

A New Proof of FDR Control Based on Forward Filtration

Ahmad Reza Ehyaei¹, Kasra Alishahi¹, and Ali Shojaie²

¹ Sharif University of Technology, Tehran, Iran.

² University of Washington, USA.

Received: 30/08/2019, Revision received: 03/04/2020, Published online: 13/05/2020

Abstract. For multiple testing problems, Benjamini and Hochberg (1995) proposed the false discovery rate (FDR) as an alternative to the family-wise error rate (FWER). Since then, researchers have provided many proofs to control the FDR under different assumptions. Storey et al. (2004) showed that the rejection threshold of a BH step-up procedure is a stopping time with respect to the reverse filtration generated by the p-values and proposed a new proof based on the martingale theory. Following this work, martingale methods have been widely used to establish FDR control in various settings, but have been primarily applied to reverse filtration only. However, forward filtration can be more amenable for generalized and adaptive FDR controlling procedures. In this paper, we present a new proof, based on forward filtration, for step-down FDR controlling procedures that start from small p-values and update the rejection regions as larger p-values are observed.

Keywords. Benjamini-Hochberg, False Discovery Rate, Step-down Methods

MSC: 62F03; 62C05.

Corresponding Author: Ahmad Reza Ehyaei(ahmad.ehyaei@gmail.com)
Kasra Alishahi(alishahi@sharif.ir)
Ali Shojaie(ashojaie@uw.edu)

1 Introduction

Multiple hypothesis testing is a key step in many large-scale inference problems in molecular biology, neuroscience, and astronomy. False discovery rate (FDR) control, introduced in the seminal paper by Benjamini and Hochberg (1995), is one of the most important methodological developments in this area. To control the FDR at a predetermined level α , Benjamini and Hochberg (BH) proposed a *step-up* procedure based on ranked p -values. For n hypotheses consisting of n_0 nulls and n_1 non-nulls, the step-up BH procedure controls the FDR at the level $n_0\alpha/n$, and continues to work under certain positive dependencies among hypotheses. A modification of the procedure proposed by Benjamini and Yekutieli (2001) also controls the FDR in general dependence settings, albeit more conservatively.

To prove that their procedure controls the FDR, Benjamini and Hochberg Benjamini and Hochberg (1995) followed the traditional hypothesis testing paradigm, where they fixed an error rate α and found a suitable rejection region $[0, p_{(k^*)}]$. In an alternative approach, the pioneering work of Storey et al. (2004) proposed a dynamic approach with a fixed rejection region of the form $[0, \alpha]$ for $\alpha \in [0, 1]$, and estimated the FDR for this region. To this end, Storey et al. (2004) used an elegant martingale argument by casting the FDR controlling procedure as a reverse filtration. Martingale arguments have also been used by Peña et al. (2011), Heesen and Janssen Heesen et al. (2015) and Benditkis Benditkis et al. (2018) to establish FDR control for multiple testing procedures. However, existing martingale arguments use a reverse filtration. In contrast, arguments based on forward filtrations, which can be more appropriate for adaptive FDR controlling procedures have not yet been explored.

In this paper, we present a new proof for *step-down* FDR controlling procedures. A step-down FDR procedure was first proposed by Benjamini and Liu (1999), and was later shown to control the FDR under positive dependence among hypotheses, Sarkar (2002). While for the same cut-offs step-up procedures are generally less conservative than their step-down counterparts, establishing Type-I error control for step-up procedures in complex hypothesis testing settings, e.g., in presence of correlation are generally more difficult than for step-down procedures. Motivated by recent efforts to establish FDR control for data-adaptive procedures Ramdas et al. (2017); Li and Barber (2019), we thus present an alternative proof for the classical step-down procedures of Benjamini and Liu (1999).

The proof of step-down procedures is in general more complicated than that of step-up procedures, which naturally form a martingale. To overcome these challenges,

here we present a new proof for the step-down procedure based on a *forward filtration*. In doing so, we have used more delicate proof techniques by recasting the problem as a martingale stopping time in order to prove the control of FDR. While we present the result and its proof under the original assumptions of Benjamini and Hochberg (1995), the new proof technique supports the development of adaptive data-adaptive procedures, e.g., for settings where external information/structure may be used to improve the power Li and Barber (2019); Ramdas et al. (2017).

2 A New Proof of FDR Control Using Step-down Procedures

In this section, we give a new proof of FDR control using step-down procedures under the classical assumptions of Benjamini and Hochberg (1995). The main result of this section (Theorem 2.1) has, in fact, been proved under weaker assumptions (see, e.g. Benditkis and Janssen, 2017), and with simpler proofs Benditkis et al. (2018). However, these, and other existing proofs, cannot be applied to data-adaptive procedures for FDR control, which is particularly beneficial in settings where incorporating additional structure can lead to improved power Li and Barber (2019). Our alternative proof based on a forward filtration can offer more flexibility in such settings.

Let H_1, \dots, H_n be n null hypotheses with p -values p_1, \dots, p_n . If H_i is true, then, $p_i \sim \text{Unif}[0, 1]$. Let $p_{(1)} \leq \dots \leq p_{(n)}$ be ordered p -values corresponding to hypotheses $H_{(1)}, \dots, H_{(n)}$, and consider the following empirical processes

$$\begin{aligned} a_t &= \#\{\text{null } p_i : p_i \leq t\}, \\ b_t &= \#\{\text{non-null } p_i : p_i \leq t\}, \\ r_t &= a_t + b_t. \end{aligned}$$

The false discovery proportion (FDP) and false discovery rate (FDR) can then be defined as

$$FDP(t) = \frac{a_t}{r_t \vee 1}, \quad FDR(t) = \mathbb{E} \left[\frac{a_t}{r_t \vee 1} \right],$$

where $a \vee b = \min(a, b)$.

It is well-known that the FDR can be controlled using the *step-up* BH procedure Benjamini and Hochberg (1995). In Storey et al. (2004), Storey *et al.* presented an alternative proof, based on an elegant martingale argument, to show that the step-up BH procedure controls the FDR. Briefly, they observe that a_t/t is a martingale as t decreases from 1 to 0. They then re-express the step-up BH procedure as rejecting all

p -values less than a threshold τ , which is a stopping time with respect to the reverse filtration as t goes from 1 to 0. Based on these observations and using the optional sampling theorem Klenke (2013), the authors prove that FDR is controlled by using the step-up BH procedure.

In this section, we modify the proof of Storey et al. (2004) to present a martingale-based argument for *step-down* FDR procedures Benjamini and Liu (1999). A step-down procedure for controlling the FDR at a user-specified level q rejects the hypotheses $H_{(1)}, \dots, H_{(i_{sd})}$ for

$$i_{sd} = \max\{i : p_{(i)} \leq \tau_{sd}\}, \quad \text{with} \quad \tau_{sd} = \inf \left\{ t : (r_t \vee 1) \leq \frac{nt}{q} \right\}, \quad (2.1)$$

and accepts the rest. For the stopping rule in a step-down procedure to be a *stopping time*, one should choose a filtration in the *forward* direction. However, in that case, a_t/t would no longer be a martingale, which complicates the proof.

We next present an alternative proof for the step-down procedure using techniques motivated by stochastic calculus. Similar to Benjamini and Hochberg (1995), we assume that the p_i s corresponding to null hypotheses are independent of each other and also independent of the p -values of non-null hypotheses. However, our proof of Theorem 2.1 uses a forward filtration, which allows us to start from small p -values and update the rejection regions as we observe larger p -values. Since it is more likely that smallest p -values are from non-null hypotheses, they can be used to a better estimation of the non-null distribution.

Theorem 2.1. *If the p -values corresponding to null hypotheses are independent of other p -values, then the step-down algorithm of (2.1) controls the FDR at the level q . More specifically,*

$$\mathbb{E} \left[\frac{a_{\tau_{sd}}}{r_{\tau_{sd}}} \right] \leq n_0 q \leq q. \quad (2.2)$$

Proof. By the definition of τ_{sd} in (2.1), $\mathbb{E} \left[\frac{a_{\tau_{sd}}}{r_{\tau_{sd}}} \right] = \mathbb{E} \left[\frac{q}{n} \frac{a_{\tau_{sd}}}{\tau_{sd}} \right]$. Thus, (2.2) is equivalent to

$$\mathbb{E} \left[\frac{1}{n} \frac{a_{\tau_{sd}}}{\tau_{sd}} \right] \leq 1. \quad (2.3)$$

We prove a stronger result, and show that

$$\mathbb{E} \left[\frac{1}{n_0} \frac{a_{\tau_{sd}}}{\tau_{sd}} \mid \{b_t\}_{t \in [0,1]} \right] \leq 1, \quad (2.4)$$

where n_0 is the number of null hypotheses. To see why it is sufficient to prove (2.4), note that

$$\mathbb{E} \left[\frac{1}{n} \frac{a_{\tau_{sd}}}{\tau_{sd}} \right] = \mathbb{E} \left(\mathbb{E} \left[\frac{1}{n} \frac{a_{\tau_{sd}}}{\tau_{sd}} \mid \{b_t\}_{t \in [0,1]} \right] \right) \leq \mathbb{E} \left(\mathbb{E} \left[\frac{1}{n_0} \frac{a_{\tau_{sd}}}{\tau_{sd}} \mid \{b_t\}_{t \in [0,1]} \right] \right).$$

Hence, (2.3) follows from (2.4).

To prove (2.4), first define $\tilde{a}_t = a_t - n_0 t$, and rewrite (2.4) as

$$\mathbb{E} \left[\frac{\tilde{a}_{\tau_{sd}}}{\tau_{sd}} \mid \{b_t\}_{t \in [0,1]} \right] \leq 0.$$

Now, fix a realization of $\{b_t\}_{t \in [0,1]}$ and note that since a and b are independent, conditioning on $\{b_t\}_{t \in [0,1]}$ does not change the distribution of $\{a_t\}_{t \in [0,1]}$, and, similarly, that of $\{\tilde{a}_t\}_{t \in [0,1]}$. Moreover, since $r_t = a_t + b_t$, the stopping time τ_{sd} can be redefined as

$$\tau_{sd} = \inf \left\{ t : \tilde{a}_t < \frac{nt}{q} - b_t - n_0 t \right\}.$$

The theorem then follows by a property of the process \tilde{a}_t , established in the following lemma. □

Lemma 2.1. *Let $\tilde{a}_t = a_t - n_0 t$; $\{\tilde{a}_t\}_{t \in [0,1]}$ be the process defined in Theorem 2.1, and g be an arbitrary deterministic function on $[0, 1]$ such that $g(1) \geq 0$. For a stopping time τ defined as*

$$\tau = \inf \{ t : \tilde{a}_t \leq g(t) \},$$

we have

$$\mathbb{E} \left[\frac{\tilde{a}_\tau}{\tau} \right] \leq 0. \tag{2.5}$$

Remark 1. Note that $\tilde{a}_1 = a_1 - n_0 = 0$ and so $\tilde{a}_1 \leq g(1)$. Therefore, $\{t : \tilde{a}_t \leq g(t)\}$ is not empty and τ is well-defined. In the proof of Theorem 2.1, $g(t) = nt/q - b_t - n_0 t$ so $g(1) = n/q - n_0 - n_1 \geq 0$.

Proof of Lemma 2.1. We first prove that for $\epsilon > 0$, $\mathbb{E} \left[\frac{\tilde{a}_\tau}{\tau + \epsilon} \right] \leq 0$. Note that \tilde{a}_t and $1/(t + \epsilon)$ are processes with bounded variation. Therefore,

$$d \left(\frac{\tilde{a}_t}{t + \epsilon} \right) = \frac{d\tilde{a}_t}{t + \epsilon} + \tilde{a}_t d \left(\frac{1}{t + \epsilon} \right), \tag{2.6}$$

and

$$\frac{\tilde{a}_\tau}{\tau + \epsilon} = \int_0^\tau \frac{1}{s + \epsilon} d\tilde{a}_s - \int_0^\tau \frac{\tilde{a}_s}{(s + \epsilon)^2} ds. \quad (2.7)$$

Now, we claim that the process η_t defined as

$$d\eta_s = d\tilde{a}_s + \frac{\tilde{a}_s}{1-s} ds, \quad (2.8)$$

is a martingale with respect to the filtration $\mathcal{F}_s = \sigma(a_u, u \leq s)$. To prove this claim, it suffices to show that $\mathbb{E}[\eta_t - \eta_s | \mathcal{F}_s] = 0$ for $t, s \in [0, 1]$ and $t \geq s$. But,

$$\begin{aligned} \mathbb{E}[\eta_t - \eta_s | \mathcal{F}_s] &= \mathbb{E}\left[\tilde{a}_t - \tilde{a}_s + \int_s^t \frac{\tilde{a}_u}{1-u} du \mid \mathcal{F}_s\right] \\ &= \mathbb{E}[\tilde{a}_t - \tilde{a}_s | \mathcal{F}_s] + \int_s^t \frac{\mathbb{E}[\tilde{a}_u | \mathcal{F}_s]}{1-u} du. \end{aligned} \quad (2.9)$$

It is then easy to see that the process \tilde{a}_t satisfies

$$\mathbb{E}[\tilde{a}_t | \mathcal{F}_s] = \frac{1-t}{1-s} \tilde{a}_s,$$

which is true, because conditioned on \mathcal{F}_s , $a_t - a_s$ counts the number of null hypothesis with p -values in $[s, t]$, among the $n_0 - a_s$ hypotheses with p -values in $[s, 1]$. Thus, $a_t - a_s$ has a binomial distribution $B(n_0 - a_s, \frac{t-s}{1-s})$. Hence,

$$\mathbb{E}[a_t - a_s | \mathcal{F}_s] = (n_0 - a_s) \frac{t-s}{1-s}. \quad (2.10)$$

Therefore,

$$\mathbb{E}[\tilde{a}_t | \mathcal{F}_s] = \mathbb{E}[a_t - n_0 t | \mathcal{F}_s] = \mathbb{E}[(a_t - a_s) + a_s - n_0 t | \mathcal{F}_s] = (n_0 - a_s) \frac{t-s}{1-s} + a_s - n_0 t = \frac{1-t}{1-s} (a_s - n_0 s). \quad (2.11)$$

The right-hand side of (2.9) can thus be simplified to

$$\mathbb{E}[\tilde{a}_t - \tilde{a}_s | \mathcal{F}_s] + \int_s^t \frac{\mathbb{E}[\tilde{a}_u | \mathcal{F}_s]}{1-u} du = \frac{s-t}{1-s} \tilde{a}_s + \int_s^t \frac{\frac{1-u}{1-s} \tilde{a}_s}{1-u} du = 0.$$

Now, substituting (2.8) into (2.7), we get

$$\begin{aligned} \frac{\tilde{a}_\tau}{\tau + \epsilon} &= \int_0^\tau \frac{1}{s + \epsilon} d\eta_s - \int_0^\tau \frac{\tilde{a}_s}{(s + \epsilon)(1-s)} ds - \int_0^\tau \frac{\tilde{a}_s}{(s + \epsilon)^2} ds \\ &= \int_0^\tau \frac{1}{s + \epsilon} d\eta_s - \int_0^\tau \frac{1 + \epsilon}{(s + \epsilon)^2(1-s)} \tilde{a}_s ds. \end{aligned}$$

Taking expectation, we see that

$$\mathbb{E} \left[\frac{\tilde{a}_\tau}{\tau + \epsilon} \right] = \mathbb{E} \left[\int_0^\tau \frac{1}{s + \epsilon} d\eta_s \right] - \mathbb{E} \left[\int_0^\tau \frac{1 + \epsilon}{(s + \epsilon)^2(1 - s)} \tilde{a}_s ds \right]. \tag{2.12}$$

The first term in (2.12) is an integral with respect to a martingale and hence a martingale. Also, τ is a stopping time. So, by the optional sampling theorem Klenke (2013), the expectation vanishes. Thus,

$$\mathbb{E} \left[\frac{\tilde{a}_\tau}{\tau + \epsilon} \right] = -\mathbb{E} \left[\int_0^\tau \frac{1 + \epsilon}{(s + \epsilon)^2(1 - s)} \tilde{a}_s ds \right].$$

From this, we have

$$\begin{aligned} \mathbb{E} \left[\int_0^\tau \frac{1 + \epsilon}{(s + \epsilon)^2(1 - s)} \tilde{a}_s ds \right] &= \mathbb{E} \left[\int_0^1 \frac{1 + \epsilon}{(s + \epsilon)^2(1 - s)} \tilde{a}_s \mathbf{1}_{\{s \leq \tau\}} ds \right], \\ &= \int_0^1 \frac{1 + \epsilon}{(s + \epsilon)^2(1 - s)} \mathbb{E} \left[\tilde{a}_s \mathbf{1}_{\{s \leq \tau\}} \right] ds, \\ &= \int_0^1 \frac{(1 + \epsilon) \mathbb{P}(\tau \leq s)}{(s + \epsilon)^2(1 - s)} \mathbb{E} [\tilde{a}_s | s \leq \tau] ds. \end{aligned}$$

However, $t \leq \tau$ means that $\tilde{a}_t \geq g(t)$ for $0 \leq s \leq t$. Moreover, for any $s, t \in [0, 1]$, $\mathbb{E}[\tilde{a}_s | \tau \geq t] \geq 0$. To see this, let p_1, \dots, p_{n_0} be the p -values corresponding to n_0 null hypothesis. Then, \tilde{a}_s and $\{\tau \geq t\}$ are decreasing random variables with respect to p_i s (by decreasing any p_i and keeping the others unchanged, the values of \tilde{a}_s and $\mathbf{1}_{\{\tau \geq t\}}$ increase). Since p_1, \dots, p_{n_0} are independent, any two decreasing random variables with respect to p_1, \dots, p_{n_0} are positively correlated. Hence,

$$\mathbb{E}[\tilde{a}_s \mathbf{1}_{\{\tau \geq t\}}] \geq \mathbb{E}[\tilde{a}_s] \mathbb{E}[\mathbf{1}_{\{\tau \geq t\}}].$$

Therefore,

$$\mathbb{E}[\tilde{a}_s | \tau \geq t] = \frac{\mathbb{E}[\tilde{a}_s \mathbf{1}_{\{\tau \geq t\}}]}{\mathbb{E}[\mathbf{1}_{\{\tau \geq t\}}]} \geq \mathbb{E}[\tilde{a}_s] \geq 0.$$

To complete the proof, it suffices to use Fatou's lemma:

$$\mathbb{E} \left[\frac{\tilde{a}_\tau}{\tau} \right] = \mathbb{E} \left[\liminf_{\epsilon \rightarrow 0} \frac{\tilde{a}_\tau}{\tau + \epsilon} \right] \leq \liminf_{\epsilon \rightarrow 0} \mathbb{E} \left[\frac{\tilde{a}_\tau}{\tau + \epsilon} \right] \leq 0.$$

□

3 Discussion

Storey's idea of using reverse filtration is a clever approach that makes the proofs of FDR control for the Benjamini-Hochberg (BH) step-up procedure and similar methods much more clear and elegant. Moreover, reverse filtration makes it possible to reject more hypotheses because after sorting p-values from smallest to largest, the stopping time with respect to reverse filtration would be the last p-value before crossing some threshold function. In contrast, the stopping time with respect to the forward filtration is the first p-value crossing the threshold function.

Despite clear advantages, proofs based on reverse filtration do not easily generalize to data-adaptive settings Li and Barber (2019); Ramdas et al. (2017), and multi-dimensional FDR control Ploner et al. (2006). Instead, FDR can be controlled in such settings by considering rejection regions of pre-defined shapes, or pre-determined weights for hypotheses. This strategy would allow us to use state-of-the-art developments for FDR control, including efficient step-up procedures, that can also accommodate dependence among hypotheses Romano et al. (2008); Blanchard and Roquain (2009). However, pre-specified regions or shapes may not efficiently capture non-null hypotheses in the presence of external structure. Therefore, this strategy can result in inferior power. As an alternative, recent procedures that assume that p-values, or test statistics, follow a mixture distribution Sun and Cai (2007) can be used for data-adaptive FDR control. This strategy can lead to 'optimal' procedures and can also handle dependence among hypotheses Cai and Sun (2009). However, these procedures only provide asymptotic FDR control, which may not be sufficient in small sample settings. The proof developed in this paper thus opens the door for new developments in FDR control for data-adaptive hypothesis testing procedures.

References

- Benditkis, J., Heesen, P., and Janssen, A. (2018), The false discovery rate (fdr) of multiple tests in a class room lecture. *Statistics & Probability Letters*, **134**, 29–35.
- Benditkis, J., and Janssen, A. (2017), Finite sample bounds for expected number of false rejections under martingale dependence with applications to FDR. *Electronic Journal of Statistics*, **11**(1), 1827–1857.
- Benjamini, Y., and Hochberg, Y. (1995), Controlling the false discovery rate: a practical

- and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B. Methodological*, **57**(1), 289–300.
- Benjamini, Y., and Liu, W. (1999), A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence. *Journal of Statistical Planning and Inference*, **82**(1), 163–170.
- Benjamini, Y., and Yekutieli, D. (2001), The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, **29**(4), 1165–1188.
- Blanchard, G., and Roquain, É. (2009), Adaptive false discovery rate control under independence and dependence. *Journal of Machine Learning Research*, **10**(Dec), 2837–2871.
- Cai, T. T., and Sun, W. (2009), Simultaneous testing of grouped hypotheses: Finding needles in multiple haystacks. *Journal of the American Statistical Association*, **104**(488), 1467–1481.
- Heesen, P., and Janssen, A. (2015), Inequalities for the false discovery rate (fdr) under dependence. *Electronic Journal of Statistics*, **9**(1), 679–716.
- Klenke, A. (2013), *Probability Theory: A Comprehensive Course*. Universitext. Springer London.
- Li, A., and Barber, R. F. (2019), Multiple testing with the structure-adaptive benjamini–hochberg algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **81**(1), 45–74.
- Peña, E. A., Habiger, J. D., and Wu, W. (2011), Power-enhanced multiple decision functions controlling family-wise error and false discovery rates. *Annals of statistics*, **39**(1), 556.
- Ploner, A., Calza, S., Gusnanto, A., and Pawitan, Y. (2006), Multidimensional local false discovery rate for microarray studies. *Bioinformatics*, **22**(5), 556–565.
- Ramdas, A., Chen, J., Wainwright, M. J., and Jordan, M. I. (2017), Dagger: A sequential algorithm for fdr control on dags. *arXiv preprint arXiv:1709.10250*.
- Romano, J. P., Shaikh, A. M., and Wolf, M. (2008), Control of the false discovery rate under dependence using the bootstrap and subsampling. *Test*, **17**(3), 417.

- Sarkar, S. K. (2002), Some results on false discovery rate in stepwise multiple testing procedures. *The Annals of Statistics*, **30**(1), 239–257.
- Storey, J. D., Taylor, J. E., and Siegmund, D. (2004), Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **66**(1), 187–205.
- Sun, W., and Cai, T. T. (2007), Oracle and adaptive compound decision rules for false discovery rate control. *Journal of the American Statistical Association*, **102**(479), 901–912.