A UNIFIED APPROACH TO FALSE DISCOVERY RATE CONTROL UNDER DEPENDENCE THAT INCORPORATES NULL DISTRIBUTION AND SHRINKAGE ESTIMATION

By

JOSEPHINE SARPONG AKOSA

Bachelor of Science in Mathematics
Kwame Nkrumah University of Science and Technology
Kumasi, Ghana
2011

Master of Science in Statistics
University of Texas at El Paso
El Paso, Texas
2014

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2018

A UNIFIED APPROACH TO FALSE DISCOVERY RATE CONTROL UNDER DEPENDENCE THAT INCORPORATES NULL DISTRIBUTION AND SHRINKAGE ESTIMATION

Dissertation Approved:

Melinda McCann, Ph.D.

Dissertation Adviser

Joshua Habiger, Ph.D.

Mark Payton, Ph.D.

Camille DeYong, Ph.D.

ACKNOWLEDGEMENTS

“I will praise you, Lord my God, with all my heart; I will glorify your name forever.” (Psalm 86:12 New International Version). I first and foremost thank the Almighty God for His guidance, protection, and wisdom, and seeing me through my Ph.D. study successfully.

To my advisor, Dr. Melinda McCann, I say a very big thank you. You were not only an adviser, but a friend, a mother and a very good mentor. Your encouragements and faith in me kept me going, especially on days I felt like quitting the Ph.D. program. It has really been an honor and a blessing to have had you as an adviser.

My profound gratitude also goes to the entire Department of Statistics at Oklahoma State University, including all of the faculty, staff, and students. I am grateful for your kindness and concern for me. In particular, I would like to thank my committee members: Dr. Joshua Habiger, Dr. Mark Payton, and Dr. Camille DeYong of the Industrial Engineering and Management Department, for their immense help, time, and insightful comments in making this project a reality.

I also thank the staff at the high performance computing centers at Oklahoma State University and the Mendoza College of Business, University of Notre Dame for allowing me use their supercomputers. Specifically, I would like to thank Jesse Schafer and Rob Easley for assisting me in setting up on the supercomputer platforms, and to Francis Bilson Darku, Gideon Mensah, and John Aidoo for assisting me in debugging my R codes.

Last but not the least, I would like to thank my family and all my friends, especially my beloved parents, siblings, Henrietta Essie Whyte, and Edwin Opata, for your support, prayers, encouragement and above all your love in writing this thesis and in my life in general. I love you all!

I dedicate this work to my beloved parents, siblings, and boyfriend for your continuous prayers, encouragement and support and to everyone who believed in me.

Josephine Sarpong Akosa

Name: Josephine Sarpong Akosa

Date of Degree: December, 2018

Title of Study: A unified approach to false discovery rate control under dependence that incorporates null distribution and shrinkage estimation

Major Field: Statistics

Abstract: FDR-controlling procedures are less stringent but powerful multiple testing procedures for large-scale inference and are therefore the preferred errror rate to control in such studies. But, the validity and accuracy of any FDR-controlling procedure is essentially determined by whether the chosen test statistic is optimal, the null distributions are correctly or conservatively specified, and whether the data are independent across tests. This study proposes two methods which provide asymptotic FDR control. The first method incorporates null distribution and shrinkage estimation into the original procedures of Benjamini and Hochberg (1995) and Benjamini et al. (2006). Extensive Monte Carlo simulations show that the proposed procedures are essentially more stable and as powerful or substantially more powerful than some procedures proposed in finite sample inferential problems, provided there are at least 30 observations in each group for a case-control experiment. The second part of the study proposes a step-down procedure that explicitly incorporates information about the dependence structure of the test statistic, thereby providing a gain in power. One main distinction of this approach from existing stepwise procedures is the null distribution used in place of the unknown distribution of the test statistics. This null distribution does not rely on the restrictive subset pivotality assumption of Westfall and Young (1993).

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure | Page

CHAPTER I

INTRODUCTION

1.1. Overview

Recent high-throughput technology has allowed a rapid increase in the amount of data collected in many scientific fields such as medicine, spatial epidemiology, genetics, biology, neuroscience, economics and finance. The analyses of such high-dimensional data sets often involve statistically testing for some behavior of interest on each of thousands or more measurements taken on the same unit. For example, in genome-wide association studies, thousands of hypotheses are tested simultaneously to identify associations between single-nucleotide polymorphism (SNPs) and some disease trait. In microarray experiments, a researcher may statistically test thousands of genes to identify which of the genes are differentially expressed. See for instance, Dudoit, Yang, Callow, and Speed (2002); Dudoit, Shaffer, and Boldrick (2003); Ge, Dudoit, and Speed (2003); Reiner, Yekutieli, and Benjamini (2003). Two types of errors can occur in any such testing situation:

i. a false positive or type I error, is committed when a variable is declared significant when it is not.

ii. a false negative or type II error, is committed when the test fails to identify a truly significant variable.

Each hypothesis test has its own type I and type II errors. When many hypotheses are tested simultaneously, as in the case of genome-wide association studies or microarray experiments, the probability that at least some type I errors are committed among a set of hypotheses may be unduly large. Multiple testing procedures (MTPs) are very useful tools for dealing with this multiplicity issue. Such procedures provide efficient methods for examining each hypothesis while also controlling for an overall error rate at a pre-specified level.

Additionally, large-scale simultaneous testing of this sort involves inference for high-dimensional multivariate distributions with complex and mostly unspecified dependencies among the genes under consideration. With the defining characteristics of such data, standard methods of multivariate analysis fail. These methods consist of matrix inversion and/or the solution of linear equations for a large number of genes. Thus, it becomes difficult, if implausible, to include all possible genes within a single model. The most common practice in such situations involves analyzing one gene at a time. In order to analyze such data, one must consider the ramifications of three choices. The first choice is that of a suitable statistic. This statistic needs to be chosen such that even though all measurements on one gene are condensed into one number, relevant information is not lost with regards to the test of interest. The second choice involves the rejection regions. Not optimizing these two choices will lead to a loss of statistical power. The third main choice is to find a method to control the inflation of error rates due to simultaneous hypotheses testing.

1.1.1. Choice of Test Statistic

A test statistic is a data-driven measurement that reduces the information in the data to one value that can be used for hypothesis testing. The most widely utilized test statistic is the standardized difference statistic given by

$$T_n = \frac{\text{Estimator}_n - \text{Null Value}}{\text{Standard Error}_n}, \tag{1.1}$$

where the subscript n emphasizes the test statistics' dependence on the sample size. For gene-specific analysis, the test statistics are computed separately for each gene. In the presence of small samples, the error variance is difficult to estimate and subject to erratic fluctuations. For instance, if the estimated variance for one gene is small by chance, the test statistic can be large even when the difference between the estimate and the null is small. For example, due to the large number of genes on each array in microarray experiments, there are usually genes with small standard errors. A common idea adopted by some researchers is to take the dependence structure between test statistics into account by borrowing information across variables rather than treating them as independence. However, these estimates are subject to bias when the error variances across the genes are not homogeneous. In seeking alternative test statistics, researchers seek a middle ground that is both powerful and less biased.

To this end, various test statistics have been suggested in the past couple of years; some of which involve modifying estimators of the error variance components. Tusher, Tibshirani, and Chu (2001) proposed the SAM t-test by adding a small constant to the gene-specific variance estimate in order to stabilize the small variances. Baldi and Long (2001) suggested the regularized t-test which substitutes the traditional variance estimate with a Bayesian estimator based on a hierarchical prior distribution. Using an empirical Bayes approach that pools information across genes, Lönnstedt and Speed (2002) proposed the B statistic. Newton, Noueiry, Sarkar, and Ahlquist (2004) and Kendziorski, Newton, Lan, and Gould (2003) pooled information across genes by considering a hierarchical gamma-gamma model. Building on the work of Lönnstedt and Speed (2002), Smyth (2004) proposed the moderated t-statistic in which posterior residual standard deviations are used in place of ordinary standard deviations. Several other information sharing methods have been proposed based on hierarchical or empirical Bayes techniques (Newton, Kendziorski, Richmond, Blattner, and Tsui (2001); Cui, Hwang, Qiu, Blades, and Churchill (2005); Fox and Dimmic (2006)). Interested readers are referred to Cui and Churchill (2003) and Smyth (2004) for an introductory review of most of these approaches.

Studies have shown that the estimation of gene-specific variances benefits considerably from pooling information across genes (Wright and Simon (2003); Smyth (2004); Cui, Hwang, Qiu, Blades, and Churchill (2005); Delmar, Robin, Tronik-Le Roux, and Daudin (2005)). Bayesian methods, though naturally allowing for information sharing across genes, can become computationally expensive. In addition, these methods rely on detailed assumptions about the underlying data and parameter-generating models. Consequently, Opgen-Rhein and Strimmer (2007) proposed the shrinkage t statistic in the framework of James-Stein-type analytic shrinkage. Since only information concerning second moments rather than fully specified distributions are utilized in the James-Stein shrinkage estimation, the method can also be considered as an empirical Bayes method. The resulting shrinkage statistic is completely analytic and requires no distributional assumptions.

1.1.1.1. Shrinkage Estimation of Covariance Matrix

Estimation of covariance matrices is normally achieved by utilizing the maximum likelihood estimate or the related unbiased sample covariance matrix. However, it is well known that if the sample size, n is small and the number of variables under consideration, m is large, these estimators are very unstable. Many techniques have been proposed to improve the estimation of the matrix; all of which rely on the concept of shrinkage. Bayesian and penalized likelihood methods incorporate shrinkage implicitly while the James-Stein-type approach does so explicitly.

A simple version of the construction of a shrinkage estimator is as follows. Suppose an unregularized estimator $\boldsymbol{U}$, and a target estimator $\boldsymbol{T}$, are available. The unregularized estimator could be either the maximum likelihood estimator or any unbiased estimator. Then, the James-Stein shrinkage estimation rule combines both estimators in a convex weighted average given by

$$\boldsymbol{U}^{\star} = \lambda \boldsymbol{T} + (1 - \lambda)\boldsymbol{U}, \tag{1.2}$$

where $\lambda \in [0, 1]$ is known as the shrinkage intensity parameter and determines the extent to which the estimates are pooled together. The search for the optimal λ is derived from a decision-theoretic perspective by minimizing a risk function, such as the mean squared error (MSE). Common approaches to estimate the minimizing λ are by utilizing MCMC, the bootstrap, cross-validation or by determining it analytically. Note that the unregularized estimator $\boldsymbol{U}$ is recovered when $\lambda = 0$, whereas the target, $\boldsymbol{T}$ dominates when $\lambda = 1$. A shrinkage estimator of this type results in a regularized estimator that typically outperforms the individual estimators, $\boldsymbol{U}$ and $\boldsymbol{T}$, both in terms of accuracy and statistical efficiency. Subsequently, utilizing this shrinkage estimator in equation (1.1) will improve the power of any multiple testing procedure. More details about this estimation procedure will be provided in chapter II.

1.1.2. Test Statistic Null Distribution

The results of any given multiple testing procedure are reported in terms of rejection regions for the test statistics, confidence regions for the parameters of interest, or adjusted p-values. Accordingly, one needs the joint distribution of the test statistics. In practice, however, the true distribution of the test statistics is often unknown. One common practice is to replace the true distribution with a theoretical null distribution, such as the standard normal or the studentized t distribution. However, the presence of correlation among the test statistics can have a significant effect on this theoretical null distribution, resulting in a distribution that is incorrect (Efron (2004, 2007a,b); Pollard and van der Laan (2004)). Even if the theoretical null is appropriate for the individual null test statistic, the effects of correlation between the genes can make the effective joint null significantly different from the theoretical null. Consequently, if these correlations are not accounted for, the multiple testing procedure can perform significantly worse. In practice, when the hope is to effectively make useful discoveries, correlation effects can play a vital role in appropriate identification.

Instead of using the theoretical null, some researchers utilize a data-generated null distribution, such as a permutation null distribution. A sufficient condition to ensure that control of the type I error rate under the assumed data-generated null distribution guarantees the desired control under the true distribution, is for the subset pivotality condition specified in Westfall and Young (1993) to be satisfied. However, in many relevant applications in biomedical problems, the subset pivotality condition is violated. This results from the fact that the data-generated null distribution may incorrectly specify the correlation structure of the true distribution of the test statistics. Efron (2007a) argued that the use of permutation null distributions does not automatically offset correlation effects since these distributions, as typically computed, tend to be similar to the theoretical null. For instance, Pollard and van der Laan (2004) showed in the two-sample problem that the permutation null distribution produces asymptotically correct null distribution if the sample sizes are equal or the covariance structure for the populations are the same.

To avoid the restrictive subset pivotality condition, various ways of estimating the empirical null distribution have been proposed. Pollard and van der Laan (2004); Dudoit, van der Laan, and Pollard (2004); Dudoit, van der Laan, and Birkner (2004) and Pollard, Birkner, van der Laan, and Dudoit (2005) proposed a generally valid null distribution by projecting the true test statistic distribution onto the space of mean zero distributions by bootstrapping centered test statistics. This proposed distribution, null shift and scale-transformed null distribution, utilizes user-supplied upper bounds for the means and variances of the test statistics corresponding to the true null hypotheses. van der Laan and Hubbard (2006) however, argued that for univariate testing, the proposed joint null distribution does not generalize the univariate null distribution one would use in univariate testing. For instance, they emphasized that the marginal distribution of a test statistic is known when the null hypothesis is true. However, the null shift and scale-transformed null distribution guarantees that the obtained marginal distributions and the known marginal distributions have equivalent mean and variance, but does not guarantee that the marginal distributions are equal. This suggests that using this null distribution does not necessarily produce optimal marginal null distributions. Subsequently, the authors proposed utilizing as the null distribution the asymptotic distribution of a vector of null quantile-transformed test statistics which is based on user-supplied marginal test statistic null distributions. Thus, adjusted p-values or rejection regions of a multiple testing procedure based on the joint null quantile-transformed null distribution capitalizes on the dependence among the test statistics to provide a better control and improvement in power than the analogue of using the procedure based only on the marginal null distributions.

The two proposed distributions are based on the notion of *null domination*, whereby the number of false rejections under test statistics' null distribution is stochastically greater than under the true test statistic distribution. This null domination condition is a weaker and less restrictive assumption compared to the assumption of subset pivotality. Unlike the data-generated null distributions, the two proposed null distributions preserve the dependence structure of the test statistics. Efron (2007a) on the other hand argued that the estimation of the null distributions discussed above are justified by asymptotic bootstrap arguments. These asymptotic assumptions raise legitimate concerns in some practical applications. Therefore, Efron (2004, 2007a) proposed an empirical estimation of the null distribution based on the notion of sparsity, (i.e., the proportion of non-null effects is small) and investigated its effect on inference. He referred to this distribution as the empirical null. Though Efron's approach does not rely on asymptotic bootstrap arguments, there are some limitations associated with this approach that can render it difficult to use in some situations. The conventional method for the estimation of the model parameters is based on moments. However, when dealing with non-sparse settings (i.e., the proportion of non-null effects does not tend to zero as the number of hypotheses tends to infinity), the empirical null estimation of Efron (2004, 2007a) does not perform well and the estimators of the null distribution are generally inconsistent. Moreover, as pointed out by Jin and Cai (2007) even when the proportion of non-null effects becomes negligible asymptotically, it still might be of interest to quantify the influence of sparsity on the estimators, in that a small error in the null may increase to large errors in subsequent studies. Additionally, one needs to specify the histogram bin width or the degrees of freedom of the spline when using the empirical null estimation of Efron (2004, 2007a). For some data, diligent adjustment may be required which may be challenging. Finally, there is no guarantee that the order of the scores is maintained in the corresponding FDR values as the approach does not place monotonicity constraints on the density.

1.1.3. Control of Error Rates

Developing MTPs has been a very active area of research. The first error rate suggested was the family-wise error rate (FWER). This measure involves controlling the probability of committing any type I error among all the hypotheses being tested. Many FWER controlling procedures involve testing of hypotheses whose statistics are multivariate normal (or t). This distributional assumption is, however, violated in many of the problems encountered in practice. In addition to this limitation, FWER procedures offer extremely stringent control of the error, which might not always be appropriate. For example, the number of tests in genome-wide studies is large and the nature of analysis is exploratory rather than confirmatory. In this case, one often wishes to make many discoveries without too many false positives, although some false positives can be accepted. Thus, the control of FWER is unnecessarily stringent and less powerful in making discoveries. In a seminal paper, Benjamini and Hochberg (1995) introduced the false discovery rate (FDR) as an alternative measure for accounting for the problem of multiplicity. The FDR, defined as the expected proportion of false positives among all those deemed significant, is a more liberal, but powerful quantity to control. Benjamini and Hochberg (1995), hereafter referred to as the BH procedure, developed a linear step-up procedure for controlling the FDR under the assumption of independence among the test statistics. There is a rich body of literature on FDR controlling procedures under various assumptions on the joint distribution of the test statistics. Many of the controlling procedures assume independent test statistics. Although some of these procedures have been shown to control the FDR under some types of dependency (Benjamini and Yekutieli (2001); Finner, Dickhaus, and Roters (2007)), these procedures were not originally designed to make use of the dependence structure of test statistics. They therefore become less powerful than a procedure which incorporates dependence in some way, especially when the test statistics are highly correlated.

In addition to the dependency issue, the BH procedure is conservative by a factor of $m_0/m = \pi_0$, the proportion of true null hypotheses among all hypotheses. Another line of research has been to utilize the data to estimate the proportion of null hypotheses and then adjust the BH procedure accordingly to provide tighter bounds (Benjamini and Hochberg (2000); Storey (2002); Storey, Taylor, and Siegmund (2004); Benjamini, Krieger, and Yekutieli (2006); Blanchard and Roquain (2009); Gavrilov, Benjamini, and Sarkar (2009); Fan, Han, and Gu (2012); He and Sarkar (2013); Heesen and Janssen (2016)). Two of such procedures with rigorously established control of FDR is the linear step-up procedure of Storey, Taylor, and Siegmund (2004) and the two-stage adaptive procedure of Benjamini, Krieger, and

Yekutieli (2006).

1.1.3.1. False Discovery Rate Control by Resampling

It has well been established that incorporating information about the dependence structure of the test statistics can improve the power of multiple testing procedures. Resampling-based procedures can provide the flexibility of accounting for the complex and unknown dependence structure among the test statistics. Controlling FDR via permutations or other types of resampling such as the bootstrap has received a lot of attention over the past two decades. Yekutieli and Benjamini (1999) initiated this subject and proposed a permutation-based procedure that offers asymptotic control of the FDR. Ge, Sealfon, and Speed (2008) also proposed three different FDR-controlling procedures, one of which has proven finite-sample control. Building on the previous work of Troendle (2000) that had restrictive parametric assumptions, Romano, Shaikh, and Wolf (2008) proposed a bootstrap procedure that controls the FDR asymptotically. Their procedure relies upon an exchangeability assumption. However, their procedure is based on a data-generated null distribution. As discussed earlier, the data-generated null may incorrectly specify the true dependence structure of the test statistics. Thus, in the presence of high correlations, their proposed procedure may undercut inferential validity.

A different option for developing resampling-based techniques is to utilize Benjamini and Hochberg's procedure on permutation p-values. Nevertheless, such permutation-based approaches do not preserve the correlation structure of the p-values.

1.2. Motivating Examples

In microarray experiments, a common goal is to identify genes that show differential expression across biological and clinical conditions. The following motivating data sets are fairly typical of data obtained in such experiments.

1.2.1. Example 1: Hereditary Breast Cancer

Consider the well known microarray experiment of Hedenfalk et al. (2001) concerning differences between two types of genetic mutations causing increased breast cancer (BRCA1 and BRCA2). The experiment consisted of $n = 15$ tumor samples from patients with primary breast cancer (7 with BRCA1 and 8 with BRCA2) to identify cases of hereditary breast cancer on the basis of $m = 3,226$ gene-expression profiles. In their analysis, the authors computed a modified F statistic and, using a threshold of $\alpha = 0.001$, identified 51 genes as differentially expressed. Following this, the authors analyzed the 15 tumor samples with a threshold of $\alpha = 0.0001$ to identify 9 to 11 genes as differentially expressed.

1.2.2. Example 2: HIV Type I Infection

The human immunodeficiency virus (HIV) data set described by Van't Wout et al. (2003) used the same RNA preparation for four experiments on four different slides. After twenty-four hours of infection with HIV virus type 1, the expression levels of cellular RNA transcripts were assessed in CD4-T-cell lines. The final dataset consisted of $n = 8$ patients (4 negative and 4 positive subjects) and $m = 7680$ gene levels. More details about the dataset are provided in Van't Wout et al. (2003) and Gottardo, Raftery, Yeung, and Bumgarner (2005).

In the two microarrray experiments described above, it is desirable to compare gene expression under the two different conditions. The aim is to identify which of the m genes have had their expression levels changed. The two experiments share the following common characteristics:

i. The dimension of the data is much larger than the sample size.

ii. There is complex, mostly uncharacterized, correlation among the genes under consideration.

iii. Error rates are inflated due to simultaneous hypotheses testing.

iv. Some proportions of the null hypotheses are expected to be true.

For the analyses of these experiments, if the standard p-value threshold of 0.05 is utilized to perform separate hypothesis tests, one would expect 161 and 384 genes to be deemed differentially expressed by chance for the breast cancer and HIV studies respectively if all null hypotheses are true. Thus, the problem of multiplicity is a major concern in performing simultaneous inference in studies of this nature. This warrants the need for multiple testing procedures in performing simultaneous inference in microarray experiments such as the two described above.

In microarray experiments, scientists measure the expression levels of hundreds or thousands of genes within a cell by measuring the amount of labeled cDNA bound to each site on an array containing many DNA samples. Unfortunately, the experimental procedure used to obtain the data induces substantial correlation among the various microarrays. This causes a major concern in the analyses of such data. For instance, utilizing the traditional approach for analyzing the two microarray experiments yields a two-sample t-statistic, t_i $(i = 1, \ldots, m)$, for each of m genes comparing the two conditions under study. Here, each t_i tests the null hypothesis that gene i is not differentially expressed under both experimental conditions (i.e., HIV positive and negative subjects for the HIV data). The $t_i's$ have been converted to z-values for easy comparison to the theoretical null distribution, $\mathcal{N}(0, 1)$. The results for these test statistics are displayed in Figure 1.1. The smooth curve is a standard normal distribution. As pointed out by Efron (2007a), more null z-values will be in the tails of the distribution for the cancer study due to the wide central histogram, so if the theoretical $\mathcal{N}(0, 1)$ distribution is used to judge significance levels, the procedure will be too liberal. On the other hand, the theoretical $\mathcal{N}(0,1)$ null is too conservative for the HIV study. This shows that utilizing an inappropriate null distribution can greatly influence inferential validity, even when multiplicity has been accounted for.

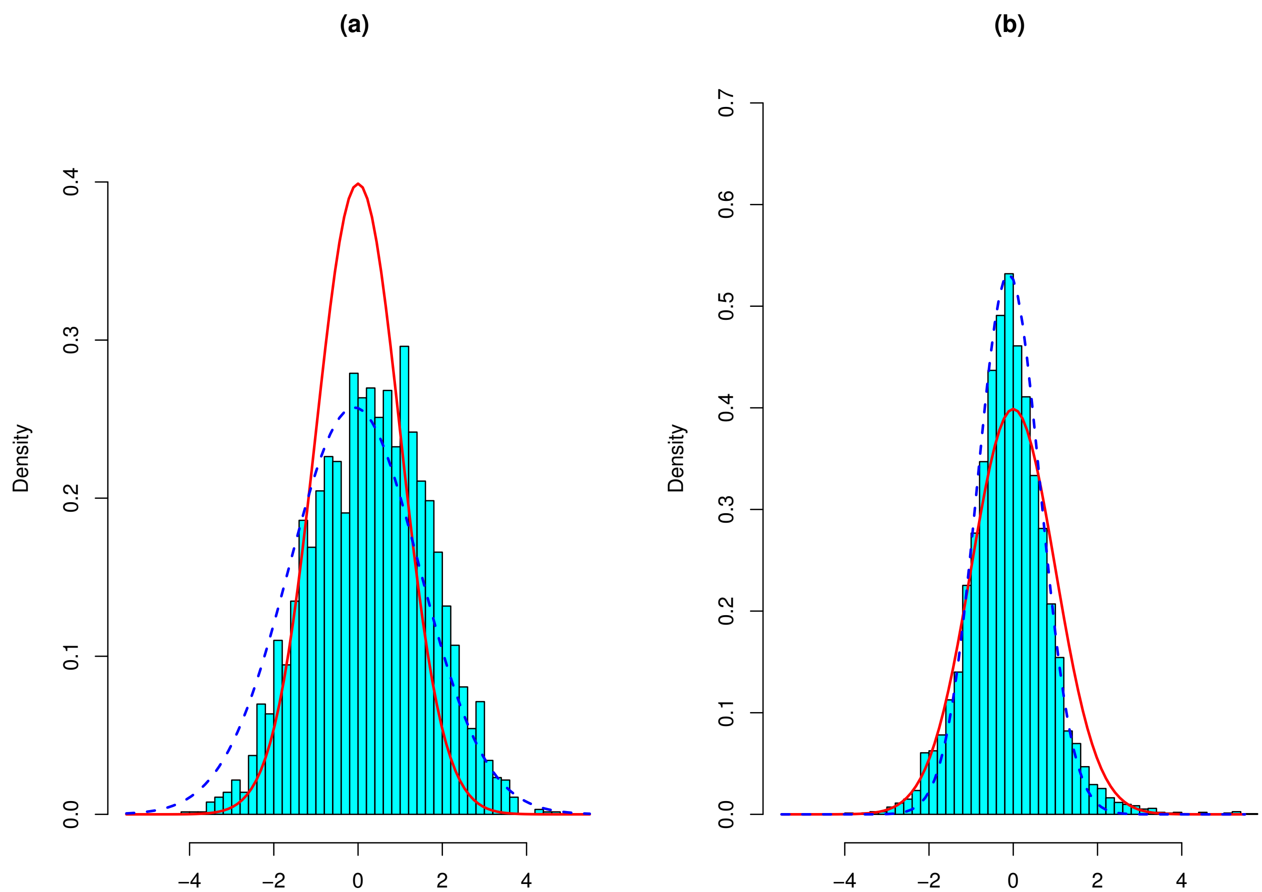

Figure 1.1. Histograms of z-values from the motivating examples. (a) Hereditary breast cancer study, 3,226 genes (Hedenfalk et al., 2001)). (b) HIV type 1 study, 7,680 genes (Van't Wout et al., 2003). Smooth red curves indicate $\mathcal{N}(0,1)$ theoretical null distribution. Dashed blue curves indicate normal empirical distribution ($\mathcal{N}(-0.09, 1.55^2)$ for (a) and $\mathcal{N}(-0.11, 0.75^2)$ for (b)) as estimated by Efron (2007a). The theoretical null is too narrow in (a) and too wide in (b).

Controlling for multiplicity by utilizing the false discovery rate at level $\alpha = 0.1$, Efron (2007a) utilized the theoretical $\mathcal{N}(0,1)$ and an estimated empirical null distribution to analyze the two microarray experiments. The breast cancer study resulted in 107 discoveries when the theoretical null distribution was utilized whereas no discoveries were made utilizing an empirical null distribution. Similarly, 22 as opposed to 180 discoveries were made for the HIV study when the theoretical null was used. Efron (2007a) explains that the discrepancies in the results do not stem from the use of the BH procedure itself, but from the unconditional use of the theoretical null distribution. Thus, blindingly utilizing the theoretical null distribution can greatly influence which cases are deemed significant, irrespective of which multiple testing procedure is employed. It is therefore essential to account for correlation when developing multiple testing procedures. Additionally, the choice of an appropriate null distribution is very crucial, as utilizing an inappropriate null can undercut inferential validity, a problem that was encountered in the analysis of the HIV study by Efron (2007a).

1.3. Contributions of this Work

In the literature, it is common to find advances in resampling-based multiple testing procedures that control FDR under dependency – albeit not in combination with estimation of test statistic null distribution and variance components. Motivated by the above applications and the limitations of existing multiple testing procedures, this study seeks to develop resampling-based procedures that brings together many aspects of multiple testing methodologies that otherwise are only considered separately. The contributions of this work are in two-fold. The first study incorporates null distribution and shrinkage estimation into the original linear step-up procedure of Benjamini and Hochberg (1995) and the two-stage adaptive procedure of Benjamini, Krieger, and Yekutieli (2006). Specifically, a James-Stein type analytic shrinkage estimation approach is first utilized to estimate the variance components. These estimates are subsequently utilized in the construction of an appropriate test statistic. After that, instead of using the theoretical null distribution or the data-generated null distribution that relies on subset pivotality to ensure type I error control, the study proposes using an empirical null distribution for the test statistics. The estimated null distribution is then utilized to obtain unadjusted p-values for use in the Benjamini and Hochberg (1995) and Benjamini, Krieger, and Yekutieli (2006) procedures.

The second part of the study proposes a step-down procedure based on the estimated shrinkage test statistic and the test statistic null distribution. One main distinction of this approach from existing stepwise FDR procedures is the null distribution used in place of the unknown joint distribution of the test statistics. This null distribution does not rely on the restrictive subset pivotality assumption of Westfall and Young (1993).

The present approach to FDR control is best described as a unified approach of multiple testing techniques with the James-Stein-type analytic shrinkage estimation of variance components of Schäfer and Strimmer (2005); Opgen-Rhein and Strimmer (2007) and the null distribution modeling of Pollard and van der Laan (2004); Dudoit, van der Laan, and Pollard (2004); Dudoit, van der Laan, and Birkner (2004); van der Laan, Dudoit, and Pollard (2004) and van der Laan and Hubbard (2006). A limitation to the methods proposed is that it is computationally intensive as compared to some of the other methods due to the resampling process. However, with modern computing power, this issue is far less important than in years past and, in general, this method is effective in many settings where traditional approaches are far too conservative.

1.4. Chapter Organization

The remaining chapters are set out as follows. Chapter 2 presents an overview of the multiple testing problem and provide some multiple testing procedures for controlling the false discovery rate. In addition, an appropriate test statistic null distribution, rather than a data-generated null distribution for large scale inference are discussed. We also discuss the general principles for the construction of James-Stein-type analytic shrinkage estimators.

In chapter 3, we present resampling-based techniques for improving the original linear step-up procedure of Benjamini and Hochberg (1995) and the two-stage adaptive linear step-up procedure of Benjamini, Krieger, and Yekutieli (2006), by incorporating shrinkage estimation of the error variance and a generally valid null distribution. Theoretical results and conditions for when the proposed resampling-based procedures provide asymptotic FDR control are also provided. Since the proposed procedures are based on asymptotic arguments, extensive Monte Carlo simulations are carried out to assess their finite sample performance. Additionally, the FDR control, power, and stability, as characterized by the standard deviations of the number of false hypotheses rejected and the total number of rejected hypotheses of the proposed resampling-based methods are compared to some existing FDR-controlling procedures.

Chapter 4 builds on the work of Romano, Shaikh, and Wolf (2008) and proposes a new step-down procedure for controlling the false discovery rate. Theoretical results and conditions for when the proposed procedure provides asymptotic control are also given. The ongoing research and possible extensions of the step-down procedure are provided as well.

Finally, chapter 5 provides discussion and concluding remarks for the methods proposed in this study. Recommendations and possible future extensions are also provided. Supplemental simulation results are provided in Appendices B and C.

CHAPTER II

PRELIMINARIES

2.1. The Problem of Multiple Testing

When performing multiple inferences, researchers normally select the statistically significant ones for emphasis, discussion, and to support conclusions of some research questions. Suppose we want to test the null hypothesis H_0 against an alternative H_1 based on a test statistic T. Then we will reject H_0 in support of H_1 if for a given rejection region Γ, $T \in \Gamma$. A type I error is committed when $T \in \Gamma$ but H_0 is really true. On the other hand, a type II error is committed when $T \notin \Gamma$ but H_1 is really true. In order to choose Γ, a pre-specified level of significance α, the acceptable type I error rate, is chosen, and all rejection regions that have a type I error rate less than or equal to α are considered. The rejection region with the smallest type II error is then chosen among the considered regions. In testing m null hypotheses, each hypotheses test will have its own type I and type II error rates, thereby making the nature of the overall error rate complicated. Consequently, an unguarded use of single-inference procedures in multiple hypothesis testing inflates the overall error rates. In an effort to address these issues, several procedures have been developed. This chapter review some advances in dealing with large-scale simultaneous hypotheses testing.

Consider testing simultaneously m null hypotheses $H_0(i);\ i = 1, \ldots, m$, based on an m-dimensional vector of test statistics, $T_n = (T_n(i) : i = 1, \ldots, m)$ with joint distribution $Q_n = Q_n(P)$, where $P \in \Omega$ is a data generating distribution. Suppose $\mathcal{H}_0 = \mathcal{H}_0(P)$ is the set of true null hypotheses and $\mathcal{H}_1 = \mathcal{H}_1(P)$, the set of false null hypotheses. Then, $m_0 =\mid \mathcal{H}_0 \mid$ is the number of true null hypotheses and $m_1 = m - m_0 =\mid \mathcal{H}_1 \mid$ is the number of false null hypotheses. Let $\mathcal{C}_n(i) = \mathcal{C}(i; T_n, Q_n, \alpha)$ denote the rejection threshold corresponding to each hypothesis test and $\mathcal{R}_n = \mathcal{R}(T_n, Q_n, \alpha)$ the set of rejected null hypotheses based on a multiple testing procedure. Denote the number of rejections and false rejections based on the procedure respectively by V and R such that,

$$
\begin{aligned}
R &=\mid \mathcal{R}(T_n, Q_n, \alpha) \mid=\mid \mathcal{R}_n \mid \\
V &=\mid \mathcal{R}(T_n, Q_n, \alpha) \cap \mathcal{H}_0(P) \mid=\mid \mathcal{R}_n \cap \mathcal{H}_0 \mid . \qquad (2.1)
\end{aligned}
$$

Most of the literature on procedures adjusting for multiple testing describe controlling one of two overall error rates: the familywise error rate or the false discovery rate.

Definition 2.1.1

The familywise error rate (FWER) is defined as the probability of making at least one type I error in a family of hypotheses. That is,

$$FWER = P(V \geq 1) \qquad (2.2)$$

Definition 2.1.2

The false discovery rate (FDR) is defined as the expected proportion of true null hypotheses among all those declared significant. The FDR is given by

$$FDR = E\left(\frac{V}{\max(1, R)}\right) \qquad (2.3)$$

2.1.1. Existing Multiple Testing Procedures

To date many FWER controlling procedures have been proposed. In these settings, instead of controlling the type I error rates at level α for each individual test, the overall FWER is controlled at level α. A rejection region is then determined that maintains level α FWER while still yielding good power. Since the FWER controls the probability of making at least one type I error, these procedures are often too stringent and might not always be the appropriate error rate to control. The interested reader is referred to Westfall and Young (1993), Hochberg and Tamhane (1987), Hsu (1996) and Shaffer (1995) for a review of some of these multiple testing procedures.

In a pioneering work, Benjamini and Hochberg (1995) proposed the FDR as an alternative measure to the FWER. In the context of analyzing a large number of variables, controlling the FDR has become increasingly popular. Choosing which overall error rate to use relies heavily on the scientific goal and expectation of the study. In high-dimensional settings, the primary aim of the initial analysis is exploratory rather than confirmatory. In such cases, one seeks a procedure with very good power in order to make as many discoveries as possible, but making some mistakes is acceptable here as these mistakes are likely to be identified in subsequent confirmatory experiments/analyses. Thus, controlling FDR seems a natural choice for such settings. Unlike the FWER, the FDR is a less stringent controlling procedure, thereby leading to an increase in statistical power.

2.1.2. Procedures for FDR Control

Benjamini and Hochberg (1995) provided a step-up p-value method for controlling FDR under the assumption of independent p-values. The authors proved their procedure controls the FDR at level $\pi_0\alpha \leq \alpha$, where π_0 is the proportion of true null hypotheses among all hypotheses. Ensuing research has shown that this procedure is still valid under some special dependencies, for example, positive regression dependence on subsets (Benjamini and Yekutieli (2001); Finner, Dickhaus, and Roters (2007)). Benjamini and Liu (1999) alternatively developed a step-down FDR procedure. The authors demonstrated that their step-down procedure controls the FDR under independence, and it neither dominates nor is dominated by the BH step-up procedure. Sarkar (2002) further showed that both the BH procedure and the procedure of Benjamini and Liu can be controlled by a generalized stepwise procedure under positive regression dependence on subsets.

The BH procedure is conservative by a factor of $\pi_0 = m_0/m$ for controlling FDR at level α when some of the hypotheses are in fact false. Knowledge of m_0, the number of true null hypotheses can be useful for improving the power of the procedure substantively. This suggests that incorporating a good estimate of π_0 into the BH procedure would result in a more powerful procedure especially when many hypotheses are false. Such procedures are referred to as adaptive procedures. There have been significant recent advances on the estimation of π_0 (Benjamini and Hochberg (2000); Storey (2002); Storey, Taylor, and Siegmund (2004); Benjamini, Krieger, and Yekutieli (2006); Blanchard and Roquain (2009); He and Sarkar (2013); Heesen and Janssen (2016)). One problem with the adaptive procedures, however, is that the estimate of π_0 can be extremely variable, especially when the p-values are highly correlated. Consequently, if this variance is not taken into account, then naive plug-in procedures will generally not offer FDR control, especially when $\pi_0 \approx 1$. In order to provide substantial improvement over the BH procedure, adaptive methods need to take into account the estimation error of π_0. One such procedure is provided by Benjamini, Krieger, and Yekutieli (2006) who adjust the α-level slightly from α to $\alpha^* = \alpha/(1+\alpha)$ to adjust for the additional variability due to the estimation of π_0. Note also that, adaptive procedures offer better performance by utilizing the difference between π_0 and 1. In the presence of small differences, these procedures offer little advantage in terms of power. Conversely, such procedures offer a more evident gain in power when the proportion is small.

The implementation of the above procedures, however, make use of the marginal distribution of the test statistics without taking into account their dependency structure. More powerful procedures can be developed if the dependency structure of the test statistics are considered. Resampling-based techniques can provide the flexibility to accomplish this. Benjamini and Yekutieli (2001) pioneered this methodological path and provided asymptotic control of FDR with a permutation-based approach. Their analysis required subset pivotality and independency between the test statistics corresponding to the true null hypotheses and those corresponding to the false null hypotheses. Troendle (2000) proposed step-up and step-down FDR procedures under the assumption of normality of the test statistics. This procedure was shown to provide asymptotic control of the FDR. Using least favorable configurations, Somerville (2004) developed both step-up and step-down FDR procedures under the assumption of a multivariate t distribution and common correlation of the test statistics. The author, however, did not provide an exhaustive proof of the validity of the assumed location of the least favorable configurations. Building on the work of Troendle (2000), Romano, Shaikh, and Wolf (2008) developed a bootstrap procedure that controls FDR asymptotically and relies upon an exchangeability assumption. Their procedure utilized a data-generated null distribution in place of the unknown joint distribution of the test statistics. However, as will be discussed in the next section, utilizing a data-generated null distribution may incorrectly specify the true dependence structure of the test statistics. Thus, in the presence of strong correlations among the test statistics, the Romano, Shaikh, and Wolf (2008) procedure may lead to misleading results.

A couple of studies have also considered the false discovery rate from different points of view, including Bayesian, empirical Bayes, as the limit of empirical process and in the context of penalized model selection. For instance, Efron, Tibshirani, Storey, and Tusher (2001) developed an empirical Bayes approach to multiple testing and made interesting connections with FDR. Storey (2002, 2003) connected the FDR concept with a certain Bayesian quantity and proposed a new FDR method which has more power than the original Benjamini and Hochberg (1995) procedure. In their paper, Abramovich, Benjamini, Donoho, and Johnstone (2006) utilized the concept of FDR in developing asymptotically minimax procedures for model selection.

2.2. Choice of Test Statistic Null Distribution

Recall from equation (2.1), the rejection region of a testing procedure is a function of the joint distribution of the test statistics, Q_n. However, in practice, the true distribution is unknown and it is normally replaced by a null distribution Q_0 or an estimator Q_{0n} thereof. The choice of an appropriate null distribution is therefore vital in order to ensure control of the type I error rate under the assumed null distribution. It is not uncommon for researchers to replace the null distribution with a theoretical null such as the standard normal distribution. Efron (2004, 2007a,b) however emphasized that even if the theoretical null is appropriate for individual null test statistics, the effects of correlation among the variables can make the effective joint null significantly narrower or wider than the theoretical null. A second choice for the null distribution is to use a data-generated null distribution such as the permutation null distribution, $Q_n(P_0)$. The validity of the permutation distribution is based on the assumption of the complete null hypotheses, i.e., that all m hypotheses are true. Pollard and van der Laan (2004); Pollard, Birkner, van der Laan, and Dudoit (2005) and Efron (2007a) argued that testing procedures based on this data-generated null distribution, $Q_n(P_0)$ do not necessarily provide good control of the type I error rate under the true distribution. In fact, the data-generated null distribution may incorrectly specify the dependence structure of the true distribution of the test statistics. Efron (2007a) further argued that the use of the permutation null distribution does not automatically offset the dependence effects since the distribution tends to be similar to the theoretical null, considering the manner in which they are estimated.

For microarray experiments, Efron (2004, 2007a,b) discussed four of many reasons why the null distribution might differ from the theoretical null. These consist of:

i. Failed assumptions: The theoretical null distribution is justified if the individual gene levels are normal or approximately normally distributed. This is however not the case for most applications in microarrays.

ii. Unmeasured covariates in an observational study: Unmeasured covariates tend to dilate the effective null distribution of the test statistics, but the theoretical null distribution does not include any dilation effects. Empirically estimating the null distribution can help account for the dilation effects. Some examples of these unmeasured covariates include age and gender.

iii. Correlations across units: Generally, theoretical null distributions for test statistics assume independence across the sampling units: for instance, across the 15 tumor samples in the hereditary breast cancer study or the 8 patients in the HIV study in the motivation examples in section 1.2 of Chapter 1. This may not always be appropriate.

iv. Correlations between genes: Independence between genes is not a requirement for the validity of some false discovery rate procedures. However, if the choice of a null distribution is inappropriate, the results of any large-scale inference can be grossly misleading. See Efron (2007a) for detailed explanation of the effect of correlation across genes.

As illustrated by Efron (2007a), a permutation null distribution deal most effectively with the first of the four reasons listed above. In testing a single hypothesis, one has no option but to use either the theoretical or a permutation null distribution. Large-scale testing, however, allows for the empirical estimation of an appropriate null distribution. An empirical null distribution uses the study's own data to estimate an appropriate null distribution.

Pollard and van der Laan (2004); Dudoit, van der Laan, and Pollard (2004); Dudoit, van der Laan, and Birkner (2004); van der Laan, Dudoit, and Pollard (2004) and van der Laan and Hubbard (2006) have developed comprehensive correlated testing solutions using resampling theory. Their proposed techniques and null distribution estimation are summarized in the following subsections.

2.2.1. Null Domination Conditions for Type I Error Rates

Suppose V_n and R_n are the number of type I errors and the number of rejected hypotheses respectively under the true distribution Q_n, and V_0 and R_0 are the number of type I errors and the number of rejected hypotheses respectively under a chosen null distribution, Q_0, by a multiple testing procedure. In order to provide proper control, the type I error rate under the null distribution, Q_0, must dominate the type I error rate under the true distribution, Q_n. That is,

$$\begin{aligned} \Theta(F_{V_n,R_n}) &\leq \Theta(F_{V_0,R_0}) && \text{(finite sample control)} \\ \limsup_{n\to\infty} \Theta(F_{V_n,R_n}) &\leq \Theta(F_{V_0,R_0}) && \text{(asymptotic control),} \end{aligned} \tag{2.4}$$

where $\Theta(\cdot)$ denotes the type I error rate and $F_{\cdot}$ is the cummulative distribution function of the number of type I errors of a given multiple testing procedure. Here, the error rate may either be the familywise error rate (FWER) or the false discovery rate (FDR) defined earlier in section 2.1. The authors, Pollard and van der Laan (2004); Dudoit, van der Laan, and Pollard (2004); Dudoit, van der Laan, and Birkner (2004); van der Laan, Dudoit, and Pollard (2004), explained that the concept of null domination differs from subset pivotality in the following two ways:

i. Unlike subset pivotality which considers all possible subsets of null hypotheses, null domination only considers the subset of true null hypotheses.

ii. Null domination requires the weaker domination of $Q_{n,\mathcal{H}_0}$ by $Q_{0,\mathcal{H}_0}$ and not the equality of the joint distributions, $Q_{n,\mathcal{H}_0}$ and $Q_{0,\mathcal{H}_0}$ for $\mathcal{H}_0$-specific test statistics.

2.2.2. Estimation of the Test Statistic Null Distribution

2.2.2.1. The Null Shift and Scale-transformed Test Statistic Null Distribution

Pollard and van der Laan (2004); Dudoit, van der Laan, and Pollard (2004); Dudoit, van der Laan, and Birkner (2004); van der Laan, Dudoit, and Pollard (2004) proposed the null distribution to be "the asymptotic distribution of a vector of null shift and scale-transformed test statistics, based on user-supplied upper bounds for the means and variances of the $\mathcal{H}_0$-specific test statistics". The general construction for this null distribution is given as follows. Assume there exists an m-dimensional known real-valued vector $\boldsymbol{\lambda}_0$ and a positive real-valued vector $\boldsymbol{\tau}_0$ of null values such that

$$\begin{aligned} \limsup_{n\to\infty} E\left(\boldsymbol{T}_n(i)\right) &\leq \boldsymbol{\lambda}_0(i) && \text{and} \\ \limsup_{n\to\infty} \operatorname{Var}\left(\boldsymbol{T}_n(i)\right) &\leq \boldsymbol{\tau}_0(i) && \text{for } \; i \in \mathcal{H}_0. \end{aligned} \tag{2.5}$$

The authors proposed the null distribution $Q_0 = Q_0(P)$ as the asymptotic distribution of the m-dimensional vector of null shift and scale-transformed test statistics

$$\boldsymbol{Z}_n^{NS}(i) = \sqrt{\min\left(1, \frac{\boldsymbol{\tau}_0(i)}{\operatorname{Var}\left(\boldsymbol{T}_n(i)\right)}\right)}\left(\boldsymbol{T}_n(i) + \boldsymbol{\lambda}_0(i) - E(\boldsymbol{T}_n(i))\right); \quad i = 1, \ldots, m. \tag{2.6}$$

For this choice of null distribution, $Q_0 = Q_0(P)$ null domination holds asymptotically (See proof of this assertion on page 34 of Dudoit, van der Laan, and Pollard (2004) or page 14 of van der Laan, Dudoit, and Pollard (2004)). Since the data generating distribution, P is unknown in practice, so is the proposed null distribution, $Q_0(P)$. Consequently, resampling techniques, such as the bootstrap can be utilized to obtain consistent estimators of the null distribution.

Remark 2.2.1

The following remarks can be made about the role of the null shift and scale values in the formulation of the null distribution.

1. ***Null shift values,*** $\boldsymbol{\lambda}_0$*: The formulation of the null distribution,* Q_0 *is based on the assumption of null domination. The null shift values,* $\boldsymbol{\lambda}_0(i)$ *are utilized to produce* $\mathcal{H}_0$*-specific statistics* $(\boldsymbol{Z}_n^{NS}(i);\ i = 1, \ldots, m)$ *that are stochastically larger than the original test statistics* $(\boldsymbol{T}_n(i);\ i = 1, \ldots, m)$*, assuming large values are evidence against* H_0*, thus, ensuring a null distribution that satisfies the null domination assumption.*

2. ***Null scale values,*** $\boldsymbol{\tau}_0$*: The null scale values,* $\boldsymbol{\tau}_0(i)$ *unlike* $\boldsymbol{\lambda}_0(i)$*, are not needed for type I error control but are rather needed to avoid the degeneration of the null distribution and infinite thresholds for the false null hypotheses. This is a vital attribute for obtaining powerful multiple testing procedures.*

2.2.2.2. The Null Quantile-transformed Test Statistic Null Distribution

Generally, the marginal distribution of the test statistics, $\boldsymbol{T}_n(i)$ for the true null hypothesis is known. Subject to this, van der Laan and Hubbard (2006) pointed out that the null shift scale-transformed distribution guarantees that the mean and variance of the marginal distribution obtained from this distribution and those of the known marginal distribution are approximately the same. On the contrary, the null shift scale-transformed distribution does not guarantee the equality of the two distributions. Hence, for univariate testing, the null shift scale-transformed distribution is not guaranteed to give the most powerful procedure. One can therefore expect to improve the power of a multiple testing procedure by utilizing a null distribution that produces the optimal marginal null distribution. Motivated by this, van der Laan and Hubbard (2006) proposed the null quantile-transformed null distribution, based on user-supplied marginal test statistic null distributions, $q_{0,i}$ $(i = 1, \ldots, m)$. For this quantile-transformed null distribution, the test statistics, $\boldsymbol{T}_n(i)$ corresponding to the set of true null hypotheses should be stochastically larger under the null distributions $q_{0,i}$ than under the true distributions $Q_{n,i}$. This condition is known as marginal null domination and is summarized as follows. For a real-valued number, z and for each $i \in \mathcal{H}_0$, the marginal null domination condition is satisfied if

$$\begin{aligned} &q_{0,i}(z) \leq Q_{n,i}(z) \qquad \text{or} \qquad &&\text{(finite sample control)} \\ &q_{0,i}(z) \leq \limsup_{n\to\infty} Q_{n,i}(z) &&\text{(asymptotic control).} \end{aligned} \tag{2.7}$$

The quantile-transformed null distribution is thus the joint distribution of the m-dimensional vector of null quantile-transformed test statistics

$$Z_n^{NQ}(i) = q_{0,i}^{-1} Q_{n,i}^{\Delta}(\boldsymbol{T}_n(i)) \qquad i = 1, \ldots, m, \tag{2.8}$$

where $Q_{n,i}^{\Delta}(z) = \Delta Q_{n,i}(z) + (1-\Delta) Q_{n,i}(z^-)$ and Δ is a uniform random variable on the interval $[0, 1]$ and independent of the data. Like the null shift scale-transformed distribution, the quantile-transformed distribution is dependent on the unknown data generating distribution. Thus, resampling techniques can also be utilized to obtain consistent estimators.

For clarity and whenever necessary, we will denote the null shift and scale-transformed null distribution as Q_0^{NS} and the null quantile-transformed null distribution as Q_0^{NQ}.

2.3. Shrinkage Estimation

Most statistical applications require an estimate of a covariance matrix and/or its inverse. The standard estimator utilized in such applications is the maximum likelihood estimate or the sample covariance matrix. However, for situations where a large number of variables but comparatively few samples are available, these estimates are unreliable and cannot be considered a good approximation to the true covariance matrix. These estimates are not even invertible in such cases. Recent advances in obtaining better estimators employ the concept of shrinkage, which is as a consequence of the work of James and Stein (1961). The general principles for the construction of James-Stein-type analytic shrinkage estimators are reviewed in the following.

2.3.1. General Concept of Shrinkage Estimation

Suppose $\boldsymbol{\Psi} = (\psi_1, \ldots, \psi_m)$ denote a set of unrestricted large-scale parameters of interest, and $\boldsymbol{\Theta} = (\theta_1, \ldots, \theta_m)$, a lower-dimensional set of parameters (target parameters). Furthermore, suppose the estimation rules $\boldsymbol{U} = \hat{\boldsymbol{\Psi}}$ and $\boldsymbol{T} = \hat{\boldsymbol{\Theta}}$ are available. Then, the James-Stein linear shrinkage suggest the estimation rule that combine both estimators in a weighted average given as

$$\boldsymbol{U}^{\star} = \lambda \boldsymbol{T} + (1 - \lambda)\boldsymbol{U}, \tag{2.9}$$

where $\lambda \in [0, 1]$ is known as the shrinkage intensity parameter and it determines the extent to which the estimates are pooled together. If $\lambda = 0$ the unrestricted estimate is recovered whereas for $\lambda = 1$, the target estimate dominates. A shrinkage estimator of this type may result in a regularized estimator that outperforms the individual estimators, $\boldsymbol{U}$ and $\boldsymbol{T}$, both in terms of accuracy and statistical efficiency.

After deciding to improve upon an unregularized estimate using the shrinkage approach of equation (2.9), the key question is how to select an optimal value for the shrinkage parameter. An appropriate approach is to choose λ from a decision-theoretic perspective by minimizing a risk function, such as the mean squared error (MSE) given by

$$R(\lambda) = E(L(\lambda)) = E\left(\sum_{i=1}^{m}(u_i^\star - \psi_i)^2\right) \tag{2.10}$$

Several techniques have been employed to estimate λ from equation (2.10). For instance, Friedman (1989) applied cross-validation techniques to estimate the optimal λ in the context of regularized classification. Morris (1983) and Greenland (2000) viewed the estimation from an empirical Bayes context. Ledoit and Wolf (2003, 2004a,b) and Schäfer and Strimmer (2005) determined the optimal λ analytically without specifying any underlying distributions or the need for computationally expensive techniques such as MCMC, bootstrap or cross-validation.

2.3.2. Analytical Determination of Shrinkage Parameter

Suppose the first two moments of the distributions of $\boldsymbol{U}$ and $\boldsymbol{T}$ exist. Schäfer and Strimmer (2005) showed that analytically minimizing the risk function of equation (2.10) with respect to λ gives the following optimal value

$$\lambda^\star = \frac{\sum_{i=1}^{m} \mathrm{Var}(u_i) - \mathrm{Cov}(t_i, u_i) - \mathrm{Bias}(u_i)E(t_i - u_i)}{\sum_{i=1}^{m} E\left((t_i - u_i)^2\right)}, \tag{2.11}$$

for which the MSE of $R(\lambda^\star)$ is minimized. Here, if $\boldsymbol{U}$ is an unbiased estimator of $\boldsymbol{\Psi}$, equation (2.11) simplifies to

$$\lambda^\star = \frac{\sum_{i=1}^{m} \mathrm{Var}(u_i) - \mathrm{Cov}(t_i, u_i)}{\sum_{i=1}^{m} E\left((t_i - u_i)^2\right)}. \tag{2.12}$$

Equation (2.11) provides a number of insights into the choice of the optimal shrinkage intensity:

1. The shrinkage parameter is directly proportional to the variance of the unregularized estimate $\boldsymbol{U}$. With increasing sample size, the variance will be expected to decrease, thereby resulting in a decrease in the shrinkage intensity. Consequently, the influence of the target estimate $\boldsymbol{T}$ on the shrinkage estimate $\boldsymbol{U}^\star$ diminishes.

2. $\lambda^\star$ is dependent on the correlation between the estimation error of $\boldsymbol{U}$ and $\boldsymbol{T}$. In the presence of positive correlation, the weight assigned to the shrinkage target decreases. Thus, the inclusion of the second term in the numerator of equation (2.11) adjusts for the fact that both estimators are inferred from the same data set.

3. $\lambda^\star$ is inversely proportional to the mean squared difference between the unregularized and target estimates, $\boldsymbol{U}$ and $\boldsymbol{T}$. Hence $\lambda^\star$ decreases with increasing mean squared difference. This penalizes against the misspecification of a target estimate.

4. The shrinkage intensity reduces if the unregularized estimator is biased towards the target.

5. In cases where the variables by design are kept identical in both the unregularized and target estimators, these variables tend not to play any vital role in the determination of the shrinkage intensity. Their contributions to the various terms in equation (2.11) cancel out.

6. $\lambda^\star$ is invariant to translations. This is, however, not true with rotation or scaling. Thus the underlying data may be centered without affecting the estimation of the optimal shrinkage intensity.

The estimation of the optimal shrinkage $\lambda^\star$ has been viewed from two different ways: (1) unbiased estimation, and (2) consistent estimation. Ledoit and Wolf (2003), based on the concept of consistency, replaced the unknown terms in equation (2.11) with consistent estimators. In their paper, Schäfer and Strimmer (2005) argued that since consistency is an asymptotic property and a basic requirement for any sensible estimator, this is only a weak requirement. Instead, the authors proposed replacing the unknown terms in equation (2.11) with their respective unbiased estimators with small adjustments made to avoid over-shrinkage or negative shrinkage in finite samples.

CHAPTER III

ON IMPROVING THE BH AND SOME ADAPTIVE BH PROCEDURES UNDER INDEPENDENCE AND DEPENDENCE

3.1. Introduction

Generally, due to dimensionality issues and the breakdown of standard methods of multivariate analysis, the classical approach to multiple testing in high-dimensional data is to first test each hypothesis individually. This usually consists of computing a one-dimensional test statistic for each hypothesis under the constraint of the null hypothesis. The observed test statistics and their corresponding null distributions are then utilized to obtain p-values for each test statistic. A multiple testing procedure is then applied to the set of p-values to determine significance thresholds that probabilistically control a measure of overall error rate at a pre-specified level α.

In the following, we will denote an m-dimensional vector of statistics, say $\boldsymbol{\theta}_n$, by $\boldsymbol{\theta}_n = \big(\boldsymbol{\theta}_n(1), \boldsymbol{\theta}_n(2), \ldots, \boldsymbol{\theta}_n(m)\big)$. As in section 2.1, consider the random sample $\mathcal{X}_n = (X_1, \ldots, X_n)$ of n independent and identically distributed (i.i.d) random variables from a data-generating distribution $P \in \Omega$. Here, Ω, may be a parametric, semiparametric or nonparametric statistical model. Consider testing m null hypotheses $H_0(i),\ i = 1, \ldots, m$, simultaneously based on a vector of test statistics, $\boldsymbol{T}_n = (\boldsymbol{T}_n(i) : i = 1, \ldots, m)$, under the data-generating distribution P. Given a test statistic null distribution Q_0, the unadjusted p-values $P_{0,n} = (P_i : i = 1, \ldots, m)$, are defined as

$$\begin{aligned} P_i &= \inf\{\alpha \in [0,1] : \text{Reject } H_0(i) \text{ at single test level } \alpha\} \\ &= \inf\{\alpha \in [0,1] : \boldsymbol{T}_n(i) \in \mathcal{C}_n(i;\alpha)\}, \quad i = 1, \ldots, m, \end{aligned} \tag{3.1}$$

where $\mathcal{C}_n(i;\alpha) = \mathcal{C}_n(T_n, Q_{0,i}, \alpha)$ are the rejection regions and are chosen such that

$$P_{Q_{0,i}}\big(\boldsymbol{T}_n(i) \in \mathcal{C}_n(i;\alpha)\big) \leq \alpha. \tag{3.2}$$

Herein, we assume the rejection regions are nested in the sense that $\mathcal{C}_n(i;\alpha) \subseteq \mathcal{C}_n(i;\alpha')$ if $\alpha \leq \alpha'$. The use of the long notation in $\mathcal{C}_n(\boldsymbol{T}_n, Q_{0,i}, \alpha)$ indicates that the unadjusted p-values are a function of the test statistics, $\boldsymbol{T}_n(i)$, the null distribution of the test statistics, Q_0, and the pre-specified significance level, α. Now, let the ordered unadjusted p-values be denoted by

$$P_{(1)} \leq \cdots \leq P_{(m)} \tag{3.3}$$

with corresponding null hypotheses $H_0^{(1)}, H_0^{(2)}, \cdots, H_0^{(m)}$. Then the linear step-up BH procedure rejects all null hypotheses $H_0^{(1)}, H_0^{(2)}, \cdots, H_0^{(k)}$, where

$$k = \max\left\{1 \le i \le m : P_{(i)} \le \frac{i}{m}\alpha\right\}.$$

This procedure does not reject any hypotheses if no such k exists. The corresponding adjusted p-values are given by

$$\tilde{P}_{(i)}^{BH} = \min_{k=i,\dots,m}\left\{\min\left\{\frac{i}{k}P_{(k)},\ 1\right\}\right\}, \quad i = 1,\dots m. \tag{3.4}$$

Thus, the linear step-up BH procedure at level q is equivalent to rejecting all hypotheses whose adjusted p-values is at most q. The BH procedure is summarized in algorithm 3.1. Benjamini and Hochberg (1995) proved that this procedure controls FDR at level α under the assumption of independent p-values. As already discussed, this procedure is conservative by a factor of $m_0/m = \pi_0$, the proportion of true null hypotheses, if there is at least one false null hypothesis. The elegant mathematical idea behind the BH procedure has drawn considerable attention from statisticians in the field of multiple testing. One line of research has been to study the robustness of the procedure to independence while another direction has been to incorporate an estimate of π_0 to improve the upper bound.

Algorithm 3.1 The BH linear step-up procedure

1. Let $p_{(1)} \le p_{(2)} \le \cdots \le p_{(m)}$ be the ordered observed p-values and let $H_0^{(1)}, H_0^{(2)}, \cdots, H_0^{(m)}$ be the corresponding null hypotheses.

2. Calculate $k = \max\left\{1 \le i \le m : p_{(i)} \le \frac{i}{m}\alpha\right\}$

3. If k exists, reject all null hypotheses corresponding to $p_{(1)} \le p_{(2)} \le \cdots \le p_{(k)}$. Otherwise reject nothing.

3.1.1. Adaptive BH Procedures

Adaptive procedures are those in which the number of true null hypotheses is estimated, and the threshold in the BH procedure is adjusted accordingly. Schweder and Spjøtvoll (1982) pioneered the estimation of m_0 from the quantile plot of the p-values against their ranks. Following the work of Schweder and Spjøtvoll (1982), Storey, Taylor, and Siegmund (2004) suggested similar estimates of m_0, and proposed one of the most widely utilized adaptive procedures. Adaptive procedures have a tremendous gain in power when many hypotheses are false. It is therefore not surprising that many statisticians in the past two decades have devoted effort to developing and analyzing estimators of m_0 and related terms. See for instance, Benjamini and Hochberg (2000); Storey (2002, 2003); Storey, Taylor, and Siegmund (2004); Langaas, Lindqvist, and Ferkingstad (2005); Meinshausen and Rice (2006); Benjamini, Krieger, and Yekutieli (2006); Gavrilov, Benjamini, and Sarkar (2009); Blanchard and Roquain (2009); Celisse and Robin (2010); Zeisel, Zuk, Domany, et al. (2011); Chen and Doerge (2012); Liang and Nettleton (2012); Heesen and Janssen (2016). However, to the best of our knowledge, control of the FDR has been rigorously established for only a few such procedures under the assumption of independent p-values, including the linear step-up procedure of Storey, Taylor, and Siegmund (2004) and the two-stage adaptive procedure of Benjamini, Krieger, and Yekutieli (2006), herein referred to as STS and BKY respectively.

3.1.1.1. STS Adaptive Linear Step-up Procedure

Suppose $\lambda \in (0, 1)$ is a tuning parameter, then Storey, Taylor, and Siegmund (2004) suggest the following estimator for m_0:

$$\widehat{m}_0(\lambda) = \frac{\#\{p_i > \lambda\} + 1}{(1 - \lambda)} \tag{3.5}$$

The rationale behind this estimator is as follows. Provided a test has reasonable power, most of the large p-values are likely to correspond to true null hypotheses. Consequently, if the true null p-values have approximately a uniform $[0, 1]$ distribution, then, one would expect about $m_0(1 - \lambda)$ of the p-values to lie in the interval $(\lambda, 1]$. The addition of one in the numerator of equation (3.5) is a small sample adjustment to avoid an estimator of zero. Having estimated m_0, one adjusts the BH procedure with the estimate accordingly. The STS adaptive linear step-up procedure is summarized in algorithm 3.2.

Algorithm 3.2 The STS adaptive linear step-up procedure

1. Estimate m_0 using equation (3.5)
2. Use the linear step-up procedure of algorithm 3.1 with α replaced by $\alpha' = \frac{i}{\widehat{m}_0}\alpha$

Storey, Taylor, and Siegmund (2004) showed that their adaptive procedure asymptotically controls the FDR under weak dependence assumptions. In the context of microarray data analysis, Qiu, Klebanov, and Yakovlev (2005) demonstrated that the variance of the number of hypotheses rejected by the STS procedure can be intolerably high, rendering the procedure unstable, especially in the presence of strong correlations between gene expression levels.

3.1.1.2. BKY Adaptive Linear Step-up Procedure

Usually, the risk involved in adaptive procedures is high due to the estimation of m_0, consequently, such procedures can become unstable if the estimation error is not taken into account. Benjamini, Krieger, and Yekutieli (2006) suggested a more stable estimator by adjusting the α-level slightly from α to $\alpha^* = \alpha/(1 + \alpha)$ to adjust for the additional variability in estimating m_0. Algorithm 3.3 summarizes the BKY procedure. Benjamini, Krieger, and Yekutieli (2006) proved that the BKY procedure controls the FDR at level α whenever the p-values are independent. They also illustrated in a simulation study that their procedure generally still controls the FDR under positive dependence.

Algorithm 3.3 The BKY two-stage adaptive linear step-up procedure

---

1. Use Algorithm 3.1 at level $\alpha^* = \alpha/(1+\alpha)$. Suppose r_1 is the number of rejected hypotheses.
   (a) If $r_1 = 0$, do not reject any hypothesis and stop.
   (b) If $r_1 = m$, reject all m hypotheses and stop; otherwise move to step 2.
2. Calculate $\hat{m}_0 = (m - r_1)$.
3. Use Algorithm 3.1 with $\alpha' = \alpha^* m/\hat{m}_0$ on all hypotheses.

However, regardless of the procedure, the validity and accuracy of these procedures are essentially determined by whether the chosen test statistic is optimal, the null distributions are correctly or conservatively specified and whether the data are independent across tests. As discussed in earlier chapters, very often, the theoretical null or the data-generated null distribution used to derive the p-values is misspecified due to dependencies among test statistics and other possible factors (Pollard and van der Laan (2004); Efron (2004, 2007a)). Correct inference depends crucially on the accurate assessment of the null distribution. Thus, misspecifying the null distribution may lead to overly pessimistic or optimistic p-values, and thus to a violation of the implicit assumption that the truly null p-values are drawn from a uniform distribution. Various attempts have been made to account for the dependencies among p-values, but it seems more natural and perhaps even easier to deal with this on the level of the original test statistic.

Multiple testing procedures with high power, good stability and good FDR control are desirable, especially in microarray data analysis. In the following, a unified approach to FDR control is described that takes into account several aspects of multiple testing methodologies that have previously only been considered separately. A notable distinction of our approach is that a generally valid null distribution is used in place of the unknown joint distribution of the test statistics.

The remainder of this chapter is set out as follows. In section 3.2, a detailed description of a shrinkage estimator of the variance components that utilizes information across all the genes in the data is provided. The shrinkage variance components are then utilized to construct the shrinkage t statistic. In addition, a choice of an appropriate null distribution and subsequently, the proposed unified approach will be discussed. Some analytical and asymptotic results for the proposed methodologies are presented in section 3.3. Conditions under which the proposed techniques provide asymptotic FDR control are also provided. Because the proposed methodologies are based on asymptotic arguments, we conduct extensive Monte Carlo simulation studies in section 3.4 to shed light on the finite sample properties of the methods. Additionally, the FDR control, power and stability of the proposed techniques are compared to some existing FDR-controlling procedures. Finally, in section 3.5, the results of the study are discussed with conclusions and recommendations provided.

3.2. A Unified Procedure to FDR Control

In order to motivate the proposed procedure, the choice of test statistic, test statistic null distribution, and the FDR-controlling procedure will be discussed.

3.2.1. The Shrinkage t Statistic

3.2.1.1. Shrinkage Estimation of Variance Components

In this section, the shrinkage t statistic of Opgen-Rhein and Strimmer (2007) developed in the framework of James-Stein-type analytic shrinkage estimation is considered. An improved estimator of the variance components will be constructed from pooling information across individual variance estimators and subsequently utilized to construct the test statistic. The goal is to find a linear combination, $\boldsymbol{S}^{\star} = \lambda \boldsymbol{T} + (1-\lambda)\boldsymbol{S}$ of a target estimator $\boldsymbol{T}$ and a matrix of unbiased sample covariances, $\boldsymbol{S}$. The entries of $\boldsymbol{S}$ are determined by

$$s_{ij}^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_{ki}-\bar{x}_i)(x_{kj}-\bar{x}_j). \tag{3.6}$$

As discussed in Chapter 2, the search for the optimal shrinkage intensity parameter, λ, in James-Stein estimation is based on minimizing a loss function such as the mean squared error. Thus, the optimal shrinkage intensity parameter is the solution that minimizes the function

$$R(L(\lambda)) = E\left(\sum_{i=1}^{m}\sum_{j=1}^{m}\left(\lambda t_{ij} + (1-\lambda)s_{ij}^2 - \sigma_{ij}^2\right)^2\right) \tag{3.7}$$

where σ_{ij}^2 are the true covariance components. Simplifying (3.7) and using the facts that $E(s_{ij}^2) = \sigma_{ij}^2$ and $\mathrm{Var}(\sigma_{ij}^2) = 0$ gives

$$\begin{aligned} R(L(\lambda)) &= \sum_{i=1}^{m}\sum_{j=1}^{m}\left(E\left(\lambda t_{ij} + (1-\lambda)s_{ij}^2 - \sigma_{ij}^2\right)^2\right) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\mathrm{Var}\left(\lambda t_{ij} + (1-\lambda)s_{ij}^2 - \sigma_{ij}^2\right) + \left[E\left(\lambda t_{ij} + (1-\lambda)s_{ij}^2 - \sigma_{ij}^2\right)\right]^2\right\} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\mathrm{Var}\left(\lambda t_{ij} + (1-\lambda)s_{ij}^2\right) + \left[E\left(\lambda t_{ij} + (1-\lambda)s_{ij}^2 - \sigma_{i}^2\right)\right]^2\right\} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\lambda^2\mathrm{Var}(t_{ij}) + (1-\lambda)^2\mathrm{Var}(s_{ij}^2) + 2\lambda(1-\lambda)\mathrm{Cov}(t_{ij}, s_{ij}^2) + \lambda^2\left[E(t_{ij} - s_{ij}^2)\right]^2\right\} \end{aligned} \tag{3.8}$$

Now, in order to minimize (3.8) with respect to λ we have

$$\frac{d(R(L(\lambda)))}{d\lambda} = \sum_{i=1}^{m}\sum_{j=1}^{m}\Big\{2\lambda\mathrm{Var}(t_{ij}) - 2(1-\lambda)\mathrm{Var}(s_{ij}^2) + 2\mathrm{Cov}(t_{ij}, s_{ij}^2)$$

$$- 4\lambda \text{Cov}(t_{ij}, s_{ij}^2) + 2\lambda \left[E(t_{ij} - s_{ij}^2)\right]^2 \Bigg\} \tag{3.9}$$

Setting (3.9) to zero and solving for λ results in

$$\begin{aligned}
\lambda^{\star} &= \frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\text{Var}(s_{ij}^2) - 2\text{Cov}(t_{ij}, s_{ij}^2)\right\}}{\sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\text{Var}(t_{ij}) + \text{Var}(s_{ij}^2) - 2\text{Cov}(t_{ij}, s_{ij}^2) + \left[E(t_{ij} - s_{ij}^2)\right]^2\right\}} \\
&= \frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\text{Var}(s_{ij}^2) - 2\text{Cov}(t_{ij}, s_{ij}^2)\right\}}{\sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\text{Var}(t_{ij} - s_{ij}^2) + \left[E(t_{ij} - s_{ij}^2)\right]^2\right\}} \\
&= \frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\text{Var}(s_{ij}^2) - 2\text{Cov}(t_{ij}, s_{ij}^2)\right\}}{\sum_{i=1}^{m}\sum_{j=1}^{m}\left\{E\left[\left(t_{ij} - s_{ij}^2\right)^2\right]\right\}}
\end{aligned} \tag{3.10}$$

To show that this is indeed the minimizing λ, it suffices to show that the second derivative of (3.7) is greater than zero. Thus,

$$\begin{aligned}
\frac{d^2(R(L(\lambda)))}{d\lambda^2} &= \sum_{i=1}^{m}\sum_{j=1}^{m}\left\{2\text{Var}(t_{ij}) + 2\text{Var}(s_{ij}^2) - 4\text{Cov}(t_{ij}, s_{ij}^2) + 2\left[E(t_{ij} - s_{ij}^2)\right]^2\right\} \\
&= 2\sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\text{Var}(t_{ij} - s_{ij}^2) + \left[E(t_{ij} - s_{ij}^2)\right]^2\right\}
\end{aligned} \tag{3.11}$$

which is positive everywhere due to the sum of two positive terms. Hence, $R(L(\lambda^{\star}))$ is a verified minimum.

3.2.1.2. Estimation of the Optimal Shrinkage Intensity Parameter

Note that $\lambda^{\star}$ is not a bona fide estimator because of its dependence on unobservable quantities. Ledoit and Wolf (2003) suggested n-consistent estimators for the unknown parameters. Utilizing the Rao-Blackwell theorem and normality assumption, Chen, Wiesel, and Hero

(2009) improved upon Ledoit and Wolf's estimators. However, rather than utilizing n-consistent estimators for the unknown parameters, Schäfer and Strimmer (2005) proposed estimating these parameters with their unbiased counterparts. The authors (Schäfer and Strimmer (2005)) argued that consistency is a weak requirement, as consistency is an asymptotic property and a basic requirement of any sensible estimator. Therefore, since interest is in small sample inference, replacing the unknown parameters with their unbiased counterparts will suffice. Each of these three estimators for the optimal intensity performs quite well as the sample size increases. Based on the arguments of Schäfer and Strimmer (2005), the unbiased estimation technique will be employed in this study. Incorporating this, the estimated shrinkage intensity is given by

$$\widehat{\lambda}^{\star} = \min\left(1, \frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left\{\widehat{\mathrm{Var}(s_{ij}^2)} - 2\widehat{\mathrm{Cov}(t_{ij}}, s_{ij}^2)\right\}}{\sum_{i=1}^{m}\sum_{j=1}^{m}\left(t_{ij} - s_{ij}^2\right)^2}\right), \tag{3.12}$$

where adjustments have been made to avoid overshrinkage or negative shrinkage in finite samples. The corpcor package in R (R Core Team (2018)) provides a fast and efficient algorithm for obtaining the shrinkage intensity estimate.

3.2.1.3. Choice of Target Matrix

The choice of an appropriate target matrix requires some diligence and has been extensively studied in the literature. See, for instance, Ledoit and Wolf (2003); Schäfer and Strimmer (2005); Warton (2008, 2010); Fisher and Sun (2011) and the references therein. The target matrix is often chosen to be positive definite and well-conditioned, and consequently, the final regularized estimate, $\boldsymbol{S}^{\star}$ is guaranteed to be positive definite and well-conditioned for any dimensionality. Some of the target matrices studied in the literature are: (i) diagonal, unit variance, (ii) diagonal, common variance, (iii) common covariance, (iv) diagonal, unequal variance, (v) perfect positive correlation and (vi) constant correlation.

3.2.1.4. Construction of the Shrinkage t Statistic

The focus for this section will be on constructing test statistics for a parameter vector

$$\boldsymbol{\theta}(P) = (\theta_1(P), \ldots, \theta_m(P)) \tag{3.13}$$

The test may be a one-sided testing problem, in which case (without loss of generality)

$$H_0(i) : \theta(i) \leq \theta_0(i) \quad \text{vs.} \quad H_1(i) : \theta(i) > \theta_0(i) \tag{3.14}$$

or a two-sided testing problem, in which case

$$H_0(i) : \theta(i) = \theta_0(i) \quad \text{vs.} \quad H_1(i) : \theta(i) \neq \theta_0(i) \tag{3.15}$$

The test statistics utilized for such analyses will be based on an estimate $\widehat{\boldsymbol{\theta}}_n = \left(\widehat{\boldsymbol{\theta}}_n(i), \ldots, \widehat{\boldsymbol{\theta}}_n(m)\right)$ computed using the data, $\mathcal{X}_n$. Then, the "studentized" test statistic for testing the one-sided and two-sided tests are given respectively by

$$\boldsymbol{T}_n(i) = \frac{\sqrt{n}\left(\widehat{\theta}_n(i) - \theta_0(i)\right)}{\widehat{\sigma}_n(i)}, \qquad \text{and}$$

$$\boldsymbol{T}_n(i) = \frac{\sqrt{n} \mid \widehat{\theta}_n(i) - \theta_0(i) \mid}{\widehat{\sigma}_n(i)} \tag{3.16}$$

where $\widehat{\sigma}_n(i)$ is an estimate of the standard deviation of $\sqrt{n}\left(\widehat{\theta}_n(i) - \theta_0(i)\right)$. Then using the estimated shrinkage variance components, the proposed modified t statistic is obtained by replacing $\widehat{\sigma}_n(i)$ with the estimated shrinkage standard deviation $s^{\star}_{ii}$ in the 'studentized' test statistic. For instance, for the two-sample t test for testing the null hypotheses of no differences in group means, the modified test statistic for testing each hypothesis becomes

$$\boldsymbol{T}_n^{\star}(i) = \frac{\bar{x}_{i1} - \bar{x}_{i2} - (\mu_{i1} - \mu_{i2})}{\sqrt{\frac{s_{i1}^{\star 2}}{n_1} + \frac{s_{i2}^{\star 2}}{n_2}}} \tag{3.17}$$

where n_1 and n_2 are the sample sizes in groups 1 and 2 respectively, and $s_{i1}^{\star 2}$ is the shrinkage estimate of the variance for group 1 for the i^{th} variable, and $s_{i2}^{\star 2}$ is the shrinkage estimate of the variance for group 2 for the i^{th} variable. It should be noted that only the diagonal elements of the shrinkage covariance matrix are being utilized, thus, in an analysis of this nature, it makes no sense to consider the estimation of the full covariance matrix.

Although different techniques to modifying the usual "studentized" test statistic have been studied in the context of differential expression, Opgen-Rhein and Strimmer (2007) were the first to propose a modified statistic that employs a variance shrinkage estimator that is fully analytic and requires no distributional assumptions. Their proposed statistic, shrinks the variance components to a common median. In the exploration for other possible shrinkage targets, the authors considered shrinking the variances against zero or towards the mean, but these other two shrinkage targets were suboptimal. Additionally, the authors approximated the covariance between the individual unregularized variances and that of the shrinkage target to be zero without any justification. This study will construct test statistic analoguous to the shrinkage statistic of Opgen-Rhein and Strimmer (2007), however, no assumptions will be made about the covariance between the individual unregularized variances and that of the shrinkage target.

3.2.2. Test Statistic Null Distribution

Recall that the p-values are functions of the distribution of the test statistics. In practice, however, this distribution is often unknown and replaced by a test statistic null distribution. The appropriate choice of null distribution is thus crucial to ensure control of the type I error rate under the assumed null distribution. Current practices utilize either a theoretical null or a data-generated null such as the permutation null. As illustrated by Efron (2004, 2007a) and Pollard and van der Laan (2004), these commonly utilized null distributions could result in misleading results, especially in the presence of strong correlations between the variables. Instead, Pollard and van der Laan (2004); Dudoit, van der Laan, and Pollard (2004); Dudoit, van der Laan, and Birkner (2004); van der Laan, Dudoit, and Pollard (2004) and van der Laan and Hubbard (2006) have provided a general characterization for a proper test statistics null distribution based on resampling theory. These distributions are briefly discussed in the following but detailed explanations are provided in section 2.2. The interested reader is referred to Dudoit and van der Laan (2008, Chapter 2) for the explicit construction and theoretical justifications of these null distributions.

3.2.2.1. Null Shift and Scale-transformed Test Statistic Null Distribution

Suppose there exists an m-dimensional known real-valued vector $\boldsymbol{\lambda}_0 = \big(\boldsymbol{\lambda}_0(i); i = 1, \ldots, m\big)$ and a positive real-valued vector $\boldsymbol{\tau}_0 = \big(\boldsymbol{\tau}_0(i); i = 1, \ldots, m\big)$ of null values such that

$$\limsup_{n\to\infty} E\big(\boldsymbol{T}_n(i)\big) \leq \boldsymbol{\lambda}_0(i) \qquad \text{and}$$
$$\limsup_{n\to\infty} \operatorname{Var}\big(\boldsymbol{T}_n(i)\big) \leq \boldsymbol{\tau}_0(i) \qquad \text{for } i \in \mathcal{H}_0. \tag{3.18}$$

The null shift and scale-transformed null distribution, $Q_0^{NS}(P)$, is defined as the asymptotic distribution of the m-dimensional vector of null shift and scale-transformed test statistics

$$\boldsymbol{Z}_n^{NS}(i) = \sqrt{\min\left(1, \frac{\boldsymbol{\tau}_0(i)}{\text{Var}\left(\boldsymbol{T}_n(i)\right)}\right)}\Big(\boldsymbol{T}_n(i) + \boldsymbol{\lambda}_0(i) - E(\boldsymbol{T}_n(i))\Big); \quad i = 1, \ldots, m. \tag{3.19}$$

3.2.2.2. Null Quantile-transformed Test Statistic Null Distribution

Suppose for a real-valued number z and for each $i \in \mathcal{H}_0$, there exits a marginal null distribution $Q_{0,i}$ such that

$$\liminf_{n\to\infty} \; Q_{0,i}^{-1} Q_{n,i}(z) \geq z \tag{3.20}$$

where $Q_{n,i}$ is the i^{th} marginal distribution of the true distribution of the test statistic, Q_n. Then, the null quantile-transformed null distribution, $Q_0^{NQ}(P)$, is defined as the joint distribution of the m-dimensional vector of null quantile-transformed test statistics

$$Z_n^{NQ}(i) = Q_{0,i}^{-1} Q_{n,i}^{\Delta}\big(\boldsymbol{T}_n(i)\big) \tag{3.21}$$

where $Q_{n,i}^{\Delta}(z) = \Delta Q_{n,i}(z) + (1-\Delta)Q_{n,i}(z^-)$ and Δ is a uniform random variable on the interval $[0, 1]$ and independent of the data.

The two distributions described above are dependent on the data-generating distribution P, which is often unknown in practice. Thus, one needs to estimate the joint distributions, Q_0^{NS} and Q_0^{NQ}. As proposed by Dudoit, van der Laan, and Birkner (2004) and van der Laan and Hubbard (2006), bootstrap techniques may be utilized to obtain consistent estimators, Q_{0n}^{NS} and Q_{0n}^{NQ} of the test statistic null distributions. The bootstrap estimation may be summarized as follows. Let P_n denote the empirical distribution of $X_{i1}, \ldots, X_{in}; \; i = 1, \ldots, m$

which assigns probability $(1/n)$ to each realization X_{ij} and let $X^{\star}_{i1}, \ldots, X^{\star}_{in}$ be i.i.d. sample observations from P_n. Generate an $m \times B$ matrix of test statistics $\boldsymbol{Z}^{\star}_n$ (either the null shift and scale-transformed test statistics, $\boldsymbol{Z}^{NS}_n(i)$, or the null quantile-transformed test statistics, $\boldsymbol{Z}^{NQ}_n(i)$) based on the bootstrap data $X^{\star}_{ij}$. Then the bootstrap estimator of the null distribution is the empirical distribution of the B columns of $\boldsymbol{Z}^{\star}_n$. The bootstrap estimation of the two null distributions based on the shrinkage t statistic are detailed in algorithms 3.4 and 3.5. In general, there is no recommendation for the number of bootstrap samples, B, to utilize. But, in order to deal with the discreteness of the bootstrap distribution, one obviously needs a very large B. In practice, however, one needs to find a balance between computational cost and estimation accuracy.

3.2.3. Proposed Unified Approach

To this end, the proposed unified approach is as follows. Without loss of generality, consider the one-sided or two-sided testing problem

$$H_0(i) : \theta(i) \leq \theta_0(i) \quad \text{vs.} \quad H_1(i) : \theta(i) > \theta_0(i) \qquad \text{or} \tag{3.27}$$

$$H_0(i) : \theta(i) = \theta_0(i) \quad \text{vs.} \quad H_1(i) : \theta(i) \neq \theta_0(i); \qquad i = 1, \ldots, m. \tag{3.28}$$

Here, the hypothesized null values, $\theta_0(i)$, are normally zero. For instance, in microarray experiments, if one is interested in looking at gene expression in cancer tumor versus normal tissue, the null hypotheses would be $H_0(i)$: the gene does not differentially express. The proposed techniques for such hypothesis testing is detailed in the following. First, calculate the test statistics of each hypothesis using the shrinkage t statistic discussed in section 3.2.1. Next, using either Algorithm 3.4 or 3.5 estimate the null shift and scale-transformed null

Algorithm 3.4 Bootstrap Estimation of the Null Distribution, $Q_0^{NS}(P)$

1. Generate $X_{i1}^\star, \ldots, X_{in}^\star$ as a random sample taken with replacement from the given data, $X_i = \{X_{i1}, \ldots, X_{in}\}$ $i = 1, \ldots m$.

2. Compute the estimate $\widehat{\boldsymbol{\theta}}_n^\star(i)$ of the same functional form as the original estimator $\widehat{\boldsymbol{\theta}}_n(i)$.

3. Compute the estimate of the shrinkage variance components using the formula

$$\sigma_{ii}^{\star\star 2} = \hat{\lambda}^{\star\star} t_{ii}^\star + (1 - \hat{\lambda}^{\star\star}) s_{ii}^{\star 2}, \tag{3.22}$$

   with the optimal estimated shrinkage intensity parameter estimated using equation (3.12) based on the bootstrap data.

4. Using the estimated quantities, compute the bootstrap shrinkage t statistic for the one-sided or two sided test respectively as given in equation (3.16)

5. Repeat steps 1 - 4 B times to obtain B bootstrap shrinkage t statistics, $\boldsymbol{T}_n^\star$, which can be arranged in an $m \times B$ matrix with each row corresponding to the m null hypotheses and each column to the B bootstrap samples.

6. Obtain $E(\boldsymbol{T}_n^\star(i))$ and $\mathrm{Var}(\boldsymbol{T}_n^\star(i))$, by computing the row means and row variances of the matrix, $\boldsymbol{T}_n^\star$.

7. Obtain an $m \times B$ matrix, $\tilde{\boldsymbol{Z}}_n^\star$ of null value shifted and scaled bootstrap test statistics

$$\tilde{\boldsymbol{Z}}_n^\star(i) = \sqrt{\min\left(1, \frac{\boldsymbol{\tau}_0(i)}{\mathrm{Var}_{P_n}\left(\boldsymbol{T}_n^\star(i)\right)}\right)} \left(\boldsymbol{T}_n^\star(i) + \boldsymbol{\lambda}_0(i) - E(\boldsymbol{T}_n^\star(i))\right); \quad i = 1, \ldots, m \tag{3.23}$$

   where $\boldsymbol{\tau}_0(i)$ and $\boldsymbol{\lambda}_0(i)$ are user-supplied null values.

8. The bootstrap estimate Q_{0n}^{NS} of $Q_0^{NS}(P)$ is the empirical distribution of the columns of $\tilde{\boldsymbol{Z}}_n^\star$.

Algorithm 3.5 Bootstrap Estimation of the Null Distribution, $Q_0^{NQ}(P)$

1. Generate $X_{i1}^\star, \ldots, X_{in}^\star$ as a random sample taken with replacement from the given data, $X_i = \{X_{i1}, \ldots, X_{in}\}$ $i = 1, \ldots m$.

2. Compute the estimate $\widehat{\boldsymbol{\theta}}_n^\star(i)$ of the same functional form as the original estimator $\widehat{\boldsymbol{\theta}}_n(i)$.

3. Compute the estimate of the shrinkage variance components using the formula

$$\sigma_{ii}^{\star\star 2} = \hat{\lambda}^{\star\star} t_{ii}^\star + (1 - \hat{\lambda}^{\star\star}) s_{ii}^{\star 2}, \tag{3.24}$$

   with the optimal estimated shrinkage intensity parameter estimated using equation (3.12) based on the bootstrap data.

4. Using the estimated quantities, compute the bootstrap shrinkage t statistic for the one-sided or two sided test respectively as given in equation (3.16)

5. Repeat steps 1 - 4 B times to obtain B bootstrap shrinkage t statistics, $\boldsymbol{T}_n^\star$, which can be arranged in an $m \times B$ matrix with each row corresponding to the m null hypotheses and each column to the B bootstrap samples.

6. Obtain an $m \times B$ matrix, $\breve{\boldsymbol{Z}}_n^\star$ of null quantile-transformed bootstrap test statistics

$$\breve{\boldsymbol{Z}}_n^\star(i) = q_{0,i}^{-1} Q_{n,i}^{\star,\Delta}(\boldsymbol{T}_n^\star(i)), \tag{3.25}$$

   based on user-supplied null distributions, $q_{0,i}$ and where $Q_{n,i}^{\star,\Delta}(z) = \Delta Q_{n,i}^\star(z) + (1 - \Delta) Q_{n,i}^\star(z^-)$, Δ is a uniform random variable on the interval $[0, 1]$, independent of the data. $Q_{n,i}^\star(z)$ is the marginal CDF defined as

$$Q_{n,i}^\star(z) = \frac{1}{B} \sum_{b=1}^{B} I\left(\boldsymbol{T}_n^\star(i) \leq z\right) \tag{3.26}$$

7. The bootstrap estimate of estimate Q_{0n}^{NQ} of $Q_0^{NQ}(P)$ is the empirical distribution of the columns of $\breve{\boldsymbol{Z}}_n^\star$.

distribution or the null quantile-transformed null distribution. Having estimated the null distributions, the unadjusted p-values are obtained by

$$p_i^\star = \Pr_{Q_{0n}^d}\big(\boldsymbol{Z}(i) \geq \boldsymbol{t}_n(i)\big) = \frac{1}{B}\sum_{b=1}^{B} I\Big(\boldsymbol{Z}_n^\star(i) \geq \boldsymbol{t}_n(i)\Big); \qquad i = 1, \ldots, m \tag{3.29}$$

for the one-sided testing problem and

$$p_i^\star = \Pr_{Q_{0n}^d}\big(\boldsymbol{Z}(i) \geq \boldsymbol{t}_n(i)\big) = \frac{1}{B}\sum_{b=1}^{B} I\Big(\big|\boldsymbol{Z}_n^\star(i)\big| \geq \big|\boldsymbol{t}_n(i)\big|\Big); \qquad i = 1, \ldots, m \tag{3.30}$$

for the two-sided testing problem and where $d =$ NS or NQ, $\boldsymbol{t}_n(i)$ is the observed statistic from the original data and $\boldsymbol{Z}_n^\star(i)$ is either the bootstrap estimate of the null shift and scale-transformed test statistics, $\tilde{\boldsymbol{Z}}_n^\star(i)$, or the null quantile-transformed test statistics, $\breve{\boldsymbol{Z}}_n^\star(i)$. Finally, apply the BH procedure of algorithm 3.1 or the BKY procedure of algorithm 3.3 utilizing the estimated unadjusted p-values.

3.3. Asymptotic Results

A formal theoretical framework to ascertain asymptotic control of FDR by the proposed methodologies are detailed in this section. In their paper, Dudoit, van der Laan, and Pollard (2004) provide four fundamental theorems that determine asymptotic control of general type I error rates defined as functions of the number of false positives for single step procedures under the null shift and scale-transformed null distribution or an estimate thereof. These theorems depend entirely on the concept of asymptotic null domination of their proposed null distribution $Q_0^{NS}(P)$ with respect to the true distribution of the test statistics $Q_n(P)$, and on convergence of an estimated null distribution $Q_{0n}^{NS}(P)$ to $Q_0^{NS}(P)$. Applying these theorems, van der Laan and Hubbard (2006) provided another theorem that determined asymptotic control of general type I error rates under the null quantile-transformed null distribution. In the following, we establish that a simple application of these theorems provides asymptotic control of FDR by the proposed unified approach.

Theorem 3.3.1 (Asymptotic Control of FDR)

Consider the problem of testing the null hypotheses $H_0(i),\ i = 1, \ldots, m$ defined by (3.14) or (3.15) based on a random m-vector of test statistics, $\boldsymbol{T}_n = \big(\boldsymbol{T}_n(i) : i = 1, \ldots, m\big)$ given by (3.16) with unknown true distribution $Q_n = Q_n(P)$. Let $\mathcal{H}_0 = \mathcal{H}_0(P)$ be the set of true null hypotheses and $\mathcal{H}_1 = \mathcal{H}_1(P)$, the set of false null hypotheses where P is the data-generating distribution. Given an m-variate test statistic null distribution Q_0, with marginal cumulative distribution functions $Q_{0,i}$, define unadjusted p-values

$$\boldsymbol{P}_{0n}(i) = 1 - Q_{0,i}\big(\boldsymbol{T}_n(i)\big), \qquad \textit{and} \tag{3.31}$$

$$\boldsymbol{P}_0(i) = 1 - Q_{0,i}\big(\boldsymbol{Z}(i)\big) \tag{3.32}$$

for the random m-vector of test statistics $\boldsymbol{T}_n \sim Q_n$ and a random m-vector $\boldsymbol{Z} \sim Q_0$. Suppose further that

i. the $\mathcal{H}_0$-specific unadjusted p-values satisfy the following null domination assumption asymptotically. For each $x \in [0, 1]$

$$\limsup_{n\to\infty} P_{Q_n}\big(\boldsymbol{P}_{0n}(i) < x\big) \le P_{Q_0}\big(\boldsymbol{P}_0(i) < x\big). \tag{3.33}$$

ii. the joint distribution of the test statistics is positive regression dependent on the subset of test statistics corresponding to the true null hypotheses.

Then, the BH linear step-up procedure (Algorithm 3.1) based on the unadjusted p-values in

(3.32) provides asymptotic control of the FDR. That is

$$\limsup_{n\to\infty} FDR_{BH} \leq \pi_0\alpha \leq \alpha. \tag{3.34}$$

Similarly, the BKY two-stage adaptive linear step-up procedure (Algorithm 3.3) based on the unadjusted p-values in (3.32) provides asymptotic control of the FDR. That is

$$\limsup_{n\to\infty} FDR_{BKY} \leq \alpha. \tag{3.35}$$

Here, Q_0 can be any of the two null distributions, Q_0^{NS} or Q_0^{NQ} described in setion 3.2.2. The proof of theorem 3.3.1 will be based on the ensuing remark.

Remark 3.3.1

Theorem 2 of Dudoit, van der Laan, and Pollard (2004) and Theorem 2.2 of Dudoit and van der Laan (2008) established that the null shift and scale-transformed null distribution Q_0^{NS} *satisfies the asymptotic joint null domination assumption for the* $\mathcal{H}_0$*-specific subvector of test statistics. That is, for each* $z \in \mathbb{R}^{m_0}$*,*

$$\limsup_{n\to\infty}\ Q_{n,\mathcal{H}_0}(z) \leq Q_{0,\mathcal{H}_0}(z). \tag{3.36}$$

Similarly, van der Laan and Hubbard (2006) established the asymptotic joint null domination assumption for the null quantile-transformed null distribution Q_0^{NS}*. The asymptotic null domination assumption of the unadjusted p-values is an immediate consequence of the joint null domination assumption since the p-values are a function of the data.*

Before proceeding to the proof of the theoretical results, we define the concept of positive regression dependence (PRD) and positive regression dependence on a subset (PRDS) established by Benjamini and Yekutieli (2001). Recall that a set D is said to be increasing if for

$x \leq y$ and $x \in D$ implies $y \in D$.

Definition 3.3.1 (Positive Regression Dependency)
Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be an n-dimensional random vector. The multivariate distribution of $\boldsymbol{X}$ is said to be positive regression dependent if for any increasing set D, $P(X \in D|X_1 = x_1, \ldots, X_i = x_i)$ is nondecreasing in x.

Definition 3.3.2 (Positive Regression Dependency on Subsets)
Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be an n-dimensional random vector. The multivariate distribution of $\boldsymbol{X}$ is said to be positive regression dependent on a subset I_0 if for any increasing set D, and for every index $i \in I_0$, $P(X \in D|X_i = x)$ is nondecreasing in x.

Benjamini and Yekutieli (2001) indicate that the PRDS property is a relaxed form of the PRD property, in that for the PRDS property, the conditioning is always on one variable, and is required to hold only for a subset I_0 of variables. The authors also point out that for dependent test statistics, the PRDS propety is a suitable technical property to prove FDR control of various stepwise procedures.

Proof of Thereom 3.3.1. Following remark 3.3.1, the asymptotic null domination assumption in (3.33) is satisfied. By theorem 1.2 of Benjamini and Yekutieli (2001), we know that the BH procedure controls the FDR at a level less than or equal to $\pi_0\alpha$ under the PRDS property. Thus, under the null distribution and the assumption of positive regression dependent on the subset of test statistics corresponding the true null hypotheses, we will have

$$E\left(\frac{V_0}{\max(1, R_0)}\right) \leq \pi_0\alpha, \tag{3.37}$$

where R_0 and V_0 are respectively, the number of null hypotheses rejected and the number of type I errors under the null distribution. Now, if we let R_n and V_n be the number of null hypotheses rejected and the number of type I errors respectively under the true distribution of the test statistics then it follows from (3.33) that

$$\limsup_{n\to\infty} E\left(\frac{V_n}{\max(1,R_n)}\right) \leq E\left(\frac{V_0}{\max(1,R_0)}\right) \leq \pi_0\alpha \leq \alpha. \tag{3.38}$$

Hence the result in (3.34). The result of (3.35) also follows from similar arguments. □

Now, since the null distribution $Q_0(P)$ is dependent on the unknown data-generating distribution, P, it is infeasible to obtain the unadjusted p-values. Consistent estimators $Q_{0n}(P)$ of $Q_0(P)$ and corresponding unadjusted p-values may be obtained by bootstrap techniques as detailed in algorithms 3.4 and 3.5. Thus, if the null distribution is consistently estimated, then the estimated null distribution will also asymptotically dominate the true distribution of the test statistics and subsequently provide asymptotic control of the FDR. The results are summarized in the following corollary.

Corollary 3.3.1 (Asymptotic Control of FDR based on Consistent Estimation of Q_0) *Let $Q_{0n}(P)$ be an estimate of $Q_0(P)$ and define the estimated unadjusted p-values $p_i^\star$ by (3.29) or (3.30). Suppose that $Q_{0n}(P)$ converges weakly to $Q_0(P)$. Then for each $x \in [0,1]$, the estimated unadjusted p-values satisfy the asymptotic null domination assumption*

$$\limsup_{n\to\infty} P_{Q_n}\big(\boldsymbol{P}_{0n}(i) < x\big) \leq P_{Q_{0n}}\big(P_i^\star < x\big) \tag{3.39}$$

and the BH and BKY procedures based on $p_i^\star$ provide asymptotic control of the FDR.

Proof of Corollary 3.3.1. To prove the corollary, it suffices to show the asymptotic consistency of the bootstrap estimate. To this end, let P_n denote the empirical distribution of $X_1, \ldots X_n$, putting mass $1/n$ on each X_i. Then, as n approaches infinity, P_n approximates the true data-generating distribution P so that $\boldsymbol{Z}_n^\star \to \boldsymbol{Z} \sim Q_0$ in distribution, conditional on P_n. Hence, Q_{0n} converges weakly to Q_0 conditional on the data (see for example Bickel and

Freedman (1981); van der Vaart and Wellner (1996)). It then follows that the bootstrap null distribution Q_{0n} asymptotically dominates the true distribution Q_n. By the continuous mapping theorem, the distribution of the $\mathcal{H}_0$-specific bootstrap unadjusted p-values, $p_i^{\star}$, converges weakly to the distribution of the $\mathcal{H}_0$-specific unadjusted p-values, $\boldsymbol{p}_0(i)$, under the null distribution. □

We note that the estimation of the p-values in the proposed procedure is based on a generally valid joint null distribution via resampling, allowing the possibility of some how accounting for dependency in the test statistics. Therefore, one may expect a gain in power by using this procedure relative to the naïve use of the BH or BKY procedure. We therefore state the following proposition without proof.

Proposition 3.3.1

Under the assumption of positive regression dependence of the joint distribution of the test statistics and as n goes to infinity, the proposed unified approach provides a gain in power relative to the BH linear step-up or the BKY two-stage adaptive linear step-up procedures.

3.4. Simulation Study

Since the proposed procedure relies on asymptotic arguments, it is essential to analyze its finite sample performance via simulations. The current section presents a Monte Carlo simulation study to compare the FDR control, power, and stability of the proposed techniques to some existing techniques in the context of testing a mean difference for two populations. It is infeasible to carry out a comprehensive simulation study capturing all possible behaviors of the hypotheses, but various different realistic scenarios which might be encountered in practice were examined. This included changing the proportion of non-null hypotheses and their dependency structure.

3.4.1. Simulation Study Design

Consider a case-control microarray experiment with m genes and n arrays of which n_1 are from the cases and $n_0 = n - n_1$ are from the controls. In the study, $m = 1,000$ was utilized and the total sample size n, the number of differentially expressed genes, and the patterns of correlations among the genes were varied. The total sample size was set as $n = 20,\ 30,\ 40,\ 50,\ 60,\ 100,\ 200,\ 300,\ 500$ with the number of cases and controls always set equal. Two different distributions were considered for the m-dimensional arrays: normal and gamma distributions.

In all, nine procedures were investigated. These procedures were:

1. The original linear step-up procedure of Benjamini and Hochberg (1995), denoted by BH. See algorithm 3.1 for details. This procedure has been shown to control the FDR at level $\alpha\pi_0 \leq \alpha$ under independence and some types of positive dependence among the test statistics.

2. The adaptive linear step-up procedure of Storey, Taylor, and Siegmund (2004), denoted by STS. The details of this procedure are provided in algorithm 3.2. This procedure has been shown to control the FDR at level α under independence and some types of weak dependence among the test statistics.

3. The linear step-up procedure of Benjamini and Yekutieli (2001), denoted by BY. The critical values for this procedure are obtained by dividing the significance level α by $\sum_{i=1}^{m}(1/i)$. This procedure has been shown to provide control of the FDR at level α under general dependence. It can, however, be extremely conservative.

4. The adaptive two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli (2006) detailed in algorithm 3.3, denoted by BKY. This procedure has been shown to

control the FDR at level α under independence and some types of positive dependence among the test statistics.

5. The BH procedure based on the shrinkage t test statistic, denoted by S-BH.

6. The BH procedure incorporating the shrinkage t test statistic and the null shift and scale-transformed test statistic null distribution, denoted by SNS-BH.

7. The BH procedure incorporating the shrinkage t test statistic and the null quantile-transformed test statistic null distribution, denoted by SNQ-BH.

8. The BKY procedure incorporating the shrinkage t test statistic and the null shift and scale-transformed test statistic null distribution, denoted by SNS-BKY.

9. The BKY procedure incorporating the shrinkage t test statistic and the null quantile-transformed test statistic null distribution, denoted by SNQ-BKY.

3.4.1.1. Simulating from the Normal Distribution

The normal random variables were generated as follows. First, to obtain a more realistic covariance matrix structure, $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_1$ were generated as block diagonal matrices such that

$$\mathbf{\Sigma}_{\boldsymbol{j}} = \begin{pmatrix} \sigma_1^2 \mathbf{\Sigma}_{\boldsymbol{\rho}} & 0 & \cdots & 0 \\ 0 & \sigma_2^2 \mathbf{\Sigma}_{\boldsymbol{\rho}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r^2 \mathbf{\Sigma}_{\boldsymbol{\rho}} \end{pmatrix}_{m \times m} \tag{3.40}$$

where $j \in \{0, 1\}$, $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_1$ are the covariance matrices for the controls and cases respectively, $r = m/b$, b is the number of blocks with $\mathbf{\Sigma}_\rho = \left(\rho^{|i-j|}\right)_{b \times b}$ following an auto-regessive structure with a variety of block correlation structures considered. Here, correlated variables within a block can be viewed as representing genes that are in the same pathway or that are co-regulated. Pairwise correlation between variables was ρ within a block and 0 between blocks. Both positive and negative correlations were considered, with values set to $\rho = 0,\ \pm 0.1,\ \pm 0.25,\ \pm 0.5,\ \pm 0.75,\ \pm 0.9$. The case where $\rho = 0$ corresponds to the case where the genes are independent. In order to account for heterogeneity of variance in the genes, $\sigma_1^2, \ldots, \sigma_r^2$ were simulated as independent and identically distributed random variables from a $\chi^2_{10}/10$ distribution.

Next, in order to obtain the genes that are differentially expressed between the case and control groups, a set of $m_0 = m\pi_0$ values, corresponding to the set of true null hypotheses, was randomly sampled from the set $\{1, \ldots, m\}$, which will be denoted by $\mathcal{I}_0$. The mean vector $\boldsymbol{\mu}_1 = (\mu_i : i = 1, \ldots, m)$, for the m-dimensional cases were then obtained by assigning $\mu_i = 0$ for each $i \in \mathcal{I}_0$ and simulating μ_i as independent and identically distributed random variables from a uniform distribution on the interval $[0.2, 1]$ for each $i \notin \mathcal{I}_0$. The proportion of true null hypotheses studied were $\pi_0 = 0.75,\ 0.8,\ 0.85,\ 0.9$.

Finally, for each combination of (n, ρ, π_0), the m-dimensional cases were generated independently from a normal distribution with mean vector $\boldsymbol{\mu}_1$ and covariance matrix $\boldsymbol{\Sigma}_1$ using the algorithm described above. The m-dimensional controls on the other hand were generated indepedently from a multivariate normal distribution with a zero mean vector and covariance matrix $\boldsymbol{\Sigma}_0$ as detailed in (3.40). A pre-specified significance level, $\alpha = 0.05$ was utilized.

3.4.1.2. Simulating from the Gamma Distribution

The gamma random variables, which are characterized by the shape and scale parameters, were generated in line with Cheriyan and Ramabhadran's mutivariate gamma distributions (Kotz, Balakrishnan, and Johnson (2004, see pages 454 through 456)) and are detailed as follows. Let $U_i,\ i = 1, \ldots, m$ be independent gamma random variables with shape parameters κ_i and a common scale parameter $\theta = 1$, that is, $U_i \sim \text{GAM}(\kappa_i, 1)$. Suppose that $U_0 \sim \text{GAM}(\kappa_0, 1)$ and let $X_i = U_0 + U_i,\ i = 1, \ldots, m$, then the m-variate random variables $\boldsymbol{X} = (X_1, X_2, \ldots, X_m)$ will be multivariate gamma random variables with pairwise correlation, $\text{corr}(X_i, X_j) = \frac{\kappa}{\sqrt{(\kappa+\kappa_i)(\kappa_0+\kappa_j)}}$. The case where $X_i = U_i$ corresponds to the case where the genes are independent. In order to account for reasonable correlation within the genes, κ_0 was set as 4. The differentially expressed genes between the cases and controls were generated in an analogous manner in which they were generated for the normal random variables. First for the cases, a set of $m_0 = m\pi_0$ values, corresponding to the set of true null hypotheses, was randomly sampled from the set $\mathcal{I}_0 = \{1, \ldots, m\}$. The values of κ_i were then set as $\kappa_i = 1$ for $i \in \mathcal{I}_0$ and for each $i \notin \mathcal{I}_0$ the values were obtained by simulating κ_i as independent and identically distributed uniform random variables on the interval $[1.5, 3]$. The value of κ_i was also set as $\kappa_i = 1$ for all the controls.The proportion of true null hypotheses studied were again $\pi_0 = 0.75,\ 0.8,\ 0.85,\ 0.9$.

3.4.1.3. Computation of Test Statistics

For each simulated data, one sided hypotheses tests were examined. For the BH, STS, BY and BKY procedures, the two-sample Welch t-test was employed while the shrinkage t statistic detailed in section 3.2.1 was utilized for the remaining procedures. The p-values for use in the BH, STS, BY, BKY and S-BH were computed as $\hat{p}_i = 1 - \Psi_\nu(T_n(i))$, where $\Psi_\nu(\cdot)$ denotes the cumulative distribution function of the studentized t-distribution with ν degrees of freedom. All simulations were carried out in R statistical language (R Core Team (2018)). The null shift and scale-transformed and the null quantile-transformed test statistics null distribution are available in the multtest package. The qvalue function with default settings in the qvalue package was utilized to obtain the estimate of π_0 as described in Storey, Taylor, and Siegmund (2004).

3.4.2. Simulation Results for Normal Variates

The simulation results are based on 1,000 replications per scenario and the number of bootstrap resamples is 10,000. The comparison of the methods are based on four performance criteria which include:

(i) the empirical FDR compared to the nominal level $\alpha = 0.05$

(ii) the empirical false non-discovery rate, defined as

$$\text{FNR} = \frac{\text{number of false non-discoveries}}{\text{total number of non-discoveries}} \tag{3.41}$$

(iii) the empirical power defined as the average number of false hypotheses rejected.

(iv) the stability of the procedures, characterized by the standard deviations of the number of false hypotheses rejected and the total number of rejected hypotheses.

Items number (ii) and (iii) are utilized to assess the empirical power of the procedures. In the following subsections, the results of the simulations are presented and analyzed. The results for the independent cases will be discussed first, followed by the general dependent cases.

3.4.2.1. Comparison of Procedures for Independent Tests

FDR Control Comparisons

Recall that under the assumption of independence of the test statistics or p-values, the BH procedure has an FDR equal to $\alpha\pi_0$ and that of the other procedures is less than or equal to α. Figure 3.1 provides graphical displays of the empirical FDR of the procedures considered. Numerical summaries are provided in Tables B.1 and B.2 in Appendix B. Here, the BH, BY, BKY, S-BH and STS procedures consistently provide satisfactory FDR control

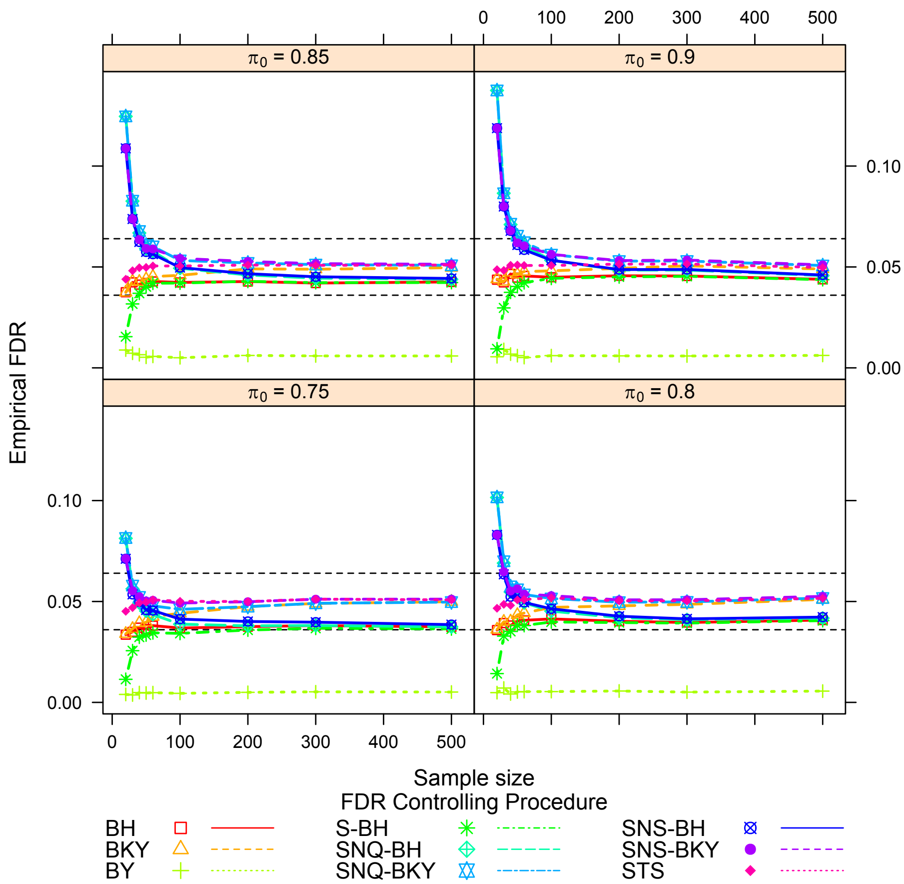

Figure 3.1. Empirical false discovery rates comparing the investigated methods under independence ($\rho = 0$) with $m = 1,000$ hypotheses for the normal variates. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is 1,000 with 10,000 bootstrap resamples. The black horizontal dashed lines depict boundaries for values within two standard deviations from the significance level. Equal sample sizes were utilized for both the cases and controls.

across all scenarios, with the STS procedure being, as expected, less conservative and the BY procedure, extremely conservative. Again, as expected, the BH, BY, S-BH, SNS-BH and SNQ-BH procedures become more conservative as the proportion of true null hypotheses π_0 decreases.The excessive conservativeness of the BY procedure is due to dividing the significance level α by $\sum_{i=1}^{m}(1/i)$ which is approximately $\ln(m)$. Conversely, according to expectation, the proposed resampling-based procedures are anti-conservative when the sample size is small but offer satisfactroy FDR control as the sample size increases. Additionally, FDR control for these resampling-based procedures varies with the proportion of true null hypotheses. For instance, when $\pi_0 = 0.9$, a minimum total sample of size, $n = 60$ was needed to achieve FDR control. However, a minimum total sample of size, $n = 30$ was needed to ensure control of the FDR when $\pi_0 = 0.75$. We also experimented with lower proportions of true null hypotheses (results not shown) and observed that as the proportion decreased to 65%, a total sample size of 20 was enough to ensure asymptotic FDR control by the proposed resampling-based methods.

Empirical Power Comparisons

Table 3.1 reports the numerical summaries of the empirical FNRs and the average number of false null hypotheses rejected, denoted by "Rejected" in the table, of the investigated FDR-controlling procedures for $n = 60, 100$, and 300. The numerical summaries for the remaining sample sizes are reported in Tables B.3 through B.5 in Appendix B. In cases where the resampling-based procedures were conservative, the power of the SNS-BH, SNQ-BH, SNS-BKY and SNQ-BKY procedures are higher in almost all instances than the power of their corresponding original procedures, with a notable gain in power when $n \leq 100$ (see Tables B.4 and B.5). Additionally, in these cases where the resampling-based procedures were conservative, these procedures had higher power (see also Tables B.3 through B.5) than the STS procedure for $n \leq 200$ when the proportion of true null hypotheses was 80% or greater and equivalent power for all other situations. Unsurprisingly, the adaptive procedures (STS, BKY, SNS-BKY and SNQ-BKY) become evidently more powerful than the other investigated procedures with an increasing proportion of true null hypotheses, π_0. As expected, in these cases, the improvement in power is due to these procedures selecting less conservative threshold values.

Table 3.1. Empirical false-non discovery rates and average number of false hypotheses rejected (in the columns "Rejected") for the investigated methods considered for the independent tests for the normal variates. Results correspond to the following simulation parameters: $n = 60, 100, 300$; $m = 1,000$; $\pi_0 = 0.75, 0.8, 0.85, 0.9$; $\alpha = 0.05$. The number of replications is 1,000 per scenario and the number of bootstrap resamples is 10,000. Equal sample sizes were utilized for both the cases and controls.

	$n = 60$							
	$\pi_0 = 0.75$		$\pi_0 = 0.8$		$\pi_0 = 0.85$		$\pi_0 = 0.9$	
	FNR	Rejected	FNR	Rejected	FNR	Rejected	FNR	Rejected
BH	0.163	104.658	0.128	82.596	0.105	51.011	0.065	37.287
STS	0.154	114.310	0.123	88.211	0.102	53.887	0.064	38.786
BY	0.208	53.356	0.163	43.796	0.129	24.477	0.084	17.294
BKY	0.160	108.366	0.126	84.787	0.104	51.874	0.065	37.815
S-BH	0.165	102.018	0.129	82.063	0.105	50.483	0.064	38.700
SNS-BH	0.157	111.754	0.123	88.116	0.100	55.606	0.061	41.910
SNQ-BH	0.158	109.751	0.123	88.389	0.100	56.074	0.060	43.045
SNS-BKY	0.153	115.533	0.121	90.255	0.099	56.440	0.060	42.317
SNQ-BKY	0.155	113.301	0.121	90.525	0.099	56.909	0.059	43.383
	$n = 100$							
BH	0.115	153.187	0.090	121.731	0.075	81.548	0.042	60.868
STS	0.107	161.032	0.085	126.343	0.072	84.244	0.041	62.056
BY	0.160	107.358	0.124	86.655	0.100	56.074	0.060	42.680
BKY	0.110	157.756	0.087	124.387	0.073	82.899	0.041	61.554
S-BH	0.116	152.385	0.089	121.862	0.074	82.028	0.041	61.805
SNS-BH	0.111	156.835	0.087	124.571	0.072	84.234	0.040	63.045
SNQ-BH	0.113	155.323	0.087	124.537	0.072	84.496	0.039	63.520
SNS-BKY	0.107	161.288	0.084	127.228	0.071	85.544	0.039	63.618
SNQ-BKY	0.108	159.734	0.084	127.113	0.071	85.743	0.039	64.086
	$n = 300$							
BH	0.037	221.128	0.028	176.836	0.024	129.007	0.011	90.216
STS	0.034	224.398	0.026	178.966	0.023	130.350	0.010	90.674
BY	0.064	198.401	0.050	158.023	0.042	112.993	0.019	82.587
BKY	0.034	224.025	0.026	178.737	0.023	130.142	0.010	90.605
S-BH	0.037	221.243	0.028	177.116	0.024	129.243	0.011	90.310
SNS-BH	0.037	221.750	0.028	177.477	0.024	129.569	0.010	90.511
SNQ-BH	0.037	221.564	0.028	177.450	0.024	129.620	0.010	90.564
SNS-BKY	0.033	224.604	0.026	179.271	0.022	130.656	0.010	90.883
SNQ-BKY	0.033	224.463	0.026	179.243	0.022	130.704	0.010	90.881

Stability of FDR-Controlling Procedures

In this subsection, we analyze the stability, characterized by the standard deviation of the number of false hypotheses rejected and the total number of rejected hypotheses, of the investigated procedures. These quantities are illustrated in Figure 3.2 and are given in parentheses in Tables B.4 and B.5. As expected, the results indicate that the non-adaptive procedures (BH, BY, S-BH, SNS-BH and SNQ-BH) have the greatest stability, followed by the BKY, SNS-BKY and SNQ-BKY procedures. The STS procedure is the least stable among all investigated procedures. The instability of the STS procedure is greater as the proportions of true null hypotheses decreases.

Generally, all investigated procedures become less conservative, more powerful, and more stable as the sample size increases. In terms of power and stability, however, the SNS-BKY and SNQ-BKY procedures are better alternatives to the original BH procedure than the STS procedure when the total sample size is 60 or greater since in such cases, the SNS-BKY and SNQ-BKY procedures are always conservative and have improved or equivalent power but better stability than the STS procedure.

3.4.2.2. Comparison of Procedures for Dependent Tests

The simulation study allows us to examine the effect of correlation between the test statistics on false discovery rate achieved by the procedures. In this subsection, we will concentrate on the cases where $\rho = 0.25, 0.5$, and 0.9. These values can be viewed as settings where there is either weak, moderate, or strong correlation among the genes. The results for the negative correlations are similar and are provided in appendix C.

FDR Control Comparisons

The empirical FDRs for the investigated procedures for $\rho = 0.25, 0.5$, and 0.9 are illustrated in

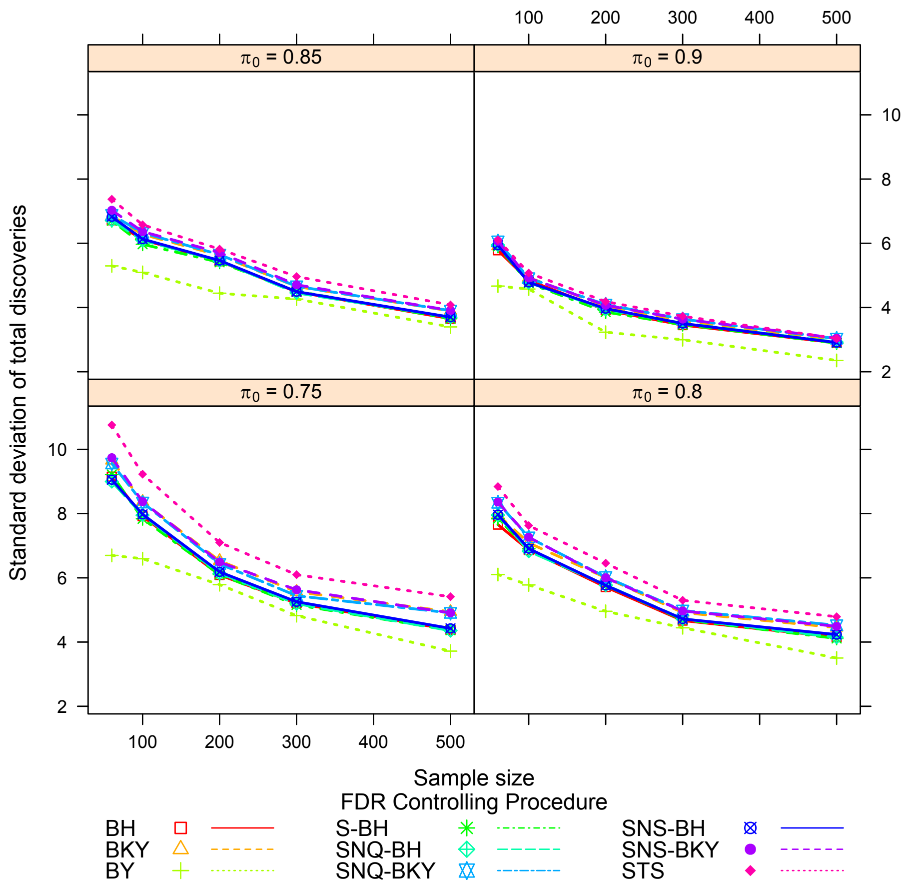

Figure 3.2. Estimated standard deviation of the total number of hypotheses rejected for the investigated methods under independence ($\rho = 0$) with $m = 1,000$ hypotheses for the normal variates. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the cases and controls. The black horizontal dashed lines depict boundaries for values within two standard deviations from the significance level.

Figure 3.3. The numerical summaries of the empirical FDRs for all other combinations of the simulation parameters are provided in Tables C.1 through C.20. The pattern of FDR control for these correlated cases are similar to the independent cases, although the resampling-based methods are somewhat less conservative when compared to the BH and BKY step-up procedures. As expected, the resampling-based methods are anti-conservative for smaller sample sizes with asymptotic control being dependent on the proprotion of true null hypotheses.

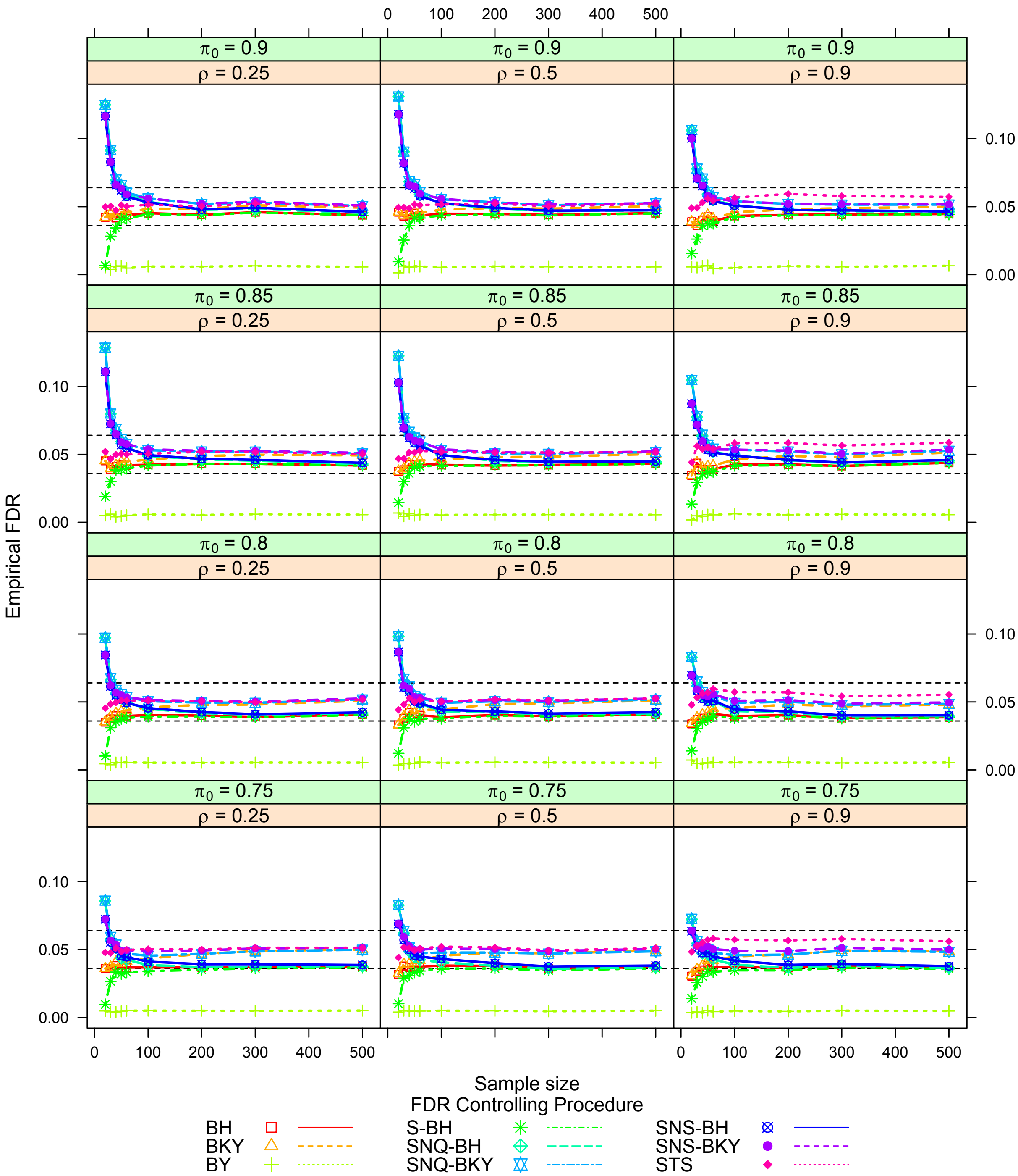

Figure 3.3. Empirical false discovery rates for the investigated methods in the presence of moderate to high correlation among the variables for the normal variates. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The number of hypotheses is $m = 1,000$ with a pre-specified significance level of $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases.

The BH, BY, S-BH, SNS-BH and SNQ-BH procedures tend to be more conservative with increasing correlation. The STS and BKY procedures consistently offer satisfactory FDR control across all investigated sample sizes while the SNS-BKY and SNQ-BKY procedures offer satisfactory FDR control for moderate to large sample sizes (i.e., $n \geq 60$), with an even smaller size needed for FDR control with a decreasing proportion of true null hypotheses. For $\rho = 0.9$, the SNS-BKY and SNQ-BKY procedures remain closer to the nominal level $\alpha = 0.05$ than the other investigated procedures.

Empirical Power Comparisons

The numerical summaries of the empirical FNRs and the average number of false hypotheses rejected for all investigated π_0, $\rho = 0.25, 0.5, 0.9$, and $n = 60$ are reported in Table 3.2. The empirical FNRs for all other configuration of simulation parameters are reported in Tables C.21 through C.30, and those for the average number of false hypotheses rejected in Tables C.31 through C.50. Here, we observe that the power of the resampling-based procedures is higher than the original BH and BKY procedures in all cases where these procedures were conservative but is equivalent to the STS procedure in such cases. Specifically, the adaptive resampling-based procedures (SNS-BKY and SNQ-BKY) are more powerful than the STS procedure for $n \leq 100$ with the SNQ-BKY being the most powerful in almost all settings.

Stability of FDR-Controlling Procedures

Finally, Figure 3.4 illustrates the estimated standard deviation of the total number of hypotheses rejected for the investigated procedures for $\rho = 0.25, 0.5$, and 0.9, with the estimated standard deviation of the number of false hypotheses rejected for all other combination of simulation parameters reported in parentheses in Tables C.31 through C.50. The stability trend observed here is similar to the observed trend for the independent case. The level

Table 3.2. Empirical false non-discovery rates and average number of false hypotheses rejected (in the columns "Rejected") for the investigated methods considered for the correlated cases with normal variates. Results correspond to the following simulation parameters: $n = 30$; $m = 1,000$; $\pi_0 = 0.75, 0.8, 0.85, 0.9$; $\alpha = 0.05$ and $\rho = 0.25, 0.5, 0.9$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The number of replications is $1,000$ per scenario and the number of bootstrap resamples is 10,000. Equal sample sizes were utilized for both the cases and controls.

	$n = 60$; $\rho = 0.25$							
	$\pi_0 = 0.75$		$\pi_0 = 0.8$		$\pi_0 = 0.85$		$\pi_0 = 0.9$	
	FNR	Rejected	FNR	Rejected	FNR	Rejected	FNR	Rejected
BH	0.163	104.309	0.129	82.229	0.104	51.121	0.066	36.970
STS	0.155	113.831	0.123	88.151	0.102	54.023	0.064	38.552
BY	0.208	53.234	0.163	43.717	0.128	24.745	0.085	16.876
BKY	0.160	107.933	0.127	84.497	0.104	52.054	0.065	37.450
S-BH	0.166	101.872	0.129	81.742	0.105	50.892	0.064	38.451
SNS-BH	0.157	111.388	0.124	87.712	0.100	55.750	0.061	41.634
SNQ-BH	0.159	109.322	0.123	88.001	0.099	56.496	0.060	42.831
SNS-BKY	0.153	115.186	0.122	89.933	0.099	56.663	0.061	42.084
SNQ-BKY	0.155	112.886	0.121	90.042	0.099	57.406	0.059	43.249
	$n = 60$; $\rho = 0.5$							
BH	0.163	104.516	0.129	82.267	0.105	50.947	0.065	37.063
STS	0.154	114.598	0.123	88.132	0.102	54.098	0.064	38.801
BY	0.208	53.466	0.163	43.702	0.128	24.773	0.084	17.008
BKY	0.160	108.181	0.127	84.456	0.104	51.866	0.065	37.538
S-BH	0.165	101.933	0.129	81.734	0.105	50.598	0.064	38.445
SNS-BH	0.157	111.609	0.124	87.611	0.100	55.573	0.061	41.695
SNQ-BH	0.159	109.418	0.123	87.942	0.100	56.208	0.060	42.740
SNS-BKY	0.153	115.271	0.122	89.709	0.099	56.467	0.061	42.105
SNQ-BKY	0.155	113.025	0.121	90.029	0.099	57.028	0.060	43.111
	$n = 60$; $\rho = 0.9$							
BH	0.163	104.623	0.128	82.826	0.104	51.131	0.066	36.906
STS	0.151	117.335	0.121	90.946	0.100	56.136	0.063	39.704
BY	0.208	53.533	0.163	44.202	0.128	24.645	0.084	17.239
BKY	0.159	108.411	0.126	84.948	0.104	52.034	0.065	37.358
S-BH	0.165	102.066	0.129	82.174	0.105	50.765	0.064	38.235
SNS-BH	0.156	111.717	0.123	88.208	0.100	55.752	0.061	41.612
SNQ-BH	0.158	109.642	0.123	88.490	0.100	56.187	0.060	42.674
SNS-BKY	0.153	115.545	0.121	90.605	0.099	56.601	0.061	42.023
SNQ-BKY	0.155	113.268	0.121	90.670	0.099	57.041	0.060	43.062

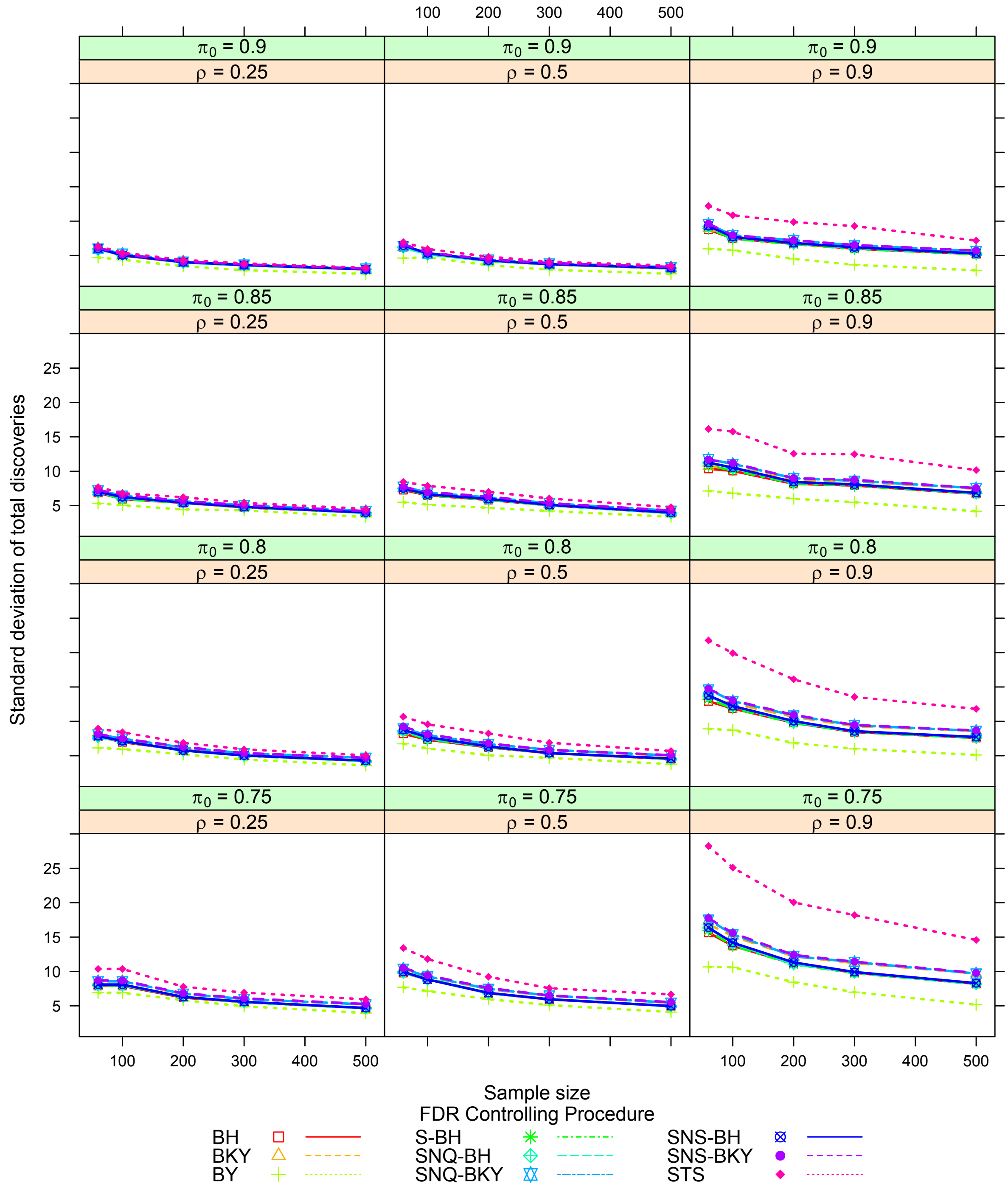

Figure 3.4. Estimated standard deviation of the total number of hypotheses rejected for the investigated methods under dependence with $\rho = 0.25, 0.5$, and 0.9 and $m = 1,000$ hypotheses for the normal variates. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases.

of stability varies with the proportion of true null hypotheses and the extent of correlation among the variables. For instance, the procedures have equivalent stability for $\rho = 0.25$ and $\pi_0 = 0.9$. However, irrespective of the proportion of false null hypotheses, substantial difference in stability is observed for the STS procedure and the other procedures for $\rho = 0.9$. In general, all investigated procedures become less stable with a decreasing proportion of true null hypotheses and with increasing pairwise correlations, with the non-adaptive procedures, especially the BY procedure, having greater stability, and the STS procedure being the least stable.

In practice, especially in microarray experiments, tests are usually correlated. In such cases, a multiple testing procedure with good FDR control, higher power and good stability is desirable. The simulation results indicate that for a total sample size of $n \geq 60$, the SNS-BKY and SNQ-BKY procedures exhibit power and stability properties intermediate between the two most commonly employed procedures, BH and STS. Particularly, the SNS-BKY and SNQ-BKY procedures have better stability and higher or equivalent power to the STS procedure and improved FDR control and higher power than the BH procedure.

3.4.3. Simulation Results for Gamma Variates

The simulation results for both the dependent and independent gamma random variables are provided in the following section. The empirical FDRs for the investigated methods are summarized in Figure 3.5. The pattern of FDR control for the resampling-based procedures, SNS-BH and SNS-BKY, for the independent cases were similar to what was observed for the normal random variables. For $\pi_0 = 0.9$ a minimum total sample size of 100 was needed to achieve FDR control while a total sample of 30 was needed when $\pi_0 = 0.75$. Interestingly, these two procedures were very conservative, with the empirical FDR equal to zero in almost all parameter configurations for the dependent cases. The SNQ-BH and

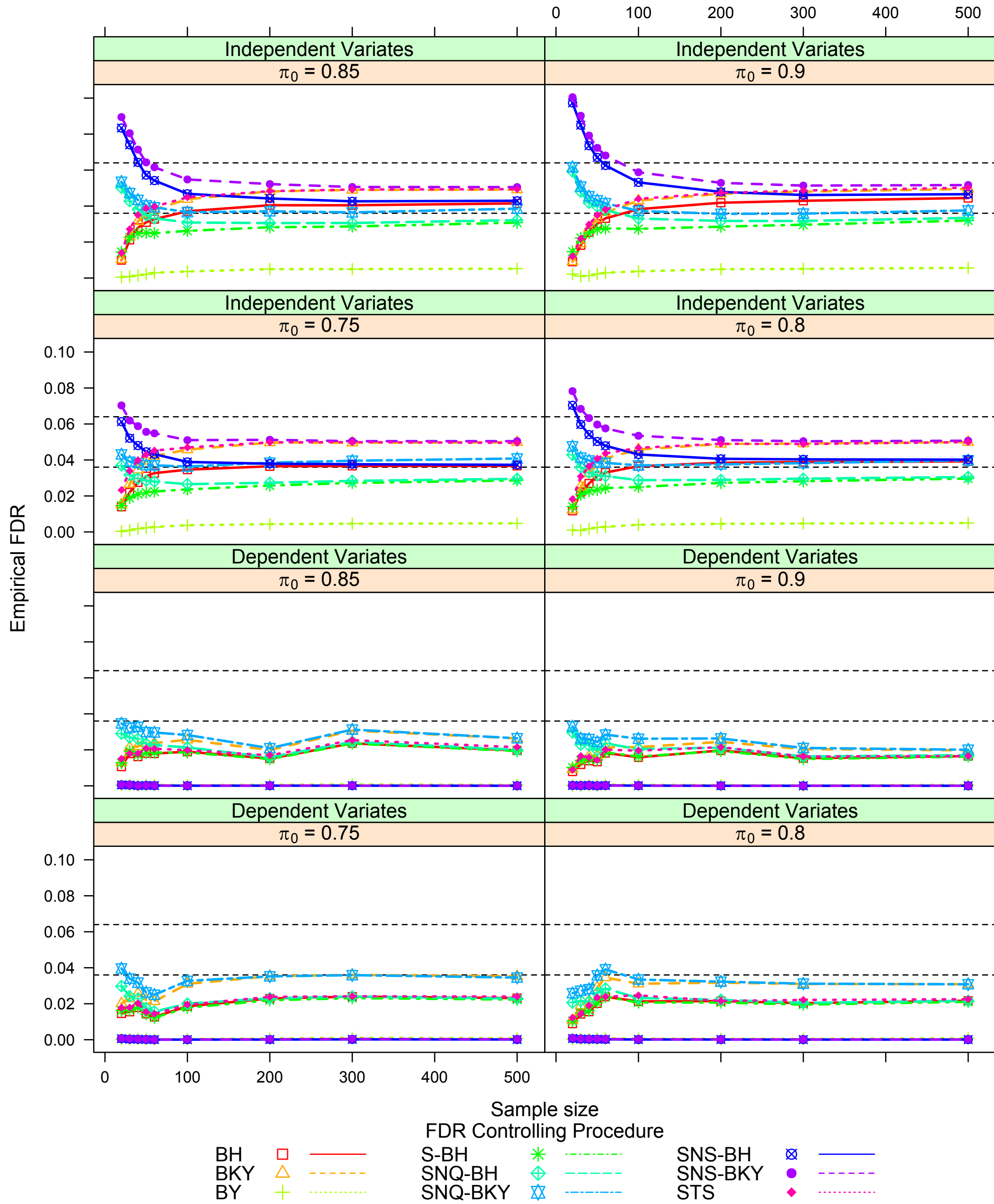

Figure 3.5. Empirical false discovery rates for the investigated methods for the gamma variates with $m = 1,000$ hypotheses. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The black horizontal dashed lines depict boundaries for values within two standard deviations from the significance level.

SNQ-BKY resampling-based procedures on the other hand, provide FDR control for all investigated parameter configurations in both the dependent and independent cases, and were less conservative compared to the SNS-BH and SNS-BKY procedures in such cases. In general, all the investigated methods were very conservative for the dependent cases with the SNQ-BKY procedure, as expected, being the least conservative in all such cases.

Numerical summaries of the empirical FNRs and the average number of false hypotheses rejected for both the dependent and independent gamma random variables for the cases where $n = 20, 60$, and 100 are reported in Tables 3.3 and 3.4 with all other results reported in Tables C.53 through C.56 in appendix C. We will exclude the power comparisons of the SNS-BH and SNS-BKY procedures with the other procedures for the cases where $n \leq 60$ (indicated by a star (*) in the tables) for the independent tests since these procedures were anti-conservative in such cases. Here, we observe that the SNQ-BKY resampling-based procedure was consistently more powerful than all the investigated methods with a significant gain in power in small to moderate sample sizes (i.e., $n \leq 100$). We re-emphasize that the SNQ-BKY resampling-based procedure was conservative for all parameter configurations for the gamma variates.

Finally, Figure 3.6 illustrates the stability of the investigated procedures, as measured by the standard deviation of the total number of hypotheses rejected across the $1,000$ simulation replications. All the procedures were quite stable when the random variables were independent, with stability trends similar to the normal random variable cases. In such cases, all the investigated procedures become more stable with an increasing proportion of true null hypotheses, with the STS procedure being the least stable. Conversely, interesting stability results were obtained for the dependent variables. All the investigated procedures were somewhat unstable with the estimated standard deviation of the number of rejected hypotheses as high as 165 for the SNQ-BKY when $\pi_0 = 0.75$ and $n = 20$. In this setting, the SNQ-BKY procedure loses its superior stability, while other proposed procedures maintain

Table 3.3. Empirical false non-discovery rates and average number of false hypotheses rejected (in the columns "Rejected") for the investigated methods considered for the independent gamma variates. Results correspond to the following simulation parameters: $n = 20, 60, 100$; $m = 1,000$; $\pi_0 = 0.75, 0.8, 0.85, 0.9$; $\alpha = 0.05$. The number of replications is $1,000$ per scenario and the number of bootstrap resamples is 10,000. Equal sample sizes were utilized for both the cases and controls. Cases where FDR control were anti-conservative are indicated with a star($\star$).

	$n = 20$							
	$\pi_0 = 0.75$		$\pi_0 = 0.8$		$\pi_0 = 0.85$		$\pi_0 = 0.9$	
	FNR	Rejected	FNR	Rejected	FNR	Rejected	FNR	Rejected
BH	0.214	45.377	0.180	24.457	0.140	11.203	0.095	6.025
STS	0.197	65.865	0.171	34.724	0.137	15.327	0.093	7.536
BY	0.249	1.658	0.199	0.961	0.149	0.592	0.100	0.418
BKY	0.212	47.919	0.179	25.227	0.140	11.465	0.094	6.088
S-BH	0.163	104.645	0.136	74.056	0.105	50.032	0.068	34.021
SNS-BH	0.134	134.759	0.111*	100.973*	0.085*	71.990*	0.053*	49.516*
SNQ-BH	0.137	131.385	0.113	98.435	0.086	70.761	0.053	49.682
SNS-BKY	0.129*	139.992*	0.108*	104.079*	0.083*	73.583*	0.053*	50.206*
SNQ-BKY	0.133	135.756	0.111	100.995	0.084	71.978	0.053	50.172
	$n = 60$							
BH	0.062	200.960	0.052	156.368	0.041	114.187	0.024	77.794
STS	0.055	207.119	0.047	160.824	0.038	116.496	0.023	78.758
BY	0.114	153.486	0.096	115.508	0.073	83.061	0.044	58.578
BKY	0.056	205.935	0.048	159.770	0.039	116.071	0.023	78.503
S-BH	0.050	210.779	0.040	166.729	0.030	123.420	0.017	84.414
SNS-BH	0.044	215.776	0.035	171.305	0.026	127.176	0.015	86.688
SNQ-BH	0.047	213.488	0.037	169.341	0.028	125.655	0.016	85.862
SNS-BKY	0.040	219.403	0.032	173.649	0.025	128.360	0.014*	87.155*
SNQ-BKY	0.042	217.265	0.034	171.835	0.027	126.936	0.015	86.309
	$n = 100$							
BH	0.026	230.499	0.020	183.775	0.016	135.916	0.010	91.319
STS	0.022	233.578	0.018	185.656	0.015	137.070	0.009	91.784
BY	0.057	204.963	0.046	161.712	0.035	119.077	0.020	82.047
BKY	0.022	233.385	0.018	185.586	0.015	137.003	0.009	91.733
S-BH	0.020	234.530	0.015	187.781	0.012	139.598	0.007	93.672
SNS-BH	0.018	236.264	0.013	189.363	0.011	140.899	0.006	94.513
SNQ-BH	0.019	235.289	0.014	188.504	0.011	140.259	0.007	94.067
SNS-BKY	0.016	238.197	0.012	190.583	0.010	141.557	0.006	94.770
SNQ-BKY	0.017	237.452	0.013	189.887	0.011	140.939	0.006	94.341

Table 3.4. Empirical false non-discovery rates and average number of false hypotheses rejected (in the columns "Rejected") for the investigated methods considered for the dependent gamma variates. Results correspond to the following simulation parameters: $n = 20, 60, 100$; $m = 1,000$; $\pi_0 = 0.75, 0.8, 0.85, 0.9$; $\alpha = 0.05$. The number of replications is $1,000$ per scenario and the number of bootstrap resamples is 10,000. Equal sample sizes were utilized for both the cases and controls.

	$n = 20$							
	$\pi_0 = 0.75$		$\pi_0 = 0.8$		$\pi_0 = 0.85$		$\pi_0 = 0.9$	
	FNR	Rejected	FNR	Rejected	FNR	Rejected	FNR	Rejected
BH	0.242	9.662	0.197	3.715	0.147	2.850	0.099	1.329
STS	0.240	11.209	0.196	4.829	0.147	3.648	0.099	1.457
BY	0.250	0.380	0.200	0.171	0.150	0.063	0.100	0.035
BKY	0.240	10.922	0.196	4.090	0.147	3.127	0.099	1.435
S-BH	0.236	15.994	0.194	7.138	0.144	6.013	0.097	3.415
SNS-BH	0.247	3.650	0.199	1.546	0.149	1.326	0.099	0.789
SNQ-BH	0.224	29.728	0.185	16.992	0.138	13.547	0.091	8.989
SNS-BKY	0.247	3.824	0.199	1.598	0.149	1.350	0.099	0.793
SNQ-BKY	0.222	31.565	0.184	17.798	0.137	14.147	0.091	9.211
	$n = 60$							
BH	0.198	58.787	0.158	46.360	0.122	29.864	0.082	19.004
STS	0.192	65.475	0.154	50.020	0.119	32.794	0.081	20.285
BY	0.234	19.314	0.186	15.383	0.142	9.191	0.094	6.196
BKY	0.195	61.951	0.156	48.148	0.121	30.766	0.081	19.359
S-BH	0.186	72.402	0.147	58.092	0.114	38.880	0.076	25.402
SNS-BH	0.231	22.696	0.184	18.831	0.139	11.652	0.092	8.177
SNQ-BH	0.178	80.524	0.140	65.283	0.109	44.299	0.071	29.773
SNS-BKY	0.230	23.482	0.183	19.266	0.139	11.816	0.092	8.237
SNQ-BKY	0.175	83.563	0.138	67.065	0.108	45.158	0.071	30.072
	$n = 100$							
BH	0.141	120.445	0.121	85.924	0.095	58.363	0.060	41.503
STS	0.133	128.552	0.114	92.775	0.091	62.610	0.057	44.990
BY	0.193	65.834	0.159	45.627	0.123	29.672	0.080	21.095
BKY	0.136	124.643	0.118	88.322	0.094	59.542	0.060	42.030
S-BH	0.130	132.361	0.111	95.995	0.087	66.880	0.054	48.049
SNS-BH	0.195	63.795	0.160	44.791	0.123	29.442	0.080	21.553
SNQ-BH	0.125	136.743	0.108	99.611	0.084	70.121	0.051	50.538
SNS-BKY	0.194	65.582	0.159	45.735	0.123	29.857	0.079	21.676
SNQ-BKY	0.121	140.594	0.105	101.838	0.083	71.186	0.051	50.946

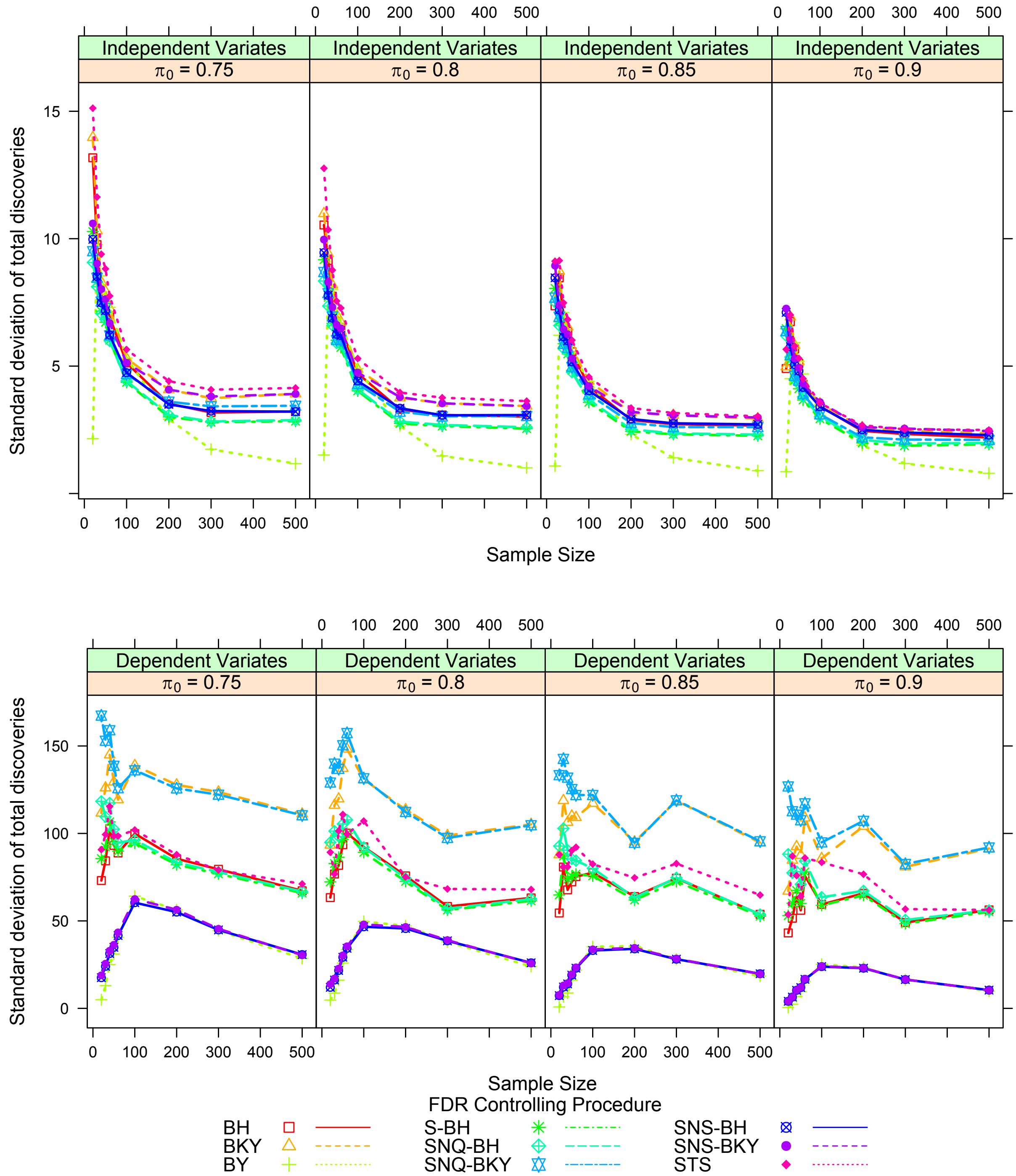

Figure 3.6. Estimated standard deviation of the total number of hypotheses rejected for the investigated methods for the gamma variates under both dependence and independence with $m = 1,000$ hypotheses. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the cases and controls.

their superior stability over the STS procedure. Recall that the investigated procedures were all conservative in this setting though. Similarly to the normally distributed variables, all the investigated procedures generally become less conservative, more powerful, and more stable with increasing sample sizes, with the SNQ-BKY being the least conservative and the most powerful for sample sizes less than or equal to 100.

To alleviate the unusual instability results obtained for the dependent cases, in practice, we recommend transforming the variables in applications where the variables are suspected to be dependent, but not approximately normally distributed before applying any of the investigated procedures to improve the consistency of the obtained results.

3.5. Discussion and Conclusions

3.5.1. Discussion

In this chapter, we proposed resampling-based procedures for multiple hypotheses testing by incorporating a generally valid null distribution and a James-Stein-type analytic shrinkage estimation of the variance components into the original Benjamini and Hochberg (1995) and Benjamini, Krieger, and Yekutieli (2006) procedures. Under the assumption of null domination and positive regression dependence, the resampling-based procedures have been shown to asymptotically control the FDR. We compared the proposed procedures with the linear step-up procedures of Benjamini and Hochberg (1995), Benjamini and Yekutieli (2001), Benjamini, Krieger, and Yekutieli (2006) and the q-value procedure of Storey (2003); Storey, Taylor, and Siegmund (2004) using extensive Monte Carlo simulations. Four different performance criteria were utilized: (i) the empirical false discovery rate, (ii) the empirical false non-discovery rate, (iii) the empirical power defined as the average number of false hypotheses rejected, and (iv) the stability of the procedures, characterized by the standard deviations of the number of false hypotheses rejected and the total number of rejected hypotheses.

We note that it is impossible to carry out a comprehensive simulation study capturing all possible behaviors of the hypotheses, but in our simulations, various different realistic scenarios that might be encountered in practice were investigated. This included varying the sample size, the proportion of non-null hypotheses, the distribution of the random variables and their dependency structure. The simulation study focused on a case-control experiment, but the proposed methodology can be extended to any hypothesis testing problem. For a variety of testing scenarios, the proposed resampling-based procedures were shown to provide satisfactory FDR control when there are at least 30 observations in each of the case and control groups (with a total sample size of 60). An even smaller sample size is needed to provide satisfactory FDR control when the proportion of true null hypotheses decreases to 65%. Specifically, when the variables are normally distributed, the resampling-based procedures consistently offer satisfactory FDR control for a total sample of size $n \geq 60$ when 85% or more of the hypotheses are truly null hypotheses. For the gamma random variables, the proposed procedures based on the null quantile-transformed null distribution (SNQ-BH and SNQ-BKY) consistently controlled the FDR at the pre-specified significance level for all parameter configurations considered. On the contrary, as in the normal case, asymptotic FDR control for the proposed procedures based on the null shift and scale-transformed null distribution (SNS-BH and SNS-BKY) for the gamma random variables is dependent on the proportion of true null hypotheses and the dependence structure of the test statistics. In particular, the procedures usually provided FDR control for total sample sizes $n \geq 60$ for the independent cases, and for all sample sizes for the dependent cases although they become extremely conservative in the dependent cases.

It is more difficult to characterize the patterns of the procedures based on the null shift and scale-transformed null distribution for the gamma random variables. Perhaps the arguments provided by van der Laan and Hubbard (2006) can provide a little insight into these results. Primarily, the marginal distribution of a test statistic is known when the null hypothesis is true. Given that, the null shift and scale-transformed null distribution ensures that the obtained marginal distribution and the known marginal distribution have equivalent mean and variance, but does not guarantee that the marginal distributions are equal. This suggests that using this null distribution does not necessarily produce optimal marginal null distributions. Additionally, van der Laan and Hubbard (2006) argued that since the marginal null distributions cannot be controlled, the null shift and scale-transformed null distributions can sometimes be problematic in finite samples although they are always asymptotically valid. In such cases, one might require a larger sample size and a larger number of bootstrap replicates to alleviate the limitations of the null shift and scale-transformed null distributions. This may explain the poor performance of the respective procedures for the dependent gamma random variables.

As earlier discussed, in practice, especially in microarray experiments, tests are often correlated. A multiple testing procedure with good FDR control, higher power and good stability is also desirable in such cases. Overall, the simulation study indicates improved FDR control and a gain in power over the original linear step-up procedures of Benjamini and Hochberg (1995) and Benjamini, Krieger, and Yekutieli (2006), even for independent tests, by incorporating a generally valid null distribution and a James-Stein-type analytic shrinkage estimation of the variance components. Specifically, the procedures based on the null shift and scale-transformed null distribution and the null quantile-transformed null distribution have better stability and higher or equivalent power to the STS, procedure and improved FDR control and higher power than the BH and the two-stage adaptive BKY procedures when the random variables are independent and $n \geq 60$. Additionally, as expected, a substantial gain in power was observed for the resampling-based procedures for the cases where the random variables were not normally distributed, with the SNQ-BKY procedure consistently outperforming all the investigated procedures in such settings. Mainly, all investigated procedures become less conservative, more powerful, and more stable as the sample size increases. Furthermore, the procedures become less stable with decreasing proportions of true null hypotheses and increasing pairwise correlations, with the non-adaptive procedures, especially the BY procedure, having greater stability, and the STS procedure being the least stable.

3.5.2. Conclusions

High-throughput gene expression experiments such as microarray experiments involve statistically testing thousands of hypotheses simultaneously to identify genes that are differentially expressed. An unguarded use of single-inference procedures for such analyses inflates the overall type I error rates. Correlation between genes and across arrays further complicates this problem. Multiple testing procedures provide efficient methods for examining each hypothesis while also controlling an overall error rate at a pre-specified level. The validity and accuracy of any such testing procedure is essentially determined by whether the chosen test statistic is optimal, the null distributions are correctly or conservatively specified, and whether the data are independent across tests. As emphasized earlier, misspecifying the null distribution may undercut inferential validity. This study proposes a new multiple testing procedure by incorporating a generally valid null distribution and a James-Stein-type analytic shrinkage estimation of the variance components into the linear step-up procedure of Benjamini and Hochberg (1995) and the two-stage adaptive step-up procedure of Benjamini, Krieger, and Yekutieli (2006). Extensive Monte Carlo simulations show that the resampling-based procedure based on the null quantile-transformed null distribution is essentially more stable and as powerful or substantially more powerful than some procedures proposed in finite sample inferential problems, provided there are at least 30 observations in both the case and control groups.

In recent years, reproducibility of statistical findings has drawn considerable attention not only from statisticians, but from all researchers engaged in empirical discovery. As noted by Stodden (2015), the reasons for irreproducibility are at least, to some extent, due to the number of false discoveries in such studies. Thus, a multiple testing procedure with improved FDR control, good power and higher stability, can help alleviate the inconsistencies in statistical findings, especially in biomedical research. Our proposed procedure based on the null quantile-transformed null distribution, SNQ-BKY, is attractive in such analyses when there are at least 30 observations in each of the case and control groups in that it has higher power than the linear step-up procedure of Benjamini and Hochberg (1995) and better FDR control, higher stability and better or equivalent power than the q-value procedure of Storey (2003); Storey, Taylor, and Siegmund (2004). The trade-off for gains in power and stability is the extra computational cost. However, with modern computing power, this issue is far less important than in years past.

3.5.3. Available Software

The null shift and scale-transformed null distribution and the null quantile-transformed null distribution are implemented in the multest package in R statistical software as part of the Bioconductor project (Pollard, Dudoit, and van der Laan (2005)). The James-Stein-type analytic shrinkage estimation of the variance components is in the corpcor package in R statistical software (Schäfer, Opgen-Rhein, Zuber, Ahdesmaki, Silva, and Strimmer. (2017)).

All simulations were carried out in R (R Core Team (2018)), using the following packages: MASS (Version 7.3-48), corpcor (Version 1.6.9), multtest (Version 2.36.0) and qvalue (Version 2.12.0)

CHAPTER IV

MODIFIED STEP-DOWN PROCEDURE THAT CONTROLS THE FALSE DISCOVERY RATE UNDER DEPENDENCE

4.1. Introduction

The linear step-up procedure of Benjamini and Hochberg (1995) (BH) and the q-value procedure of Storey, Taylor, and Siegmund (2004) (STS) are among the most commonly employed FDR controlling multiple testing procedures in practice. The STS procedure has been shown to have higher power but lower stability compared to the BH procedure (Qiu, Klebanov, and Yakovlev (2005); Li, Xie, Zand, Fogg, and Dye (2017)). In the previous chapter, we proposed resampling-based FDR controlling procedures, and showed through extensive Monte Carlo simulations that these resampling-based procedures have superior stability, measured as the standard deviations of both the number of true discoveries and the total number of discoveries, to the STS procedure, while still maintaining equivalent power to the STS procedure and higher power than the BH procedure. Although the proposed resampling-based procedures have a gain in power, these methods still rely on the marginal distribution of the test statistics and do not fully utilize the joint dependence structure of the test statistics. Additionally, the proposed methods rely on some special forms of dependency, the positive regression dependent on the subset of the test statistics corresponding to the true null hypotheses, and do not accomodate general dependence among the test statistics. Thus, one can expect methods that fully incorporate the dependence structure of the test statistics without making any assumptions about the nature of the dependence to provide an improvement in power. To this end, in this chapter we seek to develop step-down procedures for control of FDR that incorporate information about the dependence structure of the test statistics, and by so doing, improve the chances of identifying false null hypotheses.

The remainder of the chapter is set up as follows. We provide our setup and notation in section 4.2. The step-down multiple testing procedure is detailed in section 4.3. This method is dependent on the distribution of the test statistics through the data-generating distribution. However, in practice since the data-generating distribution is unknown, so is the test statistic distribution. We therefore also provide a bootstrap-based step-down procedure in section 4.3. Section 4.4 provides some theoretical results for the proposed method. The asymptotic validity of the proposed method relies on the concept of null domination. Ongoing efforts and future extensions of the work are provided in section 4.5. Finally, a summary and some conclusions are provided in section 4.6.

4.2. Setup and Notation

As in previous chapters, we will again denote an m-dimensional vector of statistics, say $\boldsymbol{\theta}_n$, by $\boldsymbol{\theta}_n = \big(\boldsymbol{\theta}_n(1),\ \boldsymbol{\theta}_n(2),\ \ldots,\ \boldsymbol{\theta}_n(m)\big)$. Consider the random sample $\mathcal{X}_n = (X_1, \ldots, X_n)$ of

n independent and identically distributed (i.i.d) random variables from a data-generating distribution $P \in \Omega$. Here, Ω, may be a parametric, semiparametric or nonparametric statistical model. Define a general hypothesis as a submodel $\omega \subseteq \Omega$. Consider the problem of testing simultaneously m hypotheses on the basis of the sample. The null hypotheses are defined as $H_0(i) = I(P \in \omega_i)$ and the corresponding alternative hypotheses as $H_1(i) = I(P \notin \omega_i)$, $i = 1, \ldots, m$. Note, $I(\cdot)$ is the indicator function, having the value of 1 when the condition in the parentheses is satisfied and 0 otherwise. Let $\mathcal{H}_0 = \mathcal{H}_0(P) = \{i : P \in \omega_i\}$ be the set of true null hypotheses, and $\mathcal{H}_1 = \mathcal{H}_1(P) = \{i : P \notin \omega_i\}$, the set of false null hypotheses. Then, $m_0 =| \mathcal{H}_0 |$ is the number of true null hypotheses, and $m_1 = m - m_0 =| \mathcal{H}_1 |$ is the number of false null hypotheses.

The aim of any multiple testing procedure is to estimate the sets $\mathcal{H}_0$ and $\mathcal{H}_1$ while controlling a measure of overall error, such as FWER or FDR, at an acceptable rate, namely α. Consequently, the decision to reject or fail to reject any null hypothesis depends on an m-dimensional vector of test statistics, $\boldsymbol{T}_n = (\boldsymbol{T}_n(i) : i = 1, \ldots, m)$; which are functions of the data, $\mathcal{X}_n$. Without loss of generality, large values of $\boldsymbol{T}_n(i)$ are assumed to indicate evidence against the null hypothesis. Let $Q_n = Q_n(P)$ denote the, typically unknown, joint distribution of the test statistics $\boldsymbol{T}_n$. As discussed in previous chapters, in practice Q_n is replaced by a null distribution, Q_0. For a given multiple testing procedure, let

$$\mathcal{R}_n = \mathcal{R}(\boldsymbol{T}_n, Q_0, \alpha) = \{i : H_0(i) \text{ is rejected}\} = \{i : \boldsymbol{T}_n(i) > C_i\}, \tag{4.1}$$

where $\mathcal{R}_n$ is a set of rejected hypotheses and $C_i = C(i; T_n, Q_0, \alpha)$ are threshold values for deciding whether or not to reject the i^{th} null hypothesis. Denote the number of rejections and the number of false rejections based on the procedure by R and V respectively such that,

$$R =| \mathcal{R}(\boldsymbol{T}_n, Q_0, \alpha) |=| \mathcal{R}_n | \qquad \text{and}$$

$$V = | \mathcal{R}(\boldsymbol{T}_n, Q_0, \alpha) \cap \mathcal{H}_0(P) | = | \mathcal{R}_n \cap \mathcal{H}_0 | . \tag{4.2}$$

The following remarks can be made about (4.2).

Remark 4.2.1

The use of the long notation in $\mathcal{R}(\boldsymbol{T}_n, Q_0, \alpha)$, the set of rejected hypotheses, indicates that $\mathcal{R}_n$ is a function of

i. the data $\mathcal{X}_n$, through an m-vector of test statistics, $\boldsymbol{T}_n$, where each $\boldsymbol{T}_n(i)$ corresponds to a null hypothesis $H_0(i)$.

ii. the null distribution of the test statistics, Q_0, for computing the threshold values, C_i, for each $\boldsymbol{T}_n(i)$.

iii. the pre-specified significance level, α.

Recall, the false discovery rate (FDR) is defined as the expected number of false rejections among those declared significant. Using the above notation, the FDR is simply

$$FDR = E\left(\frac{V}{\max(R, 1)}\right) \tag{4.3}$$

Following Remark 4.2.1, the FDR is a function of the test statistics, $\boldsymbol{T}_n(i)$, the test statistics' null distribution, Q_0, and the pre-specified significance level, α. An optimal FDR-controlling procedure requires reliable estimation of the variance components and subsequently the test statistics and the corresponding joint null distribution. However, for a large number of hypotheses with a comparatively small sample size, the traditional t-statistic is suboptimal. As previously reviewed, this is as a consequence of fluctuations in the estimation of the variance components. Additionally, the presence of correlation among the test statistics can have a significant effect on the usually employed theoretical null, resulting in a distribution that is substantially wider or narrower than optimal. Thus, an optimal FDR-controlling procedure requires a good estimate of the variance components and an accurate representation of the null distribution. Despite this insight, many FDR-controlling procedures are developed under the assumption that good estimators for the variance components and a valid approximation to the joint distribution of the test statistics are available. Practitioners are therefore faced with the challenge of selecting these two components when dealing with large-scale testing. An FDR-controlling procedure that incorporates a good estimator for the error variance and an appropriate test statistics' null distribution, while accounting for dependencies would thus be optimal. For this purpose, we will construct a step-down multiple comparison procedure for the control of FDR via resampling in the following. This procedure incorporates both estimation of an appropriate test statistics' null distribution and a James-Stein-type analytic shrinkage estimator for the variance components.

Before proceeding, it should be noted that a multiple testing procedure provides a desired finite sample control at level α over the FDR if

$$FDR_P \leq \alpha \quad \forall\, P \in \Omega. \tag{4.4}$$

However, if the procedure controls the FDR asymptotically at level α, then

$$\lim_{n\to\infty} \sup FDR_P \leq \alpha \quad \forall\, P \in \Omega, \tag{4.5}$$

where P is the data-generating distribution.

4.3. Step-down Multiple Testing Procedure

Denote the ordered test statistics by $T_{n,(1)} \leq \cdots \leq T_{n,(m)}$ with the corresponding null hypotheses $H_0^{(1)}, \ldots, H_0^{(m)}$. A step-down procedure begins with the most significant test statistic. First, the joint null hypothesis that all hypotheses, $H_0(i)$, $i = 1, \ldots, m$ are true is tested. This hypothesis is rejected if $T_{n,(m)}$ is large. If it is not large, then the procedure fails to reject all of the hypotheses; otherwise, the procedure rejects the hypothesis corresponding to the largest test statistic. Once a hypothesis is rejected, it is removed and the remaining hypotheses are tested by rejecting for large values of the maximum of the remaining test statistics, and this procedure continues until there are no more rejections. A description of this generic step-down procedure is provided in Algorithm 4.1.

Algorithm 4.1 Generic Step-down Procedure

1. If $T_{n,(m)} < c_m$, reject no hypotheses and stop. Otherwise, reject $H_0^{(m)}$ and continue.
2. If $T_{n,(m-1)} < c_{m-1}$, reject no further hypotheses and stop. Otherwise, reject $H_0^{(m-1)}$ and continue.

   ⋮

j. If $T_{n,(m-j+1)} < c_{m-j+1}$, reject no further hypotheses and stop. Otherwise, reject $H_0^{(m-j+1)}$ and continue.

   ⋮

m. If $T_{n,(1)} < c_1$, fail to reject $H_0^{(1)}$; otherwise reject $H_0^{(1)}$.

More concisely, suppose $j^\star$ is the largest integer j such that

$$T_{n,(m)} \geq c_m, \ldots, T_{n,(m-j)} \geq c_{m-j},$$

then a step-down multiple testing procedure will reject the hypotheses,

$$H_0^{(m)}, \ldots, H_0^{(m-j^\star)}.$$

However, the procedure will not reject any null hypotheses if no such j exists.

4.3.1. Calculation of the Critical Values

Suppose A_i is the probability that exactly i hypotheses are rejected for any step-down procedure. Then

$$
\begin{aligned}
A_0 &= P(T_{n,(m)} < c_m) \\
A_1 &= P(T_{n,(m)} \geq c_m, T_{n,m-1} < c_{m-1}) \\
&\vdots \\
A_r &= P(T_{n,(m)} \geq c_m, \ldots, T_{n,(m-r+1)} \geq c_{m-r+1}, T_{n,(m-r)} < c_{m-r})
\end{aligned} \tag{4.6}
$$

Following equation (4.6), the FDR of a step-down procedure can be expressed as

$$
\begin{aligned}
FDR_P = E_P\left(\frac{V}{\max(R,1)}\right) &= \sum_{r=1}^{m} \frac{1}{r} E_P(V|R=r)P(R=r) \\
&= \sum_{r=1}^{m} \frac{1}{r} E_P(V|R=r) \times P(T_{n,(m)} \geq c_m, \ldots, T_{n,(m-r+1)} \geq c_{m-r+1}, T_{n,(m-r)} < c_{m-r}),
\end{aligned} \tag{4.7}
$$

where the event $T_{n,(m-r)} < c_{m-r}$ is enforced only when $r < m$. Equation (4.7) can be shown to be asymptotically equivalent to

$$
\begin{aligned}
FDR_P = &\sum_{r=m-m_0+1}^{m} \frac{r-m+m_0}{r} \\
&\times P(T_{n,m_0:m_0} \geq c_{m_0}, \ldots, T_{n,m-r+1:m_0} \geq c_{m-r+1}, T_{n,m-r:m_0} < c_{m-r}),
\end{aligned} \tag{4.8}
$$

where $T_{n,r:m_0}$ denotes the r^{th} largest of the test statistics corresponding to the true null hypotheses and again the event $T_{n,(m-r)} < c_{m-r}$ is enforced only when $r < m$.

The aim in developing an optimal procedure is to choose $c_1, c_2, \ldots, c_m$ such that (4.8) is at least asymptotically bounded above by α for any data-generating distribution, P. The threshold values, $c = (c_i : i = 1, \dots, m)$ will be determined as follows. To obtain the first threshold value, consider any data-generating procedure such that there is only one true null hypothesis, i.e., $m_0 = 1$. Then, (4.8) reduces to

$$FDR_P = \frac{1}{m} P\left(T_{n,1:1} \geq c_1\right) \tag{4.9}$$

Subject to this, c_1 is chosen as the minimum value for which (4.9) is bounded above by α. That is,

$$c_1 := \inf\left\{x \in \mathbb{R} : \frac{1}{m} P\left(T_{n,1:1} \geq x\right) \leq \alpha\right\} \tag{4.10}$$

It should be noted that c_1 so defined is $-\infty$ when $m\alpha \geq 1$. Next, consider any data-generating procedure such that there are only two true null hypotheses, i.e., $m_0 = 2$. Then, (4.8) reduces to

$$FDR_P = \frac{1}{m-1} P\left(T_{n,2:2} \geq c_2, T_{n,1:2} < c_1\right) + \frac{2}{m} P\left(T_{n,2:2} \geq c_2, T_{n,1:2} \geq c_1\right) \tag{4.11}$$

Again, an appropriate choice of c_2 is the minimum value for which (4.11) is bounded above by α. That is,

$$c_2 := \inf\left\{x \in [c_1, \infty) : \frac{1}{m-1} P\left(T_{n,2:2} \geq x, T_{n,1:2} < c_1\right) + \frac{2}{m} P\left(T_{n,2:2} \geq x, T_{n,1:2} \geq c_1\right) \leq \alpha\right\} \tag{4.12}$$

In general, to obtain the j^{th} critical value, consider any data-generating distribution, P such that $m_0 = j$. Then, having determined $c_1, c_2, \dots, c_{j-1}$, an appropriate choice of c_j is the minimum value of c for which

$$FDR_P = \sum_{r=m-j+1}^{m} \frac{r-m+j}{r} \times P(T_{n,j:j} \geq x, \ldots, T_{n,m-r+1:j} \geq c_{m-r+1}, T_{n,m-r:j} < c_{m-r}), \tag{4.13}$$

is bounded above by α. If a solution to (4.13) exists in $[c_{j-1}, \infty)$ then that is c_j. Otherwise, c_j is set equal to c_{j-1}. The selection of the above critical values is, however, impossible due to its dependence on the unknown data-generating distribution, P, through the distribution of the test statistics. Romano, Shaikh, and Wolf (2008) (referred to as RSW hereafter) proposed a bootstrap approach that relies on an exchangeability assumption - albeit not in combination with estimation of the test statistics' null distribution and variance components. In their work, they replaced the unknown data-generating distribution with a suitable estimate $\hat{P}_n$ and then utilized bootstrap techniques to estimate the distribution of the test statistics, and subsequently the critical values. However, as emphasized by Pollard and van der Laan (2004) and Efron (2004, 2007a) utilizing a data-generating distribution to estimate the distribution of test statistics may incorrectly specify the dependence structure of the test statistics. In the presence of strong correlations among the test statistics, utilizing the RSW FDR-controlling procedure may undercut inferential validity. To this end, this study proposes constructing the critical values by first utilizing an appropriate estimate of the variance components to construct the test statistics, and then replacing the unknown joint distribution of the test statistics with an appropriate null distribution. One main distinction of this approach to that of Romano, Shaikh, and Wolf (2008) is that an appropriate null distribution is utilized in place of the unknown joint distribution. The proposed methodologies are detailed herein.

4.3.2. A Proposed Bootstrap Approach to FDR Control

As in the previous chapter, we will focus on a parameter vector,

$$\boldsymbol{\theta}(P) = (\theta_1(P), \ldots, \theta_m(P)) \tag{4.14}$$

Consider the one-sided testing problem, in which case (without loss of generality)

$$H_0(i) : \theta(i) \leq \theta_0(i) \quad \text{vs.} \quad H_1(i) : \theta(i) > \theta_0(i) \tag{4.15}$$

or the two-sided testing problem, in which case

$$H_0(i) : \theta(i) = \theta_0(i) \quad \text{vs.} \quad H_1(i) : \theta(i) \neq \theta_0(i) \tag{4.16}$$

The test statistics will be based on the shrinkage t statistic constructed in section 3.2.1. Now, as stated earlier, the selection of the critical values in (4.13) is impossible due to their dependence on the unknown distribution of the test statistics. Here, instead of utilizing a data-generated null distribution as proposed by Romano, Shaikh, and Wolf (2008), we will replace the unknown test statistic distribution by an appropriate null distribution, $Q_0(P)$. We re-emphasize that the use of an inappropriate null distribution may lead to misleading results, especially in the presence of strong correlations among the variables. Thus, one could expect to obtain an improved FDR-controlling procedure by incorporating an appropriate null distribution into the RSW procedure, especially in cases where the data-generated test statistic null distribution fails. Two different null distributions are considered; the null shift and scale-transformed null distribution, $Q_0^{NS}(P)$, proposed by Pollard and van der Laan (2004) and generalized by Dudoit, van der Laan, and Pollard (2004) and the null quantile-transformed null distribution, $Q_0^{NQ}(P)$, proposed by van der Laan and Hubbard

(2006). However, in practice, since the data-generating distribution, P, is unknown, so is the null distribution, $Q_0(P)$. Bootstrap procedures will be utilized to obtain consistent estimators, Q_{0n} (Q_{0n}^{NS} or Q_{0n}^{NQ}) of the null distributions. For this purpose, let P_n denote the empirical distribution corresponding to P, which assigns probability $(1/n)$ to each realization of X. Let $\mathcal{X}_n^\star = \{X_i^\star : i = 1, \ldots, n\}$ be distributed according to P_n and denote by $\boldsymbol{T}_n^\star$, the m-dimensional vector of test statistics computed from $\mathcal{X}_n^\star$. Then, $Q_0^{NS}(P)$ can be estimated by the distribution of the null shift and scale-transformed bootstrap test statistics

$$\tilde{\boldsymbol{Z}}_n^\star(i) = \sqrt{\min\left(1, \frac{\boldsymbol{\tau}_0(i)}{\mathrm{Var}_{P_n}\left(\boldsymbol{T}_n^\star(i)\right)}\right)}\Big(\boldsymbol{T}_n^\star(i) + \boldsymbol{\lambda}_0(i) - E\big(\boldsymbol{T}_n^\star(i)\big)\Big); \quad i = 1, \ldots, m \qquad (4.17)$$

Similarly, $Q_0^{NQ}(P)$ can be estimated by the distribution of the null quantile-transformed bootstrap test statistics

$$\breve{\boldsymbol{Z}}_n^\star(i) = q_{0,i}^{-1} Q_{n,i}^{\star,\Delta}(\boldsymbol{T}_n^\star(i)), \qquad (4.18)$$

where $Q_{n,i}^{\star,\Delta}(z) = \Delta Q_{n,i}^\star(z) + (1 - \Delta) Q_{n,i}^\star(z^-)$, Δ is a uniform random variable on the interval $[0, 1]$, independent of the data, and $Q_{n,i}^\star(z)$ is the marginal cumulative distribution function based on $\mathcal{X}_n^\star$. The bootstrap estimation of the null shift and scale-transformed null distribution $Q_{0n}^{NS}(P)$ and the null quantile-transformed null distribution, $Q_{0n}^{NQ}(P)$ based on the shrinkage t statistic are summarized in algorithms 3.4 and 3.5 in Chapter 3.

With the estimated null distribution, Q_{0n}, and matrix of test statistics, $\boldsymbol{Z}_n^\star$, (either the null shift and scale-transformed bootstrap test statistics, $\tilde{\boldsymbol{Z}}_n^\star$, or the null quantile-transformed bootstrap test statistics, $\breve{\boldsymbol{Z}}_n^\star$) from either Algorithm 3.4 or 3.5, the critical values can be defined recursively as follows. Compute the j^{th} critical value, having already determined $\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_{j-1}$ using the rule

$$\hat{c}_j = \inf\Bigg\{x \in \mathbb{R} : \sum_{r=m-j+1}^{m} \frac{r - m + j}{r}$$

$$\times P_n(Z^\star_{n,j:j} \geq x, \ldots, Z^\star_{n,m-r+1:j} \geq \hat{c}_{m-r+1}, Z^\star_{n,m-r:j} < \hat{c}_{m-r}) \leq \alpha \Big\}. \tag{4.19}$$

where the event $Z^\star_{n,(m-r):j} < \hat{c}_{m-r}$ is enforced only when $r < m$.

Remark 4.3.1

The following remarks can be made about the bootstrap approach to selecting the critical values.

1. *Some clarifications need to be provided with regards to the notation, $Z^\star_{n,r:t}$ with $r \leq t$. Note that for t true null hypotheses, $T_{n,r:t}$ corresponds to the r^{th} largest of the observations corresponding to these true hypotheses. However, the ordering of the null hypotheses in the bootstrap world is determined by the ordering of the hypotheses corresponding to the ordered test statistics, $H_{(1)}, \ldots, H_{(m)}$ from the "real" world, not according to $1, \ldots, m$. Thus, to obtain $Z^\star_{n,r:t}$, the bootstrap test statistics need to be permuted so that if $\{k_1, \ldots, k_m\}$ of $\{1, \ldots, m\}$ is such that $H_{k_1} = H_{(1)}, \ldots, H_{k_m} = H_{(m)}$, then $Z^\star_{n,r:t}$ corresponds to the r^{th} largest of the observations $Z^\star_{n,k_1}, \ldots Z^\star_{n,k_t}$.*

2. *Closed-form expressions for the probabilities in (4.19) may be typically impossible to compute. A researcher may thus use simulations to any desired degree of accuracy to compute the critical values. In practice, however, one needs to find a balance between computational cost and estimation accuracy.*

The proposed bootstrap algorithm for the estimation of the critical values in (4.13) is summarized in Algorithm 4.2. In the next section, formal theoretical justification and conditions for when the proposed step-down procedure with critical values defined by (4.13) provides asymptotic control over the FDR will be provided.

Algorithm 4.2 Proposed Bootstrap FDR-Controlling Procedure

1. Apply Algorithm 3.4 or 3.5 to generate an $m \times B$ matrix of null-transformed bootstrap test statistics, $\boldsymbol{Z}_n^\star$. The bootstrap estimator of the null distribution, Q_0 is the empirical distribution of the columns of $\boldsymbol{Z}_n^\star$.

2. Compute the j^{th} critical value, having already determined $\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_{j-1}$ using the rule

$$\hat{c}_j = \inf \Bigg\{ c \in \mathbb{R} : \frac{1}{B} \sum_{b=1}^{B} \sum_{r=m-j+1}^{m} \frac{r-m+j}{r} \times I(Z^\star_{n,j:j} \geq c, \ldots, Z^\star_{n,m-r+1:j} \geq \hat{c}_{m-r+1}, Z^\star_{n,m-r:j} < \hat{c}_{m-r}) \leq \alpha \Bigg\}. \tag{4.20}$$

3. Let $T_{n,(1)} \leq \cdots \leq T_{n,(m)}$ be the ordered test statistics with the corresponding null hypotheses $H_0^{(1)}, \ldots, H_0^{(m)}$.

4. Suppose $j^\star$ is the largest j such that $T_{n,(m)} \geq \hat{c}_m, \ldots, T_{n,(m-j)} \geq \hat{c}_{m-j}$, then, reject the hypotheses $H_0^{(m)}, \ldots, H_0^{(m-j^\star)}$. Reject nothing if no such j exists.

4.4. Some Analytical Results

In what follows, we provide conditions for when the proposed resampling-based step-down procedure provides asymptotic FDR control.

Lemma 4.4.1 (Assumption I: Asymptotic Separation of Null Hypotheses)
Consider testing the set of m null hypotheses, $H_0(i)$, against the alternative hypotheses, $H_1(i)$, $i = 1, \ldots, m$ based on the test statistics, $\boldsymbol{T}_n(i)$. Without loss of generality, we assume that m_0, $0 \leq m_0 \leq m$, hypotheses are true. Assume further that the hypotheses have been relabeled such that $\boldsymbol{T}_n(1)$, $\boldsymbol{T}_n(2), \ldots,$ $\boldsymbol{T}_n(m_0)$ correspond to the true null hypotheses. Intuitively, if

$\boldsymbol{T}_n(i)$, $i = 1, \ldots, m$ is a consistent test statistic for testing $H_0(i)$ and $T_{n,(1)} \leq \cdots \leq T_{n,(m)}$ are the corresponding ordered test statistics, then asymptotically, we will expect the ordering

$$T_{n,(1)} \leq T_{n,(2)} \leq \cdots \leq T_{n,(m_0)} \leq T_{n,(m_0+1)} \leq \cdots T_{n,(m)}$$

to correspond to

$$H_0^{(1)}, H_0^{(2)}, \ldots, H_0^{(m_0)}, H_0^{(m_0+1)}, \ldots, H_0^{(m)},$$

where $H_0^{(1)}, H_0^{(2)}, \ldots, H_0^{(m_0)}$ are true null hypotheses and $H_0^{(m_0+1)}, \ldots, H_0^{(m)}$ are false null hypotheses.

In the proposed methodology, we do not know of a general formula expressing the conditions for which the FDR will be equal to α. However, by construction, we will expect to achieve asymptotic FDR control by utilizing test statistics that satisfy Lemma 4.4.1. But, asymptotic control is based on the true data-generating distribution P through the distribution of the test statistics which is usually unknown and needs to be estimated. In the previous section, we proposed a bootstrap approach to estimate the unknown data-generating distribution and subsequently the test statistic and utilized the estimated distribution to determine the critical values. We herein provide conditions for when the proposed bootstrap approach provides asymptotic FDR control.

Lemma 4.4.2

Suppose the statistics, $\boldsymbol{T}_n(i), \ldots, \boldsymbol{T}_n(m)$ are available for testing m hypotheses and $\boldsymbol{T}_n \sim Q_n = Q_n(P)$. Let $\mathcal{H}_0 = \mathcal{H}_0(P)$ be the set of true null hypotheses and $\mathcal{H}_1 = \mathcal{H}_1(P)$, the set of false null hypotheses where P is the data-generating distribution. Assume also that there exists an m-dimensional known real-valued vector $\boldsymbol{\lambda}_0$, and a positive real-valued vector $\boldsymbol{\tau}_0$ of null *values such that*

$$\limsup_{n\to\infty} E\left(\boldsymbol{T}_n(i)\right) \leq \boldsymbol{\lambda}_0(i) \qquad \textit{and}$$
$$\limsup_{n\to\infty} Var\left(\boldsymbol{T}_n(i)\right) \leq \boldsymbol{\tau}_0(i) \qquad \textit{for} \;\; i \in \mathcal{H}_0. \tag{4.21}$$

Define an m-dimensional vector of null shift and scale-transformed test statistics whose entries are determined by

$$\tilde{\boldsymbol{Z}}_n(i) = \sqrt{\min\left(1, \frac{\boldsymbol{\tau}_0(i)}{Var\left(\boldsymbol{T}_n(i)\right)}\right)}\Big(\boldsymbol{T}_n(i) + \boldsymbol{\lambda}_0(i) - E(\boldsymbol{T}_n(i))\Big); \quad i = 1, \ldots, m. \tag{4.22}$$

Suppose that the m-dimensional vector of test statistics, $\tilde{\boldsymbol{Z}}_n$, weakly converges to a random m-dimensional vector $\tilde{\boldsymbol{Z}}$ such that $\tilde{\boldsymbol{Z}}$ has a continuous joint distribuiton $Q_0 = Q_0(P)$, i.e.,

$$\tilde{Z}_n \xrightarrow{\mathcal{L}} \tilde{Z} \sim Q_0. \tag{4.23}$$

Then, for all $c = (c_i : i = 1, \ldots, m) \in \mathbb{R}$ and $c_1 \leq c_2 \leq \cdots \leq c_m$ and this choice of null distribution, $Q_0(P)$ and for all $x \in \mathbb{R}$

$$\limsup_{n\to\infty} P_{Q_n}\Big(\big(T_{n,j:j} \geq c_j, \ldots, T_{n,m-r+1:j} \geq c_{m-r+1}, T_{n,m-r:j} < c_{m-r}\big) \leq x\Big)$$
$$\leq P_{Q_0}\Big(\big(\tilde{Z}_{j:j} \geq c_j, \ldots, \tilde{Z}_{m-r+1:j} \geq c_{m-r+1}, \tilde{Z}_{m-r:j} < c_{m-r}\big) \leq x\Big) \tag{4.24}$$

where $j \in m_0$ and where $T_{n,r:j}$ and $Z_{r:j}$ denotes the r^{th} largest of the test statistics corresponding to the true null hypotheses for their respective test statistics. That is, the joint distribution of the $\mathcal{H}_0$-specific test statistics under the null distribution, Q_0, is asymptotically stochastically larger than under the true distribution, Q_n.

We will prove Lemma 4.4.2 in an analoguous manner to the proof of the joint null domination assumption provided in Dudoit and van der Laan (2008, Chapter 2).

Proof of Lemma 4.4.2. First, define an intermediate random vector $\boldsymbol{S}_n = (\boldsymbol{S}_n(i) : i = 1, \ldots, m)$ as

$$\boldsymbol{S}_n(i) = \boldsymbol{T}_n(i) + \max\left\{0, \boldsymbol{\lambda}_0(i) - E\big(\boldsymbol{T}_n(i)\big)\right\}, \quad i = 1, \ldots, m. \tag{4.25}$$

Here, we note that since the second term in $\boldsymbol{S}_n(i)$ is always zero or greater, we have $\boldsymbol{S}_n(i) \geq \boldsymbol{T}_n(i)$ for each $i = 1, \ldots, m$. Now, by (4.21) we have

$$\lim_{n\to\infty} \sqrt{\min\left\{1, \frac{\boldsymbol{\tau}_0(i)}{\mathrm{Var}\,(\boldsymbol{T}_n(i))}\right\}} = 1 \qquad \text{and} \qquad \lim_{n\to\infty} \big(\boldsymbol{\lambda}_0(i) - E\big(\boldsymbol{T}_n(i)\big)\big) = 0. \tag{4.26}$$

It then follows that

$$\lim_{n\to\infty} \tilde{\boldsymbol{Z}}_n(i) = \boldsymbol{T}_n(i) \qquad \text{and} \qquad \lim_{n\to\infty} \boldsymbol{S}_n(i) = \boldsymbol{T}_n(i) \tag{4.27}$$

Thus, the null specific subvectors $(\boldsymbol{S}_n(i) : i \in \mathcal{H}_0)$ and $(\tilde{\boldsymbol{Z}}_n(i) : i \in \mathcal{H}_0)$ have the same asymptotic joint null distribution. So by assumption, $\boldsymbol{S}_n$ also converges weakly to $\tilde{\boldsymbol{Z}}$, that is,

$$(\boldsymbol{S}_n(i) : i \in \mathcal{H}_0) \xrightarrow{\mathcal{L}} (\tilde{\boldsymbol{Z}}(i) : i \in \mathcal{H}_0) \sim Q_{0,\mathcal{H}_0}. \tag{4.28}$$

Now, by the continuous mapping theorem and for each $x \in \mathbb{R}$ we have

$$\begin{aligned}
&\limsup_{n\to\infty} P_{Q_n}\Big(\big(T_{n,j:j} \geq c_j, \ldots, T_{n,m-r+1:j} \geq c_{m-r+1}, T_{n,m-r:j} < c_{m-r}\big) \leq x\Big) \\
&\quad\leq \limsup_{n\to\infty} P_{Q_0}\Big(\big(S_{n,j:j} \geq c_j, \ldots, S_{n,m-r+1:j} \geq c_{m-r+1}, S_{n,m-r:j} < c_{m-r}\big) \leq x\Big) \\
&\quad= P_{Q_0}\Big(\big(Z_{j:j} \geq c_j, \ldots, Z_{m-r+1:j} \geq c_{m-r+1}, Z_{m-r:j} < c_{m-r}\big) \leq x\Big)
\end{aligned} \tag{4.29}$$

where $j \in m_0$ and $T_{n,r:j}$, $S_{n,r:j}$, and $Z_{r:j}$ denotes the r^{th} largest of the test statistics corresponding to the true null hypotheses for their respective test statistics. □

The fundamental results underlining the proposed resampling-based procedure are summarized in the following theorem.

Theorem 4.4.1

Consider testing the set of m null hypotheses, $H_0(i)$, against the alternative hypotheses, $H_1(i)$, $i = 1, \ldots, m$ based on the shrinkage t statistics, $\boldsymbol{T}_n(i)$ given in section 3.2.1.4. Suppose that the conditions in Lemma 4.4.2 are satisfied. Suppose further that the $\mathcal{H}_0$-specific joint distribution of the test statistic

i. has continuous marginal distributions.

ii. has connected support.

iii. satisfies the asymptotic null domination assumption, that is, the joint distribution of the $\mathcal{H}_0$-specific test statistics is asymptotically stochastically larger under the null distribution Q_0 than under the true distribution Q_n.

Then the step-down procedure with critical values described in Algorithm 4.2 provides asymptotic control over the FDR. That is,

$$\limsup_{n \to \infty} FDR \leq \alpha. \tag{4.30}$$

4.5. Ongoing Efforts

Preliminary simulation results have shown that the proposed step-down procedure provides asymptotic FDR control. In the next stage of this work, Monte Carlo simulation studies will be carried out to assess finite sample performance of the proposed procedure. Additionally, FDR control, power, and stability as characterized by the standard deviations of the number of false hypotheses rejected and the total number of rejected hypotheses of the proposed step-down procedure will be compared with some existing and commonly employed FDR-controlling procedures.

Our ongoing efforts also include providing complete theoretical proofs of asymptotic FDR control for the proposed step-down procedure based on the James-Stein-type analytic shrinkage estimation and the null distribution.

4.6. Summary and Conclusions

Modern statistical inference problems in areas such as medicine, spatial epidemiology, genomics, and marketing, routinely involve statistically testing for some behavior of interest on each of thousands or more measurements taken on the same unit. This usually involves inference for high-dimensional multivariate distributions with complex and mostly unspecified dependencies among the variables under consideration. The nature of analysis of such data in the initial stages is normally exploratory so the false discovery rate is the commonly employed measure to control the inflation of type I errors. To date, various FDR-controlling procedures have been proposed for the analysis of such high-dimensional data. However, many of the existing procedures are based on the marginal distributions, failing to account for the dependence structure of the test statistics. Moreover, some of these marginal procedures are developed under specific assumptions about the joint distribution of the test statistics, such as independence or some form of weak dependence. Consequently, these methods tend to lose power when the test statistics are highly correlated. FDR-controlling procedures that incorporate information about the dependence structure of the test statistics remain limited.

In addition to the above, developing cut-off values for FDR-controlling procedures require knowledge of the distribution of the test statistics. In practice, however, the true distribution of the test statistic is unknown and is usually replaced by a theoretical or data-generated null distribution. Resampling techniques provide the flexibility of estimating the distribution of the test statistics, and by so doing, account for the complex and unknown dependence structure among the test statistics. Romano, Shaikh, and Wolf (2008) provided a bootstrap step-down procedure that was shown to provide asymptotic FDR control under fairly weak assumptions, but required an exchangeability assumption for the joint limiting distribution of the null-specific test statistics. In their procedure, the authors replaced the unknown data-generating distribution with a suitable estimate and subsequently utilized the estimated data-generating distribution to estimate the joint distribution of the test statistics. However, as pointed out by Pollard and van der Laan (2004) and Efron (2004, 2007a), utilizing a data-generating distribution to estimate the distribution of the test statistics may incorrectly specify the dependence structure of the test statistics. Thus in the presence of high correlations, the step-down procedure of Romano, Shaikh, and Wolf (2008) may undercut inferential validity.

We also note that for a large number of hypotheses with comparatively small sample sizes, the traditional t-statistic is suboptimal. This is normally due to fluctuations in the estimation of the variance components. To this end, an optimal FDR-controlling procedure requires a good estimate of the variance components and an accurate representation of the test statistic null distribution. Here, we extended the step-down procedure of Romano, Shaikh, and Wolf (2008) by incorporating a James-Stein-type analytic shrinkage estimation of the error variance and a generally valid null distribution. Asymptotic validity of the proposed method holds under the asymptotic null domination assumption for the null-specific test statistic distribution.

Since the proposed procedure is based on asymptotic arguments, it is necessary to shed light on its finite sample performance. Future work will address this, in addition to providing complete theoretical proofs for asymptotic FDR control.

CHAPTER V

CONCLUDING REMARKS

5.1. Summary

Rapid advancement in technology, especially in genomics and imaging has redefined how statisticians approach simultaneous inference problems. For example, microarray experiments generate large multiplicity problems in which a researcher may statistically test thousands of hypotheses simultaneously to identify which of the genes are differentially expressed. This usually involves inference for high-dimensional multivariate distributions with complex and mostly unspecified dependencies among the variables under consideration. In these situations, an unguarded use of single-inference procedures inflates the overall type I error rates. This has lead to an explosion in multiple testing literature on statistical methods for large-scale inference. Multiple testing procedures provide efficient methods for examining each hypothesis while also controlling for an overall error rate at a pre-specified level. An optimal multiple testing procedure needs to take into account the ramifications of three choices: (i) choice of a suitable test statistic, (ii) choice of a test statistic null distribution, and (iii) control of an overall error rate.

One line of research in the field of multiple testing deals with the control of an overall error measure. Classical approach to simultaneous inference controls the family-wise error rate (FWER), defined as the probability of making at least one type I error. However, FWER procedures offer extremely stringent control of the error which might not always be appropriate. For instance, the number of tests in most large-scale inference is large and the nature of analysis is exploratory rather than confirmatory. In such cases, one often wishes to make as many discoveries as possible without resulting in too many false discoveries, although some false discoveries can be tolerated. Benjamini and Hochberg (1995) introduced the false discovery rate (FDR) as an alternative to the FWER. FDR-controlling procedures are less stringent but powerful multiple testing procedures for large-scale inference than the FWER-controlling procedures, and are therefore the preferred error rate to control in such studies. Thus far, there is a rich body of literature on FDR-controlling procedures. Many of the existing procedures are based on the marginal distributions of the test statistics without taking into account their dependence structure. Moreover, some of these marginal procedures are developed under specific assumptions about the joint distribution of the test statistics, such as independence or some form of weak dependence. Even so, these procedures still do not account for the assumed dependence structure. They therefore become less powerful than a procedure which incorporates dependence in some way, especially when the test statistics are highly correlated.

Another line of research has been to develop optimal test statistics for large-scale inference. For a large number of hypotheses with comparatively small sample sizes, the traditional t-statistic is suboptimal. As discussed earlier, this is normally due to fluctuations in the estimation of the variance components. Accordingly, various test statistics have been suggested in the past couple of years; some of which involve modifying estimators of the error variance components.

A third line of research has been to develop alternative null distributions for use in large-scale inference. FDR procedures are based on cut-off values which are derived from the joint distribution of the test statistics. In practice, this distribution is unknown and is often replaced by a theoretical null distribution or a data-generated null distribution. However, as pointed out in Pollard and van der Laan (2004) and Efron (2004, 2007a), the usually employed theoretical null or the data-generated null distribution can misspecify the dependence structure of the test statistic. Thus, multiple testing procedures can perform significantly worse if an inappropriate null distribution is utilized.

In this study, we have developed a unified approach to FDR control that takes into account all the three aspects of developing an optimal multiple testing procedures, that have otherwise been considered separately. Our contribution is in two-fold. In the first part of the study, we proposed improved resampling-based procedures by incorporating a generally valid null distribution and a James-Stein-type analytic shrinkage estimation of the variance components into the original linear step-up procedure of Benjamini and Hochberg (1995) and the two-stage linear adaptive procedure of Benjamini, Krieger, and Yekutieli (2006). Two null distributions were considered: the null shift and scale-transformed null distribution, and the null quantile-transformed null distribution. Under the assumptions of null domination and positive regression dependent on the subset of test statistics corresponding to the true null hypotheses, the resampling-based procedures were shown to provide asymptotic FDR control. We also compared the proposed procedures with the linear step-up procedures of Benjamini and Hochberg (1995), Benjamini and Yekutieli (2001), Benjamini, Krieger, and Yekutieli (2006) and the q-value procedure of Storey (2003); Storey, Taylor, and Siegmund (2004) using extensive Monte Carlo simulations for case-control experiments. The simulation results indicated that the resampling-based procedure based on the null quantile-transformed null distribution is essentially more stable and as powerful or substantially more powerful than the q-value procedure of Storey (2003); Storey, Taylor, and Siegmund (2004), provided there are at least 30 observations in both the case and control groups.

In the second part of the study, we extended the step-down procedure of Romano, Shaikh, and Wolf (2008) by incorporating a James-Stein-type analytic shrinkage estimation of the error variance and a generally valid null distribution. Asymptotic validity of the proposed method holds under the asymptotic null domination assumption for the null-specific test statistic distribution.

5.2. Future Directions

The proposed step-down procedure is justified by asymptotic arguments, it is therefore important to investigate its finite sample performance using Monte Carlo simulation studies. Future work will address this, in addition to providing complete theoretical proofs of asymptotic FDR control.

The null shift and scale-transformed null distribution and the null quantile-transformed null distribution employed in this study are justified by bootstrap arguments. In general, there is no recommendations for the number of bootstrap resamples to utilize. However, in order to deal with the discreteness of the bootstrap distribution, especially for estimating very small p-values, one obviously needs a very large bootstrap resample. An alternative is to replace the marginal null distributions obtained from the null distributions with Gaussian approximations or smoothed estimation methods. Another future direction will be to explore specific algorithms for accurate estimation of the tail probabilities.

5.3. Conclusions

In recent years, reproducibility of statistical findings has drawn considerable attention not only from statisticians, but from all researchers engaged in empirical discovery. As noted by Stodden (2015), the reasons for irreproducibility are at least, to some extent, due to the number of false discoveries in such studies. Thus, a multiple testing procedure with improved FDR control, good power and higher stability, can help alleviate the inconsistencies in statistical findings, especially in biomedical research. The improved FDR control, power and stability of the proposed resampling-based procedures under various testing scenarios allow the procedures to be very competitive with or outperform many procedures proposed in finite sample inferential problems, even under independence. This makes the proposed procedures very attractive in large-scale inference and a better alternative to the classical Benjamini and Hochberg (1995) approach.

5.4. Software Availability

The null shift and scale-transformed null distribution and the null quantile-transformed null distribution are implemented in the multest package in R statistical software as part of the Bioconductor project (Pollard, Dudoit, and van der Laan (2005)). The James-Stein-type analytic shrinkage estimation of the variance components is in the corpcor package in R statistical software (Schäfer, Opgen-Rhein, Zuber, Ahdesmaki, Silva, and Strimmer. (2017)). The R code for obtaining the step-down critical values discussed in chapter 4 are provided in Appendix D.

REFERENCES

Abramovich, F., Y. Benjamini, D. L. Donoho, and I. M. Johnstone (2006, 04). Adapting to unknown sparsity by controlling the false discovery rate. *Ann. Statist. 34*(2), 584–653.

Baldi, P. and A. D. Long (2001). A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes. *Bioinformatics 17*(6), 509–519.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological) 57*(1), 289–300.

Benjamini, Y. and Y. Hochberg (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics 25*(1), 60.

Benjamini, Y., A. M. Krieger, and D. Yekutieli (2006, sep). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika 93*(3), 491–507.

Benjamini, Y. and W. Liu (1999, dec). A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence. *Journal of Statistical Planning and Inference 82*(1-2), 163–170.

Benjamini, Y. and D. Yekutieli (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of statistics*, 1165–1188.

Bickel, P. J. and D. A. Freedman (1981). Some asymptotic theory for the bootstrap. *The annals of statistics*, 1196–1217.

Blanchard, G. and É. Roquain (2009). Adaptive false discovery rate control under independence and dependence. *Journal of Machine Learning Research 10*(Dec), 2837–2871.

Celisse, A. and S. Robin (2010). A cross-validation based estimation of the proportion of true null hypotheses. *Journal of Statistical Planning and Inference 140*(11), 3132 – 3147.

Chen, X. and R. W. Doerge (2012). Generalized estimators for multiple testing: proportion of true nulls and false discovery rate. In *Technical report*. Department of Statistics, Purdue University West Lafayette, USA.

Chen, Y., A. Wiesel, and A. O. Hero (2009, apr). Shrinkage estimation of high dimensional covariance matrices. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE.

Cui, X. and G. A. Churchill (2003, Mar). Statistical tests for differential expression in cDNA microarray experiments. *Genome Biology 4*(4), 210.

Cui, X., J. T. G. Hwang, J. Qiu, N. J. Blades, and G. A. Churchill (2005). Improved statistical tests for differential gene expression by shrinking variance components estimates. *Biostatistics 6*(1), 59–75.

Delmar, P., S. Robin, D. Tronik-Le Roux, and J. J. Daudin (2005). Mixture model on the variance for the differential analysis of gene expression data. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 54*(1), 31–50.

Dudoit, S., J. P. Shaffer, and J. C. Boldrick (2003). Multiple hypothesis testing in microarray experiments. *Statistical Science 18*(1), 71–103.

Dudoit, S. and M. J. van der Laan (2008). *Multiple testing procedures with applications to genomics*. Springer-Verlag New York.

Dudoit, S., M. J. van der Laan, and M. D. Birkner (2004). Multiple testing procedures for controlling tail probability error rates.

Dudoit, S., M. J. van der Laan, and K. S. Pollard (2004). Multiple testing. Part I. Single-step procedures for control of general type i error rates. *Statistical Applications in Genetics and Molecular Biology 3*(1), 1–69.

Dudoit, S., Y. H. Yang, M. J. Callow, and T. P. Speed (2002). Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica 12*(1), 111–139.

Efron, B. (2004). Large-scale simultaneous hypothesis testing. *Journal of the American Statistical Association 99*(465), 96–104.

Efron, B. (2007a). Correlation and large-scale simultaneous significance testing. *Journal of the American Statistical Association 102*(477), 93–103.

Efron, B. (2007b, 08). Size, power and false discovery rates. *Annals of Statistics 35*(4), 1351–1377.

Efron, B., R. Tibshirani, J. D. Storey, and V. Tusher (2001). Empirical Bayes analysis of a microarray experiment. *Journal of the American Statistical Association 96*(456), 1151–1160.

Fan, J., X. Han, and W. Gu (2012). Estimating false discovery proportion under arbitrary covariance dependence. *Journal of the American Statistical Association 107*(499), 1019–1035. PMID: 24729644.

Finner, H., T. Dickhaus, and M. Roters (2007, aug). Dependency and false discovery rate: Asymptotics. *The Annals of Statistics 35*(4), 1432–1455.

Fisher, T. J. and X. Sun (2011, may). Improved Stein-type shrinkage estimators for the high-dimensional multivariate normal covariance matrix. *Computational Statistics & Data Analysis 55*(5), 1909–1918.

Fox, R. J. and M. W. Dimmic (2006, Mar). A two-sample Bayesian t-test for microarray data. *BMC Bioinformatics 7*(1), 126.

Friedman, J. H. (1989). Regularized discriminant analysis. *Journal of the American Statistical Association 84*(405), 165–175.

Gavrilov, Y., Y. Benjamini, and S. K. Sarkar (2009, 04). An adaptive step-down procedure with proven FDR control under independence. *Annals of Statistics 37*(2), 619–629.

Ge, Y., S. Dudoit, and T. P. Speed (2003, Jun). Resampling-based multiple testing for microarray data analysis. *Test 12*(1), 1–77.

Ge, Y., S. C. Sealfon, and T. P. Speed (2008). Some step-down procedures controlling the false discovery rate under dependence. *Statistica Sinica 18*(3), 881.

Gottardo, R., A. E. Raftery, K. Y. Yeung, and R. E. Bumgarner (2005, jul). Bayesian robust inference for differential gene expression in microarrays with multiple samples. *Biometrics 62*(1), 10–18.

Greenland, S. (2000, feb). Principles of multilevel modelling. *International Journal of Epidemiology 29*(1), 158–167.

He, L. and S. K. Sarkar (2013). On improving some adaptive BH procedures controlling the FDR under dependence. *Electronic Journal of Statistics 7*(0), 2683–2701.

Hedenfalk, I., D. Duggan, Y. Chen, M. Radmacher, M. Bittner, R. Simon, P. Meltzer, B. Gusterson, M. Esteller, M. Raffeld, Z. Yakhini, A. Ben-Dor, E. Dougherty, J. Kononen, L. Bubendorf, W. Fehrle, S. Pittaluga, S. Gruvberger, N. Loman, O. Johannsson, H. Olsson, B. Wilfond, G. Sauter, O.-P. Kallioniemi, Å. Borg, and J. Trent (2001, feb). Gene-expression profiles in hereditary breast cancer. *New England Journal of Medicine 344*(8), 539–548.

Heesen, P. and A. Janssen (2016, jan). Dynamic adaptive multiple tests with finite sample FDR control. *Journal of Statistical Planning and Inference 168*, 38–51.

Hochberg, Y. and A. Tamhane (1987). *Multiple comparison procedures*. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley.

Hsu, J. (1996, feb). *Multiple Comparisons*. Chapman and Hall/CRC.

James, W. and C. Stein (1961). Estimation with quadratic loss. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, Volume 1, pp. 361–379.

Jin, J. and T. T. Cai (2007). Estimating the null and the proportion of nonnull effects in large-scale multiple comparisons. *Journal of the American Statistical Association 102*(478), 495–506.

Kendziorski, C. M., M. A. Newton, H. Lan, and M. N. Gould (2003). On parametric empirical Bayes methods for comparing multiple groups using replicated gene expression profiles. *Statistics in Medicine 22*(24), 3899–3914.

Kotz, S., N. Balakrishnan, and N. Johnson (2004). *Continuous Multivariate Distributions, Volume 1: Models and Applications*. Wiley Series in Probability and Statistics. John Wiley & Sons.

Langaas, M., B. H. Lindqvist, and E. Ferkingstad (2005). Estimating the proportion of true null hypotheses, with application to DNA microarray data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67*(4), 555–572.

Ledoit, O. and M. Wolf (2003, dec). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance 10*(5), 603–621.

Ledoit, O. and M. Wolf (2004a, jul). Honey, i shrunk the sample covariance matrix. *The Journal of Portfolio Management 30*(4), 110–119.

Ledoit, O. and M. Wolf (2004b, feb). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis 88*(2), 365–411.

Li, D., Z. Xie, M. Zand, T. Fogg, and T. Dye (2017, Jan). Bon-EV: an improved multiple testing procedure for controlling false discovery rates. *BMC Bioinformatics 18*(1), 1.

Liang, K. and D. Nettleton (2012). Adaptive and dynamic adaptive procedures for false discovery rate control and estimation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 74*(1), 163–182.

Lönnstedt, I. and T. Speed (2002). Replicated microarray data. *Statistica Sinica 12*(1), 31–46.

Meinshausen, N. and J. Rice (2006, 02). Estimating the proportion of false null hypotheses among a large number of independently tested hypotheses. *Annals of Statistics 34*(1), 373–393.

Morris, C. N. (1983, mar). Parametric empirical bayes inference: Theory and applications. *Journal of the American Statistical Association 78*(381), 47.

Newton, M. A., C. M. Kendziorski, C. S. Richmond, F. R. Blattner, and K. W. Tsui (2001). On differential variability of expression ratios: Improving statistical inference about gene expression changes from microarray data. *Journal of Computational Biology 8*(1), 37–52. PMID: 11339905.

Newton, M. A., A. Noueiry, D. Sarkar, and P. Ahlquist (2004). Detecting differential gene expression with a semiparametric hierarchical mixture method. *Biostatistics 5*(2), 155–176.

Opgen-Rhein, R. and K. Strimmer (2007). Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach. *Statistical applications in genetics and molecular biology 6*(1).

Pollard, K. S., M. D. Birkner, M. J. van der Laan, and S. Dudoit (2005). Test statistics null distributions in multiple testing: Simulation studies and applications to genomics.

Pollard, K. S., S. Dudoit, and M. J. van der Laan (2005). *Multiple Testing Procedures: R multtest Package and Applications to Genomics, in Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer.

Pollard, K. S. and M. J. van der Laan (2004). Choice of a null distribution in resampling-based multiple testing. *Journal of Statistical Planning and Inference 125*(1), 85 – 100. The Third International Conference on Multiple Comparisons.

Qiu, X., L. Klebanov, and A. Yakovlev (2005). Correlation between gene expression levels and limitations of the empirical bayes methodology for finding differentially expressed genes. *Statistical applications in genetics and molecular biology 4*(1).

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reiner, A., D. Yekutieli, and Y. Benjamini (2003, feb). Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics 19*(3), 368–375.

Romano, J. P., A. M. Shaikh, and M. Wolf (2008, Oct). Control of the false discovery rate under dependence using the bootstrap and subsampling. *TEST 17*(3), 417.

Sarkar, S. K. (2002). Some results on false discovery rate in stepwise multiple testing procedures. *Annals of statistics*, 239–257.

Schäfer, J., R. Opgen-Rhein, V. Zuber, M. Ahdesmaki, A. P. D. Silva, and K. Strimmer. (2017). *corpcor: Efficient Estimation of Covariance and (Partial) Correlation*. R package version 1.6.9.

Schäfer, J. and K. Strimmer (2005, jan). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology 4*(1).

Schweder, T. and E. Spjøtvoll (1982). Plots of p-values to evaluate many tests simultaneously. *Biometrika 69*(3), 493–502.

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

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology 3*(1), 1–25.

Somerville, P. N. (2004). FDR step-down and step-up procedures for the correlated case. In *Recent Developments in Multiple Comparison Procedures*, pp. 100–118. Institute of Mathematical Statistics.

Stodden, V. (2015). Reproducing statistical results. *Annual Review of Statistics and Its Application 2*(1), 1–19.

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(3), 479–498.

Storey, J. D. (2003, 12). The positive false discovery rate: a Bayesian interpretation and the q -value. *The Annals of Statistics 31*(6), 2013–2035.

Storey, J. D., J. E. Taylor, and D. Siegmund (2004). Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 66*(1), 187–205.

Troendle, J. F. (2000). Stepwise normal theory multiple test procedures controlling the false discovery rate. *Journal of Statistical Planning and Inference 84*(1-2), 139–158.

Tusher, V. G., R. Tibshirani, and G. Chu (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences 98*(9), 5116–5121.

van der Laan, M. J., S. Dudoit, and K. S. Pollard (2004, jan). Multiple testing. Part II. Step-down procedures for control of the family-wise error rate. *Statistical Applications in Genetics and Molecular Biology 3*(1), 1–33.

van der Laan, M. J. and A. E. Hubbard (2006). Quantile-function based null distribution in resampling based multiple testing. *Statistical Applications in Genetics and Molecular Biology 5*(1).

van der Vaart, A. W. and J. A. Wellner (1996). *Weak convergence and empirical processes : with applications to statistics*. Springer series in statistics. Springer.

Van't Wout, A. B., G. K. Lehrman, S. A. Mikheeva, G. C. O'Keeffe, M. G. Katze, R. E. Bumgarner, G. K. Geiss, and J. I. Mullins (2003). Cellular gene expression upon human immunodeficiency virus type 1 infection of CD+-t-cell lines. *Journal of Virology 77*(2), 1392–1402.

Warton, D. I. (2008, mar). Penalized normal likelihood and ridge regularization of correlation and covariance matrices. *Journal of the American Statistical Association 103*(481), 340–349.

Warton, D. I. (2010, may). Regularized sandwich estimators for analysis of high-dimensional data using generalized estimating equations. *Biometrics 67*(1), 116–123.

Westfall, P. H. and S. S. Young (1993, January). *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment (Wiley Series in Probability and Statistics)* (1 ed.). Wiley-Interscience.

Wright, G. W. and R. M. Simon (2003). A random variance model for detection of differential gene expression in small microarray experiments. *Bioinformatics 19*(18), 2448–2455.

Yekutieli, D. and Y. Benjamini (1999). Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical Planning and Inference 82*(1-2), 171–196.

Zeisel, A., O. Zuk, E. Domany, et al. (2011). FDR control with adaptive procedures and FDR monotonicity. *The Annals of Applied Statistics 5*(2A), 943–968.

APPENDICES

APPENDIX A

LIST OF ABBREVIATIONS AND NOTATIONS

BH – linear step-up procedure of Benjamini and Hochberg (1995)

BKY – two-stage adaptive step-up procedure of Benjamini, Krieger, and Yekutieli (2006)

BY – linear step-up procedure of Benjamini and Yekutieli (2001)

FDR – false discovery rate

FWER – family-wise error rate

PRD – positive regression dependency

PRDS – Positive regression dependency on subsets

STS – q-value procedure of Storey (2003); Storey, Taylor, and Siegmund (2004)

S-BH – BH procedure based on the shrinkage t statistic

SNS-BH – BH procedure based on the shrinkage t statistic and the null shift and scale-transformed test statistic null distribution

SNQ-BH – BH procedure based on the shrinkage t statistic and the null quantile-transformed test statistic null distribution

SNS-BKY – BKY procedure based on the shrinkage t statistic and the null shift and scale-transformed test statistic null distribution

SNQ-BKY – BKY procedure based on the shrinkage t statistic and the null quantile-transformed test statistic null distribution

$Q_0(P)$ – null distribution

$Q_0^{NS}(P)$ – null shift and scale-transformed null distribution

$Q_0^{NQ}(P)$ – null quantile-transformed null distribution

$Q_n(P)$ – true distribution of the test statistics

$Q_{0n}(P)$ – bootstrap estimate of the null distribution

$Q_{0n}^{NS}(P)$ – bootstrap estimate of the null shift and scale-transformed null distribution

$Q_{0n}^{NQ}(P)$ – bootstrap estimate of the null quantile-transformed null distribution

R_0 – number of hypotheses rejected under the null distribution

R_n – number of hypotheses rejected under the true distribution of the test statistics

$\mathcal{X}_n = (X_1, \ldots, X_n)$ – random sample of n independent and identically distributed random variables

V_0 – number of false discoveries under the null hypotheses

V_n – number of false discoveries under the true distribution of the test statistics

$\boldsymbol{\theta}_n = \big(\boldsymbol{\theta}_n(1),\ \boldsymbol{\theta}_n(2),\ \ldots,\ \boldsymbol{\theta}_n(m)\big)$ – m-dimensional vector of statistics

Ω – a statistical model

$\mathcal{H}_0 = \mathcal{H}_0(P)$ – the set of true null hypotheses

$\mathcal{H}_1 = \mathcal{H}_1(P)$ – the set of false null hypotheses

$\mathcal{C}_n(i;\alpha) = \mathcal{C}_n(T_n, Q_{0,i}, \alpha)$ – set of rejection regions

m – number of hypotheses

$m_0 =\mid \mathcal{H}_0 \mid$ – the number of true null hypotheses

$m_1 = m - m_0 =\mid \mathcal{H}_1 \mid$ – the number of false null hypotheses

n – total sample size

α – pre-specified significance level

APPENDIX B

SUPPLEMENTAL SIMULATION RESULTS FOR INDEPENDENT TESTS

B.1. Normally Distributed Random Variables

The numerical summaries for Monte Carlo simulation studies for the normal independent random variables are provided in this section.

Table B.1. Empirical FDRs for the investigated methods for the independent tests for $\pi_0 = 0.85$ and 0.9. The number of replications is 1,000 per scenario and the number of bootstrap resamples is 10,000. The standard errors are provided in parenthesis.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	20	0.044	0.049	0.006	0.044	0.009	0.119	0.137	0.119	0.137
		(0.157)	(0.161)	(0.072)	(0.157)	(0.076)	(0.154)	(0.135)	(0.154)	(0.135)
	30	0.042	0.048	0.009	0.043	0.030	0.080	0.087	0.080	0.087
		(0.086)	(0.084)	(0.078)	(0.086)	(0.077)	(0.076)	(0.073)	(0.076)	(0.073)
	40	0.047	0.051	0.007	0.047	0.037	0.068	0.071	0.068	0.071
		(0.051)	(0.053)	(0.042)	(0.051)	(0.045)	(0.051)	(0.049)	(0.051)	(0.050)
	50	0.045	0.051	0.006	0.046	0.040	0.061	0.064	0.062	0.065
		(0.040)	(0.041)	(0.024)	(0.040)	(0.038)	(0.042)	(0.042)	(0.043)	(0.042)
	60	0.046	0.051	0.005	0.048	0.042	0.058	0.060	0.060	0.062
		(0.033)	(0.035)	(0.017)	(0.034)	(0.031)	(0.035)	(0.036)	(0.035)	(0.036)
	100	0.045	0.051	0.006	0.048	0.044	0.053	0.053	0.056	0.056
		(0.026)	(0.027)	(0.012)	(0.027)	(0.026)	(0.027)	(0.027)	(0.028)	(0.028)
	200	0.046	0.051	0.006	0.049	0.045	0.049	0.048	0.053	0.053
		(0.023)	(0.024)	(0.009)	(0.024)	(0.022)	(0.023)	(0.023)	(0.024)	(0.024)
	300	0.046	0.051	0.006	0.051	0.045	0.049	0.048	0.053	0.053
		(0.022)	(0.023)	(0.008)	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)	(0.023)
	500	0.044	0.050	0.006	0.049	0.044	0.046	0.046	0.051	0.051
		(0.021)	(0.022)	(0.008)	(0.022)	(0.021)	(0.021)	(0.021)	(0.023)	(0.022)
0.85	20	0.037	0.044	0.009	0.037	0.015	0.109	0.125	0.109	0.125
		(0.129)	(0.127)	(0.088)	(0.128)	(0.100)	(0.123)	(0.114)	(0.123)	(0.114)
	30	0.042	0.048	0.007	0.043	0.032	0.074	0.083	0.074	0.083
		(0.064)	(0.062)	(0.056)	(0.064)	(0.065)	(0.060)	(0.059)	(0.060)	(0.059)
	40	0.041	0.050	0.007	0.043	0.037	0.062	0.067	0.064	0.068
		(0.039)	(0.041)	(0.029)	(0.040)	(0.039)	(0.042)	(0.042)	(0.043)	(0.043)
	50	0.042	0.050	0.005	0.044	0.040	0.057	0.059	0.059	0.061
		(0.031)	(0.033)	(0.018)	(0.032)	(0.031)	(0.034)	(0.034)	(0.034)	(0.034)
	60	0.043	0.050	0.006	0.045	0.042	0.056	0.058	0.059	0.060
		(0.028)	(0.030)	(0.015)	(0.029)	(0.028)	(0.031)	(0.031)	(0.032)	(0.031)
	100	0.042	0.050	0.005	0.046	0.042	0.050	0.049	0.054	0.053
		(0.022)	(0.024)	(0.010)	(0.023)	(0.021)	(0.023)	(0.023)	(0.024)	(0.024)
	200	0.043	0.051	0.006	0.049	0.043	0.047	0.046	0.053	0.052
		(0.019)	(0.021)	(0.008)	(0.020)	(0.019)	(0.020)	(0.020)	(0.021)	(0.020)
	300	0.042	0.051	0.006	0.049	0.042	0.045	0.044	0.052	0.051
		(0.018)	(0.021)	(0.007)	(0.019)	(0.017)	(0.018)	(0.018)	(0.019)	(0.019)
	500	0.043	0.051	0.006	0.050	0.042	0.044	0.044	0.051	0.051
		(0.017)	(0.019)	(0.007)	(0.018)	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)

Table B.2. Empirical FDRs for the investigated methods for the independent tests for $\pi_0 = 0.75$ and 0.8. The number of replications is 1,000 per scenario and the number of bootstrap resamples is 10,000. The standard errors are provided in parenthesis. Cases where FDR control was anti-conservative are indicated with a star ($\star$).

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	20	0.036	0.047	0.005	0.036	0.014	0.083	0.102	0.083	0.102
		(0.086)	(0.092)	(0.066)	(0.086)	(0.078)	(0.077)	(0.067)	(0.077)	(0.067)
	30	0.039	0.049	0.007	0.040	0.033	0.063	0.069	0.065	0.070
		(0.040)	(0.040)	(0.055)	(0.040)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
	40	0.038	0.048	0.004	0.040	0.035	0.052	0.055	0.055	0.057
		(0.027)	(0.030)	(0.017)	(0.028)	(0.027)	(0.030)	(0.030)	(0.030)	(0.030)
	50	0.040	0.051	0.005	0.043	0.038	0.053	0.053	0.056	0.056
		(0.024)	(0.026)	(0.013)	(0.024)	(0.024)	(0.026)	(0.026)	(0.026)	(0.026)
	60	0.041	0.051	0.005	0.044	0.038	0.050	0.049	0.054	0.053
		(0.021)	(0.024)	(0.011)	(0.022)	(0.021)	(0.022)	(0.023)	(0.023)	(0.023)
	100	0.041	0.052	0.005	0.047	0.040	0.046	0.045	0.053	0.051
		(0.018)	(0.020)	(0.008)	(0.019)	(0.017)	(0.018)	(0.018)	(0.020)	(0.019)
	200	0.040	0.050	0.006	0.048	0.040	0.043	0.042	0.051	0.050
		(0.015)	(0.018)	(0.006)	(0.016)	(0.015)	(0.015)	(0.015)	(0.017)	(0.017)
	300	0.040	0.050	0.005	0.049	0.039	0.041	0.041	0.051	0.050
		(0.014)	(0.017)	(0.006)	(0.016)	(0.014)	(0.015)	(0.015)	(0.016)	(0.016)
	500	0.041	0.052	0.006	0.051	0.041	0.042	0.042	0.053	0.052
		(0.015)	(0.017)	(0.006)	(0.016)	(0.014)	(0.015)	(0.014)	(0.016)	(0.016)
0.75	20	0.034	0.045	0.004	0.034	0.011	0.071	0.081	0.071	0.081
		(0.078)	(0.069)	(0.045)	(0.078)	(0.079)	(0.060)	(0.060)	(0.061)	(0.060)
	30	0.035	0.047	0.004	0.036	0.026	0.054	0.056	0.055	0.058
		(0.032)	(0.034)	(0.024)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
	40	0.037	0.049	0.005	0.039	0.032	0.050	0.049	0.052	0.052
		(0.024)	(0.026)	(0.016)	(0.024)	(0.024)	(0.025)	(0.025)	(0.025)	(0.026)
	50	0.036	0.050	0.005	0.040	0.033	0.046	0.045	0.050	0.049
		(0.020)	(0.023)	(0.012)	(0.021)	(0.020)	(0.021)	(0.022)	(0.022)	(0.022)
	60	0.038	0.051	0.005	0.042	0.034	0.046	0.044	0.051	0.048
		(0.019)	(0.022)	(0.009)	(0.019)	(0.017)	(0.019)	(0.019)	(0.020)	(0.020)
	100	0.037	0.050	0.004	0.044	0.034	0.041	0.039	0.049	0.046
		(0.016)	(0.018)	(0.006)	(0.017)	(0.015)	(0.016)	(0.016)	(0.017)	(0.017)
	200	0.038	0.050	0.005	0.047	0.036	0.040	0.038	0.050	0.047
		(0.013)	(0.016)	(0.005)	(0.015)	(0.013)	(0.013)	(0.013)	(0.015)	(0.014)
	300	0.038	0.051	0.005	0.049	0.037	0.040	0.038	0.051	0.049
		(0.013)	(0.015)	(0.005)	(0.014)	(0.012)	(0.013)	(0.013)	(0.014)	(0.014)
	500	0.037	0.051	0.005	0.050	0.036	0.039	0.037	0.051	0.050
		(0.012)	(0.015)	(0.005)	(0.014)	(0.012)	(0.012)	(0.012)	(0.014)	(0.014)

Table B.3. Empirical false non-discovery rates for the investigated methods for the normal independent tests. The number of replications is 1,000 per scenario and the number of bootstrap resmaples is 10,000. The standard errors of the estimated false non-discovery rate is of the order of 0.006 or less for all the methods.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	20	0.099	0.099	0.100	0.099	0.100	0.095	0.093	0.095	0.093
	30	0.094	0.093	0.099	0.094	0.095	0.088	0.086	0.088	0.086
	40	0.085	0.083	0.096	0.084	0.084	0.078	0.076	0.078	0.076
	50	0.074	0.073	0.091	0.074	0.073	0.069	0.068	0.069	0.067
	60	0.065	0.064	0.084	0.065	0.064	0.061	0.060	0.060	0.059
	100	0.042	0.041	0.060	0.041	0.041	0.040	0.039	0.039	0.039
	200	0.019	0.019	0.031	0.019	0.019	0.019	0.019	0.018	0.018
	300	0.011	0.010	0.019	0.010	0.011	0.010	0.010	0.010	0.010
	500	0.004	0.003	0.008	0.004	0.004	0.004	0.004	0.003	0.003
0.85	20	0.148	0.148	0.150	0.148	0.150	0.144	0.142	0.144	0.142
	30	0.140	0.138	0.148	0.140	0.143	0.134	0.132	0.134	0.132
	40	0.127	0.125	0.144	0.127	0.129	0.122	0.120	0.121	0.120
	50	0.115	0.112	0.137	0.114	0.116	0.110	0.109	0.110	0.109
	60	0.105	0.102	0.129	0.104	0.105	0.100	0.100	0.099	0.099
	100	0.075	0.072	0.100	0.073	0.074	0.072	0.072	0.071	0.071
	200	0.040	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
	300	0.024	0.023	0.042	0.023	0.024	0.024	0.024	0.022	0.022
	500	0.010	0.009	0.020	0.009	0.010	0.010	0.010	0.009	0.009
0.8	20	0.196	0.194	0.200	0.196	0.199	0.188	0.185	0.188	0.185
	30	0.179	0.175	0.196	0.179	0.183	0.171	0.169	0.170	0.168
	40	0.160	0.155	0.187	0.159	0.162	0.153	0.152	0.152	0.151
	50	0.143	0.138	0.175	0.141	0.144	0.137	0.136	0.135	0.135
	60	0.128	0.123	0.163	0.126	0.129	0.123	0.123	0.121	0.121
	100	0.090	0.085	0.124	0.087	0.089	0.087	0.087	0.084	0.084
	200	0.047	0.044	0.075	0.045	0.047	0.046	0.046	0.043	0.043
	300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.026	0.026
	500	0.012	0.011	0.025	0.011	0.012	0.012	0.012	0.010	0.010
0.75	20	0.245	0.242	0.250	0.245	0.249	0.236	0.234	0.236	0.234
	30	0.224	0.218	0.245	0.224	0.232	0.215	0.215	0.214	0.214
	40	0.201	0.193	0.235	0.200	0.206	0.193	0.194	0.191	0.192
	50	0.181	0.172	0.222	0.178	0.184	0.173	0.175	0.171	0.173
	60	0.163	0.154	0.208	0.160	0.165	0.157	0.158	0.153	0.155
	100	0.115	0.107	0.160	0.110	0.116	0.111	0.113	0.107	0.108
	200	0.060	0.055	0.096	0.056	0.060	0.059	0.060	0.055	0.055
	300	0.037	0.034	0.064	0.034	0.037	0.037	0.037	0.033	0.033
	500	0.018	0.015	0.035	0.016	0.018	0.017	0.017	0.015	0.015

Table B.4. Average number of false hypotheses rejected for the investigated methods for the independent tests for $\pi_0 = 0.85$ and 0.9. The number of replications is 1,000 per scenario and the number of bootstrap resamples is 10,000. The standard errors are provided in parenthesis. Cases where FDR control was anti-conservative are indicated with a star ($\star$).

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	20	1.240	1.485	0.150	1.244	0.555	5.253*	7.727*	5.253*	7.727*
		(1.660)	(1.897)	(0.456)	(1.665)	(1.106)	(3.318)	(3.843)	(3.318)	(3.843)
	30	6.703	7.552	0.923	6.746	6.019	13.054*	15.782*	13.054*	15.787*
		(4.118)	(4.356)	(1.273)	(4.156)	(4.219)	(4.890)	(4.929)	(4.890)	(4.943)
	40	16.808	18.135	3.949	17.013	17.540	23.503*	25.641*	23.555*	25.724*
		(5.487)	(5.683)	(2.872)	(5.557)	(5.744)	(5.567)	(5.378)	(5.641)	(5.487)
	50	27.806	29.194	9.855	28.148	29.079	33.320	34.935*	33.543	35.216*
		(5.357)	(5.466)	(4.100)	(5.448)	(5.449)	(5.329)	(5.102)	(5.450)	(5.210)
	60	37.287	38.786	17.294	37.815	38.700	41.910	43.045	42.317	43.383
		(5.337)	(5.524)	(4.614)	(5.433)	(5.440)	(5.341)	(5.309)	(5.411)	(5.345)
	100	60.868	62.056	42.680	61.554	61.805	63.045	63.520	63.618	64.086
		(4.281)	(4.334)	(4.494)	(4.307)	(4.182)	(4.133)	(4.096)	(4.195)	(4.146)
	200	82.350	83.041	71.616	82.816	82.593	82.939	82.994	83.440	83.427
		(3.054)	(3.106)	(3.129)	(3.070)	(3.034)	(2.988)	(2.995)	(2.994)	(3.012)
	300	90.216	90.674	82.587	90.605	90.310	90.511	90.564	90.883	90.881
		(2.476)	(2.522)	(2.880)	(2.487)	(2.486)	(2.456)	(2.468)	(2.456)	(2.469)
	500	96.658	96.867	92.442	96.845	96.699	96.727	96.743	96.919	96.927
		(1.631)	(1.580)	(2.203)	(1.588)	(1.620)	(1.598)	(1.594)	(1.571)	(1.571)
0.85	20	2.182	2.701	0.276	2.191	0.575	7.007*	9.238*	7.007*	9.238*
		(2.397)	(2.804)	(0.647)	(2.412)	(1.152)	(3.848)	(4.273)	(3.848)	(4.273)
	30	11.770	13.658	1.786	11.876	8.135	18.502*	20.513*	18.518*	20.559*
		(5.340)	(5.757)	(1.939)	(5.419)	(5.274)	(5.817)	(5.890)	(5.850)	(5.976)
	40	26.090	28.807	7.014	26.511	23.838	32.711	34.012*	33.078	34.348*
		(6.044)	(6.482)	(3.775)	(6.155)	(6.416)	(6.202)	(6.190)	(6.430)	(6.357)
	50	39.728	42.691	15.226	40.449	38.755	45.147	46.007	45.782	46.627
		(5.932)	(6.396)	(4.561)	(6.108)	(6.289)	(6.173)	(6.118)	(6.357)	(6.238)
	60	51.011	53.887	24.477	51.874	50.483	55.606	56.074	56.440	56.909
		(6.096)	(6.502)	(5.209)	(6.162)	(6.113)	(6.056)	(5.954)	(6.173)	(6.047)
	100	81.548	84.244	56.074	82.899	82.028	84.234	84.496	85.544	85.743
		(5.402)	(5.585)	(5.038)	(5.461)	(5.359)	(5.313)	(5.301)	(5.421)	(5.384)
	200	114.555	116.342	94.315	115.937	114.935	115.592	115.715	116.985	117.008
		(4.445)	(4.528)	(4.339)	(4.489)	(4.459)	(4.450)	(4.437)	(4.484)	(4.459)
	300	129.007	130.350	112.993	130.142	129.243	129.569	129.620	130.656	130.704
		(3.603)	(3.569)	(4.140)	(3.544)	(3.608)	(3.543)	(3.551)	(3.513)	(3.507)
	500	141.408	142.102	132.288	142.040	141.552	141.622	141.610	142.211	142.205
		(2.477)	(2.463)	(3.201)	(2.467)	(2.468)	(2.471)	(2.480)	(2.439)	(2.441)

Table B.5. Average number of false hypotheses rejected for the investigated methods for the independent tests for $\pi_0 = 0.75$ and 0.8. The number of replications is 1,000 per scenario and the number of bootstrap resamples is 10,000. The standard errors are provided in parenthesis. Cases where FDR control was anti-conservative are indicated with a star ($\star$).

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	20	5.388	7.281	0.463	5.417	1.725	14.553*	18.688*	14.561*	18.728*
		(4.202)	(5.135)	(0.853)	(4.232)	(2.633)	(6.133)	(6.203)	(6.155)	(6.279)
	30	25.230	30.192	4.531	25.757	21.143	35.525	38.072*	36.061*	38.628*
		(7.416)	(8.021)	(3.378)	(7.581)	(7.893)	(7.666)	(7.329)	(7.942)	(7.531)
	40	47.957	53.540	15.417	49.134	45.453	56.182	57.597	57.379	58.736
		(7.489)	(8.019)	(5.142)	(7.714)	(7.726)	(7.398)	(7.145)	(7.618)	(7.277)
	50	66.969	72.742	29.768	68.810	65.895	73.741	74.437	75.444	76.225
		(6.966)	(7.397)	(6.027)	(7.189)	(7.039)	(7.080)	(7.031)	(7.312)	(7.288)
	60	82.596	88.211	43.796	84.787	82.063	88.116	88.389	90.255	90.525
		(6.909)	(7.692)	(6.041)	(7.098)	(7.109)	(7.102)	(7.082)	(7.366)	(7.356)
	100	121.731	126.343	86.655	124.387	121.862	124.571	124.537	127.228	127.113
		(6.091)	(6.403)	(5.697)	(6.176)	(6.141)	(6.062)	(6.049)	(6.175)	(6.223)
	200	160.684	163.676	135.430	163.036	161.052	161.813	161.785	164.245	164.135
		(4.677)	(4.905)	(4.878)	(4.710)	(4.730)	(4.659)	(4.677)	(4.644)	(4.650)
	300	176.836	178.966	158.023	178.737	177.116	177.477	177.450	179.271	179.243
		(3.685)	(3.663)	(4.366)	(3.577)	(3.687)	(3.642)	(3.637)	(3.560)	(3.572)
	500	190.454	191.589	179.359	191.530	190.594	190.754	190.696	191.743	191.740
		(2.724)	(2.673)	(3.347)	(2.620)	(2.704)	(2.709)	(2.691)	(2.594)	(2.599)
0.75	20	6.965	10.698	0.642	7.008	1.453	18.680*	21.000*	18.732*	21.096*
		(5.119)	(6.728)	(1.059)	(5.158)	(2.475)	(7.164)	(7.062)	(7.263)	(7.218)
	30	33.524	41.601	6.162	34.341	23.885	45.819	45.332	46.901	46.333
		(8.214)	(9.014)	(3.889)	(8.445)	(9.015)	(8.554)	(8.311)	(8.967)	(8.642)
	40	61.344	71.099	19.549	63.427	55.572	71.986	70.158	74.093	72.290
		(8.422)	(9.414)	(5.864)	(8.757)	(9.069)	(8.440)	(8.401)	(8.824)	(8.760)
	50	84.840	95.004	36.245	87.869	81.260	93.500	91.526	96.670	94.390
		(8.517)	(9.554)	(6.650)	(8.900)	(8.647)	(8.253)	(8.234)	(8.768)	(8.615)
	60	104.658	114.310	53.356	108.366	102.018	111.754	109.751	115.533	113.301
		(8.008)	(8.821)	(6.998)	(8.339)	(8.061)	(7.958)	(7.813)	(8.262)	(8.080)
	100	153.187	161.032	107.358	157.756	152.385	156.835	155.323	161.288	159.734
		(6.992)	(7.624)	(6.501)	(7.200)	(7.011)	(7.003)	(7.036)	(7.121)	(7.160)
	200	202.261	206.948	170.472	206.096	202.203	203.504	203.009	207.241	206.739
		(5.130)	(5.299)	(5.668)	(5.104)	(5.163)	(5.138)	(5.142)	(5.043)	(5.106)
	300	221.128	224.398	198.401	224.025	221.243	221.750	221.564	224.604	224.463
		(4.067)	(4.173)	(4.657)	(4.005)	(4.016)	(4.057)	(4.005)	(4.005)	(3.957)
	500	236.653	238.428	223.213	238.364	236.778	236.917	236.910	238.612	238.561
		(2.945)	(2.888)	(3.477)	(2.882)	(2.906)	(2.926)	(2.934)	(2.835)	(2.858)

B.2. Gamma Distributed Random Variables

Table B.6. Empirical FDRs for the investigated methods for the independent tests for $\pi_0 = 0.85$ and 0.9 for the gamma variates. The number of replications is 1,000 per scenario and the number of bootstrap resamples is 10,000. The standard errors are provided in parenthesis.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.90	20	0.009	0.012	0.002	0.009	0.015	0.097	0.059	0.101	0.062
		(0.047)	(0.051)	(0.039)	(0.047)	(0.021)	(0.042)	(0.034)	(0.042)	(0.035)
	30	0.018	0.022	0.001	0.019	0.022	0.085	0.048	0.090	0.051
		(0.022)	(0.024)	(0.011)	(0.022)	(0.019)	(0.034)	(0.026)	(0.035)	(0.027)
	40	0.025	0.030	0.001	0.027	0.025	0.074	0.042	0.079	0.045
		(0.021)	(0.022)	(0.007)	(0.021)	(0.019)	(0.031)	(0.024)	(0.032)	(0.025)
	50	0.030	0.035	0.002	0.033	0.027	0.067	0.040	0.072	0.043
		(0.020)	(0.022)	(0.007)	(0.021)	(0.018)	(0.028)	(0.022)	(0.029)	(0.023)
	60	0.033	0.038	0.003	0.036	0.028	0.063	0.038	0.068	0.042
		(0.020)	(0.022)	(0.007)	(0.021)	(0.018)	(0.026)	(0.020)	(0.027)	(0.021)
	100	0.038	0.044	0.004	0.043	0.027	0.053	0.033	0.059	0.037
		(0.021)	(0.023)	(0.007)	(0.022)	(0.018)	(0.024)	(0.019)	(0.025)	(0.021)
	200	0.042	0.047	0.005	0.047	0.029	0.048	0.032	0.053	0.036
		(0.020)	(0.022)	(0.007)	(0.021)	(0.016)	(0.021)	(0.018)	(0.022)	(0.018)
	300	0.043	0.049	0.005	0.048	0.030	0.046	0.032	0.051	0.036
		(0.021)	(0.022)	(0.007)	(0.022)	(0.017)	(0.021)	(0.018)	(0.022)	(0.019)
	500	0.044	0.050	0.006	0.050	0.032	0.047	0.033	0.052	0.038
		(0.020)	(0.022)	(0.008)	(0.021)	(0.018)	(0.021)	(0.018)	(0.022)	(0.019)
0.85	20	0.010	0.014	0.001	0.010	0.014	0.083	0.050	0.089	0.053
		(0.029)	(0.031)	(0.016)	(0.029)	(0.017)	(0.034)	(0.027)	(0.035)	(0.028)
	30	0.021	0.027	0.001	0.023	0.023	0.074	0.043	0.080	0.047
		(0.019)	(0.021)	(0.008)	(0.019)	(0.016)	(0.027)	(0.022)	(0.028)	(0.023)
	40	0.028	0.035	0.001	0.031	0.025	0.064	0.038	0.071	0.043
		(0.017)	(0.020)	(0.006)	(0.018)	(0.015)	(0.023)	(0.018)	(0.024)	(0.019)
	50	0.031	0.039	0.002	0.036	0.025	0.057	0.035	0.064	0.040
		(0.017)	(0.020)	(0.006)	(0.018)	(0.015)	(0.021)	(0.017)	(0.023)	(0.018)
	60	0.033	0.040	0.003	0.038	0.025	0.054	0.033	0.062	0.039
		(0.017)	(0.019)	(0.006)	(0.018)	(0.014)	(0.020)	(0.016)	(0.021)	(0.018)
	100	0.037	0.045	0.004	0.044	0.026	0.047	0.031	0.055	0.037
		(0.016)	(0.018)	(0.005)	(0.017)	(0.014)	(0.018)	(0.015)	(0.019)	(0.016)
	200	0.040	0.048	0.005	0.048	0.028	0.044	0.031	0.052	0.037
		(0.016)	(0.019)	(0.006)	(0.018)	(0.014)	(0.017)	(0.014)	(0.018)	(0.016)
	300	0.040	0.049	0.005	0.049	0.029	0.043	0.031	0.051	0.036
		(0.016)	(0.019)	(0.006)	(0.018)	(0.014)	(0.016)	(0.014)	(0.018)	(0.016)
	500	0.041	0.049	0.005	0.049	0.031	0.043	0.032	0.051	0.039
		(0.016)	(0.018)	(0.006)	(0.018)	(0.014)	(0.016)	(0.014)	(0.018)	(0.016)

Table B.7. Empirical FDRs for the investigated methods for the independent tests for $\pi_0 = 0.8$ and 0.75 for the gamma variates. The number of replications is 1,000 per scenario and the number of bootstrap resamples is 10,000. The standard errors are provided in parenthesis.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.80	20	0.012	0.018	0.001	0.012	0.014	0.070	0.043	0.078	0.048
		(0.027)	(0.028)	(0.032)	(0.027)	(0.014)	(0.026)	(0.020)	(0.028)	(0.021)
	30	0.022	0.031	0.001	0.025	0.021	0.060	0.036	0.068	0.041
		(0.016)	(0.018)	(0.006)	(0.017)	(0.013)	(0.020)	(0.016)	(0.022)	(0.017)
	40	0.027	0.036	0.002	0.032	0.023	0.054	0.033	0.063	0.040
		(0.015)	(0.018)	(0.005)	(0.016)	(0.012)	(0.018)	(0.015)	(0.019)	(0.016)
	50	0.031	0.041	0.003	0.038	0.023	0.050	0.031	0.060	0.039
		(0.015)	(0.017)	(0.005)	(0.016)	(0.012)	(0.017)	(0.014)	(0.019)	(0.015)
	60	0.033	0.044	0.003	0.041	0.024	0.048	0.031	0.058	0.038
		(0.014)	(0.017)	(0.005)	(0.015)	(0.012)	(0.017)	(0.013)	(0.018)	(0.015)
	100	0.037	0.047	0.004	0.046	0.025	0.043	0.029	0.053	0.037
		(0.014)	(0.016)	(0.005)	(0.015)	(0.011)	(0.014)	(0.012)	(0.016)	(0.014)
	200	0.038	0.049	0.005	0.049	0.027	0.041	0.029	0.051	0.038
		(0.014)	(0.016)	(0.005)	(0.016)	(0.012)	(0.014)	(0.012)	(0.016)	(0.014)
	300	0.039	0.049	0.005	0.049	0.028	0.040	0.030	0.050	0.038
		(0.014)	(0.017)	(0.005)	(0.016)	(0.012)	(0.014)	(0.012)	(0.016)	(0.014)
	500	0.039	0.050	0.005	0.050	0.030	0.040	0.031	0.051	0.040
		(0.014)	(0.016)	(0.005)	(0.015)	(0.012)	(0.014)	(0.012)	(0.015)	(0.014)
0.75	20	0.014	0.023	0.000	0.015	0.015	0.061	0.037	0.070	0.043
		(0.018)	(0.019)	(0.011)	(0.018)	(0.012)	(0.020)	(0.017)	(0.022)	(0.018)
	30	0.022	0.034	0.001	0.027	0.019	0.052	0.032	0.062	0.039
		(0.013)	(0.016)	(0.005)	(0.014)	(0.011)	(0.017)	(0.013)	(0.019)	(0.015)
	40	0.028	0.040	0.002	0.035	0.021	0.048	0.030	0.059	0.038
		(0.012)	(0.015)	(0.004)	(0.014)	(0.010)	(0.015)	(0.012)	(0.017)	(0.014)
	50	0.031	0.043	0.002	0.039	0.022	0.045	0.029	0.056	0.037
		(0.013)	(0.016)	(0.004)	(0.014)	(0.010)	(0.014)	(0.012)	(0.016)	(0.013)
	60	0.033	0.045	0.003	0.042	0.022	0.043	0.028	0.055	0.037
		(0.013)	(0.015)	(0.004)	(0.014)	(0.011)	(0.014)	(0.012)	(0.016)	(0.013)
	100	0.035	0.047	0.004	0.046	0.024	0.039	0.027	0.051	0.036
		(0.012)	(0.015)	(0.004)	(0.013)	(0.010)	(0.012)	(0.010)	(0.014)	(0.012)
	200	0.036	0.050	0.004	0.050	0.026	0.038	0.027	0.051	0.038
		(0.012)	(0.015)	(0.004)	(0.014)	(0.010)	(0.012)	(0.011)	(0.014)	(0.013)
	300	0.037	0.050	0.005	0.050	0.027	0.038	0.028	0.050	0.040
		(0.011)	(0.014)	(0.004)	(0.013)	(0.010)	(0.012)	(0.010)	(0.014)	(0.012)
	500	0.037	0.050	0.005	0.050	0.029	0.037	0.029	0.050	0.041
		(0.012)	(0.015)	(0.004)	(0.014)	(0.011)	(0.012)	(0.011)	(0.014)	(0.013)

Table B.8. Empirical false non-discovery rates for the investigated methods for the gamma independent variates. The number of replications is 1,000 per scenario and the number of bootstrap resmaples is 10,000. The standard errors of the estimated false non-discovery rate is of the order of 0.008 or less for all the methods.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.90	20	0.095	0.093	0.100	0.094	0.068	0.053	0.053	0.053	0.053
	30	0.066	0.064	0.094	0.065	0.044	0.036	0.037	0.035	0.036
	40	0.045	0.043	0.074	0.044	0.031	0.026	0.027	0.025	0.026
	50	0.032	0.031	0.056	0.031	0.022	0.019	0.020	0.018	0.019
	60	0.024	0.023	0.044	0.023	0.017	0.015	0.016	0.014	0.015
	100	0.010	0.009	0.020	0.009	0.007	0.006	0.007	0.006	0.006
	200	0.001	0.001	0.004	0.001	0.001	0.001	0.001	0.001	0.001
	300	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.85	20	0.140	0.137	0.149	0.140	0.105	0.085	0.086	0.083	0.084
	30	0.099	0.094	0.140	0.097	0.072	0.060	0.062	0.058	0.061
	40	0.071	0.067	0.113	0.068	0.052	0.044	0.047	0.042	0.045
	50	0.053	0.050	0.090	0.051	0.039	0.034	0.036	0.032	0.034
	60	0.041	0.038	0.073	0.039	0.030	0.026	0.028	0.025	0.027
	100	0.016	0.015	0.035	0.015	0.012	0.011	0.011	0.010	0.011
	200	0.002	0.002	0.007	0.002	0.001	0.001	0.001	0.001	0.001
	300	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.80	20	0.180	0.171	0.199	0.179	0.136	0.111	0.113	0.108	0.111
	30	0.124	0.115	0.182	0.121	0.094	0.079	0.083	0.076	0.079
	40	0.090	0.083	0.145	0.085	0.069	0.059	0.062	0.056	0.059
	50	0.067	0.062	0.117	0.063	0.052	0.045	0.048	0.042	0.045
	60	0.052	0.047	0.096	0.048	0.040	0.035	0.037	0.032	0.034
	100	0.020	0.018	0.046	0.018	0.015	0.013	0.014	0.012	0.013
	200	0.002	0.002	0.008	0.002	0.002	0.001	0.002	0.001	0.001
	300	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.75	20	0.214	0.197	0.249	0.212	0.163	0.134	0.137	0.129	0.133
	30	0.145	0.131	0.221	0.139	0.113	0.097	0.101	0.092	0.096
	40	0.105	0.094	0.174	0.098	0.083	0.072	0.076	0.067	0.071
	50	0.080	0.071	0.139	0.073	0.064	0.056	0.059	0.052	0.055
	60	0.062	0.055	0.114	0.056	0.050	0.044	0.047	0.040	0.042
	100	0.026	0.022	0.057	0.022	0.020	0.018	0.019	0.016	0.017
	200	0.004	0.003	0.013	0.003	0.003	0.002	0.003	0.002	0.002
	300	0.001	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000
	500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table B.9. Average number of false hypotheses rejected for the investigated methods for the gamma independent variates for $\pi_0 = 0.9$ and 0.85. The number of replications is 1,000 per scenario and the number of bootstrap resamples is 10,000. The standard errors are provided in parenthesis. Cases where FDR control was anti-conservative are indicated with a star ($\star$).

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.90	20	6.025	7.536	0.418	6.088	34.021	49.516*	49.682	50.206*	50.172
		(4.821)	(5.533)	(0.849)	(4.885)	(6.243)	(5.692)	(5.346)	(5.741)	(5.463)
	30	36.660	38.988	6.972	37.371	58.868	66.885*	65.629	67.555*	66.163
		(6.489)	(6.648)	(4.499)	(6.640)	(4.867)	(4.449)	(4.539)	(4.529)	(4.549)
	40	57.539	59.238	27.641	58.435	71.479	76.413*	75.158	77.012*	75.688
		(5.227)	(5.286)	(5.900)	(5.307)	(4.013)	(3.797)	(3.785)	(3.792)	(3.800)
	50	70.187	71.395	46.142	71.030	79.464	82.748*	81.764	83.238*	82.261
		(4.534)	(4.569)	(5.159)	(4.530)	(3.566)	(3.491)	(3.514)	(3.491)	(3.525)
	60	77.794	78.758	58.578	78.503	84.414	86.688	85.862	87.155*	86.309
		(3.710)	(3.779)	(4.644)	(3.754)	(3.194)	(2.987)	(3.110)	(2.999)	(3.079)
	100	91.319	91.784	82.047	91.733	93.672	94.513	94.067	94.770	94.341
		(2.493)	(2.464)	(2.960)	(2.479)	(2.202)	(2.101)	(2.156)	(2.082)	(2.103)
	200	98.663	98.761	96.358	98.759	99.066	99.185	99.107	99.235	99.169
		(1.085)	(1.054)	(1.688)	(1.048)	(0.900)	(0.851)	(0.891)	(0.841)	(0.875)
	300	99.756	99.778	99.103	99.783	99.834	99.853	99.843	99.864	99.850
		(0.505)	(0.485)	(0.917)	(0.484)	(0.430)	(0.402)	(0.420)	(0.389)	(0.412)
	500	99.993	99.993	99.949	99.993	99.993	99.994	99.995	99.995	99.996
		(0.083)	(0.083)	(0.220)	(0.083)	(0.083)	(0.077)	(0.071)	(0.071)	(0.063)
0.85	20	11.203	15.327	0.592	11.465	50.032	71.990*	70.761	73.583*	71.978
		(7.241)	(8.894)	(1.080)	(7.510)	(7.765)	(6.951)	(6.654)	(7.161)	(6.727)
	30	56.989	62.022	11.727	58.643	84.435	96.449*	93.905	98.080*	95.454
		(8.058)	(8.514)	(6.199)	(8.278)	(6.227)	(5.887)	(5.854)	(5.924)	(5.999)
	40	85.661	89.406	41.942	87.698	103.439	111.092*	108.712	112.649*	110.214
		(6.311)	(6.577)	(7.072)	(6.327)	(5.350)	(4.947)	(5.007)	(4.989)	(4.948)
	50	102.707	105.802	65.558	104.741	115.273	120.568	118.687	121.925*	120.052
		(5.784)	(5.844)	(5.999)	(5.841)	(4.914)	(4.741)	(4.778)	(4.735)	(4.753)
	60	114.187	116.496	83.061	116.071	123.420	127.176	125.655	128.360	126.936
		(4.861)	(4.985)	(5.473)	(4.826)	(4.250)	(4.074)	(4.076)	(3.987)	(4.094)
	100	135.916	137.070	119.077	137.003	139.598	140.899	140.259	141.557	140.939
		(3.285)	(3.226)	(4.340)	(3.224)	(2.859)	(2.738)	(2.782)	(2.671)	(2.769)
	200	148.239	148.427	144.251	148.457	148.796	148.950	148.830	149.068	148.985
		(1.347)	(1.271)	(2.197)	(1.269)	(1.103)	(1.032)	(1.085)	(0.969)	(1.015)
	300	149.748	149.791	148.836	149.792	149.852	149.871	149.867	149.891	149.886
		(0.474)	(0.438)	(1.041)	(0.437)	(0.382)	(0.353)	(0.360)	(0.327)	(0.333)
	500	149.997	149.998	149.960	149.997	149.999	149.999	149.998	149.999	149.999
		(0.055)	(0.045)	(0.196)	(0.055)	(0.032)	(0.032)	(0.045)	(0.032)	(0.032)

Table B.10. Average number of false hypotheses rejected for the investigated methods for the gamma independent variates for $\pi_0 = 0.75$ and 0.8. The number of replications is 1,000 per scenario and the number of bootstrap resamples is 10,000. The standard errors are provided in parenthesis. Cases where FDR control was anti-conservative are indicated with a star ($\star$).

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.80	20	24.457	34.724	0.961	25.227	74.056	100.973⋆	98.435	104.079⋆	100.995
		(10.327)	(12.313)	(1.513)	(10.750)	(8.844)	(8.008)	(7.570)	(8.237)	(7.741)
	30	87.011	96.116	21.549	90.518	117.384	131.894	128.330	134.990⋆	131.377
		(8.623)	(9.358)	(8.117)	(8.747)	(7.163)	(6.574)	(6.581)	(6.760)	(6.827)
	40	121.628	128.207	63.733	125.587	141.347	150.635	147.472	153.506	150.481
		(7.164)	(7.612)	(7.962)	(7.371)	(6.179)	(5.685)	(5.834)	(5.764)	(5.894)
	50	142.576	147.875	93.969	146.399	156.376	162.876	160.284	165.426	163.003
		(6.026)	(6.290)	(6.981)	(6.020)	(5.381)	(5.086)	(5.158)	(5.134)	(5.107)
	60	156.368	160.824	115.508	159.770	166.729	171.305	169.341	173.649	171.835
		(5.666)	(5.827)	(6.650)	(5.651)	(4.993)	(4.848)	(4.968)	(4.813)	(4.883)
	100	183.775	185.656	161.712	185.586	187.781	189.363	188.504	190.583	189.887
		(3.525)	(3.497)	(4.838)	(3.374)	(3.175)	(3.010)	(3.092)	(2.912)	(2.988)
	200	198.062	198.412	193.187	198.415	198.685	198.825	198.731	199.033	198.946
		(1.378)	(1.306)	(2.432)	(1.278)	(1.155)	(1.112)	(1.142)	(1.021)	(1.055)
	300	199.767	199.812	198.763	199.814	199.853	199.872	199.855	199.898	199.891
		(0.501)	(0.457)	(1.117)	(0.447)	(0.397)	(0.377)	(0.398)	(0.325)	(0.334)
	500	199.997	199.997	199.964	199.998	200.000	200.000	200.000	200.000	200.000
		(0.055)	(0.055)	(0.192)	(0.045)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
0.75	20	45.377	65.865	1.658	47.919	104.645	134.759	131.385	139.992⋆	135.756
		(12.836)	(14.319)	(2.151)	(13.583)	(9.879)	(8.699)	(8.311)	(8.964)	(8.596)
	30	123.236	137.271	37.424	129.521	154.609	170.488	166.137	175.553	171.248
		(9.146)	(10.397)	(9.752)	(9.500)	(7.790)	(7.238)	(7.221)	(7.386)	(7.428)
	40	162.836	172.912	92.287	169.201	182.494	192.461	188.919	196.989	193.532
		(7.195)	(7.909)	(8.500)	(7.363)	(6.548)	(6.192)	(6.296)	(6.281)	(6.411)
	50	185.575	193.126	128.670	191.186	198.829	205.673	202.901	209.723	207.110
		(6.513)	(7.079)	(7.666)	(6.626)	(6.050)	(5.803)	(5.893)	(5.743)	(5.949)
	60	200.960	207.119	153.486	205.935	210.779	215.776	213.488	219.403	217.265
		(6.008)	(6.028)	(7.260)	(5.941)	(5.284)	(4.929)	(5.103)	(4.843)	(4.983)
	100	230.499	233.578	204.963	233.385	234.530	236.264	235.289	238.197	237.452
		(3.951)	(3.766)	(5.196)	(3.783)	(3.561)	(3.411)	(3.509)	(3.218)	(3.309)
	200	247.316	247.847	240.490	247.856	247.987	248.198	248.061	248.561	248.454
		(1.623)	(1.444)	(2.711)	(1.421)	(1.383)	(1.313)	(1.345)	(1.184)	(1.227)
	300	249.567	249.692	247.893	249.688	249.719	249.740	249.719	249.805	249.788
		(0.639)	(0.560)	(1.404)	(0.559)	(0.528)	(0.505)	(0.533)	(0.433)	(0.449)
	500	249.988	249.991	249.891	249.992	249.994	249.995	249.995	249.997	249.997
		(0.109)	(0.094)	(0.327)	(0.089)	(0.077)	(0.071)	(0.071)	(0.055)	(0.055)

APPENDIX C

SUPPLEMENTAL SIMULATION RESULTS FOR DEPENDENT TESTS

C.1. Normally Distributed Random Variables

C.1.1. Numerical Summaries of Empirical False Discovery Rates

The numerical summaries of the empirical false discovery rates corresponding to all the simulation parameters for the various FDR controlling procedures considered are reported in the following.

Table C.1. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.1	20	0.040	0.047	0.007	0.040	0.009	0.121	0.135	0.121	0.135
			(0.151)	(0.158)	(0.076)	(0.151)	(0.075)	(0.164)	(0.137)	(0.164)	(0.137)
		30	0.039	0.045	0.005	0.039	0.026	0.077	0.086	0.077	0.086
			(0.073)	(0.077)	(0.056)	(0.072)	(0.064)	(0.076)	(0.070)	(0.076)	(0.070)
		40	0.045	0.050	0.007	0.046	0.037	0.067	0.072	0.067	0.073
			(0.050)	(0.050)	(0.039)	(0.050)	(0.046)	(0.052)	(0.051)	(0.052)	(0.052)
		50	0.046	0.052	0.006	0.047	0.041	0.065	0.066	0.066	0.067
			(0.040)	(0.041)	(0.028)	(0.040)	(0.036)	(0.042)	(0.042)	(0.042)	(0.042)
		60	0.045	0.051	0.006	0.047	0.042	0.059	0.060	0.061	0.062
			(0.034)	(0.036)	(0.018)	(0.035)	(0.032)	(0.035)	(0.036)	(0.036)	(0.036)
		100	0.045	0.051	0.006	0.049	0.044	0.053	0.053	0.056	0.056
			(0.027)	(0.028)	(0.012)	(0.028)	(0.026)	(0.028)	(0.028)	(0.029)	(0.028)
		200	0.045	0.051	0.006	0.049	0.045	0.049	0.049	0.053	0.053
			(0.023)	(0.026)	(0.009)	(0.024)	(0.023)	(0.024)	(0.024)	(0.025)	(0.025)
		300	0.044	0.050	0.006	0.049	0.044	0.047	0.047	0.052	0.052
			(0.022)	(0.024)	(0.008)	(0.023)	(0.022)	(0.023)	(0.022)	(0.023)	(0.023)
		500	0.045	0.051	0.006	0.050	0.045	0.047	0.047	0.052	0.052
			(0.021)	(0.023)	(0.008)	(0.022)	(0.021)	(0.022)	(0.021)	(0.022)	(0.023)
	-0.1	20	0.045	0.055	0.009	0.045	0.013	0.081	0.106	0.081	0.106
			(0.161)	(0.172)	(0.092)	(0.161)	(0.089)	(0.183)	(0.153)	(0.183)	(0.153)
		30	0.043	0.046	0.007	0.043	0.031	0.041	0.055	0.041	0.055
			(0.080)	(0.076)	(0.063)	(0.080)	(0.082)	(0.077)	(0.070)	(0.077)	(0.070)
		40	0.043	0.049	0.007	0.044	0.035	0.035	0.042	0.035	0.042
			(0.049)	(0.051)	(0.040)	(0.049)	(0.043)	(0.045)	(0.045)	(0.045)	(0.045)
		50	0.044	0.051	0.006	0.046	0.039	0.032	0.036	0.033	0.036
			(0.038)	(0.039)	(0.025)	(0.038)	(0.036)	(0.035)	(0.035)	(0.035)	(0.035)
		60	0.045	0.050	0.005	0.047	0.040	0.030	0.033	0.031	0.033
			(0.034)	(0.035)	(0.017)	(0.035)	(0.033)	(0.029)	(0.030)	(0.030)	(0.030)
		100	0.046	0.052	0.007	0.049	0.045	0.027	0.028	0.028	0.030
			(0.026)	(0.028)	(0.012)	(0.027)	(0.025)	(0.021)	(0.022)	(0.022)	(0.022)
		200	0.045	0.050	0.007	0.049	0.044	0.025	0.025	0.027	0.027
			(0.023)	(0.025)	(0.010)	(0.024)	(0.023)	(0.018)	(0.018)	(0.018)	(0.019)
		300	0.044	0.049	0.006	0.048	0.043	0.024	0.024	0.026	0.026
			(0.021)	(0.023)	(0.008)	(0.023)	(0.021)	(0.016)	(0.016)	(0.017)	(0.017)
		500	0.044	0.050	0.006	0.049	0.044	0.023	0.023	0.026	0.026
			(0.021)	(0.022)	(0.008)	(0.022)	(0.021)	(0.015)	(0.015)	(0.016)	(0.016)

Table C.2. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.25$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.25	20	0.042	0.050	0.005	0.042	0.007	0.117	0.125	0.117	0.125
			(0.155)	(0.160)	(0.071)	(0.155)	(0.065)	(0.156)	(0.130)	(0.156)	(0.130)
		30	0.045	0.050	0.004	0.045	0.028	0.083	0.091	0.083	0.092
			(0.086)	(0.085)	(0.038)	(0.086)	(0.063)	(0.077)	(0.071)	(0.077)	(0.071)
		40	0.042	0.047	0.006	0.043	0.034	0.066	0.069	0.066	0.070
			(0.049)	(0.051)	(0.038)	(0.049)	(0.045)	(0.051)	(0.050)	(0.052)	(0.050)
		50	0.045	0.051	0.006	0.046	0.041	0.062	0.065	0.063	0.066
			(0.040)	(0.043)	(0.025)	(0.040)	(0.037)	(0.041)	(0.042)	(0.042)	(0.042)
		60	0.044	0.050	0.005	0.046	0.041	0.057	0.059	0.059	0.060
			(0.034)	(0.036)	(0.017)	(0.034)	(0.032)	(0.036)	(0.036)	(0.037)	(0.036)
		100	0.045	0.052	0.006	0.049	0.045	0.053	0.054	0.056	0.056
			(0.027)	(0.030)	(0.012)	(0.028)	(0.028)	(0.029)	(0.029)	(0.030)	(0.030)
		200	0.044	0.050	0.006	0.048	0.044	0.048	0.048	0.052	0.052
			(0.023)	(0.026)	(0.009)	(0.024)	(0.022)	(0.024)	(0.023)	(0.025)	(0.025)
		300	0.046	0.052	0.006	0.051	0.046	0.049	0.048	0.054	0.053
			(0.022)	(0.024)	(0.009)	(0.023)	(0.022)	(0.023)	(0.023)	(0.024)	(0.023)
		500	0.044	0.050	0.006	0.049	0.044	0.046	0.045	0.051	0.051
			(0.021)	(0.024)	(0.008)	(0.023)	(0.021)	(0.022)	(0.022)	(0.023)	(0.023)
	-0.25	20	0.038	0.042	0.002	0.038	0.009	0.110	0.122	0.110	0.122
			(0.148)	(0.149)	(0.035)	(0.148)	(0.071)	(0.153)	(0.119)	(0.153)	(0.119)
		30	0.043	0.049	0.007	0.044	0.030	0.080	0.088	0.080	0.088
			(0.080)	(0.081)	(0.066)	(0.080)	(0.070)	(0.078)	(0.070)	(0.078)	(0.070)
		40	0.043	0.051	0.004	0.044	0.037	0.069	0.073	0.070	0.073
			(0.050)	(0.052)	(0.026)	(0.050)	(0.045)	(0.053)	(0.051)	(0.053)	(0.052)
		50	0.045	0.051	0.007	0.046	0.040	0.062	0.064	0.063	0.065
			(0.039)	(0.040)	(0.026)	(0.039)	(0.036)	(0.041)	(0.042)	(0.042)	(0.042)
		60	0.046	0.051	0.006	0.048	0.043	0.060	0.061	0.062	0.062
			(0.034)	(0.035)	(0.018)	(0.034)	(0.032)	(0.037)	(0.036)	(0.037)	(0.037)
		100	0.044	0.050	0.007	0.047	0.044	0.052	0.052	0.055	0.055
			(0.027)	(0.028)	(0.012)	(0.027)	(0.026)	(0.028)	(0.027)	(0.029)	(0.028)
		200	0.045	0.050	0.006	0.049	0.045	0.049	0.049	0.053	0.053
			(0.022)	(0.024)	(0.009)	(0.024)	(0.022)	(0.023)	(0.023)	(0.025)	(0.024)
		300	0.044	0.050	0.006	0.049	0.044	0.047	0.047	0.052	0.051
			(0.021)	(0.023)	(0.009)	(0.023)	(0.021)	(0.022)	(0.022)	(0.023)	(0.023)
		500	0.046	0.051	0.006	0.051	0.046	0.048	0.047	0.053	0.053
			(0.021)	(0.023)	(0.008)	(0.022)	(0.021)	(0.022)	(0.022)	(0.023)	(0.023)

Table C.3. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.5$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.5	20	0.046	0.049	0.001	0.046	0.010	0.118	0.131	0.118	0.131
			(0.160)	(0.160)	(0.033)	(0.160)	(0.078)	(0.167)	(0.134)	(0.167)	(0.134)
		30	0.043	0.049	0.006	0.044	0.025	0.082	0.091	0.082	0.091
			(0.077)	(0.078)	(0.061)	(0.077)	(0.062)	(0.077)	(0.076)	(0.077)	(0.076)
		40	0.044	0.049	0.005	0.045	0.036	0.066	0.068	0.066	0.069
			(0.051)	(0.052)	(0.036)	(0.051)	(0.047)	(0.051)	(0.051)	(0.051)	(0.051)
		50	0.045	0.052	0.006	0.046	0.042	0.063	0.065	0.065	0.067
			(0.041)	(0.043)	(0.024)	(0.042)	(0.038)	(0.044)	(0.043)	(0.044)	(0.044)
		60	0.044	0.051	0.006	0.045	0.042	0.058	0.059	0.060	0.061
			(0.036)	(0.039)	(0.019)	(0.036)	(0.034)	(0.039)	(0.039)	(0.040)	(0.039)
		100	0.045	0.052	0.005	0.048	0.044	0.053	0.052	0.056	0.055
			(0.028)	(0.031)	(0.011)	(0.029)	(0.027)	(0.030)	(0.029)	(0.031)	(0.030)
		200	0.045	0.052	0.006	0.049	0.044	0.049	0.049	0.053	0.053
			(0.024)	(0.028)	(0.009)	(0.026)	(0.024)	(0.025)	(0.025)	(0.026)	(0.027)
		300	0.044	0.050	0.006	0.049	0.044	0.047	0.047	0.052	0.051
			(0.024)	(0.027)	(0.009)	(0.025)	(0.024)	(0.025)	(0.024)	(0.026)	(0.026)
		500	0.045	0.052	0.006	0.050	0.045	0.047	0.047	0.053	0.053
			(0.023)	(0.026)	(0.008)	(0.024)	(0.023)	(0.024)	(0.023)	(0.025)	(0.025)
	-0.5	20	0.041	0.044	0.007	0.041	0.014	0.118	0.129	0.118	0.129
			(0.146)	(0.148)	(0.081)	(0.146)	(0.095)	(0.156)	(0.128)	(0.156)	(0.128)
		30	0.047	0.053	0.003	0.047	0.030	0.081	0.088	0.081	0.088
			(0.085)	(0.087)	(0.030)	(0.085)	(0.068)	(0.073)	(0.070)	(0.073)	(0.070)
		40	0.045	0.050	0.005	0.046	0.038	0.068	0.074	0.068	0.075
			(0.051)	(0.051)	(0.034)	(0.051)	(0.045)	(0.053)	(0.052)	(0.054)	(0.052)
		50	0.043	0.048	0.005	0.044	0.040	0.062	0.063	0.064	0.064
			(0.039)	(0.040)	(0.023)	(0.039)	(0.036)	(0.042)	(0.041)	(0.042)	(0.041)
		60	0.046	0.051	0.006	0.047	0.043	0.060	0.061	0.062	0.063
			(0.035)	(0.036)	(0.019)	(0.035)	(0.032)	(0.036)	(0.036)	(0.037)	(0.037)
		100	0.046	0.051	0.006	0.049	0.045	0.053	0.053	0.056	0.057
			(0.028)	(0.028)	(0.012)	(0.028)	(0.027)	(0.029)	(0.029)	(0.030)	(0.030)
		200	0.045	0.050	0.006	0.048	0.044	0.048	0.048	0.052	0.052
			(0.024)	(0.024)	(0.009)	(0.024)	(0.023)	(0.024)	(0.024)	(0.025)	(0.026)
		300	0.044	0.049	0.006	0.048	0.044	0.047	0.047	0.051	0.051
			(0.022)	(0.023)	(0.009)	(0.023)	(0.022)	(0.023)	(0.022)	(0.024)	(0.023)
		500	0.046	0.052	0.006	0.052	0.046	0.048	0.048	0.054	0.053
			(0.020)	(0.021)	(0.008)	(0.022)	(0.020)	(0.021)	(0.021)	(0.022)	(0.022)

Table C.4. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.75$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.75	20	0.034	0.039	0.006	0.034	0.013	0.105	0.116	0.105	0.116
			(0.128)	(0.133)	(0.070)	(0.128)	(0.077)	(0.160)	(0.136)	(0.160)	(0.136)
		30	0.045	0.052	0.003	0.045	0.031	0.076	0.081	0.076	0.081
			(0.090)	(0.093)	(0.031)	(0.091)	(0.083)	(0.091)	(0.084)	(0.091)	(0.084)
		40	0.046	0.055	0.005	0.047	0.038	0.069	0.073	0.070	0.074
			(0.059)	(0.063)	(0.035)	(0.059)	(0.054)	(0.064)	(0.064)	(0.065)	(0.065)
		50	0.045	0.055	0.006	0.047	0.042	0.063	0.064	0.064	0.065
			(0.049)	(0.054)	(0.026)	(0.049)	(0.046)	(0.054)	(0.054)	(0.054)	(0.055)
		60	0.045	0.055	0.006	0.047	0.042	0.059	0.060	0.061	0.061
			(0.041)	(0.047)	(0.019)	(0.042)	(0.041)	(0.046)	(0.046)	(0.046)	(0.047)
		100	0.045	0.055	0.006	0.049	0.045	0.054	0.054	0.057	0.057
			(0.035)	(0.041)	(0.014)	(0.036)	(0.035)	(0.037)	(0.037)	(0.039)	(0.039)
		200	0.045	0.053	0.006	0.049	0.044	0.049	0.049	0.053	0.053
			(0.029)	(0.034)	(0.010)	(0.031)	(0.029)	(0.031)	(0.031)	(0.032)	(0.032)
		300	0.044	0.053	0.006	0.049	0.044	0.047	0.047	0.052	0.051
			(0.029)	(0.035)	(0.010)	(0.031)	(0.029)	(0.030)	(0.030)	(0.032)	(0.032)
		500	0.043	0.052	0.006	0.048	0.043	0.046	0.045	0.050	0.050
			(0.029)	(0.035)	(0.010)	(0.031)	(0.029)	(0.029)	(0.030)	(0.031)	(0.032)
	-0.75	20	0.036	0.038	0.007	0.036	0.011	0.111	0.128	0.111	0.128
			(0.133)	(0.135)	(0.075)	(0.133)	(0.081)	(0.152)	(0.136)	(0.152)	(0.136)
		30	0.042	0.047	0.005	0.042	0.027	0.079	0.085	0.079	0.085
			(0.083)	(0.081)	(0.056)	(0.083)	(0.068)	(0.081)	(0.076)	(0.081)	(0.076)
		40	0.045	0.050	0.006	0.046	0.038	0.069	0.074	0.069	0.074
			(0.053)	(0.054)	(0.035)	(0.053)	(0.049)	(0.057)	(0.057)	(0.058)	(0.058)
		50	0.043	0.049	0.005	0.044	0.039	0.060	0.063	0.061	0.064
			(0.041)	(0.042)	(0.023)	(0.042)	(0.038)	(0.045)	(0.045)	(0.045)	(0.045)
		60	0.045	0.050	0.006	0.047	0.042	0.059	0.061	0.060	0.063
			(0.037)	(0.037)	(0.018)	(0.037)	(0.035)	(0.040)	(0.040)	(0.041)	(0.041)
		100	0.045	0.050	0.006	0.048	0.044	0.053	0.054	0.057	0.057
			(0.030)	(0.030)	(0.013)	(0.031)	(0.029)	(0.032)	(0.032)	(0.033)	(0.033)
		200	0.045	0.051	0.006	0.050	0.045	0.050	0.049	0.054	0.054
			(0.025)	(0.025)	(0.010)	(0.027)	(0.025)	(0.026)	(0.026)	(0.027)	(0.027)
		300	0.044	0.050	0.006	0.049	0.044	0.047	0.047	0.052	0.052
			(0.024)	(0.024)	(0.009)	(0.025)	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)
		500	0.045	0.050	0.006	0.050	0.045	0.047	0.047	0.052	0.052
			(0.024)	(0.024)	(0.009)	(0.026)	(0.024)	(0.025)	(0.025)	(0.026)	(0.026)

Table C.5. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.9$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is 1,000 with 10,000 bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.9	20	0.039	0.049	0.006	0.039	0.016	0.100	0.106	0.100	0.106
			(0.135)	(0.142)	(0.071)	(0.135)	(0.099)	(0.164)	(0.158)	(0.164)	(0.159)
		30	0.036	0.049	0.005	0.037	0.026	0.071	0.078	0.071	0.078
			(0.087)	(0.097)	(0.050)	(0.088)	(0.072)	(0.103)	(0.104)	(0.103)	(0.104)
		40	0.040	0.053	0.006	0.041	0.036	0.065	0.070	0.066	0.071
			(0.070)	(0.079)	(0.042)	(0.071)	(0.068)	(0.083)	(0.083)	(0.083)	(0.083)
		50	0.043	0.056	0.007	0.044	0.037	0.057	0.059	0.058	0.061
			(0.064)	(0.075)	(0.033)	(0.065)	(0.060)	(0.070)	(0.070)	(0.071)	(0.071)
		60	0.040	0.054	0.004	0.041	0.038	0.054	0.055	0.056	0.057
			(0.054)	(0.066)	(0.021)	(0.055)	(0.052)	(0.061)	(0.061)	(0.062)	(0.063)
		100	0.043	0.057	0.005	0.046	0.042	0.051	0.051	0.054	0.054
			(0.049)	(0.061)	(0.019)	(0.051)	(0.048)	(0.052)	(0.052)	(0.054)	(0.054)
		200	0.044	0.059	0.006	0.048	0.044	0.048	0.048	0.052	0.052
			(0.045)	(0.060)	(0.015)	(0.047)	(0.045)	(0.046)	(0.046)	(0.049)	(0.049)
		300	0.044	0.058	0.006	0.049	0.044	0.047	0.047	0.052	0.052
			(0.044)	(0.058)	(0.013)	(0.047)	(0.044)	(0.046)	(0.045)	(0.048)	(0.047)
		500	0.045	0.057	0.007	0.050	0.044	0.046	0.046	0.052	0.052
			(0.042)	(0.053)	(0.014)	(0.045)	(0.042)	(0.043)	(0.043)	(0.046)	(0.046)
	-0.9	20	0.029	0.035	0.007	0.029	0.014	0.101	0.115	0.101	0.115
			(0.110)	(0.120)	(0.076)	(0.110)	(0.085)	(0.152)	(0.140)	(0.152)	(0.140)
		30	0.038	0.041	0.004	0.038	0.029	0.074	0.085	0.074	0.085
			(0.086)	(0.082)	(0.045)	(0.086)	(0.081)	(0.088)	(0.088)	(0.088)	(0.088)
		40	0.040	0.044	0.004	0.041	0.033	0.062	0.066	0.062	0.067
			(0.060)	(0.059)	(0.027)	(0.060)	(0.056)	(0.064)	(0.064)	(0.064)	(0.065)
		50	0.040	0.047	0.006	0.041	0.037	0.058	0.062	0.059	0.063
			(0.048)	(0.048)	(0.030)	(0.048)	(0.045)	(0.053)	(0.054)	(0.053)	(0.054)
		60	0.044	0.049	0.006	0.045	0.041	0.057	0.058	0.058	0.060
			(0.047)	(0.047)	(0.021)	(0.048)	(0.046)	(0.052)	(0.052)	(0.053)	(0.052)
		100	0.045	0.050	0.006	0.048	0.045	0.053	0.054	0.057	0.057
			(0.038)	(0.038)	(0.015)	(0.039)	(0.037)	(0.040)	(0.040)	(0.041)	(0.041)
		200	0.046	0.051	0.006	0.050	0.046	0.050	0.049	0.054	0.054
			(0.034)	(0.034)	(0.011)	(0.036)	(0.034)	(0.036)	(0.036)	(0.038)	(0.038)
		300	0.046	0.051	0.006	0.050	0.045	0.048	0.048	0.053	0.052
			(0.031)	(0.031)	(0.011)	(0.032)	(0.031)	(0.032)	(0.032)	(0.034)	(0.033)
		500	0.044	0.049	0.006	0.049	0.044	0.046	0.046	0.051	0.051
			(0.031)	(0.030)	(0.011)	(0.033)	(0.032)	(0.032)	(0.032)	(0.034)	(0.034)

Table C.6. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.1	20	0.034	0.041	0.003	0.034	0.011	0.112	0.137	0.112	0.137
			(0.112)	(0.122)	(0.050)	(0.112)	(0.082)	(0.130)	(0.115)	(0.130)	(0.115)
		30	0.039	0.047	0.007	0.039	0.027	0.069	0.079	0.069	0.079
			(0.056)	(0.057)	(0.057)	(0.056)	(0.056)	(0.060)	(0.059)	(0.061)	(0.059)
		40	0.042	0.051	0.005	0.044	0.038	0.065	0.068	0.065	0.068
			(0.039)	(0.042)	(0.026)	(0.039)	(0.039)	(0.042)	(0.042)	(0.042)	(0.043)
		50	0.041	0.050	0.006	0.043	0.039	0.057	0.060	0.059	0.062
			(0.031)	(0.034)	(0.019)	(0.031)	(0.032)	(0.034)	(0.035)	(0.035)	(0.035)
		60	0.042	0.050	0.005	0.044	0.040	0.054	0.055	0.057	0.057
			(0.029)	(0.030)	(0.015)	(0.029)	(0.028)	(0.031)	(0.031)	(0.032)	(0.031)
		100	0.043	0.050	0.005	0.046	0.041	0.049	0.048	0.053	0.052
			(0.023)	(0.026)	(0.010)	(0.024)	(0.023)	(0.024)	(0.024)	(0.025)	(0.025)
		200	0.042	0.050	0.006	0.048	0.042	0.046	0.046	0.052	0.051
			(0.019)	(0.021)	(0.008)	(0.020)	(0.019)	(0.019)	(0.019)	(0.021)	(0.020)
		300	0.042	0.050	0.005	0.049	0.042	0.044	0.044	0.051	0.051
			(0.017)	(0.020)	(0.007)	(0.018)	(0.017)	(0.018)	(0.017)	(0.019)	(0.019)
		500	0.043	0.051	0.006	0.050	0.043	0.045	0.044	0.052	0.051
			(0.018)	(0.020)	(0.006)	(0.019)	(0.018)	(0.018)	(0.018)	(0.019)	(0.019)
	-0.1	20	0.031	0.040	0.004	0.031	0.009	0.096	0.124	0.096	0.124
			(0.107)	(0.117)	(0.053)	(0.107)	(0.077)	(0.119)	(0.119)	(0.119)	(0.119)
		30	0.038	0.046	0.003	0.039	0.027	0.072	0.081	0.072	0.081
			(0.060)	(0.058)	(0.034)	(0.060)	(0.064)	(0.058)	(0.057)	(0.058)	(0.057)
		40	0.042	0.050	0.005	0.043	0.038	0.063	0.068	0.064	0.069
			(0.038)	(0.041)	(0.030)	(0.039)	(0.039)	(0.042)	(0.042)	(0.042)	(0.043)
		50	0.043	0.049	0.006	0.044	0.039	0.058	0.059	0.060	0.061
			(0.031)	(0.032)	(0.019)	(0.031)	(0.031)	(0.033)	(0.034)	(0.033)	(0.034)
		60	0.042	0.050	0.005	0.044	0.040	0.055	0.056	0.057	0.058
			(0.028)	(0.030)	(0.015)	(0.029)	(0.028)	(0.030)	(0.030)	(0.030)	(0.030)
		100	0.043	0.051	0.006	0.047	0.043	0.051	0.050	0.055	0.054
			(0.022)	(0.024)	(0.010)	(0.023)	(0.021)	(0.024)	(0.023)	(0.024)	(0.025)
		200	0.043	0.051	0.006	0.049	0.043	0.047	0.047	0.053	0.052
			(0.019)	(0.021)	(0.008)	(0.020)	(0.019)	(0.019)	(0.019)	(0.020)	(0.020)
		300	0.042	0.050	0.005	0.049	0.042	0.045	0.045	0.052	0.051
			(0.017)	(0.020)	(0.007)	(0.019)	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)
		500	0.043	0.050	0.006	0.049	0.043	0.044	0.044	0.051	0.051
			(0.017)	(0.019)	(0.007)	(0.018)	(0.017)	(0.017)	(0.017)	(0.019)	(0.018)

Table C.7. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.25$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.25	20	0.045	0.052	0.005	0.045	0.019	0.111	0.129	0.111	0.129
			(0.141)	(0.144)	(0.071)	(0.141)	(0.123)	(0.139)	(0.125)	(0.139)	(0.125)
		30	0.039	0.047	0.006	0.039	0.030	0.072	0.080	0.072	0.080
			(0.057)	(0.057)	(0.044)	(0.057)	(0.059)	(0.060)	(0.061)	(0.060)	(0.061)
		40	0.041	0.050	0.004	0.042	0.038	0.064	0.068	0.065	0.069
			(0.040)	(0.043)	(0.022)	(0.041)	(0.040)	(0.044)	(0.046)	(0.044)	(0.046)
		50	0.042	0.050	0.004	0.044	0.039	0.057	0.059	0.059	0.061
			(0.031)	(0.033)	(0.017)	(0.032)	(0.031)	(0.033)	(0.034)	(0.034)	(0.035)
		60	0.042	0.050	0.005	0.044	0.040	0.055	0.055	0.057	0.058
			(0.028)	(0.031)	(0.014)	(0.029)	(0.027)	(0.031)	(0.031)	(0.032)	(0.032)
		100	0.042	0.050	0.006	0.046	0.042	0.049	0.049	0.054	0.053
			(0.022)	(0.025)	(0.009)	(0.023)	(0.022)	(0.024)	(0.023)	(0.024)	(0.025)
		200	0.043	0.052	0.005	0.049	0.043	0.047	0.046	0.053	0.052
			(0.019)	(0.022)	(0.007)	(0.020)	(0.019)	(0.019)	(0.019)	(0.021)	(0.020)
		300	0.043	0.052	0.006	0.050	0.043	0.046	0.045	0.052	0.052
			(0.018)	(0.020)	(0.007)	(0.019)	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)
		500	0.042	0.050	0.006	0.049	0.042	0.044	0.043	0.051	0.051
			(0.017)	(0.020)	(0.006)	(0.019)	(0.017)	(0.017)	(0.017)	(0.019)	(0.019)
	-0.25	20	0.038	0.044	0.005	0.038	0.011	0.107	0.123	0.107	0.123
			(0.125)	(0.127)	(0.055)	(0.125)	(0.086)	(0.136)	(0.111)	(0.136)	(0.111)
		30	0.042	0.050	0.005	0.043	0.029	0.071	0.080	0.071	0.080
			(0.063)	(0.061)	(0.047)	(0.063)	(0.068)	(0.061)	(0.060)	(0.061)	(0.061)
		40	0.042	0.048	0.004	0.043	0.037	0.062	0.068	0.063	0.069
			(0.039)	(0.040)	(0.023)	(0.039)	(0.038)	(0.041)	(0.043)	(0.041)	(0.044)
		50	0.042	0.050	0.005	0.044	0.039	0.057	0.059	0.059	0.061
			(0.031)	(0.033)	(0.017)	(0.032)	(0.031)	(0.034)	(0.034)	(0.034)	(0.035)
		60	0.041	0.049	0.006	0.043	0.039	0.053	0.055	0.056	0.058
			(0.027)	(0.029)	(0.016)	(0.028)	(0.027)	(0.028)	(0.029)	(0.029)	(0.030)
		100	0.042	0.050	0.006	0.046	0.041	0.049	0.048	0.053	0.052
			(0.022)	(0.024)	(0.010)	(0.023)	(0.022)	(0.023)	(0.024)	(0.024)	(0.024)
		200	0.043	0.051	0.006	0.048	0.042	0.047	0.046	0.052	0.051
			(0.020)	(0.021)	(0.008)	(0.021)	(0.019)	(0.020)	(0.020)	(0.021)	(0.021)
		300	0.043	0.051	0.006	0.050	0.043	0.046	0.045	0.052	0.052
			(0.018)	(0.019)	(0.007)	(0.019)	(0.018)	(0.018)	(0.018)	(0.020)	(0.020)
		500	0.042	0.050	0.006	0.050	0.042	0.044	0.044	0.052	0.051
			(0.017)	(0.019)	(0.007)	(0.018)	(0.016)	(0.017)	(0.017)	(0.018)	(0.018)

Table C.8. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.5$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.5	20	0.037	0.047	0.007	0.038	0.014	0.103	0.122	0.103	0.122
			(0.122)	(0.130)	(0.075)	(0.122)	(0.089)	(0.128)	(0.118)	(0.128)	(0.118)
		30	0.039	0.047	0.004	0.039	0.030	0.069	0.077	0.069	0.077
			(0.058)	(0.060)	(0.038)	(0.058)	(0.063)	(0.061)	(0.062)	(0.061)	(0.062)
		40	0.043	0.051	0.006	0.043	0.036	0.062	0.066	0.063	0.067
			(0.040)	(0.043)	(0.033)	(0.040)	(0.040)	(0.044)	(0.044)	(0.044)	(0.044)
		50	0.043	0.051	0.005	0.045	0.041	0.058	0.061	0.060	0.062
			(0.034)	(0.036)	(0.018)	(0.034)	(0.034)	(0.037)	(0.038)	(0.038)	(0.038)
		60	0.043	0.053	0.006	0.045	0.041	0.056	0.057	0.059	0.060
			(0.031)	(0.035)	(0.015)	(0.032)	(0.031)	(0.034)	(0.035)	(0.035)	(0.035)
		100	0.042	0.052	0.005	0.046	0.042	0.049	0.050	0.054	0.053
			(0.024)	(0.027)	(0.010)	(0.025)	(0.023)	(0.025)	(0.025)	(0.026)	(0.026)
		200	0.042	0.051	0.006	0.048	0.042	0.046	0.045	0.052	0.051
			(0.021)	(0.025)	(0.008)	(0.022)	(0.021)	(0.022)	(0.021)	(0.024)	(0.023)
		300	0.042	0.051	0.006	0.048	0.042	0.045	0.044	0.051	0.050
			(0.020)	(0.024)	(0.007)	(0.021)	(0.020)	(0.020)	(0.020)	(0.022)	(0.022)
		500	0.043	0.052	0.006	0.050	0.043	0.045	0.044	0.052	0.052
			(0.018)	(0.022)	(0.007)	(0.020)	(0.018)	(0.019)	(0.018)	(0.020)	(0.020)
	-0.5	20	0.041	0.048	0.005	0.041	0.015	0.102	0.126	0.102	0.126
			(0.132)	(0.130)	(0.067)	(0.132)	(0.100)	(0.123)	(0.115)	(0.123)	(0.115)
		30	0.040	0.048	0.005	0.041	0.029	0.072	0.081	0.072	0.081
			(0.058)	(0.057)	(0.045)	(0.058)	(0.058)	(0.060)	(0.060)	(0.060)	(0.060)
		40	0.043	0.050	0.007	0.044	0.037	0.064	0.068	0.065	0.069
			(0.040)	(0.042)	(0.032)	(0.041)	(0.039)	(0.043)	(0.043)	(0.043)	(0.043)
		50	0.044	0.051	0.006	0.046	0.041	0.059	0.060	0.061	0.062
			(0.032)	(0.034)	(0.020)	(0.033)	(0.032)	(0.034)	(0.034)	(0.035)	(0.035)
		60	0.044	0.051	0.006	0.046	0.042	0.056	0.057	0.059	0.060
			(0.029)	(0.029)	(0.016)	(0.029)	(0.029)	(0.031)	(0.030)	(0.032)	(0.031)
		100	0.042	0.050	0.006	0.045	0.041	0.049	0.048	0.053	0.053
			(0.023)	(0.024)	(0.010)	(0.024)	(0.023)	(0.024)	(0.024)	(0.025)	(0.025)
		200	0.042	0.050	0.005	0.048	0.042	0.046	0.045	0.052	0.051
			(0.019)	(0.020)	(0.007)	(0.020)	(0.019)	(0.020)	(0.020)	(0.021)	(0.021)
		300	0.042	0.050	0.006	0.048	0.042	0.044	0.044	0.051	0.050
			(0.017)	(0.018)	(0.007)	(0.018)	(0.017)	(0.018)	(0.018)	(0.019)	(0.019)
		500	0.043	0.051	0.006	0.050	0.042	0.045	0.044	0.052	0.051
			(0.017)	(0.018)	(0.007)	(0.019)	(0.017)	(0.018)	(0.017)	(0.019)	(0.019)

Table C.9. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.75$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.75	20	0.039	0.047	0.010	0.039	0.013	0.105	0.123	0.105	0.123
			(0.121)	(0.125)	(0.088)	(0.120)	(0.084)	(0.140)	(0.132)	(0.140)	(0.132)
		30	0.039	0.049	0.003	0.040	0.025	0.068	0.080	0.069	0.080
			(0.068)	(0.071)	(0.028)	(0.068)	(0.060)	(0.071)	(0.071)	(0.071)	(0.072)
		40	0.040	0.049	0.005	0.041	0.036	0.061	0.064	0.062	0.066
			(0.046)	(0.050)	(0.026)	(0.047)	(0.046)	(0.053)	(0.053)	(0.053)	(0.053)
		50	0.040	0.050	0.005	0.041	0.039	0.056	0.057	0.058	0.059
			(0.039)	(0.045)	(0.018)	(0.040)	(0.039)	(0.045)	(0.045)	(0.046)	(0.046)
		60	0.043	0.054	0.006	0.045	0.041	0.056	0.056	0.058	0.059
			(0.036)	(0.041)	(0.018)	(0.037)	(0.036)	(0.040)	(0.040)	(0.041)	(0.041)
		100	0.041	0.052	0.005	0.045	0.040	0.048	0.048	0.053	0.052
			(0.028)	(0.035)	(0.011)	(0.030)	(0.028)	(0.031)	(0.030)	(0.032)	(0.032)
		200	0.043	0.054	0.005	0.049	0.043	0.046	0.046	0.053	0.052
			(0.025)	(0.031)	(0.009)	(0.028)	(0.026)	(0.027)	(0.026)	(0.029)	(0.029)
		300	0.043	0.055	0.006	0.050	0.043	0.046	0.046	0.053	0.052
			(0.024)	(0.030)	(0.009)	(0.026)	(0.024)	(0.025)	(0.025)	(0.026)	(0.027)
		500	0.043	0.053	0.006	0.050	0.043	0.045	0.044	0.052	0.051
			(0.023)	(0.030)	(0.008)	(0.026)	(0.023)	(0.024)	(0.024)	(0.026)	(0.026)
	-0.75	20	0.035	0.044	0.007	0.035	0.012	0.103	0.127	0.103	0.127
			(0.119)	(0.135)	(0.074)	(0.119)	(0.084)	(0.129)	(0.116)	(0.129)	(0.116)
		30	0.042	0.050	0.004	0.043	0.031	0.074	0.082	0.074	0.083
			(0.062)	(0.063)	(0.034)	(0.063)	(0.067)	(0.064)	(0.063)	(0.064)	(0.063)
		40	0.042	0.049	0.007	0.043	0.037	0.064	0.068	0.065	0.069
			(0.042)	(0.043)	(0.041)	(0.042)	(0.042)	(0.047)	(0.048)	(0.047)	(0.048)
		50	0.042	0.049	0.006	0.044	0.040	0.057	0.061	0.059	0.062
			(0.035)	(0.035)	(0.020)	(0.036)	(0.035)	(0.038)	(0.039)	(0.038)	(0.039)
		60	0.043	0.049	0.005	0.045	0.041	0.054	0.056	0.056	0.058
			(0.031)	(0.032)	(0.015)	(0.032)	(0.031)	(0.034)	(0.035)	(0.035)	(0.035)
		100	0.042	0.049	0.005	0.046	0.042	0.049	0.049	0.053	0.053
			(0.024)	(0.024)	(0.009)	(0.025)	(0.024)	(0.026)	(0.026)	(0.027)	(0.027)
		200	0.044	0.051	0.006	0.049	0.043	0.047	0.047	0.053	0.053
			(0.021)	(0.021)	(0.008)	(0.023)	(0.021)	(0.022)	(0.022)	(0.024)	(0.023)
		300	0.044	0.051	0.006	0.050	0.043	0.046	0.046	0.053	0.052
			(0.019)	(0.019)	(0.008)	(0.021)	(0.019)	(0.020)	(0.020)	(0.022)	(0.021)
		500	0.043	0.050	0.006	0.050	0.043	0.045	0.044	0.052	0.052
			(0.019)	(0.018)	(0.007)	(0.020)	(0.019)	(0.019)	(0.019)	(0.021)	(0.020)

Table C.10. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.9$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.9	20	0.034	0.044	0.002	0.034	0.013	0.087	0.104	0.087	0.105
			(0.116)	(0.122)	(0.028)	(0.116)	(0.086)	(0.145)	(0.145)	(0.145)	(0.145)
		30	0.044	0.056	0.005	0.044	0.029	0.072	0.078	0.072	0.078
			(0.080)	(0.087)	(0.041)	(0.079)	(0.076)	(0.090)	(0.095)	(0.090)	(0.095)
		40	0.039	0.054	0.004	0.040	0.035	0.059	0.064	0.060	0.064
			(0.059)	(0.071)	(0.026)	(0.060)	(0.060)	(0.069)	(0.072)	(0.070)	(0.073)
		50	0.039	0.055	0.006	0.041	0.037	0.053	0.055	0.055	0.057
			(0.051)	(0.064)	(0.026)	(0.052)	(0.051)	(0.056)	(0.058)	(0.057)	(0.059)
		60	0.039	0.054	0.005	0.042	0.037	0.051	0.052	0.054	0.054
			(0.046)	(0.057)	(0.020)	(0.047)	(0.046)	(0.052)	(0.053)	(0.054)	(0.055)
		100	0.042	0.058	0.006	0.046	0.042	0.049	0.049	0.053	0.053
			(0.043)	(0.056)	(0.017)	(0.046)	(0.043)	(0.046)	(0.046)	(0.049)	(0.049)
		200	0.043	0.058	0.005	0.049	0.042	0.046	0.045	0.053	0.052
			(0.036)	(0.047)	(0.012)	(0.039)	(0.035)	(0.037)	(0.037)	(0.040)	(0.040)
		300	0.041	0.056	0.006	0.048	0.041	0.044	0.044	0.050	0.050
			(0.036)	(0.050)	(0.013)	(0.039)	(0.036)	(0.037)	(0.037)	(0.040)	(0.040)
		500	0.044	0.059	0.006	0.051	0.044	0.046	0.045	0.053	0.053
			(0.035)	(0.048)	(0.011)	(0.039)	(0.035)	(0.036)	(0.036)	(0.040)	(0.039)
	-0.9	20	0.037	0.043	0.007	0.037	0.014	0.099	0.115	0.099	0.115
			(0.126)	(0.131)	(0.068)	(0.126)	(0.086)	(0.139)	(0.131)	(0.139)	(0.131)
		30	0.041	0.047	0.005	0.042	0.028	0.072	0.079	0.072	0.079
			(0.071)	(0.071)	(0.040)	(0.071)	(0.069)	(0.076)	(0.077)	(0.076)	(0.078)
		40	0.039	0.046	0.006	0.040	0.035	0.061	0.063	0.062	0.064
			(0.049)	(0.051)	(0.031)	(0.050)	(0.049)	(0.055)	(0.056)	(0.055)	(0.056)
		50	0.041	0.048	0.006	0.043	0.039	0.056	0.058	0.058	0.060
			(0.040)	(0.041)	(0.021)	(0.041)	(0.040)	(0.046)	(0.046)	(0.046)	(0.046)
		60	0.041	0.048	0.005	0.043	0.039	0.053	0.054	0.056	0.057
			(0.036)	(0.037)	(0.016)	(0.037)	(0.037)	(0.041)	(0.042)	(0.042)	(0.042)
		100	0.041	0.048	0.006	0.045	0.041	0.048	0.047	0.052	0.052
			(0.032)	(0.032)	(0.013)	(0.034)	(0.032)	(0.034)	(0.034)	(0.035)	(0.035)
		200	0.042	0.049	0.005	0.048	0.042	0.046	0.046	0.052	0.052
			(0.028)	(0.027)	(0.010)	(0.029)	(0.028)	(0.029)	(0.028)	(0.030)	(0.030)
		300	0.042	0.050	0.006	0.049	0.042	0.045	0.044	0.051	0.051
			(0.027)	(0.026)	(0.010)	(0.029)	(0.027)	(0.028)	(0.027)	(0.030)	(0.030)
		500	0.043	0.050	0.005	0.051	0.043	0.045	0.044	0.052	0.052
			(0.026)	(0.024)	(0.009)	(0.028)	(0.025)	(0.026)	(0.026)	(0.028)	(0.028)

Table C.11. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is 1,000 with 10,000 bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.1	20	0.044	0.050	0.004	0.044	0.018	0.089	0.100	0.089	0.100
			(0.119)	(0.111)	(0.053)	(0.119)	(0.087)	(0.085)	(0.072)	(0.085)	(0.072)
		30	0.039	0.049	0.006	0.040	0.031	0.062	0.068	0.064	0.070
			(0.039)	(0.041)	(0.048)	(0.039)	(0.039)	(0.041)	(0.040)	(0.041)	(0.040)
		40	0.038	0.048	0.005	0.040	0.033	0.054	0.055	0.056	0.057
			(0.027)	(0.030)	(0.017)	(0.028)	(0.027)	(0.029)	(0.028)	(0.029)	(0.029)
		50	0.040	0.051	0.005	0.043	0.037	0.052	0.052	0.055	0.056
			(0.024)	(0.026)	(0.013)	(0.024)	(0.023)	(0.026)	(0.026)	(0.026)	(0.026)
		60	0.039	0.050	0.006	0.043	0.037	0.049	0.049	0.053	0.052
			(0.022)	(0.024)	(0.011)	(0.023)	(0.021)	(0.023)	(0.023)	(0.024)	(0.024)
		100	0.040	0.050	0.005	0.046	0.039	0.045	0.044	0.051	0.050
			(0.018)	(0.020)	(0.008)	(0.018)	(0.017)	(0.018)	(0.018)	(0.020)	(0.019)
		200	0.040	0.052	0.005	0.049	0.040	0.043	0.043	0.052	0.051
			(0.015)	(0.018)	(0.006)	(0.017)	(0.015)	(0.016)	(0.016)	(0.017)	(0.017)
		300	0.041	0.052	0.005	0.050	0.041	0.043	0.042	0.052	0.051
			(0.015)	(0.017)	(0.006)	(0.016)	(0.015)	(0.015)	(0.015)	(0.017)	(0.016)
		500	0.040	0.051	0.005	0.050	0.040	0.042	0.041	0.052	0.051
			(0.014)	(0.017)	(0.006)	(0.016)	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)
	-0.1	20	0.041	0.049	0.003	0.041	0.014	0.087	0.101	0.087	0.101
			(0.094)	(0.098)	(0.050)	(0.094)	(0.072)	(0.076)	(0.067)	(0.076)	(0.067)
		30	0.042	0.051	0.008	0.043	0.033	0.065	0.070	0.066	0.071
			(0.040)	(0.040)	(0.038)	(0.040)	(0.041)	(0.041)	(0.039)	(0.042)	(0.040)
		40	0.040	0.050	0.005	0.042	0.036	0.056	0.058	0.058	0.061
			(0.029)	(0.031)	(0.016)	(0.029)	(0.028)	(0.032)	(0.032)	(0.032)	(0.032)
		50	0.041	0.051	0.005	0.044	0.038	0.052	0.054	0.056	0.056
			(0.024)	(0.026)	(0.013)	(0.025)	(0.023)	(0.025)	(0.025)	(0.026)	(0.026)
		60	0.041	0.052	0.005	0.044	0.039	0.050	0.050	0.054	0.054
			(0.022)	(0.024)	(0.011)	(0.023)	(0.021)	(0.023)	(0.023)	(0.024)	(0.023)
		100	0.040	0.050	0.005	0.045	0.039	0.045	0.044	0.051	0.050
			(0.018)	(0.020)	(0.008)	(0.018)	(0.017)	(0.018)	(0.018)	(0.019)	(0.019)
		200	0.040	0.050	0.006	0.047	0.039	0.042	0.042	0.050	0.049
			(0.015)	(0.017)	(0.006)	(0.017)	(0.015)	(0.016)	(0.015)	(0.017)	(0.017)
		300	0.040	0.051	0.005	0.049	0.040	0.042	0.042	0.052	0.051
			(0.015)	(0.017)	(0.006)	(0.016)	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)
		500	0.040	0.050	0.005	0.050	0.040	0.041	0.041	0.051	0.050
			(0.014)	(0.016)	(0.005)	(0.016)	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)

Table C.12. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.25$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.25	20	0.035	0.045	0.005	0.035	0.010	0.085	0.097	0.085	0.097
			(0.084)	(0.083)	(0.060)	(0.083)	(0.069)	(0.074)	(0.069)	(0.074)	(0.069)
		30	0.038	0.048	0.004	0.039	0.031	0.061	0.066	0.062	0.068
			(0.039)	(0.041)	(0.024)	(0.040)	(0.040)	(0.041)	(0.042)	(0.042)	(0.042)
		40	0.040	0.050	0.005	0.042	0.035	0.055	0.056	0.057	0.059
			(0.029)	(0.030)	(0.021)	(0.029)	(0.028)	(0.030)	(0.030)	(0.030)	(0.031)
		50	0.040	0.051	0.005	0.043	0.038	0.052	0.053	0.055	0.056
			(0.023)	(0.026)	(0.013)	(0.024)	(0.023)	(0.025)	(0.025)	(0.025)	(0.025)
		60	0.040	0.051	0.006	0.044	0.038	0.049	0.050	0.053	0.054
			(0.022)	(0.025)	(0.011)	(0.023)	(0.022)	(0.024)	(0.024)	(0.024)	(0.024)
		100	0.040	0.051	0.005	0.046	0.039	0.045	0.045	0.051	0.051
			(0.019)	(0.022)	(0.008)	(0.020)	(0.019)	(0.019)	(0.019)	(0.021)	(0.020)
		200	0.040	0.051	0.005	0.048	0.039	0.043	0.042	0.051	0.050
			(0.016)	(0.019)	(0.006)	(0.017)	(0.015)	(0.016)	(0.016)	(0.018)	(0.017)
		300	0.039	0.050	0.005	0.048	0.039	0.041	0.041	0.050	0.049
			(0.015)	(0.018)	(0.006)	(0.016)	(0.015)	(0.015)	(0.015)	(0.017)	(0.017)
		500	0.041	0.052	0.005	0.051	0.040	0.042	0.042	0.053	0.052
			(0.014)	(0.018)	(0.005)	(0.016)	(0.014)	(0.015)	(0.015)	(0.017)	(0.017)
	-0.25	20	0.034	0.044	0.003	0.034	0.017	0.084	0.098	0.084	0.098
			(0.080)	(0.081)	(0.044)	(0.080)	(0.086)	(0.075)	(0.068)	(0.075)	(0.068)
		30	0.040	0.051	0.004	0.041	0.033	0.064	0.070	0.065	0.072
			(0.040)	(0.040)	(0.030)	(0.039)	(0.039)	(0.039)	(0.040)	(0.040)	(0.041)
		40	0.039	0.049	0.005	0.041	0.036	0.054	0.057	0.057	0.059
			(0.028)	(0.030)	(0.017)	(0.028)	(0.028)	(0.029)	(0.030)	(0.030)	(0.030)
		50	0.039	0.050	0.005	0.042	0.037	0.051	0.052	0.054	0.055
			(0.023)	(0.025)	(0.013)	(0.024)	(0.023)	(0.025)	(0.024)	(0.025)	(0.025)
		60	0.040	0.050	0.005	0.044	0.038	0.049	0.049	0.053	0.053
			(0.023)	(0.025)	(0.011)	(0.024)	(0.022)	(0.024)	(0.024)	(0.025)	(0.025)
		100	0.040	0.051	0.005	0.046	0.039	0.045	0.044	0.052	0.050
			(0.017)	(0.020)	(0.008)	(0.018)	(0.017)	(0.018)	(0.018)	(0.019)	(0.019)
		200	0.039	0.050	0.005	0.047	0.038	0.042	0.041	0.050	0.049
			(0.015)	(0.017)	(0.006)	(0.016)	(0.015)	(0.015)	(0.015)	(0.017)	(0.017)
		300	0.040	0.050	0.005	0.049	0.039	0.042	0.041	0.051	0.050
			(0.015)	(0.017)	(0.006)	(0.016)	(0.015)	(0.015)	(0.015)	(0.017)	(0.017)
		500	0.040	0.050	0.005	0.049	0.039	0.041	0.040	0.051	0.050
			(0.013)	(0.015)	(0.005)	(0.015)	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)

Table C.13. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.5$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.5	20	0.033	0.044	0.004	0.033	0.012	0.087	0.098	0.087	0.098
			(0.077)	(0.079)	(0.051)	(0.077)	(0.055)	(0.085)	(0.073)	(0.085)	(0.074)
		30	0.038	0.048	0.005	0.039	0.031	0.061	0.066	0.062	0.067
			(0.039)	(0.041)	(0.031)	(0.040)	(0.040)	(0.042)	(0.041)	(0.042)	(0.041)
		40	0.042	0.053	0.005	0.044	0.038	0.058	0.059	0.060	0.061
			(0.029)	(0.032)	(0.017)	(0.030)	(0.030)	(0.032)	(0.033)	(0.033)	(0.034)
		50	0.039	0.050	0.005	0.042	0.036	0.050	0.051	0.054	0.054
			(0.026)	(0.030)	(0.013)	(0.027)	(0.025)	(0.029)	(0.028)	(0.029)	(0.029)
		60	0.040	0.051	0.006	0.044	0.038	0.049	0.049	0.053	0.053
			(0.024)	(0.027)	(0.012)	(0.025)	(0.023)	(0.025)	(0.025)	(0.026)	(0.026)
		100	0.039	0.050	0.005	0.045	0.038	0.044	0.043	0.050	0.049
			(0.019)	(0.024)	(0.008)	(0.021)	(0.019)	(0.020)	(0.020)	(0.022)	(0.021)
		200	0.040	0.052	0.006	0.048	0.040	0.043	0.042	0.051	0.051
			(0.017)	(0.022)	(0.007)	(0.019)	(0.017)	(0.018)	(0.018)	(0.020)	(0.019)
		300	0.040	0.051	0.005	0.049	0.039	0.041	0.041	0.051	0.050
			(0.016)	(0.022)	(0.006)	(0.019)	(0.017)	(0.017)	(0.017)	(0.019)	(0.019)
		500	0.041	0.053	0.005	0.051	0.041	0.042	0.042	0.053	0.052
			(0.016)	(0.020)	(0.006)	(0.019)	(0.016)	(0.017)	(0.017)	(0.019)	(0.019)
	-0.5	20	0.041	0.049	0.007	0.041	0.013	0.085	0.099	0.085	0.099
			(0.110)	(0.098)	(0.072)	(0.110)	(0.069)	(0.075)	(0.069)	(0.075)	(0.069)
		30	0.038	0.048	0.005	0.039	0.031	0.060	0.066	0.062	0.067
			(0.038)	(0.038)	(0.033)	(0.038)	(0.039)	(0.040)	(0.040)	(0.040)	(0.040)
		40	0.040	0.050	0.007	0.042	0.036	0.056	0.058	0.059	0.060
			(0.029)	(0.030)	(0.021)	(0.030)	(0.028)	(0.031)	(0.031)	(0.031)	(0.031)
		50	0.038	0.048	0.005	0.041	0.036	0.050	0.051	0.053	0.054
			(0.025)	(0.025)	(0.013)	(0.025)	(0.024)	(0.026)	(0.026)	(0.026)	(0.026)
		60	0.040	0.050	0.005	0.043	0.038	0.049	0.049	0.053	0.053
			(0.021)	(0.022)	(0.010)	(0.022)	(0.021)	(0.023)	(0.023)	(0.024)	(0.023)
		100	0.041	0.050	0.005	0.046	0.039	0.046	0.045	0.052	0.050
			(0.019)	(0.020)	(0.008)	(0.020)	(0.019)	(0.020)	(0.020)	(0.021)	(0.021)
		200	0.040	0.050	0.005	0.048	0.039	0.043	0.042	0.051	0.050
			(0.015)	(0.017)	(0.006)	(0.017)	(0.015)	(0.016)	(0.016)	(0.017)	(0.017)
		300	0.040	0.050	0.005	0.049	0.040	0.042	0.042	0.051	0.050
			(0.015)	(0.017)	(0.006)	(0.017)	(0.015)	(0.015)	(0.015)	(0.017)	(0.017)
		500	0.040	0.051	0.005	0.050	0.040	0.041	0.041	0.052	0.051
			(0.015)	(0.015)	(0.006)	(0.016)	(0.014)	(0.015)	(0.015)	(0.016)	(0.016)

Table C.14. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.75$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.75	20	0.035	0.046	0.004	0.035	0.015	0.076	0.091	0.076	0.091
			(0.094)	(0.092)	(0.046)	(0.094)	(0.075)	(0.085)	(0.081)	(0.085)	(0.081)
		30	0.037	0.050	0.004	0.039	0.031	0.060	0.066	0.062	0.067
			(0.045)	(0.047)	(0.035)	(0.045)	(0.045)	(0.050)	(0.049)	(0.050)	(0.050)
		40	0.039	0.051	0.004	0.041	0.034	0.055	0.056	0.057	0.058
			(0.035)	(0.040)	(0.018)	(0.036)	(0.035)	(0.039)	(0.040)	(0.040)	(0.040)
		50	0.039	0.052	0.005	0.041	0.036	0.050	0.051	0.054	0.055
			(0.030)	(0.037)	(0.014)	(0.031)	(0.030)	(0.033)	(0.033)	(0.034)	(0.034)
		60	0.040	0.054	0.005	0.044	0.039	0.049	0.049	0.054	0.054
			(0.029)	(0.036)	(0.012)	(0.030)	(0.028)	(0.031)	(0.031)	(0.032)	(0.032)
		100	0.038	0.051	0.005	0.044	0.037	0.044	0.043	0.050	0.049
			(0.023)	(0.030)	(0.009)	(0.025)	(0.022)	(0.024)	(0.024)	(0.027)	(0.026)
		200	0.041	0.055	0.006	0.049	0.040	0.044	0.043	0.052	0.051
			(0.022)	(0.030)	(0.008)	(0.025)	(0.022)	(0.023)	(0.023)	(0.025)	(0.025)
		300	0.040	0.054	0.005	0.049	0.040	0.042	0.042	0.051	0.051
			(0.022)	(0.029)	(0.007)	(0.024)	(0.022)	(0.022)	(0.022)	(0.025)	(0.024)
		500	0.040	0.052	0.006	0.049	0.039	0.041	0.040	0.051	0.050
			(0.020)	(0.027)	(0.007)	(0.023)	(0.020)	(0.021)	(0.020)	(0.023)	(0.023)
	-0.75	20	0.036	0.047	0.007	0.036	0.015	0.081	0.095	0.081	0.095
			(0.092)	(0.098)	(0.075)	(0.092)	(0.081)	(0.080)	(0.072)	(0.080)	(0.073)
		30	0.038	0.048	0.005	0.039	0.031	0.064	0.070	0.065	0.071
			(0.040)	(0.041)	(0.028)	(0.040)	(0.042)	(0.046)	(0.044)	(0.046)	(0.044)
		40	0.039	0.048	0.005	0.040	0.035	0.053	0.055	0.056	0.057
			(0.030)	(0.030)	(0.018)	(0.030)	(0.030)	(0.032)	(0.032)	(0.033)	(0.032)
		50	0.039	0.049	0.005	0.042	0.038	0.051	0.052	0.054	0.055
			(0.024)	(0.025)	(0.014)	(0.025)	(0.024)	(0.026)	(0.026)	(0.027)	(0.027)
		60	0.040	0.051	0.005	0.044	0.038	0.050	0.049	0.054	0.054
			(0.023)	(0.024)	(0.011)	(0.024)	(0.023)	(0.025)	(0.025)	(0.026)	(0.026)
		100	0.041	0.050	0.005	0.046	0.039	0.045	0.045	0.052	0.051
			(0.020)	(0.020)	(0.008)	(0.021)	(0.020)	(0.021)	(0.020)	(0.022)	(0.022)
		200	0.040	0.050	0.005	0.048	0.039	0.043	0.042	0.051	0.050
			(0.017)	(0.017)	(0.006)	(0.019)	(0.017)	(0.018)	(0.017)	(0.019)	(0.019)
		300	0.040	0.050	0.005	0.049	0.040	0.042	0.041	0.051	0.050
			(0.017)	(0.017)	(0.006)	(0.019)	(0.016)	(0.017)	(0.017)	(0.019)	(0.019)
		500	0.040	0.050	0.005	0.050	0.040	0.042	0.041	0.052	0.051
			(0.016)	(0.016)	(0.006)	(0.019)	(0.016)	(0.016)	(0.016)	(0.019)	(0.019)

Table C.15. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.9$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.9	20	0.034 (0.090)	0.048 (0.095)	0.007 (0.069)	0.034 (0.091)	0.014 (0.065)	0.070 (0.096)	0.083 (0.100)	0.070 (0.096)	0.083 (0.100)
		30	0.037 (0.058)	0.053 (0.066)	0.005 (0.033)	0.038 (0.058)	0.031 (0.060)	0.058 (0.066)	0.064 (0.069)	0.059 (0.067)	0.065 (0.069)
		40	0.038 (0.046)	0.056 (0.058)	0.005 (0.022)	0.041 (0.048)	0.035 (0.046)	0.052 (0.053)	0.054 (0.053)	0.055 (0.054)	0.057 (0.055)
		50	0.039 (0.043)	0.057 (0.057)	0.005 (0.018)	0.042 (0.044)	0.037 (0.043)	0.050 (0.048)	0.051 (0.049)	0.054 (0.050)	0.055 (0.051)
		60	0.041 (0.040)	0.060 (0.054)	0.005 (0.017)	0.045 (0.041)	0.040 (0.039)	0.051 (0.044)	0.050 (0.044)	0.056 (0.046)	0.054 (0.046)
		100	0.040 (0.034)	0.057 (0.048)	0.006 (0.014)	0.045 (0.037)	0.038 (0.033)	0.045 (0.036)	0.043 (0.035)	0.051 (0.039)	0.049 (0.038)
		200	0.040 (0.031)	0.057 (0.044)	0.006 (0.011)	0.048 (0.034)	0.040 (0.031)	0.043 (0.032)	0.042 (0.031)	0.051 (0.035)	0.050 (0.035)
		300	0.038 (0.029)	0.054 (0.042)	0.005 (0.010)	0.047 (0.033)	0.038 (0.029)	0.040 (0.030)	0.039 (0.029)	0.049 (0.034)	0.048 (0.033)
		500	0.039 (0.029)	0.055 (0.041)	0.005 (0.010)	0.048 (0.033)	0.039 (0.029)	0.040 (0.029)	0.040 (0.029)	0.050 (0.033)	0.049 (0.033)
	-0.9	20	0.034 (0.084)	0.042 (0.088)	0.005 (0.060)	0.034 (0.085)	0.017 (0.076)	0.078 (0.086)	0.092 (0.085)	0.078 (0.086)	0.092 (0.086)
		30	0.038 (0.047)	0.046 (0.047)	0.006 (0.032)	0.039 (0.048)	0.031 (0.049)	0.060 (0.052)	0.065 (0.053)	0.061 (0.052)	0.066 (0.053)
		40	0.040 (0.038)	0.049 (0.038)	0.006 (0.023)	0.042 (0.038)	0.037 (0.037)	0.056 (0.042)	0.058 (0.042)	0.059 (0.043)	0.060 (0.043)
		50	0.037 (0.033)	0.047 (0.033)	0.005 (0.016)	0.040 (0.034)	0.037 (0.033)	0.050 (0.036)	0.050 (0.037)	0.053 (0.037)	0.054 (0.038)
		60	0.040 (0.030)	0.049 (0.030)	0.005 (0.012)	0.043 (0.031)	0.038 (0.030)	0.049 (0.032)	0.049 (0.032)	0.053 (0.034)	0.053 (0.034)
		100	0.039 (0.025)	0.049 (0.024)	0.005 (0.010)	0.045 (0.027)	0.038 (0.025)	0.044 (0.027)	0.044 (0.027)	0.050 (0.028)	0.049 (0.028)
		200	0.040 (0.023)	0.050 (0.021)	0.006 (0.009)	0.048 (0.025)	0.039 (0.023)	0.043 (0.024)	0.042 (0.024)	0.051 (0.026)	0.050 (0.026)
		300	0.040 (0.021)	0.050 (0.020)	0.005 (0.007)	0.049 (0.024)	0.040 (0.021)	0.042 (0.022)	0.041 (0.022)	0.051 (0.024)	0.050 (0.024)
		500	0.040 (0.022)	0.050 (0.020)	0.005 (0.007)	0.049 (0.024)	0.040 (0.022)	0.041 (0.022)	0.041 (0.022)	0.051 (0.025)	0.050 (0.024)

Table C.16. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.1	20	0.034	0.043	0.005	0.034	0.009	0.071	0.082	0.071	0.083
			(0.077)	(0.066)	(0.061)	(0.077)	(0.049)	(0.060)	(0.061)	(0.060)	(0.061)
		30	0.036	0.049	0.005	0.037	0.028	0.055	0.059	0.057	0.061
			(0.034)	(0.034)	(0.028)	(0.033)	(0.034)	(0.034)	(0.035)	(0.034)	(0.035)
		40	0.037	0.050	0.005	0.040	0.033	0.051	0.050	0.054	0.053
			(0.024)	(0.027)	(0.018)	(0.025)	(0.024)	(0.026)	(0.027)	(0.027)	(0.028)
		50	0.038	0.050	0.005	0.041	0.034	0.047	0.045	0.051	0.049
			(0.020)	(0.022)	(0.011)	(0.020)	(0.020)	(0.021)	(0.022)	(0.022)	(0.022)
		60	0.037	0.050	0.005	0.042	0.034	0.045	0.043	0.050	0.048
			(0.018)	(0.021)	(0.009)	(0.019)	(0.018)	(0.019)	(0.019)	(0.020)	(0.020)
		100	0.038	0.051	0.005	0.045	0.035	0.042	0.039	0.050	0.046
			(0.015)	(0.019)	(0.007)	(0.016)	(0.015)	(0.016)	(0.016)	(0.017)	(0.017)
		200	0.038	0.051	0.005	0.048	0.037	0.040	0.038	0.051	0.048
			(0.013)	(0.016)	(0.006)	(0.014)	(0.013)	(0.014)	(0.013)	(0.015)	(0.015)
		300	0.038	0.051	0.005	0.049	0.037	0.039	0.038	0.051	0.049
			(0.013)	(0.016)	(0.005)	(0.015)	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)
		500	0.037	0.050	0.005	0.049	0.036	0.038	0.037	0.051	0.049
			(0.012)	(0.016)	(0.005)	(0.014)	(0.012)	(0.013)	(0.012)	(0.014)	(0.014)
	-0.1	20	0.030	0.042	0.005	0.031	0.007	0.070	0.084	0.070	0.084
			(0.066)	(0.067)	(0.064)	(0.065)	(0.041)	(0.064)	(0.062)	(0.064)	(0.062)
		30	0.036	0.049	0.004	0.037	0.028	0.055	0.059	0.057	0.060
			(0.033)	(0.034)	(0.023)	(0.033)	(0.035)	(0.034)	(0.033)	(0.034)	(0.033)
		40	0.037	0.049	0.004	0.039	0.032	0.049	0.049	0.052	0.052
			(0.024)	(0.027)	(0.014)	(0.024)	(0.024)	(0.026)	(0.025)	(0.026)	(0.026)
		50	0.037	0.050	0.006	0.041	0.033	0.046	0.045	0.051	0.049
			(0.021)	(0.023)	(0.012)	(0.021)	(0.020)	(0.022)	(0.022)	(0.023)	(0.022)
		60	0.037	0.050	0.005	0.042	0.034	0.045	0.042	0.050	0.047
			(0.018)	(0.021)	(0.010)	(0.019)	(0.018)	(0.019)	(0.019)	(0.020)	(0.020)
		100	0.037	0.050	0.005	0.044	0.035	0.042	0.040	0.049	0.046
			(0.014)	(0.017)	(0.007)	(0.015)	(0.014)	(0.015)	(0.015)	(0.016)	(0.016)
		200	0.037	0.050	0.005	0.047	0.036	0.040	0.038	0.050	0.047
			(0.013)	(0.016)	(0.006)	(0.015)	(0.013)	(0.014)	(0.013)	(0.015)	(0.015)
		300	0.038	0.051	0.005	0.049	0.037	0.040	0.038	0.051	0.049
			(0.013)	(0.016)	(0.005)	(0.014)	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)
		500	0.038	0.051	0.005	0.050	0.037	0.039	0.038	0.051	0.050
			(0.012)	(0.015)	(0.005)	(0.014)	(0.012)	(0.013)	(0.013)	(0.014)	(0.014)

Table C.17. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.25$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.25	20	0.036	0.048	0.005	0.036	0.010	0.072	0.086	0.072	0.086
			(0.077)	(0.067)	(0.051)	(0.077)	(0.064)	(0.060)	(0.059)	(0.060)	(0.059)
		30	0.036	0.048	0.004	0.037	0.027	0.055	0.058	0.057	0.059
			(0.033)	(0.034)	(0.027)	(0.033)	(0.036)	(0.034)	(0.036)	(0.035)	(0.036)
		40	0.038	0.051	0.004	0.041	0.033	0.050	0.051	0.054	0.054
			(0.025)	(0.027)	(0.014)	(0.025)	(0.024)	(0.027)	(0.027)	(0.026)	(0.027)
		50	0.036	0.049	0.005	0.040	0.032	0.045	0.044	0.049	0.047
			(0.021)	(0.024)	(0.011)	(0.022)	(0.020)	(0.022)	(0.021)	(0.023)	(0.022)
		60	0.037	0.050	0.005	0.041	0.033	0.044	0.042	0.050	0.047
			(0.019)	(0.022)	(0.010)	(0.020)	(0.018)	(0.020)	(0.020)	(0.020)	(0.020)
		100	0.037	0.050	0.005	0.044	0.034	0.041	0.039	0.049	0.045
			(0.016)	(0.020)	(0.007)	(0.017)	(0.016)	(0.017)	(0.016)	(0.018)	(0.017)
		200	0.037	0.050	0.005	0.047	0.035	0.039	0.037	0.049	0.047
			(0.013)	(0.017)	(0.006)	(0.015)	(0.013)	(0.014)	(0.013)	(0.016)	(0.015)
		300	0.038	0.051	0.005	0.049	0.036	0.039	0.038	0.051	0.049
			(0.013)	(0.017)	(0.005)	(0.015)	(0.013)	(0.014)	(0.013)	(0.016)	(0.015)
		500	0.038	0.051	0.005	0.050	0.037	0.039	0.038	0.051	0.050
			(0.013)	(0.017)	(0.005)	(0.015)	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)
	-0.25	20	0.035	0.048	0.005	0.035	0.011	0.071	0.085	0.071	0.086
			(0.069)	(0.068)	(0.062)	(0.070)	(0.062)	(0.060)	(0.065)	(0.060)	(0.065)
		30	0.037	0.050	0.004	0.038	0.028	0.056	0.058	0.058	0.060
			(0.033)	(0.033)	(0.023)	(0.033)	(0.034)	(0.034)	(0.035)	(0.034)	(0.036)
		40	0.036	0.049	0.004	0.039	0.032	0.049	0.050	0.052	0.052
			(0.024)	(0.026)	(0.013)	(0.024)	(0.024)	(0.025)	(0.026)	(0.025)	(0.026)
		50	0.037	0.049	0.005	0.040	0.032	0.046	0.044	0.050	0.048
			(0.020)	(0.022)	(0.012)	(0.021)	(0.021)	(0.022)	(0.022)	(0.023)	(0.023)
		60	0.036	0.049	0.005	0.041	0.033	0.044	0.042	0.049	0.046
			(0.018)	(0.021)	(0.009)	(0.019)	(0.018)	(0.019)	(0.019)	(0.020)	(0.020)
		100	0.037	0.050	0.005	0.044	0.035	0.042	0.039	0.049	0.046
			(0.014)	(0.017)	(0.007)	(0.015)	(0.014)	(0.015)	(0.015)	(0.016)	(0.016)
		200	0.038	0.051	0.005	0.048	0.037	0.040	0.039	0.051	0.048
			(0.013)	(0.016)	(0.005)	(0.015)	(0.013)	(0.014)	(0.013)	(0.015)	(0.015)
		300	0.038	0.051	0.005	0.050	0.037	0.040	0.038	0.051	0.050
			(0.013)	(0.016)	(0.005)	(0.015)	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)
		500	0.038	0.050	0.005	0.050	0.037	0.039	0.038	0.051	0.049
			(0.013)	(0.015)	(0.005)	(0.014)	(0.012)	(0.013)	(0.012)	(0.014)	(0.014)

Table C.18. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.5$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.5	20	0.032	0.044	0.004	0.032	0.010	0.069	0.083	0.069	0.083
			(0.071)	(0.063)	(0.053)	(0.072)	(0.065)	(0.065)	(0.066)	(0.065)	(0.066)
		30	0.038	0.052	0.005	0.039	0.029	0.057	0.062	0.059	0.064
			(0.034)	(0.038)	(0.029)	(0.034)	(0.035)	(0.035)	(0.037)	(0.035)	(0.037)
		40	0.037	0.051	0.005	0.039	0.031	0.050	0.049	0.053	0.052
			(0.025)	(0.029)	(0.017)	(0.026)	(0.025)	(0.027)	(0.027)	(0.028)	(0.028)
		50	0.037	0.049	0.005	0.041	0.033	0.046	0.044	0.050	0.048
			(0.022)	(0.026)	(0.012)	(0.023)	(0.022)	(0.023)	(0.023)	(0.024)	(0.024)
		60	0.037	0.051	0.005	0.042	0.034	0.045	0.043	0.050	0.047
			(0.020)	(0.024)	(0.010)	(0.021)	(0.019)	(0.021)	(0.020)	(0.022)	(0.022)
		100	0.038	0.052	0.005	0.045	0.036	0.043	0.041	0.051	0.048
			(0.017)	(0.022)	(0.007)	(0.019)	(0.017)	(0.018)	(0.018)	(0.020)	(0.019)
		200	0.038	0.052	0.005	0.048	0.036	0.040	0.038	0.050	0.048
			(0.014)	(0.019)	(0.006)	(0.016)	(0.014)	(0.015)	(0.014)	(0.017)	(0.016)
		300	0.036	0.049	0.005	0.047	0.035	0.038	0.036	0.049	0.047
			(0.014)	(0.018)	(0.005)	(0.016)	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)
		500	0.037	0.051	0.005	0.049	0.036	0.038	0.037	0.050	0.049
	-0.5	20	0.029	0.042	0.008	0.029	0.013	0.066	0.076	0.066	0.077
			(0.065)	(0.067)	(0.083)	(0.065)	(0.080)	(0.062)	(0.062)	(0.062)	(0.062)
		30	0.037	0.049	0.005	0.038	0.028	0.055	0.057	0.056	0.058
			(0.033)	(0.034)	(0.033)	(0.033)	(0.035)	(0.034)	(0.035)	(0.035)	(0.035)
		40	0.038	0.050	0.005	0.040	0.032	0.050	0.050	0.053	0.052
			(0.024)	(0.025)	(0.016)	(0.024)	(0.024)	(0.025)	(0.026)	(0.025)	(0.027)
		50	0.038	0.051	0.006	0.042	0.034	0.048	0.046	0.053	0.050
			(0.022)	(0.023)	(0.012)	(0.022)	(0.021)	(0.023)	(0.023)	(0.023)	(0.023)
		60	0.036	0.049	0.005	0.041	0.034	0.045	0.042	0.050	0.047
			(0.018)	(0.020)	(0.010)	(0.019)	(0.018)	(0.020)	(0.019)	(0.020)	(0.019)
		100	0.037	0.049	0.005	0.044	0.034	0.042	0.039	0.049	0.046
			(0.014)	(0.017)	(0.007)	(0.016)	(0.015)	(0.016)	(0.015)	(0.017)	(0.016)
		200	0.038	0.051	0.005	0.048	0.036	0.040	0.038	0.051	0.048
			(0.013)	(0.015)	(0.005)	(0.015)	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)
		300	0.038	0.051	0.005	0.049	0.037	0.040	0.038	0.051	0.049
			(0.013)	(0.014)	(0.005)	(0.015)	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)
		500	0.037	0.050	0.005	0.049	0.036	0.038	0.037	0.051	0.049
			(0.013)	(0.015)	(0.005)	(0.015)	(0.012)	(0.013)	(0.013)	(0.015)	(0.014)

Table C.19. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.75$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.75	20	0.036	0.046	0.005	0.036	0.013	0.072	0.083	0.072	0.083
			(0.086)	(0.070)	(0.057)	(0.085)	(0.069)	(0.073)	(0.075)	(0.073)	(0.075)
		30	0.035	0.050	0.005	0.036	0.026	0.052	0.056	0.054	0.058
			(0.037)	(0.043)	(0.028)	(0.038)	(0.041)	(0.040)	(0.042)	(0.041)	(0.043)
		40	0.036	0.051	0.004	0.038	0.030	0.049	0.048	0.052	0.051
			(0.031)	(0.038)	(0.015)	(0.032)	(0.029)	(0.034)	(0.034)	(0.034)	(0.034)
		50	0.039	0.054	0.005	0.042	0.035	0.049	0.047	0.053	0.051
			(0.028)	(0.034)	(0.014)	(0.029)	(0.028)	(0.031)	(0.031)	(0.032)	(0.032)
		60	0.038	0.053	0.005	0.042	0.034	0.045	0.043	0.050	0.048
			(0.025)	(0.032)	(0.011)	(0.026)	(0.024)	(0.027)	(0.026)	(0.029)	(0.028)
		100	0.037	0.052	0.005	0.044	0.034	0.041	0.039	0.049	0.046
			(0.021)	(0.028)	(0.008)	(0.023)	(0.020)	(0.022)	(0.021)	(0.024)	(0.024)
		200	0.037	0.053	0.005	0.047	0.036	0.040	0.038	0.050	0.047
			(0.018)	(0.026)	(0.006)	(0.021)	(0.018)	(0.018)	(0.018)	(0.021)	(0.021)
		300	0.038	0.053	0.005	0.049	0.036	0.039	0.038	0.051	0.049
			(0.018)	(0.026)	(0.006)	(0.021)	(0.018)	(0.018)	(0.018)	(0.021)	(0.021)
		500	0.037	0.051	0.005	0.049	0.036	0.038	0.037	0.050	0.049
			(0.018)	(0.025)	(0.006)	(0.021)	(0.017)	(0.018)	(0.018)	(0.021)	(0.021)
	-0.75	20	0.036	0.044	0.004	0.036	0.009	0.069	0.080	0.069	0.080
			(0.080)	(0.070)	(0.053)	(0.080)	(0.050)	(0.064)	(0.063)	(0.063)	(0.063)
		30	0.035	0.047	0.004	0.036	0.028	0.052	0.056	0.054	0.058
			(0.034)	(0.034)	(0.024)	(0.035)	(0.038)	(0.036)	(0.038)	(0.037)	(0.038)
		40	0.037	0.049	0.005	0.040	0.032	0.050	0.050	0.053	0.053
			(0.025)	(0.026)	(0.017)	(0.026)	(0.024)	(0.027)	(0.027)	(0.028)	(0.028)
		50	0.037	0.049	0.005	0.040	0.033	0.047	0.045	0.051	0.049
			(0.022)	(0.022)	(0.012)	(0.023)	(0.021)	(0.024)	(0.023)	(0.024)	(0.024)
		60	0.037	0.049	0.005	0.041	0.033	0.044	0.042	0.050	0.047
			(0.020)	(0.021)	(0.009)	(0.021)	(0.020)	(0.021)	(0.021)	(0.022)	(0.022)
		100	0.038	0.050	0.005	0.045	0.035	0.042	0.040	0.050	0.047
			(0.017)	(0.018)	(0.007)	(0.019)	(0.016)	(0.018)	(0.018)	(0.020)	(0.019)
		200	0.038	0.051	0.005	0.048	0.036	0.040	0.038	0.051	0.048
			(0.015)	(0.015)	(0.006)	(0.017)	(0.015)	(0.015)	(0.015)	(0.017)	(0.017)
		300	0.038	0.051	0.005	0.049	0.037	0.040	0.038	0.051	0.049
			(0.014)	(0.014)	(0.005)	(0.016)	(0.014)	(0.014)	(0.014)	(0.017)	(0.016)
		500	0.038	0.050	0.005	0.050	0.037	0.039	0.038	0.051	0.050
			(0.014)	(0.014)	(0.005)	(0.016)	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)

Table C.20. Empirical FDRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.9$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.9	20	0.030	0.048	0.004	0.031	0.014	0.064	0.072	0.064	0.073
			(0.070)	(0.079)	(0.040)	(0.071)	(0.067)	(0.081)	(0.085)	(0.081)	(0.085)
		30	0.034	0.053	0.004	0.036	0.026	0.051	0.055	0.053	0.056
			(0.048)	(0.061)	(0.026)	(0.050)	(0.047)	(0.055)	(0.057)	(0.056)	(0.058)
		40	0.035	0.055	0.004	0.038	0.030	0.048	0.047	0.051	0.050
			(0.039)	(0.052)	(0.017)	(0.041)	(0.040)	(0.044)	(0.044)	(0.045)	(0.046)
		50	0.038	0.057	0.005	0.042	0.034	0.048	0.046	0.052	0.050
			(0.037)	(0.051)	(0.017)	(0.039)	(0.036)	(0.040)	(0.040)	(0.042)	(0.042)
		60	0.037	0.058	0.004	0.042	0.034	0.045	0.043	0.050	0.048
			(0.034)	(0.049)	(0.014)	(0.036)	(0.033)	(0.037)	(0.036)	(0.040)	(0.039)
		100	0.037	0.057	0.005	0.045	0.035	0.042	0.039	0.049	0.046
			(0.031)	(0.046)	(0.011)	(0.035)	(0.030)	(0.032)	(0.031)	(0.036)	(0.035)
		200	0.037	0.057	0.005	0.046	0.035	0.039	0.037	0.049	0.046
			(0.027)	(0.043)	(0.009)	(0.031)	(0.026)	(0.027)	(0.026)	(0.031)	(0.030)
		300	0.038	0.058	0.005	0.049	0.037	0.040	0.038	0.051	0.049
			(0.026)	(0.043)	(0.008)	(0.031)	(0.026)	(0.027)	(0.026)	(0.032)	(0.031)
		500	0.037	0.056	0.005	0.049	0.036	0.038	0.036	0.050	0.048
			(0.025)	(0.040)	(0.008)	(0.030)	(0.025)	(0.025)	(0.025)	(0.030)	(0.030)
	-0.9	20	0.033	0.040	0.004	0.033	0.013	0.069	0.080	0.069	0.081
			(0.071)	(0.067)	(0.041)	(0.071)	(0.064)	(0.072)	(0.074)	(0.072)	(0.074)
		30	0.036	0.046	0.005	0.037	0.026	0.054	0.056	0.056	0.058
			(0.041)	(0.041)	(0.027)	(0.041)	(0.041)	(0.045)	(0.046)	(0.046)	(0.047)
		40	0.039	0.051	0.005	0.042	0.035	0.051	0.051	0.054	0.054
			(0.033)	(0.033)	(0.019)	(0.034)	(0.033)	(0.037)	(0.037)	(0.037)	(0.038)
		50	0.036	0.048	0.005	0.040	0.033	0.047	0.045	0.051	0.049
			(0.026)	(0.026)	(0.013)	(0.028)	(0.026)	(0.029)	(0.029)	(0.030)	(0.031)
		60	0.037	0.048	0.005	0.041	0.034	0.045	0.043	0.050	0.047
			(0.024)	(0.024)	(0.012)	(0.026)	(0.024)	(0.027)	(0.026)	(0.028)	(0.027)
		100	0.037	0.049	0.005	0.044	0.035	0.042	0.039	0.049	0.046
			(0.022)	(0.021)	(0.009)	(0.024)	(0.021)	(0.022)	(0.022)	(0.025)	(0.024)
		200	0.037	0.049	0.005	0.046	0.035	0.039	0.037	0.049	0.047
			(0.020)	(0.018)	(0.007)	(0.023)	(0.019)	(0.020)	(0.020)	(0.023)	(0.022)
		300	0.037	0.049	0.005	0.047	0.036	0.038	0.037	0.049	0.048
			(0.019)	(0.017)	(0.007)	(0.021)	(0.019)	(0.019)	(0.019)	(0.022)	(0.021)
		500	0.038	0.050	0.005	0.050	0.037	0.039	0.038	0.051	0.050
			(0.018)	(0.017)	(0.006)	(0.021)	(0.018)	(0.019)	(0.018)	(0.022)	(0.021)

C.1.2. Numerical Summaries of Empirical False Non-discovery Rates

Table C.21. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.1$ and ± 0.25. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.1	20	0.099	0.099	0.100	0.099	0.099	0.096$\star$	0.093$\star$	0.096$\star$	0.093$\star$
		30	0.094	0.093	0.099	0.094	0.094	0.088$\star$	0.085 $\star$	0.088$\star$	0.085$\star$
		40	0.084	0.083	0.096	0.084	0.084	0.078$\star$	0.076$\star$	0.078$\star$	0.076$\star$
		50	0.074	0.073	0.091	0.074	0.073	0.069$\star$	0.067$\star$	0.069$\star$	0.067$\star$
		60	0.065	0.064	0.084	0.065	0.064	0.061	0.060	0.060	0.059
		100	0.042	0.041	0.060	0.041	0.041	0.040	0.039	0.039	0.039
		200	0.019	0.019	0.031	0.019	0.019	0.019	0.018	0.018	0.018
		300	0.011	0.010	0.019	0.010	0.011	0.010	0.010	0.010	0.010
		500	0.004	0.003	0.008	0.003	0.004	0.004	0.004	0.003	0.003
	-0.1	20	0.099	0.099	0.100	0.099	0.099	0.098$\star$	0.096$\star$	0.098$\star$	0.096$\star$
		30	0.094	0.093	0.099	0.094	0.094	0.093	0.090	0.093	0.090
		40	0.084	0.083	0.096	0.084	0.083	0.085	0.082	0.085	0.082
		50	0.075	0.073	0.091	0.074	0.073	0.077	0.075	0.077	0.075
		60	0.066	0.064	0.084	0.065	0.064	0.069	0.067	0.069	0.067
		100	0.042	0.041	0.060	0.041	0.041	0.046	0.045	0.045	0.045
		200	0.019	0.018	0.031	0.019	0.019	0.022	0.022	0.022	0.022
		300	0.011	0.010	0.019	0.010	0.011	0.013	0.013	0.013	0.013
		500	0.004	0.004	0.008	0.004	0.004	0.005	0.005	0.005	0.005
	0.25	20	0.099	0.099	0.100	0.099	0.099	0.095$\star$	0.093$\star$	0.095$\star$	0.093$\star$
		30	0.094	0.093	0.099	0.094	0.094	0.088$\star$	0.085$\star$	0.088$\star$	0.085$\star$
		40	0.084	0.083	0.096	0.084	0.084	0.078$\star$	0.076$\star$	0.078$\star$	0.076$\star$
		50	0.075	0.073	0.091	0.074	0.073	0.069$\star$	0.068$\star$	0.069$\star$	0.067$\star$
		60	0.066	0.064	0.085	0.065	0.064	0.061	0.060	0.061	0.059
		100	0.042	0.041	0.060	0.041	0.041	0.040	0.039	0.039	0.039
		200	0.019	0.018	0.030	0.019	0.019	0.018	0.018	0.018	0.018
		300	0.011	0.010	0.019	0.010	0.010	0.010	0.010	0.010	0.010
		500	0.004	0.004	0.008	0.004	0.004	0.004	0.004	0.003	0.004
	-0.25	20	0.099	0.099	0.100	0.099	0.099	0.095$\star$	0.093$\star$	0.095$\star$	0.093$\star$
		30	0.094	0.093	0.099	0.094	0.094	0.088$\star$	0.086$\star$	0.088$\star$	0.086$\star$
		40	0.085	0.083	0.096	0.084	0.084	0.078$\star$	0.076$\star$	0.078$\star$	0.076$\star$
		50	0.075	0.073	0.091	0.074	0.073	0.069$\star$	0.068$\star$	0.069$\star$	0.067$\star$
		60	0.065	0.064	0.084	0.065	0.064	0.061	0.060	0.061	0.059
		100	0.041	0.040	0.060	0.041	0.040	0.039	0.039	0.039	0.038
		200	0.019	0.018	0.031	0.019	0.019	0.018	0.018	0.018	0.018
		300	0.011	0.010	0.019	0.010	0.011	0.010	0.010	0.010	0.010
		500	0.004	0.004	0.008	0.004	0.004	0.004	0.004	0.003	0.004

Table C.22. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.5$ and ± 0.75. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ number of bootstrap samples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.5	20	0.099	0.099	0.100	0.099	0.100	0.096⋆	0.094⋆	0.096⋆	0.094⋆
		30	0.094	0.093	0.099	0.094	0.094	0.088⋆	0.086⋆	0.088⋆	0.086⋆
		40	0.084	0.083	0.096	0.084	0.084	0.078⋆	0.076⋆	0.078⋆	0.076⋆
		50	0.074	0.073	0.091	0.074	0.073	0.069	0.068⋆	0.069⋆	0.067⋆
		60	0.065	0.064	0.084	0.065	0.064	0.061	0.060	0.061	0.060
		100	0.042	0.040	0.060	0.041	0.041	0.039	0.039	0.039	0.038
		200	0.019	0.018	0.031	0.019	0.019	0.018	0.018	0.018	0.018
		300	0.011	0.010	0.019	0.010	0.011	0.010	0.010	0.010	0.010
		500	0.004	0.003	0.008	0.003	0.004	0.004	0.004	0.003	0.003
	-0.5	20	0.099	0.099	0.100	0.099	0.099	0.095⋆	0.093⋆	0.095⋆	0.093⋆
		30	0.094	0.093	0.099	0.094	0.094	0.088⋆	0.085⋆	0.088⋆	0.085⋆
		40	0.084	0.083	0.096	0.084	0.084	0.078⋆	0.076⋆	0.078⋆	0.076⋆
		50	0.074	0.073	0.091	0.074	0.073	0.069	0.068⋆	0.069⋆	0.067⋆
		60	0.066	0.064	0.085	0.065	0.064	0.061	0.060	0.061	0.060
		100	0.042	0.041	0.060	0.041	0.041	0.039	0.039	0.039	0.038
		200	0.019	0.019	0.031	0.019	0.019	0.019	0.018	0.018	0.018
		300	0.010	0.010	0.019	0.010	0.010	0.010	0.010	0.010	0.010
		500	0.004	0.004	0.008	0.004	0.004	0.004	0.004	0.003	0.003
	0.75	20	0.099	0.099	0.100	0.099	0.099	0.096⋆	0.093⋆	0.096⋆	0.093⋆
		30	0.094	0.093	0.099	0.094	0.094	0.088⋆	0.086⋆	0.088⋆	0.086⋆
		40	0.084	0.083	0.096	0.084	0.084	0.078⋆	0.076⋆	0.078⋆	0.076⋆
		50	0.074	0.072	0.091	0.074	0.073	0.069	0.068⋆	0.069⋆	0.067⋆
		60	0.065	0.063	0.084	0.065	0.064	0.061	0.060	0.060	0.059
		100	0.041	0.040	0.060	0.041	0.041	0.039	0.039	0.039	0.038
		200	0.019	0.018	0.031	0.019	0.019	0.018	0.018	0.018	0.018
		300	0.011	0.010	0.019	0.010	0.011	0.010	0.010	0.010	0.010
		500	0.004	0.004	0.008	0.004	0.004	0.004	0.004	0.003	0.003
	-0.75	20	0.099	0.099	0.100	0.099	0.099	0.095⋆	0.093⋆	0.095⋆	0.093⋆
		30	0.094	0.093	0.099	0.094	0.094	0.088⋆	0.086⋆	0.088⋆	0.086⋆
		40	0.085	0.083	0.096	0.084	0.084	0.079⋆	0.077⋆	0.079⋆	0.077⋆
		50	0.075	0.073	0.091	0.074	0.073	0.069	0.068	0.069	0.068⋆
		60	0.066	0.064	0.085	0.065	0.064	0.061	0.060	0.061	0.060
		100	0.042	0.040	0.060	0.041	0.041	0.039	0.039	0.039	0.038
		200	0.019	0.018	0.031	0.019	0.019	0.019	0.018	0.018	0.018
		300	0.010	0.010	0.019	0.010	0.010	0.010	0.010	0.010	0.010
		500	0.004	0.003	0.008	0.004	0.004	0.004	0.004	0.003	0.003

Table C.23. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9; \rho = \pm 0.9$ and $\pi = 0.85; \rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.9	20	0.099	0.098	0.100	0.099	0.099	0.095⋆	0.093⋆	0.095⋆	0.093⋆
		30	0.094	0.092	0.099	0.094	0.094	0.088⋆	0.085⋆	0.088⋆	0.085⋆
		40	0.085	0.082	0.096	0.084	0.084	0.078⋆	0.076⋆	0.078⋆	0.076⋆
		50	0.075	0.072	0.091	0.074	0.073	0.069	0.068	0.069	0.068
		60	0.066	0.063	0.084	0.065	0.064	0.061	0.060	0.061	0.060
		100	0.042	0.040	0.060	0.041	0.041	0.039	0.039	0.039	0.039
		200	0.019	0.018	0.030	0.019	0.019	0.018	0.018	0.018	0.018
		300	0.011	0.010	0.019	0.010	0.011	0.010	0.010	0.010	0.010
		500	0.004	0.003	0.008	0.004	0.004	0.004	0.004	0.003	0.003
	-0.9	20	0.099	0.099	0.100	0.099	0.099	0.095⋆	0.093⋆	0.095⋆	0.093⋆
		30	0.094	0.093	0.099	0.094	0.094	0.088⋆	0.086⋆	0.088⋆	0.086⋆
		40	0.085	0.083	0.096	0.084	0.084	0.078	0.076⋆	0.078	0.076⋆
		50	0.075	0.073	0.091	0.074	0.074	0.069	0.068	0.069	0.068
		60	0.065	0.064	0.085	0.065	0.064	0.061	0.060	0.061	0.059
		100	0.042	0.040	0.060	0.041	0.041	0.039	0.039	0.039	0.038
		200	0.019	0.018	0.031	0.019	0.019	0.019	0.018	0.018	0.018
		300	0.011	0.010	0.019	0.010	0.010	0.010	0.010	0.010	0.010
		500	0.004	0.004	0.008	0.004	0.004	0.004	0.004	0.003	0.003
0.85	0.1	20	0.148	0.148	0.150	0.148	0.150	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.138	0.149	0.140	0.143	0.134⋆	0.133⋆	0.134⋆	0.132⋆
		40	0.128	0.125	0.144	0.127	0.130	0.122⋆	0.121⋆	0.122⋆	0.121⋆
		50	0.115	0.112	0.137	0.114	0.116	0.110	0.109	0.109	0.109
		60	0.104	0.102	0.129	0.104	0.105	0.100	0.100	0.099	0.099
		100	0.075	0.072	0.100	0.074	0.074	0.072	0.072	0.071	0.071
		200	0.041	0.039	0.062	0.039	0.040	0.040	0.039	0.038	0.038
		300	0.024	0.023	0.042	0.023	0.024	0.024	0.024	0.023	0.022
		500	0.010	0.009	0.021	0.009	0.010	0.010	0.010	0.009	0.009
	-0.1	20	0.148	0.148	0.150	0.148	0.150	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.138	0.149	0.140	0.143	0.134⋆	0.132⋆	0.134⋆	0.132⋆
		40	0.128	0.125	0.144	0.127	0.130	0.122	0.121⋆	0.122⋆	0.121⋆
		50	0.115	0.112	0.136	0.114	0.116	0.110	0.109	0.109	0.108
		60	0.104	0.102	0.128	0.104	0.105	0.100	0.100	0.099	0.099
		100	0.075	0.072	0.100	0.073	0.074	0.072	0.072	0.071	0.071
		200	0.040	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
		300	0.024	0.023	0.042	0.023	0.024	0.024	0.024	0.022	0.022
		500	0.010	0.009	0.020	0.009	0.010	0.010	0.010	0.009	0.009

Table C.24. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.25$ and ± 0.5. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.25	20	0.148	0.148	0.150	0.148	0.150	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.138	0.148	0.140	0.143	0.134⋆	0.132⋆	0.134⋆	0.132⋆
		40	0.128	0.125	0.144	0.127	0.129	0.122⋆	0.121⋆	0.122⋆	0.121⋆
		50	0.115	0.112	0.137	0.114	0.116	0.110	0.109	0.110	0.109
		60	0.104	0.102	0.128	0.104	0.105	0.100	0.099	0.099	0.099
		100	0.075	0.072	0.100	0.073	0.074	0.072	0.072	0.071	0.071
		200	0.040	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
		300	0.024	0.023	0.042	0.023	0.024	0.024	0.024	0.022	0.022
		500	0.010	0.009	0.021	0.009	0.010	0.010	0.010	0.009	0.009
	-0.25	20	0.148	0.148	0.150	0.148	0.150	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.138	0.148	0.140	0.143	0.134⋆	0.132⋆	0.134⋆	0.132⋆
		40	0.127	0.125	0.144	0.127	0.129	0.122	0.121⋆	0.121	0.120⋆
		50	0.115	0.112	0.137	0.114	0.116	0.110	0.109	0.109	0.109
		60	0.105	0.102	0.129	0.104	0.105	0.100	0.100	0.099	0.099
		100	0.075	0.072	0.100	0.074	0.074	0.072	0.072	0.071	0.071
		200	0.041	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
		300	0.024	0.023	0.042	0.023	0.024	0.023	0.023	0.022	0.022
		500	0.010	0.009	0.021	0.009	0.010	0.010	0.010	0.009	0.009
	0.5	20	0.148	0.148	0.150	0.148	0.150	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.138	0.149	0.140	0.143	0.134⋆	0.133⋆	0.134⋆	0.133⋆
		40	0.127	0.125	0.144	0.127	0.130	0.122	0.121⋆	0.122	0.120⋆
		50	0.115	0.112	0.137	0.114	0.116	0.110	0.109	0.109	0.109
		60	0.105	0.102	0.128	0.104	0.105	0.100	0.100	0.099	0.099
		100	0.075	0.072	0.100	0.074	0.075	0.072	0.072	0.071	0.071
		200	0.040	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
		300	0.024	0.023	0.042	0.023	0.024	0.024	0.024	0.022	0.022
		500	0.010	0.009	0.021	0.009	0.010	0.010	0.010	0.009	0.009
	-0.5	20	0.148	0.148	0.150	0.148	0.150	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.138	0.148	0.140	0.143	0.134⋆	0.132⋆	0.134⋆	0.132⋆
		40	0.127	0.125	0.144	0.127	0.129	0.121⋆	0.120⋆	0.121⋆	0.120⋆
		50	0.115	0.112	0.137	0.114	0.116	0.110	0.109	0.109	0.109
		60	0.105	0.102	0.129	0.104	0.105	0.100	0.100	0.100	0.099
		100	0.075	0.072	0.100	0.073	0.074	0.072	0.072	0.071	0.071
		200	0.040	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
		300	0.024	0.023	0.042	0.023	0.024	0.024	0.024	0.022	0.022
		500	0.010	0.009	0.021	0.009	0.010	0.010	0.010	0.009	0.009

Table C.25. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.75$ and ± 0.9. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.75	20	0.148	0.147	0.150	0.148	0.150	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.138	0.149	0.140	0.143	0.134⋆	0.132⋆	0.134⋆	0.132⋆
		40	0.127	0.124	0.144	0.127	0.129	0.122	0.120⋆	0.121	0.120⋆
		50	0.115	0.112	0.137	0.114	0.116	0.110	0.109	0.109	0.109
		60	0.104	0.101	0.128	0.103	0.104	0.100	0.099	0.099	0.099
		100	0.075	0.072	0.100	0.073	0.074	0.072	0.072	0.071	0.071
		200	0.040	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
		300	0.024	0.022	0.042	0.023	0.024	0.023	0.023	0.022	0.022
		500	0.010	0.009	0.021	0.009	0.010	0.010	0.010	0.009	0.009
	-0.75	20	0.148	0.148	0.150	0.148	0.150	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.138	0.148	0.140	0.143	0.134⋆	0.133⋆	0.134⋆	0.133⋆
		40	0.127	0.125	0.144	0.127	0.129	0.122⋆	0.121⋆	0.122⋆	0.120⋆
		50	0.115	0.112	0.137	0.114	0.116	0.110	0.109	0.110	0.109
		60	0.104	0.101	0.128	0.103	0.104	0.100	0.099	0.099	0.098
		100	0.075	0.072	0.100	0.074	0.074	0.072	0.072	0.071	0.071
		200	0.040	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
		300	0.024	0.023	0.041	0.023	0.024	0.023	0.023	0.022	0.022
		500	0.010	0.009	0.021	0.009	0.010	0.010	0.010	0.009	0.009
	0.9	20	0.148	0.147	0.150	0.148	0.150	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.137	0.148	0.140	0.143	0.134⋆	0.132⋆	0.134⋆	0.132⋆
		40	0.127	0.123	0.144	0.127	0.129	0.122	0.121⋆	0.122	0.121⋆
		50	0.115	0.111	0.137	0.115	0.116	0.110	0.110	0.110	0.109
		60	0.104	0.100	0.128	0.104	0.105	0.100	0.100	0.099	0.099
		100	0.075	0.071	0.100	0.073	0.074	0.072	0.072	0.071	0.071
		200	0.040	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
		300	0.024	0.022	0.042	0.023	0.024	0.024	0.024	0.023	0.022
		500	0.010	0.009	0.021	0.010	0.010	0.010	0.010	0.009	0.009
	-0.9	20	0.148	0.148	0.150	0.148	0.149	0.144⋆	0.142⋆	0.144⋆	0.142⋆
		30	0.140	0.138	0.148	0.140	0.143	0.134⋆	0.132⋆	0.134⋆	0.132⋆
		40	0.127	0.125	0.144	0.127	0.129	0.122	0.121⋆	0.121	0.120⋆
		50	0.115	0.112	0.137	0.114	0.116	0.110	0.109	0.109	0.109
		60	0.104	0.101	0.128	0.103	0.105	0.100	0.099	0.099	0.099
		100	0.074	0.072	0.100	0.073	0.074	0.072	0.072	0.071	0.071
		200	0.040	0.038	0.062	0.039	0.040	0.039	0.039	0.038	0.038
		300	0.024	0.023	0.042	0.023	0.024	0.024	0.024	0.022	0.022
		500	0.010	0.009	0.021	0.009	0.010	0.010	0.010	0.009	0.009

Table C.26. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.1$ and ± 0.25. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap samples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.1	20	0.196	0.194	0.200	0.196	0.199	0.189⋆	0.185⋆	0.189⋆	0.185⋆
		30	0.180	0.175	0.197	0.179	0.183	0.171	0.169⋆	0.171⋆	0.168⋆
		40	0.160	0.155	0.188	0.159	0.162	0.153	0.151	0.151	0.150
		50	0.143	0.137	0.175	0.141	0.144	0.137	0.136	0.135	0.135
		60	0.129	0.123	0.163	0.127	0.129	0.124	0.123	0.122	0.121
		100	0.090	0.085	0.125	0.087	0.090	0.087	0.087	0.084	0.085
		200	0.047	0.044	0.074	0.045	0.047	0.046	0.046	0.043	0.043
		300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.026	0.026
		500	0.012	0.011	0.025	0.011	0.012	0.012	0.012	0.010	0.010
	-0.1	20	0.196	0.194	0.200	0.196	0.199	0.188⋆	0.185⋆	0.188⋆	0.185⋆
		30	0.179	0.175	0.197	0.179	0.183	0.171⋆	0.169⋆	0.170⋆	0.168⋆
		40	0.160	0.155	0.188	0.159	0.162	0.153	0.151	0.152	0.150
		50	0.143	0.137	0.175	0.141	0.144	0.137	0.136	0.135	0.135
		60	0.129	0.123	0.163	0.127	0.129	0.124	0.123	0.122	0.121
		100	0.090	0.085	0.124	0.087	0.090	0.087	0.087	0.084	0.084
		200	0.047	0.044	0.075	0.045	0.047	0.046	0.046	0.043	0.043
		300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.026	0.026
		500	0.012	0.011	0.025	0.011	0.012	0.012	0.012	0.010	0.010
	0.25	20	0.196	0.195	0.200	0.196	0.199	0.188⋆	0.185⋆	0.188⋆	0.185⋆
		30	0.179	0.175	0.196	0.179	0.183	0.171	0.169⋆	0.171	0.168⋆
		40	0.160	0.155	0.188	0.159	0.162	0.153	0.151	0.152	0.150
		50	0.143	0.138	0.175	0.141	0.144	0.137	0.136	0.135	0.135
		60	0.129	0.123	0.163	0.127	0.129	0.124	0.123	0.122	0.121
		100	0.090	0.085	0.124	0.087	0.090	0.087	0.087	0.084	0.085
		200	0.047	0.044	0.075	0.045	0.047	0.046	0.046	0.043	0.043
		300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.026	0.026
		500	0.012	0.011	0.025	0.011	0.012	0.012	0.012	0.010	0.010
	-0.25	20	0.196	0.194	0.200	0.196	0.199	0.188⋆	0.185⋆	0.188⋆	0.185⋆
		30	0.179	0.175	0.196	0.179	0.183	0.171⋆	0.169⋆	0.171⋆	0.168⋆
		40	0.160	0.155	0.187	0.159	0.162	0.153	0.152	0.152	0.150
		50	0.143	0.138	0.175	0.141	0.144	0.137	0.137	0.136	0.135
		60	0.129	0.123	0.163	0.127	0.129	0.124	0.123	0.121	0.121
		100	0.090	0.085	0.124	0.087	0.089	0.087	0.087	0.084	0.084
		200	0.047	0.044	0.074	0.044	0.047	0.046	0.046	0.043	0.043
		300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.025	0.025
		500	0.012	0.010	0.025	0.011	0.012	0.011	0.011	0.010	0.010

Table C.27. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.5$ and ± 0.75. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.5	20	0.196	0.194	0.200	0.196	0.199	0.188⋆	0.185⋆	0.188⋆	0.185⋆
		30	0.180	0.175	0.197	0.179	0.183	0.171	0.169⋆	0.171	0.168⋆
		40	0.160	0.155	0.187	0.159	0.162	0.153	0.152	0.152	0.150
		50	0.143	0.138	0.176	0.142	0.144	0.137	0.137	0.136	0.135
		60	0.129	0.123	0.163	0.127	0.129	0.124	0.123	0.122	0.121
		100	0.090	0.086	0.125	0.088	0.090	0.087	0.087	0.085	0.085
		200	0.047	0.044	0.075	0.045	0.047	0.046	0.046	0.044	0.044
		300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.026	0.026
		500	0.012	0.011	0.025	0.011	0.012	0.012	0.012	0.010	0.010
	-0.5	20	0.196	0.195	0.200	0.196	0.199	0.189⋆	0.185⋆	0.189⋆	0.185⋆
		30	0.179	0.175	0.197	0.179	0.183	0.171	0.169⋆	0.170	0.168⋆
		40	0.160	0.155	0.187	0.159	0.162	0.153	0.151	0.152	0.150
		50	0.143	0.138	0.176	0.141	0.144	0.137	0.136	0.135	0.135
		60	0.129	0.123	0.163	0.127	0.129	0.124	0.123	0.122	0.121
		100	0.090	0.086	0.124	0.088	0.090	0.088	0.088	0.085	0.085
		200	0.048	0.044	0.075	0.045	0.047	0.046	0.046	0.044	0.044
		300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.026	0.026
		500	0.012	0.011	0.025	0.011	0.012	0.012	0.012	0.010	0.010
	0.75	20	0.196	0.194	0.200	0.196	0.199	0.189⋆	0.185⋆	0.189⋆	0.185⋆
		30	0.179	0.175	0.196	0.179	0.183	0.171	0.169⋆	0.171	0.168⋆
		40	0.160	0.155	0.187	0.159	0.162	0.153	0.152	0.152	0.151
		50	0.143	0.137	0.176	0.142	0.144	0.137	0.137	0.136	0.135
		60	0.128	0.122	0.163	0.127	0.129	0.123	0.123	0.121	0.121
		100	0.090	0.085	0.124	0.087	0.090	0.087	0.087	0.084	0.085
		200	0.047	0.043	0.075	0.044	0.047	0.046	0.046	0.043	0.043
		300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.025	0.025
		500	0.012	0.011	0.025	0.011	0.012	0.012	0.011	0.010	0.010
	-0.75	20	0.196	0.194	0.200	0.196	0.199	0.189⋆	0.185⋆	0.189⋆	0.185⋆
		30	0.179	0.176	0.196	0.179	0.183	0.171⋆	0.169⋆	0.171⋆	0.168⋆
		40	0.160	0.155	0.187	0.159	0.162	0.153	0.152	0.152	0.151
		50	0.143	0.138	0.175	0.141	0.144	0.137	0.136	0.135	0.135
		60	0.129	0.123	0.163	0.127	0.129	0.124	0.123	0.122	0.121
		100	0.090	0.085	0.124	0.087	0.090	0.087	0.087	0.084	0.085
		200	0.047	0.044	0.075	0.044	0.047	0.046	0.046	0.043	0.043
		300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.026	0.026
		500	0.012	0.011	0.025	0.011	0.012	0.012	0.012	0.010	0.010

Table C.28. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8; \rho = \pm 0.9$ and $\pi_0 = 0.75; \rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.9	20	0.196	0.192	0.200	0.196	0.198	0.188⋆	0.185⋆	0.188⋆	0.185⋆
		30	0.180	0.174	0.197	0.179	0.183	0.171	0.169⋆	0.171	0.168⋆
		40	0.159	0.152	0.187	0.158	0.161	0.152	0.151	0.151	0.150
		50	0.143	0.135	0.175	0.141	0.144	0.137	0.136	0.135	0.135
		60	0.128	0.121	0.163	0.126	0.129	0.123	0.123	0.121	0.121
		100	0.090	0.084	0.124	0.087	0.090	0.087	0.087	0.084	0.085
		200	0.047	0.043	0.075	0.044	0.047	0.046	0.046	0.043	0.043
		300	0.029	0.026	0.050	0.027	0.028	0.028	0.028	0.026	0.026
		500	0.012	0.010	0.025	0.011	0.012	0.012	0.012	0.011	0.011
	-0.9	20	0.196	0.194	0.200	0.196	0.199	0.189⋆	0.186⋆	0.189⋆	0.185⋆
		30	0.179	0.175	0.196	0.179	0.183	0.171	0.169⋆	0.170	0.168⋆
		40	0.160	0.155	0.188	0.159	0.162	0.153	0.152	0.152	0.151
		50	0.143	0.137	0.176	0.141	0.144	0.137	0.136	0.135	0.135
		60	0.129	0.123	0.163	0.127	0.129	0.124	0.123	0.122	0.121
		100	0.090	0.085	0.124	0.087	0.090	0.087	0.087	0.084	0.084
		200	0.047	0.043	0.074	0.044	0.047	0.046	0.046	0.043	0.043
		300	0.028	0.026	0.050	0.026	0.028	0.028	0.028	0.026	0.025
		500	0.012	0.011	0.025	0.011	0.012	0.012	0.012	0.011	0.011
0.75	0.1	20	0.244	0.242	0.249	0.244	0.249	0.236⋆	0.234⋆	0.236⋆	0.234⋆
		30	0.225	0.218	0.245	0.224	0.232	0.214	0.215	0.214	0.214
		40	0.202	0.194	0.235	0.200	0.206	0.193	0.194	0.191	0.193
		50	0.181	0.172	0.222	0.178	0.184	0.173	0.175	0.171	0.173
		60	0.163	0.154	0.208	0.159	0.165	0.156	0.158	0.153	0.155
		100	0.115	0.107	0.160	0.110	0.116	0.111	0.113	0.107	0.108
		200	0.060	0.055	0.096	0.056	0.060	0.059	0.059	0.055	0.055
		300	0.037	0.034	0.064	0.034	0.037	0.037	0.037	0.033	0.033
		500	0.018	0.015	0.034	0.016	0.017	0.017	0.017	0.015	0.015
	-0.1	20	0.245	0.242	0.249	0.244	0.249	0.236⋆	0.234⋆	0.236⋆	0.234⋆
		30	0.225	0.218	0.245	0.224	0.232	0.215	0.215	0.214	0.214
		40	0.202	0.194	0.235	0.200	0.206	0.193	0.194	0.191	0.192
		50	0.181	0.172	0.222	0.178	0.184	0.173	0.175	0.170	0.172
		60	0.163	0.154	0.208	0.160	0.166	0.157	0.159	0.154	0.156
		100	0.115	0.107	0.160	0.110	0.116	0.111	0.113	0.107	0.108
		200	0.060	0.055	0.096	0.056	0.060	0.059	0.059	0.054	0.055
		300	0.038	0.034	0.064	0.034	0.037	0.037	0.037	0.033	0.034
		500	0.018	0.016	0.035	0.016	0.018	0.018	0.018	0.016	0.016

Table C.29. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.25$ and ± 0.5. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.25	20	0.244	0.242	0.249	0.244	0.249	0.236*	0.235*	0.236*	0.234*
		30	0.224	0.218	0.245	0.224	0.232	0.215	0.215	0.214	0.214
		40	0.201	0.193	0.235	0.200	0.206	0.193	0.194	0.191	0.193
		50	0.181	0.172	0.222	0.178	0.184	0.173	0.175	0.170	0.172
		60	0.163	0.155	0.208	0.160	0.166	0.157	0.159	0.153	0.155
		100	0.115	0.107	0.160	0.110	0.116	0.111	0.113	0.107	0.108
		200	0.060	0.055	0.096	0.056	0.061	0.059	0.060	0.055	0.055
		300	0.037	0.034	0.065	0.034	0.037	0.037	0.037	0.033	0.033
		500	0.018	0.016	0.035	0.016	0.018	0.018	0.018	0.016	0.016
	-0.25	20	0.245	0.242	0.249	0.245	0.249	0.236*	0.234*	0.236*	0.234*
		30	0.225	0.218	0.246	0.224	0.232	0.215	0.215	0.214	0.215
		40	0.201	0.193	0.235	0.199	0.206	0.192	0.194	0.191	0.192
		50	0.181	0.173	0.222	0.179	0.184	0.174	0.175	0.171	0.173
		60	0.163	0.154	0.207	0.160	0.165	0.156	0.158	0.153	0.155
		100	0.115	0.107	0.160	0.110	0.116	0.111	0.113	0.106	0.108
		200	0.060	0.055	0.096	0.056	0.060	0.059	0.059	0.054	0.055
		300	0.038	0.034	0.064	0.034	0.037	0.037	0.037	0.033	0.034
		500	0.018	0.016	0.035	0.016	0.018	0.018	0.018	0.016	0.016
	0.5	20	0.244	0.241	0.250	0.244	0.249	0.236*	0.234*	0.236*	0.234*
		30	0.224	0.218	0.245	0.224	0.232	0.215	0.215	0.214	0.215*
		40	0.202	0.193	0.235	0.200	0.207	0.193	0.195	0.191	0.193
		50	0.181	0.172	0.222	0.178	0.184	0.174	0.175	0.171	0.173
		60	0.163	0.154	0.208	0.160	0.165	0.157	0.159	0.153	0.155
		100	0.114	0.106	0.159	0.110	0.115	0.111	0.112	0.106	0.108
		200	0.061	0.055	0.096	0.056	0.061	0.059	0.060	0.055	0.055
		300	0.037	0.034	0.065	0.034	0.037	0.037	0.037	0.033	0.033
		500	0.018	0.016	0.035	0.016	0.018	0.018	0.018	0.015	0.015
	-0.5	20	0.244	0.242	0.250	0.244	0.249	0.236*	0.234*	0.236*	0.234*
		30	0.224	0.218	0.245	0.223	0.231	0.214	0.215	0.214	0.214
		40	0.202	0.194	0.235	0.200	0.207	0.193	0.194	0.191	0.193
		50	0.181	0.173	0.222	0.179	0.184	0.174	0.175	0.171	0.173
		60	0.163	0.154	0.208	0.160	0.166	0.157	0.159	0.154	0.156
		100	0.115	0.107	0.160	0.110	0.116	0.111	0.112	0.106	0.108
		200	0.061	0.055	0.096	0.056	0.061	0.059	0.060	0.055	0.055
		300	0.037	0.033	0.064	0.034	0.037	0.036	0.037	0.033	0.033
		500	0.018	0.016	0.035	0.016	0.018	0.018	0.018	0.016	0.016

Table C.30. Empirical FNRs for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.75$ and ± 0.9. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors of the estimated quantities were of the order of 0.008 or less for all the methods. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.75	20	0.244	0.241	0.249	0.244	0.249	0.236*	0.234*	0.236*	0.234*
		30	0.225	0.217	0.245	0.224	0.232	0.215	0.215	0.214	0.215
		40	0.201	0.192	0.235	0.200	0.206	0.193	0.194	0.191	0.192
		50	0.181	0.171	0.222	0.178	0.184	0.174	0.175	0.171	0.173
		60	0.163	0.153	0.208	0.160	0.165	0.157	0.159	0.153	0.155
		100	0.115	0.107	0.160	0.110	0.116	0.111	0.113	0.107	0.108
		200	0.060	0.054	0.096	0.056	0.060	0.059	0.059	0.055	0.055
		300	0.038	0.034	0.065	0.034	0.037	0.037	0.037	0.034	0.034
		500	0.018	0.016	0.035	0.016	0.018	0.018	0.018	0.016	0.016
	-0.75	20	0.245	0.242	0.249	0.245	0.249	0.236*	0.234*	0.236*	0.234*
		30	0.224	0.218	0.245	0.224	0.232	0.215	0.215	0.214	0.214
		40	0.202	0.193	0.235	0.200	0.207	0.193	0.194	0.191	0.193
		50	0.181	0.172	0.222	0.178	0.184	0.174	0.175	0.171	0.173
		60	0.163	0.154	0.208	0.159	0.165	0.157	0.158	0.153	0.155
		100	0.115	0.107	0.160	0.110	0.116	0.111	0.113	0.107	0.108
		200	0.061	0.055	0.096	0.056	0.061	0.059	0.060	0.055	0.055
		300	0.037	0.033	0.064	0.034	0.037	0.037	0.037	0.033	0.033
		500	0.018	0.016	0.035	0.016	0.018	0.018	0.018	0.015	0.015
	0.9	20	0.244	0.239	0.249	0.244	0.248	0.236*	0.234*	0.236*	0.234*
		30	0.224	0.214	0.245	0.223	0.232	0.215	0.215	0.214	0.214
		40	0.201	0.190	0.235	0.200	0.206	0.192	0.194	0.190	0.192
		50	0.181	0.169	0.222	0.178	0.184	0.173	0.175	0.171	0.173
		60	0.163	0.151	0.208	0.159	0.165	0.156	0.158	0.153	0.155
		100	0.115	0.105	0.160	0.110	0.116	0.111	0.113	0.107	0.108
		200	0.060	0.053	0.096	0.056	0.060	0.059	0.059	0.054	0.055
		300	0.037	0.033	0.064	0.034	0.037	0.037	0.037	0.033	0.033
		500	0.018	0.015	0.035	0.016	0.018	0.017	0.017	0.015	0.015
	-0.9	20	0.244	0.242	0.249	0.244	0.249	0.236*	0.234*	0.236*	0.234*
		30	0.225	0.218	0.245	0.224	0.232	0.215	0.215	0.214	0.215
		40	0.201	0.193	0.235	0.200	0.206	0.193	0.194	0.191	0.192
		50	0.181	0.172	0.222	0.179	0.184	0.174	0.175	0.171	0.173
		60	0.163	0.154	0.208	0.160	0.166	0.157	0.159	0.154	0.155
		100	0.115	0.106	0.160	0.110	0.116	0.111	0.113	0.107	0.108
		200	0.061	0.054	0.096	0.056	0.061	0.059	0.060	0.055	0.055
		300	0.038	0.033	0.064	0.034	0.037	0.037	0.037	0.033	0.034
		500	0.018	0.016	0.035	0.016	0.018	0.017	0.018	0.015	0.015

C.1.3. Numerical Summaries of Average Number of False Hypotheses Rejected

Table C.31. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.1	20	1.146	1.349	0.171	1.150	0.572	5.020⋆	7.479⋆	5.020⋆	7.479⋆
			(1.595)	(1.765)	(0.475)	(1.601)	(1.121)	(3.294)	(3.783)	(3.294)	(3.783)
		30	6.919	7.725	1.057	6.969	6.460	13.380⋆	16.049⋆	13.380⋆	16.050⋆
			(4.373)	(4.737)	(1.399)	(4.424)	(4.499)	(5.125)	(4.829)	(5.125)	(4.831)
		40	17.028	18.409	4.078	17.223	17.652	23.561⋆	25.612⋆	23.617⋆	25.680⋆
			(5.512)	(5.739)	(2.959)	(5.590)	(5.689)	(5.347)	(5.202)	(5.426)	(5.285)
		50	27.868	29.285	9.959	28.206	29.033	33.467⋆	35.077⋆	33.764⋆	35.356⋆
			(5.241)	(5.498)	(4.021)	(5.373)	(5.413)	(5.316)	(5.274)	(5.439)	(5.379)
		60	37.274	38.709	17.143	37.766	38.667	41.913	42.993	42.293	43.429
			(5.338)	(5.474)	(4.602)	(5.371)	(5.385)	(5.184)	(5.192)	(5.287)	(5.259)
		100	60.948	62.083	42.727	61.566	61.848	63.106	63.485	63.619	64.032
			(4.370)	(4.486)	(4.393)	(4.411)	(4.329)	(4.319)	(4.304)	(4.355)	(4.309)
		200	82.416	83.102	71.670	82.912	82.661	83.024	83.135	83.472	83.554
			(3.049)	(3.088)	(3.294)	(3.070)	(3.041)	(3.021)	(3.008)	(2.998)	(2.989)
		300	90.254	90.713	82.677	90.631	90.385	90.569	90.566	90.906	90.899
			(2.502)	(2.515)	(2.878)	(2.497)	(2.511)	(2.501)	(2.463)	(2.499)	(2.487)
		500	96.679	96.872	92.544	96.867	96.707	96.769	96.763	96.963	96.937
			(1.600)	(1.594)	(2.288)	(1.581)	(1.602)	(1.586)	(1.604)	(1.555)	(1.559)
	-0.1	20	1.243	1.438	0.156	1.245	0.606	2.451⋆	4.643⋆	2.451⋆	4.643⋆
			(1.698)	(1.827)	(0.443)	(1.702)	(1.222)	(2.184)	(2.972)	(2.184)	(2.972)
		30	6.873	7.745	1.024	6.918	6.338	7.701	10.845	7.701	10.845
			(4.160)	(4.451)	(1.355)	(4.190)	(4.389)	(3.918)	(4.261)	(3.918)	(4.261)
		40	17.440	18.755	4.030	17.610	18.123	16.287	19.446	16.287	19.459
			(5.406)	(5.500)	(2.910)	(5.451)	(5.522)	(4.944)	(4.806)	(4.944)	(4.837)
		50	27.398	28.949	9.539	27.720	28.740	24.736	27.249	24.784	27.346
			(5.519)	(5.527)	(3.913)	(5.564)	(5.601)	(5.166)	(5.099)	(5.240)	(5.209)
		60	36.845	38.220	17.100	37.306	38.209	33.327	35.290	33.539	35.521
			(5.100)	(5.217)	(4.605)	(5.200)	(5.100)	(4.957)	(4.776)	(5.044)	(4.844)
		100	60.892	62.040	42.688	61.534	61.716	56.801	57.571	57.282	58.001
			(4.074)	(4.096)	(4.466)	(4.085)	(4.043)	(4.057)	(3.985)	(4.126)	(4.050)
		200	82.574	83.166	71.660	83.075	82.810	79.503	79.668	79.912	80.064
			(3.015)	(3.047)	(3.361)	(2.989)	(2.980)	(3.015)	(3.036)	(3.021)	(3.070)
		300	90.370	90.784	82.678	90.709	90.462	88.097	88.143	88.438	88.441
			(2.547)	(2.533)	(2.869)	(2.526)	(2.520)	(2.663)	(2.667)	(2.651)	(2.689)
		500	96.612	96.790	92.533	96.790	96.640	95.378	95.392	95.559	95.572
			(1.633)	(1.587)	(2.185)	(1.597)	(1.630)	(1.819)	(1.816)	(1.769)	(1.758)

Table C.32. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.25$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.25	20	1.230	1.465	0.140	1.232	0.573	5.170⋆	7.763⋆	5.170⋆	7.763⋆
			(1.708)	(1.897)	(0.406)	(1.714)	(1.176)	(3.475)	(3.878)	(3.475)	(3.878)
		30	6.755	7.667	0.973	6.819	6.357	13.411⋆	16.073⋆	13.412⋆	16.075⋆
			(4.223)	(4.578)	(1.284)	(4.275)	(4.386)	(5.056)	(4.973)	(5.058)	(4.977)
		40	17.234	18.553	4.160	17.432	17.829	23.581⋆	25.631⋆	23.622⋆	25.727⋆
			(5.647)	(5.990)	(3.069)	(5.753)	(5.721)	(5.645)	(5.537)	(5.704)	(5.638)
		50	27.526	29.080	9.545	27.883	28.932	33.330	34.875⋆	33.581	35.170⋆
			(5.505)	(5.732)	(4.076)	(5.586)	(5.785)	(5.558)	(5.408)	(5.670)	(5.509)
		60	36.970	38.552	16.876	37.450	38.451	41.634	42.831	42.084	43.249
			(5.372)	(5.631)	(4.679)	(5.432)	(5.365)	(5.238)	(5.054)	(5.309)	(5.167)
		100	60.801	61.885	42.656	61.412	61.622	62.912	63.409	63.518	63.975
			(4.233)	(4.406)	(4.312)	(4.282)	(4.235)	(4.191)	(4.205)	(4.231)	(4.257)
		200	82.488	83.194	71.771	82.971	82.752	83.111	83.189	83.570	83.645
			(3.142)	(3.128)	(3.380)	(3.124)	(3.122)	(3.111)	(3.065)	(3.098)	(3.062)
		300	90.431	90.909	82.763	90.812	90.538	90.737	90.734	91.101	91.096
			(2.470)	(2.464)	(2.711)	(2.422)	(2.462)	(2.442)	(2.452)	(2.419)	(2.439)
		500	96.622	96.817	92.627	96.793	96.648	96.717	96.681	96.867	96.849
			(1.655)	(1.616)	(2.188)	(1.605)	(1.650)	(1.620)	(1.623)	(1.594)	(1.598)
	-0.25	20	1.187	1.386	0.155	1.194	0.599	5.061⋆	7.713⋆	5.061⋆	7.713⋆
			(1.621)	(1.789)	(0.449)	(1.636)	(1.145)	(3.198)	(3.588)	(3.198)	(3.588)
		30	6.671	7.468	1.055	6.723	6.321	13.371 ⋆	15.875⋆	13.371⋆	15.875⋆
			(4.024)	(4.233)	(1.403)	(4.064)	(4.167)	(4.632)	(4.750)	(4.632)	(4.750)
		40	16.801	18.171	3.977	16.988	17.672	23.641⋆	25.605⋆	23.702⋆	25.687⋆
			(5.176)	(5.426)	(2.857)	(5.271)	(5.601)	(5.460)	(5.307)	(5.543)	(5.415)
		50	27.577	29.149	9.691	27.975	29.014	33.392	34.910⋆	33.651	35.167⋆
			(5.466)	(5.551)	(3.944)	(5.540)	(5.453)	(5.273)	(5.159)	(5.374)	(5.273)
		60	37.164	38.518	17.099	37.640	38.541	41.723	42.898	42.102	43.249
			(4.974)	(5.104)	(4.691)	(5.067)	(5.200)	(5.138)	(4.993)	(5.221)	(5.044)
		100	61.270	62.352	42.768	61.870	62.122	63.411	63.799	63.989	64.341
			(4.043)	(4.039)	(4.306)	(4.069)	(4.014)	(3.988)	(3.969)	(4.024)	(3.973)
		200	82.531	83.142	71.629	83.054	82.788	83.180	83.215	83.601	83.625
			(3.012)	(2.967)	(3.304)	(2.975)	(2.962)	(2.925)	(2.914)	(2.929)	(2.955)
		300	90.404	90.811	82.488	90.758	90.483	90.651	90.702	90.989	91.022
			(2.535)	(2.535)	(2.737)	(2.527)	(2.525)	(2.533)	(2.528)	(2.530)	(2.515)
		500	96.631	96.794	92.390	96.793	96.640	96.691	96.681	96.870	96.853
			(1.685)	(1.665)	(2.129)	(1.653)	(1.683)	(1.657)	(1.666)	(1.612)	(1.635)

Table C.33. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.5$. Fifty blocks are utilized with pairwise correlation between variables within a block to be ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.5	20	1.119	1.341	0.120	1.122	0.510	4.787⋆	7.246⋆	4.787⋆	7.246⋆
			(1.576)	(1.785)	(0.384)	(1.580)	(1.091)	(3.224)	(3.767)	(3.224)	(3.767)
		30	6.898	7.907	0.998	6.954	6.231	13.176⋆	15.723⋆	13.176⋆	15.725⋆
			(4.087)	(4.681)	(1.319)	(4.132)	(4.339)	(4.986)	(5.101)	(4.986)	(5.107)
		40	17.122	18.606	4.020	17.325	17.845	23.775⋆	25.967⋆	23.830⋆	26.046⋆
			(5.604)	(6.049)	(2.905)	(5.686)	(5.889)	(5.646)	(5.562)	(5.720)	(5.644)
		50	27.703	29.497	9.680	28.042	28.909	33.242	34.846⋆	33.488⋆	35.108⋆
			(5.541)	(5.874)	(4.112)	(5.642)	(5.781)	(5.499)	(5.473)	(5.640)	(5.580)
		60	37.063	38.801	17.008	37.538	38.445	41.695	42.740	42.105	43.111
			(5.577)	(6.030)	(4.575)	(5.650)	(5.636)	(5.598)	(5.341)	(5.714)	(5.469)
		100	61.006	62.229	42.757	61.681	61.920	63.233	63.601	63.779	64.155
			(4.398)	(4.757)	(4.668)	(4.474)	(4.392)	(4.336)	(4.232)	(4.365)	(4.260)
		200	82.534	83.250	71.669	83.078	82.765	83.147	83.246	83.560	83.665
			(3.242)	(3.368)	(3.470)	(3.244)	(3.232)	(3.202)	(3.149)	(3.178)	(3.189)
		300	90.379	90.837	82.649	90.722	90.452	90.644	90.640	90.994	91.003
			(2.513)	(2.539)	(2.865)	(2.469)	(2.499)	(2.465)	(2.478)	(2.468)	(2.468)
		500	96.719	96.919	92.643	96.887	96.743	96.801	96.790	96.967	96.971
			(1.672)	(1.662)	(2.214)	(1.644)	(1.670)	(1.636)	(1.658)	(1.623)	(1.611)
	-0.5	20	1.234	1.392	0.136	1.239	0.575	5.230⋆	7.889⋆	5.230⋆	7.889⋆
			(1.661)	(1.819)	(0.405)	(1.679)	(1.126)	(3.296)	(3.769)	(3.296)	(3.769)
		30	6.736	7.504	1.038	6.773	6.394	13.502⋆	16.062⋆	13.502⋆	16.066⋆
			(4.089)	(4.258)	(1.366)	(4.127)	(4.216)	(4.941)	(4.883)	(4.941)	(4.895)
		40	17.045	18.325	3.964	17.232	17.968	23.570⋆	25.652⋆	23.645⋆	25.739⋆
			(5.345)	(5.301)	(2.917)	(5.416)	(5.668)	(5.406)	(5.282)	(5.516)	(5.390)
		50	27.695	29.159	9.792	28.011	29.079	33.429	34.950	33.667⋆	35.221⋆
			(5.378)	(5.415)	(3.878)	(5.448)	(5.432)	(5.223)	(5.092)	(5.318)	(5.166)
		60	36.839	38.282	16.914	37.314	38.188	41.438	42.570	41.808	42.980
			(5.146)	(5.114)	(4.591)	(5.227)	(5.036)	(5.048)	(4.906)	(5.131)	(5.004)
		100	60.991	62.134	42.727	61.638	61.916	63.183	63.557	63.764	64.118
			(4.308)	(4.221)	(4.422)	(4.307)	(4.302)	(4.193)	(4.238)	(4.233)	(4.286)
		200	82.454	83.101	71.588	82.939	82.721	83.075	83.195	83.515	83.657
			(3.063)	(3.045)	(3.243)	(3.044)	(3.048)	(3.032)	(3.037)	(3.039)	(3.042)
		300	90.581	91.049	82.772	90.968	90.688	90.875	90.903	91.233	91.270
			(2.466)	(2.415)	(2.748)	(2.420)	(2.426)	(2.440)	(2.421)	(2.396)	(2.390)
		500	96.619	96.828	92.409	96.802	96.658	96.729	96.698	96.900	96.871
			(1.577)	(1.538)	(2.099)	(1.550)	(1.570)	(1.561)	(1.562)	(1.529)	(1.549)

Table C.34. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.75$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.75	20	1.133	1.363	0.157	1.134	0.582	5.045⋆	7.528⋆	5.045⋆	7.530⋆
			(1.653)	(1.953)	(0.452)	(1.656)	(1.180)	(3.485)	(4.362)	(3.485)	(4.369)
		30	6.662	7.916	1.053	6.707	6.179	13.211⋆	15.587⋆	13.216⋆	15.597⋆
			(4.393)	(5.311)	(1.396)	(4.444)	(4.590)	(5.352)	(5.366)	(5.364)	(5.389)
		40	17.068	19.023	4.148	17.230	17.890	23.745⋆	25.768⋆	23.820⋆	25.842⋆
			(5.730)	(6.621)	(3.078)	(5.804)	(5.933)	(5.851)	(5.856)	(5.950)	(5.940)
		50	27.917	30.100	9.674	28.325	29.169	33.442	34.962⋆	33.710⋆	35.254⋆
			(6.177)	(7.129)	(4.185)	(6.314)	(6.418)	(6.393)	(6.183)	(6.535)	(6.335)
		60	37.135	39.220	16.977	37.657	38.679	41.883	43.040	42.263	43.417
			(5.931)	(7.023)	(4.936)	(6.093)	(6.056)	(5.833)	(5.811)	(5.935)	(5.902)
		100	61.209	62.676	43.026	61.854	62.068	63.324	63.772	63.885	64.320
			(4.533)	(5.138)	(4.883)	(4.604)	(4.498)	(4.490)	(4.471)	(4.535)	(4.511)
		200	82.589	83.379	71.568	83.057	82.800	83.181	83.258	83.625	83.705
			(3.411)	(3.592)	(3.667)	(3.400)	(3.345)	(3.342)	(3.308)	(3.325)	(3.284)
		300	90.407	90.948	82.666	90.748	90.471	90.673	90.698	90.999	91.026
			(2.596)	(2.720)	(3.094)	(2.603)	(2.593)	(2.584)	(2.564)	(2.574)	(2.542)
		500	96.603	96.843	92.520	96.795	96.637	96.691	96.685	96.879	96.855
			(1.740)	(1.753)	(2.248)	(1.686)	(1.732)	(1.717)	(1.710)	(1.668)	(1.677)
	-0.75	20	1.232	1.367	0.147	1.236	0.611	5.226⋆	7.711⋆	5.226⋆	7.711⋆
			(1.737)	(1.848)	(0.424)	(1.748)	(1.226)	(3.501)	(4.005)	(3.501)	(4.005)
		30	6.814	7.704	1.061	6.850	6.225	13.298⋆	15.770⋆	13.298⋆	15.772⋆
			(4.052)	(4.192)	(1.395)	(4.083)	(4.152)	(4.705)	(4.821)	(4.705)	(4.825)
		40	16.822	18.145	3.957	17.019	17.599	23.205⋆	25.332⋆	23.248⋆	25.416⋆
			(5.144)	(5.129)	(2.780)	(5.217)	(5.427)	(5.298)	(5.064)	(5.361)	(5.165)
		50	27.625	29.202	9.634	27.969	28.868	33.042	34.657	33.258	34.913⋆
			(5.402)	(5.335)	(3.955)	(5.462)	(5.491)	(5.325)	(5.241)	(5.426)	(5.318)
		60	36.824	38.409	16.748	37.291	38.127	41.289	42.560	41.675	42.924
			(5.141)	(5.116)	(4.528)	(5.201)	(5.029)	(4.975)	(4.880)	(5.019)	(4.957)
		100	61.117	62.264	42.763	61.757	61.999	63.270	63.721	63.839	64.248
			(4.149)	(4.138)	(4.173)	(4.146)	(4.101)	(4.109)	(4.020)	(4.083)	(4.068)
		200	82.510	83.211	71.656	83.006	82.738	83.103	83.171	83.570	83.651
			(2.921)	(2.937)	(3.258)	(2.906)	(2.891)	(2.916)	(2.903)	(2.886)	(2.838)
		300	90.508	90.965	82.751	90.878	90.575	90.784	90.780	91.110	91.090
			(2.337)	(2.327)	(2.645)	(2.330)	(2.347)	(2.347)	(2.363)	(2.342)	(2.352)
		500	96.647	96.860	92.584	96.819	96.665	96.718	96.718	96.894	96.877
			(1.546)	(1.523)	(2.130)	(1.507)	(1.539)	(1.526)	(1.508)	(1.489)	(1.493)

Table C.35. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.9$ and $\rho = \pm 0.9$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.9	0.9	20	1.373	2.006	0.178	1.377	0.670	5.387⋆	7.794⋆	5.388⋆	7.801⋆
			(2.083)	(3.087)	(0.522)	(2.092)	(1.476)	(4.412)	(5.130)	(4.415)	(5.151)
		30	6.903	9.140	1.075	6.962	6.632	13.436⋆	16.041⋆	13.451⋆	16.055⋆
			(5.265)	(7.061)	(1.582)	(5.335)	(5.526)	(6.390)	(6.557)	(6.428)	(6.584)
		40	16.968	19.857	4.094	17.137	17.647	23.708⋆	25.817⋆	23.796⋆	25.926⋆
			(6.695)	(8.824)	(3.288)	(6.814)	(7.333)	(7.335)	(7.281)	(7.447)	(7.403)
		50	27.607	30.611	9.925	27.970	28.854	33.345	34.704	33.601	34.980
			(7.439)	(9.532)	(4.896)	(7.559)	(7.714)	(7.591)	(7.422)	(7.750)	(7.588)
		60	36.906	39.704	17.239	37.358	38.235	41.612	42.674	42.023	43.062
			(7.256)	(9.084)	(5.835)	(7.335)	(7.407)	(7.159)	(7.053)	(7.325)	(7.196)
		100	60.927	62.946	42.483	61.520	61.803	63.113	63.498	63.692	63.992
			(5.646)	(7.015)	(5.616)	(5.688)	(5.641)	(5.499)	(5.491)	(5.538)	(5.525)
		200	82.623	83.840	71.698	83.091	82.893	83.232	83.257	83.700	83.719
			(3.749)	(4.384)	(4.156)	(3.754)	(3.688)	(3.692)	(3.672)	(3.704)	(3.667)
		300	90.392	91.068	82.685	90.728	90.476	90.630	90.696	90.980	91.001
			(2.879)	(3.183)	(3.386)	(2.864)	(2.860)	(2.868)	(2.853)	(2.859)	(2.850)
		500	96.623	96.936	92.553	96.809	96.651	96.710	96.714	96.889	96.894
			(1.792)	(1.834)	(2.361)	(1.749)	(1.784)	(1.775)	(1.760)	(1.747)	(1.749)
	-0.9	20	1.218	1.412	0.142	1.221	0.649	5.130⋆	7.516⋆	5.130⋆	7.516⋆
			(1.747)	(1.931)	(0.422)	(1.751)	(1.387)	(3.606)	(3.924)	(3.606)	(3.924)
		30	6.892	7.751	0.920	6.929	6.332	13.337⋆	15.732⋆	13.339⋆	15.743⋆
			(4.413)	(4.550)	(1.310)	(4.443)	(4.624)	(5.069)	(5.146)	(5.074)	(5.168)
		40	16.864	18.693	3.990	17.059	17.473	23.557	25.642⋆	23.627	25.728⋆
			(5.562)	(5.388)	(3.012)	(5.670)	(5.740)	(5.646)	(5.503)	(5.745)	(5.597)
		50	27.367	29.425	9.539	27.748	28.625	33.103	34.663	33.345	34.913
			(5.462)	(5.406)	(4.083)	(5.592)	(5.539)	(5.416)	(5.189)	(5.544)	(5.305)
		60	37.021	38.910	16.727	37.551	38.479	41.663	42.837	42.055	43.244
			(5.292)	(5.318)	(4.646)	(5.390)	(5.204)	(5.056)	(4.991)	(5.101)	(5.098)
		100	61.132	62.611	42.725	61.793	62.002	63.342	63.728	63.911	64.263
			(4.085)	(4.181)	(4.401)	(4.007)	(3.968)	(3.874)	(3.852)	(3.880)	(3.830)
		200	82.525	83.262	71.594	83.008	82.734	83.111	83.196	83.583	83.634
			(2.782)	(2.835)	(3.034)	(2.801)	(2.791)	(2.788)	(2.776)	(2.792)	(2.797)
		300	90.395	90.957	82.746	90.798	90.499	90.676	90.696	91.038	91.042
			(2.284)	(2.333)	(2.753)	(2.240)	(2.251)	(2.245)	(2.229)	(2.201)	(2.197)
		500	96.637	96.849	92.444	96.816	96.663	96.700	96.678	96.862	96.858
			(1.577)	(1.592)	(2.021)	(1.562)	(1.574)	(1.560)	(1.569)	(1.548)	(1.545)

Table C.36. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.1	20	2.021	2.530	0.198	2.029	0.492	6.747⋆	9.010⋆	6.747⋆	9.010⋆
			(2.343)	(2.771)	(0.499)	(2.364)	(1.083)	(3.971)	(4.438)	(3.971)	(4.438)
		30	11.585	13.566	1.726	11.694	7.975	18.519⋆	20.418⋆	18.535⋆	20.442⋆
			(5.374)	(5.883)	(1.905)	(5.421)	(5.180)	(5.883)	(5.868)	(5.918)	(5.908)
		40	25.840	28.588	6.898	26.252	23.555	32.086⋆	33.448⋆	32.415⋆	33.822⋆
			(6.264)	(6.681)	(3.589)	(6.369)	(6.336)	(6.504)	(6.222)	(6.745)	(6.471)
		50	39.762	42.824	15.535	40.519	38.841	45.302	46.029	45.958	46.670
			(6.062)	(6.464)	(5.048)	(6.216)	(6.484)	(6.236)	(6.252)	(6.376)	(6.394)
		60	51.119	54.083	24.574	52.068	50.713	55.623	56.300	56.469	57.142
			(5.827)	(6.296)	(5.056)	(5.972)	(5.926)	(5.824)	(5.921)	(6.068)	(6.057)
		100	81.520	84.113	55.982	82.825	81.944	84.156	84.318	85.409	85.553
			(5.391)	(5.548)	(5.057)	(5.464)	(5.356)	(5.244)	(5.307)	(5.275)	(5.349)
		200	114.122	115.983	94.252	115.498	114.493	115.209	115.291	116.546	116.549
			(4.509)	(4.541)	(4.451)	(4.550)	(4.527)	(4.477)	(4.458)	(4.454)	(4.426)
		300	128.944	130.258	112.914	130.027	129.153	129.470	129.473	130.573	130.595
			(3.592)	(3.662)	(4.060)	(3.591)	(3.608)	(3.578)	(3.572)	(3.576)	(3.544)
		500	141.357	142.038	132.092	141.947	141.435	141.554	141.544	142.155	142.132
			(2.466)	(2.402)	(3.228)	(2.373)	(2.453)	(2.454)	(2.407)	(2.386)	(2.383)
	-0.1	20	2.141	2.697	0.249	2.146	0.543	6.981⋆	9.127⋆	6.981⋆	9.127⋆
			(2.397)	(2.839)	(0.588)	(2.407)	(1.160)	(4.052)	(4.537)	(4.052)	(4.537)
		30	11.761	13.758	1.713	11.884	7.995	18.719⋆	20.716⋆	18.728⋆	20.758⋆
			(5.303)	(5.635)	(1.832)	(5.392)	(5.117)	(5.721)	(5.668)	(5.739)	(5.748)
		40	25.900	28.448	6.830	26.251	23.553	32.037	33.393⋆	32.383⋆	33.683⋆
			(6.093)	(6.469)	(3.821)	(6.201)	(6.558)	(6.385)	(6.279)	(6.570)	(6.452)
		50	39.926	42.839	15.637	40.635	39.030	45.462	46.297	46.065	46.929
			(6.103)	(6.487)	(4.720)	(6.234)	(6.243)	(6.145)	(6.049)	(6.329)	(6.193)
		60	51.068	54.015	24.678	52.036	50.666	55.773	56.383	56.700	57.204
			(6.195)	(6.472)	(5.188)	(6.315)	(6.283)	(5.962)	(6.018)	(6.173)	(6.174)
		100	81.534	84.132	55.904	82.909	81.991	84.238	84.398	85.495	85.663
			(5.312)	(5.286)	(4.648)	(5.365)	(5.302)	(5.315)	(5.200)	(5.369)	(5.314)
		200	114.344	116.249	94.012	115.791	114.742	115.498	115.517	116.905	116.918
			(4.481)	(4.483)	(4.588)	(4.451)	(4.453)	(4.433)	(4.388)	(4.423)	(4.406)
		300	129.068	130.407	113.112	130.241	129.313	129.674	129.670	130.785	130.790
			(3.519)	(3.514)	(4.116)	(3.490)	(3.505)	(3.501)	(3.460)	(3.465)	(3.463)
		500	141.350	142.056	132.310	142.033	141.487	141.607	141.615	142.208	142.193
			(2.527)	(2.429)	(3.227)	(2.416)	(2.526)	(2.471)	(2.472)	(2.404)	(2.387)

Table C.37. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.25$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.25	20	2.073	2.608	0.208	2.077	0.508	6.890⋆	9.173⋆	6.890⋆	9.173⋆
			(2.343)	(2.693)	(0.570)	(2.353)	(1.129)	(4.098)	(4.503)	(4.098)	(4.503)
		30	11.613	13.557	1.840	11.736	8.120	18.465⋆	20.453⋆	18.494⋆	20.500⋆
			(5.464)	(5.987)	(1.921)	(5.548)	(5.160)	(5.982)	(5.897)	(6.046)	(5.977)
		40	25.785	28.484	6.760	26.161	23.659	32.161⋆	33.508⋆	32.463⋆	33.818⋆
			(6.159)	(6.601)	(3.722)	(6.307)	(6.556)	(6.319)	(6.275)	(6.480)	(6.438)
		50	39.850	42.738	15.438	40.573	38.705	45.118	46.081	45.762	46.675
			(6.399)	(6.806)	(4.660)	(6.526)	(6.518)	(6.374)	(6.231)	(6.507)	(6.417)
		60	51.121	54.023	24.745	52.054	50.892	55.750	56.496	56.663	57.406
			(6.214)	(6.699)	(5.242)	(6.339)	(6.346)	(6.156)	(6.078)	(6.323)	(6.206)
		100	81.631	84.411	55.723	82.975	82.001	84.184	84.359	85.517	85.695
			(5.283)	(5.700)	(4.968)	(5.368)	(5.305)	(5.387)	(5.290)	(5.537)	(5.446)
		200	114.347	116.260	93.877	115.750	114.736	115.507	115.486	116.795	116.837
			(4.469)	(4.744)	(4.347)	(4.501)	(4.505)	(4.441)	(4.431)	(4.494)	(4.429)
		300	129.030	130.478	113.101	130.188	129.290	129.621	129.620	130.712	130.775
			(3.673)	(3.830)	(4.206)	(3.676)	(3.652)	(3.661)	(3.648)	(3.635)	(3.639)
		500	141.387	142.060	132.119	142.022	141.486	141.610	141.582	142.173	142.207
			(2.583)	(2.577)	(3.171)	(2.541)	(2.584)	(2.569)	(2.569)	(2.523)	(2.501)
	-0.25	20	2.137	2.646	0.235	2.142	0.558	7.062⋆	9.185⋆	7.062⋆	9.185⋆
			(2.409)	(2.795)	(0.564)	(2.418)	(1.115)	(4.009)	(4.335)	(4.009)	(4.335)
		30	11.795	13.759	1.845	11.965	8.162	18.644⋆	20.732⋆	18.660⋆	20.758⋆
			(5.318)	(5.716)	(1.896)	(5.391)	(5.198)	(5.769)	(5.793)	(5.800)	(5.841)
		40	26.164	28.851	6.778	26.575	23.823	32.514	33.838⋆	32.828	34.170⋆
			(6.250)	(6.399)	(3.612)	(6.390)	(6.651)	(6.253)	(6.235)	(6.443)	(6.389)
		50	39.956	42.809	15.489	40.659	38.777	45.217	46.148	45.865	46.730
			(6.157)	(6.337)	(4.770)	(6.323)	(6.393)	(6.082)	(5.993)	(6.252)	(6.174)
		60	50.973	54.143	24.648	51.899	50.659	55.830	56.384	56.657	57.251
			(5.855)	(6.134)	(4.863)	(6.015)	(5.931)	(5.885)	(5.808)	(5.978)	(5.918)
		100	81.432	84.090	56.010	82.790	81.950	84.214	84.366	85.493	85.647
			(5.201)	(5.237)	(5.035)	(5.197)	(5.245)	(5.143)	(5.080)	(5.250)	(5.225)
		200	114.317	116.214	93.889	115.735	114.702	115.472	115.444	116.790	116.767
			(4.403)	(4.368)	(4.357)	(4.355)	(4.401)	(4.314)	(4.320)	(4.318)	(4.335)
		300	129.085	130.416	112.883	130.214	129.371	129.685	129.695	130.705	130.745
			(3.543)	(3.469)	(4.036)	(3.460)	(3.543)	(3.511)	(3.453)	(3.427)	(3.420)
		500	141.491	142.132	132.213	142.097	141.579	141.700	141.705	142.277	142.304
			(2.522)	(2.429)	(3.167)	(2.452)	(2.508)	(2.490)	(2.493)	(2.442)	(2.435)

Table C.38. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.5$. Fifty blocks are utilized with pairwise correlation between variables within a block to be ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.5	20	2.046	2.585	0.240	2.054	0.522	6.933*	9.083*	6.933*	9.083*
			(2.390)	(2.884)	(0.565)	(2.400)	(1.177)	(4.165)	(4.679)	(4.165)	(4.679)
		30	11.726	13.787	1.735	11.833	7.911	18.522*	20.314*	18.549*	20.371*
			(5.543)	(6.375)	(1.885)	(5.596)	(5.178)	(6.277)	(6.213)	(6.331)	(6.311)
		40	25.966	28.795	6.954	26.347	23.558	32.382	33.714*	32.699	34.058*
			(6.593)	(7.310)	(3.970)	(6.721)	(6.955)	(6.787)	(6.854)	(7.008)	(7.040)
		50	39.758	42.916	15.481	40.421	38.635	45.231	46.107	45.842	46.795
			(6.450)	(7.259)	(5.044)	(6.571)	(6.719)	(6.471)	(6.581)	(6.698)	(6.738)
		60	50.947	54.098	24.773	51.866	50.598	55.573	56.208	56.467	57.028
			(6.536)	(7.213)	(5.429)	(6.632)	(6.758)	(6.527)	(6.527)	(6.643)	(6.666)
		100	81.412	84.258	55.595	82.770	81.832	84.064	84.266	85.325	85.473
			(5.520)	(6.320)	(5.081)	(5.596)	(5.540)	(5.512)	(5.583)	(5.710)	(5.712)
		200	114.534	116.511	94.163	115.979	114.913	115.637	115.692	117.000	117.000
			(4.659)	(5.002)	(4.539)	(4.690)	(4.655)	(4.661)	(4.650)	(4.685)	(4.661)
		300	129.078	130.474	113.138	130.125	129.297	129.638	129.593	130.624	130.582
			(3.751)	(3.938)	(4.118)	(3.727)	(3.744)	(3.710)	(3.710)	(3.677)	(3.700)
		500	141.452	142.101	132.066	142.042	141.559	141.673	141.634	142.264	142.221
			(2.519)	(2.521)	(3.226)	(2.441)	(2.507)	(2.492)	(2.482)	(2.412)	(2.427)
	-0.5	20	2.123	2.618	0.224	2.132	0.526	7.022*	9.295*	7.022*	9.295*
			(2.310)	(2.637)	(0.546)	(2.329)	(1.146)	(3.998)	(4.426)	(3.998)	(4.426)
		30	11.796	13.651	1.764	11.924	7.964	18.576*	20.704*	18.599*	20.759*
			(5.417)	(5.591)	(1.848)	(5.479)	(5.175)	(5.834)	(5.619)	(5.880)	(5.717)
		40	26.207	28.842	6.962	26.624	23.941	32.816*	33.886*	33.081*	34.213*
			(6.171)	(6.385)	(3.635)	(6.287)	(6.571)	(6.555)	(6.401)	(6.724)	(6.591)
		50	39.759	42.583	15.417	40.445	38.818	45.288	46.131	45.916	46.795
			(6.197)	(6.144)	(4.758)	(6.313)	(6.410)	(6.035)	(6.227)	(6.220)	(6.378)
		60	50.858	53.896	24.484	51.803	50.566	55.524	56.245	56.396	57.099
			(6.016)	(6.183)	(5.053)	(6.218)	(6.176)	(6.072)	(6.031)	(6.201)	(6.156)
		100	81.784	84.465	56.018	83.140	82.199	84.378	84.572	85.666	85.836
			(5.364)	(5.434)	(4.908)	(5.462)	(5.475)	(5.451)	(5.428)	(5.562)	(5.516)
		200	114.484	116.341	94.166	115.870	114.868	115.570	115.640	116.955	116.943
			(4.204)	(4.218)	(4.294)	(4.213)	(4.236)	(4.147)	(4.175)	(4.188)	(4.188)
		300	128.914	130.274	113.046	130.056	129.149	129.549	129.521	130.629	130.594
			(3.488)	(3.500)	(3.981)	(3.485)	(3.542)	(3.499)	(3.512)	(3.435)	(3.463)
		500	141.378	142.039	132.017	141.985	141.458	141.576	141.577	142.170	142.176
			(2.507)	(2.468)	(3.130)	(2.437)	(2.489)	(2.498)	(2.463)	(2.419)	(2.403)

Table C.39. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.75$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.75	20	2.186	3.016	0.242	2.195	0.580	7.055⋆	9.261⋆	7.056⋆	9.264⋆
			(2.679)	(3.557)	(0.586)	(2.696)	(1.281)	(4.802)	(5.267)	(4.805)	(5.277)
		30	11.662	14.325	1.750	11.770	8.002	18.599⋆	20.670⋆	18.627⋆	20.730⋆
			(6.336)	(7.778)	(2.012)	(6.402)	(5.934)	(7.090)	(6.974)	(7.144)	(7.066)
		40	26.459	30.031	7.191	26.883	24.188	32.675	34.064⋆	33.009	34.402⋆
			(7.109)	(8.717)	(4.217)	(7.325)	(7.601)	(7.567)	(7.507)	(7.808)	(7.687)
		50	39.825	43.531	15.132	40.501	38.740	45.189	46.106	45.866	46.708
			(7.338)	(9.131)	(5.227)	(7.514)	(7.625)	(7.537)	(7.608)	(7.787)	(7.791)
		60	51.434	55.120	24.851	52.321	51.084	55.986	56.529	56.809	57.358
			(7.053)	(8.753)	(5.716)	(7.238)	(7.217)	(7.138)	(7.033)	(7.358)	(7.269)
		100	81.602	84.691	55.957	82.896	82.024	84.187	84.347	85.461	85.524
			(6.054)	(7.630)	(5.746)	(6.199)	(6.249)	(6.144)	(6.134)	(6.299)	(6.298)
		200	114.483	116.713	94.207	115.836	114.835	115.495	115.602	116.863	116.900
			(4.942)	(5.797)	(5.200)	(4.966)	(4.952)	(4.883)	(4.886)	(4.931)	(4.921)
		300	129.264	130.845	113.190	130.413	129.474	129.857	129.822	130.927	130.927
			(3.951)	(4.395)	(4.520)	(3.866)	(3.903)	(3.866)	(3.889)	(3.842)	(3.852)
		500	141.443	142.113	132.145	142.052	141.538	141.635	141.662	142.250	142.239
			(2.616)	(2.770)	(3.428)	(2.540)	(2.610)	(2.610)	(2.595)	(2.517)	(2.563)
	-0.75	20	2.074	2.446	0.255	2.077	0.561	6.978⋆	9.305⋆	6.978⋆	9.305⋆
			(2.269)	(2.470)	(0.604)	(2.275)	(1.165)	(4.110)	(4.563)	(4.110)	(4.563)
		30	11.879	13.801	1.782	12.003	8.150	18.635⋆	20.385⋆	18.659⋆	20.418⋆
			(5.479)	(5.690)	(1.989)	(5.535)	(5.463)	(6.011)	(6.097)	(6.060)	(6.154)
		40	25.981	28.667	6.827	26.391	23.851	32.376⋆	33.734⋆	32.694⋆	34.086⋆
			(6.131)	(6.222)	(3.752)	(6.250)	(6.526)	(6.355)	(6.275)	(6.519)	(6.450)
		50	39.746	42.543	15.441	40.420	38.639	45.113	46.081	45.731	46.686
			(6.418)	(6.367)	(5.019)	(6.541)	(6.577)	(6.496)	(6.279)	(6.620)	(6.421)
		60	51.397	54.521	24.806	52.291	51.188	56.049	56.700	56.950	57.555
			(6.142)	(6.061)	(5.018)	(6.236)	(6.212)	(6.172)	(6.120)	(6.335)	(6.216)
		100	81.462	84.143	55.734	82.761	81.872	83.985	84.294	85.314	85.477
			(5.201)	(5.329)	(5.127)	(5.315)	(5.252)	(5.229)	(5.188)	(5.328)	(5.306)
		200	114.656	116.522	94.155	116.011	115.017	115.784	115.750	117.117	117.093
			(4.272)	(4.338)	(4.223)	(4.350)	(4.335)	(4.292)	(4.265)	(4.366)	(4.309)
		300	129.243	130.551	113.215	130.340	129.461	129.807	129.804	130.891	130.915
			(3.578)	(3.549)	(3.968)	(3.512)	(3.559)	(3.502)	(3.486)	(3.464)	(3.469)
		500	141.487	142.216	132.139	142.155	141.620	141.719	141.711	142.328	142.327
			(2.342)	(2.338)	(3.204)	(2.318)	(2.332)	(2.329)	(2.330)	(2.304)	(2.300)

Table C.40. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.85$ and $\rho = \pm 0.9$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.85	0.9	20	2.199	3.510	0.234	2.212	0.602	7.084⋆	9.126⋆	7.085⋆	9.139⋆
			(3.033)	(4.981)	(0.619)	(3.073)	(1.493)	(5.138)	(5.903)	(5.140)	(5.948)
		30	11.739	15.471	1.850	11.927	8.260	18.852⋆	20.677⋆	18.915⋆	20.744⋆
			(7.244)	(10.368)	(2.351)	(7.380)	(6.982)	(8.516)	(8.549)	(8.627)	(8.651)
		40	26.056	30.674	6.951	26.449	23.664	32.405	33.473⋆	32.699	33.822⋆
			(8.412)	(11.331)	(4.796)	(8.577)	(8.835)	(8.757)	(8.725)	(8.992)	(8.970)
		50	39.325	44.450	15.423	39.990	38.276	44.926	45.712	45.591	46.349
			(8.542)	(12.494)	(6.141)	(8.808)	(9.030)	(9.049)	(9.065)	(9.318)	(9.290)
		60	51.131	56.136	24.645	52.034	50.765	55.752	56.187	56.601	57.041
			(8.638)	(12.303)	(6.993)	(8.826)	(9.009)	(8.912)	(8.956)	(9.145)	(9.185)
		100	81.552	85.587	56.003	82.877	81.910	84.171	84.317	85.476	85.635
			(7.465)	(10.304)	(6.518)	(7.653)	(7.514)	(7.463)	(7.420)	(7.649)	(7.607)
		200	114.318	117.048	93.993	115.641	114.613	115.389	115.463	116.663	116.753
			(5.564)	(7.293)	(5.687)	(5.661)	(5.610)	(5.616)	(5.565)	(5.649)	(5.657)
		300	128.823	130.675	112.985	129.985	129.097	129.440	129.469	130.518	130.574
			(4.421)	(5.503)	(4.998)	(4.398)	(4.420)	(4.434)	(4.387)	(4.386)	(4.317)
		500	141.190	142.109	131.983	141.858	141.316	141.437	141.405	142.049	142.055
			(2.886)	(3.215)	(3.775)	(2.783)	(2.876)	(2.856)	(2.831)	(2.756)	(2.756)
	-0.9	20	2.067	2.528	0.226	2.072	0.599	6.938⋆	9.090⋆	6.938⋆	9.094⋆
			(2.554)	(2.797)	(0.594)	(2.561)	(1.402)	(4.482)	(4.982)	(4.482)	(4.993)
		30	11.854	14.016	1.771	11.969	8.103	18.721⋆	20.573⋆	18.756⋆	20.623⋆
			(5.440)	(5.434)	(2.043)	(5.552)	(5.484)	(6.156)	(6.104)	(6.219)	(6.178)
		40	26.259	29.231	7.003	26.661	24.000	32.565	33.796	32.847	34.145⋆
			(6.554)	(6.258)	(3.853)	(6.643)	(6.842)	(6.631)	(6.417)	(6.807)	(6.631)
		50	39.727	43.140	15.379	40.458	38.716	45.205	46.062	45.894	46.778
			(6.298)	(6.344)	(5.064)	(6.470)	(6.510)	(6.348)	(6.226)	(6.591)	(6.449)
		60	51.277	54.907	24.745	52.225	50.814	55.908	56.421	56.815	57.253
			(6.209)	(6.123)	(5.512)	(6.288)	(6.356)	(6.050)	(5.953)	(6.217)	(6.097)
		100	81.883	84.787	55.986	83.200	82.230	84.365	84.574	85.647	85.734
			(5.272)	(5.465)	(5.060)	(5.338)	(5.344)	(5.301)	(5.401)	(5.428)	(5.439)
		200	114.523	116.556	94.233	115.918	114.915	115.569	115.641	116.893	116.944
			(4.227)	(4.443)	(4.234)	(4.252)	(4.218)	(4.182)	(4.232)	(4.191)	(4.284)
		300	129.008	130.429	113.081	130.090	129.212	129.589	129.587	130.628	130.647
			(3.575)	(3.758)	(3.968)	(3.564)	(3.571)	(3.564)	(3.536)	(3.535)	(3.536)
		500	141.395	142.167	132.112	142.078	141.499	141.623	141.638	142.251	142.241
			(2.469)	(2.521)	(3.142)	(2.395)	(2.459)	(2.451)	(2.451)	(2.369)	(2.367)

Table C.41. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.1	20	5.199	7.098	0.465	5.233	1.781	14.045⋆	18.310⋆	14.053⋆	18.347⋆
			(4.243)	(5.213)	(0.881)	(4.293)	(2.569)	(6.147)	(6.293)	(6.165)	(6.367)
		30	25.094	30.083	4.324	25.562	20.860	35.381	38.144⋆	35.867⋆	38.703⋆
			(7.162)	(8.155)	(3.267)	(7.345)	(7.602)	(7.601)	(7.404)	(7.862)	(7.660)
		40	48.256	54.087	15.231	49.440	45.821	56.402	57.803	57.735	59.099
			(7.689)	(8.571)	(5.193)	(7.945)	(8.145)	(7.760)	(7.752)	(8.073)	(7.838)
		50	67.233	73.179	30.024	69.073	66.025	73.829	74.540	75.675	76.321
			(7.377)	(8.098)	(5.933)	(7.608)	(7.657)	(7.411)	(7.314)	(7.625)	(7.636)
		60	82.336	88.053	43.716	84.393	81.815	87.687	88.110	89.906	90.106
			(7.139)	(7.762)	(6.066)	(7.401)	(7.133)	(7.188)	(7.015)	(7.431)	(7.288)
		100	121.237	125.916	86.128	123.958	121.529	124.156	124.120	126.772	126.662
			(5.818)	(6.261)	(5.923)	(5.928)	(5.969)	(5.841)	(5.851)	(5.955)	(5.909)
		200	160.614	163.674	135.848	162.965	161.006	161.767	161.803	164.177	164.049
			(4.620)	(4.744)	(4.848)	(4.629)	(4.629)	(4.581)	(4.542)	(4.595)	(4.625)
		300	176.837	178.869	158.025	178.665	177.107	177.489	177.443	179.284	179.227
			(3.788)	(3.869)	(4.450)	(3.746)	(3.812)	(3.767)	(3.816)	(3.680)	(3.691)
		500	190.382	191.513	179.322	191.480	190.505	190.646	190.639	191.703	191.662
			(2.705)	(2.628)	(3.496)	(2.569)	(2.677)	(2.650)	(2.649)	(2.571)	(2.575)
	-0.1	20	5.312	7.127	0.475	5.359	1.701	14.491⋆	18.785⋆	14.498⋆	18.825⋆
			(4.220)	(5.214)	(0.868)	(4.290)	(2.641)	(6.038)	(6.352)	(6.055)	(6.425)
		30	25.545	30.452	4.320	26.088	21.059	35.512⋆	38.277⋆	36.060⋆	38.825⋆
			(7.132)	(7.840)	(3.298)	(7.330)	(7.573)	(7.129)	(7.090)	(7.403)	(7.373)
		40	48.010	53.803	15.271	49.224	45.918	56.380	57.761	57.718	58.971
			(7.119)	(8.091)	(5.239)	(7.337)	(7.730)	(7.497)	(7.509)	(7.836)	(7.701)
		50	67.346	73.108	30.043	69.160	66.054	73.798	74.561	75.630	76.299
			(7.146)	(7.696)	(5.963)	(7.338)	(7.383)	(7.249)	(7.204)	(7.481)	(7.446)
		60	82.275	88.088	44.021	84.446	81.690	87.848	88.060	90.009	90.140
			(7.126)	(7.472)	(6.138)	(7.324)	(7.415)	(7.169)	(7.089)	(7.320)	(7.352)
		100	121.415	126.032	86.681	124.077	121.574	124.228	124.161	126.920	126.740
			(5.820)	(6.123)	(5.869)	(5.996)	(5.918)	(5.851)	(5.792)	(5.928)	(5.943)
		200	160.581	163.570	135.571	162.934	160.955	161.713	161.695	164.077	163.984
			(4.363)	(4.493)	(4.554)	(4.378)	(4.377)	(4.387)	(4.392)	(4.434)	(4.408)
		300	176.823	179.014	158.077	178.664	177.068	177.435	177.435	179.217	179.213
			(3.823)	(3.802)	(4.321)	(3.818)	(3.819)	(3.798)	(3.826)	(3.838)	(3.796)
		500	190.415	191.504	179.370	191.467	190.547	190.652	190.665	191.686	191.661
			(2.695)	(2.608)	(3.323)	(2.583)	(2.703)	(2.660)	(2.682)	(2.563)	(2.560)

Table C.42. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.25$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.25	20	4.986	6.896	0.414	5.027	1.814	14.541⋆	18.752⋆	14.545⋆	18.803⋆
			(4.170)	(5.176)	(0.788)	(4.217)	(2.767)	(6.218)	(6.445)	(6.226)	(6.550)
		30	25.337	30.108	4.550	25.836	20.978	35.370	38.007⋆	35.934	38.596⋆
			(7.662)	(8.678)	(3.461)	(7.839)	(7.846)	(7.815)	(7.624)	(8.100)	(7.850)
		40	48.074	53.700	15.211	49.206	45.706	56.369	57.879	57.647	59.119
			(7.571)	(8.665)	(5.498)	(7.868)	(7.940)	(7.408)	(7.357)	(7.730)	(7.650)
		50	66.826	72.996	29.730	68.600	65.855	73.503	74.253	75.385	75.952
			(7.596)	(8.501)	(6.095)	(7.812)	(7.875)	(7.553)	(7.503)	(7.863)	(7.752)
		60	82.229	88.151	43.717	84.497	81.742	87.712	88.001	89.933	90.042
			(7.154)	(7.773)	(6.076)	(7.289)	(7.290)	(7.067)	(6.918)	(7.231)	(7.198)
		100	121.301	126.023	86.388	123.998	121.504	124.155	124.093	126.857	126.697
			(6.134)	(6.881)	(5.903)	(6.299)	(6.245)	(6.154)	(6.121)	(6.326)	(6.328)
		200	160.601	163.640	135.346	162.938	160.958	161.759	161.676	164.057	164.001
			(4.763)	(4.994)	(5.063)	(4.798)	(4.711)	(4.641)	(4.631)	(4.774)	(4.728)
		300	176.771	178.749	157.940	178.546	177.021	177.380	177.364	179.116	179.075
			(3.820)	(3.876)	(4.337)	(3.809)	(3.791)	(3.800)	(3.799)	(3.743)	(3.739)
		500	190.380	191.454	179.368	191.447	190.521	190.624	190.608	191.675	191.625
			(2.797)	(2.696)	(3.444)	(2.631)	(2.761)	(2.759)	(2.764)	(2.604)	(2.615)
	-0.25	20	5.365	7.245	0.486	5.407	1.799	14.695⋆	18.816⋆	14.710⋆	18.855⋆
			(4.151)	(5.040)	(0.895)	(4.200)	(2.592)	(5.896)	(6.125)	(5.936)	(6.204)
		30	25.215	30.128	4.360	25.765	20.463	35.455⋆	38.318⋆	35.957⋆	38.903⋆
			(7.297)	(7.813)	(3.198)	(7.482)	(7.527)	(7.314)	(7.148)	(7.584)	(7.419)
		40	48.191	53.974	15.459	49.327	45.829	56.396	57.711	57.700	58.982
			(7.290)	(7.718)	(5.389)	(7.540)	(7.507)	(7.270)	(7.190)	(7.547)	(7.476)
		50	66.876	72.718	30.023	68.686	65.766	73.332	74.085	75.084	75.762
			(6.903)	(7.446)	(5.687)	(7.226)	(7.100)	(7.143)	(6.891)	(7.279)	(7.133)
		60	82.395	87.906	43.869	84.489	81.873	87.871	88.189	90.031	90.242
			(6.673)	(7.073)	(6.057)	(6.914)	(6.860)	(6.802)	(6.636)	(7.015)	(6.922)
		100	121.696	126.201	86.358	124.340	121.885	124.448	124.410	127.082	126.930
			(5.713)	(5.894)	(5.500)	(5.788)	(5.857)	(5.786)	(5.811)	(5.820)	(5.865)
		200	160.754	163.858	135.861	163.149	161.180	161.916	161.844	164.282	164.244
			(4.455)	(4.426)	(4.721)	(4.414)	(4.486)	(4.419)	(4.476)	(4.454)	(4.486)
		300	176.917	179.035	157.889	178.791	177.192	177.552	177.533	179.313	179.317
			(3.806)	(3.794)	(4.302)	(3.781)	(3.805)	(3.790)	(3.800)	(3.737)	(3.750)
		500	190.572	191.636	179.409	191.593	190.709	190.799	190.826	191.823	191.805
			(2.646)	(2.616)	(3.314)	(2.566)	(2.657)	(2.634)	(2.629)	(2.576)	(2.586)

Table C.43. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.5$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.5	20	5.388	7.427	0.486	5.439	1.846	14.730⋆	19.054⋆	14.743⋆	19.138⋆
			(4.477)	(5.767)	(0.965)	(4.527)	(2.687)	(6.571)	(6.940)	(6.600)	(7.088)
		30	25.144	30.517	4.350	25.653	20.982	35.416	38.173⋆	35.912	38.799⋆
			(7.369)	(9.382)	(3.351)	(7.566)	(8.030)	(7.919)	(7.837)	(8.217)	(8.168)
		40	47.952	54.203	15.515	49.120	45.829	56.318	57.721	57.602	58.958
			(8.237)	(9.910)	(5.769)	(8.454)	(8.645)	(8.397)	(8.351)	(8.720)	(8.576)
		50	66.768	72.731	29.639	68.511	65.650	73.282	74.054	75.166	75.732
			(7.905)	(9.489)	(6.238)	(8.218)	(7.920)	(7.917)	(7.841)	(8.284)	(8.143)
		60	82.267	88.132	43.702	84.456	81.734	87.611	87.942	89.709	90.029
			(7.250)	(8.914)	(6.652)	(7.464)	(7.602)	(7.596)	(7.420)	(7.829)	(7.627)
		100	120.993	125.749	85.996	123.686	121.272	123.998	123.888	126.609	126.426
			(6.363)	(7.486)	(5.989)	(6.633)	(6.430)	(6.500)	(6.399)	(6.616)	(6.582)
		200	160.515	163.640	135.413	162.872	160.875	161.682	161.590	163.964	163.886
			(4.948)	(5.590)	(4.958)	(4.980)	(5.000)	(4.972)	(4.974)	(4.959)	(4.942)
		300	176.797	178.900	157.896	178.580	177.065	177.391	177.398	179.125	179.148
			(3.932)	(4.117)	(4.538)	(3.841)	(3.927)	(3.854)	(3.901)	(3.782)	(3.811)
		500	190.366	191.506	179.605	191.449	190.536	190.628	190.627	191.665	191.661
			(2.701)	(2.678)	(3.573)	(2.543)	(2.668)	(2.647)	(2.638)	(2.529)	(2.510)
	-0.5	20	4.909	6.660	0.472	4.960	1.582	14.155⋆	18.482⋆	14.164⋆	18.508⋆
			(4.053)	(4.838)	(0.888)	(4.108)	(2.383)	(5.994)	(6.088)	(6.018)	(6.143)
		30	25.458	30.235	4.326	25.997	21.211	35.660	38.349⋆	36.141	38.867⋆
			(6.984)	(7.451)	(3.370)	(7.171)	(7.634)	(7.436)	(7.241)	(7.702)	(7.433)
		40	48.090	53.863	15.494	49.355	45.984	56.400	57.827	57.658	58.972
			(7.176)	(7.406)	(5.225)	(7.466)	(7.512)	(7.181)	(7.117)	(7.346)	(7.353)
		50	67.034	72.824	29.546	68.722	65.829	73.726	74.300	75.427	76.062
			(7.080)	(7.252)	(6.150)	(7.329)	(7.394)	(7.013)	(7.041)	(7.227)	(7.236)
		60	82.333	88.049	43.835	84.463	81.703	87.796	87.994	89.875	90.101
			(6.439)	(6.548)	(5.884)	(6.624)	(6.736)	(6.607)	(6.610)	(6.816)	(6.817)
		100	121.065	125.509	86.654	123.651	121.113	123.812	123.793	126.486	126.319
			(5.813)	(5.985)	(5.563)	(5.944)	(5.897)	(5.776)	(5.850)	(5.922)	(5.916)
		200	160.270	163.342	135.399	162.761	160.626	161.402	161.398	163.852	163.757
			(4.655)	(4.677)	(5.035)	(4.687)	(4.687)	(4.656)	(4.695)	(4.690)	(4.742)
		300	176.798	178.998	157.782	178.706	177.097	177.462	177.465	179.269	179.242
			(3.587)	(3.589)	(4.284)	(3.584)	(3.577)	(3.635)	(3.595)	(3.635)	(3.641)
		500	190.393	191.542	179.392	191.484	190.517	190.622	190.627	191.699	191.675
			(2.738)	(2.637)	(3.527)	(2.630)	(2.731)	(2.706)	(2.716)	(2.638)	(2.605)

Table C.44. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.75$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.75	20	5.048	7.674	0.443	5.084	1.686	14.428⋆	18.651⋆	14.447⋆	18.715⋆
			(4.262)	(6.301)	(0.904)	(4.316)	(2.651)	(6.649)	(7.297)	(6.688)	(7.402)
		30	25.226	30.990	4.487	25.778	20.940	35.309	38.112⋆	35.857	38.715⋆
			(8.245)	(11.079)	(3.510)	(8.580)	(8.710)	(9.114)	(8.915)	(9.517)	(9.252)
		40	47.777	54.222	15.391	49.034	45.412	56.012	57.216	57.159	58.461
			(8.845)	(11.711)	(5.982)	(9.191)	(9.387)	(9.220)	(9.034)	(9.517)	(9.373)
		50	66.716	73.399	29.527	68.468	65.556	73.404	73.996	75.176	75.652
			(9.205)	(12.319)	(7.173)	(9.590)	(9.341)	(9.308)	(9.095)	(9.651)	(9.513)
		60	82.497	89.236	44.011	84.617	82.121	87.925	88.233	90.141	90.354
			(8.505)	(11.622)	(7.137)	(8.753)	(8.577)	(8.481)	(8.405)	(8.868)	(8.724)
		100	121.157	126.546	86.258	123.869	121.317	124.149	124.063	126.853	126.701
			(6.839)	(9.096)	(6.628)	(7.075)	(6.973)	(6.876)	(6.814)	(7.002)	(7.098)
		200	160.716	164.072	135.462	163.178	161.088	161.877	161.899	164.294	164.290
			(5.488)	(6.604)	(5.407)	(5.471)	(5.509)	(5.420)	(5.399)	(5.453)	(5.433)
		300	176.906	179.237	158.139	178.788	177.197	177.572	177.572	179.412	179.386
			(4.278)	(4.995)	(4.711)	(4.267)	(4.254)	(4.251)	(4.252)	(4.179)	(4.291)
		500	190.561	191.602	179.203	191.562	190.692	190.776	190.782	191.782	191.792
			(2.973)	(3.155)	(3.761)	(2.884)	(2.984)	(2.974)	(2.967)	(2.877)	(2.890)
	-0.75	20	5.197	6.942	0.462	5.227	1.817	14.204⋆	18.473⋆	14.210⋆	18.508⋆
			(4.299)	(4.897)	(0.867)	(4.336)	(2.798)	(6.390)	(6.736)	(6.406)	(6.800)
		30	25.268	29.912	4.355	25.759	20.777	35.316⋆	37.907⋆	35.894⋆	38.526⋆
			(7.687)	(7.612)	(3.353)	(7.834)	(8.181)	(7.801)	(7.688)	(8.145)	(8.016)
		40	47.903	53.707	15.627	49.056	45.587	56.033	57.263	57.297	58.406
			(7.559)	(7.411)	(5.312)	(7.708)	(7.652)	(7.419)	(7.125)	(7.786)	(7.349)
		50	66.954	72.961	29.729	68.694	65.768	73.671	74.190	75.430	75.914
			(7.277)	(7.406)	(6.022)	(7.462)	(7.322)	(7.390)	(7.181)	(7.593)	(7.401)
		60	82.418	87.938	43.900	84.538	81.922	87.824	88.135	89.970	90.331
			(6.666)	(6.773)	(6.279)	(6.935)	(6.821)	(6.729)	(6.665)	(6.991)	(6.867)
		100	121.260	126.047	86.329	123.933	121.521	124.188	124.138	126.847	126.713
			(6.022)	(6.109)	(5.699)	(6.173)	(6.035)	(5.894)	(5.981)	(6.040)	(6.089)
		200	160.794	163.977	135.425	163.206	161.125	161.902	161.911	164.294	164.281
			(4.628)	(4.826)	(5.026)	(4.685)	(4.673)	(4.644)	(4.652)	(4.718)	(4.720)
		300	176.774	178.930	157.984	178.635	177.074	177.372	177.365	179.235	179.162
			(3.767)	(3.696)	(4.201)	(3.639)	(3.736)	(3.654)	(3.730)	(3.575)	(3.596)
		500	190.356	191.463	179.280	191.442	190.508	190.612	190.623	191.657	191.617
			(2.763)	(2.754)	(3.343)	(2.688)	(2.744)	(2.739)	(2.755)	(2.649)	(2.677)

Table C.45. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.8$ and $\rho = \pm 0.9$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is 1,000 with 10,000 bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.8	0.9	20	5.543	9.439	0.521	5.581	2.084	14.637⋆	18.661⋆	14.701⋆	18.748⋆
			(5.760)	(10.722)	(1.188)	(5.834)	(4.098)	(9.047)	(9.816)	(9.200)	(9.966)
		30	24.968	32.269	4.307	25.476	20.395	35.200	37.943⋆	35.801	38.569⋆
			(9.869)	(14.796)	(4.066)	(10.133)	(10.733)	(10.988)	(11.023)	(11.535)	(11.507)
		40	48.682	57.537	15.989	50.002	46.419	56.807	58.069	58.112	59.340
			(11.841)	(17.275)	(7.247)	(12.211)	(12.473)	(12.194)	(12.019)	(12.667)	(12.474)
		50	67.027	75.245	29.995	68.729	66.109	73.648	74.358	75.451	76.070
			(11.083)	(16.660)	(8.273)	(11.366)	(11.434)	(11.201)	(11.258)	(11.658)	(11.622)
		60	82.826	90.946	44.202	84.948	82.174	88.208	88.490	90.605	90.670
			(10.524)	(15.785)	(8.709)	(10.867)	(10.906)	(10.736)	(10.678)	(11.212)	(11.175)
		100	121.115	127.456	86.572	123.776	121.361	124.148	123.976	126.804	126.624
			(8.929)	(13.148)	(8.310)	(9.086)	(9.028)	(8.967)	(8.958)	(9.147)	(9.143)
		200	160.838	164.715	135.598	163.277	161.219	161.987	161.971	164.351	164.278
			(6.625)	(8.725)	(6.524)	(6.640)	(6.633)	(6.593)	(6.593)	(6.588)	(6.590)
		300	176.525	178.975	157.913	178.386	176.822	177.186	177.159	178.913	178.932
			(5.004)	(6.256)	(5.608)	(4.953)	(5.008)	(4.975)	(4.978)	(4.926)	(4.952)
		500	190.330	191.625	179.277	191.328	190.454	190.570	190.577	191.569	191.556
			(3.507)	(3.946)	(4.469)	(3.410)	(3.504)	(3.481)	(3.475)	(3.351)	(3.366)
	-0.9	20	5.451	6.977	0.445	5.497	1.888	14.361⋆	18.234⋆	14.387⋆	18.286⋆
			(4.787)	(5.235)	(0.898)	(4.856)	(3.051)	(6.742)	(7.333)	(6.801)	(7.427)
		30	25.493	30.712	4.425	25.944	20.861	35.715	38.191⋆	36.261	38.747⋆
			(8.107)	(7.901)	(3.535)	(8.331)	(8.725)	(8.327)	(8.159)	(8.627)	(8.440)
		40	47.582	53.768	15.250	48.747	45.139	55.684	57.183	57.009	58.425
			(8.033)	(7.941)	(5.958)	(8.282)	(8.309)	(8.069)	(7.716)	(8.341)	(8.019)
		50	66.999	73.536	29.655	68.777	65.825	73.801	74.321	75.529	76.009
			(7.344)	(7.465)	(6.357)	(7.483)	(7.354)	(7.398)	(7.245)	(7.675)	(7.450)
		60	82.225	88.553	43.833	84.433	81.751	87.757	88.054	89.888	90.113
			(7.122)	(7.470)	(6.336)	(7.390)	(7.354)	(7.213)	(7.143)	(7.452)	(7.336)
		100	121.349	126.389	86.495	124.032	121.594	124.237	124.259	126.999	126.857
			(5.976)	(6.520)	(5.632)	(6.098)	(6.087)	(6.023)	(6.019)	(6.176)	(6.191)
		200	160.766	164.049	135.777	163.134	161.095	161.872	161.837	164.275	164.241
			(4.544)	(5.001)	(4.934)	(4.632)	(4.553)	(4.561)	(4.561)	(4.624)	(4.588)
		300	176.880	179.164	158.342	178.698	177.137	177.531	177.489	179.266	179.312
			(3.616)	(3.936)	(4.241)	(3.662)	(3.661)	(3.648)	(3.680)	(3.597)	(3.638)
		500	190.321	191.529	179.457	191.375	190.459	190.577	190.557	191.612	191.577
			(2.895)	(2.966)	(3.367)	(2.804)	(2.898)	(2.897)	(2.882)	(2.772)	(2.741)

Table C.46. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.1$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.1	20	7.503	11.101	0.722	7.587	1.508	19.170⋆	21.313⋆	19.229⋆	21.378⋆
			(5.339)	(6.739)	(1.162)	(5.426)	(2.542)	(7.172)	(7.173)	(7.285)	(7.274)
		30	32.866	41.591	6.252	33.743	23.266	45.926	45.350	47.085	46.307
			(8.334)	(9.672)	(3.882)	(8.654)	(8.965)	(8.658)	(8.457)	(9.121)	(8.833)
		40	61.073	70.811	19.015	63.162	55.337	71.545	70.021	73.834	72.027
			(8.605)	(10.020)	(5.694)	(8.891)	(9.111)	(8.529)	(8.533)	(9.038)	(8.931)
		50	85.023	95.232	36.256	88.034	81.405	93.505	91.766	96.671	94.575
			(8.241)	(9.290)	(6.871)	(8.697)	(8.527)	(8.393)	(8.272)	(8.760)	(8.666)
		60	104.760	114.618	53.227	108.537	102.340	111.765	109.804	115.515	113.258
			(8.082)	(9.315)	(7.382)	(8.517)	(8.293)	(8.097)	(8.015)	(8.360)	(8.276)
		100	153.358	161.223	107.547	157.927	152.382	156.889	155.426	161.348	159.831
			(6.630)	(7.210)	(6.596)	(6.730)	(6.611)	(6.435)	(6.523)	(6.741)	(6.638)
		200	202.333	207.137	170.340	206.106	202.222	203.572	203.054	207.310	206.816
			(5.068)	(5.263)	(5.392)	(5.216)	(5.107)	(5.080)	(5.078)	(5.123)	(5.173)
		300	221.237	224.391	198.752	224.013	221.316	221.842	221.653	224.553	224.393
			(4.051)	(4.275)	(4.829)	(4.073)	(4.089)	(4.070)	(4.092)	(4.071)	(4.090)
		500	236.674	238.379	223.258	238.319	236.804	236.947	236.905	238.550	238.529
			(3.136)	(3.099)	(3.674)	(3.032)	(3.147)	(3.130)	(3.130)	(3.006)	(3.030)
	-0.1	20	7.290	10.515	0.695	7.357	1.493	18.502⋆	20.939⋆	18.552⋆	21.038⋆
			(5.025)	(6.541)	(1.120)	(5.108)	(2.432)	(7.279)	(7.268)	(7.366)	(7.422)
		30	33.112	41.388	6.151	33.877	23.950	45.787	45.482	46.763	46.490
			(8.272)	(9.154)	(3.681)	(8.570)	(8.812)	(8.424)	(8.426)	(8.788)	(8.635)
		40	60.985	70.689	19.499	62.960	55.469	71.517	69.971	73.807	72.131
			(8.568)	(9.308)	(5.682)	(8.961)	(9.051)	(8.662)	(8.612)	(8.951)	(8.974)
		50	85.143	94.897	36.331	88.105	81.626	93.780	91.947	97.084	94.846
			(8.114)	(8.994)	(6.891)	(8.500)	(8.495)	(7.977)	(8.056)	(8.511)	(8.373)
		60	104.222	114.218	53.172	107.842	101.525	111.087	109.134	114.873	112.654
			(7.985)	(8.531)	(6.833)	(8.314)	(8.227)	(7.811)	(7.916)	(8.107)	(8.116)
		100	153.257	161.132	107.423	157.857	152.427	156.938	155.492	161.380	159.814
			(6.615)	(7.132)	(6.614)	(6.793)	(6.601)	(6.508)	(6.573)	(6.738)	(6.805)
		200	202.429	207.294	170.399	206.263	202.334	203.738	203.185	207.470	206.902
			(5.179)	(5.151)	(5.547)	(5.115)	(5.212)	(5.106)	(5.121)	(5.023)	(5.144)
		300	221.068	224.259	198.450	223.943	221.177	221.807	221.520	224.527	224.307
			(4.175)	(4.157)	(4.482)	(4.113)	(4.228)	(4.154)	(4.200)	(4.120)	(4.096)
		500	236.352	238.085	222.765	238.072	236.474	236.603	236.598	238.306	238.260
			(2.873)	(2.766)	(3.707)	(2.736)	(2.875)	(2.853)	(2.894)	(2.752)	(2.717)

Table C.47. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.25$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.25	20	7.393	10.960	0.702	7.507	1.558	18.568*	20.788*	18.628*	20.858*
			(5.310)	(6.907)	(1.065)	(5.457)	(2.563)	(7.459)	(7.438)	(7.576)	(7.560)
		30	33.336	41.567	6.153	34.349	23.558	45.462	45.092	46.582	46.045
			(8.256)	(10.159)	(3.871)	(8.591)	(9.354)	(8.955)	(8.743)	(9.424)	(9.088)
		40	61.478	71.202	19.846	63.539	55.873	71.606	69.875	73.964	71.901
			(9.006)	(10.765)	(6.058)	(9.483)	(9.293)	(9.138)	(9.020)	(9.657)	(9.392)
		50	85.096	95.175	36.370	88.138	81.499	93.615	91.760	96.943	94.678
			(8.653)	(10.167)	(7.036)	(9.061)	(8.950)	(8.672)	(8.748)	(9.090)	(9.068)
		60	104.309	113.831	53.234	107.933	101.872	111.388	109.322	115.186	112.886
			(8.219)	(9.852)	(7.112)	(8.617)	(8.403)	(8.094)	(8.301)	(8.462)	(8.533)
		100	153.305	161.318	107.587	157.846	152.244	156.886	155.466	161.348	159.745
			(6.846)	(8.115)	(6.794)	(7.196)	(6.937)	(6.915)	(6.855)	(7.112)	(7.108)
		200	202.240	207.006	170.048	206.128	202.128	203.538	203.000	207.297	206.795
			(5.168)	(5.571)	(5.658)	(5.170)	(5.268)	(5.170)	(5.212)	(5.149)	(5.113)
		300	221.098	224.363	198.326	223.977	221.223	221.793	221.610	224.560	224.385
			(4.124)	(4.245)	(4.832)	(4.032)	(4.141)	(4.102)	(4.134)	(4.006)	(4.022)
		500	236.321	238.219	222.961	238.104	236.434	236.570	236.548	238.348	238.312
			(3.117)	(3.115)	(3.765)	(3.023)	(3.146)	(3.131)	(3.134)	(2.960)	(3.012)
	-0.25	20	7.296	11.020	0.673	7.362	1.505	19.120*	21.287*	19.155*	21.365*
			(5.105)	(6.641)	(1.053)	(5.173)	(2.628)	(7.103)	(7.047)	(7.169)	(7.182)
		30	32.975	41.294	5.916	33.836	23.337	45.442	44.816	46.489	45.850
			(8.084)	(9.044)	(3.681)	(8.340)	(9.084)	(8.304)	(8.155)	(8.663)	(8.498)
		40	61.640	71.048	19.399	63.715	55.908	72.171	70.539	74.363	72.642
			(8.407)	(9.000)	(5.894)	(8.732)	(9.308)	(8.416)	(8.449)	(8.763)	(8.815)
		50	84.529	94.363	36.354	87.499	81.015	93.142	91.256	96.404	94.130
			(8.323)	(8.899)	(6.771)	(8.697)	(8.692)	(8.205)	(8.404)	(8.677)	(8.717)
		60	104.711	114.205	53.647	108.241	102.151	111.759	109.718	115.411	113.073
			(7.982)	(8.022)	(7.104)	(8.282)	(8.074)	(7.884)	(7.826)	(8.155)	(8.043)
		100	153.535	161.283	107.328	158.121	152.613	157.103	155.528	161.617	159.963
			(6.558)	(7.100)	(6.310)	(6.892)	(6.769)	(6.620)	(6.738)	(6.893)	(6.905)
		200	202.454	207.364	170.421	206.376	202.347	203.714	203.210	207.569	207.049
			(5.064)	(5.152)	(5.267)	(5.032)	(5.070)	(5.014)	(4.990)	(4.949)	(4.946)
		300	221.038	224.277	198.377	223.922	221.162	221.677	221.489	224.493	224.318
			(4.237)	(4.244)	(4.714)	(4.154)	(4.193)	(4.182)	(4.188)	(4.096)	(4.119)
		500	236.385	238.190	222.940	238.123	236.527	236.708	236.652	238.364	238.326
			(2.956)	(2.875)	(3.609)	(2.878)	(2.932)	(2.906)	(2.910)	(2.870)	(2.867)

Table C.48. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.5$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.5	20	7.551	11.424	0.629	7.639	1.564	18.711⋆	20.901⋆	18.747⋆	20.961⋆
			(5.270)	(7.330)	(1.039)	(5.367)	(2.683)	(7.329)	(7.288)	(7.400)	(7.380)
		30	33.440	41.875	6.013	34.376	23.694	45.772	44.963	46.830	45.967⋆
			(8.817)	(11.599)	(3.958)	(9.158)	(9.428)	(9.638)	(9.717)	(10.143)	(10.221)
		40	60.623	70.864	19.369	62.669	54.797	71.269	69.608	73.557	71.545
			(9.472)	(12.221)	(6.095)	(9.923)	(10.211)	(9.904)	(9.819)	(10.472)	(10.272)
		50	84.920	94.829	36.460	87.899	81.252	93.268	91.390	96.381	94.185
			(9.231)	(11.491)	(7.294)	(9.619)	(9.468)	(9.084)	(8.953)	(9.591)	(9.450)
		60	104.516	114.598	53.466	108.181	101.933	111.609	109.418	115.271	113.025
			(8.872)	(11.399)	(7.613)	(9.245)	(9.106)	(8.866)	(8.802)	(9.146)	(9.155)
		100	153.785	161.707	107.785	158.327	152.875	157.372	155.903	161.744	160.324
			(7.555)	(9.071)	(7.001)	(7.727)	(7.705)	(7.433)	(7.511)	(7.598)	(7.635)
		200	202.072	206.773	170.186	205.876	201.977	203.299	202.819	207.046	206.506
			(5.395)	(6.214)	(5.853)	(5.419)	(5.424)	(5.275)	(5.359)	(5.441)	(5.452)
		300	221.100	224.335	198.247	224.031	221.218	221.788	221.605	224.651	224.431
			(4.342)	(4.623)	(4.932)	(4.275)	(4.327)	(4.301)	(4.325)	(4.311)	(4.298)
		500	236.369	238.288	222.881	238.197	236.510	236.670	236.618	238.426	238.377
			(3.218)	(3.296)	(3.942)	(3.054)	(3.217)	(3.204)	(3.207)	(3.035)	(3.044)
	-0.5	20	7.402	10.592	0.663	7.462	1.554	18.896⋆	20.933⋆	18.906⋆	20.990⋆
			(5.075)	(6.262)	(1.084)	(5.146)	(2.464)	(6.714)	(6.927)	(6.732)	(7.024)
		30	33.609	41.787	6.146	34.530	24.242	46.054	45.439	47.082	46.325
			(7.821)	(8.384)	(3.694)	(8.100)	(8.724)	(8.102)	(8.126)	(8.547)	(8.423)
		40	61.252	70.784	19.610	63.362	55.175	71.895	70.088	74.179	72.067
			(8.460)	(9.109)	(5.988)	(8.767)	(9.102)	(8.660)	(8.512)	(9.040)	(8.780)
		50	84.537	94.586	36.075	87.598	81.118	93.327	91.447	96.470	94.349
			(8.547)	(8.689)	(6.444)	(8.920)	(8.657)	(8.510)	(8.320)	(8.889)	(8.620)
		60	104.136	114.022	53.239	107.836	101.540	111.112	108.976	114.707	112.400
			(7.597)	(7.867)	(7.069)	(7.965)	(7.591)	(7.407)	(7.343)	(7.698)	(7.658)
		100	153.485	161.434	107.620	158.133	152.572	157.168	155.718	161.676	160.232
			(6.540)	(6.895)	(6.670)	(6.738)	(6.587)	(6.514)	(6.581)	(6.751)	(6.739)
		200	202.127	207.171	170.189	206.000	202.056	203.407	202.904	207.225	206.689
			(5.173)	(5.188)	(5.399)	(5.179)	(5.251)	(5.159)	(5.222)	(5.131)	(5.212)
		300	221.322	224.530	198.566	224.186	221.418	221.953	221.773	224.754	224.577
			(4.047)	(3.952)	(4.579)	(4.023)	(4.077)	(3.993)	(4.051)	(4.006)	(3.949)
		500	236.250	238.096	222.878	237.964	236.384	236.495	236.469	238.204	238.206
			(3.004)	(2.965)	(3.714)	(2.949)	(3.020)	(2.985)	(3.038)	(2.913)	(2.922)

Table C.49. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.75$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.75	20	7.792	12.352	0.752	7.870	1.697	19.204⋆	21.241⋆	19.310⋆	21.387⋆
			(6.118)	(9.364)	(1.248)	(6.253)	(2.976)	(9.201)	(9.389)	(9.372)	(9.615)
		30	32.950	42.238	6.025	33.880	23.480	45.285	44.709	46.369	45.690
			(10.219)	(14.440)	(4.304)	(10.635)	(10.874)	(11.081)	(10.958)	(11.649)	(11.436)
		40	61.408	72.035	19.398	63.424	55.409	71.941	70.214	74.281	72.147
			(10.657)	(15.343)	(6.719)	(11.042)	(11.224)	(10.819)	(10.822)	(11.470)	(11.280)
		50	84.787	96.035	36.157	87.837	81.207	93.464	91.520	96.656	94.476
			(10.486)	(14.802)	(7.780)	(11.014)	(10.863)	(10.669)	(10.695)	(11.159)	(11.129)
		60	104.601	115.021	53.536	108.233	102.001	111.344	109.458	115.190	112.899
			(10.310)	(14.515)	(8.708)	(10.925)	(10.445)	(10.436)	(10.314)	(11.003)	(10.811)
		100	153.298	161.485	106.959	157.889	152.388	156.928	155.455	161.391	159.819
			(8.666)	(11.634)	(8.359)	(8.908)	(8.755)	(8.625)	(8.639)	(8.877)	(8.778)
		200	202.342	207.520	170.379	206.201	202.269	203.593	203.145	207.279	206.898
			(6.029)	(7.712)	(6.410)	(5.872)	(6.097)	(6.011)	(6.018)	(5.872)	(5.984)
		300	221.021	224.241	198.290	223.852	221.127	221.642	221.477	224.377	224.235
			(4.734)	(5.659)	(5.544)	(4.637)	(4.743)	(4.742)	(4.758)	(4.604)	(4.622)
		500	236.394	238.161	222.960	238.098	236.491	236.677	236.621	238.345	238.293
			(3.375)	(3.733)	(4.120)	(3.256)	(3.390)	(3.355)	(3.374)	(3.241)	(3.252)
	-0.75	20	7.303	10.875	0.670	7.357	1.501	18.913⋆	21.058⋆	18.988⋆	21.152⋆
			(5.007)	(6.261)	(1.116)	(5.072)	(2.612)	(7.300)	(7.326)	(7.451)	(7.474)
		30	33.415	41.536	6.155	34.306	23.628	45.640	45.352	46.761	46.254
			(8.249)	(8.585)	(3.765)	(8.578)	(8.957)	(8.880)	(8.594)	(9.358)	(8.919)
		40	61.016	70.915	19.668	63.067	55.178	71.488	69.834	73.702	71.860
			(8.657)	(8.569)	(6.170)	(9.019)	(9.058)	(8.782)	(8.614)	(9.103)	(9.042)
		50	84.775	94.889	36.112	87.863	81.182	93.417	91.601	96.690	94.548
			(8.215)	(8.340)	(6.873)	(8.441)	(8.373)	(7.917)	(7.790)	(8.274)	(8.107)
		60	104.787	114.591	53.540	108.509	101.923	111.722	109.673	115.562	113.314
			(7.960)	(8.224)	(7.110)	(8.224)	(8.229)	(8.047)	(8.170)	(8.498)	(8.394)
		100	153.346	161.016	107.630	157.817	152.405	156.869	155.428	161.384	159.785
			(6.569)	(6.977)	(6.294)	(6.813)	(6.521)	(6.472)	(6.503)	(6.736)	(6.832)
		200	202.179	207.285	170.324	206.110	202.126	203.473	202.944	207.281	206.790
			(4.882)	(5.166)	(5.404)	(4.960)	(5.002)	(4.896)	(4.964)	(4.920)	(4.998)
		300	221.260	224.540	198.502	224.116	221.323	221.842	221.666	224.668	224.526
			(4.077)	(4.070)	(4.795)	(4.017)	(4.108)	(4.084)	(4.025)	(3.988)	(3.985)
		500	236.495	238.288	223.058	238.189	236.614	236.776	236.773	238.451	238.386
			(3.080)	(3.008)	(3.579)	(2.906)	(3.084)	(3.034)	(3.049)	(2.931)	(2.917)

Table C.50. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the normal variates with $\pi_0 = 0.75$ and $\rho = \pm 0.9$. Fifty blocks are utilized with pairwise correlation between the variables within a block of ρ. The pre-specified significance level is $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis. Cases where FDR control were anti-conservative are indicated with a star($\star$).

π_0	ρ	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.75	0.9	20	7.913	14.298	0.789	8.039	2.253	19.218⋆	21.210⋆	19.365⋆	21.406⋆
			(7.436)	(14.116)	(1.449)	(7.630)	(4.390)	(11.069)	(11.288)	(11.312)	(11.584)
		30	33.420	45.796	6.342	34.421	23.824	45.726	44.948	46.881	46.008
			(13.504)	(22.659)	(4.902)	(14.033)	(14.017)	(14.586)	(14.397)	(15.356)	(15.054)
		40	61.404	74.626	19.659	63.552	55.752	72.021	70.395	74.370	72.467
			(13.922)	(22.550)	(8.185)	(14.563)	(14.963)	(14.289)	(14.322)	(15.006)	(15.051)
		50	84.765	97.728	36.028	87.800	81.013	93.411	91.369	96.751	94.375
			(13.690)	(21.842)	(9.720)	(14.378)	(14.317)	(13.947)	(13.836)	(14.760)	(14.639)
		60	104.623	117.335	53.533	108.411	102.066	111.717	109.642	115.545	113.268
			(13.358)	(21.311)	(10.460)	(14.012)	(13.691)	(13.460)	(13.402)	(14.083)	(14.006)
		100	153.179	162.837	107.115	157.779	152.207	156.827	155.293	161.360	159.730
			(10.529)	(16.574)	(10.282)	(10.885)	(10.734)	(10.530)	(10.624)	(10.912)	(10.980)
		200	202.514	208.156	170.518	206.289	202.438	203.754	203.241	207.512	206.987
			(7.299)	(10.270)	(7.989)	(7.326)	(7.325)	(7.265)	(7.219)	(7.288)	(7.301)
		300	221.110	224.771	198.629	223.996	221.186	221.749	221.570	224.576	224.365
			(5.704)	(7.571)	(6.448)	(5.664)	(5.709)	(5.606)	(5.676)	(5.591)	(5.671)
		500	236.631	238.631	223.093	238.293	236.735	236.896	236.855	238.575	238.501
			(3.848)	(4.493)	(4.625)	(3.727)	(3.827)	(3.824)	(3.803)	(3.711)	(3.685)
	-0.9	20	7.528	10.722	0.701	7.610	1.796	18.842⋆	20.874⋆	18.925⋆	20.962⋆
			(5.794)	(6.501)	(1.247)	(5.911)	(3.121)	(8.093)	(8.124)	(8.235)	(8.255)
		30	33.186	41.597	6.117	34.164	23.350	45.427	44.855	46.539	45.815
			(9.012)	(8.634)	(4.085)	(9.338)	(10.558)	(9.687)	(9.855)	(10.172)	(10.168)
		40	61.350	71.660	19.607	63.504	55.620	71.945	70.316	74.172	72.445
			(9.155)	(9.059)	(6.309)	(9.475)	(9.918)	(9.361)	(9.047)	(9.822)	(9.538)
		50	84.610	95.115	36.155	87.474	81.003	93.275	91.303	96.428	94.204
			(8.678)	(8.793)	(7.083)	(8.948)	(8.878)	(8.400)	(8.531)	(8.791)	(8.815)
		60	104.163	114.447	53.221	107.887	101.511	111.223	109.235	114.972	112.898
			(8.001)	(8.634)	(7.412)	(8.239)	(8.223)	(7.946)	(7.829)	(8.279)	(8.240)
		100	153.335	161.716	107.657	158.067	152.533	157.025	155.587	161.489	159.856
			(6.833)	(7.907)	(6.775)	(7.025)	(6.765)	(6.709)	(6.765)	(6.915)	(6.919)
		200	202.017	207.401	170.289	205.913	201.914	203.270	202.753	207.098	206.619
			(5.038)	(6.100)	(5.369)	(5.090)	(5.055)	(5.066)	(5.058)	(5.067)	(5.090)
		300	220.988	224.489	198.347	223.910	221.091	221.672	221.486	224.514	224.332
			(4.193)	(4.602)	(4.710)	(4.108)	(4.236)	(4.174)	(4.238)	(4.081)	(4.100)
		500	236.506	238.340	222.843	238.251	236.625	236.817	236.763	238.491	238.433
			(3.047)	(3.200)	(3.723)	(2.956)	(3.058)	(3.033)	(3.040)	(2.969)	(2.985)

C.2. Gamma Distributed Random Variables

C.2.1. Numerical Summaries of Empirical False Discovery Rates

Table C.51. Empirical FDRs for the investigated methods for the correlated cases for the gamma variates with $\pi_0 = 0.85$ and 0.9. The number of hypotheses is $m = 1,000$ with a pre-specified significance level of $\alpha = 0.05$. The number of replications for each scenario is 1,000 with 10,000 bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.90	20	0.008	0.009	0.000	0.009	0.010	0.000	0.030	0.000	0.033
		(0.070)	(0.075)	(0.004)	(0.079)	(0.077)	(0.003)	(0.129)	(0.003)	(0.142)
	30	0.012	0.017	0.000	0.014	0.014	0.000	0.023	0.000	0.026
		(0.085)	(0.106)	(0.003)	(0.099)	(0.092)	(0.003)	(0.114)	(0.003)	(0.130)
	40	0.014	0.016	0.000	0.016	0.015	0.000	0.022	0.000	0.026
		(0.090)	(0.099)	(0.005)	(0.101)	(0.092)	(0.003)	(0.111)	(0.003)	(0.126)
	50	0.013	0.014	0.000	0.017	0.014	0.000	0.020	0.000	0.024
		(0.087)	(0.091)	(0.002)	(0.106)	(0.090)	(0.001)	(0.106)	(0.001)	(0.126)
	60	0.018	0.020	0.001	0.021	0.019	0.000	0.024	0.000	0.028
		(0.103)	(0.111)	(0.008)	(0.116)	(0.104)	(0.003)	(0.114)	(0.003)	(0.132)
	100	0.016	0.020	0.001	0.022	0.016	0.000	0.020	0.000	0.026
		(0.086)	(0.102)	(0.009)	(0.110)	(0.087)	(0.002)	(0.100)	(0.003)	(0.125)
	200	0.020	0.022	0.000	0.024	0.019	0.000	0.021	0.000	0.026
		(0.107)	(0.114)	(0.004)	(0.127)	(0.107)	(0.001)	(0.112)	(0.001)	(0.132)
	300	0.015	0.016	0.000	0.020	0.015	0.000	0.016	0.000	0.021
		(0.090)	(0.094)	(0.003)	(0.113)	(0.088)	(0.000)	(0.093)	(0.001)	(0.115)
	500	0.017	0.017	0.000	0.020	0.016	0.000	0.017	0.000	0.020
		(0.099)	(0.099)	(0.005)	(0.113)	(0.098)	(0.001)	(0.099)	(0.001)	(0.114)
0.85	20	0.011	0.015	0.000	0.013	0.013	0.001	0.029	0.001	0.035
		(0.079)	(0.099)	(0.007)	(0.094)	(0.084)	(0.006)	(0.121)	(0.007)	(0.142)
	30	0.018	0.018	0.001	0.021	0.019	0.000	0.027	0.001	0.032
		(0.102)	(0.102)	(0.010)	(0.118)	(0.104)	(0.008)	(0.122)	(0.012)	(0.141)
	40	0.016	0.018	0.000	0.022	0.017	0.000	0.026	0.000	0.033
		(0.087)	(0.095)	(0.002)	(0.112)	(0.090)	(0.001)	(0.112)	(0.002)	(0.136)
	50	0.018	0.021	0.000	0.024	0.018	0.000	0.023	0.000	0.030
		(0.094)	(0.105)	(0.004)	(0.119)	(0.095)	(0.003)	(0.109)	(0.003)	(0.133)
	60	0.018	0.021	0.001	0.024	0.018	0.000	0.023	0.000	0.030
		(0.092)	(0.103)	(0.006)	(0.116)	(0.092)	(0.002)	(0.104)	(0.003)	(0.130)
	100	0.019	0.020	0.000	0.026	0.019	0.000	0.021	0.000	0.028
		(0.097)	(0.100)	(0.003)	(0.123)	(0.095)	(0.001)	(0.103)	(0.001)	(0.129)
	200	0.015	0.017	0.001	0.020	0.015	0.000	0.016	0.000	0.021
		(0.087)	(0.094)	(0.006)	(0.107)	(0.085)	(0.002)	(0.088)	(0.002)	(0.108)
	300	0.024	0.025	0.001	0.031	0.023	0.000	0.024	0.000	0.031
		(0.113)	(0.119)	(0.005)	(0.138)	(0.113)	(0.001)	(0.115)	(0.001)	(0.139)
	500	0.020	0.021	0.000	0.026	0.019	0.000	0.020	0.000	0.026
		(0.097)	(0.104)	(0.005)	(0.123)	(0.096)	(0.001)	(0.098)	(0.001)	(0.123)

Table C.52. Empirical FDRs for the investigated methods for the correlated cases for the gamma variates with $\pi_0 = 0.75$ and 0.8. The number of hypotheses is $m = 1,000$ with a pre-specified significance level of $\alpha = 0.05$. The number of replications for each scenario is 1,000 with 10,000 bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.80	20	0.009	0.012	0.000	0.011	0.010	0.001	0.021	0.001	0.026
		(0.071)	(0.087)	(0.009)	(0.084)	(0.075)	(0.013)	(0.099)	(0.016)	(0.119)
	30	0.014	0.014	0.000	0.019	0.015	0.000	0.022	0.001	0.027
		(0.084)	(0.084)	(0.004)	(0.107)	(0.085)	(0.005)	(0.103)	(0.006)	(0.124)
	40	0.016	0.019	0.000	0.022	0.016	0.000	0.021	0.000	0.028
		(0.085)	(0.098)	(0.004)	(0.110)	(0.085)	(0.003)	(0.099)	(0.004)	(0.124)
	50	0.020	0.023	0.000	0.029	0.021	0.000	0.026	0.000	0.036
		(0.095)	(0.107)	(0.005)	(0.126)	(0.096)	(0.003)	(0.108)	(0.004)	(0.138)
	60	0.024	0.024	0.001	0.035	0.024	0.000	0.028	0.000	0.039
		(0.104)	(0.104)	(0.006)	(0.140)	(0.103)	(0.002)	(0.115)	(0.004)	(0.149)
	100	0.021	0.025	0.001	0.031	0.021	0.000	0.023	0.000	0.033
		(0.094)	(0.106)	(0.005)	(0.127)	(0.092)	(0.002)	(0.098)	(0.003)	(0.131)
	200	0.021	0.022	0.001	0.032	0.021	0.000	0.022	0.000	0.032
		(0.093)	(0.093)	(0.004)	(0.125)	(0.091)	(0.001)	(0.095)	(0.002)	(0.125)
	300	0.020	0.022	0.000	0.031	0.020	0.000	0.020	0.000	0.031
		(0.088)	(0.094)	(0.003)	(0.123)	(0.086)	(0.001)	(0.088)	(0.001)	(0.122)
	500	0.022	0.023	0.001	0.031	0.021	0.000	0.022	0.000	0.031
		(0.099)	(0.102)	(0.006)	(0.129)	(0.097)	(0.001)	(0.099)	(0.001)	(0.129)
0.75	20	0.015	0.018	0.000	0.020	0.016	0.001	0.030	0.001	0.040
		(0.075)	(0.087)	(0.005)	(0.101)	(0.080)	(0.006)	(0.112)	(0.008)	(0.142)
	30	0.016	0.018	0.000	0.023	0.017	0.000	0.024	0.000	0.034
		(0.078)	(0.087)	(0.005)	(0.107)	(0.079)	(0.003)	(0.097)	(0.005)	(0.128)
	40	0.018	0.020	0.001	0.026	0.018	0.000	0.023	0.000	0.032
		(0.088)	(0.096)	(0.005)	(0.117)	(0.088)	(0.003)	(0.099)	(0.005)	(0.128)
	50	0.014	0.015	0.000	0.022	0.014	0.000	0.018	0.000	0.026
		(0.074)	(0.078)	(0.004)	(0.104)	(0.072)	(0.002)	(0.082)	(0.003)	(0.113)
	60	0.013	0.014	0.000	0.021	0.012	0.000	0.016	0.000	0.025
		(0.060)	(0.069)	(0.001)	(0.095)	(0.059)	(0.001)	(0.069)	(0.001)	(0.103)
	100	0.019	0.020	0.000	0.031	0.018	0.000	0.020	0.000	0.033
		(0.080)	(0.083)	(0.003)	(0.121)	(0.077)	(0.001)	(0.083)	(0.002)	(0.124)
	200	0.023	0.024	0.001	0.035	0.022	0.000	0.023	0.000	0.035
		(0.095)	(0.097)	(0.005)	(0.130)	(0.092)	(0.001)	(0.095)	(0.002)	(0.130)
	300	0.024	0.024	0.001	0.036	0.023	0.000	0.024	0.000	0.036
		(0.100)	(0.100)	(0.007)	(0.132)	(0.098)	(0.002)	(0.100)	(0.003)	(0.131)
	500	0.023	0.024	0.001	0.035	0.022	0.000	0.023	0.000	0.035
		(0.097)	(0.099)	(0.006)	(0.128)	(0.095)	(0.001)	(0.096)	(0.002)	(0.127)

C.2.2. Numerical Summaries of Empirical False Non-discovery Rates

Table C.53. Empirical FNRs for the investigated methods for the correlated cases for the gamma variates with $\pi_0 = 0.85$ and 0.9. The number of hypotheses is $m = 1,000$ with a pre-specified significance level of $\alpha = 0.05$. The number of replications for each scenario is 1,000 with 10,000 bootstrap resamples. Equal sample sizes were utilized for both the controls and cases.The standard errors are provided in parenthesis.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.90	20	0.099	0.099	0.100	0.099	0.097	0.099	0.091	0.099	0.091
		(0.008)	(0.009)	(0.000)	(0.009)	(0.013)	(0.004)	(0.020)	(0.004)	(0.021)
	30	0.096	0.096	0.100	0.096	0.093	0.099	0.089	0.099	0.089
		(0.014)	(0.016)	(0.002)	(0.015)	(0.018)	(0.006)	(0.022)	(0.006)	(0.022)
	40	0.092	0.092	0.099	0.092	0.088	0.097	0.083	0.097	0.083
		(0.019)	(0.020)	(0.006)	(0.020)	(0.023)	(0.010)	(0.026)	(0.010)	(0.026)
	50	0.087	0.087	0.097	0.087	0.081	0.095	0.077	0.095	0.077
		(0.023)	(0.023)	(0.009)	(0.023)	(0.026)	(0.011)	(0.028)	(0.011)	(0.029)
	60	0.082	0.081	0.094	0.081	0.076	0.092	0.071	0.092	0.071
		(0.027)	(0.028)	(0.014)	(0.028)	(0.030)	(0.016)	(0.030)	(0.016)	(0.031)
	100	0.060	0.057	0.080	0.060	0.054	0.080	0.051	0.079	0.051
		(0.031)	(0.030)	(0.024)	(0.031)	(0.030)	(0.023)	(0.029)	(0.023)	(0.030)
	200	0.028	0.026	0.044	0.027	0.025	0.047	0.024	0.047	0.024
		(0.021)	(0.021)	(0.024)	(0.021)	(0.019)	(0.023)	(0.019)	(0.023)	(0.019)
	300	0.015	0.014	0.027	0.015	0.014	0.030	0.014	0.030	0.013
		(0.014)	(0.014)	(0.017)	(0.014)	(0.013)	(0.017)	(0.013)	(0.017)	(0.013)
	500	0.006	0.005	0.012	0.006	0.005	0.014	0.005	0.014	0.005
		(0.008)	(0.008)	(0.011)	(0.008)	(0.007)	(0.011)	(0.007)	(0.011)	(0.007)
0.85	20	0.147	0.147	0.150	0.147	0.144	0.149	0.138	0.149	0.137
		(0.015)	(0.018)	(0.001)	(0.017)	(0.021)	(0.006)	(0.030)	(0.006)	(0.032)
	30	0.143	0.143	0.149	0.143	0.139	0.147	0.133	0.147	0.132
		(0.024)	(0.025)	(0.005)	(0.026)	(0.030)	(0.011)	(0.034)	(0.011)	(0.036)
	40	0.137	0.135	0.148	0.136	0.130	0.146	0.124	0.146	0.123
		(0.031)	(0.032)	(0.008)	(0.033)	(0.036)	(0.013)	(0.039)	(0.013)	(0.041)
	50	0.131	0.129	0.146	0.130	0.124	0.143	0.118	0.143	0.117
		(0.035)	(0.036)	(0.015)	(0.037)	(0.039)	(0.017)	(0.041)	(0.018)	(0.042)
	60	0.122	0.119	0.142	0.121	0.114	0.139	0.109	0.139	0.108
		(0.040)	(0.040)	(0.020)	(0.041)	(0.042)	(0.021)	(0.043)	(0.022)	(0.044)
	100	0.095	0.091	0.123	0.094	0.087	0.123	0.084	0.123	0.083
		(0.045)	(0.045)	(0.033)	(0.046)	(0.044)	(0.031)	(0.044)	(0.032)	(0.045)
	200	0.048	0.044	0.075	0.046	0.043	0.080	0.042	0.079	0.041
		(0.035)	(0.035)	(0.036)	(0.035)	(0.033)	(0.034)	(0.033)	(0.035)	(0.033)
	300	0.026	0.024	0.046	0.025	0.023	0.052	0.023	0.051	0.022
		(0.025)	(0.025)	(0.030)	(0.024)	(0.023)	(0.030)	(0.023)	(0.030)	(0.023)
	500	0.010	0.009	0.022	0.009	0.009	0.027	0.009	0.026	0.008
		(0.014)	(0.014)	(0.020)	(0.013)	(0.013)	(0.022)	(0.013)	(0.022)	(0.013)

Table C.54. Empirical FNRs for the investigated methods for the correlated cases for the gamma variates with $\pi_0 = 0.75$ and 0.8. The number of hypotheses is $m = 1,000$ with a pre-specified significance level of $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ number of bootstrap resamples. Equal sample sizes were utilized for both the controls and cases.The standard errors are provided in parenthesis.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.80	20	0.197	0.196	0.200	0.196	0.194	0.199	0.185	0.199	0.184
		(0.020)	(0.023)	(0.003)	(0.023)	(0.026)	(0.009)	(0.037)	(0.009)	(0.039)
	30	0.190	0.190	0.199	0.189	0.184	0.197	0.175	0.197	0.174
		(0.032)	(0.032)	(0.007)	(0.035)	(0.039)	(0.014)	(0.045)	(0.015)	(0.047)
	40	0.182	0.180	0.197	0.181	0.173	0.194	0.165	0.194	0.164
		(0.041)	(0.044)	(0.014)	(0.044)	(0.047)	(0.019)	(0.051)	(0.020)	(0.053)
	50	0.167	0.164	0.192	0.165	0.157	0.188	0.149	0.188	0.148
		(0.053)	(0.054)	(0.023)	(0.056)	(0.057)	(0.026)	(0.059)	(0.027)	(0.061)
	60	0.158	0.154	0.186	0.156	0.147	0.184	0.140	0.183	0.138
		(0.057)	(0.056)	(0.030)	(0.059)	(0.059)	(0.031)	(0.059)	(0.032)	(0.061)
	100	0.121	0.114	0.159	0.118	0.111	0.160	0.108	0.159	0.105
		(0.061)	(0.061)	(0.046)	(0.063)	(0.060)	(0.043)	(0.059)	(0.045)	(0.061)
	200	0.060	0.055	0.097	0.058	0.056	0.104	0.054	0.102	0.052
		(0.047)	(0.046)	(0.048)	(0.047)	(0.044)	(0.046)	(0.044)	(0.047)	(0.044)
	300	0.033	0.030	0.061	0.031	0.030	0.070	0.030	0.067	0.028
		(0.034)	(0.034)	(0.041)	(0.034)	(0.033)	(0.041)	(0.032)	(0.042)	(0.032)
	500	0.011	0.010	0.027	0.010	0.010	0.033	0.010	0.031	0.009
		(0.018)	(0.018)	(0.027)	(0.017)	(0.017)	(0.030)	(0.017)	(0.029)	(0.016)
0.75	20	0.242	0.240	0.250	0.240	0.236	0.247	0.224	0.247	0.222
		(0.033)	(0.037)	(0.004)	(0.039)	(0.044)	(0.014)	(0.056)	(0.015)	(0.061)
	30	0.233	0.230	0.248	0.231	0.223	0.244	0.212	0.244	0.210
		(0.047)	(0.049)	(0.010)	(0.052)	(0.056)	(0.020)	(0.062)	(0.021)	(0.066)
	40	0.221	0.217	0.244	0.218	0.209	0.240	0.199	0.240	0.196
		(0.057)	(0.059)	(0.021)	(0.062)	(0.064)	(0.027)	(0.067)	(0.028)	(0.071)
	50	0.211	0.206	0.241	0.208	0.199	0.238	0.190	0.237	0.187
		(0.062)	(0.063)	(0.027)	(0.066)	(0.067)	(0.030)	(0.069)	(0.032)	(0.072)
	60	0.198	0.192	0.234	0.195	0.186	0.231	0.178	0.230	0.175
		(0.069)	(0.070)	(0.035)	(0.073)	(0.072)	(0.036)	(0.073)	(0.038)	(0.076)
	100	0.141	0.133	0.193	0.136	0.130	0.195	0.125	0.194	0.121
		(0.075)	(0.075)	(0.058)	(0.078)	(0.072)	(0.055)	(0.072)	(0.057)	(0.074)
	200	0.073	0.067	0.116	0.068	0.068	0.125	0.067	0.122	0.063
		(0.055)	(0.055)	(0.058)	(0.055)	(0.053)	(0.056)	(0.053)	(0.058)	(0.053)
	300	0.041	0.037	0.074	0.037	0.038	0.084	0.037	0.080	0.034
		(0.041)	(0.041)	(0.049)	(0.041)	(0.040)	(0.049)	(0.040)	(0.050)	(0.039)
	500	0.016	0.014	0.035	0.014	0.015	0.043	0.015	0.040	0.013
		(0.024)	(0.024)	(0.034)	(0.022)	(0.023)	(0.036)	(0.023)	(0.036)	(0.022)

C.2.3. Numerical Summaries of Average Number of False Hypotheses Rejected

Table C.55. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the gamma variates with $\pi_0 = 0.85$ and 0.9. The number of hypotheses is $m = 1,000$ with a pre-specified significance level of $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.90	20	1.329	1.457	0.035	1.435	3.415	0.789	8.989	0.793	9.211
		(8.888)	(9.459)	(0.480)	(9.687)	(13.445)	(3.977)	(21.058)	(4.016)	(21.736)
	30	3.702	4.520	0.298	3.880	6.859	1.533	11.761	1.543	11.921
		(14.284)	(16.064)	(2.329)	(15.075)	(18.681)	(6.386)	(22.692)	(6.452)	(23.117)
	40	8.134	8.721	1.395	8.394	12.696	3.412	17.289	3.430	17.509
		(19.965)	(20.590)	(6.532)	(20.560)	(24.218)	(10.232)	(26.616)	(10.307)	(27.054)
	50	13.172	14.062	2.854	13.478	19.443	4.931	24.001	4.970	24.310
		(23.522)	(23.840)	(9.531)	(24.103)	(27.268)	(11.911)	(28.888)	(12.060)	(29.350)
	60	19.004	20.285	6.196	19.359	25.402	8.177	29.773	8.237	30.072
		(28.044)	(28.319)	(15.175)	(28.526)	(30.482)	(16.545)	(30.996)	(16.736)	(31.393)
	100	41.503	44.990	21.095	42.030	48.049	21.553	50.538	21.676	50.946
		(31.072)	(30.743)	(25.015)	(31.402)	(30.010)	(23.848)	(29.385)	(24.040)	(29.629)
	200	73.863	75.493	57.865	74.403	76.678	54.814	77.353	55.213	77.795
		(20.319)	(20.329)	(23.371)	(20.300)	(18.630)	(22.970)	(18.260)	(23.120)	(18.273)
	300	85.652	86.697	75.041	86.072	87.151	71.884	87.455	72.259	87.771
		(12.955)	(13.031)	(16.296)	(12.858)	(12.010)	(16.481)	(11.818)	(16.533)	(11.771)
	500	94.590	94.969	88.985	94.827	95.140	86.656	95.216	86.938	95.382
		(7.094)	(7.156)	(9.796)	(6.919)	(6.559)	(10.444)	(6.463)	(10.390)	(6.376)
0.85	20	2.850	3.648	0.063	3.127	6.013	1.326	13.547	1.350	14.147
		(16.618)	(19.664)	(0.749)	(18.307)	(22.595)	(7.000)	(31.832)	(7.202)	(33.608)
	30	7.407	7.763	0.796	7.777	11.968	2.928	18.717	3.004	19.210
		(25.760)	(25.986)	(5.640)	(27.074)	(31.089)	(11.566)	(35.836)	(11.965)	(36.954)
	40	14.511	15.852	2.090	15.247	21.267	4.905	28.474	4.982	29.191
		(32.529)	(33.471)	(8.644)	(34.296)	(37.550)	(13.960)	(40.990)	(14.244)	(42.201)
	50	20.488	22.760	4.931	21.237	28.379	7.562	34.575	7.666	35.392
		(36.790)	(38.036)	(15.875)	(38.177)	(40.804)	(18.868)	(42.809)	(19.255)	(43.929)
	60	29.864	32.794	9.191	30.766	38.880	11.652	44.299	11.816	45.158
		(41.331)	(42.125)	(21.486)	(42.539)	(43.946)	(22.887)	(44.754)	(23.323)	(45.708)
	100	58.363	62.610	29.672	59.542	66.880	29.442	70.121	29.857	71.186
		(46.062)	(45.828)	(35.163)	(46.881)	(44.814)	(33.070)	(44.138)	(33.695)	(44.829)
	200	106.144	109.760	80.179	107.530	110.316	75.128	111.293	76.055	112.426
		(33.625)	(33.671)	(35.392)	(33.786)	(31.646)	(33.993)	(31.242)	(34.547)	(31.302)
	300	126.838	128.965	107.769	128.031	129.107	102.086	129.511	103.108	130.484
		(22.615)	(22.612)	(27.925)	(22.352)	(21.219)	(28.101)	(20.960)	(28.334)	(20.710)
	500	141.516	142.259	130.798	142.136	142.317	126.184	142.443	127.067	142.950
		(12.348)	(12.487)	(18.344)	(11.973)	(11.630)	(19.787)	(11.514)	(19.651)	(11.143)

Table C.56. Average number of false hypotheses rejected for the investigated methods for the correlated cases for the gamma variates with $\pi_0 = 0.75$ and 0.8. The number of hypotheses is $m = 1,000$ with a pre-specified significance level of $\alpha = 0.05$. The number of replications for each scenario is $1,000$ with $10,000$ bootstrap resamples. Equal sample sizes were utilized for both the controls and cases. The standard errors are provided in parenthesis.

π_0	n	BH	STS	BY	BKY	S-BH	SNS-BH	SNQ-BH	SNS-BKY	SNQ-BKY
0.80	20	3.715	4.829	0.171	4.090	7.138	1.546	16.992	1.598	17.798
		(21.912)	(25.484)	(3.520)	(24.049)	(28.326)	(9.960)	(39.858)	(10.486)	(42.124)
	30	10.780	11.353	1.231	11.559	17.883	3.928	27.689	4.041	28.687
		(34.340)	(34.539)	(8.458)	(36.841)	(41.603)	(15.779)	(48.282)	(16.433)	(50.224)
	40	20.198	22.607	3.651	21.330	29.611	7.103	38.496	7.310	39.783
		(43.969)	(46.359)	(15.734)	(46.378)	(50.034)	(21.412)	(53.675)	(22.325)	(55.665)
	50	36.533	39.394	9.800	38.217	47.209	13.941	55.217	14.318	56.919
		(56.022)	(56.963)	(25.514)	(58.490)	(59.922)	(29.034)	(61.670)	(30.042)	(63.649)
	60	46.360	50.020	15.383	48.148	58.092	18.831	65.283	19.266	67.065
		(59.267)	(59.103)	(33.153)	(61.353)	(61.061)	(34.215)	(61.434)	(35.217)	(63.151)
	100	85.924	92.775	45.627	88.322	95.995	44.791	99.611	45.735	101.838
		(62.179)	(62.352)	(49.384)	(63.619)	(60.540)	(46.566)	(59.738)	(47.850)	(60.895)
	200	146.362	151.370	111.718	148.967	151.058	104.578	152.140	106.457	154.415
		(43.473)	(43.455)	(46.950)	(43.569)	(41.100)	(45.622)	(40.645)	(46.486)	(40.496)
	300	171.597	174.545	145.992	173.607	173.906	138.396	174.399	140.470	176.122
		(30.601)	(30.632)	(37.920)	(30.080)	(29.060)	(38.591)	(28.708)	(38.954)	(28.166)
	500	190.540	191.253	177.458	191.524	191.283	171.701	191.424	173.476	192.310
		(15.436)	(15.577)	(23.785)	(14.628)	(14.615)	(26.029)	(14.442)	(25.734)	(13.707)
0.75	20	9.662	11.209	0.380	10.922	15.994	3.650	29.728	3.824	31.565
		(37.884)	(41.589)	(4.420)	(42.930)	(48.813)	(16.871)	(61.354)	(17.894)	(65.498)
	30	19.979	22.365	2.197	21.781	30.303	7.279	42.909	7.604	44.978
		(51.584)	(53.861)	(12.476)	(56.246)	(60.610)	(23.720)	(67.274)	(25.008)	(70.796)
	40	33.455	37.587	6.721	35.728	46.726	11.702	57.729	12.127	60.139
		(62.293)	(64.307)	(24.411)	(66.189)	(68.817)	(31.091)	(72.369)	(32.492)	(75.343)
	50	44.214	49.872	10.520	46.880	58.078	15.022	67.559	15.516	70.308
		(66.963)	(68.440)	(30.695)	(70.612)	(71.688)	(34.761)	(73.464)	(36.135)	(76.540)
	60	58.787	65.475	19.314	61.951	72.402	22.696	80.524	23.482	83.563
		(74.087)	(75.355)	(40.482)	(77.864)	(76.503)	(41.624)	(77.033)	(43.384)	(79.971)
	100	120.445	128.552	65.834	124.643	132.361	63.795	136.743	65.582	140.594
		(76.507)	(75.673)	(63.916)	(78.623)	(72.259)	(60.388)	(70.830)	(62.418)	(72.430)
	200	188.184	193.316	147.498	192.089	192.678	138.942	193.717	142.070	197.212
		(50.193)	(50.349)	(55.982)	(50.163)	(47.920)	(54.991)	(47.536)	(56.283)	(47.354)
	300	216.819	219.858	187.589	219.778	218.972	178.951	219.431	182.262	222.159
		(35.444)	(35.498)	(44.096)	(34.651)	(34.015)	(44.782)	(33.722)	(45.318)	(32.801)
	500	237.683	238.707	222.094	239.213	238.367	215.284	238.487	217.985	239.858
		(19.190)	(19.314)	(28.264)	(18.106)	(18.519)	(30.709)	(18.374)	(30.333)	(17.367)

APPENDIX D

R CODE FOR ESTIMATION OF STEP-DOWN CRITICAL VALUES

```
################################################################################
#                                crit.val function                             #
#                                -----------------                             #
# The function crit.val is used to obtain the critical values described in the #
# manuscript. The inputs are Data (B x s matrix of bootstrap test statistics   #
# where B is the number of bootstrap samples and s is the number of hypotheses #
# being tested), alpha (desired level of significance) and start (at which     #
# the search algorithm should start searching for the critical value)          #
#                                                                              #
# Required Packages: MASS 7.3-48                                               #
#                    matrixStats 0.53.0                                        #
################################################################################
```

```
crit.val <- function(Data, alpha = 0.10, start = 1){
cVal <- c()  ## initialize the vector of critical values
alphaVal <- c()
B <- nrow(Data) ## number of bootsrap replicates
s <- ncol(Data)  ## number of hypotheses being tested

for(j in 1:s){

## The ordering of the true null hypotheses in the bootstrap world is not
## 1,2,...,s, but it is instead determined by the ordering H_(1),...,H_(s)
## from the real world.So obtain the permutation {k_1,...,k_s} of {1,..,s}
## defined such that H_k_1 = H_(1),..., H_k_s = H_(s)
if (j==1) {
DataB <- as.matrix(Data[, j])
} else {
DataB <- t(apply(Data[, 1:j], 1, sort))
start <- which(t.dat >= cVal[1])[1]
}

t.dat <- sort(DataB[, j]) ## sort the B jth ordered test statistics

t.alpha <- lapply(start:B, function(i){

tval <- t.dat[i]
```

```
##Extract B by j sub block matrix from DataB and replicate the sub block j
## times and stack them up. Here j represents the jth critical value being
## sought given the 1, 2, ... j-1th critical values
DataBlock <- do.call("rbind", rep(list(DataB[ ,j:1]), j))

## If j =1, DataBlock will be a 1 x B matrix, convert it to a B x 1 matrix
if(j==1) DataBlock <- t(DataBlock)

## Create a matrix equivalent to DataBlock where the first column
## corresponds to the proposed jth critical value and the remaining
## columns correspond to the already computed critical values
CMatBlock <- matrix(rep(c(tval, cVal), j*B), byrow = TRUE, ncol = j)

## For the jth column and ith row, select cases where the test statistics
## exceed the previous computed critical values
IndBlock <- DataBlock >= CMatBlock

## We need the last inequality in each summand of the probabilities to be
## '<' instead of '>=', so we negate the result for '>=' to obtain the
## result.
COLS <- matrix(rep(1:j, each=j*B), byrow=F, ncol = j)
ROWS <- matrix(rep(1:j, each=j*B), byrow=T, ncol = j)
IndBlock[ROWS == (COLS-1)] <- !IndBlock[ROWS == (COLS-1)]
```

```
## To eliminate the blocks of the matrices not needed in the summand of
## the probabilities, we set indicators of those blocks to be true so that
## they don't affect the results (Recall TRUE*x = x and FALSE*x = 0, where
## x can be any value). The rows are repeated B times so we adjust for
## that.
IndBlock[ROWS <= (COLS-2)] <- TRUE

## Take the row product of the Indicator variable to eliminate blocks of
## the matrices not needed in computing the probabilities.
Indicators <- rowProds(IndBlock)

## Obtain the probabilities in finding the critical value
pVec <- colMeans(matrix(Indicators, byrow = FALSE, nrow = B))
#print(pVec)

## Find all the c_j's
alpha.hat <- sum((1:j)/((s-j+1):s) * pVec)
return(alpha.hat)
})

t.alpha <- unlist(t.alpha)

## Obtain the critical value by finding the min of c_j's
c_min <- min(t.dat[start:B][t.alpha <= alpha])

## Obtain the corresponding alpha values of the critical values
```

```
c_alpha <- t.alpha[t.dat[start:B] == c_min]

## Find the critical values
cVal <- c(c_min, cVal)

## Obtain the corresponding alpha values of the critical values
alphaVal <-c(c_alpha, alphaVal)

#cVal; alphaVal
}

return(data.frame(j=s:1, c_j=cVal, alpha_j=alphaVal))
}
```

VITA

Josephine Sarpong Akosa

Candidate for the Degree of

Doctor of Philosophy

Dissertation: A UNIFIED APPROACH TO FALSE DISCOVERY RATE CONTROL UNDER DEPENDENCE THAT INCORPORATES NULL DISTRIBUTION AND SHRINKAGE ESTIMATION

Major Field: Statistics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Statistics at Oklahoma State University, Stillwater, Oklahoma, in December, 2018.

Completed the requirements for the Master of Science in Statistics at University of Texas at El Paso, El Paso, Texas, in 2014.

Completed the requirements for the Bachelor of Science in Mathematics at Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, in 2011.

Experience:

Graduate Teaching Associate, Oklahoma State University, Aug. 2014 – Dec. 2018. Course Instructor for Elementary Statistics, Intermediate Statistics and SAS Programming.

Graduate and Professional Student Government Association, Representative, Aug. 2016 – May 2017.

Data Science Intern, Nationwide Insurance, May 2016 – Aug. 2016.

Graduate Teaching Assistant, University of Texas at El Paso, Aug. 2012 – Aug. 2014.

Professional Memberships:

American Statistical Association (ASA).