Quick multiple test procedures and p-value adjustments (DRAFT)

(short title: Quick multiple test procedures)

Jiangtao Gou
Department of Mathematics and Statistics, Villanova University

August 24, 2019

[ABSTRACT]

We develop a family of logarithmic time stepwise multiple test procedures, called the Quick procedures. Comparing with existing linear time and quadratic time procedures, the Quick procedure not only greatly simplifies and speeds up the decision process but also significantly enhance the statistical power. An R package elitism to implement the newly proposed logarithmic time and linear time procedures is made available on CRAN.

[KEY WORDS] closure principle, partition testing principle, power, stepwise multiple test procedure, time complexity

1 Introduction

P-value based multiple test procedures are the most widely used multiple test procedures, where the decisions are made only based on marginal p-values. A good p-value based multiple test procedure usually are (1) reasonably conservative, (2) easy to understand and (3) time efficient. Equivalently, this procedure satisfies three conditions as follows: (1) The type I error is correctly controlled at a given level. Meanwhile, the procedure should have enough statistical power to detect true significances. (2) Researchers, especially non-statisticians, can understand and use this multiplicity adjustment straightforwardly. (3) The number of comparisons between p-values and their corresponding critical values should be small or moderate.

We propose a set of quick multiple test procedures based on p-values which (1) strongly control the familywise error rate and are more powerful than the commonly used Hochberg and Hommel procedures, (2) follow a simple routine with only two search steps and are easier than the Hochberg procedure, (3) have a logarithmic time complexity, comparing with the Hochberg procedure that runs in linear time.

In addition to newly proposed quick procedures, we also develop modified versions of the Hommel (1988) procedure, Rom (1990) procedure, and Gou et al. (2014) procedure. These modifications improve the time efficiency or increase the statistical power or both. The original versions of these procedures have been widely applied in various research areas, including educational research (Porter, 2018), medicine (Wang and Cui, 2012; Tamhane et al., 2018; Gou and Xi, 2019), neuroimaging (Zhang and Gou, 2019), and public health (Aickin and Gensler, 1996). Statistical softwares, including `R` (Bretz et al., 2010), `SAS` (Westfall et al., 2011), `Stata` (Newson, 2010), and `MATLAB` (Metcalfe et al., 2019), have developed specific packages or functions to implement these multiple test procedures.

The paper is organized as follows. Section 2 provides new modifications of Hommel's, Rom's and Gou et al.'s procedures, introduces the logarithmic time Quick procedures, and discusses the relations among these procedures. Two theorems in Section 3 provide sufficient conditions of existence of linear time and logarithmic time multiple test procedures. Methods to calculate adjusted p-values are included in Section 4. A simulation study is performed to compare the familywise error rate, power and number of comparisons of newly proposed procedures with those of the existing procedures. Results from that investigation are reported in Section 5. Section 6 gives concluding remarks. All proofs are in the Appendix.

2 Logarithmic and linear time multiple test procedures

In this paper we consider testing a family of hypotheses $\{H_i\}_{i=1}^n$ with controlling the familywise error rate (FWER). Let $p_i,\ i = 1, \ldots, n$ denote the marginal p-values of tests for $H_i,\ i = 1, \ldots, n$. The ordered p-values are $p_{(1)} \leq \cdots \leq p_{(n)}$, and the corresponding hypotheses are $\{H_{(i)}\}_{i=1}^n$. The strong control of the FWER at level α requires that $\sup_I \Pr(\text{reject one or more } H_i, i \in I | H_I) \leq \alpha$ for all possible index set I, where $H_I = \cap_{i \in I} H_i$ (Hochberg and Tamhane, 1987; Dmitrienko et al., 2009). Following a brief summary of widely used p-value based procedures, new modifications of these existing procedures and the Quick procedure are developed in this section.

2.1 Holm and Hochberg procedures

The most popular FWER controlling procedures are the Holm (1979) and Hochberg (1988) procedures, along with their variations (Tamhane and Liu, 2008; Bretz et al., 2009). They are both closed testing methods constructed from a certain intersection hypothesis test based on the closure principle (Marcus et al., 1976) or partition testing principle (Hsu, 1996). The Holm (1979) procedure is a step-down procedure based on the Bonferroni test (Dunn, 1961), and the type I error control is guaranteed by the Bonferroni inequality $\Pr\left(\cup_{i=1}^n \{p_i \leq \alpha/n\}\right) \leq \alpha$. Meanwhile, the Hochberg (1988) procedure is a step-up procedure based on the Simes (1986) test, and the type I error control is guaranteed by the Simes inequality $\Pr\left(\cup_{i=1}^n \left\{p_{(i)} \leq i\alpha/n\right\}\right) \leq \alpha$, which holds under independence and certain types of positive dependence (Sarkar and Chang, 1997; Sarkar, 1998). The Hochberg procedure starts from the largest p-value. If $p_{(n)} \leq \alpha$, then all hypotheses are rejected; otherwise, $H_{(n)}$ is accepted and $p_{(n-1)}$ is compared with $\alpha/2$. If $p_{(n-1)} \leq \alpha/2$, then all the remaining hypotheses $\{H_{(k)}\}_{k=1}^{n-1}$ are rejected; otherwise, $H_{(n-1)}$ is accepted. Similarly, in Step i, $p_{(n-i+1)}$ is compared with α/i. If smaller, than all the remaining hypotheses $\{H_{(k)}\}_{k=1}^{n-i+1}$ are rejected; otherwise $H_{(n-i+1)}$ is accepted and one goes to Step $i+1$. In fact, the Hochberg procedure does not apply the Simes test exactly. Instead, a conservative Simes test is actually used based on the Simes-Hochberg inequality $\Pr\left(\cup_{i=1}^n \left\{p_{(n-i+1)} \leq \alpha/i\right\}\right) \leq \alpha$ (Huang and Hsu, 2007; Gou and Tamhane, 2018b).

2.2 Hommel procedure

Using the exact α-level Simes (1986) test as a local test for all intersection hypotheses, Hommel (1988) provided an exact shortcut to the corresponding closed procedure. The Hommel procedure starts from the largest p-value. If $p_{(n)} \leq \alpha$, all hypotheses are rejected; otherwise, let j be the smallest integer for which $p_{(n-j+k)} \leq k\alpha/j$ for at least one $k = 1, \ldots, j$, and reject all hypotheses whose p-values are less than or equal to $\alpha/(j-1)$. If no such j exists, accept all hypotheses. Hommel (1989) showed that his procedure is more powerful than the Hochberg procedure since a conservative Simes test is used as a local test in the Hochberg procedure instead of an exact Simes test. However, the Hommel procedure is less popular than the Hochberg procedure in practice, because it is not as easy to understand and apply as the Hochberg procedure (Shaffer, 1995). Meanwhile, the method of Hommel (1988) takes quadratic time in the number of hypotheses n, where the number of comparisons is $(n^2-n+2)/2$ in the worst-case scenario, comparing with the Hochberg (1988) procedure which only takes linear time with at most n comparisons.

Hommel (1988)'s method potentially compares the ith largest p-value $p_{(n-i+1)}$ with $n-i+1$ different critical values, which are $(k-i+1)\alpha/k$ where k runs from i to n. This method includes some redundant comparisons. For example, if $p_{(n-i+1)}$ is greater than $(k'-i+1)\alpha/k'$, then $p_{(n-i+1)} > (k-i+1)\alpha/k$ holds for all $k < k'$ and these comparisons are considered to be redundant. If we only keep the minimum necessary comparisons in the original Hommel (1988)'s method, we can achieve a better time complexity.

Meijer et al. (2019) recently developed a linear time algorithm for the Hommel procedure and its robust variant (Hommel, 1983, 1986). This method can be treated as a shortcut of Hommel's shortcut procedure. The key component of Meijer et al.'s method is the application of a linear time algorithm for determining the convex hull of a planar set (Graham, 1972; Andrew, 1979; Fortune, 1989), where the finite set is $\{(1,p_{(1)}),\ldots,(n,p_{(n)})\}$. In this section, we will summarize Meijer et al.'s algorithm in a concise way and present our modification for further improvement. For the sake of simplicity, we name the original Hommel (1988) procedure as the Hommel-Q procedure, where the letter Q stands for quadratic time, and Meijer et al. (2019)'s method as the Hommel-L procedure, where the letter L stands for linear time. Meanwhile, we also omit the parentheses and use the capital letter for the ordered p-values in this section on the Hommel procedure. Equivalently, we assume that the p-values are sorted so that $P_1 \le \cdots \le P_n$.

Meijer et al.'s method includes three consecutive routines. The first routine finds the convex hull $\{(I(k),P_{I(k)})\}_{k=1}^{K} \in \{(1,P_1),\ldots,(n,P_n)\}$ of the planar set $\{(i,P_i)\}_{i=1}^{n}$, where indices I satisfy $I(1) < \cdots < I(K)$ and K is the total number of of points of the convex hull. For convenience, we usually add two ancillary points $(I(-1),P_{I(-1)}) = (-1,P_{-1}=1)$ and $(I(0),P_{I(0)}) = (0,P_0=0)$ into the planar set in programming realization, as shown in Figure 1. These two points automatically belongs to the extended convex hull of set$\{(i,P_i)\}_{i=-1}^{n}$. In each step, we consider three consecutive points $(I_j(k-2),P_{I_j(k-2)})$, $(I_j(k-1),P_{I_j(k-1)})$ and $(I_j(k),P_{I_j(k)})$ in a temporary convex hull with indices $\{I_j(l)\}_{l=-1}^{k}$, where j indicates the step number. This routine starts from the temporary convex hull with indices $\{I_1(l)) = l\}_{l=-1}^{2}$, using three consecutive points $(0,P_0)$, $(1,P_1)$ and $(2,P_2)$. In step j, if three points $(I_j(k-2),P_{I_j(k-2)})$, $(I_j(k-1),P_{I_j(k-1)})$ and $(I_j(k),P_{I_j(k)})$ are locally convex, as shown in Figure 1-a, which satisfy

$$\frac{P_{I_j(k)}-P_{I_j(k-1)}}{I_j(k)-I_j(k-1)} > \frac{P_{I_j(k-1)}-P_{I_j(k-2)}}{I_j(k-1)-I_j(k-2)}, \tag{1}$$

then a new point will be added in step $j+1$. The temporary convex hull in step $j+1$ will include points with indices $\{I_{j+1}(l)\}_{l=-1}^{k+1}$ which are

$$I_{j+1}(k+1) = I_j(k)+1 \ \text{ and } \ I_{j+1}(s) = I_j(s), 1 \le s \le k, \tag{2}$$

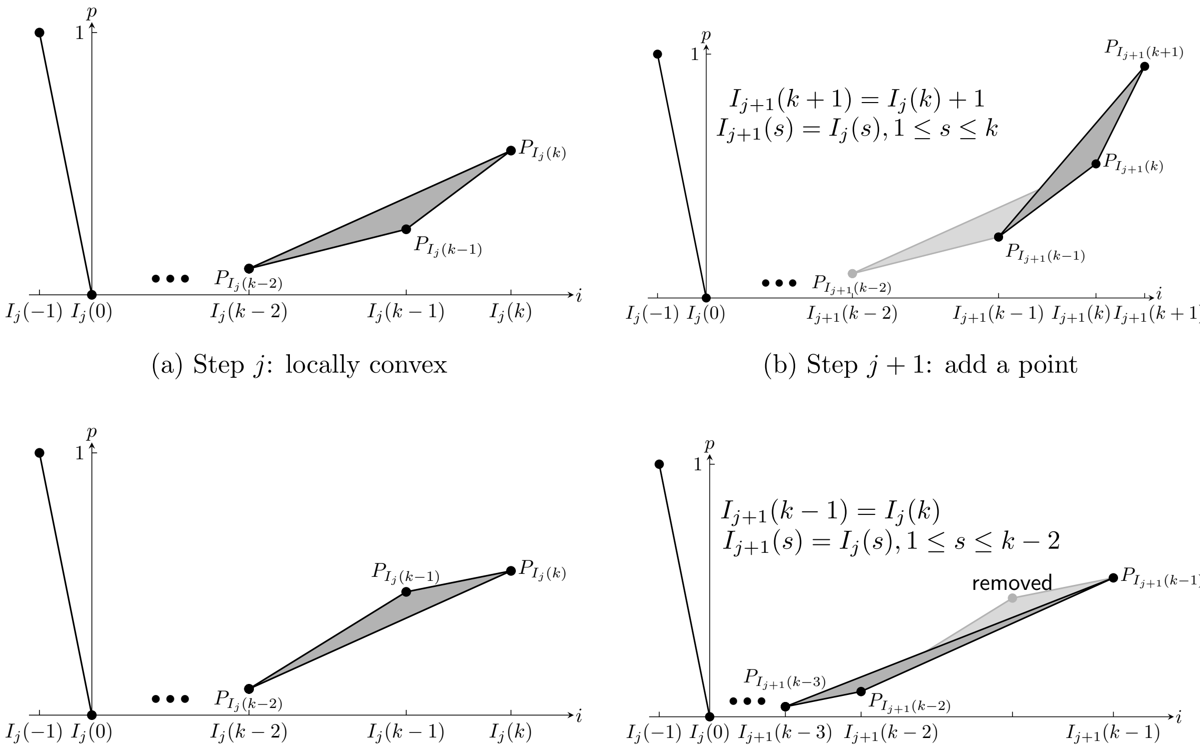
as shown in Figure 1-b. If three points $(I_j(k-2), P_{I_j(k-2)})$, $(I_j(k-1), P_{I_j(k-1)})$ and $(I_j(k), P_{I_j(k)})$ are locally concave in step j, as shown in Figure 1-c, which satisfy

$$\frac{P_{I_j(k)} - P_{I_j(k-1)}}{I_j(k) - I_j(k-1)} \leq \frac{P_{I_j(k-1)} - P_{I_j(k-2)}}{I_j(k-1) - I_j(k-2)}, \tag{3}$$

then $(I_j(k-1), P_{I_j(k-1)})$ will be removed and we next test whether three points $(I_j(k-3), P_{I_j(k-3)})$, $(I_j(k-2), P_{I_j(k-2)})$ and $(I_j(k), P_{I_j(k)})$ are locally convex or concave. The corresponding temporary convex hull in step $j+1$ in this case will include points with indices $\{I_{j+1}(l)\}_{l=-1}^{k-1}$ which are

$$I_{j+1}(k-1) = I_j(k) \ \text{ and } \ I_{j+1}(s) = I_j(s), 1 \leq s \leq k-2, \tag{4}$$

as shown in Figure 1-d. Noting that each step either includes one new point into consideration or reduces the number of possible points of the convex hull by one, we can show that this search takes at least $n-1$ comparisons and at most $2n-2$ comparisons to find the convex hull.



(a) Step j: locally convex

(b) Step $j+1$: add a point

(c) Step j: locally concave

(d) Step $j+1$: remove a point

Figure 1: Routine One: Steps of the convex hull searching algorithm

We develop a method using only simple operations to find the convex hull by hand. Starting from a table of function values with two rows, we list all p-values in an ascending order from right to left in the top row as the dependent variable, and list the corresponding orders of these p-values in the bottom row as the independent variable. The calculation starts from the right and checks three consecutive p-values and their orders each time, as indicated by solid circles, from the list of possible points of the convex hull. Taking the point in the middle as the reference point, if the approximate derivative by the forward difference is greater than that by the backward difference, as shown in (1), we move to the left and include a new p-value that is just greater than the current three p-values. The smallest p-value among the current three p-values is still in the list of possible points of the convex hull, but is not involved in the calculation in the next step, as indicated by dashed circles. Otherwise, if (3) holds, the point in the middle is removed from the list of possible points of the convex hull, and we move to the right and include the p-value that is just less than the current three p-values.

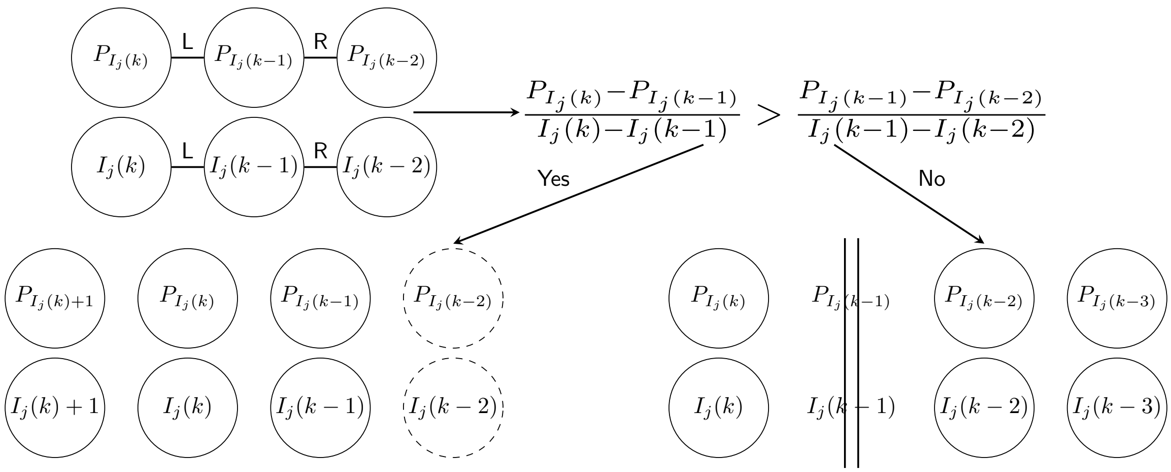


Figure 2: Illustration of the algorithm for determining the convex hull

Consider an example with four p-values: $P_1 = 0.01$, $P_2 = 0.03$, $P_3 = 0.04$, and $P_4 = 0.20$, as shown in Figure 3. Starting from three points $(0, P_0)$, $(1, P_1)$ and $(2, P_2)$, since $(0.03 - 0.01)/(2 - 1) > (0.01 - 0)/(1 - 0)$, we move to the left and consider $(1, P_1)$, $(2, P_2)$ and $(3, P_3)$. Note that $(0.04 - 0.03)/(3 - 2) < (0.03 - 0.01)/(2 - 1)$, we permanently remove $(2, P_2)$, move to the right and evaluate the triplet $(0, P_0)$, $(1, P_1)$ and $(3, P_3)$. Because $(0.04-0.01)/(3-1) > (0.01-0)/(1-0)$, we move to the left again and check $(1, P_1)$, $(3, P_3)$ and $(4, P_4)$. Since $(0.20-0.04)/(4-3) > (0.04-0.01)/(3-1)$ and we have already moved to the left end of the table, the searching process is completed and all the points left constitute the convex hull, which are $(1, P_1)$, $(3, P_3)$ and $(4, P_4)$.

Based on the convex hull determined in the first routine, the second routine finds the smallest integer j for which $\cup_{k=1}^{j}\{P_{n-j+k} \leq k\alpha/j\}$ is true. This routine is equivalent to search discrete tangent lines for all extreme points of the convex hull, which pass grid points $(i, 0)$ on the index axis where $i \in 1, \cdots, n$. Suppose the extreme points we found using the first routine are $\{(I(k), P_{I(k)})\}_{k=1}^{K}$. Starting from $I(K)$, for each extreme point $(I(k), P_{I(k)})$, we find the discrete tangent line that passes $(I(k), P_{I(k)})$ and $(J(k), 0)$, where $J(k)$ is given by

$$J(k) = \min\left\{j : \frac{P_{I(k)}}{I(k) - j} < \frac{P_{I(k-1)}}{I(k-1) - j}\right\}, \quad k = 2, \ldots, K, \tag{5}$$

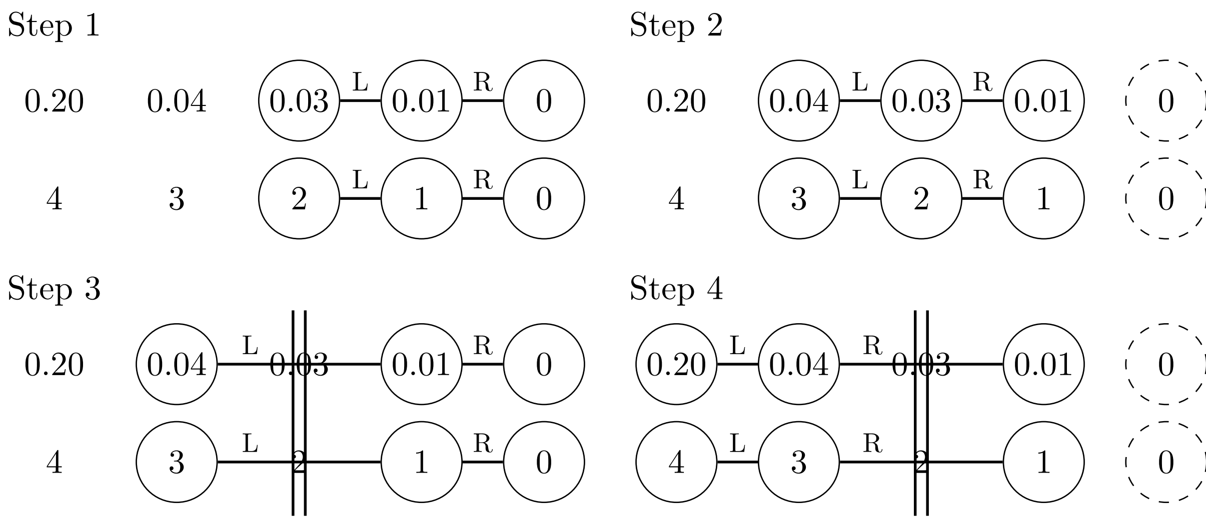


Figure 3: An example for illustrating the convex hull searching algorithm

and $J(1) = 0$, as shown in Figure 4. Note that $J(k) \leq I(k-1)$, and the searching process of $J(k)$ can start from $I(k-1)-1$. This routine takes at most $n-2$ comparisons to determine the intercept set $\{J(k)\}_{k=1}^{K}$. The dashed lines in Figure 4 indicate the convex hull that is determined in the first routine.

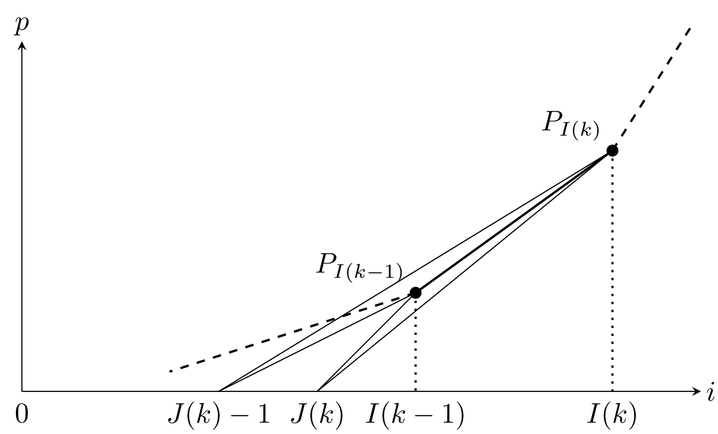


Figure 4: Routine Two: Steps of the discrete tangent line searching algorithm

We continue the example in Figure 3 where the extreme points of the convex hull are $(1, 0.01)$, $(3, 0.04)$ and $(4, 0.20)$. In this example, $I(1) = 1$, $I(2) = 3$, $I(3) = 4$, and $K = 3$. Starting from $(I(3), P_{I(3)}) = (4, 0.20)$, we consider $I(2) - 1 = 2$ for searching $J(3)$. Since $P_{I(3)}/(I(3) - 2) = 0.10 > 0.04 = P_{I(2)}/(I(2) - 2)$, the line passes through $(4, 0.20)$ and $(2, 0)$ is not a discrete tangent line. Since $J(3) \leq I(2)$ and always exists, we know that $J(3) = I(2) = 3$. Similarly, for $(I(2), P_{I(2)}) = (3, 0.04)$, we first check if the line passes through $(I(2), P_{I(2)})$ and $(I(1) - 1, 0)$ is a discrete tangent line. Since $P_{I(2)}/(I(2) - 0) = 0.013 > 0.010 = P_{I(1)}/(I(1) - 0)$, we achieve $J(2) = 1$. We complete the second routine for this example and have $(J(1), J(2), J(3)) = (0, 1, 3)$.

Using $\{I(k)\}_{k=1}^K$ from the first routine and $\{J(k)\}_{k=1}^K$ from the second routine, we locate the set of hypotheses to be rejected in the last routine. The third routine includes two subroutines. The first subroutine finds the index $j_\star = \min\{j : P_{n-j+k} \leq k\alpha/j, \exists k \in \{1, \dots, j\}\}$. Starting from the extreme point $(I(K), P_{I(K)})$, we search

$$J_\star = \min\left\{j : J(k) \leq j \leq J(k+1) - 1, \frac{P_{I(k)}}{I(k) - j} \leq \frac{\alpha}{n-j}, k \in \{1, \dots, K\}\right\}. \tag{6}$$

This process is equivalent to search a line that passes $(J_\star, 0)$ and $(I(k), P_{I(k)})$ and this line is just below point (n, α), as shown in Figure 5. Suppose $J(k) \leq J_\star \leq J(k+1) - 1$ and let $j_\star = n - J_\star$ and $k_\star = I(k) - J_\star$. We obtain $P_{n-j_\star+k_\star} \leq k_\star\alpha/j_\star$ and $j_\star$ is the smallest integer that makes this inequality holds.

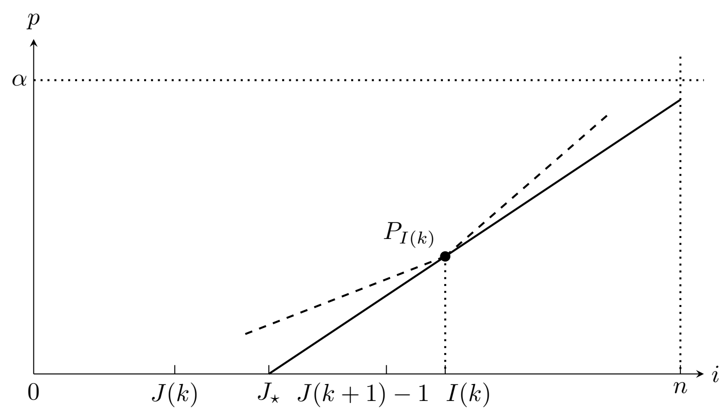


Figure 5: Routine Three: Steps of the critical value searching algorithm

The second part of the third routine is to find all hypotheses whose p-values are less than or equal to $\alpha/(j_\star - 1)$. In Hommel (1988)'s Hommel-Q and Meijer et al. (2019)'s Hommel-L methods, the searching process takes linear time from $P_{n-j_\star+k_\star-1}$ to the smallest p-value P_1, where indices $j_\star$ and $k_\star$ are determined in the first part of the third routine. Since p-values have been sorted, instead using linear search, a more efficient option is binary search. We consider using two variations of binary search. One is the standard binary search or binary chop, that eliminates the half in which the target value $P_m = \max_{1 \leq i \leq n-j_\star+k_\star-1}\{P_i \leq \alpha/(j_\star - 1)\}$ cannot lie in each iteration. This half-interval search runs in logarithmic time, and we call the corresponding Hommel procedure the Hommel-LS-B method. The other variation is interpolation search, where linear interpolation is used to estimate the position of the target value. When the p-values are distributed uniformly or almost uniformly, the interpolation search makes double logarithmic time complexity. We name the linear time Hommel procedure with interpolation search in the last routine as Hommel-LS-I.

In the worst case, Hommel-L/-LS-B/-LS-I methods take $2n - 2$ comparisons in the first routine for convex hull searching, $n-2$ comparisons in the second routine for discrete tangent line searching, and n comparisons in the last routine. So the total number of comparisons is bounded from above by $4n - 4$, where n is the number of hypotheses.

2.3 Rom procedure

The Hochberg procedure is a conservative shortcut procedure using the Simes test as the local test in a closed testing procedure. For an intersection hypothesis with k elementary hypotheses, the Simes test compares the ith largest p-value $p_{(n-i+1)}$ with the critical value $(n-i+1)\alpha/n$. In order to keep a simple sequential way of making inferences on individual hypotheses, the Hochberg procedure has to use the minimum critical value of all of the local Simes tests $\min_{k\geq i}\{(k-i+1)\alpha/k\} = \alpha/i$ for the ith largest p-value $p_{(n-i+1)}$. This makes the Hochberg procedure conservative, and the type I error is less than the nominal level α.

The conservativeness of the Hochberg procedure can be reduced if we can have a local test where the critical value $c_i\alpha$ for $p_{(n-i+1)}$ only depends on i. Since $c_i\alpha$ does not depend on the number of elementary hypotheses k in the local test, it follows that $\min_{k\geq i}\{c_i\alpha\} = c_i\alpha$, and the type I error is exactly at the nominal level α if all the local tests are exact.

Rom (1990) proposed a local test and its corresponding multple testing procedure. The critical constants c_i's satisfy the equation

$$\Pr\left(P_{(n)} > c_1\alpha, \cdots, P_{(n-i+1)} > c_i\alpha, \cdots, P_{(1)} > c_n\alpha\right) = 1-\alpha,$$

where c_i does not depend on the number of hypotheses n. Rom (1990) further obtained a recurrence formula to calculate the critical constant c_i based on $c_{i-1}, \ldots, c_1$, that is

$$c_i = \frac{1}{i}\left[\frac{1-\alpha^{i-1}}{1-\alpha} - \sum_{k=2}^{i-1}\binom{i}{k-1} c_k^{i-k+1}\alpha^{i-k}\right], \tag{7}$$

where $c_1 = 1$ and $c_2 = 1/2$. The Rom procedure follows the similar way of the Hochberg procedure to make inferences. The only difference is that the Rom procedure uses more liberal critical constants than the Hochberg procedure does. In order to explicitly show the difference between the Rom and Hochberg procedures, we calculate the asymptotic expansion of c_i, as shown in (8). The deviation details are included in the appendix.

$$c_i = \frac{1}{i}\left[1 + \frac{i-2}{2(i-1)}\alpha + \frac{(2i^2-7i+8)(i-3)}{6(i-1)(i-2)^2}\alpha^2\right] + o(\alpha^2) \tag{8}$$

We also give a simple asymptotic form of the critical values of the Rom procedure when i is large and α is small, as shown in (9). We call a Rom-type procedure with the critical values in (9) the Rom-A procedure, where the letter A stands for asymptotic.

$$\text{Rom-A:} \quad c_i{\cdot}\alpha = \frac{\log(1/(1-\alpha))}{i}. \tag{9}$$

This asymptotic expression in (9) also shows that the Hochberg procedure becomes more conservative when the nominal level α is larger, since $\log(1/(1-\alpha))/\alpha$ is an increasing function of α. Similarly, we can provide another Rom-type procedure based on the expansion in (8), where only the linear term of α is kept. We call this procedure with the critical values in (10) the Rom-1 procedure, where the number 1 means we use the first order approximations of the exact critical constants of the Rom procedure.

$$\text{Rom-1:} \quad c_i = \frac{1}{i}\left[1 + \frac{i-2}{2(i-1)}\alpha\right]. \tag{10}$$

Meanwhile, we refer to the original Rom procedure as the Rom-X procedure, where the letter X stands for exact. The critical constants of the Rom-1, Rom-X and Rom-A procedures are compared in Table 1, where $c_{i,\mathrm{Rom}}/c_{i,\mathrm{Hoch}} = i{\cdot}c_{i,\mathrm{Rom}}$ are reported for i ranging from 3 to 1000.

Table 1: Ratios of the critical constants of the Rom-1, Rom-X and Rom-A procedures to those of the Hochberg procedure for various i with $\alpha = 0.05$

i	3	10	30	100	300	1000
$c_{\text{Rom-1}}/c_{\text{Hoch}}$	1.0125	1.0222	1.0241	1.0247	1.0249	1.0250
$c_{\text{Rom-X}}/c_{\text{Hoch}}$	1.0125	1.0229	1.0250	1.0256	1.0258	1.0258
$c_{\text{Rom-A}}/c_{\text{Hoch}}$	1.0259	1.0259	1.0259	1.0259	1.0259	1.0259

Table 1 shows that the Rom-1 procedure is slightly more conservative than the original Rom procedure, and the Rom-A procedure with simple critical values $c_i\alpha = \log(1/(1-\alpha))/i$ is more liberal than the original Rom procedure, but the difference diminishes for large i.

The Rom procedure is less popular than the Hochberg procedure in practice, part of the reason is that the calculation of the exact critical constants in (7) is too complex for practitioners. To implement the Rom procedure, we recommend to use the Rom-1 procedure or a combination of the Rom-1 and Rom-A procedures, for example, a procedure with $c_i = 1/i{\cdot}\left[1 + (i-2)\alpha/(2(i-1))\right]$ for $i \le 5$ and $c_i = \log(1/(1-\alpha))/(i\alpha)$ for $i > 5$, which we refer to as the Rom-1A procedure.

2.4 Gou-Tamhane-Xi-Rom (GTXR) procedure

Gou et al. (2014) proposed a class of linear time stepwise multiple test procedures with prespecified critical constant $\{c_i\}_{i=1}^n$ and $\{d_i\}_{i=1}^n$ which satisfy $c_i \ge d_i$, $i = 1, \ldots, n$: let j be the smallest integer such that $p_{(n-j+1)} \le c_j\alpha$ and reject all hypotheses $H_{(i)}$ corresponding to $p_{(i)} \le d_j\alpha$. If no such j exists, accept all hypotheses. A procedure with $c_i = (i+1)/(2i)$ and $d_i = 1/i$ is named the GTXR-0 procedure, since neither c_i nor d_i depends on α and they are the zeroth-order approximations of the exact critical constants (Rom, 2013; Gou et al., 2014). Gou and Tamhane (2018a) discussed various choices of c_i and d_i, and the Hochberg (1988) and Rom (1990) procedures also belong to this class with critical constants $c_i = d_i$.

The GTXR procedure is a closed testing procedure that uses a local test for every intersection hypothesis. This local test rejects $\cap_{i=1}^n H_{(i)}$ if at least one $p_{(n-i+1)} \le c_i\alpha$ and at the same time $p_{(1)} \le d_i\alpha$. The probability of rejection is

$$P(\mathbf{c}, \mathbf{d}) = \Pr\left(\cup_{i=1}^n \left\{\left\{p_{(n-i+1)} \le c_i\alpha\right\} \cap \left\{p_{(1)} \le d_i\alpha\right\}\right\}\right), \tag{11}$$

and a proper choice of c_i and d_i should satisfy $P(\mathbf{c}, \mathbf{d}) \le \alpha$ for all $n \in \mathbb{Z}^+$.

Gou et al. (2014) provided an exact α-level local test with $d_i = 1/i$ prespecified, and c_i's are found by solving $P(\mathbf{c}, \mathbf{d}) = \alpha$ starting from $c_1 = 1$. The corresponding testing procedure is named as the GTXR-Xc procedure, where Xc stands for exact c_i critical constants. The expansion of the exact c_i is

$$c_i = \frac{i+1}{2i}{\cdot}\left[1 + \frac{\alpha}{6}\left(1 - \frac{1}{i+1} - \frac{1}{(i-1)^2} + \frac{1}{(i-1)^2(i+1)}\right)\right] + o(\alpha), \tag{12}$$

and the procedure using the first-order approximation to the exact c_i is called the GTXR-1c procedure correspondingly.

Besides prespecifying $d_i = 1/i$ and calculating c_i, an alternative way is to prespecify $c_i = (i+1)/(2i)$ and determine d_i via solving $P(\mathbf{c}, \mathbf{d}) = \alpha$. In this section, we propose two new GTXR procedures with $c_i = (i+1)/(2i)$. We first introduce a different notation for ordered p-values in order to indicate the sample size, where $p_{1:n} \le \cdots \le p_{n:n}$ denote the ordered p-values for a sample of size n. Then probabilities $C(i|n)$'s $(1 \le i \le n+1)$ involving critical constants c_i's are defined by

$$\begin{cases} C(i|n) = \Pr\left(\{\cap_{j<i} \{p_{n-j+1:n} > c_j\alpha\}\} \cap \{p_{n-i+1:n} \le c_i\alpha\}\right), \quad i = 1, \ldots, n, \\ C(n+1|n) = \Pr\left(\cap_{j=1}^{n} \{p_{n-j+1:n} > c_j\alpha\}\right), \end{cases} \tag{13}$$

where the probability $C(n+1|n) = 1 - \sum_{i=1}^{n} C(i|n)$. In addition, we define the probabilities $D(i|n)$'s $(1 \le i \le n)$ involving both c_i's and d_i's by

$$D(i|n) = \Pr\left(\{\cap_{j<i} \{p_{n-j+1:n} > c_j\alpha\}\} \cap \{p_{n-i+1:n} \le c_i\alpha\} \cap \{p_{1:n} \le d_i\alpha\}\right), \tag{14}$$

where $i = 1, \ldots, n$. It is clear that $P(\mathbf{c}, \mathbf{d}) = \sum_{i=1}^{n} D(i|n)$, where $P(\mathbf{c}, \mathbf{d})$ is defined in (11).

Exact local tests require that $\sum_{i=1}^{n} D(i|n) = \alpha$ for all n's. When the critical constants c_i's are prespecified, we can find the critical constant d_n when d_i's $(1 \le n-1)$ have been calculated:

$$d_n = \frac{\alpha - \sum_{i=1}^{n-1} D(i|n)}{\alpha n(1 - \sum_{i=1}^{n-1} C(i|n-1))}, \tag{15}$$

where the terms $C(i|n-1)$'s and $D(i|n)$'s can be calculated using the recursive relation by Finner and Roters (1994) and Cai and Sarkar (2008)

$$C(i|n) = \frac{\alpha n c_i}{n-i+1} \cdot C(i|n-1), \quad i = 1, \ldots, n, \tag{16}$$

and the recursive relation by Gou et al. (2014)

$$D(i|n) = \frac{\alpha n \left(c_i^{n-i+1} - (c_i - d_i)^{n-i+1}\right)}{(n-i+1)c_i^{n-i}} \cdot C(i|n-1), \quad i = 1, \ldots, n, \tag{17}$$

starting from $c_1 = d_1 = 1$ and $C(1|1) = D(1|1) = \alpha$. Moreover, these recursive relations can be generalized for other applications (Zhang and Gou, 2016). The procedure with d_i in (15) is called the GTXR-Xd procedure, where Xd stands for exact d_i critical constants. The expansion of d_i is

$$d_i = \frac{1}{i} \cdot \left[1 + \frac{\alpha^2}{12}\left(1 - \frac{1}{(i-2)^2}\right)\right] + o(\alpha^2), \quad i \ge 3, \tag{18}$$

where the coefficient of the linear term of α is zero. The procedure using the second-order approximation to the exact d_i is called the GTXR-2d procedure. Details of deviations of (15) and (18) are provided in the appendix.

GTXR procedure makes $n+1$ comparisons in the worst case. The search of $j_\star = \min\{j : p_{(n-j+1)} \le c_j\alpha\}$ takes linear time starting from the largest p-value. For the other search of $i_\star = \max_{1 \le i \le n - j_\star + 1}\{i : p_{(i)} \le d_{j_\star}\alpha\}$, either standard binary search or interpolation search can be applied, and the corresponding procedures are denoted by GTXR-B or GTXR-I.

2.5 Quick procedure

The implementation of the GTXR procedures includes two stages: the comparison and the determination stages. In the comparison stage, the ordered p-values are compared with the corresponding critical values from the largest to the smallest to find the smallest index i which satisfies $p_{(n-i+1)} \leq c_i\alpha$. The determination stage is followed, rejecting all p-values which are less than $d_i\alpha$, where the index i is selected in the comparison step. The comparison step takes linear time, and the determination step takes logarithmic time with a binary search or an interpolation search algorithm.

We propose a new family of stepwise multiple test procedures to further reduce the time complexity from linear time to logarithmic time. These procedures include a logarithmic comparison step and a logarithmic determination step, which we refer to as the Quick procedures.

With the Quick procedure, we first reject all hypotheses if all p-values are less than the nominal level α. Next, we find the index $i = \min\{i' : p_{(n-i'+1)} \leq c\alpha\}$. Lastly, we reject all hypotheses whose p-values are less than or equal to α/i. Using binary search, both the comparison and determination steps can be completed in logarithmic time. The Quick procedure with standard binary search is named as the Quick-B procedure, and the number of comparison is $2\lfloor\log_2(n-1)\rfloor+3$ in the worst case, where $\lfloor\cdot\rfloor$ is the floor function. Alternatively, the Quick-I procedure with interpolation search algorithm can even be completed in double logarithmic time under some perfect or nearly perfect circumstances.

The Quick procedures belong to the family of the closed testing procedures. The type I error rate of the local test with n elementary hypotheses is

$$P_n(c) = \Pr\left(\left\{p_{(n)} \leq \alpha\right\} \cup \left[\cup_{i=2}^{n} \left\{\left\{p_{(n-i+1)} \leq c\alpha\right\} \cap \left\{p_{(1)} \leq \alpha/i\right\}\right\}\right]\right), \tag{19}$$

which is an increasing function of n when c is fixed. In order to test individual hypotheses and control the familywise error rate (FWER) under the nominal level α, the critical constant c need to satisfy $P_n(c) \leq \alpha$, where n is the number of individual hypotheses. Omitting the subscript n, the expression of $P(c)$ is

$$P(c) = \alpha^n + \sum_{r=1}^{n-1} (1-\alpha)^r \alpha^{n-r} \binom{n}{r} \sum_{j=r+1}^{n} \binom{n-r}{j-r-1} (1-c)^{j-r-1} \left[c^{n-j+1} - \left(c - \frac{1}{j}\right)^{n-j+1}\right], \tag{20}$$

and the deviation details are provided in the appendix. We refer to the Quick procedure using the largest possible constant c as the Quick-X procedure, where the letter X stands for exact. Therefore, the critical constant c in the Quick-X procedure can be found by solving

$$c = \arg\left\{P(c) = \alpha\right\}. \tag{21}$$

Although there is no closed-form solution for the exact c value, we compute its asymptotic expansion, as shown in (22).

$$c = \frac{n}{2(n-1)} + \frac{\alpha}{12} \cdot \left[1 + \frac{3}{n-1} + \frac{2}{(n-2)^2} - \frac{6}{(n-1)(n-2)^2}\right] + o(\alpha). \tag{22}$$

Details of deviations are presented in the appendix. Besides using the exact c value from (21), we can apply some approximate c's with explicit expressions based on the expansion in (22). For example, we can take

$$c = 1/2$$

and name the corresponding procedure as the Quick-00 procedure, where the double zeros indicate the critical constant c depends neither on the nominal level α nor on the number of hypotheses n. Similarly, we have the Quick-01 procedure, where the critical constant

$$c = n/(2(n-1))$$

does not depend on α, and the Quick-10 procedure, where the critical constant

$$c = 1/2 + \alpha/12$$

does not depend on n. The Quick-11 procedure is defined with the critical constant

$$\begin{cases} c = 1, \ n = 2, \quad c = 3/4, \ n = 3, \quad c = 2/3 + \alpha/12, \ n = 4, \\ c = \frac{n}{2(n-1)} + \frac{\alpha}{12} \cdot \left[1 + \frac{3}{n-1} + \frac{2}{(n-2)^2} - \frac{6}{(n-1)(n-2)^2} \right], \ n \geq 5, \end{cases}$$

using the linear terms of α in the expansion in (22).

2.6 Relations among procedures

A general power comparisons for the stepwise multiple testing procedures are summarized in Figure 6, where an arrow from procedure A to B indicates that A is uniformly more powerful than B. Meanwhile, we underline the newly proposed stepwise procedures in this paper.

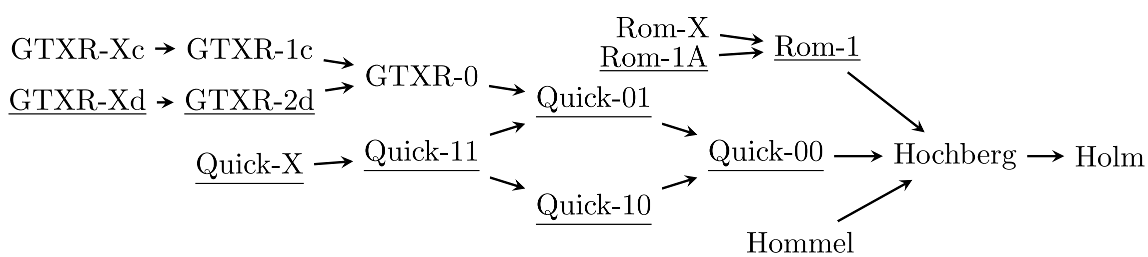


Figure 6: Procedure relationship diagram for power comparisons.

The time complexities of these procedures include quadratic time, linear time and logarithmic time. The numbers of comparisons of the Holm, Hochberg, Rom and GTXR procedures are about n. The Hommel-Q procedure takes about $n^2/2$ comparisons, and the Hommel-L procedure takes about $4n$. The Quick procedures only need about $2 \log_2 n$ comparisons.

3 General quick procedures

In this section we discuss general linear and logarithmic time p-value based multiple test procedures. The following theorem first asserts a sufficient condition for linear time closed testing procedure.

Theorem 1 (Existence of linear time shortcuts)**.** *Consider a closed testing procedure based on the local test with the type I error control* $\Pr\left(\cup_{i=1}^{n}\left\{p_{(i)} \leq a_{in}\right\}\right) \leq \alpha$. *This multiple test procedure has an exact shortcut in linear time if any of the following conditions is satisfied: (1)* $a_{in} = g(n)$, *(2)* $a_{in} = g(n-i)$, *or (3)* $a_{in} = (g(n) - g(n-i)) \cdot h(n)$, *where* $g(\cdot)$ *is a univariate monotonic function.*

For example, the Holm procedure satisfies $a_{in} = g(n)$ where $g(k) = \alpha/k$ and the Hochberg procedure satisfies another condition $a_{in} = g(n-i)$ where $g(k) = \alpha/(k+1)$. The Hommel (1988) and robust Hommel (1983) procedures both satisfy the last condition $a_{in} = (g(n) - g(n-i)) \cdot h(n)$ where $g(k) = k$ and $h(k) = \alpha/k$ for the Hommel (1988) procedure and $h(k) = \alpha/(kH_k)$ for the robust Hommel (1983) procedure, where H_k is the k-th harmonic number that $H_k = \sum_{i=1}^{k}(1/i)$.

Cai and Sarkar (2008) and Gou and Tamhane (2014) have provided several generalized Simes tests, and some Hommel-like procedures in quadratic time can be constructed based on these tests (Gou and Sarkar, 2019). In order to achieve linear time Hommel-type multiple testing procedures, a starting point is to extend the Simes test whose critical values follow the conditions in Theorem 1.

Quick procedures run logarithmic time. The following theorem provides a broad class of logarithmic-time p-value based multiple test procedures based on the Quick procedures. For instance, the Quick-00 procedure is a special case of the general procedure in Theorem 2, using $i_\star = 1$, $a_1 = \alpha$, $a = \alpha/2$ and $b_i = \alpha/i$.

Theorem 2 (Existence of logarithmic time shortcuts)**.** *A multiple test procedure has a logarithmic shortcut if its intersection hypothesis test is based on the inequality of the type I error control* $\Pr\left(\cup_{i=1}^{n}\left\{\left\{p_{(n-i+1)} \leq a_i\right\} \cap \left\{p_{(1)} \leq b_i\right\}\right\}\right) \leq \alpha$, *where critical values* $a_i \equiv a$ *are constant for all* $i > i_\star$, *and* $i_\star$ *is a finite integer that does not depend on* n.

4 Adjusted p-values

Comparing with the significant-nonsignificant dichotomy for each hypothesis under a given significance level, the adjusted p-values for simultaneous inference are considered to be more informative (Rosenthal and Rubin, 1983). The adjusted p-values for the Holm procedure are $\widetilde{p}_{(n-i+1)} = \max_{i \leq j \leq n}\{j \cdot p_{(n-j+1)}\}$, and the adjusted p-values for the Hochberg procedure are $\widetilde{p}_{(n-i+1)} = \min_{1 \leq j \leq i}\{j \cdot p_{(n-j+1)}\}$. Wright (1992) provided a cubic time and a quadratic time methods to calculate the adjusted p-values for the Hommel procedure. The linear time method for the Hommel procedure has recently been developed by Meijer et al. (2019). For the Rom procedures, there is no closed-form solution of the adjusted p-values for the Rom-X procedure. The newly proposed Rom-1 and Rom-A procedures have the closed-form adjusted

p-values:

$$\text{Rom-1:}\;\; \widetilde{p}_{(n-i+1)} = \begin{cases} p_{(n)} \;\text{ for } i=1, \quad \min\{p_{(n)}, 2p_{(n-1)}\} \;\text{ for } i=2, \\ \min\left\{\widetilde{p}_{(n-i+2)}, \frac{i-1}{i-2}\left(\sqrt{1+\frac{i-2}{i-1}\cdot 2i\cdot p_{(n-i+1)}}-1\right)\right\} \;\text{ for } i=3,\ldots,n, \end{cases}$$
$$\text{Rom-A:}\;\; \widetilde{p}_{(n-i+1)} = \min_{1\le j\le i}\left\{1-\exp\left(-jp_{(n-j+1)}\right)\right\} \;\text{ for } i=1,\ldots,n. \tag{23}$$

The adjusted p-values for the GTXR-0 procedure given by Gou et al. (2014) are $\widetilde{p}_{(n-i+1)} = \min_{1\le j\le i}\{\max\{2jp_{(n-j+1)}/(j+1), jp_{(n-i+1)}\}\}$. When the level of significance is small, the adjusted p-values for the GTXR-1c and GTXR-2d procedures are:

$$\text{GTXR-1c:}\;\; \widetilde{p}_{(n-i+1)} = \min_{1\le j\le i}\left\{\max\left\{\frac{2jp_{(n-j+1)}}{j+1} - \frac{2(j-2)j^4p^2_{(n-j+1)}}{3(j-1)^2(j+1)^3}, jp_{(n-i+1)}\right\}\right\}, \tag{24}$$
$$\text{GTXR-2d:}\;\; \widetilde{p}_{(n-i+1)} = \min_{1\le j\le i}\left\{\max\left\{\frac{2jp_{(n-j+1)}}{j+1}, jp_{(n-i+1)} - \frac{(j-3)(j-1)j^3p^3_{(n-i+1)}}{12(j-2)^2}\right\}\right\}.$$

For the Quick-00 procedure, the adjusted p-values are:

$$\text{Quick-00:}\;\; \widetilde{p}_{(n-i+1)} = \min\left\{p_{(n)}, \min_{2\le j\le i}\left\{\max\left\{2p_{(n-j+1)}, jp_{(n-i+1)}\right\}\right\}\right\}. \tag{25}$$

The expressions of the adjusted p-values for the Quick-01, Quick-10, and Quick-11 procedures are listed in the appendix.

5 Numerical comparisons

We conduct a simulation study for studying the FWER, statistical power and number of comparisons under independence and dependence. We first generate n normal random variables $\{z_i\}_{i=1}^n$ with common correlation ρ, where n_0 normally distributed z_is follow $N(0,1^2)$ and the other $n-n_0$ random variables z_is follow $N(\delta,1^2)$. Next the marginal p-values are calculated by $p_i = 1-\Phi(z_i)$, where $\Phi(\cdot)$ denotes the standard normal distribution function. The Holm, Hochberg, Hommel, Rom, GTXR and Quick procedures are evaluated for sample sizes of $n=10$ and 1000 at a nominal $\alpha=0.05$, where the proportions of true null hypotheses n_0/n are 0.2 and 0.8, and the common correlation coefficients ρ are 0 for independence and 0.5 for dependence. The number of replica is 10^8 if the sample size $n=10$ or the test statistics are independent. Only when $n=1000$ and $\rho=0.5$ do we run 10^5 simulations instead of 10^8.

Table 2 and 3 indicate that the actual level of significance of the GXTR and Quick procedures are greater than that of the Holm, Hochberg, Hommel and Rom procedures. This can be interpreted as the GTXR and Quick procedures are less conservative than the other procedures. Simulated FWERs are all below the nominal level $\alpha=5\%$.

There are various definitions of statistical power when applying multiple test procedures. We evaluate these procedures in terms of the expected true positive rate, which is the expected proportion of rejected true significances. This expected value is also called the

Table 2: Familywise error rates (%) of Holm, Hochberg, Hommel, Rom, GTXR, and Quick procedures under independence ($\alpha = 5\%$)

$\rho = 0$	$n = 10$				$n = 1000$			
	$n_0/n = 0.2$		$n_0/n = 0.8$		$n_0/n = 0.2$		$n_0/n = 0.8$	
Procedure	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$
Holm	1.399	4.541	4.185	4.814	1.020	1.932	3.949	4.399
Hochberg	1.470	4.729	4.198	4.822	1.020	1.933	3.949	4.399
Hommel	1.704	4.802	4.294	4.874	1.178	3.491	3.977	4.606
Rom-1	1.505	4.736	4.291	4.925	1.045	1.990	4.043	4.508
Rom-X	1.506	4.736	4.294	4.928	1.046	1.992	4.046	4.512
Rom-1A	1.511	4.736	4.308	4.945	1.046	1.992	4.046	4.512
GTXR-0	1.939	4.828	4.507	4.978	1.707	4.621	4.459	4.973
GTXR-1c/Xc	1.943	4.828	4.509	4.979	1.710	4.625	4.461	4.974
GTXR-2d/Xd	1.940	4.828	4.508	4.979	1.707	4.622	4.460	4.974
Quick-00	1.820	4.738	4.474	4.957	1.706	4.619	4.458	4.973
Quick-01	1.873	4.764	4.503	4.973	1.706	4.620	4.459	4.973
Quick-10	1.824	4.740	4.476	4.958	1.709	4.623	4.461	4.974
Quick-11/X	1.878	4.767	4.506	4.975	1.710	4.624	4.461	4.974

Table 3: Familywise error rates (%) of Holm, Hochberg, Hommel, Rom, GTXR, and Quick procedures under common positive correlation ($\alpha = 5\%$)

$\rho = 0.5$	$n = 10$				$n = 1000$			
	$n_0/n = 0.2$		$n_0/n = 0.8$		$n_0/n = 0.2$		$n_0/n = 0.8$	
Procedure	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$
Holm	3.183	4.529	3.559	3.749	0.829	2.158	1.572	1.784
Hochberg	3.675	4.741	3.583	3.761	0.829	2.158	1.572	1.784
Hommel	3.908	4.742	3.776	3.900	1.666	2.223	1.709	1.799
Rom-1	3.691	4.741	3.654	3.834	0.851	2.202	1.611	1.821
Rom-X	3.692	4.741	3.656	3.836	0.853	2.204	1.612	1.822
Rom-1A	3.692	4.741	3.667	3.849	0.853	2.205	1.612	1.822
GTXR-0	4.010	4.743	4.091	4.156	2.357	2.627	2.138	2.151
GTXR-1c/Xc	4.010	4.743	4.095	4.159	2.362	2.633	2.145	2.155
GTXR-2d/Xd	4.010	4.743	4.092	4.157	2.357	2.629	2.139	2.151
Quick-00	3.803	4.741	4.009	4.080	2.354	2.627	2.138	2.151
Quick-01	3.863	4.742	4.063	4.129	2.354	2.627	2.138	2.151
Quick-10	3.808	4.741	4.014	4.084	2.358	2.627	2.144	2.154
Quick-11/X	3.869	4.742	4.069	4.134	2.359	2.627	2.145	2.155

average power or per-pair power (Horn and Dunnett, 2004; Porter, 2018). We compare the average power of the Holm, Hochberg, Hommel, Rom, GTXR and Quick procedures under independence and dependence in Table 4 and 5. The GTXR and Quick procedures, which have similar average power, are always more powerful than the other procedures.

Table 4: Powers (%) of Holm, Hochberg, Hommel, Rom, GTXR, and Quick procedures under independence ($\alpha = 5\%$)

$\rho = 0$	$n = 10$				$n = 1000$			
	$n_0/n = 0.2$		$n_0/n = 0.8$		$n_0/n = 0.2$		$n_0/n = 0.8$	
Procedure	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$
Holm	31.49	96.72	28.65	92.78	2.97	60.80	2.94	55.49
Hochberg	31.82	96.80	28.67	92.79	2.97	60.80	2.94	55.49
Hommel	33.24	96.88	28.82	92.83	3.21	66.51	2.96	55.93
Rom-1	32.12	96.84	28.93	92.89	3.01	61.08	2.98	55.73
Rom-X	32.13	96.84	28.94	92.90	3.02	61.09	2.99	55.74
Rom-1A	32.18	96.84	28.98	92.91	3.02	61.09	2.99	55.74
GTXR-0	34.90	96.92	29.22	92.91	3.94	69.26	3.15	56.69
GTXR-1c/Xc	34.92	96.92	29.22	92.92	3.95	69.27	3.15	56.70
GTXR-2d/Xd	34.90	96.92	29.22	92.92	3.94	69.26	3.15	56.69
Quick-00	34.36	96.87	29.16	92.90	3.94	69.25	3.15	56.69
Quick-01	34.66	96.88	29.21	92.91	3.94	69.26	3.15	56.69
Quick-10	34.38	96.87	29.16	92.90	3.95	69.26	3.15	56.69
Quick-11/X	34.69	96.89	29.22	92.91	3.95	69.26	3.15	56.69

Table 6 and 7 summarize the average numbers of comparisons of all multiple test procedures. For the procedures involving binary search algorithm, two versions of methods are provided, indicated by the letter I for interpolation search and by letter B for standard binary search. Although an interpolation search algorithm can achieve an $O(\log \log n)$ average performance in some ideal situations, the worst-case performance of the interpolation search is $O(n)$. Alternatively, the average and worst-case performances of a standard binary search are both $O(\log n)$. So the standard binary search has a more stable time performance comparing with the interpolation search. Results show that the Quick-B procedure needs much fewer number of comparisons in most situations. The reduction of number of comparisons is notably significant when the number of hypotheses n is large.

6 Discussion and concluding remarks

Liu (1996) gave a representation for the stepwise multiple test procedures using an $n \times n$ lower-triangular matrix called C-matrix. For a step-up procedure, a matrix element c_{ij} is the critical constant used in step i and compared with the jth largest p-value $p_{(n-j+1)}$. The C-matrices of the Hochberg, Hommel, GTXR-0 and Quick-00 procedures with four hypotheses ($n = 4$) are given in Table 8. The number of distinct elements in the C-matrix indicates the time complexity of the corresponding multiple test procedure. For a C-matrix with $O(n^2)$

Table 5: Powers (%) of Holm, Hochberg, Hommel, Rom, GTXR, and Quick procedures under common positive correlation ($\alpha = 5\%$)

$\rho = 0.5$	$n = 10$				$n = 1000$			
	$n_0/n = 0.2$		$n_0/n = 0.8$		$n_0/n = 0.2$		$n_0/n = 0.8$	
Procedure	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$
Holm	32.95	95.96	28.74	92.69	3.18	60.48	2.97	55.16
Hochberg	33.60	96.11	28.77	92.70	3.18	60.48	2.97	55.16
Hommel	35.39	96.38	28.96	92.73	4.58	66.70	3.09	55.70
Rom-1	33.87	96.17	29.04	92.80	3.22	60.75	3.01	55.40
Rom-X	33.88	96.17	29.04	92.80	3.23	60.75	3.01	55.41
Rom-1A	33.92	96.17	29.08	92.82	3.23	60.75	3.01	55.41
GTXR-0	37.01	96.53	29.41	92.77	5.83	68.99	3.50	56.28
GTXR-1c/Xc	37.03	96.54	29.41	92.77	5.84	68.99	3.50	56.28
GTXR-2d/Xd	37.01	96.53	29.41	92.77	5.83	68.99	3.50	56.28
Quick-00	36.28	96.43	29.33	92.77	5.83	68.98	3.50	56.28
Quick-01	36.63	96.48	29.40	92.77	5.83	68.98	3.50	56.28
Quick-10	36.31	96.44	29.34	92.77	5.84	68.99	3.50	56.28
Quick-11/X	36.67	96.48	29.40	92.77	5.84	68.99	3.50	56.28

distinct elements, the time complexity of its corresponding procedure is at most $O(n^2)$. For a C-matrix with $O(n)$ distinct elements, its corresponding procedure takes linear time $O(n)$. A procedure can be achieved in logarithmic time $O(\log n)$ if its C-matrix has only $O(1)$ distinct elements. The Hommel-L method is an example of a linear time procedure with an $O(n^2)$-element C-matrix. It is also worth determining whether a logarithmic time procedure with an $O(n)$-element C-matrix exists or not.

We discuss the linear time FWER-control stepwise multiple test procedures and develop their new extensions in this paper. Other types of p-value based procedures can be found in a general review paper by Tamhane and Gou (2018). The procedures of Holm, Hochberg, Hommel-Q have been included in an R function `p.adjust` from package `stats` and also in another R function `mt.rawp2adjp` from package `multtest`. The R function `rom` in package `mutoss` has applied the Rom-X procedure, and the R function `hommel` in package `hommel` has provided the Hommel-L method. For the aim of forming a comprehensive collection, we implement the existing and newly proposed p-value based multiple testing procedures, including the GTXR and Quick procedure families, in the R package `elitism` (equipment for logarithmic and linear time stepwise multiple hypothesis testing) which is available on CRAN.

Appendix

Derivations of Equation (8), (15), (18), (20), (22), (23) and (24) and proofs of Theorems 1 and 2 are included in this appendix.

Derivation of Equation (8). Since the critical constants in the Rom procedure is $c_i = 1/i +$

Table 6: Numbers of comparisons of Holm, Hochberg, Hommel, Rom, GTXR, and Quick procedures under independence ($\alpha = 5\%$)

$\rho = 0$	$n = 10$				$n = 1000$			
	$n_0/n = 0.2$		$n_0/n = 0.8$		$n_0/n = 0.2$		$n_0/n = 0.8$	
Procedure	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$
Holm	3.5	8.8	1.6	2.9	24.8	487.4	6.9	112.0
Hochberg	8.4	3.2	9.9	9.1	977.2	514.6	995.1	890.0
Hommel-Q	30.8	4.7	45.0	37.8	363730.8	40745.0	487417.6	361527.1
Hommel-L	28.4	21.3	26.9	25.3	3908.0	3322.5	3393.0	3257.3
Hommel-LS-I	28.3	21.3	26.9	25.3	3700.3	3162.5	3379.5	3224.2
Hommel-LS-B	28.3	21.3	26.9	25.3	3695.0	3111.5	3379.6	3209.9
Rom-1/X/1A	8.3	3.2	9.9	9.1	976.9	512.3	995.0	889.5
GTXR-0-I	8.7	4.3	10.6	10.1	602.8	340.5	885.1	886.2
GTXR-1c/Xc-I	8.7	4.3	10.6	10.1	601.7	341.0	884.6	886.7
GTXR-2d/Xd-I	8.7	4.3	10.6	10.1	602.8	340.5	885.1	886.2
GTXR-0-B	8.7	4.4	10.6	10.1	590.8	221.2	883.9	791.8
GTXR-1c/Xc-B	8.7	4.4	10.6	10.1	589.7	221.0	883.5	791.6
GTXR-2d/Xd-B	8.7	4.4	10.6	10.1	590.8	221.2	883.9	791.8
Quick-00-I	7.5	8.9	4.0	5.4	62.3	192.9	19.3	114.2
Quick-01-I	7.6	8.7	4.2	5.4	62.3	192.9	19.3	114.3
Quick-10-I	7.5	8.9	4.1	5.4	62.2	193.1	19.3	114.8
Quick-11/X-I	7.6	8.7	4.2	5.4	62.2	193.1	19.3	114.9
Quick-00-B	5.8	5.2	4.5	5.4	17.9	18.8	16.0	16.6
Quick-01-B	5.9	5.2	4.6	5.4	17.9	18.8	16.0	16.6
Quick-10-B	5.8	5.2	4.5	5.4	17.9	18.8	16.0	16.6
Quick-11/X-B	5.9	5.2	4.6	5.4	17.9	18.8	16.0	16.6

Table 7: Numbers of comparisons of Holm, Hochberg, Hommel, Rom, GTXR, and Quick procedures under common positive correlation ($\alpha = 5\%$)

$\rho = 0.5$	$n = 10$				$n = 1000$			
	$n_0/n = 0.2$		$n_0/n = 0.8$		$n_0/n = 0.2$		$n_0/n = 0.8$	
Procedure	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$	$\delta = 2$	$\delta = 4$
Holm	3.7	8.7	1.6	2.9	26.4	484.9	7.0	111.4
Hochberg	8.0	3.2	9.8	9.1	975.2	517.1	994.5	890.6
Hommel-Q	29.9	5.0	44.1	37.6	353498.1	43167.2	479870.8	359813.7
Hommel-L	28.0	21.4	25.7	24.5	3881.7	3346.8	3571.1	3464.2
Hommel-LS-I	27.9	21.4	25.7	24.5	3657.6	3159.3	3546.4	3421.5
Hommel-LS-B	27.9	21.4	25.7	24.5	3652.6	3128.9	3545.7	3411.7
Rom-1/X/1A	7.9	3.2	9.8	9.1	974.9	514.9	994.5	890.1
GTXR-0-I	8.2	4.2	10.4	10.0	598.8	339.5	884.3	880.6
GTXR-1c/Xc-I	8.1	4.2	10.4	10.0	597.7	339.9	883.8	880.9
GTXR-2d/Xd-I	8.1	4.2	10.4	10.0	598.8	339.5	884.3	880.6
GTXR-0-B	8.2	4.3	10.3	10.0	591.0	221.1	882.4	791.8
GTXR-1c/Xc-B	8.2	4.3	10.3	10.0	589.9	221.0	881.9	791.6
GTXR-2d/Xd-B	8.2	4.3	10.3	10.0	591.0	221.1	882.4	791.8
Quick-00-I	6.3	9.1	3.7	5.2	56.7	187.2	17.7	109.7
Quick-01-I	6.4	8.9	3.8	5.2	56.6	187.3	17.7	109.8
Quick-10-I	6.3	9.1	3.7	5.2	56.6	187.4	17.7	110.1
Quick-11/X-I	6.4	8.9	3.8	5.2	56.5	187.5	17.7	110.1
Quick-00-B	5.3	5.1	4.0	5.2	16.1	18.8	14.1	16.9
Quick-01-B	5.3	5.1	4.1	5.3	16.1	18.8	14.1	16.9
Quick-10-B	5.3	5.1	4.0	5.2	16.1	18.8	14.1	16.9
Quick-11/X-B	5.3	5.1	4.1	5.3	16.1	18.8	14.1	16.9

Table 8: C-matrix representation of stepwise multiple test procedures ($n = 4$)

Hochberg

$$\begin{pmatrix} 1 & & & \\ 1 & 1/2 & & \\ 1 & 1/2 & 1/3 & \\ 1 & 1/2 & 1/3 & 1/4 \end{pmatrix}$$

Hommel

$$\begin{pmatrix} 1 & & & \\ 1 & 1/2 & & \\ 1 & 2/3 & 1/3 & \\ 1 & 3/4 & 1/2 & 1/4 \end{pmatrix}$$

GTXR-0

$$\begin{pmatrix} 1 & & & \\ 1 & 3/4 & & \\ 1 & 3/4 & 2/3 & \\ 1 & 3/4 & 2/3 & 5/8 \end{pmatrix}$$

Quick-00

$$\begin{pmatrix} 1 & & & \\ 1 & 1/2 & & \\ 1 & 1/2 & 1/2 & \\ 1 & 1/2 & 1/2 & 1/2 \end{pmatrix}$$

$o(1)$, based on (7), we obtain

$$c_i = (1/i)\cdot[1+\alpha - i(i-1)/2\cdot(1/(i-1))^2\alpha] + o(\alpha)$$
$$= \frac{1}{i}\cdot\left[1+\frac{i-2}{2(i-1)}\alpha\right] + o(\alpha).$$

From the equation above, we further obtain

$$c_i = \frac{1}{i}\cdot\left[1+\alpha+\alpha^2 - \binom{i}{i-2}\left(\frac{1}{i-1}\left(1+\frac{i-3}{2(i-2)}\alpha\right)\right)^2\alpha - \binom{i}{i-3}\frac{1}{(i-2)^3}\alpha^2\right] + o(\alpha^2)$$
$$= \frac{1}{i}\cdot\left[1+\left(1-\frac{i}{2(i-1)}\right)\alpha + \left(1-\frac{i(i-3)}{2(i-1)(i-2)} - \frac{i(i-1)}{6(i-2)^2}\right)\alpha^2\right] + o(\alpha^2),$$

which yields Equation (8). □

Derivation of Equation (15). The recursive relation (17) gives that $D(n|n) = \alpha n d_n \cdot C(n|n-1)$. We rewrite this recursive formula and have

$$d_n = \frac{D(n|n)}{\alpha n C(n|n-1)}.$$

To get (15), we substitute $D(n|n) = \alpha - \sum_{i=1}^{n-1} D(i|n)$ and $C(n|n-1) = 1 - \sum_{i=1}^{n-1} C(i|n-1)$ into the equation above. □

Derivation of Equation (18). We first note that $C(i|n) = O(\alpha^{n-i+1})$, $D(i|n) = O(\alpha^{n-i+1})$, and $d_n = 1/n + o(1)$ when $c_i = (i+1)/(2i)$. In addition, we find that $C(n-1|n-1) = \alpha_{n-1}\cdot(n-1)\cdot C(n-1|n-2) = n\alpha/2 + o(\alpha)$. Using the recursive formula (16) and (17), we can get

$$d_n = \frac{\alpha - D(n-1|n)}{\alpha n\left[1 - C(n-1|n-1)\right]} + o(\alpha) = \frac{1}{n}\cdot\frac{1 - n\alpha/2 + o(\alpha)}{1 - n\alpha/2 + o(\alpha)} + o(\alpha) = \frac{1}{n} + o(\alpha)$$

based on (15). Next we calculate

$$C(n-1|n-1) = \frac{n\alpha}{2}C(n-1|n-2) = \frac{n\alpha}{2}\left[1 - \frac{(n-1)\alpha}{2}\right] + o(\alpha^2),$$
$$C(n-2|n-1) = \alpha\cdot\frac{n-1}{2(n-2)}\cdot\frac{n-1}{2}\cdot C(n-2|n-2) = \frac{(n-1)^3\alpha^2}{8(n-1)} + o(\alpha^2),$$
$$D(n-1|n) = \alpha\cdot C(n-1|n-1) = \frac{n\alpha^2}{2}\left[1 - \frac{(n-1)\alpha}{2}\right] + o(\alpha^3),$$
$$D(n-2|n) = \alpha\cdot\frac{n}{3}\cdot\frac{c_{n-2}^3 - (c_{n-2} - d_{n-2})^3}{c_{n-2}^2}\cdot C(n-2|n-1)$$
$$= \frac{\alpha n(3n^2 - 12n + 13)}{3(n-1)^2(n-2)}\cdot C(n-2|n-1)$$
$$= \frac{\alpha^3 n(n-1)(3n^2 - 12n + 13)}{24(n-2)^2} + o(\alpha^3).$$

Thus we finally have

$$\begin{aligned}
d_n &= \frac{\alpha - D(n-1|n) - D(n-2|n)}{\alpha n\left[1 - C(n-1|n-1) - C(n-2|n-1)\right]} + o(\alpha^2) \\
&= \frac{1}{n}\cdot \frac{1 - n\alpha/2 + n(n-1)\alpha^2/8\cdot(1 - 1/(3(n-1)^2))}{1 - n\alpha/2 + n(n-1)\alpha^2/8\cdot(1 - 1/(n(n-2)))} + o(\alpha^2) \\
&= \frac{1}{n}\cdot\left[1 + \frac{\alpha^2\cdot(n-1)(n-3)}{12(n-2)^2}\right] + o(\alpha^2),
\end{aligned}$$

using equation (15). This gives the result in equation (18). □

Derivation of Equation (20). Assume $d_i = 1/i$, $i = 1, \ldots, n$. The probability of rejecting the null hypothesis using the Quick procedure is

$$P(c) = \Pr\left(p_{(n)} \le \alpha\right) + \sum_{j=1}^{n-1} \Pr\left(p_{(n)} > \alpha, p_{(n-1)} > c\alpha, \cdots, p_{(n-j+1)} > c\alpha, p_{(n-j)} \le c\alpha, p_{(1)} \le d_{j+1}\alpha\right),$$

where

$$\Pr\left(p_{(n)} \le \alpha\right) = \alpha^n,$$

and

$$\begin{aligned}
&\Pr\left(p_{(n)} > \alpha, p_{(n-j+1)} > c\alpha, p_{(n-j)} \le c\alpha, p_{(1)} \le d_{j+1}\alpha\right) \\
&= \Pr\left(p_{(n-j+1)} > \alpha, p_{(n-j)} \le d_{j+1}\alpha\right) + \cdots \\
&\quad + \Pr\left(p_{(n-j+1)} > \alpha, \left\{p_{(n-j)}, \cdots, p_{(2)}\right\} \in (d_{j+1}\alpha, c\alpha], p_{(1)} \le d_{j+1}\alpha\right) \\
&\quad + \cdots \\
&\quad + \Pr\left(p_{(n)} > \alpha, \left\{p_{(n-1)}, \cdots, p_{(n-j+1)}\right\} \in (c\alpha, \alpha], p_{(n-j)} \le d_{j+1}\alpha\right) + \cdots \\
&\quad + \Pr\left(p_{(n)} > \alpha, \left\{p_{(n-1)}, \cdots, p_{(n-j+1)}\right\} \in (c\alpha, \alpha], \left\{p_{(n-j)}, \cdots, p_{(2)}\right\} \in (d_{j+1}\alpha, c\alpha], p_{(1)} \le d_{j+1}\alpha\right) \\
&= \sum_{i=0}^{n-1-j} \binom{n}{j}\binom{n-j}{i} (1-\alpha)^j \left[c\alpha - d_{j+1}\alpha\right]^i \left[d_{j+1}\alpha\right]^{n-j-i} \\
&\quad + \sum_{i=0}^{n-1-j} \binom{n}{j-1}\binom{n-j+1}{1}\binom{n-j}{i} (1-\alpha)^{j-1} (\alpha - c\alpha) \left[c\alpha - d_{j+1}\alpha\right]^i \left[d_{j+1}\alpha\right]^{n-j-i} \\
&\quad + \cdots \\
&\quad + \sum_{i=0}^{n-1-j} \binom{n}{1}\binom{n-1}{j-1}\binom{n-j}{i} (1-\alpha) (\alpha - c\alpha)^{j-1} \left[c\alpha - d_{j+1}\alpha\right]^i \left[d_{j+1}\alpha\right]^{n-j-i} \\
&= \binom{n}{j} (1-\alpha)^j \alpha^{n-j} \left[c^{n-j} - (c - d_{j+1})^{n-j}\right] \\
&\quad + \binom{n}{j-1}\binom{n-j+1}{1} (1-\alpha)^{j-1} \alpha^{n-j+1} (1-c) \left[c^{n-j} - (c - d_{j+1})^{n-j}\right] \\
&\quad + \cdots \\
&\quad + \binom{n}{1}\binom{n-1}{j-1} (1-\alpha) \alpha^{n-1} (1-c)^{j-1} \left[c^{n-j} - (c - d_{j+1})^{n-j}\right]
\end{aligned}$$

$$= \sum_{k=0}^{j-1} \binom{n}{j-k} \binom{n-j+k}{k} (1-\alpha)^{j-k} \alpha^{n-j+k} (1-c)^k \left[c^{n-j} - (c-d_{j+1})^{n-j}\right].$$

Therefore,

$$P(c) = \alpha^n + \sum_{j=1}^{n-1} \sum_{k=0}^{j-1} \binom{n}{j-k} \binom{n-j+k}{k} (1-\alpha)^{j-k} \alpha^{n-j+k} (1-c)^k \left[c^{n-j} - (c-d_{j+1})^{n-j}\right]$$

$$(r \triangleq j-k)$$

$$= \alpha^n + \sum_{j=1}^{n-1} \sum_{r=1}^{j} \binom{n}{r} \binom{n-r}{j-r} (1-\alpha)^r \alpha^{n-r} (1-c)^{j-r} \left[c^{n-j} - (c-d_{j+1})^{n-j}\right]$$

$$= \alpha^n + \sum_{r=1}^{n-1} \sum_{j=r}^{n-1} \binom{n}{r} \binom{n-r}{j-r} (1-\alpha)^r \alpha^{n-r} (1-c)^{j-r} \left[c^{n-j} - (c-d_{j+1})^{n-j}\right],$$

and the result in Equation (20) is yielded by replacing the index of summation j with $j-1$.

In addition, by verifying

$$\sum_{j=r+1}^{n} \binom{n-r}{j-r-1} (1-c)^{j-r-1} \left[c^{n-j+1} - (c-1/j)^{m-j+1}\right] \leq \frac{n-r}{n}, \text{ when } c = 1/2,$$

we obtain that the probability $P(c = 1/2)$ is less than α. □

Derivation of Equation (22). For large n, we have, from equation 20,

$$P(c) = \sum_{r=n-3}^{n-1} (1-\alpha)^r \alpha^{n-r} \binom{n}{r} \sum_{j=r+1}^{n} \binom{n-r}{j-r-1} (1-c)^{j-r-1} \left[c^{n-j+1} - \left(c - \frac{1}{j}\right)^{n-j+1}\right] + o(\alpha^3)$$

$$\stackrel{\text{def}}{=} P_{r=n-1}(c) + P_{r=n-2}(c) + P_{r=n-3}(c) + o(\alpha^3).$$

Let $c = c_0 + c_1\alpha + o(\alpha)$. We calculate $P_r(c)$, $r = n-3, n-2, n-1$, as shown below.

$$P_{n-1}(c) = \alpha - (n-1)\alpha^2 + \frac{(n-1)(n-2)}{2}\alpha^3 + o(\alpha^3),$$

$$P_{n-2}(c) = \frac{\alpha^2}{2(n-1)} \left[(2n-1)(n-2) + 2(n-1)c_0\right]$$
$$+ \frac{\alpha^3}{2(n-1)} \left[2(n-1)c_1 - (n-2)\left[(2n-1)(n-2) + 2(n-1)c_0\right]\right] + o(\alpha^3),$$

$$P_{n-3}(c) = \frac{\alpha^3}{2} \left[\frac{n(n-1)}{3(n-2)^2} + \frac{(n-2)(n^2-3n+1)}{n-1} + \left[2(n-2) - \frac{n(2n-3)}{(n-1)(n-2)}\right] c_0 + 2c_0^2\right] + o(\alpha^3).$$

Let $P_{n-1}(c) + P_{n-2}(c) + P_{n-3}(c)$ be equal to $\alpha + o(\alpha^3)$. We obtain

$$\begin{cases} c_0 = \frac{n}{2(n-1)}, \\ c_1 = \frac{n(n^2-2n+2)}{12(n-1)(n-2)^2}, \end{cases}$$

giving the result in equation (22). □

Proof of Theorem 1. Liu (1996) defined a lower triangular matrix called the critical matrix to present the closed test with type I error control $\Pr\left(\cup_{i=1}^{n}\left\{p_{(i)} \le a_{in}\right\}\right) \le \alpha$. Using Meijer et al. (2019)'s algorithm, the key component is to compare a column of critical values $(a_{1,1+k}, \ldots, a_{n-k,n})^\intercal$ with another column $(a_{1+l,1+k}, \ldots, a_{n-k+l,n})^\intercal$, where $1 \le l \le k$.

$$
\begin{array}{cccccccc}
a_{1,1} & & & & & & & \\
\vdots & \ddots & & & & & & \\
a_{1+k,1+k} & \cdots & \boxed{\begin{array}{c} a_{1+l,1+k} \\ a_{2+l,2+k} \\ \vdots \\ a_{n-k+l,n} \end{array}} & \cdots & \boxed{\begin{array}{c} a_{1,1+k} \\ a_{2,2+k} \\ \vdots \\ a_{n-k,n} \end{array}} & & & \\
a_{2+k,2+k} & \cdots & & \cdots & & & & \\
\vdots & & & & & \ddots & & \\
a_{n,n} & \cdots & & \cdots & & \cdots & a_{1,n} &
\end{array}
$$

Meijer et al. (2019)'s algorithm performs in linear time if and only if

$$\frac{a_{i,i+k}}{a_{i+l,i+k}} = \frac{\text{distance between } k \text{ and } k+i}{\text{distance between } k-l \text{ and } k+i}, \text{ for } i = 1, \cdots, n-k \text{ and } l = 1, \cdots, k.$$

This is equivalent to $a_{i,n}/a_{j,n} = [g(n) - g(n-i)]/[g(n) - g(n-j)]$, where $g : \{1, \ldots, n\} \to \mathbb{R}$ is a monotonic function. We arrive at $a_{i,n} = [g(n) - g(n-i)] \cdot h(n)$ as $h(n)$ only depends on n.

Meanwhile, when all critical values in the same row $(a_{m,m}, \ldots, a_{1,m})$ are the same, a step-down multiple test procedure can be constructed. In other words, if $a_{in} = g(n)$ that only depends on n, a linear-time step-down procedure exists. Similarly, a step-up procedure exists when all critical values in a column $(a_{1,n-m+1}, \ldots, a_{m,n})^\intercal$ are the same. This condition is equivalent to $a_{in} = g(n-i)$. □

Proof of Theorem 2. The critical matrix of the closed test in Theorem 2 is

$$
\begin{array}{cccccc}
a_1 & & & & & \\
\vdots & \ddots & & & & \\
a_1 & \cdots & a_{i_\star} & & & \\
a_1 & \cdots & a_{i_\star} & a & & \\
\vdots & & \vdots & \vdots & \ddots & \\
a_1 & \cdots & a_{i_\star} & a & \cdots & a
\end{array}
$$

Therefore, the corresponding procedure first behaves as a version of Hochberg's step-up method for the first $i_\star$ hypotheses which are sorted by their p-values from the largest to smallest. Next, the rest $n - i_\star$ hypotheses are tested using the Quick procedure. Since $i_\star$ does not depend on n, the time complexity of the procedure in Theorem 2 is $O(\log n)$. □

Derivation of Equation (23) *and* (24). By solving $\widetilde{p}$ from

$$p = \frac{\widetilde{p}}{i}\left(1 + \frac{(i-2)\widetilde{p}}{2(i-1)}\right) \quad \text{and} \quad p = -\frac{\log(1-\widetilde{p})}{i},$$

we obtain the formulas in (23). For the GTXR-1c procedure, based on (12) we have

$$p = \frac{i+1}{2i} \cdot \widetilde{p} + \frac{\widetilde{p}^2}{12}\left(1 - \frac{1}{(i-1)^2}\right).$$

Assuming $\widetilde{p} = \lambda_1 p + \lambda_2 p^2 + o(p^2)$, this equation yields $\lambda_1 = (2i)/(i+1)$ and $\lambda_2 = -(2i/(i+1))^3(1-1/(i-1)^2)/12$. For the GTXR-2d procedure, we assume that $\widetilde{p} = \lambda_1 p + \lambda_2 p^2 + \lambda_3 p^3 + o(p^3)$. Solving

$$p = \frac{\widetilde{p}}{i}\left(1 + \frac{\widetilde{p}^2}{12}\left(1 - \frac{1}{(i-2)^2}\right)\right)$$

yields $\lambda_1 = i$, $\lambda_2 = 0$ and $\lambda_3 = -i^3(1-1/(i-2)^2)/12$. □

Derivation of Equation (25). Adjusted p-values of Quick-00, Quick-01, Quick-10 and Quick-11 procedure can be found in a similar way. Equation (25) that gives the adjusted p-values of Quick-00 procedure is obtained by solving $\widetilde{p}$ from $p = \widetilde{p}/2$ and $p = \widetilde{p}/j$. The adjusted p-values of Quick-01 procedure is

$$\text{Quick-01:}\quad \widetilde{p}_{(n-i+1)} = \min\left\{p_{(n)}, \min_{2\le j\le i}\left\{\max\left\{2(n-1)p_{(n-j+1)}/n, jp_{(n-i+1)}\right\}\right\}\right\}, \tag{26}$$

by solving $p = \widetilde{p}n/(2(n-1))$. Similarly, the adjusted p-values of Quick-10 procedure is

$$\text{Quick-10:}\quad \widetilde{p}_{(n-i+1)} = \min\left\{p_{(n)}, \min_{2\le j\le i}\left\{\max\left\{3\sqrt{1+\frac{4}{3}p_{(n-j+1)}} - 1, jp_{(n-i+1)}\right\}\right\}\right\}, \tag{27}$$

by solving $p = \widetilde{p}/2 + \widetilde{p}^2/12$. For the Quick-11 procedure with critical constant $c = c_A(n) + c_B(n)\alpha$, the adjusted p-value is

$$\text{Quick-11:}\quad \widetilde{p}_{(n-i+1)} = \min\left\{p_{(n)}, \min_{2\le j\le i}\left\{\max\left\{\frac{c_A(n)}{2c_B(n)}\sqrt{1+\frac{4c_B(n)p_{(n-j+1)}}{c_A(n)^2}} - 1, jp_{(n-i+1)}\right\}\right\}\right\}, \tag{28}$$

where $c_A(n) = n/(2(n-1))$ and $c_B(n) = 1/12 + 1/4/(n-1) + 1/6/(n-2)^2 - 1/2/(n-1)/(n-2)^2$ for $n \ge 5$. This formula is obtained by solving $p = c_A(n)\widetilde{p} + c_B(n)\widetilde{p}^2$. □

References

Aickin, M. and Gensler, H. (1996). Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods. *American Journal of Public Health*, 86(5):726–728.

Andrew, A. (1979). Another efficient algorithm for convex hulls in two dimensions. *Information Processing Letters*, 9(5):216–219.

Bretz, F., Hothorn, T., and Westfall, P. (2010). *Multiple Comparisons Using R*. Chapman and Hall/CRC, Boca Raton, Florida.

Bretz, F., Maurer, W., Brannath, W., and Posch, M. (2009). A graphical approach to sequentially rejective multiple test procedures. *Statistics in Medicine*, 28(4):586–604.

Cai, G. and Sarkar, S. K. (2008). Modified Simes' critical values under independence. *Statistics & Probability Letters*, 78(12):1362–1368.

Dmitrienko, A., Tamhane, A. C., and Bretz, F. (2009). *Multiple Testing Problems in Pharmaceutical Statistics*. Taylor & Francis, Boca Raton, Florida.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293):52–64.

Finner, H. and Roters, M. (1994). On the limit behaviour of the joint distribution function of order statistics. *Annals of the Institute of Statistical Mathematics*, 46(2):343–349.

Fortune, S. (1989). Stable maintenance of point set triangulations in two dimensions. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 494–499, Los Alamitos, CA, USA. IEEE Computer Society.

Gou, J. and Sarkar, S. K. (2019+). Incorporating pairwise correlation information in Simes test and related FWER and FDR controlling multiple testing procedures. *manuscript submitted*.

Gou, J. and Tamhane, A. C. (2014). On generalized simes critical constants. *Biometrical Journal*, 56(6):1035–1054.

Gou, J. and Tamhane, A. C. (2018a). A flexible choice of critical constants for the improved hybrid Hochberg-Hommel procedure. *Sankhya B*, 80(1):85–97.

Gou, J. and Tamhane, A. C. (2018b). Hochberg procedure under negative dependence. *Statistica Sinica*, 28(1):339–362.

Gou, J., Tamhane, A. C., Xi, D., and Rom, D. M. (2014). A class of improved hybrid Hochberg–Hommel type step-up multiple test procedures. *Biometrika*, 101(4):899–911.

Gou, J. and Xi, D. (2019). Hierarchical testing of a primary and a secondary endpoint in a group sequential design with different information times. *Statistics in Biopharmaceutical Research*, 0(ja):1–26.

Graham, R. (1972). An efficient algorith for determining the convex hull of a finite planar set. *Information Processing Letters*, 1(4):132–133.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4):800–802.

Hochberg, Y. and Tamhane, A. C. (1987). *Multiple Comparison Procedures*. John Wiley and Sons, New York.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6:65–70.

Hommel, G. (1983). Tests of the overall hypothesis for arbitrary dependence structures. *Biometrical Journal*, 25:423–430.

Hommel, G. (1986). Multiple test procedures for arbitrary dependence structures. *Metrika*, 33(1):321–336.

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, 75(2):383–386.

Hommel, G. (1989). A comparison of two modified Bonferroni procedures. *Biometrika*, 76(3):624–625.

Horn, M. and Dunnett, C. W. (2004). *Power and sample size comparisons of stepwise FWE and FDR controlling test procedures in the normal many-one case*, volume 47 of *Lecture Notes–Monograph Series*, pages 48–64. Institute of Mathematical Statistics, Beachwood, Ohio, USA.

Hsu, J. C. (1996). *Multiple Comparisons – Theory and Methods*. Chapman and Hall, London.

Huang, Y. and Hsu, J. C. (2007). Hochberg's step-up method: Cutting corners off Holm's step-down method. *Biometrika*, 94(4):965–975.

Liu, W. (1996). On multiple tests of a non-hierarchical finite family of hypotheses. *Journal of the Royal Statistical Society, Series B*, 58(2):455–461.

Marcus, R., Peritz, E., and Gabriel, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, 63:655–660.

Meijer, R. J., Krebs, T. J. P., and Goeman, J. J. (2019). Hommel's procedure in linear time. *Biometrical Journal*, 61(1):73–82.

Metcalfe, A., Green, D., Greenfield, T., Mansor, M., Smith, A., and Tuke, J. (2019). *Statistics in Engineering: With Examples in MATLAB and R*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition.

Newson, R. B. (2010). Frequentist q-values for multiple-test procedures. *Stata Journal*, 10(4):568–584.

Porter, K. E. (2018). Statistical power in evaluations that investigate effects on multiple outcomes: A guide for researchers. *Journal of Research on Educational Effectiveness*, 11(2):267–295.

Rom, D. M. (1990). A sequentially rejective test procedure based on a modified Bonferroni inequality. *Biometrika*, 77(3):663–665.

Rom, D. M. (2013). An improved Hochberg procedure for multiple tests of significance. *British Journal of Mathematical and Statistical Psychology*, 66(1):189–196.

Rosenthal, R. and Rubin, D. B. (1983). Ensemble-adjusted p values. *Psychological Bulletin*, 94(3):540–541.

Sarkar, S. K. (1998). Some probability inequalities for ordered MTP_2 random variables: a proof of the Simes conjecture. *The Annals of Statistics*, (2):494–504.

Sarkar, S. K. and Chang, C. K. (1997). The Simes method for multiple hypothesis testing with positively dependent test statistics. *Journal of the American Statistical Asociation*, 92:1601–1608.

Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology*, 46(1):561–584.

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, 73:751–754.

Tamhane, A. C. and Gou, J. (2018). Advances in p-value based multiple test procedures. *Journal of Biopharmaceutical Statistics*, 28(1):10–27.

Tamhane, A. C., Gou, J., Jennison, C., Mehta, C. R., and Curto, T. (2018). A gatekeeping procedure to test a primary and a secondary endpoint in a group sequential design with multiple interim looks. *Biometrics*, 74(1):40–48.

Tamhane, A. C. and Liu, L. (2008). On weighted Hochberg procedures. *Biometrika*, 95(2):279–294.

Wang, B. and Cui, X. (2012). A new partition testing strategy for multiple endpoints. *Statistics in Medicine*, 31(20):2151–2168.

Westfall, P. H., Tobias, R. D., Rom, D. M., Wolfinger, R. D., and Hochberg, Y. (2011). *Multiple comparisons and multiple tests using the SAS system*. SAS Institute, Cary, North Carolina.

Wright, S. P. (1992). Adjusted p-values for simultaneous inference. *Biometrics*, 48(4):1005–1013.

Zhang, F. and Gou, J. (2016). A p-value model for theoretical power analysis and its applications in multiple testing procedures. *BMC Medical Research Methodology*, 16(1):135.

Zhang, F. and Gou, J. (2019). Control of false positive rates in clusterwise fMRI inferences. *Journal of Applied Statistics*, 46(11):1956–1972.