Journal of Privacy and Confidentiality
Vol. 11 (2) 2021

Submitted Aug 2020
Published Aug 2021

DIFFERENTIALLY PRIVATE FALSE DISCOVERY RATE CONTROL

CYNTHIA DWORK[†], WEIJIE J. SU[‡], AND LI ZHANG[‡‡]

[†] Harvard University

[‡] University of Pennsylvania

[‡‡] Google Research, Inc.

Abstract. Differential privacy provides a rigorous framework for privacy-preserving data analysis. This paper proposes the first differentially private procedure for controlling the false discovery rate (FDR) in multiple hypothesis testing. Inspired by the Benjamini-Hochberg procedure (BHq), our approach is to first repeatedly add noise to the logarithms of the p-values to ensure differential privacy and to select an approximately smallest p-value serving as a promising candidate at each iteration; the selected p-values are further supplied to the BHq and our private procedure releases only the rejected ones. Moreover, we develop a new technique that is based on a backward submartingale for proving FDR control of a broad class of multiple testing procedures, including our private procedure, and both the BHq step-up and step-down procedures. As a novel aspect, the proof works for arbitrary dependence between the true null and false null test statistics, while FDR control is maintained up to a small multiplicative factor.

1. Introduction

With the growing availability of large-scale datasets, decision-making in healthcare, information technology, and government agencies is increasingly driven by data analyses. This data-driven paradigm, however, comes with great risk if the databases contain sensitive information of individuals such as health records or financial data. Without appropriate adjustments, statistical analysis applied to these databases can lead to privacy violation. For example, Homer et al. demonstrate that, under certain conditions, it is possible to determine whether an individual with a known genotype is in a genome-wide association study (GWAS) even when only minor allele frequencies are revealed [38]. Such privacy issues have serious implications: at best, individuals and agencies are discouraged from sharing their data for research purposes due to the concern of privacy leakage, impeding scientific progress [41]; at worst, potential adversaries could make use of sensitive information to jeopardize the social foundations of liberal democracy [64].

Key words and phrases: Differential Privacy, Report Noisy Max, False Discovery Rate, Benjamini–Hochberg Procedure, Positive Regression Dependence on Subset, Submartingale.

* The authors would like to thank Michal Linial, Daniel Rubin, Abba Krieger, Xinran Li, and Sanat Sarkar for helpful discussions and useful comments about an early version of the manuscript. C. D. and W. S. were supported in part by the NSF via grant CCF-1763314.

www.journalprivacyconfidentiality.org
DOI:10.29012/jpc.755

Being able to conduct data analysis in a way that preserves privacy, therefore, is key to removing barriers to scientific research while preventing breaches of personal data. First introduced by Dwork et al. [25], *differential privacy* (Definition 2.1) has put private data analysis on a rigorous foundation. A differentially private algorithm is required to hide the presence or absence of any individual or small group of individuals, the intuition being that an adversary unable to tell whether or not a given individual is even a member of the database surely cannot glean information specific to this individual. In computer science, considerable efforts have been made to develop private data release mechanisms [25, 47, 2] and private machine learning algorithms under differential privacy constraints, for example, boosting [35], empirical risk minimization [15], private PAC learning [4], and deep learning [1, 12]. On the statistical front, differential privacy has been added to and incorporated into many statistical methods in areas of robust statistics [24], nonparametric density estimation [63], hypothesis testing [61, 31], finite-sample confidence intervals [40], functional data analysis [34], network data analysis [39], and linear regression [43, 62].

In this paper, we provide the first differentially private multiple testing procedure. The problem of multiple testing arises in many privacy-sensitive applications such as a GWAS, where a large number of single-nucleotide polymorphisms (SNPs) are tested simultaneously for an association with a disease and the hope is to control some error rate for the significant SNPs. Perhaps the most popular error rate is the *false discovery rate* (FDR), which, roughly speaking, is the expected fraction of erroneously rejected hypotheses among all rejected hypotheses. This notion of type I error rate was introduced in the seminal work of Benjamini and Hochberg [5], along with the Benjamini–Hochberg procedure (BHq) that controls the FDR under certain conditions. This procedure is detailed in Algorithm 1.

Our interest in privacy-preserving FDR control arose as a group of researchers showed how to use one-way marginals, specifically, allele frequency statistics, together with the DNA of a target individual and allele frequency statistics for the general population, to determine the target's presence or absence in the study [38]. In response, the US National Institutes of Health and the Wellcome Trust changed the access policy to statistics of this type in the studies they fund. Although differential privacy has been shown to permit nontrivial estimates of very large numbers of statistical queries [10, 36], the errors introduced in these techniques are – and *must be* [14, 29] – too large for the (typical) setting, where the number of alleles exceeds the square of the number of data subjects.

Our procedure, which is referred to as PrivateBHq henceforth (Algorithm 4), is derived by recognizing the iterative nature of the BHq procedure and making each iteration differentially private. PrivateBHq provides *unconditional* end-to-end privacy. For now, regarding p-values as functions of a dataset, our proof of the privacy guarantees of PrivateBHq relies on a new definition of sensitivity that is tailor-made for p-values (Definition 2.5). Loosely speaking, this definition evaluates how *insensitive* p-values are to perturbations of any individual record in the database. All computations satisfy the definition for *some* choice of the privacy parameters, but not all choices of these parameters yield useful results when we enforce privacy. Popular examples of p-values are described in terms of these privacy parameters in Section 2.

Another contribution of this paper lies in our proof of the FDR control of PrivateBHq and beyond. In short, all existing proof strategies for FDR control are invalid for PrivateBHq. Thus, a new technique for proving FDR control is needed. To this end, we

(1) Develop a novel proof of FDR control for a class of multiple testing procedures, including the original (non-private) BHq and many of its variants – a proof requiring different assumptions than those found in the vast literature on this topic (see Section 3) – and
(2) Relate the FDR control and power properties of PrivateBHq to the corresponding properties of the non-private version.

The outline of the remainder of the paper is as follows. The next two subsections elucidate the two contributions, namely developing PrivateBHq and proving FDR control for a class of procedures, and the following subsection consolidates privacy and inferential properties together for PrivateBHq. To make this paper self-contained, in Section 2 we give a brief introduction to differential privacy, followed by the complete development of the PrivateBHq procedure. Section 3 is devoted to establishing FDR control of a broad class of multiple testing procedures and, as an application, Section 4 proves FDR control of PrivateBHq and argues its power as well. The paper is concluded by a discussion in Section 5.

Algorithm 1 BHq (Step-Up) Procedure

Input: nominal level $0 < q < 1$ and p-values $p_1, \ldots, p_m$
Output: a set of rejected hypotheses

```
sort the p-values in increasing order: p_(1) ≤ p_(2) ≤ ··· ≤ p_(m)
for j = m to 1 do
    if p_(j) > qj/m then
        continue
    else
        reject p_(1), ..., p_(j) and halt
    end if
end for
```

In words, the BHq (step-up) procedure finds the largest $j^\star$ such that $p_{(j^\star)} \le qj^\star/m$ and rejects all p-values below $qj^\star/m$.

1.1. **Making BHq private.** The original BHq is our starting point in developing the PrivateBHq procedure. The original procedure is non-private because the data of a single individual can affect the p-values of all hypotheses simultaneously, possibly changing the outcome of the BHq procedure dramatically.

To make the BHq private, for now we need two facts about differential privacy: (1) differential privacy is closed under composition, permitting us to bound the cumulative privacy loss over multiple differentially private computations. This allows us to build complex differentially private algorithms from simple differentially private primitives, and (2) we will make use of the well-known Report Noisy Max (respectively, Report Noisy Min) primitive [26], in which appropriately distributed fresh random noise is added to the result of each computation, and the index of the computation yielding the maximum (respectively, minimum) noisy value is returned. By returning only one index the procedure allows us to pay an accuracy price for a single computation rather than all computations.

A natural approach to obtaining a private version of BHq is by repeated use of Report Noisy Max: Starting with $j = m$ and decreasing: use Report Noisy Max to find the (approximately) largest p-value; estimate that p-value and, if the estimate is above a certain

more conservative critical value than qj/m, accept the corresponding null hypothesis, remove it from consideration, and repeat. Once a hypothesis is found with its p-value below the threshold, reject all the remaining hypotheses. The principal difficulty with this approach is that every iteration of the algorithm incurs a privacy loss, which can be mitigated only by increasing the magnitude of the noise used by Report Noisy Max. Since each iteration corresponds to the failure of rejecting a null hypothesis, this step-up procedure is paying in privacy precisely for all null hypotheses accepted, which are by definition not the "interesting" ones. Moreover, recognizing that most null hypotheses in a typical GWAS would be accepted, it is fundamentally difficult to preserve information content while protecting individual privacy by emulating the step-up procedure.

Instead of starting with the largest p-value and considering the values in decreasing order, another approach is to start with the smallest p-value and consider the values in increasing order, rejecting hypotheses one by one until we find a p-value above some threshold. This widely studied variant is called the BHq step-down procedure, which, in contrast to the aforementioned BHq step-up procedure, finds the largest j such that $p_{(i)} \leq qi/m$ for all $i \leq j$ and then rejects $p_{(1)}, \ldots, p_{(j)}$. Their definitions reveal that the step-down procedure shall be more conservative than its step-up counterpart. This variant, however, can assume less stringent critical values than the BHq critical values while still offering FDR control, often allowing more discoveries than the step-up counterpart [32].

If we make the natural modifications to the step-down procedure using Report Noisy Min, also known as the *Private Min* (Algorithm 2), instead of Report Noisy Max, then we pay a privacy cost only for nulls rejected in favor of the corresponding alternative hypotheses, which by definition are the "interesting" ones. Since the driving application of BHq is to select promising directions for future investigation that have a decent chance of panning out, we can view its outcome as advice for allocating resources. Thus, a procedure that finds a relatively small number of high-quality hypotheses, still achieving FDR control, may be as useful as a procedure that finds a much larger set.

1.2. **A new technique for proving FDR control.** While various techniques have been developed in the literature for proving FDR control, they are not applicable to privacy-preserving procedures. Any privacy-preserving procedure is necessarily randomized. Consequently, the jth most significant noisy p-value may not necessarily correspond to the jth most significant true p-value. Even worse, PrivateBHq may compare a noisy p-value to a critical value with a different rank and, as an inevitable result, a larger p-value may be rejected while a smaller p-value is accepted. This is in stark contrast to the (non-private) BHq and most of its variants, which reject p-values that are contiguous in sorted order.

These facts about the PrivateBHq procedure destroy some crucial properties for proving FDR control in existing approaches. For example, it is not clear how to adapt the elegant martingale technique for FDR control, proposed by Storey, Taylor, and Siegmund [57]. In essence, this approach is to construct an empirical process indexed by a threshold under which a p-value is rejected. In the case of PrivateBHq, unfortunately, no such threshold exists for singling out p-values for declaring significance. Another technique that appears frequently in the FDR control literature (see, for example, [7, 53, 30, 50, 11, 37]) is based on a crucial property of BHq: provided that a p-value is rejected, the effective threshold for declaring significance is completely determined by the remaining p-values. Unfortunately, this property is not satisfied by PrivateBHq either.

To pursue a new strategy for PrivateBHq, we observe that, although PrivateBHq might skip some of the minimum p-values, nevertheless it preserves a key property with high probability: if R rejections are made, the largest rejected p-value is roughly upper bounded by qR/m. This motivates us to give the following definition.

Definition 1.1. Given any cutoffs $0 < q_1 \leq q_2 \leq \cdots \leq q_m$, a multiple testing procedure is said to be *compliant* with $\{q_j\}_{j=1}^m$, if all rejected p-values are always bounded above by q_R, where R is the number of rejections.

In the case of no rejections ($R = 0$), as a convention, the (non-existent) rejected p-value is considered to be bounded above by q_R. Compliance is an instance of a more general condition termed *self-consistency* [8, 9], which, roughly speaking, requires that any rejected p-value be upper bounded by a general function of the total number of rejections. Interestingly, the compliance condition as a special instance of self-consistency has not been considered in the literature. Here, we prefer to use the compliance condition as self-consistency further allows a procedure to incorporate prior information about each hypothesis into the cutoffs, which is beyond the scope of this paper. Using the BHq critical values $\{qj/m\}$ as the cutoffs (referred to as BHq-compliance henceforth), however, our condition is sufficiently general to cover many classical multiple testing procedures, including both the step-down and step-up procedures, the generalized step-up-step-down procedures [60, 51] and particularly the PrivateBHq procedure (Proposition 4.1). The compliance condition is solely determined by the number of rejected p-values and the size of the largest one, without requiring that each rejected p-value be below its associated critical value. As a consequence, this condition permits skipping the smallest p-values and this is well-suited for differentially private procedures.

As revealed by this work, FDR control, roughly speaking, is a consequence of BHq-compliance together with the *independence with the null* condition (Definition 1.2). As such, our finding offers more than expected, applying to far more examples than PrivateBHq. In detail, we consider a generalized FDR [52, 54] defined as

$$\mathrm{FDR}_k := \mathbb{E}\left[\frac{V}{R}; V \geq k\right],$$

where V denotes the number of true null hypotheses that are falsely rejected (false discoveries). The present paper primarily focuses on the case of $k \geq 2$ and, whenever clear from the context, the term FDR control in this paper stands for FDR_k control. Note that FDR_k reduces to the usual FDR if the positive integer k is set to 1. This slightly relaxed FDR permits no more than $k-1$ false discoveries without any penalty, trading off for more power improvement while still maintaining a meaningful interpretation of the rejected hypotheses. The difference between the original FDR and FDR_k becomes negligible if the number of discoveries R is large. As an aside, we remark that the compliance condition is not satisfied by the FDR_k-controlling procedures developed in [52, 54].

Now we introduce the independence with the null condition, which is concerned with the distribution of the p-values. This condition is satisfied by the three examples in Appendix B.

Definition 1.2. A set of m test statistics are said to satisfy a condition referred to as *independence within the null*, or IWN for short, if the true null test statistics are jointly independent.

More elaboration on this new condition is carried out following Theorem 2 below.

With the two preparatory definitions in place, we offer the following theorem. Let m_0 denote the number of true null hypotheses and $\pi_0 := m_0/m$ be the true null proportion.

Theorem 1 . *If the test statistics obey the IWN condition, then any procedure that is compliant with the BHq critical values* $\{qj/m\}_{j=1}^{m}$ *must satisfy*

$$\mathrm{FDR}_k \le C_k \pi_0 q \tag{1.1}$$

for every $k \ge 2$*, where* C_k *is a universal constant.*

We immediately obtain the following corollary.

Corollary 1.3 . *If the test statistics obey the IWN condition, both the BHq step-up and step-down procedures satisfy* (1.1) *for* $k \ge 2$.

This bound involves an additional factor C_k, compared with the usual bound $\pi_0 q$ in the FDR literature. Explicitly, letting $\{\xi_j\}_{j=1}^{\infty}$ be i. i. d. exponential random variables with mean 1, the constant is given as

$$C_k = \mathbb{E}\left[\max_{j \ge k} \frac{j}{\xi_1 + \cdots + \xi_j}\right]. \tag{1.2}$$

For example, $C_2 \approx 2.41, C_3 \approx 1.85, C_{10} \approx 1.32$, and C_k tends to 1 as $k \to \infty$. In particular, C_1 defined in (1.2) is infinite, and this is exactly why Theorem 1 does not apply to the usual FDR.

Theorem 1 is optimal for all $k \ge 2$ as we show next.

Theorem 2 . *Given any* $C < C_k$*, if* q *is sufficiently small and* m *is sufficiently large, then there exists a BHq-compliant procedure applied to a set of IWN p-values such that*

$$\mathrm{FDR}_k > Cq.$$

In the literature, existing FDR-controlling procedures often assume independence between the true null and false null test statistics (see [5, 6]) or certain sophisticated correlation structures between these two sets of test statistics, such as the positive regression dependent on subset (PRDS) property [7, 42, 55] (see also [51, 9]). Roughly speaking, the PRDS property holds if the test statistics exhibit certain positive dependence on each true null test statistic. In particular, the dependence between true and false nulls cannot be arbitrary. For the sake of completeness, we emphasize that the literature has considered a few cases for FDR control with an arbitrary correlation between the two sets of test statistics [7, 8], but, unfortunately, the associated procedures are often extremely conservative. As a well-known example, Benjamini and Yekutieli showed in Theorem 1.3 of [7] that the BHq procedure gives FDR control using critical values at level $q/(1 + \frac{1}{2} + \cdots + \frac{1}{m}) \approx q/(\log m + 0.577)$ in place of q. In fact, BHq with this log-factor correction could be even more conservative than the Bonferroni method [45].

In contrast, Theorem 1 makes no assumptions regarding the dependence between the true nulls and false nulls while still controlling the FDR up to a small multiplicative factor, as the IWN condition is concerned only with the true nulls. As such, Theorem 1 is a contribution of independent interest to the vast FDR literature. Notably, the dependence can even be "adversarial" in the sense that the false null p-values can even be constructed as arbitrary functions of the true null p-values. This provides positive evidence toward understanding the robustness of the BHq procedure observed in a wide range of theoretical and empirical studies [56, 33, 16].

2. The PrivateBHq Procedure

In this section, we first introduce the differential privacy machinery at a minimal level and then focus on developing the PrivateBHq procedure.

2.1. Preliminaries on differential privacy.

A database $D = (d_1, d_2, \ldots, d_n) \in \mathcal{X}^n$ consists of n data items (for example, health records of n individuals), where $\mathcal{X}$ is a sample universe. Data items need not be independent (for example, health records of siblings). Two databases $D, D' = (d'_1, d'_2, \ldots, d'_n)$ are said to be *neighbors*, or *adjacent*, if they differ only in one data item. That is, there is exactly one j such that $d_j \neq d'_j$. A (randomized) mechanism $\mathcal{M}$ is an algorithm that takes a database as input and releases some (randomized) response of interest. We denote by range($\mathcal{M}$) the collection of all possible outputs of the mechanism $\mathcal{M}$. In the context of genome-wide association studies, a database D records genotypes of individuals, and $\mathcal{M}$, for example, is a mechanism that releases the minor allele frequency of a SNP plus some random noise.

Differential privacy, now sometimes called *pure differential privacy*, was defined and first constructed in [25]. The relaxation defined next is sometimes referred to as *approximate differential privacy*.

Definition 2.1 (Differential Privacy [25, 23])**.** A (randomized) mechanism $\mathcal{M}$ is (ε, δ)-differentially private for some nonnegative ε, δ if for all adjacent databases D, D' and for any measurable event $S \subset \text{range}(\mathcal{M})$,

$$\mathbb{P}(\mathcal{M}(D) \in S) \leq \mathrm{e}^{\varepsilon}\, \mathbb{P}(\mathcal{M}(D') \in S) + \delta.$$

Pure differential privacy is the special case where $\delta = 0$. In the definition above, both databases D, D' are fixed and the probabilities are taken over the randomness of the mechanism $\mathcal{M}$. The parameters ε and δ measure the desired privacy protection. With small ε and δ, this definition states that the likelihood of the released response is indifferent to changing a single individual in the database, thus leaking little indication of whether a particular individual is in the database even if all the other individuals are known. This provides strong privacy protection for each individual in and outside the database.

To report a statistic $f = f(D)$ in a differentially private manner, it is necessary to randomize the mechanism. As its name suggests, the *Laplace mechanism* preserves privacy by perturbing f with noise generated from the Laplace distribution $\text{Lap}(\lambda)$, whose probability density is $\exp(-|x|/\lambda)/(2\lambda)$. The scale $\lambda > 0$ should be calibrated to the *sensitivity* of the statistic f, defined as follows.

Definition 2.2. Let f be a real or vector valued function that takes as input a database. The sensitivity of f, denoted as Δf, is the supremum of $\|f(D) - f(D')\|_1$ over all adjacent D, D', where $\|\cdot\|_1$ denotes the ℓ_1 norm.

Formally, for any function f that maps databases to $\mathbb{R}^r$ for some positive integer r, we have the following result.

Lemma 2.3 Laplace Mechanism [25]**.** *The Laplace mechanism $\mathcal{M}_L$ that outputs*

$$\mathcal{M}_L(D; f) = f(D) + (Z_1, \ldots, Z_r)$$

preserves $(\varepsilon, 0)$-differential privacy, where Z_j are i. i. d. draws from $\text{Lap}(\Delta f/\varepsilon)$.

Intuitively, sensitivity quantifies the effect of any individual in the dataset on the outcome of the analysis. In this mechanism, Laplace noise with magnitude proportional to the sensitivity has the effect of masking the characteristics of any individual, thereby preserving privacy.

A simple algorithm that integrates the Laplace mechanism is the Private Min, which is better known as the Report Noisy Min in the literature [26] and will be the building block of PrivateBHq, introduced in Section 2.3. Consider a collection of scalar functions $f_1, \ldots, f_m$. The Private Min adds Laplace noise to each f_j and then reports the smallest noisy count (with fresh noise added) and its index. A formal description of Private Min is given in Algorithm 2. The following lemma concerns its privacy property.

Lemma 2.4 . *The Private Min, as detailed in Algorithm 2, is* $(\varepsilon, 0)$*-differentially private.*

A peek at the proof of this well known lemma, which for completeness appears in the appendix, reveals that reporting each of $j^\star$ and $f_{j^\star}(D) + Z$ is $(\varepsilon/2, 0)$-differentially private, hence leading to a total privacy loss of $(\varepsilon/2, 0) + (\varepsilon/2, 0) = (\varepsilon, 0)$. Here we have used the simple fact that differential privacy loss adds up under the composition of sequential mechanisms, that is, the union of the outputs of a sequence of mechanisms that each preserve $(\varepsilon_j, \delta_j)$-differential privacy is $(\sum \varepsilon_j, \sum \delta_j)$-differentially private [25]. As an aside, the Advanced Composition Theorem [27] (see Lemma 2.9 in Section 2.4), provides a much tighter bound on this privacy degradation.

Algorithm 2 Private Min (Report Noisy Min)

Input: database D, functions $f_1, \ldots, f_m$ each with sensitivity at most Δ, and privacy parameter ε

Output: index $j^\star$ and approximation to $f_{j^\star}(D)$

1: **for** $j = 1$ to m **do**
2: set $\widetilde{f_j} = f_j(D) + Z_j$, where Z_j is independently sampled from $\mathrm{Lap}(2\Delta/\varepsilon)$
3: **end for**
4: return $j^\star = \underset{j}{\operatorname{argmin}}\ \widetilde{f_j}$ and $f_{j^\star}(D) + Z$, where Z is a fresh draw from $\mathrm{Lap}(2\Delta/\varepsilon)$

Looking ahead, and omitting some technicalities, PrivateBHq will operate on differentially private approximations to the logarithms of p-values, returned by multiple invocations of Private Min. Since differential privacy is closed under post-processing [26], any subsequent computation on these differentially privately obtained values can never increase privacy loss. Thus, PrivateBHq is indeed differentially private, for all p-value functions. Its statistical properties will depend on the kinds of p-value computations that are performed, which we turn to next.

2.2. **Multiplicative sensitivity of p-values.** Multiple testing procedures ubiquitously act on a set of p-values that are computed by functions that operate on databases. A p-value in our context is frequently referred to as the function on databases for computing the p-value instead of its numerical value, in contrast with the vast statistical literature that often does not distinguish between the function that maps a database to a p-value and the result of the mapping.

We now consider making p-value computations private as the first step toward developing a private multiple testing procedure. In many important p-value computations (see Example

2.6), a larger p-value is affected more in magnitude by the change of a single data item than a smaller p-value. As a result, directly adding noise to the p-values may overprotect privacy and completely overwhelm signals in small p-values. This would inevitably lead to significant detection power loss as the smallest p-values are more likely to correspond to promising hypotheses.

Our solution will be to (very carefully) work with the logarithms of the p-values. This strategy is motivated by the observation that, although the (additive) sensitivity of a p-value may vary greatly, oftentimes the relative change (that is, the ratio) of a p-value on two neighboring databases is very stable, regardless of the magnitude of the p-value, unless it is extremely small. In light of this observation, the sensitivity of a p-value, that is, the worst-case change due to the replacement of an individual in the database, is best measured multiplicatively. Below, η and ν are nonnegative.

Definition 2.5 (Multiplicative Sensitivity)**.** A p-value function p is said to be (η, ν)-*multiplicatively sensitive*, or (η, ν)-*sensitive* for short, if for all adjacent databases D and D', either both $p(D), p(D') \leq \nu$ or

$$\mathrm{e}^{-\eta} p(D) \leq p(D') \leq \mathrm{e}^{\eta} p(D).$$

Our PrivateBHq algorithm will make explicit use of both parameters in ensuring privacy. The parameter ν is introduced in recognition of the fact that a very small p-value may jump or fall by a relatively large multiplicative factor between adjacent databases. This parameter is normally much less than the Bonferroni level q/m (see, for example, [19]), resulting in essentially no power loss for truncating p-values at ν. A p-value can satisfy different pairs of (η, ν)-multiplicative sensitivities. In short, the two parameters η and ν exhibit a certain trade-off relationship in the sense that one can increase (resp. decrease) η and decrease (resp. increase) ν in a careful way such that a p-value still satisfies this condition. Every p-value satisfies (η, ν)-sensitivity for *some* values of the parameters. Moreover, given p-value functions p_1, p_2 with multiplicative sensitivities (η_1, ν_1) and (η_2, ν_2) respectively, it is immediate that both functions satisfy $(\max\{\eta_1, \eta_2\}, \max\{\nu_1, \nu_2\})$-sensitivity, so given a collection of p-values there always exist η, ν so that all of the p-values in the collection are (η, ν)-sensitive.

Given an (η, ν)-sensitivity p-value function p and a database D, we work with the logarithmic mapping

$$\theta(D; p, \nu) = \log \max\{\nu, p(D)\}$$

This statistic satisfies $\theta(D) - \eta \leq \theta(D') \leq \theta(D) + \eta$ for all neighboring databases D, D'. In other words, θ has an *additive* sensitivity bounded by η. Hence, Lemma 2.2 ensures that adding Laplace noise $\mathrm{Lap}(\eta/\varepsilon)$ to $\theta(D)$ preserves $(\varepsilon, 0)$-differential privacy.

We will see below via examples that two large and important classes of p-value computations are (η, ν)-sensitive for some small η and ν, with rigorous proofs given in Appendix A; as a consequence of this, preserving privacy for these p-values only requires a small amount of noise, leading to negligible accuracy loss. Recall that m denotes the total number of hypotheses.

Example 2.6 (Binomial Distribution)**.** Suppose the n individuals in D are, respectively, associated with n i. i. d. Bernoulli variables $\xi_1, \ldots, \xi_n$, each of which takes the value 1 with probability α and the value 0 otherwise. Let T denote the sum. A p-value $p(D)$ for testing

$H_0 : \alpha \le \frac{1}{2}$ against the alternative $H_1 : \alpha > \frac{1}{2}$ is defined as

$$p(D) = \sum_{i=t}^{n} \frac{1}{2^n} \binom{n}{i},$$

where t is the realization of T on the database D. Denote by t' the counterpart of t on a neighboring database D'. Without loss of generality, assume $t' = t + 1$. The difference between the two p-values, $|p(D) - p'(D)| = \frac{1}{2^n}\binom{n}{t}$, attains its maximum at $t = \lfloor n/2 \rfloor$ or $\lfloor (n+1)/2 \rfloor$ ($\lfloor x \rfloor$ denotes the greatest integer that is less than or equal to x) and decays rapidly as t deviates from $n/2$. This implies that additive sensitivity is not a good measure of the variability of this p-value construction.

Instead, we fix a (very) small ν and denote by η the maximum of $\log \frac{p(D)}{p(D')}$ subject to the constraint $p(D') \ge \nu$. The p-value by definition is (η, ν)-sensitive. To evaluate η, observe that the log-likelihood ratio

$$\log \frac{p(D)}{p(D')} = \log \frac{\sum_{i=t}^{n} \frac{1}{2^n}\binom{n}{i}}{\sum_{i=t+1}^{n} \frac{1}{2^n}\binom{n}{i}} = \log \left[1 + \frac{\binom{n}{t}}{\sum_{i=t+1}^{n} \binom{n}{i}} \right] \le \frac{\binom{n}{t}}{\sum_{i=t+1}^{n} \binom{n}{i}}.$$

In the appendix, it is shown that $\binom{n}{t} / \sum_{i=t+1}^{n} \binom{n}{i} \lesssim \sqrt{\frac{\log n}{n}}$ under the constraint $p(D') \ge m^{-1-c}$ for any small constant $c > 0$ if $m \le \texttt{poly}(n)$ (that is, m grows at most polynomially in n) as $n \to \infty$. Therefore, we can set $\nu = m^{-1-c}$ and $\eta \asymp \sqrt{\frac{\log n}{n}}$. Note that this choice of ν is much below the Bonferroni level q/m.

Example 2.7 (Truncated Exponential Distribution)**.** Let $\zeta_1, \ldots, \zeta_n$ be i. i. d. random variables sampled from the density $\frac{\lambda e^{-\lambda x}}{1 - e^{-A\lambda}} \cdot \mathbf{1}(0 \le x \le A)$ for positive A and λ, an exponential distribution truncated at A. Denote by $T = \zeta_1 + \cdots + \zeta_n$ the sum (T is a sufficient statistic for λ). To test $H_0 : \lambda = 1$ against the alternative $H_1 : \lambda > 1$, we consider the p-value $p(D) = \mathbb{P}_{\lambda=1}(T \ge t)$, where t is the realization of T (note that the value t differs at most by A between adjacent databases). With the same notations as in Example 2.6, this p-value is (η, ν)-multiplicatively sensitive with $\nu = m^{-1-c}$ and $\eta \asymp \sqrt{\frac{\log n}{n}}$ for any small constant $c > 0$. Similarly, the analysis applies to the case of a Gaussian distribution. In short, consider i. i. d. random variables $\xi_1, \ldots, \xi_n$ drawn from the normal distribution $\mathcal{N}(\mu, 1)$ truncated at $-A$ and A, which has density $e^{-(x-\mu)^2/2} / \int_{-A}^{A} e^{-(u-\mu)^2/2} du$. Writing $T = \xi_1 + \cdots + \xi_n$, we use the p-value $p(D) = \mathbb{P}_{\mu=0}(|T| \ge |t|)$ to test $H_0 : \mu = 0$ against $H_1 : \mu \ne 0$ (t is the realization of T). Using the same proof strategy as for the exponential distribution, one can show that this p-value strategy is (η, ν)-multiplicatively sensitive with some $\nu = m^{-1-c}$ and $\eta \asymp \sqrt{\frac{\log n}{n}}$.

We remark that (η, ν)-sensitivity is a worst-case guarantee on the sensitivity of a p-value function. Only the interpretation of the p-value requires the i.i.d. assumption. Regarding the above-mentioned two examples, the asymptotic expressions of the privacy parameter η can be easily made precise.

As seen in both examples, the parameter η vanishes roughly at the rate $O(n^{-1/2})$, implying that less noise is required for privacy protection as the sample size becomes larger. This appealing feature is impossible without the restriction $p \ge \nu$ for some appropriate choice of ν. Specifically, in the absence of this constraint, or equivalently by setting $\nu = 0$,

we shall have $\eta = n + 1$ in the first example and $\eta = \infty$ in the second, requiring a vast or even an infinite amount of multiplicative noise for preserving privacy. This would completely dilute any signal of interest. To be complete, we note that not all p-value computations necessarily lead to vanishing η and ν as $n \to \infty$. An example from [61, 65] considers a privacy-preserving release of χ^2-statistics computed from allelic contingency tables. For the sake of simplicity, here we consider 2×2 contingency tables with $n/2$ cases and $n/2$ controls:

	allele type	
	major	minor
case	a	$\frac{n}{2} - a$
control	a	$\frac{n}{2} - a$

	allele type	
	major	minor
case	$a + 1$	$\frac{n}{2} - a - 1$
control	a	$\frac{n}{2} - a$

TABLE 1. Two neighboring allelic contingency tables.

In the case of a fixed $a > 5$, one can show that the two p-values computed from the two tables neither differ by a negligible factor nor both tend to zero as $n \to \infty$. This fact is elaborated in detail in the appendix.

2.3. **Developing PrivateBHq.** The PrivateBHq procedure (Algorithm 4) is the sequential composition of Algorithm 3, which we refer to as the peeling mechanism, denoted as `peeling`. In a little more detail, given (non-private) p-value functions $p_1, \ldots, p_m$ and a prescribed number of invocations $m' \leq m$, PrivateBHq first applies Private Min m' times to the logarithms of the p-values, "peeling off" and removing from further consideration the approximately smallest element with each new invocation of Private Min. These m' pre-selected hypotheses are thought of as promising hypotheses. In particular, the number m' as an upper bound on the total number of discoveries shall be much less than m. This recognizes that, in many application scenarios, much fewer are truly significant in an ocean of mediocre hypotheses.

During the peeling procedure, in order to keep track of indices within the original set, `peeling` removes a function from further consideration by redefining it to be $+\infty$, ensuring that it will not be returned by future invocations of the Private Min. The Laplace noise scale λ shall be chosen to adjust for the privacy protection target, factoring in the multiplicative sensitivities of $p_1, \ldots, p_m$ and the number of invocations m'.

Algorithm 3 Peeling Mechanism `peeling`

Input: database D, functions $f_1, \ldots, f_m$, number of invocations m', and Laplace noise scale λ

Output: indices $i_1, \ldots, i_{m'}$ and approximations to $f_{i_1}(D), \ldots, f_{i_{m'}}(D)$

1: **for** $j = 1$ to m' **do**
2: let $(i_j, \tilde{f}_{i_j}(D))$ be returned by Private Min applied to $(D, f_1, \ldots, f_m)$ with Laplace noise scale λ
3: set $f_{i_j} \equiv +\infty$
4: **end for**
5: return the m'-tuple $\{(i_1, \tilde{f}_{i_1}(D)), \ldots, (i_{m'}, \tilde{f}_{i_{m'}}(D))\}$

With m' hypotheses yielded by `peeling` in place, PrivateBHq supplies quantities in logarithmic scale instead of, in the conventional setting, the m' raw p-values and critical values to the (step-up) BHq procedure. This difference however does not affect the way BHq proceeds. To be concrete, BHq first orders the noisy values $\tilde{\theta}_{i_1}, \ldots, \tilde{\theta}_{i_{m'}}$ as $\tilde{\theta}_{(i_1)} \leq \cdots \leq \tilde{\theta}_{(i_{m'})}$, and then rejects any corresponding hypotheses if $\tilde{\theta}_{i_j}$ is below $\max\{\gamma_j : \tilde{\theta}_{(i_j)} \leq \gamma_j\}$, with the convention that $\max \emptyset = -\infty$. As we will see in Section 4, the cutoffs $\gamma_1, \ldots, \gamma_{m'}$ are chosen specifically to ensure FDR control of PrivateBHq; roughly speaking, γ_j is slightly below the logarithm of the corresponding BHq critical value qj/m, where the gap between the two accounts for the multiplicative sensitivity of the p-values and the uncertainty brought by the Laplace mechanism.

Algorithm 4 The PrivateBHq Procedure

Input: database D, parameters $\varepsilon, \delta, \eta, \nu$, (η, ν)-multiplicatively sensitive p-value functions $p_1, \ldots, p_m$, number of invocations m', Laplace noise scale $\lambda = \lambda(\varepsilon, \delta, \eta, m')$, and cutoffs $\gamma_1 < \cdots < \gamma_{m'}$

Output: a set of up to m' rejected hypotheses

1: set $\theta_j = \log \max\{\nu, p_j(D)\}$ for $1 \leq j \leq m$
2: obtain $(i_1, \tilde{\theta}_{i_1}), \ldots, (i_{m'}, \tilde{\theta}_{i_{m'}})$ by applying `peeling` to $\theta_1, \ldots, \theta_m$ with noise scale λ
3: apply (step-up) BHq to $\tilde{\theta}_{i_1}, \ldots, \tilde{\theta}_{i_{m'}}$ with cutoffs $\gamma_1, \ldots, \gamma_{m'}$
4: return the indices of rejected hypotheses

2.4. **Preserving privacy.** The proof that PrivateBHq is differentially private relies on the fact that the algorithm only accesses the data through the values returned by `peeling`. Thus, intuitively, the final results reported by BHq shall release no more privacy than the intermediate results yielded by `peeling`. This intuition is indeed true, that is, differential privacy is closed under *post-processing*, as shown by the following lemma.

Lemma 2.8 [23, 63]**.** *Let $\mathcal{M}$ be an (ε, δ)-differentially private mechanism and g be any (measurable) function. Then $g(\mathcal{M})$ also preserves (ε, δ)-differential privacy.*

This lemma implicitly assumes the range of the mechanism $\mathcal{M}$ falls into the domain of g. In our context, taking g to be step-up BHq, Lemma 2.8 shows that it suffices to establish the differential privacy property of `peeling`. By construction, each θ_j has sensitivity no more than η. Lemma 2.4 then immediately ensures that the Private Min, which is invoked sequentially m' times in PrivateBHq, guarantees on its own $(2\eta/\lambda, 0)$-differential privacy. Making use of the fact that, at worst "(ε, δ)'s add up" (see the discussion right below Lemma 2.4), one can conclude that the peeling mechanism is $(2m'\eta/\lambda, 0)$-differentially private. Equivalently, to achieve $(\varepsilon, 0)$-differential privacy for `peeling`, and therefore also for PrivateBHq, we can set the Laplace noise scale to be $\lambda = 2m'\eta/\varepsilon$. In this way, the noise level grows linearly with m'.

Surprisingly, we can trade a little bit of δ for a significant improvement on ε, as shown by the lemma below.

Lemma 2.9 Advanced Composition [27]**.** *For all $\varepsilon, \delta \geq 0$ and $\delta' > 0$, running l mechanisms sequentially that are each (ε, δ)-differentially private preserves $(\varepsilon\sqrt{2l\log(1/\delta')} + l\varepsilon(\mathrm{e}^{\varepsilon} - 1), l\delta + \delta')$-differential privacy.*

This lemma holds no matter how each mechanism adaptively depends on information released by prior mechanisms. Taking $\delta = 0$ in Lemma 2.9, we easily obtain the main theorem of this section, with its proof deferred to the appendix. This theorem shows adding Laplace noise with scale of order roughly $O(\sqrt{m'})$ is sufficient for protecting privacy of PrivateBHq.

Theorem 3 . *Let η, ν be chosen so that all the p-value functions input to PrivateBHq are (η, ν)-sensitive. Given $\varepsilon \le 0.5, \delta \le 0.1$ and $m' \ge 10$, PrivateBHq with Laplace noise scale $\lambda = \eta\sqrt{10m' \log(1/\delta)}/\varepsilon$, or larger, is (ε, δ)-differentially private.*

We remark that the constraints on ε, δ, and m' are used to optimize the constants for practical use.

3. Proving FDR Control Using a Submartingale

The main purpose of this section is to prove Theorem 1. The proof strategy contains two novel elements: an upper bound on FDR_k involving only true null p-values (Equation (3.1) below) and a backward submartingale that allows us to use a martingale maximal inequality. In addition, this section attempts to obtain the optimal constant C_k for Theorem 1 in Section 3.2, where we give some intuition behind Theorem 2, and considers a new variant of the FDR in Section 3.3.

Throughout the section, we focus on an arbitrary BHq-compliant procedure. That is, any p-value rejected by the procedure is not greater than qR/m, where R denotes the total number of rejections.

3.1. Controlling FDR_k.

In this subsection, we prove Theorem 1. However, the proof presented here does not seek to optimize the constant C_k in Theorem 1. We consider

$$\mathrm{FDP}_k := \frac{V\mathbf{1}_{V \ge k}}{R},$$

which gives $\mathrm{FDR}_k \equiv \mathbb{E}\,\mathrm{FDP}_k$ by taking expectation. The following upper bound on the FDP_k for $k \ge 2$ of the BHq-compliant procedure serves as the basis for our analysis:

$$\mathrm{FDP}_k \le \max_{k \le j \le m_0} \frac{qj}{mp^0_{(j)}}. \tag{3.1}$$

Above, $p^0_{(1)} \le p^0_{(2)} \le \cdots \le p^0_{(m_0)}$ are the order statistics of the m_0 true null p-values. To prove (3.1), denote by V the number of false rejections. If $V \le k-1$, (3.1) holds since $\mathrm{FDP}_k = 0$. Otherwise, the largest rejected true null p-value is at least $p^0_{(V)}$ and, therefore, one must have $p^0_{(V)} \le qR/m$ due to the compliance condition. As a consequence, we get

$$\mathrm{FDP}_k = \frac{V}{R} \le \frac{V}{mp^0_{(V)}/q} \le \max_{k \le j \le m_0} \frac{qj}{mp^0_{(j)}}. \tag{3.2}$$

The IWN condition imposed in Theorem 1 ensures the joint independence of the true null p-values, each of which is, by definition, stochastically larger than or equal to $U(0,1)$. Thus, the ordered true null p-values can be replaced by the order statistics $U_{(1)} \le U_{(2)} \le \cdots \le U_{(m_0)}$

of m_0 i. i. d. uniform random variables on $(0,1)$, while (3.2) remains true in the expectation sense (recall that $\pi_0 = m_0/m$):

$$\mathrm{FDR}_k \le \mathbb{E}\left[\max_{k\le j\le m_0} \frac{qj}{mU_{(j)}}\right] = q\pi_0\,\mathbb{E}\left[\max_{k\le j\le m_0} \frac{j}{m_0 U_{(j)}}\right].$$

Therefore, Theorem 1 follows from the lemma below.

Lemma 3.1 . *Let $U_{(1)} \le \cdots \le U_{(n)}$ denote the order statistics of n i. i. d. uniform variables on $(0,1)$. There exists an absolute constant c_k such that*

$$\sup_{n\ge k} \mathbb{E}\left[\max_{k\le j\le n} \frac{j}{nU_{(j)}}\right] \le c_k$$

for $k \ge 2$.

The proof of this lemma starts by recognizing a well-known representation in law for uniform order statistics:

$$(U_{(1)},\ldots,U_{(n)}) \overset{d}{=} \left(\frac{T_1}{T_{n+1}},\ldots,\frac{T_n}{T_{n+1}}\right), \tag{3.3}$$

where $T_j = \xi_1 + \cdots + \xi_j$ and $\xi_1,\ldots,\xi_{n+1}$ are i. i. d. exponential random variables with mean 1. Writing

$$W_j = \frac{jT_{n+1}}{T_j},$$

Lemma 3.1 is equivalent to showing

$$\mathbb{E}\left[\max_{k\le j\le n} \frac{W_j}{n}\right] \le c_k. \tag{3.4}$$

Intuitively, the maximum is likely to be attained at some small index j as W_j/n is close to 1 for a large value of j, due to the law of large numbers. This intuition can be indeed made rigorous by the fact that $W_1,\ldots,W_{n+1}$ is a backward submartingale, as shown by the following lemma.

Lemma 3.2 . *With respect to the filtration $\mathcal{F}_j := \sigma(T_j, T_{j+1},\ldots,T_{n+1})$ for $j = 1,\ldots,n+1$, the stochastic process $W_1,\ldots,W_{n+1}$ is a backward submartingale. That is, $\mathbb{E}(W_j|\mathcal{F}_{j+1}) \ge W_{j+1}$ for $j = 1,\ldots,n$.*

The proof of Lemma 3.2 is deferred to the appendix. Next, we apply this lemma to prove (3.4) (hence Lemma 3.1 follows immediately) using the following martingale maximal inequality (for a proof, see pages 71–73 of [48]).

Lemma 3.3 ℓ_1 Martingale Maximal Inequality**.** *Let $X_1,\ldots,X_n$ be a (forward) submartingale. Then,*

$$\mathbb{E}\left(\max_{1\le j\le n} X_j\right) \le \frac{\mathrm{e}}{\mathrm{e}-1}\left[1 + \mathbb{E}\left(X_n \log X_n; X_n \ge 1\right)\right].$$

Proof of Lemma 3.1. Since Lemma 3.2 asserts that W_j/n is a backward submartingale, Lemma 3.3 concludes

$$\begin{aligned}\mathbb{E}\left(\max_{k\le j\le n} \frac{W_j}{n}\right) &\le \frac{\mathrm{e}}{\mathrm{e}-1}\left[1 + \mathbb{E}\left(\frac{W_k}{n}\log\frac{W_k}{n}; \frac{W_k}{n}\ge 1\right)\right] \\ &= \frac{\mathrm{e}}{\mathrm{e}-1}\left[1 + \mathbb{E}\left(\frac{k}{nU_{(k)}}\log\frac{k}{nU_{(k)}}; \frac{k}{nU_{(k)}}\ge 1\right)\right].\end{aligned}$$

To complete the proof, it suffices to show that for a fixed k the expectation above involving $k/(nU_{(k)})$ is uniformly bounded for all $n \geq k$. To this end, observe that $U_{(k)}$ is distributed as $\mathrm{Beta}(k, n+1-k)$, and this allows us to evaluate the expectation as

$$\begin{aligned}
\mathbb{E}\left(\frac{k}{nU_{(k)}}\log\frac{k}{nU_{(k)}}; \frac{k}{nU_{(k)}} \geq 1\right) &= \int_0^{\frac{k}{n}} \frac{x^{k-1}(1-x)^{n-k}}{\mathrm{B}(k, n+1-k)} \frac{k}{nx}\log\frac{k}{nx}\,\mathrm{d}x \\
&\leq \int_0^{\frac{k}{n}} \frac{x^{k-1}}{\mathrm{B}(k, n+1-k)} \frac{k}{nx}\log\frac{k}{nx}\,\mathrm{d}x \\
&= \frac{1}{n^k \mathrm{B}(k, n+1-k)} \int_0^k k y^{k-2}\log\frac{k}{y}\,\mathrm{d}y \\
&= \frac{1}{n^k \mathrm{B}(k, n+1-k)} \cdot \frac{k^k}{(k-1)^2}.
\end{aligned}$$

To obtain an upper bound that is independent of n, it suffices to show that $n^k \mathrm{B}(k, n+1-k)$ has a lower bound depending only on k. Indeed, this is the case:

$$\begin{aligned}
n^k \mathrm{B}(k, n+1-k) &= n^k \frac{\Gamma(k)\Gamma(n+1-k)}{\Gamma(n+1)} \\
&= \frac{n^k (k-1)!}{n(n-1)\cdots(n-k+1)} \geq (k-1)!.
\end{aligned}$$

□

3.2. **Optimizing the bounds.** The constant C_k in Theorem 1 matters from a practical perspective. This section is aimed at finding the *optimal* constants for all $k \geq 2$. Compared with what has been performed in Section 3.1, this improvement is based on a delicate property about the expectation in (3.4), as detailed by the following lemma.

Lemma 3.4 . *Define* $C_k^{(n)} = \mathbb{E}\left[\max_{k\leq j\leq n} \frac{j}{nU_{(j)}}\right]$ *for* $n \geq k \geq 2$*, where* $U_{(j)}$*'s are the order statistics of* n *i. i. d. uniform variables on* $(0,1)$*. Then,* $C_k^{(n)} \leq C_k^{(n+1)}$*.*

The monotonicity in Lemma 3.4 reveals that the optimal C_k in (3.4) takes the form (recall that $T_j = \xi_1 + \cdots + \xi_j$ is defined in (3.3))

$$C_k := \lim_{n\to\infty} C_k^{(n)} = \lim_{n\to\infty} \mathbb{E}\left[\max_{k\leq j\leq n} \frac{jT_{n+1}}{nT_j}\right]. \tag{3.5}$$

Note that C_k does not seem to admit a closed-form expression. Nevertheless, this optimal constant can be easily computed via simulations.

While relegating the full proof of Theorem 2 to Appendix A, here we provide a proof sketch based on the construction of a BHq-compliant procedure and a set of p-values satisfying the IWN condition to show the optimality of C_k. Explicitly, let the true null p-values be m_0 i. i. d. uniform variables $U_1, \ldots, U_{m_0}$ between 0 and 1, and let all the $m - m_0$ false null p-values be 0. Denote by $j^\star$ the index $k \leq j \leq m_0$ that maximizes $j/U_{(j)}$. The BHq-compliant procedure rejects the $j^\star$ smallest true null p-values and any $\max\{\lceil mU_{(j^\star)}/q\rceil - j^\star, 0\}$ of the false null p-values ($\lceil x\rceil$ denotes the least integer that is greater than or equal to x), which by construction are all 0. This procedure is compliant (self-consistent) but not *nonincreasing* (a procedure is called nonincreasing if it never rejects more if some p-value gets larger),

so the FDR-controlling results in [9] do not apply to our case. Taking q sufficiently small and assuming that $m - m_0$ is sufficiently large, we get $\mathrm{FDP}_k \approx qj^\star/(mU_{(j^\star)})$ with high probability. Consequently, we get

$$\mathrm{FDR}_k \approx \mathbb{E}\left[\frac{qj^\star}{mU_{(j^\star)}}\right] = \mathbb{E}\left[\max_{k\le j\le m_0} \frac{j}{m_0 U_{(j)}}\right] \pi_0 q = C_k^{(m_0)} \pi_0 q,$$

which tends to $C_k q$ by taking $m_0 \to \infty$ and $m_0/m \to 1$.

For the moment, suppose the limit can be taken under the expectation in (3.5). As such, the optimal constant for FDR_k is

$$C_k = \mathbb{E}\left[\lim_{n\to\infty} \max_{k\le j\le n} \frac{jT_{n+1}}{nT_j}\right] = \mathbb{E}\left[\max_{k\le j<\infty} \frac{j}{T_j}\right], \tag{3.6}$$

where the last equality results from applying the strong law of large numbers to T_n/n. Recognizing that the integrable random variable $\max_{k\le j<\infty} j/T_j$ decreases to 1 almost surely as k increases to infinity, Lebesgue's dominated convergence theorem readily asserts that $C_k = 1 + o_k(1)$, where $o_k(1)$ denotes a sequence of numbers tending to 0 as $k \to \infty$. This is formally stated in the proposition below, where we consider a sequence of multiple testing problems indexed by l such that both $m_l, k_l \to \infty$ as $l \to \infty$.

Proposition 3.5 . *Under the assumptions of Theorem 1, as $k \to \infty$, we have* $\mathrm{FDR}_k \le (1 + o_k(1))q$.

To make the derivation of the optimal C_k above rigorous, we must validate (3.6). In fact, the Vitali convergence theorem together with the following lemma ensures that the limit $\lim_{n\to\infty}$ and expectation $\mathbb{E}$ can be interchanged.

Lemma 3.6 . *For a fixed $k \ge 2$, the sequence of random variables*

$$\max_{k\le j\le n} \frac{jT_{n+1}}{nT_j}$$

are uniformly integrable for $n \ge k$.

While the proof of Lemma 3.6 is deferred to the appendix, the proof of Lemma 3.4 is given below.

Proof of Lemma 3.4. Denote by $U_{(1)} \le \cdots \le U_{(n)} \le U_{(n+1)}$ the order statistics of $n+1$ i. i. d. uniform random variables on $(0,1)$. Then, $U_{(1)}/U_{(n+1)} \le \cdots \le U_{(n)}/U_{(n+1)}$ are distributed the same as the order statistics of n i. i. d. uniform random variables on $(0,1)$

and, moreover, are independent of $U_{(n+1)}$. Making use of this fact, we get

$$\begin{aligned} C_k^{(n+1)} &= \mathbb{E}\left[\max_{k\le j\le n+1} \frac{j}{(n+1)U_{(j)}}\right] \\ &\ge \mathbb{E}\left[\max_{k\le j\le n} \frac{j}{(n+1)U_{(j)}}\right] \\ &= \mathbb{E}\left[\frac{n}{(n+1)U_{(n+1)}} \cdot \max_{k\le j\le n} \frac{j}{nU_{(j)}/U_{(n+1)}}\right] \\ &= \mathbb{E}\left[\frac{n}{(n+1)U_{(n+1)}}\right] \mathbb{E}\left[\max_{k\le j\le n} \frac{j}{nU_{(j)}/U_{(n+1)}}\right] \\ &= \mathbb{E}\left[\frac{n}{(n+1)U_{(n+1)}}\right] C_k^{(n)}. \end{aligned}$$

Since the density of $U_{(n+1)}$ is $(n+1)x^n$ for $0 < x < 1$, we readily see that

$$\mathbb{E}\left[\frac{n}{(n+1)U_{(n+1)}}\right] = 1.$$

This completes the last step in certifying $C_k^{(n+1)} \ge C_k^{(n)}$. □

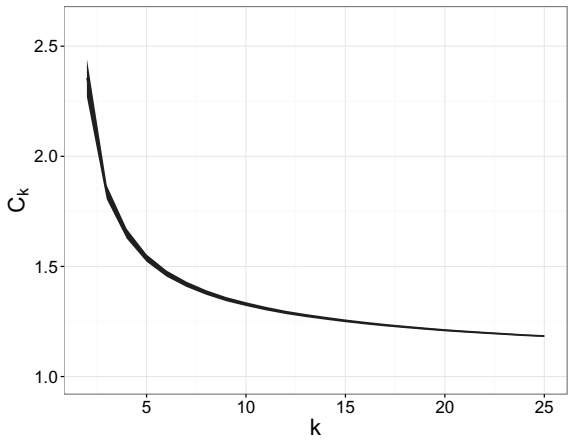


FIGURE 1. Monte Carlo simulated values of C_k using (3.6). The solid line indicates the maximum of j/T_j over $k \le j \le 10^5$, averaged over 10^4 runs. The (tiny) shaded band illustrates the 99%-coverage confidence interval for each k using normal approximation.

Now, we turn to numerically evaluate C_k using the expression (3.6). Although the distribution of each j/T_j admits an analytical expression, it is however not clear how to calculate the distribution of the maximum of j/T_j over j. In view of this difficulty, we resort to Monte Carlo simulations, and Figure 1 presents the results that are averaged over 10^4 independent replicates. For instance, $C_2 \approx 2.41, C_3 \approx 1.85, C_4 \approx 1.65, C_5 \approx 1.54$, and $C_{25} \approx 1.18$. In passing, we remark that the estimated values of C_k as a function of k are fairly accurate as indicated by the uniformly short widths of the confidence intervals for all k.

3.3. **Controlling** FDR^k**.** To further leverage the martingale-based proof idea, we consider a variant of the FDR defined as

$$\mathrm{FDR}^k := \mathbb{E}\left[\frac{V}{R}; R \geq k\right],$$

which includes the usual FDR as an example by taking $k = 1$. This relaxed FDR differs insignificantly from the usual FDR if a large number of discoveries are expected, which is often the case in modern multiple testing applications such as genome-wide association studies. For the moment, we do not intend to advocate the use of this new FDR definition in practice as it is clear that future investigation is needed.

In the following, we aim to prove Theorem 4, a counterpart of Theorem 1 for the FDR^k. A similarity between the two theorems lies in that their proofs both make use of martingale arguments. That being said, the bound on the FDR_k in Theorem 1 cannot carry over to the FDR^k because $\mathrm{FDR}_k \leq \mathrm{FDR}^k$.

Theorem 4 . *If the test statistics obey the IWN condition, then any BHq-compliant procedure satisfies*

$$\mathrm{FDR}^k \leq \left(1 + \frac{2}{\sqrt{qk}}\right) q.$$

for any $k \geq 1$.

A number of remarks are as follows. This theorem allows us to take $k = 1$, thus giving a bound on the usual FDR: $\mathrm{FDR} \leq q + 2\sqrt{q}$. For example, we can set $q = 0.0024$ if the FDR is aimed to be controlled at 10%. Such a bound is not available in Theorem 1. For completeness, the bound for $k = 1$ might not be sharp since Doob's ℓ^2 martingale maximal inequality used in the proof of Theorem 4 is generally not sharp. Indeed, this bound can be improved using a careful treatment of (3.1) (see [59]). For $k \geq 2$, the bound here is larger than that in Theorem 1, namely $2/\sqrt{qk} \geq C_k - 1$, due to the optimality of C_k and the fact $\mathrm{FDR}^k \geq \mathrm{FDR}_k$. The following proof actually establishes a stronger bound, $\pi_0 q + 2\sqrt{\pi_0 q/k}$, on the FDR^k. Recall that π_0 is the true null proportion m_0/m.

Proof of Theorem 4. Due to the compliance condition, the number of false discoveries satisfies

$$V \leq \#\left\{i \text{ is true null} : p_i \leq \frac{qR}{m}\right\}.$$

Thus, we get an upper bound on $\mathrm{FDP} := \frac{V}{R}$ (with the convention $0/0 = 0$) that takes the following form:

$$\mathrm{FDP} \leq \max_{R \leq j \leq m} \frac{\#\{i \text{ is true null} : p_i \leq qj/m\}}{j}.$$

Consequently, we get

$$\mathrm{FDP}^k := \frac{V \mathbf{1}_{R \geq k}}{R} \leq \max_{k \leq j \leq m} \frac{\#\{i \text{ is true null} : p_i \leq qj/m\}}{j}. \tag{3.7}$$

Similar to what has been argued in Section 3.1, the inequality (3.7) still holds if all true null p-values are replaced by m_0 i. i. d. uniform variables $U_1, \ldots, U_{m_0}$ on $(0, 1)$. This observation shows that it suffices to prove

$$\mathbb{E}\left[\max_{k \leq j \leq m} \frac{\#\{1 \leq i \leq m_0 : U_i \leq qj/m\}}{j}\right] \leq \left(1 + 2/\sqrt{qk}\right) q. \tag{3.8}$$

To show (3.8), denote by $V_j = \#\{1 \le i \le m_0 : U_i \le qj/m\}$ and $Y_j = V_j/j$. Conditional on Y_{j+1}, for every $i \in \{1 \le i \le m_0 : U_i \le q(j+1)/m\}$ the random variable U_i is uniformly distributed on $[0, q(j+1)/m]$. Hence, the conditional expectation of V_j given Y_{j+1} is

$$\mathbb{E}(V_j|Y_{j+1}) = \frac{V_{j+1}\frac{qj}{m}}{\frac{q(j+1)}{m}} = \frac{jV_{j+1}}{j+1},$$

which is equivalent to

$$\mathbb{E}(Y_j|Y_{j+1}) = Y_{j+1}\,.$$

In words, Y_j is a backward martingale and, as a consequence, $(Y_j - qm_0/m)_+$ is a backward submartingale. This fact allows us to apply Doob's ℓ^2 martingale maximal inequality to $(Y_j - qm_0/m)_+$, yielding

$$\begin{aligned}
\mathbb{E}\left[\max_{k\le j\le m}\left(Y_j - \frac{qm_0}{m}\right)_+^2\right] &\le \left(\frac{2}{2-1}\right)^2 \mathbb{E}\left(Y_k - \frac{qm_0}{m}\right)_+^2 \\
&\le 4\,\mathbb{E}\left(Y_k - \frac{qm_0}{m}\right)^2 \\
&= \frac{4qm_0(1-qk/m)}{km} \\
&< \frac{4\pi_0 q}{k}.
\end{aligned}$$

Using Jensen's inequality, the left-hand side of (3.8) satisfies

$$\begin{aligned}
\mathbb{E}\left[\max_{k\le j\le m} Y_j\right] &\le \frac{qm_0}{m} + \mathbb{E}\left[\max_{k\le j\le m}\left(Y_j - \frac{qm_0}{m}\right)_+\right] \\
&\le \pi_0 q + \sqrt{\mathbb{E}\left[\max_{k\le j\le m}\left(Y_j - \frac{qm_0}{m}\right)_+^2\right]} \\
&\le \pi_0 q + 2\sqrt{\frac{\pi_0 q}{k}}.
\end{aligned}$$

□

4. FDR Control and Power of PrivateBHq

As an application of Theorem 1, this section considers FDR control and power of PrivateBHq. Throughout this process, we take the assumptions of Theorem 3 as given. That is, we assume that each p_i is (η, ν)-sensitive and the parameters satisfy $\varepsilon \le 0.5, \delta \le 0.1$, and $m' \ge 10$. From Theorem 3, PrivateBHq preserves (ε, δ)-differential privacy, and for brevity this fact will not be reiterated in this section.

The proposition below demonstrates that the PrivateBHq is indeed compliant by making the cutoffs $\{\gamma_j\}$ in Algorithm 4 slightly more stringent than the logarithms of the BHq critical values.

Proposition 4.1 . *For any $0 < q < 1$, use the cutoffs*

$$\gamma_j = \log\frac{qj}{m} - \frac{\eta\sqrt{10m'\log(1/\delta)}\log(6m'/q)}{\varepsilon} \tag{4.1}$$

for $j = 1, \ldots, m'$ in PrivateBHq. Under the assumptions of Theorem 3, this procedure is compliant with the BHq critical values qj/m with probability at least $1 - 0.1q$.

As a remark, the first term $\log \frac{qj}{m}$ in (4.1) corresponds to the non-private cutoff and the second term $-\frac{\eta\sqrt{10m' \log(1/\delta)}\log(6m'/q)}{\varepsilon}$ is used to handle the added noise. Notably, the constant 0.1 above can be replaced by any positive constant provided that the second term is appropriately scaled. The proof of Proposition 4.1 is given later after Theorem 5.

The compliance condition shown in Proposition 4.1 together with Theorem 1 implies FDR control of PrivateBHq. More precisely, letting $\mathcal{C}$ denote the event that the rejected p-values are compliant, we have

$$\begin{aligned} \mathrm{FDR}_k &= \mathbb{E}\left(\mathrm{FDP}_k; \mathcal{C}\right) + \mathbb{E}\left(\mathrm{FDP}_k; \overline{\mathcal{C}}\right) \\ &\le C_k q + \mathbb{P}(\overline{\mathcal{C}}) \le (C_k + 0.1)q \end{aligned}$$

for every $k \ge 2$. As such, to control the FDR at level, say 10% (a common level used in practice), we can set $q = 0.1/(C_k + 0.1)$ in PrivateBHq. This proves the following theorem.

Theorem 5 . *Under the same assumptions as in Proposition 4.1 and if the test statistics satisfy the IWN condition, the PrivateBHq procedure gives*

$$\mathrm{FDR}_k \le (C_k + 0.1)q$$

for all $k \ge 2$.

To prove Proposition 4.1, we first present a simple lemma that gives a concentration bound on Laplace random variables, and its proof can be found in the appendix.

Lemma 4.2 . *Let $Z_1, \ldots, Z_n$ be i. i. d.* $\mathrm{Lap}(\lambda)$ *random variables. For any $0 < \alpha < 1$, the following two statements are true:*

(1) *With probability at least $1 - \alpha$, all Z_j are larger than $-\lambda \log \frac{n}{2\alpha}$.*
(2) *With probability at least $1 - \alpha$, all $|Z_j|$ are smaller than $\lambda \log \frac{n}{\alpha}$.*

Proof of Proposition 4.1. Let $\tilde{\theta}_{i_j} = \log \max\{\nu, p_{i_j}\} + Z_{i_j}$ be yielded by `peeling` in Algorithm 4, where Z_{i_j} follows $\mathrm{Lap}(\lambda)$ for $j = 1, \ldots, m'$. The parameter $\lambda = \eta\sqrt{10m' \log(1/\delta)}/\varepsilon$ is as in Theorem 3. Taking $\alpha = 0.1q$, Lemma 4.2 shows that

$$Z_{i_j} > -\lambda \log \frac{m'}{2 \times 0.1q} > -\frac{\eta\sqrt{10m' \log(1/\delta)}\log(6m'/q)}{\varepsilon}. \tag{4.2}$$

uniformly for $j = 1, \ldots, m'$ with probability at least $1 - 0.1q$.

Next, we show that on the event (4.2), PrivateBHq is compliant. Denote by R_{Pt} the number of rejections made by this procedure. If $\tilde{\theta}_{i_j}$ is rejected, it must satisfy

$$\log \max\{\nu, p_{i_j}\} + Z_{i_j} \le \gamma_{R_{\mathrm{Pt}}} = \log \frac{qR_{\mathrm{Pt}}}{m} - \frac{\eta\sqrt{10m' \log(1/\delta)}\log(6m'/q)}{\varepsilon}.$$

Plugging (4.2) into this display gives

$$\log \max\{\nu, p_{i_j}\} \le \log \frac{qR_{\mathrm{Pt}}}{m}.$$

Thus, $p_{i_j} \le qR_{\mathrm{Pt}}/m$ for all rejected p_{i_j} on the event (4.2), which happens with probability at least $1 - 0.1q$. This completes the proof.

□

Next, Theorem 6 shows that the PrivateBHq procedure with a slightly inflated nominal level is at least as powerful as the BHq step-down procedure. The proofs of this theorem and its corollary are deferred to the appendix. To state the theorem, let R_{SD} denote the number of rejections made by the (non-private) step-down procedure.

Theorem 6 . *Fix q and assume $\nu \leq q/m$. Under the assumptions of Theorem 5, run the PrivateBHq procedure at level*

$$q' = q\mathrm{e}^{\frac{24\eta\sqrt{m'\log(1/\delta)}\log m}{\varepsilon}}$$

and the BHq step-down procedure at level q. Then, the numbers of rejections satisfy

$$R_{\mathrm{Pt}} \geq \min\{R_{\mathrm{SD}}, m'\} \tag{4.3}$$

with probability tending to one as $m \to \infty$.

When $R_{\mathrm{SD}} \geq m'$ and the event (4.3) happens, PrivateBHq must reject all p-values passing through `peeling`. In the case where non-null p-values are significant enough to pass through `peeling`, this fact suggests that PrivateBHq achieves high power. This high-power property, however, is appealing if q' is only slightly larger than q or, put more simply, the number $24\eta\sqrt{m'\log(1/\delta)}\log m/\varepsilon$ is small. With regard to Examples 2.6 and 2.7, this is equivalent to having a sufficiently large sample size n. The following corollary formalizes this point.

Corollary 4.3 . *In Examples 2.6 and 2.7, fix ε, δ and assume $m' \leq \min\{n^{1-c}, m\}$ for constant $c > 0$. Under the assumptions of Theorem 6, the claims of both Theorems 5 and 6 hold as $m, n \to \infty$ if PrivateBHq is performed at level $(1 + c')q$ for a sufficiently small constant $c' > 0$.*

4.1. **Empirical evaluation.** In this subsection, we evaluate the price paid for privacy in terms of FDR control and power in the PrivateBHq procedure. The aim is to provide a better picture of how much detection power would be compromised due to privacy guarantees for FDR control. For completeness, this comparison includes a private version of Bonferroni's method, which is referred to as PrivateBonf in this paper. PrivateBonf is perhaps the simplest baseline for private multiple hypothesis testing. This procedure is detailed as follows. As in Algorithm 4, let $p_1, \ldots, p_m$ be (η, ν)-sensitive p-values and set $\theta_j = \log\max\{p_j, \nu\}$ for all j. The PrivateBonf procedure adds independent $\mathrm{Lap}(\tilde{\lambda})$ noise to all θ_j, where $\tilde{\lambda} = \eta\sqrt{10m\log(1/\delta)}/(2\varepsilon)$, and rejects those with noisy counts below

$$\log\frac{q}{m} - \frac{\eta\sqrt{10m\log(1/\delta)}\log(5m/q)}{2\varepsilon}.$$

The following result is concerned with privacy and family-wise error rate (FWER) control of PrivateBonf. Note that the FWER denotes the probability that at least one false positive is made. The proof is deferred to the appendix. As an aside, the privacy guarantee in this result might be improved by using the sparse vector technique [36] and this is left for future investigation.

Proposition 4.4 . *Under the assumptions of Theorem 3, the following two statements are true:*

(1) *PrivateBonf is (ε, δ)-differentially private;*

(2) *PrivateBonf satisfies* FWER $\leq 1.1q$.

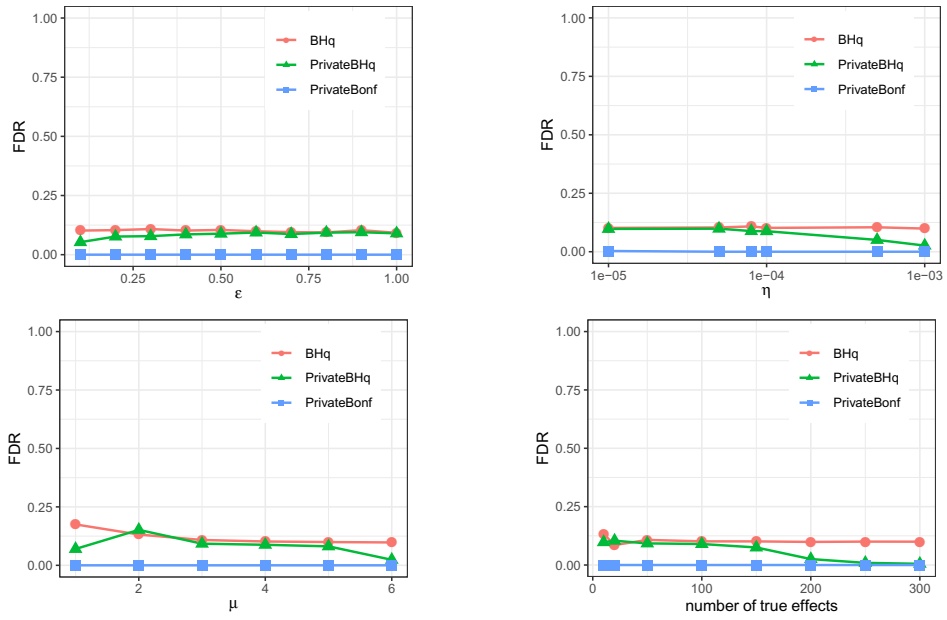


FIGURE 2. The FDR of BHq, PrivateBHq, and PrivateBonf, plotted against varying $\varepsilon, \eta, \mu, m_1$, respectively, and averaged over 100 independent replicates. Note that PrivateBHq in general discovers fewer than BHq and has a smaller FDR than BHq as well.

Figures 2 and 3 present, respectively, the FDR and the power of PrivateBHq, the (non-private) BHq step-up procedure, and PrivateBonf by simulations. Unless specified, we set $m = 10^5, m' = 100, q = 0.1, \eta = 10^{-4}, \nu = 0.5q/m, \varepsilon = 0.5$, and $\delta = 0.001$. To construct the p-values, we let $p_i = \Phi(\xi_i - \mu)$ for $i = 1, \ldots, m_1$ and $p_{m_1+1}, \ldots, p_m$ be i.i.d. uniform variables on $(0, 1)$, where Φ is the CDF of $\mathcal{N}(0, 1)$, $\xi_1, \ldots, \xi_{m_1}$ are i.i.d. copies of $\mathcal{N}(0, 1)$, and the default values of μ and m_1 (the number of true effects) are set to 4 and 100, respectively. In summary, the FDR of PrivateBHq is empirically controlled at q in almost all scenarios, though Theorem 5 is only concerned with FDR_k for $k \geq 2$. Moreover, PrivateBonf is uniformly the least powerful among the three procedures. This is not surprising given that PrivateBonf is inherently developed for FWER control. Looking closely, the performance of PrivateBHq is comparable to that of BHq when ε is not too small and η is not too large. Notably, the power of PrivateBHq deteriorates when the number of true effects exceeds 150, which is due to the truncation of the PrivateBHq procedure at $m' = 100$.

For completeness, we refer interested readers to a set of numerical comparisons of an FDR_k-controlling procedure [52, 54], PrivateBHq, and PrivateBonf in the appendix.

5. DISCUSSION

This paper has developed a privacy-preserving multiple testing procedure termed PrivateBHq for FDR control. On the privacy side, we propose a new notion of sensitivity tailored to p-values and recognize the sequential nature of the BHq (step-down) procedure so as to keep

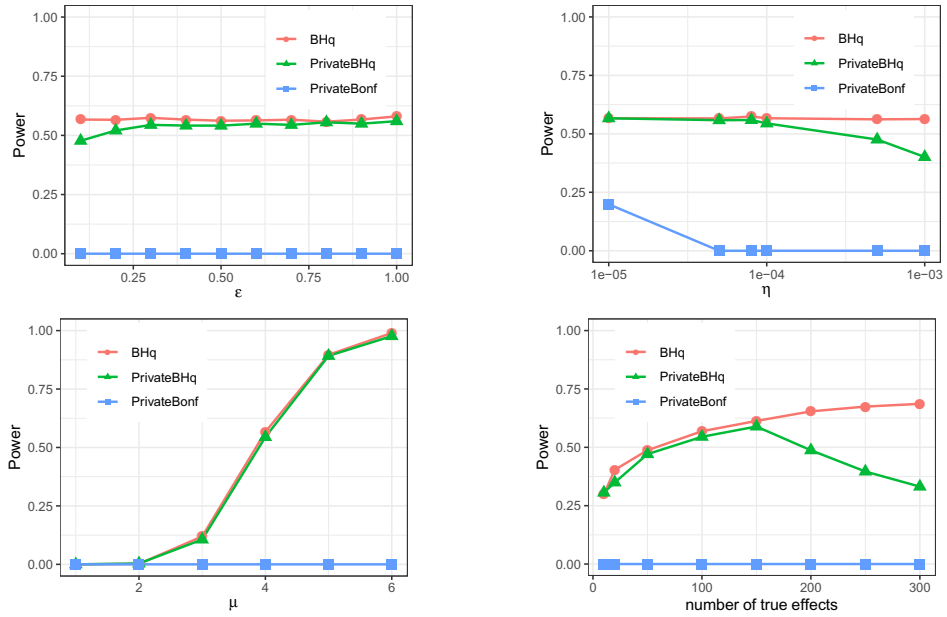


FIGURE 3. The power of BHq, PrivateBHq, and PrivateBonf, plotted against varying $\varepsilon, \eta, \mu, m_1$, respectively, and averaged over 100 independent replicates.

PrivateBHq efficient under the differential privacy constraint. Differential privacy of this whole pipeline follows from the composition nature of differential privacy. On the statistical side, as a major contribution of the paper, it is proved that a large class of multiple testing procedures, including the step-up, step-down, and PrivateBHq procedures, control the FDR_k only provided the joint independence of the true null test statistics. A novel aspect of this result lies in the absence of any assumption on the dependence between the true nulls and false nulls. Notably, some recent progress has been made along this direction using the the FDR-linking technique [59].

Looking forward, our work raises a number of open questions. First, it would be interesting to take into account prior knowledge, such as the importance of hypotheses and beliefs about which are true nulls, into the design of a differentially private procedure. Second, it would be of interest to develop private procedures for control of other popular error rates such as the q-value [58]. Moreover, it is natural to wonder if the bound in Theorem 1 can improve by imposing some structure on the dependence between the true null and false null test statistics. Third, recognizing the vital importance of `peeling` in our PrivateBHq, an interesting direction is to investigate alternatives to `peeling`, such as the oneshot approach to the problem of private top-k selection [49]. Last, it would be interesting to consider other notions of privacy such as concentrated differential privacy and Gaussian differential privacy [28, 13, 18].

Finally, we wish to make a connection to a remarkable property of differential privacy: it protects against false discoveries due to adaptive data analysis, where an analysis is informed by prior interactions with the same database [22, 21, 3]. Adaptivity is ubiquitous in practice as the analyst is often not clear a priori what are the right questions to ask about a database. In the multiple testing context, this issue arises when hypotheses are adaptively selected

based on prior discoveries. A question of great interest is to develop a multiple testing procedure that continues to preserve privacy in the presence of adaptivity.

Disclosure / Competing Interests

C. Dwork is Editor-in-Chief of this Journal. As per the Journal's policies, she did not participate in the editorial process, and all aspects of the editorial workflow were blinded (not accessible) to her.

References

[1] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang. Deep learning with differential privacy. In *ACM SIGSAC Conference on Computer and Communications Security*, pages 308–318, 2016, pages 308–318. doi:10.1145/2976749.2978318.

[2] B. Barak, K. Chaudhuri, C. Dwork, S. Kale, F. McSherry, and K. Talwar. Privacy, accuracy, and consistency too: a holistic solution to contingency table release. In *Proceedings of the twenty-sixth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pages 273–282. ACM, 2007, pages 273–282. doi:10.1145/1265530.1265569.

[3] R. Bassily, K. Nissim, A. Smith, T. Steinke, U. Stemmer, and J. Ullman. Algorithmic stability for adaptive data analysis. In *Proceedings of the 48th annual ACM Symposium on Theory of Computing*, pages 1046–1059, 2016, pages 1046–1059. doi:10.1137/16M1103646.

[4] A. Beimel, S. P. Kasiviswanathan, and K. Nissim. Bounds on the sample complexity for private learning and private data release. In *Theory of Cryptography Conference*, pages 437–454. Springer, 2010, pages 437–454. doi:10.1007/978-3-642-11799-2_26.

[5] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate – A practical and powerful approach to multiple testing. *Journal of the Royal Statistics Society: Series B (Statistical Methodology)*, 57(1):289–300, 1995. doi:10.1111/j.2517-6161.1995.tb02031.x.

[6] Y. Benjamini, A. M. Krieger, and D. Yekutieli. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507, 2006. doi:10.1093/biomet/93.3.491.

[7] Y. Benjamini and D. Yekutieli. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4):1165–1188, 2001. doi:10.1214/aos/1013699998.

[8] G. Blanchard and E. Roquain. Two simple sufficient conditions for FDR control. *Electronic Journal of Statistics*, 2:963–992, 2008. doi:10.1214/08-EJS180.

[9] G. Blanchard and E. Roquain. Adaptive false discovery rate control under independence and dependence. *Journal of Machine Learning Research*, 10(12):2837–2871, 2009. URL: https://www.jmlr.org/papers/volume10/blanchard09a/blanchard09a.pdf.

[10] A. Blum, K. Ligett, and A. Roth. A learning theory approach to noninteractive database privacy. *Journal of the ACM (JACM)*, 60(2):12, 2013. doi:10.1145/2450142.2450148.

[11] M. Bogdan, E. van den Berg, C. Sabatti, W. J. Su, and E. J. Candès. SLOPE – adaptive variable selection via convex optimization. *The Annals of Applied Statistics*, 9(3):1103, 2015. doi:10.1214/15-AOAS842.

[12] Z. Bu, J. Dong, Q. Long, and W. J. Su. Deep learning with Gaussian differential privacy. *Harvard Data Science Review*, 2020(23), 2020. doi:10.1162/99608f92.cfc5dd25.

[13] M. Bun and T. Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography Conference*, pages 635–658. Springer, 2016, pages 635–658. doi:10.1007/978-3-662-53641-4_24.

[14] M. Bun, J. Ullman, and S. P. Vadhan. Fingerprinting codes and the price of approximate differential privacy. *SIAM J. Comput.*, 47(5):1888–1938, 2018. doi:10.1137/15M1033587.

[15] K. Chaudhuri, C. Monteleoni, and A. D. Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12(Mar):1069–1109, 2011. URL: https://jmlr.org/papers/volume12/chaudhuri11a/chaudhuri11a.pdf.

[16] S. Clarke and P. Hall. Robustness of multiple testing procedures against dependence. *The Annals of Statistics*, 37(1):332–358, 2009. doi:10.1214/07-AOS557.

[17] L. De Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Science & Business Media, 2007. doi:10.1007/0-387-34471-3.

[18] J. Dong, A. Roth, and W. J. Su. Gaussian differential privacy. *arXiv preprint arXiv:1905.02383*, 2019. URL: https://arxiv.org/abs/1905.02383.

[19] S. Dudoit, M. J. Van Der Laan, and M. J. van der Laan. *Multiple testing procedures with applications to genomics*. Springer, 2008. doi:10.1007/978-0-387-49317-6.

[20] R. Durrett. *Probability: Theory and Examples*. Cambridge University Press, 2010. doi:10.1017/9781108591034.

[21] C. Dwork, V. Feldman, M. Hardt, T. Pitassi, O. Reingold, and A. Roth. Generalization in adaptive data analysis and holdout reuse. In *Advances in Neural Information Processing Systems*, pages 2350–2358, 2015, pages 2350–2358. doi:10.5555/2969442.2969502.

[22] C. Dwork, V. Feldman, M. Hardt, T. Pitassi, O. Reingold, and A. Roth. The reusable holdout: Preserving validity in adaptive data analysis. *Science*, 349(6248):636–638, 2015. doi:10.1126/science.aaa9375.

[23] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. In *Proceedings of EUROCRYPT*, pages 486–503, 2006, pages 486–503. doi:10.1007/11761679_29.

[24] C. Dwork and J. Lei. Differential privacy and robust statistics. In *Proceedings of the 41st annual ACM Symposium on Theory of Computing*, pages 371–380. ACM, 2009, pages 371–380. doi:10.1145/1536414.1536466.

[25] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography*, pages 265–284. Springer, 2006. pages 265–284. doi:10.1007/11681878_14.

[26] C. Dwork and A. Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3-4):211–407, 2014. doi:10.1561/0400000042.

[27] C. Dwork, G. Rothblum, and S. Vadhan. Boosting and differential privacy. In *Foundations of Computer Science (FOCS)*, 2010. doi:10.1109/FOCS.2010.12.

[28] C. Dwork and G. N. Rothblum. Concentrated differential privacy. *arXiv preprint arXiv:1603.01887*, 2016. URL: https://arxiv.org/abs/1603.01887.

[29] C. Dwork, A. Smith, T. Steinke, J. Ullman, and S. Vadhan. Robust traceability from trace amounts. In *IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 650–669. IEEE, 2015, pages 650–669. doi:10.1109/FOCS.2015.46.

[30] H. Finner, T. Dickhaus, and M. Roters. On the false discovery rate and an asymptotically optimal rejection curve. *The Annals of Statistics*, 37(2):596–618, 2009. doi:10.1214/07-AOS569.

[31] M. Gaboardi, H. Lim, R. Rogers, and S. Vadhan. Differentially private Chi-squared hypothesis testing: Goodness of fit and independence testing. In *International Conference on Machine Learning*, pages 2111–2120, 2016, pages 2111–2120. doi:10.5555/3045390.3045613.

[32] Y. Gavrilov, Y. Benjamini, and S. K. Sarkar. An adaptive step-down procedure with proven FDR control under independence. *The Annals of Statistics*, 37(2):619–629, 2009. doi:10.1214/07-AOS586.

[33] Y. Ge, S. C. Sealfon, and T. P. Speed. Some step-down procedures controlling the false discovery rate under dependence. *Statistica Sinica*, 18(3):881–904, 2008. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2583793/.

[34] R. Hall, A. Rinaldo, and L. Wasserman. Differential privacy for functions and functional data. *Journal of Machine Learning Research*, 14(2):703–727, 2013. URL: https://www.jmlr.org/papers/volume14/hall13a/hall13a.pdf.

[35] M. Hardt, K. Ligett, and F. McSherry. A simple and practical algorithm for differentially private data release. In *Advances in Neural Information Processing Systems*, pages 2339–2347, 2012, pages 2339–2347. doi:10.5555/2999325.2999396.

[36] M. Hardt and G. N. Rothblum. A multiplicative weights mechanism for privacy-preserving data analysis. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 61–70. IEEE, 2010, pages 61–70. doi:10.1109/FOCS.2010.85.

[37] P. Heesen and A. Janssen. Inequalities for the false discovery rate (FDR) under dependence. *Electronic Journal of Statistics*, 9(1):679–716, 2015. doi:10.1214/15-EJS1016.

[38] N. Homer, S. Szelinger, M. Redman, D. Duggan, W. Tembe, J. Muehling, J. V. Pearson, D. A. Stephan, S. F. Nelson, and D. W. Craig. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS genetics*, 4(8):e1000167, 2008. doi:10.1371/journal.pgen.1000167.

[39] V. Karwa and A. Slavković. Inference using noisy degrees: Differentially private β-model and synthetic graphs. *The Annals of Statistics*, 44(1):87–112, 2016. doi:10.1214/15-AOS1358.

[40] V. Karwa and S. Vadhan. Finite sample differentially private confidence intervals. In *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018. doi:10.4230/LIPIcs.ITCS.2018.44.

[41] J. Kaye. The tension between data sharing and the protection of privacy in genomics research. *Annual Review of Genomics and Human Genetics*, 13(1):415–431, 2012. doi:10.1146/annurev-genom-082410-101454.

[42] E. L. Lehmann. Some concepts of dependence. *The Annals of Mathematical Statistics*, 37(5):1137–1153, 1966. doi:10.1214/aoms/1177699260.

[43] J. Lei, A.-S. Charest, A. Slavković, A. Smith, and S. Fienberg. Differentially private model selection with penalized and constrained likelihood. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2016. doi:10.1111/rssa.12324.

[44] J. Littlewood. On the probability in the tail of a binomial distribution. *Advances in Applied Probability*, 1(1):43–72, 1969. doi:10.2307/1426408.

[45] T. McDaneld, L. Kuehn, M. Thomas, W. Snelling, T. Smith, E. Pollak, J. Cole, and J. Keele. Genomewide association study of reproductive efficiency in female cattle. *Journal of Animal Science*, 92(5):1945–1957, 2014. doi:10.2527/jas.2012-6807.

[46] B. D. McKay. On littlewood's estimate for the binomial distribution. *Advances in Applied Probability*, 21(2):475–478, 1989. doi:10.2307/1427172.

[47] F. McSherry and K. Talwar. Mechanism design via differential privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science*, pages 94–103, 2007, pages 94–103. doi:10.1109/FOCS.2007.66.

[48] J. Neveu. *Discrete-parameter martingales*. North Holland, Amsterdam, 1975. doi:10.2307/2344365.

[49] G. Qiao, W. J. Su, and L. Zhang. Oneshot differentially private top-k selection. In *International Conference on Machine Learning*, 2021. URL: https://arxiv.org/abs/2105.08233.

[50] E. Roquain and F. Villers. Exact calculations for false discovery proportion with application to least favorable configurations. *The Annals of Statistics*, 39(1):584–612, 2011. doi:10.1214/10-AOS847.

[51] S. K. Sarkar. Some results on false discovery rate in stepwise multiple testing procedures. *The Annals of Statistics*, 30(1):239–257, 2002. doi:10.1214/AOS/1015362192.

[52] S. K. Sarkar. Stepup procedures controlling generalized FWER and generalized FDR. *The Annals of Statistics*, 35(6):2405–2420, 2007. doi:10.1214/009053607000000398.

[53] S. K. Sarkar. On methods controlling the false discovery rate. *Sankhyā: The Indian Journal of Statistics, Series A*, 70:135–168, 2008. doi:10.11/495.181.

[54] S. K. Sarkar and W. Guo. On a generalized false discovery rate. *The Annals of Statistics*, 37(3):1545–1565, 2009. doi:10.1214/08-AOS617.

[55] T. K. Sarkar. Some lower bounds of reliability. Technical report, DTIC Document, 1969. URL: https://statistics.stanford.edu/research/some-lower-bounds-reliability.

[56] J. D. Storey. The positive false discovery rate: a Bayesian interpretation and the q-value. *The Annals of Statistics*, 31(6):2013–2035, 2003. doi:10.1214/AOS/1074290335.

[57] J. D. Storey, J. E. Taylor, and D. Siegmund. 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, 2004. doi:10.1111/j.1467-9868.2004.00439.x.

[58] J. D. Storey and R. Tibshirani. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences*, 100(16):9440–9445, 2003. doi:10.1073/pnas.1530509100.

[59] W. J. Su. The FDR-linking theorem. *arXiv preprint arXiv:1812.08965*, 2018. URL: https://arxiv.org/abs/1812.08965.

[60] A. C. Tamhane, W. Liu, and C. W. Dunnett. A generalized step-up-down multiple test procedure. *The Canadian Journal of Statistics*, 26(2):353–363, 1998. doi:10.2307/3315516.

[61] C. Uhler, A. Slavković, and S. E. Fienberg. Privacy-preserving data sharing for genome-wide association studies. *The Journal of Privacy and Confidentiality*, 5(1):137, 2013. doi:10.29012/JPC.V5I1.629.

[62] Y.-X. Wang. Revisiting differentially private linear regression: optimal and adaptive prediction & estimation in unbounded domain. *arXiv preprint arXiv:1803.02596*, 2018. URL: https://arxiv.org/abs/1803.02596.

[63] L. Wasserman and S. Zhou. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489):375–389, 2010. doi:10.1198/jasa.2009.tm08651.

[64] Wikipedia. Facebook–Cambridge Analytica data scandal. https://en.wikipedia.org/wiki/Facebook2018.

[65] F. Yu, S. E. Fienberg, A. Slavković, and C. Uhler. Scalable privacy-preserving data sharing methodology for genome-wide association studies. *Journal of Biomedical Informatics*, 50:133–141, 2014. doi:10.1016/j.jbi.2014.01.008.

Appendix A. Proofs

This section proves all results made without proof in the main text. Below, the proofs are listed in order of appearance of their corresponding results.

Proof of Lemma 2.4. For an arbitrary index $1 \le j \le m$ and a measurable set $S \subset \mathbb{R}$, it suffices to prove that

$$\frac{\mathbb{P}(\tilde{f}_j \text{ is the smallest and } f_j(D) + Z \in S)}{\mathbb{P}(\tilde{f}'_j \text{ is the smallest and } f_j(D') + Z \in S)} \le \mathrm{e}^{\varepsilon},$$

where $\tilde{f}'_j$ is the counterpart of $\tilde{f}_j$ evaluated on an adjacent database D'. This inequality is equivalent to

$$\frac{\mathbb{P}(\tilde{f}_j \text{ is smallest})\,\mathbb{P}(f_j(D) + Z \in S | \tilde{f}_j \text{ is smallest})}{\mathbb{P}(\tilde{f}'_j \text{ is smallest})\,\mathbb{P}(f_j(D') + Z \in S | \tilde{f}'_j \text{ is smallest})} \le \mathrm{e}^{\varepsilon}. \tag{A.1}$$

First, releasing the index of the smallest noisy count is $(\varepsilon/2, 0)$-differentially private, which has been proven by Claim 3.9 in Section 3.3 of [26]. That is,

$$\frac{\mathbb{P}(\tilde{f}_j \text{ is smallest})}{\mathbb{P}(\tilde{f}'_j \text{ is smallest})} \le \mathrm{e}^{\varepsilon/2}.$$

Second, observe that by assumption $|f_j(D) - f_j(D')| \le \Delta$. Then Lemma 2.3 shows that

$$\frac{\mathbb{P}(f_j(D) + Z \in S | \tilde{f}_j \text{ is smallest})}{\mathbb{P}(f_j(D') + Z \in S | \tilde{f}'_j \text{ is smallest})} \le \mathrm{e}^{\varepsilon/2}.$$

Combining the last two displays concludes that (A.1) is bounded by e^{ε}. This finishes the proof.

□

Proof of Example 2.6. We use $t-1$ in place of t. Under the constraint that $p(D), p(D') \ge \nu$, we aim to prove that

$$\frac{\binom{n}{t-1}}{\sum_{i=t}^{n}\binom{n}{i}} \le \eta$$

if $\eta \asymp \sqrt{(\log n)/n}$. Without loss of generality, assume $t \ge n/2$, where a well-known result is

$$\sum_{i=t}^{n} \frac{1}{2^n}\binom{n}{i} \le \mathrm{e}^{-n\,\mathrm{KL}(\frac{t}{n}, \frac{1}{2})}. \tag{A.2}$$

Above, the Kullback–Leibler divergence is defined as

$$\mathrm{KL}(a, b) = a \log \frac{a}{b} + (1-a) \log \frac{1-a}{1-b}.$$

It is easy to show that

$$\mathrm{KL}\left(a, \frac{1}{2}\right) \ge 2\left(a - \frac{1}{2}\right)^2.$$

Therefore, plugging

$$\sum_{i=t}^{n} \frac{1}{2^n}\binom{n}{i} \ge \nu = m^{-1-c} = \frac{1}{\mathtt{poly}(n)}$$

into (A.2), we get

$$\frac{t}{n} \leq \frac{1}{2} + O\left(\sqrt{\frac{\log n}{n}}\right)$$

or, put differently,

$$t \leq \frac{n}{2} + O\left(\sqrt{n \log n}\right).$$

Therefore, we can assume $t \leq 7n/8$. Provided that $n/2 \leq t \leq 7n/8$, we can apply Littlewood's theorem [44, 46]. Letting $u = (2t-n)/\sqrt{n}, \rho = 1 - t/n$ and $\Xi(x) = \Phi(-x)/\phi(x)$, where $\Phi(x)$ and $\phi(x)$ are the cumulative distribution function and density function of $\mathcal{N}(0,1)$ respectively, this theorem gives

$$\sum_{i=t}^{n} \frac{1}{2^n} \binom{n}{i} = (1 + O(1/n))\, \Phi(-u) \mathrm{e}^{A_1 + A_2/\sqrt{\rho(1-\rho)n}},$$

where

$$A_1 = \frac{u^2}{2} - \left(t - \frac{1}{2}\right) \log \frac{2t}{n} - \left(n - t + \frac{1}{2}\right) \log \frac{2(n-t)}{n}$$

and

$$A_2 = \frac{1-2\rho}{6}\left[\frac{1-u^2}{\Xi(u)} + u^3\right] + \frac{1/\Xi(u) - u}{2}.$$

Because $t \leq n/2 + O(\sqrt{n \log n})$ as $n \to \infty$, we have $u = O(\sqrt{\log n}), \rho = \frac{1}{2} - o(1)$. Making use of the fact that $\Xi(u) = (1 + o(1))/u$, we see that

$$\begin{aligned}
\sum_{i=t}^{n} \frac{1}{2^n} \binom{n}{i} &= (1 + O(1/n))\, \Phi(-u) \mathrm{e}^{A_1} (1 + o(1)) \\
&= (1 + o(1)) \frac{\Phi(-u)}{\sqrt{2\theta}\phi(u)} \mathrm{e}^{-(t-\frac{1}{2}) \log \frac{2t}{n} - (n-t+\frac{1}{2}) \log \frac{2(n-t)}{n}} \\
&= (1 + o(1)) \frac{\Phi(-u)}{\sqrt{2\theta}\phi(u)} \mathrm{e}^{-t \log \frac{2t}{n} - (n-t) \log \frac{2(n-t)}{n}} \\
&= (1 + o(1)) \frac{\Phi(-u)}{\sqrt{2\theta}\phi(u)} \mathrm{e}^{-n\, \mathrm{KL}(\frac{t}{n}, \frac{1}{2})} \\
&= (1 + o(1)) \cdot (1 + o(1)) \frac{1}{\sqrt{2\theta} u} \mathrm{e}^{-n\, \mathrm{KL}(\frac{t}{n}, \frac{1}{2})} \\
&\geq \frac{O(1)}{\sqrt{2\theta \log n}} \mathrm{e}^{-n\, \mathrm{KL}(\frac{t}{n}, \frac{1}{2})}.
\end{aligned} \tag{A.3}$$

Next, we consider

$$\frac{1}{2^n} \binom{n}{t-1}$$

By Stirling's formula and using the fact that $t = (0.5 + o(1))n$, we get

$$\begin{aligned}\frac{1}{2^n}\binom{n}{t-1} &= \frac{t}{n+1-t}\frac{1}{2^n}\binom{n}{t} \\ &= (1+o(1))\frac{1}{2^n}\binom{n}{t} \\ &= (1+o(1))\sqrt{\frac{2}{\theta n}}\cdot\frac{n^n}{2^n t^t (n-t)^{n-t}} \\ &= (1+o(1))\sqrt{\frac{2}{\theta n}}\cdot \mathrm{e}^{-n\,\mathrm{KL}(\frac{t}{n},\frac{1}{2})}\end{aligned}$$

Thus, we get

$$\begin{aligned}\frac{\frac{1}{2^n}\binom{n}{t-1}}{\sum_{i=t}^n \frac{1}{2^n}\binom{n}{i}} &\le O(1)\frac{(1+o(1))\sqrt{\frac{2}{\theta n}}\cdot \mathrm{e}^{-n\,\mathrm{KL}(\frac{t}{n},\frac{1}{2})}}{\frac{1}{\sqrt{2\theta\log n}}\mathrm{e}^{-n\,\mathrm{KL}(\frac{t}{n},\frac{1}{2})}} \\ &= O\left(\sqrt{\frac{\log n}{n}}\right).\end{aligned}$$

Thus, with $\nu = m^{-1-c}$, we can choose $\eta \asymp \sqrt{\frac{\log n}{n}}$.

□

Proof of Example 2.7. Let $\zeta, \zeta_1, \ldots, \zeta_n$ be i. i. d. exponential variable with $\lambda = 1$ truncated at A. Consider the cumulant generating function

$$\kappa(\theta) = \log \mathbb{E}\, \mathrm{e}^{\theta\zeta}.$$

As in the proof of Example 2.6, it does not lose generality by assuming $t > n\,\mathbb{E}\,\zeta$. Write $a = t/n > \mathbb{E}\,\zeta$ and let θ_a be the root of the saddle-point equation

$$\kappa'(\theta_a) = a. \tag{A.4}$$

In particular, $\mathbb{E}_{\theta_a}\,\zeta = a$. Note that under $\mathbb{E}_\theta$ the density of ζ is

$$\frac{\lambda \mathrm{e}^{-\lambda x}}{1-\mathrm{e}^{-A\lambda}}\mathrm{e}^{\theta x - \kappa(\theta)}\cdot \mathbf{1}(0 \le x \le A).$$

Through exponential tilting, we get

$$\begin{aligned}\mathbb{P}(T \ge na) &= \mathbb{P}\left(\zeta_1 + \cdots + \zeta_n \ge na\right) \\ &= \mathrm{e}^{-n(a\theta_a - \kappa(\theta_a))}\,\mathbb{E}_{\theta_a}\,\mathrm{e}^{-\theta_a(T-na)}\mathbf{1}(T \ge na).\end{aligned}$$

Using saddle point approximation, we get

$$\mathbb{E}_{\theta_a}\,\mathrm{e}^{-\theta_a(T-na)}\mathbf{1}(T \ge na) = \frac{1+o(1)}{\sqrt{2\theta\kappa''(\theta_a)n}\theta_a}$$

Thus, we have

$$\mathbb{P}(T \ge na) = (1+o(1))\frac{\mathrm{e}^{-n(a\theta_a - \kappa(\theta_a))}}{\sqrt{2\theta\kappa''(\theta_a)n}\theta_a}. \tag{A.5}$$

Next, we evaluate $\kappa''(\theta_a)$ and θ_a. Denote by μ and σ^2 the mean and variance of ζ, respectively. We get $\theta_a = o(1)$ and $\kappa''(\theta_a) = \mathrm{Var}_{\theta_a}(\zeta) = \sigma^2 + o(1)$. In particular, from (A.4) we get

$$\theta_a = (1 + o(1))\frac{a - \mu}{\sigma^2},$$

which gives

$$a\theta_a - \kappa(\theta_a) = (1 + o(1))\frac{(a - \mu)^2}{2\sigma^2}.$$

Plugging this display into

$$\mathrm{e}^{-n(a\theta_a - \kappa(\theta_a))} \geq \nu = m^{-1-c} = \frac{1}{\mathtt{poly}(n)}.$$

gives

$$t = n\mu + O(\sqrt{n \log n}).$$

Therefore, we get

$$\mathbb{E}_{\theta_a} \mathrm{e}^{-\theta_a(T - na)}\mathbf{1}(T \geq na) = \frac{1 + o(1)}{\sqrt{2\theta\kappa''(\theta_a)n}\theta_a} = \frac{O(1)}{\sqrt{\log n}},$$

which together with (A.5) yields

$$\mathbb{P}\left(T \geq na\right) = \frac{\mathrm{e}^{-n(a\theta_a - \kappa(\theta_a))}}{\sqrt{\log n}}. \tag{A.6}$$

To evaluate the ratio

$$\frac{\mathbb{P}\left(na - A \leq T < na\right)}{\mathbb{P}\left(T \geq na\right)},$$

it remains to approximate $\mathbb{P}\left(na - A \leq T < na\right)$. We use the local central limit theorem to do this. Explicitly, using the local central limit theorem, we get

$$\begin{aligned}
\mathbb{P}\left(t - A \leq T < t\right) &= \mathrm{e}^{-n(a\theta_a - \kappa(\theta_a))}\, \mathbb{E}_{\theta_a} \mathrm{e}^{-\theta_a(T - na)}\mathbf{1}(na - A \leq T < na) \\
&\leq \mathrm{e}^{-n(a\theta_a - \kappa(\theta_a))}\, \mathbb{E}_{\theta_a} \mathrm{e}^{\theta_a A}\mathbf{1}(na - A \leq T < na) \\
&= \mathrm{e}^{\theta_a A}\mathrm{e}^{-n(a\theta_a - \kappa(\theta_a))}\, \mathbb{P}_{\theta_a}(na - A \leq T < na) \\
&= \mathrm{e}^{\theta_a A}\mathrm{e}^{-n(a\theta_a - \kappa(\theta_a))}\left(\frac{A}{\sqrt{2\theta n}\sigma_a} + o(1/\sqrt{n})\right) \\
&= (1 + o(1))\frac{A\mathrm{e}^{\theta_a A}\mathrm{e}^{-n(a\theta_a - \kappa(\theta_a))}}{\sqrt{2\theta n}\sigma_a},
\end{aligned} \tag{A.7}$$

where σ_a is the standard deviation of ζ tilted at θ_a. That is, $\sigma_a = \sqrt{\mathrm{Var}_{\theta_a}\,\zeta} = \sigma + o(1)$.

Finally, combing (A.6) and (A.7) gives

$$\frac{\mathbb{P}\left(na - A \leq T < na\right)}{\mathbb{P}\left(T \geq na\right)} \leq O(1)\sqrt{\frac{\log n}{n}}.$$

As such, we can choose

$$\eta = O(1)\sqrt{\frac{\log n}{n}}.$$

□

An example of p-value computation from [61, 65]. Now we show that the contingency table example in Section 2.2 does not give p-values that are (η, ν)-sensitive with some $\eta, \nu \to 0$ even if $n \to \infty$. In particular, we focus on two adjacent tables as shown in Table 1. The χ^2-statistic of the left table is

$$\chi^2_L = 0$$

because $a \times (n/2 - a) - a \times (n/2 - a) = 0$ and, as a consequence, the corresponding p-value is

$$p_L \approx \mathbb{P}(\chi^2_1 \geq \chi^2_L) = 1.$$

Next, for the right table the statistic equals

$$\begin{aligned}
\chi^2_R &= \frac{[(a+1)(n/2-a) - a(n/2-a-1)]^2}{n} \\
&\quad \times \left[\frac{1}{\frac{n}{2}(2a+1)} + \frac{1}{\frac{n}{2}(2a+1)} + \frac{1}{\frac{n}{2}(n-2a-1)} + \frac{1}{\frac{n}{2}(n-2a-1)}\right] \\
&= \frac{n}{4} \times \left[\frac{1}{\frac{n}{2}(2a+1)} + \frac{1}{\frac{n}{2}(2a+1)} + \frac{1}{\frac{n}{2}(n-2a-1)} + \frac{1}{\frac{n}{2}(n-2a-1)}\right] \\
&= \frac{1}{2}\left[\frac{1}{2a+1} + \frac{1}{2a+1} + \frac{1}{n-2a-1} + \frac{1}{n-2a-1}\right] \\
&= \frac{1}{2a+1} + \frac{1}{n-2a-1}.
\end{aligned}$$

Now, assuming $5 \leq a \ll n$, we get

$$\chi^2_R = \frac{1}{2a+1} + o(1),$$

leading to

$$p_R \approx \mathbb{P}(\chi^2_1 \geq \chi^2_R) \approx 2\Phi\left(-\frac{1}{\sqrt{2a+1}}\right).$$

Thus, in this example both p_R and p_L are bounded below away from 0 and the ratio $2\Phi\left(-\frac{1}{\sqrt{2a+1}}\right)$ does not tend to 1 as $n \to \infty$. As a consequence of this, it is impossible to have both vanishing η and ν for this p-value computation.

□

Proof of Theorem 3. The PrivateBHq procedure acts on the intermediate results

$$(i_1, \tilde{\theta}_{i_1}), \ldots, (i_k, \tilde{\theta}_{i_k})$$

provided by the `peeling`. Hence, Lemma 2.8 implies that it suffices to establish the (ε, δ)-differential privacy for `peeling` as a part of PrivateBHq. By Lemma 2.4, each Private Min in `peeling` is $(\frac{2\varepsilon}{\sqrt{10k\log(1/\delta)}}, 0)$-differentially private. Then Lemma 2.9 immediately asserts that `peeling` preserves $(\tilde{\varepsilon}, \delta)$-differential privacy, where

$$\begin{aligned}
\tilde{\varepsilon} &= \frac{2\varepsilon}{\sqrt{10k\log(1/\delta)}}\sqrt{2k\log(1/\delta)} + k\frac{2\varepsilon}{\sqrt{10k\log(1/\delta)}}(\mathrm{e}^{\frac{2\varepsilon}{\sqrt{10k\log(1/\delta)}}} - 1) \\
&= \frac{2\varepsilon}{\sqrt{5}} + \frac{2\varepsilon\sqrt{k}}{\sqrt{10\log(1/\delta)}}(\mathrm{e}^{\frac{2\varepsilon}{\sqrt{10k\log(1/\delta)}}} - 1).
\end{aligned} \tag{A.8}$$

Recognizing that $\frac{2\varepsilon}{\sqrt{10k\log(1/\delta)}} \leq 0.0659$ under the assumptions $\varepsilon \leq 0.5, \delta \leq 0.1$ and $k \geq 10$, we get

$$\begin{aligned}\frac{2\varepsilon\sqrt{k}}{\sqrt{10\log(1/\delta)}}\left[\mathrm{e}^{\frac{2\varepsilon}{\sqrt{10k\log(1/\delta)}}}-1\right] &\leq \frac{2\varepsilon\sqrt{k}}{\sqrt{10\log(1/\delta)}} \times 1.034 \times \frac{2\varepsilon}{\sqrt{10k\log(1/\delta)}} \\ &\leq 0.0899\varepsilon.\end{aligned}$$

Substituting this line into (A.8) yields $\tilde{\varepsilon} \leq 2\varepsilon/\sqrt{5} + 0.0899\varepsilon \leq \varepsilon$, as desired.

□

Proof of Lemma 3.2. The proof is similar to that of Example 5.6.1 in [20]. By scaling, assume that ξ_i are exponential random variables with parameter 1, i.e, $\mathbb{E}\,\xi_i = 1$. Note that W_j is measurable with respect to $\mathcal{F}_j$. In the proof, we first consider the conditional expectation $\mathbb{E}(W_j^{-1}|\mathcal{F}_{j+1})$, then return to $\mathbb{E}(W_j|\mathcal{F}_{j+1})$ by applying Jensen's inequality. Specifically, we have

$$\mathbb{E}\left[\frac{\xi_l}{jT_{m+1}}\Big|\mathcal{F}_{j+1}\right] = \frac{1}{jT_{m+1}}\,\mathbb{E}(\xi_l|\mathcal{F}_{j+1})$$

because T_{m+1} is measurable in $\mathcal{F}_{j+1}$. Next, observe that by symmetry we get

$$\mathbb{E}(\xi_l|\mathcal{F}_{j+1}) = \mathbb{E}(\xi_k|\mathcal{F}_{j+1})$$

for any $l, k \leq j+1$. Combining the last two displays gives

$$\begin{aligned}\mathbb{E}\left[\frac{T_j}{jT_{m+1}}\Big|\mathcal{F}_{j+1}\right] &= \mathbb{E}\left[\frac{T_j}{(j+1)T_{m+1}}\Big|\mathcal{F}_{j+1}\right] + \sum_{l=1}^{j}\mathbb{E}\left[\frac{\xi_l}{j(j+1)T_{m+1}}\Big|\mathcal{F}_{j+1}\right] \\ &= \mathbb{E}\left[\frac{T_j}{(j+1)T_{m+1}}\Big|\mathcal{F}_{j+1}\right] + \sum_{l=1}^{j}\mathbb{E}\left[\frac{\xi_{j+1}}{j(j+1)T_{m+1}}\Big|\mathcal{F}_{j+1}\right] \\ &= \mathbb{E}\left[\frac{T_{j+1}}{(j+1)T_{m+1}}\Big|\mathcal{F}_{j+1}\right] \\ &= \frac{T_{j+1}}{(j+1)T_{m+1}}\,.\end{aligned}$$

To complete the proof, note that Jensen's inequality asserts that

$$\mathbb{E}(W_j|\mathcal{F}_{j+1}) \geq \frac{1}{\mathbb{E}(W_j^{-1}|\mathcal{F}_{j+1})} = \frac{(j+1)T_{m+1}}{T_{j+1}} = W_{j+1}\,,$$

as desired. □

Proof of Theorem 2. In addition to the proof sketch in Section 3.2, it remains to show that

$$\frac{1}{q}\,\mathbb{E}\left[\frac{j^\star}{j^\star + \max\{\lceil mU_{(j^\star)}/q\rceil - j^\star, 0\}}\right] \to C_k$$

as $q \to 0, m \to \infty, m - m_0 \to \infty$ and $m_0/m \to 1$. Since $j^\star = O_{\mathbb{P}}(1)$ and $mU_{(j^\star)}$ is bounded below away from 0 as $m_0 \to \infty$ and $m_0/m \to 1$, one can show that[1]

$$\max\{\lceil mU_{(j^\star)}/q\rceil - j^\star, 0\} = mU_{(j^\star)}/q - j^\star + O_{\mathbb{P}}(1).$$

[1]One needs to ensure that $m - m_0$ is larger than $\max\{\lceil mU_{(j^\star)}/q\rceil - j^\star, 0\}$ with high probability. Thus, q should tend to 0 slowly as $m - m_0 \to \infty$.

Thus, we get

$$\frac{1}{q}\cdot\frac{j^\star}{j^\star+\max\{\lceil mU_{(j^\star)}/q\rceil-j^\star,0\}}=\frac{j^\star}{mU_{(j^\star)}}+o_{\mathbb{P}}(1)$$
$$=\max_{k\le j\le m_0}\frac{j}{mU_{(j)}}+o_{\mathbb{P}}(1).$$

By assumption, we have

$$\mathbb{E}\left[\max_{k\le j\le m_0}\frac{j}{mU_{(j)}}\right]=(1+o(1))\,\mathbb{E}\left[\max_{k\le j\le m_0}\frac{j}{m_0U_{(j)}}\right]\to C_k$$

as $m_0\to\infty$.

To complete the proof, the last step is to show that

$$\frac{1}{q}\cdot\frac{j^\star}{j^\star+\max\{\lceil mU_{(j^\star)}/q\rceil-j^\star,0\}}$$

is bounded by an integrable random variable. To this end, we note that if $mU_{(j^\star)}/q\ge j^\star$, then

$$\frac{1}{q}\cdot\frac{j^\star}{j^\star+\max\{\lceil mU_{(j^\star)}/q\rceil-j^\star,0\}}\le\frac{j^\star}{mU_{(j^\star)}},$$

and otherwise $q\ge mU_{(j^\star)}/j^\star$, yielding

$$\frac{1}{q}\cdot\frac{j^\star}{j^\star+\max\{\lceil mU_{(j^\star)}/q\rceil-j^\star,0\}}\le\frac{1}{q}\cdot\frac{j^\star}{j^\star+0}\le\frac{j^\star}{mU_{(j^\star)}}.$$

Note that $\frac{j^\star}{mU_{(j^\star)}}$ is bounded by an integrable random variable by resorting the representation using T_j. □

Proof of Lemma 3.6. We first prove the case where $k\ge 3$. The uniform integrability follows if we show

$$\sup_{n\ge k}\mathbb{E}\left(\max_{k\le j\le n}\frac{jT_{n+1}}{nT_j}\right)^2<\infty. \tag{A.9}$$

As proved by Lemma 3.2, $\frac{jT_{n+1}}{nT_j}$ is a backward submartingale for $j=k,k+1,\dots,n$. Thus, by Doob's maximal inequality, we get

$$\mathbb{E}\left(\max_{k\le j\le n}\frac{jT_{n+1}}{nT_j}\right)^2\le 4\,\mathbb{E}\left(\frac{kT_{n+1}}{nT_k}\right)^2$$

So, the proof would be completed if we verify

$$\sup_{n\ge k}\mathbb{E}\left(\frac{kT_{n+1}}{nT_k}\right)^2<\infty.$$

To this end, note that

$$\begin{aligned}
\mathbb{E}\,\frac{k^2T_{n+1}^2}{n^2T_k^2} &= \frac{k^2}{n^2}\,\mathbb{E}\,\frac{(T_k+T_{n+1}-T_k)^2}{T_k^2}\\
&= \frac{k^2}{n^2}\,\mathbb{E}\,\frac{T_k^2+2T_k(n+1-k)+(n+1-k)^2+(n+1-k)}{T_k^2}\\
&\le \frac{k^2}{n^2}\,\mathbb{E}\,\frac{T_k^2+2nT_k+n^2}{T_k^2}\\
&\le \frac{k^2}{n^2}+\frac{2k^2}{n}\,\mathbb{E}\,\frac{1}{T_k}+k^2\,\mathbb{E}\,\frac{1}{T_k^2}\\
&\le 1+2k\,\mathbb{E}\,\frac{1}{T_k}+k^2\,\mathbb{E}\,\frac{1}{T_k^2},
\end{aligned}$$

which is finite if $k \ge 3$. Thus, (A.9) holds for $k \ge 3$.

Next, we turn to the case of $k = 2$. Recognize that

$$\begin{aligned}
\max_{2\le j\le n}\frac{jT_{n+1}}{nT_j} &\le \frac{2T_{n+1}}{nT_2}+\max_{3\le j\le n}\frac{jT_{n+1}}{nT_j}\\
&= \frac{2}{n}+\frac{2(T_{n+1}-T_2)}{nT_2}+\max_{3\le j\le n}\frac{jT_{n+1}}{nT_j}
\end{aligned}$$

Since $\frac{2}{n}$ and $\max_{3\le j\le n}\frac{jT_{n+1}}{nT_j}$ are both uniformly integrable, it is sufficiently to show the uniform integrability of $\frac{2(T_{n+1}-T_2)}{nT_2}$ for $n \ge 2$. To this end, note that

$$\begin{aligned}
\mathbb{E}\left[\frac{2(T_{n+1}-T_2)}{nT_2}\right]^{1.5} &= \frac{2^{1.5}}{n^{1.5}}\,\mathbb{E}(T_{n+1}-T_2)^{1.5}\,\mathbb{E}\,T_2^{-1.5}\\
&\le \frac{2^{1.5}}{n^{1.5}}\left[\mathbb{E}(T_{n+1}-T_2)^2\right]^{\frac{3}{4}}\mathbb{E}\,T_2^{-1.5}\\
&= \frac{2^{1.5}}{n^{1.5}}\left[(n-1)^2+n-1\right]^{\frac{3}{4}}\mathbb{E}\,T_2^{-1.5}\\
&< 2^{1.5}\,\mathbb{E}\,T_2^{-1.5},
\end{aligned}$$

which is finite. The proof is complete. □

Proof of Lemma 4.2. We first consider part one. Note that

$$\mathbb{P}\left(Z_j \le -\lambda\log\frac{n}{2\alpha}\right) = \frac{1}{2}\times\frac{2\alpha}{n} = \frac{\alpha}{n}. \tag{A.10}$$

Hence, taking a union bound, we get

$$\begin{aligned}
\mathbb{P}\left(\text{all } Z_j > -\lambda\log\frac{n}{2\alpha}\right) &= 1-\mathbb{P}\left(\min Z_j \le -\lambda\log\frac{n}{2\alpha}\right)\\
&\ge 1-\sum_{j=1}^{n}\mathbb{P}\left(Z_j \le -\lambda\log\frac{n}{2\alpha}\right)\\
&= 1-n\frac{\alpha}{n}\\
&= 1-\alpha.
\end{aligned}$$

The proof of part two is the follows the same reasoning except using

$$\mathbb{P}\left(|Z_j| \geq \lambda \log \frac{n}{\alpha}\right) = \frac{\alpha}{n}$$

in place of (A.10).

□

Proof of Theorem 6. In the proof below, we replace the assumption on the nominal level with the relaxed assumption that $q \geq 6m^{-1.5}$. Let $0 < \alpha, \alpha' < 1$ be specified later. Denote by $R'_{\text{SD}} = \min\{R_{\text{SD}}, m'\}$ and let $p_{j_1}, \ldots, p_{j_{R'_{\text{SD}}}}$ be the R'_{SD} smallest p-values. By the construction of the step-down procedure, we get

$$\max\{p_{j_1}, \ldots, p_{j_{R'_{\text{SD}}}}\} \leq \frac{qR'_{\text{SD}}}{m}.$$

First, we point out that the first R'_{SD} selections (without added noise) in the peeling stage of PrivateBHq, denoted as $\theta_{i_1}, \ldots, \theta_{i_{R'_{\text{SD}}}}$, obey

$$\max\{\theta_{i_1}, \ldots, \theta_{i_{R'_{\text{SD}}}}\} \leq \log \frac{qR'_{\text{SD}}}{m} + 2\lambda \log \frac{m^2}{\alpha} \tag{A.11}$$

with probability at least $1 - \alpha$. To show this, we recognize that, with probability at least $1 - \alpha$, all the mm' noise terms added by PrivateBHq are bounded in absolute value by

$$\lambda \log \frac{mm'}{\alpha} \leq \lambda \log \frac{m^2}{\alpha} \tag{A.12}$$

by using Lemma 4.2. Now, consider the $l$the step of invoking the peeling, where $1 \leq l \leq R'_{\text{SD}}$. Note that at least one of $\theta_{j_1}, \ldots, \theta_{j_{R'_{\text{SD}}}}$ remains on the list. Hence, at least one candidate for Report Noisy Min is no greater than

$$\log \max\left\{\frac{qR'_{\text{SD}}}{m}, \nu\right\} + \lambda \log \frac{m^2}{\alpha} = \log \frac{qR'_{\text{SD}}}{m} + \lambda \log \frac{m^2}{\alpha}.$$

Then, it must hold that

$$\theta_{i_l} + Z'_{i_l} \leq \log \frac{qR'_{\text{SD}}}{m} + \lambda \log \frac{m^2}{\alpha},$$

where on the event (A.12) $Z'_{i_l} \geq -\lambda \log \frac{m^2}{\alpha}$. Therefore, we get

$$\theta_{i_l} \leq \log \frac{qR'_{\text{SD}}}{m} + 2\lambda \log \frac{m^2}{\alpha},$$

thus confirming (A.11).

With added noise, the counts satisfy

$$\begin{aligned} \tilde{\theta}_{i_l} &\leq \log \frac{qR'_{\text{SD}}}{m} + 2\lambda \log \frac{m^2}{\alpha} + \lambda \log \frac{m'}{2\alpha'} \\ &\leq -\log \frac{m}{qR'_{\text{SD}}} + 2\lambda \log \frac{m^2}{\alpha} + \lambda \log \frac{m}{2\alpha'} \end{aligned} \tag{A.13}$$

with probability at least $1 - \alpha - \alpha'$ for all $l = 1, \ldots, R'_{\text{SD}}$. Next, take

$$2\lambda \log \frac{m^2}{\alpha} + \lambda \log \frac{m}{2\alpha'} \leq \frac{16\eta\sqrt{k \log(1/\delta)} \log m}{\varepsilon} \tag{A.14}$$

as given for the moment. Then, from (A.13) we get

$$\tilde{\theta}_{i_l} \leq -\log \frac{m}{qR'_{\mathrm{SD}}} + \frac{16\eta\sqrt{k\log(1/\delta)}\log m}{\varepsilon} \tag{A.15}$$

for all $i = 1, \ldots, R'_{\mathrm{SD}}$ with probability at least $1 - \alpha - \alpha'$. Now, we turn to verify (A.14), which is equivalent to

$$2\log\frac{m^2}{\alpha} + \log\frac{m}{2\alpha'} \leq \frac{16}{\sqrt{10}}\log m.$$

To this end, it suffices to set $\alpha = m^{-0.014}$ and $\alpha' = m^{-0.029}/2$. Since both $\alpha, \alpha' \to 0$ as $m \to \infty$, we see that (A.15) holds with probability tending to one.

Recognizing (A.15), to reject all of these R'_{SD} hypotheses using PrivateBHq, it is sufficient to have

$$-\log\frac{m}{qR'_{\mathrm{SD}}} + \frac{16\eta\sqrt{k\log(1/\delta)}\log m}{\varepsilon} \leq -\log\frac{m}{q'R'_{\mathrm{SD}}} - \frac{\eta\sqrt{10k\log\frac{1}{\delta}}\log\frac{6m'}{q'}}{\varepsilon},$$

which is equivalent to

$$\begin{aligned}\log\frac{q'}{q} &\geq \frac{\eta\sqrt{10k\log\frac{1}{\delta}}\log\frac{6m'}{q'}}{\varepsilon} + \frac{16\eta\sqrt{k\log(1/\delta)}\log m}{\varepsilon} \\ &= \frac{\eta\sqrt{k\log(1/\delta)}}{\varepsilon}\left[\sqrt{10}\log\frac{6m'}{q'} + 16\log m\right].\end{aligned} \tag{A.16}$$

Since $q \geq 6m^{-1.5}$, we get

$$\begin{aligned}\sqrt{10}\log\frac{6m'}{q'} &\leq \sqrt{10}\log\frac{6m}{q} \\ &\leq \sqrt{10}\log\frac{6m}{6m^{-1.5}} \\ &< 8\log m.\end{aligned}$$

Hence, (A.16) is implied by

$$\log\frac{q'}{q} \geq \frac{24\eta\sqrt{k\log(1/\delta)}\log m}{\varepsilon},$$

which is in fact an equality by assumption. Thus, the proof is complete.

□

Proof of Corollary 4.3. A careful look at the proof of Theorem 5 reveals that the event in Proposition 4.1 holds with probability at least $1 - q/12$. As such, the bound on the FDR_k in Theorem 5 can be strengthened to $(C_k + 1/12)q$.

In light of the above, we set c' such that

$$(C_k + 1/12)(1 + c') = C_k + 0.1.$$

Then, PrivateBHq at level $(1 + c')q$ controls the FDR_k at level $(C_k + 0.1)q$ as ensured by Theorem 5.

It remains to prove that the claim of Theorem 6 also holds. To this end, we only need to show that q' that is given in the statement of Theorem 6 is less than $(1 + c')q$ for sufficiently large m, n. That is,

$$(1 + c')q > q' \equiv q\mathrm{e}^{\frac{24\eta\sqrt{m'\log(1/\delta)}\log m}{\varepsilon}}.$$

In both examples, $\eta = n^{-0.5+o(1)}$ and $\log m = \log \texttt{poly}(n) = n^{o(1)}$. Thus, we have

$$\mathrm{e}^{\frac{24\eta\sqrt{m'\log(1/\delta)}\log m}{\varepsilon}} \to 0$$

due to $m' \le n^{1-c}$. Therefore, in words, PrivateBHq at level $(1+c')q$ should be at least as powerful as the truncated BHq step-down procedure with probability tending to one.

□

Proof of Proposition 4.4. We first prove that PrivateBonf is (ε, δ)-differentially private. To this end, we start by observing that each θ_j corrupted by $\mathrm{Lap}(\tilde{\lambda})$ noise is ε'-differentially private, where

$$\varepsilon' = \frac{\eta}{\tilde{\lambda}} = \frac{2\varepsilon}{\sqrt{10m\log(1/\delta)}}.$$

Using the Advanced Composition Theorem, therefore, it suffices to show that

$$\varepsilon'\sqrt{2m\log(1/\delta)} + m\varepsilon'(\mathrm{e}^{\varepsilon'} - 1) \le \varepsilon,$$

Under the assumptions of Theorem 3, we have

$$\mathrm{e}^{\varepsilon'} - 1 \le 1.034\varepsilon'$$

Thus, the proof would be completed once we show

$$\varepsilon'\sqrt{2m\log(1/\delta)} + 1.034m\varepsilon'^2 \le \varepsilon,$$

which is equivalent to

$$\frac{2}{\sqrt{5}} + 1.034 \times \frac{2\varepsilon}{5\log(1/\delta)} \le 1.$$

This inequality can be easily verified.

Next, we turn to show the second statement. As with the proof of Theorem 5, we only need to show that, with probability at least $1 - 0.1q$, all noisy counts $\theta_j + \mathrm{Lap}(\tilde{\lambda})$ with $p_j > q/m$ are above

$$\log\frac{q}{m} - \frac{\eta\sqrt{10m\log(1/\delta)}\log(5m/q)}{2\varepsilon}.$$

This statement is implied if m i. i. d. $\mathrm{Lap}(\tilde{\lambda})$ noise terms are all above $-\frac{\eta\sqrt{10m\log(1/\delta)}\log(5m/q)}{2\varepsilon}$ with probability at least $1 - 0.1q$. This claim is true by invoking Lemma 4.2.

□

Appendix B. More Simulation Results

B.1. **Comparions of FDR_k-controlling procedures with others.** In this section, we follow the setting of Figures 2 and 3, and compare the FDR_k and power of a step-up procedure [52, 54], PrivateBHq, and PrivateBonf. Specifically, we consider the case of $k = 3$, and use the critical values

$$q_j = \left(\frac{q\max\{j,k\}}{m}\prod_{i=1}^{k-1}\frac{i}{m - \max\{j,k\} + i}\right)^{\frac{1}{k}}$$

for $j = 1, \ldots, m$, where $k = 3$ (see Eqn. (5.3) in [54]). For simplicity, we refer to this procedure as BHq_k. Figures 4 and 5 display the results.

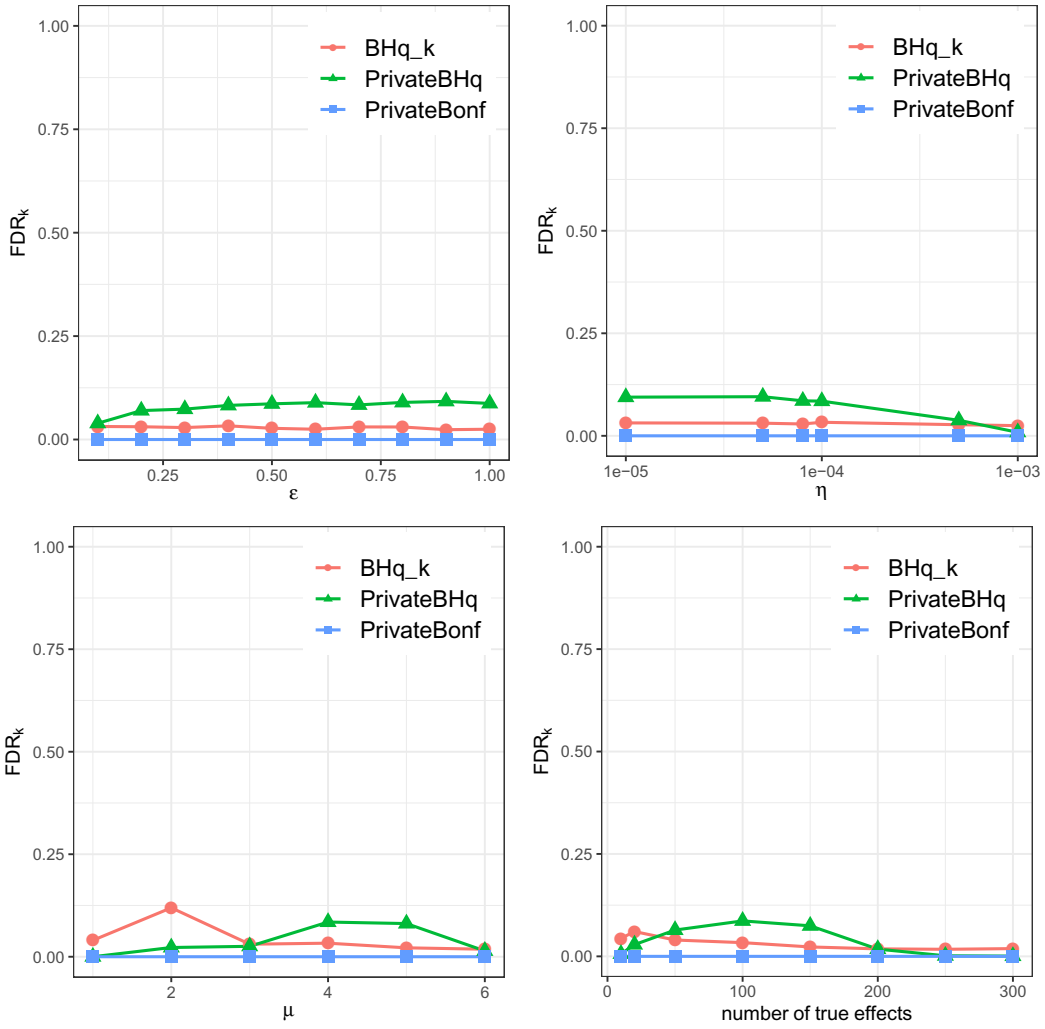


FIGURE 4. The FDR of BHq$_k$, PrivateBHq, and PrivateBonf, plotted against varying $\varepsilon, \eta, \mu, m_1$, respectively, and averaged over 100 independent replicates.

B.2. **BHq under Negative Dependence.** This section features three simulated examples with certain negative dependence between the true null and false null test statistics. The simulation results empirically show that the BHq step-up procedure controls FDR_2 and FDR_5, and this is consistent with Theorem 1. Throughout, $\mathcal{N}_0$ with cardinality m_0 and $\mathcal{N}_1$ with cardinality $m_1 \equiv m - m_0$ denote the set of true null hypotheses and the set of false null hypotheses, respectively.

Example B.1 (Multivariate Normal)**.** Consider observing $X \sim \mathcal{N}(\mu, \Sigma)$. The covariance Σ is constructed as follows: $\Sigma_{ii} = 1$ for all $1 \le i \le m$, $\Sigma_{ij} = 0$ if $i \ne j$ and both $i, j \in \mathcal{N}_0$ or $i, j \in \mathcal{N}_1$, and $\Sigma_{ij} = -1/\sqrt{m_0 m_1}$ if one of i, j belongs to $\mathcal{N}_0$ and the other belongs to $\mathcal{N}_1$ (if this value is set to be smaller than $-1/\sqrt{m_0 m_1}$, the covariance Σ is not positive semidefinite). The distribution of X satisfies the IWN condition and, therefore, Theorem 1 guarantees FDR control of the BHq procedure used to test $\mu_i = 0$ against the one-sided

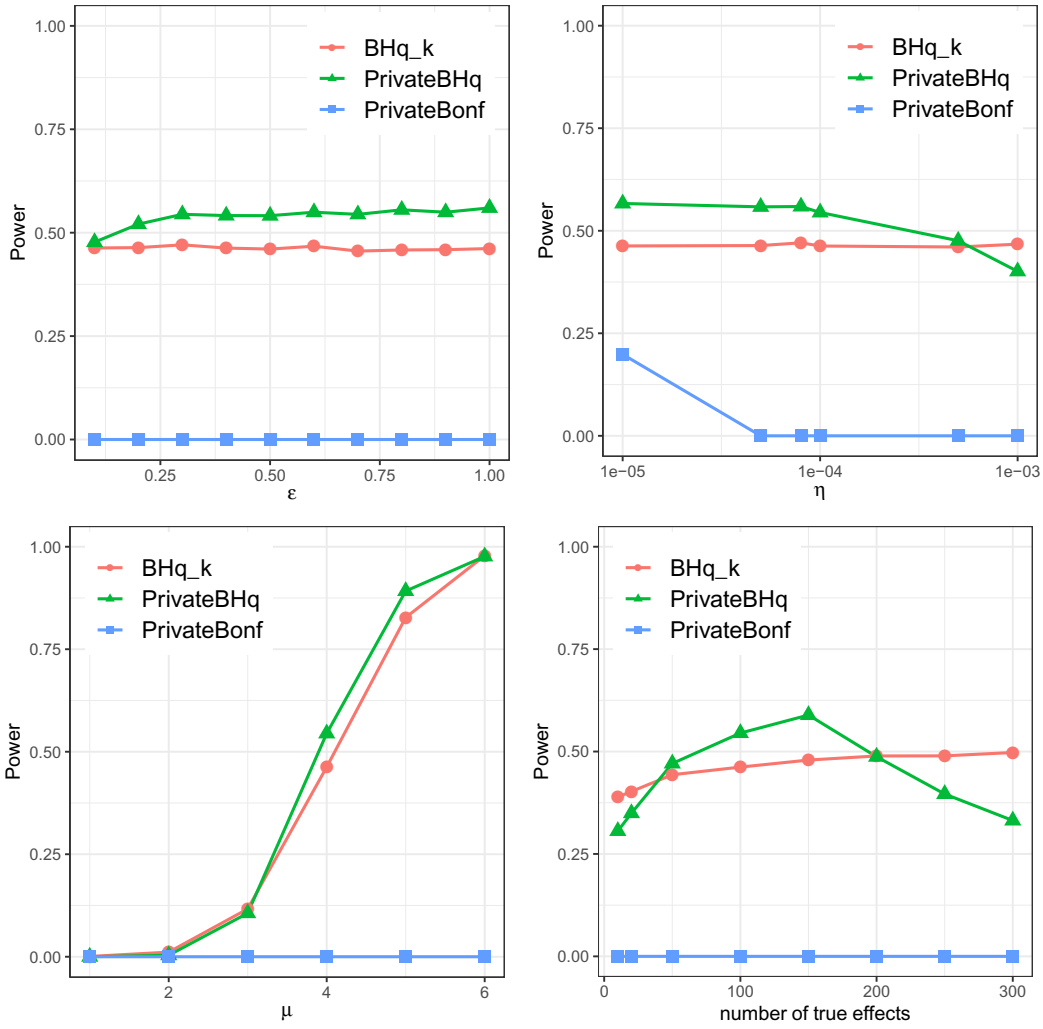
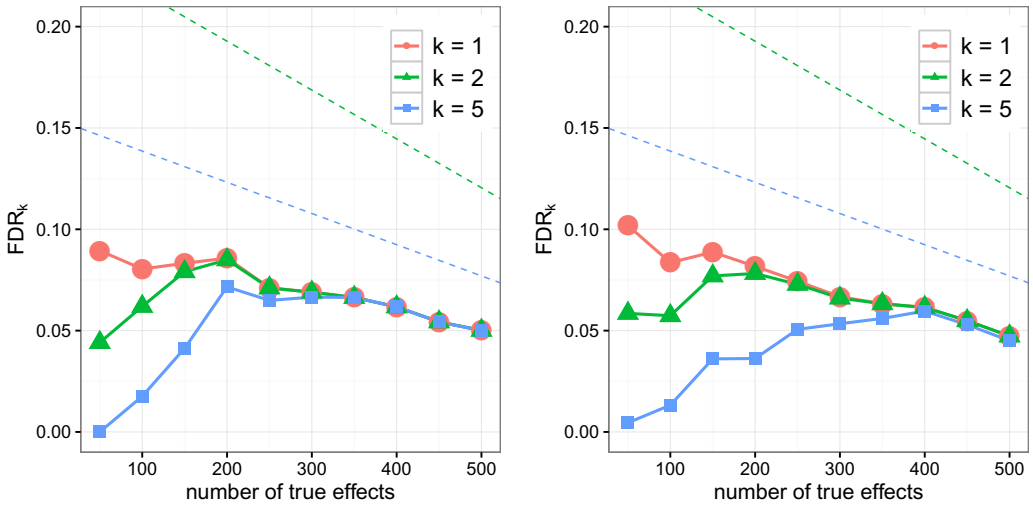


FIGURE 5. The power of BHq_k, PrivateBHq, and PrivateBonf, plotted against varying $\varepsilon, \eta, \mu, m_1$, respectively, and averaged over 100 independent replicates.

alternative $\mu_i > 0$. In contrast, the results of [7] are not applicable because the PRDS property does not hold due to $-1/\sqrt{m_0 m_1} < 0$. Furthermore, Theorem 1 is still valid for testing against the two-sided alternatives $\mu_i \neq 0$. In general, the PRDS property is not satisfied for two-sided tests (see discussion in Section 3.1 of [7]).

Figure 6 presents the empirical FDR, FDR_2, and FDR_5 of the BHq procedure for both one-sided and two-sided alternatives in this example and, in addition, the bound $C_k\theta_0 q$ in Theorem 1 for $k = 2, 5$ in dashed lines. As predicted by Theorem 1, the empirical FDR_2 and FDR_5 are indeed below their corresponding dashed lines. In fact, the empirical values are much below the bounds in Theorem 1, which are derived by assuming a least favorable dependence structure between the nulls and non-nulls. This pattern is also observed in the following two plots. Moreover, these empirical error rates decrease eventually as the number of true effects m_1 increases, which reflects the presence of the true null proportion θ_0 in

the bound $C_k\theta_0 q$. Notably, this bound can be smaller than the nominal level q if m_1 is sufficiently large.



(A) one-sided tests (B) two-sided tests

FIGURE 6. FDR_k of the BHq for $k = 1, 2, 5$ in Example B.1, with level $q = 0.1$. The FDR_1 is just the usual FDR. We set m to 1000, vary m_1 from 50 to 500, and let $\mu_i = 2$ for $1 \le i \le m_1$ and $\mu_i = 0$ otherwise. The covariance matrix Σ has ones on the diagonal; $\Sigma_{ij} = \Sigma_{ji} = -1/\sqrt{m_0 m_1}$ for all $1 \le i \le m_1$ and $m_1 + 1 \le j \le m$; all the rest entries are zero. The results are averaged over 100 replicates. The upper and lower dashed lines denote the bounds $C_2\theta_0 q$ and $C_5\theta_0 q$, respectively.

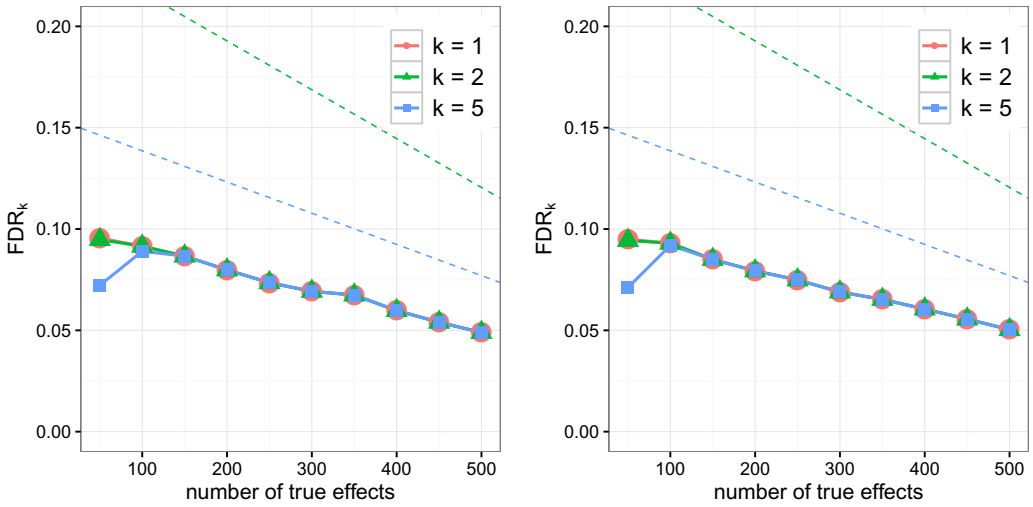
Example B.2 (Multivariate t-Distribution with Different Denominators)**.** Consider observing i. i. d. vectors $X^{(1)}, \ldots, X^{(n)}$ from $\mathcal{N}(\mu, \Sigma)$, where both μ and Σ are the same as the previous example. To test $\mu_i = 0$ against $\mu_i > 0$ or $\mu_i \neq 0$, we use the t-test statistics

$$t_i = \frac{\sqrt{n}\bar{X}_i}{\sqrt{\frac{1}{n-1}\sum_{l=1}^{n}(X_i^{(l)} - \bar{X}_i)^2}},$$

where $\bar{X}_i = (X_i^{(1)} + \cdots + X_i^{(n)})/n$ for $i = 1, \ldots, m$. As earlier, Theorem 1 applies to this example, as opposed to the existing FDR literature, which fails to ensure FDR control of the BHq procedure in this example.

Numerical results for Example B.2 are displayed in Figure 7. The setup follows Example B.1, with n being set to 10. While the behavior of the BHq procedure in Figure 6 basically remains the same in the present plot, we wish to point out that the effect of the true null proportion θ_0 is more pronounced in the present simulation study and the three error rates coincide exactly once m_1 exceeds 100 as the BHq in this setting always rejects a substantial number of true nulls. The latter shows the difference between the FDR and FDR_k is inconsequential in this example.

Example B.3 (Multivariate Normal with Block-Diagonal Covariance)**.** Consider bivariate normal variables $X_i, \widetilde{X}_i$ with means $\mu_i = 0$ and $\widetilde{\mu}_i \neq 0$, respectively, for $i = 1, \ldots, m$. Let



(A) one-sided tests (B) two-sided tests

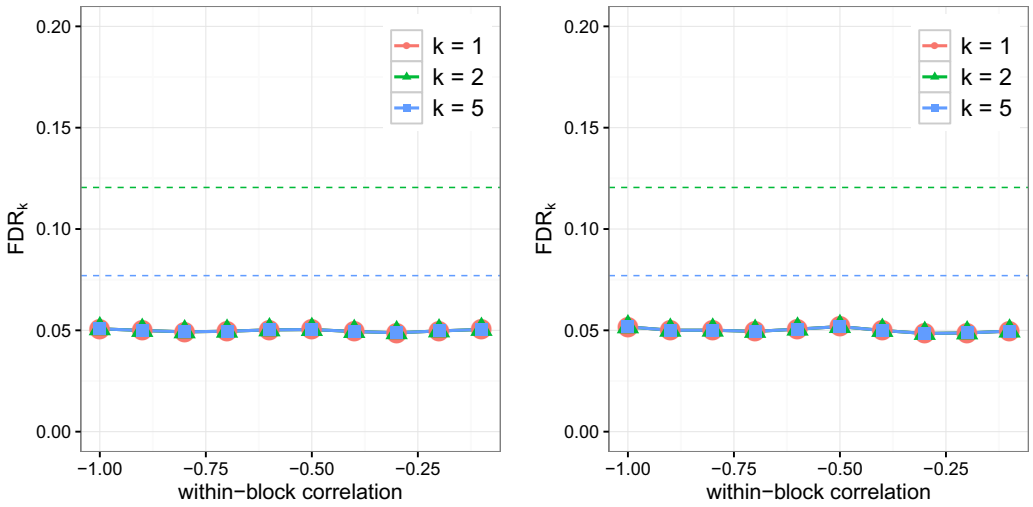
FIGURE 7. FDR$_k$ of the BHq for $k = 1, 2, 5$ in Example B.2, with level $q = 0.1$. The experimental setup is the same as Figure 6. The parameter n is set to 10.

Var X_i = Var $\widetilde{X}_i = 1$ for all i and the m pairs $(X_i, \widetilde{X}_i)$ be jointly independent. Thus, the $2m$ normal variables exhibit a diagonal-block covariance matrix that is formed by m 2×2 blocks on the diagonal. The correlation corr$(X_i, \widetilde{X}_i)$ within every block varies from -1 to -0.1. Note that there are m true nulls among the $2m$ hypotheses and, therefore, $\theta_0 = 0.5$. The IWN condition is satisfied because all true nulls are located in different blocks. Consequently, the BHq procedure maintains FDR$_k$ control in this example by applying Theorem 1, as opposed to existing results in the literature, which to our knowledge are not capable of confirming the FDR control for this example. Moreover, the usual FDR control follows from Theorem 1 as a corollary, whose proof can be found in the appendix. As an appealing feature of this result, the dependence within each block can be arbitrary and even be different across blocks.

Corollary B.4 . *Fix $0 < q < 1$. Assume that $\{1 \le i \le m : \widetilde{\mu}_i \ge c_1\}/m \ge c_2$ for positive constants c_1, c_2 in Example B.3. For both one-sided and two-sided alternatives, the BHq procedure controls the usual FDR in an asymptotic sense. That is, as $m \to \infty$, we get* FDR $\le (1 + o_m(1))q$.

The numerical results are summarized in Figure 8. Note that the three FDR variants coincide through the range of within-block correlations. Interestingly, the bound corresponding to $k = 5$ (the lower dashed line) is below the nominal level $q = 0.1$.

Proof of Corollary B.4. The proof idea is to apply Theorem 1 or Corollary 1.3 and recognize the number of rejections R in Example B.3 tends to infinity. We only consider the case of one-sided alternatives and the proof of the two-sided alternatives case is very similar.



(A) one-sided tests (B) two-sided tests

FIGURE 8. FDR_k of the BHq for $k = 1, 2, 5$ in Example B.3, with level $q = 0.1$. Here, m is set to 5000, $\widetilde{\mu}_i$ is set to 1.5 for all i. Note that the true null proportion $\theta_0 = 0.5$. All points represent the average of 100 independent runs. The correlation between X_i and $\widetilde{X}_i$ is set to be the same across all i, varying from -1 to -0.1.

Assume for the moment that $R \to \infty$ with probability tending to one as $m \to \infty$. Then, we have

$$\begin{aligned} |\mathrm{FDP} - \mathrm{FDP}_k| &= \frac{V \mathbf{1}_{V<k}}{\max\{R, 1\}} \\ &\le \frac{k-1}{\max\{R, 1\}} \\ &= o_{\mathbb{P}}(1). \end{aligned}$$

Thus, from Theorem 1, we get

$$\begin{aligned} \mathrm{FDR} &= \mathrm{FDR}_k + o_m(1) \\ &\le C_k \theta_0 q + o_m(1) \\ &\le C_k q + o_m(1). \end{aligned} \tag{B.1}$$

Letting $k \to \infty$, one gets $C_k \to 1$. Thus, from (B.1) it follows that

$$\mathrm{FDR} \le q + o_m(1).$$

Now, we aim to complete the proof by showing $R \to \infty$ in probability. In fact, it will be shown that $\mathbb{P}(R \ge \lfloor \sqrt{m} \rfloor) \to 1$. By the construction of the step-up procedure, $R \ge \lfloor \sqrt{m} \rfloor$ if

$$\#\left\{ i : \Phi(-X_i) \le \frac{q\lfloor \sqrt{m} \rfloor}{2m} \right\} + \#\left\{ i : \Phi(-\widetilde{X}_i) \le \frac{q\lfloor \sqrt{m} \rfloor}{2m} \right\} \ge \lfloor \sqrt{m} \rfloor,$$

which is implied by

$$\#\left\{ i : \Phi(-\widetilde{X}_i) \le \frac{q\lfloor \sqrt{m} \rfloor}{2m} \right\} \ge \lfloor \sqrt{m} \rfloor. \tag{B.2}$$

Now, we aim to show (B.2) holds with probability tending to one. Denote by $A = \{i : \widetilde{\mu}_i \geq c_1\}$, which, by assumption, satisfies $\#A \geq c_2 m$. Consider $W_i := \widetilde{X}_i - \widetilde{\mu}_i \sim \mathcal{N}(0,1)$ for $i \in A$ and let $W_{(1)} \geq \cdots \geq W_{(\#A)}$ be the order statistics. Note that (B.2) simply follows from

$$\Phi\left(-c_1 - W_{(\lfloor\sqrt{m}\rfloor)}\right) \leq \frac{q\lfloor\sqrt{m}\rfloor}{2m},$$

which is equivalent to

$$-c_1 - W_{(\lfloor\sqrt{m}\rfloor)} \leq \Phi^{-1}\left(\frac{q\lfloor\sqrt{m}\rfloor}{2m}\right). \tag{B.3}$$

To prove (B.3), we make two observations:

$$\Phi^{-1}\left(\frac{q\lfloor\sqrt{m}\rfloor}{2m}\right) = -\sqrt{2\log\frac{2m}{q\lfloor\sqrt{m}\rfloor}} + o(1) \tag{B.4}$$

$$W_{(\lfloor\sqrt{m}\rfloor)} = \sqrt{2\log\frac{\#A}{\lfloor\sqrt{m}\rfloor}} + o_{\mathbb{P}}(1) \tag{B.5}$$

as both $q\lfloor\sqrt{m}\rfloor/(2m)$ tends to zero and $\#A/\lfloor\sqrt{m}\rfloor$ diverges to infinity. Above, (B.4) is a standard result and a proof of (B.5) can be found in Chapter 2 of [17]. Hence, it suffices to show

$$-\frac{c_1}{2} - \sqrt{2\log\frac{c_2 m}{\lfloor\sqrt{m}\rfloor}} \leq -\sqrt{2\log\frac{2m}{q\lfloor\sqrt{m}\rfloor}}$$

for sufficiently large m, which is true since

$$\begin{aligned}\sqrt{2\log\frac{c_2 m}{\lfloor\sqrt{m}\rfloor}} - \sqrt{2\log\frac{2m}{q\lfloor\sqrt{m}\rfloor}} &= \sqrt{\log m + 2\log c_2 + o(1)} - \sqrt{\log m + 2\log(2/q) + o(1)} \\ &= \sqrt{\log m} + o(1) - \sqrt{\log m} - o(1) \\ &= o(1).\end{aligned}$$

□