashlaq, A., Hassan, B., Ramadan, A., Raslan, B.: Using design of experiment to predict concert compressive strength using fly ash and ggbfs as cement alternative. In: Proceedings of the International Conference on Industrial Engineering and Operations Management, pp. 741–749 (2017)
19. Khreis, H., Deflorio, A., Lee, W.L., De Larramendi, M.R., Schmuelling, B.: Sensitivity analysis for induction machine manufacturing tolerances: Modeling of electrical parameters deviation. In: Ecological Vehicles and Renewable Energies (EVER), 2017 Twelfth International Conference on, pp. 1–9. IEEE (2017)
20. Kruskal, W.H.: A nonparametric test for the several sample problem. The Annals of Mathematical Statistics pp. 525–540 (1952)
21. Kruskal, W.H., Wallis, W.A.: Use of ranks in one-criterion variance analysis. Journal of the American statistical Association **47**(260), 583–621 (1952)
22. Lakens, D.: Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and anovas. Frontiers in psychology **4**, 863 (2013)
23. Liptak, T.: On the combination of independent tests. Magyar Tud Akad Mat Kutato Int Kozl **3**, 171–197 (1958)
24. Liu, Q., Chen, X., Gindy, N.: Evaluation of superalloy heavy-duty grinding based on multivariate tests. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture **221**(9), 1421–1430 (2007)
25. Melas, V.B., Pepelyshev, A., Shpilev, P., Salmaso, L., Corain, L., Arboretti, R.: On the optimal choice of the number of empirical fourier coefficients for comparison of regression curves. Statistical Papers **56**(4), 981–997 (2015)
26. Munzel, U., Brunner, E.: Nonparametric methods in multivariate factorial designs. Journal of Statistical Planning and Inference **88**(1), 117–132 (2000)
27. Pesarin, F.: A resampling procedure for nonparametric combination of several dependent tests. Journal of the Italian Statistical Society **1**(1), 87–101 (1992)
28. Pesarin, F.: Permutation testing of multidimensional hypotheses by nonparametric combination of dependent tests. Cleup (1999)
29. Pesarin, F.: Multivariate permutation tests: with applications in biostatistics, vol. 240. Wiley Chichester (2001)
30. Pesarin, F., Salmaso, L.: Finite-sample consistency of combination-based permutation tests with application to repeated measures designs. Journal of Nonparametric Statistics **22**(5), 669–684 (2010)
31. Pesarin, F., Salmaso, L.: Permutation Tests for Complex Data: Theory, Applications and Software. Wiley Series in Probability and Statistics. Wiley (2010)
32. Ronchi, F., Salmaso, L., De Dominicis, M., Illert, J., Krajewski, K., Link, R., Monti, L.: Optimal designs to develop and support an experimental strategy on innovation of thermoforming production process. Statistica **77**(2), 109 (2017)
33. Salmaso, L.: Synchronized permutation tests in 2k factorial designs. Communications in Statistics - Theory and Methods **32**(7), 1419–1437 (2003)
34. Sangnuan, K., Laosiritaworn, W.S.: Determining the optimal parameter of coordinate measuring machine with design of experiment. In: MATEC Web of Conferences, vol. 68, p. 06009. EDP Sciences (2016)
35. Wiemer, H., Schwarzenberger, M., Dietz, G., Juhrisch, M., Ihlenfeldt, S.: A holistic and doe-based approach to developing and putting into operation complex manufacturing process chains of composite components. Procedia CIRP **66**, 147–152 (2017)

36. Zhang, Q., Zhang, T., Zhang, Z., Zhou, X., Wang, Y., He, C.: Experiment of integrated fermentation hydrogen production by photosynthetic bacteria cooperating with enterobacter aerogenes. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering **33**(9), 243–249 (2017)