

Feedback-Enhanced Online Multiple Testing with Applications to Conformal Selection

Lin Lu

LINLU102099@GMAIL.COM

*School of Statistics and Data Science, LPMC, KLMDASR and LEBPS
Nankai University
Tianjin, P.R. China*

Yuyang Huo

HUOYYNK@GMAIL.COM

*Department of Statistics and Finance, School of Management
University of Science and Technology of China
Hefei, P.R. China*

Haojie Ren

HAOJIEREN@SJTU.EDU.CN

*School of Mathematical Sciences
Shanghai Jiao Tong University
Shanghai, P.R. China*

Zhaojun Wang

ZJWANGNK@126.COM

*School of Statistics and Data Science, LPMC, KLMDASR and LEBPS
Nankai University
Tianjin, P.R. China*

Changliang Zou

NK.CHLZOU@GMAIL.COM

*School of Statistics and Data Science, LPMC, KLMDASR and LEBPS
Nankai University
Tianjin, P.R. China*

Editor: Chris Oates

Abstract

This work studies online multiple testing with feedback, where decisions are made sequentially, and the true state of the hypothesis is revealed after decisions are made, either instantly or with a delay, and under either full or bandit feedback. We propose Generalized alpha-investing with feedback (GAIF) along with its adaptive variants, a feedback-enhanced framework that dynamically adjusts thresholds using revealed outcomes, ensuring finite-sample false discovery rate (FDR)/marginal FDR (mFDR) control. We further extend GAIF to online conformal testing by constructing valid conformal p -values and developing feedback-enhanced testing rules with finite-sample mFDR control. We also propose a feedback-driven score selection criterion to adaptively choose the candidate score that is most effective for the testing procedure, together with a theoretical analysis of its optimality. Numerical simulations and real-data applications demonstrate the effectiveness of our methods.

Keywords: Conformal prediction, distribution shifts, generalized alpha-investing procedure, model selection, online conformal p -value, online FDR control

1. Introduction

Real-time decision making plays a critical role in a growing number of modern applications, such as online recruitment for job hiring (Faliagka et al., 2014), real-time alignment of large language models (Huang et al., 2025), and time-series anomaly detection (Rebjock et al., 2021). These tasks can be naturally formulated as online multiple testing problems (Foster and Stine, 2008). Consider a potentially infinite stream of null hypotheses $\{\mathbb{H}_{01}, \mathbb{H}_{02}, \dots, \mathbb{H}_{0t}, \dots\}$ that are tested sequentially based on incoming test statistics such as p -values $\{p_1, p_2, \dots, p_t, \dots\}$. At each time t , a real-time decision must be made based on the current test statistic. Let θ_t denote the true state of the hypothesis at time t , where $\theta_t = 0$ if $\mathbb{H}_{0t}$ is true and $\theta_t = 1$ otherwise. The testing problem at time t is

$$\mathbb{H}_{0t} : \theta_t = 0 \quad \text{vs.} \quad \mathbb{H}_{1t} : \theta_t = 1.$$

Let $\delta_t \in \{0, 1\}$ be the decision at time t , with $\delta_t = 1$ indicating rejection of $\mathbb{H}_{0t}$. To ensure the reliability of the testing procedure, it is essential to control a well-defined error rate. Define the false discovery proportion (FDP) and false discovery rate (FDR) (Benjamini and Hochberg, 1995) at time t by

$$\text{FDP}(t) = \frac{V(t)}{1 \vee R(t)} := \frac{\sum_{j=1}^t \delta_j (1 - \theta_j)}{1 \vee \sum_{j=1}^t \delta_j}, \quad \text{FDR}(t) = \mathbb{E} \{ \text{FDP}(t) \},$$

where $V(t)$ and $R(t)$ represent the numbers of false rejections (discoveries) and rejections at time t , respectively. Given a target level $\alpha \in (0, 1)$, classical work on online multiple testing (Ramdas et al., 2017, 2018; Tian and Ramdas, 2019) aims to guarantee $\sup_t \text{FDR}(t) \leq \alpha$ or its weaker variant (Foster and Stine, 2008; Zrnic et al., 2021), the marginal FDR (mFDR), i.e., $\text{mFDR}(t) = \mathbb{E}\{V(t)\}/\mathbb{E}\{1 \vee R(t)\}$.

Unlike the classical setting, we consider another realistic scenario in which feedback on past hypotheses is revealed after decisions are made. Specifically, after issuing δ_t , the corresponding state θ_t may be observed either immediately or after a delay, and in some settings it may be observed only partially. For instance, in the bandit setting (Neu et al., 2010), feedback θ_t is observed only when the decision $\delta_t = 1$ is made. Our work encompasses all these feedback-available settings, including full or partial feedback, as well as immediate or delayed feedback.

Such feedback-available settings naturally arise in many practical applications. To demonstrate its relevance, we present three motivating examples: online conformal selection, real-time alignment of large language models, and time-series anomaly detection.

- **Online conformal selection.** Conformal selection aims to identify valuable individuals whose unknown label Y satisfies a pre-specified requirement by leveraging machine learning predictions (Wu et al., 2024; Jin and Candès, 2025). The online conformal selection setting naturally aligns with the feedback-available online multiple testing framework. For example, in diabetes risk prediction, a patient's true condition may later be confirmed by an expert, offering feedback that could improve future decisions. Some works (Huo et al., 2024; Xu and Ramdas, 2024) considered online conformal testing problems but did not exploit feedback information.

- **Real-time LLM alignment.** Large language models (LLMs) are increasingly used in high-stakes domains such as healthcare, finance, and law, where outputs must be reliable and factual. However, LLMs can hallucinate—producing plausible but incorrect content (Huang et al., 2025). A remedy is to filter or certify LLM outputs (Gui et al., 2024; Bai and Jin, 2024). Some alignment approaches based on conformal testing are proposed for this use but are generally designed for offline environments, whereas many applications demand immediate and trustworthy screening. In such applications, follow-up feedback is usually available, which, if incorporated, could continuously improve alignment.
- **Time series anomaly detection.** Detecting anomalies in time series is crucial for industrial monitoring, fraud detection, and healthcare analytics. To ensure reliability, prior work (Rebjock et al., 2021; Gang et al., 2023; Krönert et al., 2023) addresses online FDR control but typically ignores real-time feedback. In practice, such feedback is often available: once an anomaly is flagged, a subsequent system failure, user verification of a fraudulent transaction, or expert annotation may confirm or refute the alarm.

These observations underscore that feedback is a foundational element of adaptive decision-making, playing a pivotal role across a wide range of online applications. Despite its importance, current online multiple testing procedures seldom incorporate feedback in a systematic way. This naturally raises a fundamental and compelling question: *Can available feedback be effectively incorporated into online multiple testing procedures in a way that allows us to enhance statistical power while still ensuring valid error rate control?*

To this end, we develop a feedback-enhanced framework that systematically integrates revealed feedback into the generalized alpha-investing (GAI) framework (Aharoni and Rosset, 2014; Javanmard and Montanari, 2018), achieving significant performance gains without compromising statistical validity.

1.1 Our contributions

To the best of our knowledge, this is the first work to incorporate feedback information directly into the construction of testing thresholds for online FDR procedures and to employ this idea to online conformal testing. Our main contributions are summarized as follows:

- **GAIF and its variants.** We introduce Generalized Alpha-Investing with Feedback (GAIF), a feedback-enhanced extension of the GAI framework, together with its adaptive variants. The key idea is to use revealed feedback to sharpen the FDP estimator by discounting confirmed non-null hypotheses, and, in the adaptive version, to recycle the α -wealth that would otherwise be permanently lost whenever a non-null hypothesis yields a large p -value. The framework applies to a broad range of feedback settings, including full or partial feedback, as well as immediate or delayed feedback, and it also extends to locally dependent scenarios. We establish online FDR control under independence and online mFDR control under local dependence for GAIF and its variants.
- **Online conformal testing with feedback.** We extend the GAIF to *online conformal testing*. We first provide an explicit construction of valid, independent p -values (for null

hypotheses) by dynamically updating the calibration dataset. We then combine them with suitably modified GAIF rules to obtain *online conformal testing with feedback* (OCTF), which enjoys finite-sample online mFDR control. This extension bridges the gap between traditional online multiple testing and conformal inference, yielding distribution-free, model-agnostic tools for real-time decision-making.

- **Feedback-driven score selection.** We further introduce a strategy to adaptively select among candidate predictive models or conformity scores. This is particularly important under non-stationary non-null distributions, where the optimal score may vary over time. By locally averaging recent non-null p -values to track the best-performing score, our selection criterion can accommodate such shifts while maintaining validity.

We provide rigorous proofs for all proposed methods. Extensive simulations and real-data experiments demonstrate that our procedures substantially outperform existing approaches while effectively controlling the online FDR when feedback information is provided.

1.2 Related works

Our work is situated at the intersection of online multiple testing and conformal inference. We review key developments in each area and highlight gaps that motivate our contribution.

Online multiple testing under independence. Early works on online multiple testing began with the alpha-investing strategy of Foster and Stine (2008), later generalized by Aharoni and Rosset (2014) and Javanmard and Montanari (2018) into the GAI framework, which led to the LORD algorithm. Building on this line of work, Ramdas et al. (2017) introduced LORD++, an improved version of GAI tailored for online FDR control. Subsequent refinements include SAFFRON (Ramdas et al., 2018), which adapts to the proportion of non-nulls, and ADDIS (Tian and Ramdas, 2019), which adjusts for conservative null p -values. These methods, including LORD++, SAFFRON, and ADDIS, achieve online FDR control when null p -values are independent of all other p -values. Separately, Gang et al. (2023) developed structure-adaptive rules based on local FDR, which improve power but only ensure asymptotic FDR control under correct model specification. For a comprehensive review, see Robertson et al. (2023). All of these methods determine thresholds solely from past rejections, without considering real-time feedback.

Online multiple testing under dependence. In practice, hypotheses often exhibit dependence, and applying methods designed for independence can lead to inflated error rates. To address arbitrary dependence, Xu and Ramdas (2024) proposed e -LOND, an FDR-controlling procedure based on e -values. Zhang et al. (2025) extended this approach to e -GAI, achieving improved power by dynamically allocating the testing levels. Alternatively, research has focused on local dependence structures. Zrnic et al. (2021) introduced LORD_{dep} and SAFFRON_{dep}, establishing mFDR control under local dependence; Rebjock et al. (2021) later adapted these methods to time-series anomaly detection. Recently, Fisher (2024) showed that LORD++ with suitable local modifications can maintain FDR control under certain dependence, while Fischer et al. (2024) proposed an online procedure for p -values satisfying positive regression dependence on a subset (PRDS). Despite these advances, existing dependence-aware methods do not incorporate any real-time feedback.

Conformal inference and conformal multiple testing. Conformal inference (Vovk et al., 2005) offers a model-agnostic way to quantify prediction uncertainty. In the multiple testing setting, early works constructed conformal p -values and applied the Benjamini–Hochberg (BH) procedure (Benjamini and Hochberg, 1995) to achieve finite-sample FDR control (Bates et al., 2023; Jin and Candès, 2025). Subsequent extensions addressed covariate shift (Jin and Candès, 2023), constrained selection (Wu et al., 2024; Nair et al., 2025), and conditional testing (Wu et al., 2025), as well as model selection (Bai and Jin, 2024; Gui et al., 2025). However, these contributions remain confined to offline settings. The few existing efforts to extend conformal multiple testing to the online domain (Huo et al., 2024; Xu and Ramdas, 2024; Liu et al., 2026) have not taken feedback information into account. Although related research on the construction of online conformal prediction sets (Gibbs and Candès, 2021, 2024; Xi et al., 2025) has considered feedback, multiple testing problems with FDR control remain largely unexplored. One very recent work, by Humbert et al. (2025), establishes asymptotic online FDP control through an online learning strategy, while we achieve feedback-enhanced testing based on the GAI framework.

1.3 Organization of the paper

The remainder of this paper is organized as follows. Section 2 introduces the GAIF and establishes FDR control under independence. Section 3 extends the GAIF to achieve adaptive α -wealth allocation and to address local dependence. In Section 4, we construct explicit online conformal p -values and apply modified GAIF rules to online conformal testing, providing finite-sample theoretical guarantees. We also extend the proposed framework to achieve score selection and address distribution shifts. Simulation and real-data experiment results are presented in Sections 5 and 6, respectively. Finally, we conclude the paper in Section 7.

2. Generalized Alpha-Investing with Feedback (GAIF)

In this section, we begin by revisiting the generalized alpha-investing (GAI) framework in Section 2.1. In Section 2.2, we introduce our feedback-enhanced framework, termed generalized alpha-investing with feedback (GAIF), which incorporates feedback information to improve FDP estimation. In Section 2.3, we establish finite-sample online FDR control for GAIF procedures.

2.1 Recap: Generalized alpha-investing

Early approaches to online error-rate control were based on so-called *alpha-investing* strategies (Foster and Stine, 2008), which were later generalized into the GAI framework (Aharoni and Rosset, 2014; Javanmard and Montanari, 2018). These procedures are based on the principle of *alpha-wealth*: the testing process starts with an initial wealth, spends wealth to conduct hypothesis tests, and earns wealth whenever a rejection is made, thereby generating resources for future tests. Accordingly, each incoming p -value p_t is compared against a dynamically chosen threshold α_t , and the testing decision is given by $\delta_t = \mathbb{I}\{p_t \leq \alpha_t\}$.

Based on GAI, Ramdas et al. (2017) proposed controlling online FDR by ensuring that an estimator of the FDP remains below a pre-specified level α . A specific example is the

LORD++ algorithm, where the estimated FDP at time t is given by

$$\widehat{\text{FDP}}_{\text{LORD}}(t) = \frac{\sum_{j \leq t} \alpha_j}{1 \vee \sum_{j \leq t} \delta_j} \stackrel{(i)}{\geq} \text{FDP}^*(t) = \frac{\sum_{j \leq t, j \in \mathcal{H}_0} \alpha_j}{1 \vee \sum_{j \leq t} \delta_j} \approx \text{FDP}(t) := \frac{\sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{I}\{p_j \leq \alpha_j\}}{1 \vee \sum_{j \leq t} \delta_j}, \quad (1)$$

where $\mathcal{H}_0$ is the index set of null hypotheses. To ensure that $\text{FDR}(t) \leq \alpha$, it suffices to enforce $\widehat{\text{FDP}}_{\text{LORD}}(t) = \frac{\sum_{j \leq t} \alpha_j}{R(t) \vee 1} \leq \alpha$. The corresponding testing level of LORD++ is

$$\alpha_t^{\text{LORD++}} = \gamma_t s_0 + (\alpha - s_0) \gamma_{t-\tau_1} \mathbb{I}\{\tau_1 < t\} + \alpha \sum_{j: \tau_j < t, \tau_j \neq \tau_1} \gamma_{t-\tau_j},$$

where $\{\gamma_t\}_{t=1}^\infty$ is a given infinite non-increasing sequence of positive constants that sums to one, τ_j is the time of the j -th rejection, and $s_0 > 0$ is the pre-specified initial wealth.

Building on this, some adaptive versions of LORD++ were proposed subsequently, such as SAFFRON (Ramdas et al., 2018) and ADDIS (Tian and Ramdas, 2019). Specifically, SAFFRON introduces a user-specified parameter $\lambda \in [0, 1]$ to estimate the proportion of true nulls in the online testing stream, and obtains a less conservative estimated FDP as

$$\widehat{\text{FDP}}_{\text{SAFFRON}}(t) = \frac{\sum_{j \leq t} \alpha_j \frac{\mathbb{I}\{p_j > \lambda\}}{(1-\lambda)}}{1 \vee \sum_{j \leq t} \delta_j}. \quad (2)$$

These GAI-based procedures guarantee online mFDR control under the *conditional super-uniformity* assumption for null p -values:

$$\Pr(p_t \leq \alpha_t \mid \mathcal{G}_{t-1}) \leq \alpha_t \quad \text{for all } t \in \mathcal{H}_0, \quad (3)$$

where $\mathcal{G}_t = \sigma(\delta_{1:t})$ and each threshold $\alpha_t = f_t(\delta_{1:t-1})$ is measurable with respect to past decisions. Here $\delta_{1:t} := (\delta_1, \dots, \delta_t)$. Furthermore, if all null p -values are independent of other p -values, these procedures also ensure strict FDR control.

Note that the inequality (i) in (1) is tight if and only if all past hypotheses are null. This conservative step is one of the major sources of the gap between the realized FDR of LORD++ and the target level α . When feedback becomes available over time, it allows us to sharpen the FDP estimator and accordingly reduce the resulting slack. Intuitively, once the j th test is confirmed non-null, its contribution α_j can be removed from the FDP bound and reallocated. Our proposed GAIF and its adaptive variant extend LORD++ and SAFFRON by systematically recycling such recovered wealth and reinvesting it into future tests.

2.2 Boosting GAI via feedback: GAIF

We now propose generalized alpha-investing with feedback (GAIF), which leverages revealed feedback information to improve FDP estimation. The key idea is that once some past hypotheses are known to be non-nulls, their contribution to the FDP upper bound can be reduced accordingly, leaving more α -wealth available for future tests.

Formally, let

$$\mathcal{I}_t = \{j \in [t] : \theta_j \text{ is revealed by time } t\}$$

denote the index set of hypotheses whose labels are available by time t , and let $\bar{\mathcal{I}}_t = [t] \setminus \mathcal{I}_t$, where $[t] = \{1, 2, \dots, t\}$. Realistic scenarios can lead to diverse feedback structures. For

example, feedback may be available only when a rejection is made or may arrive after a delay $d \geq 0$. Our framework encompasses a variety of practically relevant cases:

- Full and instant feedback: $\mathcal{I}_t = \{j \in [t] : j \leq t - 1\}$;
- Bandit and instant feedback: $\mathcal{I}_t = \{j \in [t] : j \leq t - 1, \delta_j = 1\}$;
- Full and delayed feedback: $\mathcal{I}_t = \{j \in [t] : j \leq t - d - 1\}$, where d is the delay;
- Bandit and delayed feedback: $\mathcal{I}_t = \{j \in [t] : j \leq t - d - 1, \delta_j = 1\}$.

At time t , the available feedback reveals the true values of $\{\theta_j\}_{j \in \mathcal{I}_t}$. We leverage this information to refine the FDP estimator as follows.

Definition 1 (GAIF). *Define the feedback-enhanced FDP estimator as*

$$\widehat{\text{FDP}}_{\text{GAIF}}(t) := \frac{\sum_{j \in \mathcal{I}_t} (1 - \theta_j) \alpha_j + \sum_{j \in \bar{\mathcal{I}}_t} \alpha_j}{1 \vee \sum_{j=1}^t \delta_j}.$$

A procedure is said to be a GAI with Feedback (GAIF) procedure if it assigns test levels $\{\alpha_t\}$ such that, for every t ,

$$\widehat{\text{FDP}}_{\text{GAIF}}(t) \leq \alpha.$$

The GAIF framework aligns with intuition. For $j \in \mathcal{I}_t$, feedback reveals whether the corresponding hypothesis is null or non-null, so their contribution to the FDP bound can be accounted for precisely through $(1 - \theta_j) \alpha_j$. For indices not yet revealed, i.e., $j \in \bar{\mathcal{I}}_t$, we keep the contribution α_j . Thus, GAIF improves the FDP estimator relative to the no-feedback case.

Importantly, this refinement leads directly to a less conservative constraint on the test levels α_t . Recall that LORD++ enforces $\alpha_t \leq \alpha(1 \vee \sum_{j \leq t} \delta_j) - \sum_{j \leq t-1} \alpha_j$. In contrast, a GAIF procedure selects α_t to satisfy

$$\alpha_t \leq \alpha(1 \vee \sum_{j=1}^t \delta_j) - \sum_{j \in \mathcal{I}_t} \alpha_j(1 - \theta_j) - \sum_{j \in \bar{\mathcal{I}}_t \setminus \{t\}} \alpha_j. \quad (4)$$

The presence of revealed feedback enlarges the admissible upper bound for α_t . Indeed, since $1 - \theta_j \leq 1$ for every $j \in \mathcal{I}_t$, (4) implies

$$\sum_{j \in \mathcal{I}_t} \alpha_j(1 - \theta_j) + \sum_{j \in \bar{\mathcal{I}}_t \setminus \{t\}} \alpha_j \leq \sum_{j \in \mathcal{I}_t} \alpha_j + \sum_{j \in \bar{\mathcal{I}}_t \setminus \{t\}} \alpha_j = \sum_{j=1}^{t-1} \alpha_j.$$

Therefore, the upper bound of α_t under GAIF is always at least as large as that under LORD++. In this sense, an appropriate procedure with feedback is expected to improve power by sharpening the FDP estimator and releasing more α -wealth for future allocation.

We next present a concrete instance of GAIF as follows:

$$\alpha_t^{\text{GAIF}} = \gamma_t s_0 + (\alpha - s_0) \gamma_{t-\tau_1} \mathbb{I}\{\tau_1 < t\} + \alpha \sum_{j: \tau_j < t, \tau_j \neq \tau_1} \gamma_{t-\tau_j} + \sum_{j: j \in \mathcal{I}_t} \gamma_{t-j} \alpha_j \theta_j. \quad (5)$$

When no feedback is observed, i.e., $\mathcal{I}_t = \emptyset$, the above reduces exactly to the LORD++ update. Consequently, the test levels α_t^{GAIF} are never smaller than those of LORD++ whenever feedback is available. The last term represents the additional α -wealth recovered from revealed non-nulls and reallocated to future tests.

The following result states that the α_t^{GAIF} construction in Eq. (5) satisfies Definition 1.

Proposition 2. *Assume the testing levels α_t^{GAIF} are generated by Eq. (5), the sequence $\{\gamma_j\}_{j=1}^\infty$ is non-negative and satisfies $\sum_{j=1}^\infty \gamma_j \leq 1$, the initial wealth parameter satisfies $s_0 \leq \alpha$. Then, the following is satisfied:*

$$\widehat{\text{FDP}}_{\text{GAIF}}(t) \leq \alpha.$$

Algorithm 1 summarizes the implementation of this concrete GAIF procedure under a general feedback regime. The only difference across settings lies in how the feedback-available set $\mathcal{I}_t$ is updated. We denote the resulting procedures under full and instant feedback, bandit and instant feedback, full and delayed feedback, and bandit and delayed feedback by LF, LF-BI, LF-FD, and LF-BD, respectively.

Algorithm 1 GAIF procedures

Input: Target FDR level α ; pre-specified parameters for constructing test levels.

- 1: **for** $t = 1, 2, \dots$ **do**
- 2: Observe the p -value p_t .
- 3: Update $\mathcal{I}_t$ according to the specific feedback regime.
- 4: Construct $\alpha_t = \alpha_t^{\text{GAIF}}$ using Definition 1 with the current $\mathcal{I}_t$.
- 5: **if** $p_t \leq \alpha_t$ **then** set $\delta_t = 1$; **else** set $\delta_t = 0$.
- 6: Reveal feedback by obtaining θ_j for all $j \in \mathcal{I}_{t+1} \setminus \mathcal{I}_t$.
- 7: **end for**

Output: Rejection set $\mathcal{R} = \{t : \delta_t = 1\}$.

2.3 Theoretical guarantee under independence

The following result states that GAIF procedures guarantee online mFDR control under conditional super-uniformity of null p -values

$$\Pr(p_t \leq \alpha_t \mid \mathcal{F}_{t-1}) \leq \alpha_t \quad \text{for all } t \in \mathcal{H}_0, \quad (6)$$

where the enlarged filtration $\mathcal{F}_{t-1} := \sigma(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})$. Under stronger independence and monotonicity assumptions, GAIF further achieves online FDR control. Here, the test level sequence $\{\alpha_t\}_{t \in \mathbb{N}}$ is said to be a monotonic function of the past if, for all $t \in \mathbb{N}$, α_t is coordinate-wise non-decreasing in the past decisions $\{\delta_j : j < t\}$ and revealed feedback $\{\theta_j : j \in \mathcal{I}_t\}$ for the GAIF procedure.

Theorem 3 (Online mFDR and FDR control for GAIF). *Let $\{\alpha_t\}_{t \in \mathbb{N}}$ be a sequence of test levels satisfying $\widehat{\text{FDP}}_{\text{GAIF}}(t) \leq \alpha$. Then:*

- (a) *If the null p -values are conditionally super-uniformly distributed as in (6), then the procedure guarantees*

$$\text{mFDR}(t) \leq \alpha \quad \text{for all } t \in \mathbb{N}.$$

- (b) *If the null p -values are mutually independent and also independent of the non-null p -values, and if the test levels $\{\alpha_t\}$ are a monotone sequence of functions of the past for all t , then the procedure satisfies*

$$\text{FDR}(t) \leq \alpha \quad \text{for all } t \in \mathbb{N}.$$

The test levels of GAIF in (5) inherit the monotonicity structure from LORD++, and therefore satisfy the monotonicity condition needed for Theorem 3 (b). The proof of Theorem 3 is provided in Appendix F.2.

3. Extensions of GAIF

In this section, we discuss two important extensions of GAIF. We first develop the Adaptive GAIF framework in Section 3.1, which further enhances GAIF by enabling adaptive α -wealth allocation. We then extend our methods to the settings with local dependence in Section 3.2.

3.1 Adaptive GAIF: improve GAIF via adaptive α -wealth allocation

Section 2.2 shows that feedback improves power by sharpening the FDP estimate. In online testing, power also depends on how the available α -wealth is allocated over time, since the test levels α_t must be chosen sequentially. This motivates an improvement direction beyond feedback-enhanced FDP estimation, namely adaptive α -wealth allocation.

To motivate this idea, we revisit the difference between LORD++ and SAFFRON. The SAFFRON weighting function $\kappa(p) := \frac{\mathbb{I}\{p > \lambda\}}{1 - \lambda}$ has a well-known interpretation as a null proportion estimator (Storey et al., 2004; Döhler and Meah, 2023). Here we highlight a complementary perspective: $\kappa(\cdot)$ also functions as an α -wealth allocation rule, concentrating testing budget on hypotheses with small p -values. Indeed, under LORD++, the test levels satisfy $\alpha_t \leq \alpha \sum_{j=1}^t \delta_j - \sum_{j=1}^{t-1} \alpha_j$. By contrast, SAFFRON enforces

$$\alpha_t \leq \alpha \frac{1 - \lambda}{\mathbb{I}\{p_t > \lambda\}} \sum_{j=1}^t \delta_j - \frac{1}{\mathbb{I}\{p_t > \lambda\}} \sum_{j=1}^{t-1} \alpha_j \mathbb{I}\{p_j > \lambda\}. \quad (7)$$

When $p_t \leq \lambda$, the upper bound is interpreted as $+\infty$, permitting aggressive investment at promising tests; when $p_t > \lambda$ and $p_j > \lambda$ for all $j \leq t - 1$, the bound is strictly smaller than the LORD++ bound, ensuring less wealth is spent on unpromising tests. This observation clarifies the distinct roles of FDP estimation and α -wealth allocation in improving power, thereby helping to explain how GAIF and SAFFRON achieve power gains over LORD++ through different mechanisms.

However, SAFFRON may allocate α -wealth inefficiently when a non-null hypothesis yields a large p -value. Specifically, when $p_j > \lambda$, the hypothesis contributes to the FDP estimator through $\kappa(p_j) = 1/(1 - \lambda)$. In the absence of feedback, this conservative penalty is unavoidable. Once feedback reveals $\theta_j = 1$, we know retrospectively that this test carried no false discovery risk. The wealth α_j could have been reallocated to future tests. Adaptive GAIF is designed to address this issue. It retains the feedback-based refinement of GAIF while incorporating the SAFFRON-style allocation rule through $\kappa(\cdot)$.

Definition 4 (Adaptive GAIF). Define the adaptive feedback-enhanced FDP estimator as

$$\widehat{\text{FDP}}_{\text{Ada-GAIF}}(t) := \frac{\sum_{j \in \mathcal{I}_t} (1 - \theta_j) \alpha_j \kappa(p_j) + \sum_{j \in \bar{\mathcal{I}}_t} \alpha_j \kappa(p_j)}{1 \vee \sum_{j=1}^t \delta_j},$$

where $\kappa(p) = \frac{\mathbb{I}\{p > \lambda\}}{1 - \lambda}$ and $\lambda \in [0, 1)$ is a user-chosen parameter for identifying large p -values. A procedure is called an Adaptive GAIF procedure if it assigns test levels $\{\alpha_t\}$ such that, for every t ,

$$\widehat{\text{FDP}}_{\text{Ada-GAIF}}(t) \leq \alpha. \quad (8)$$

The key difference between GAIF and Adaptive GAIF is the introduction of the $\kappa(p)$ in the FDP estimate. If feedback is ignored so that $\mathcal{I}_t = \emptyset$, the FDP estimator of Adaptive GAIF reduces to that of SAFFRON.

Remark 5. In principle, other constructions of $\kappa(\cdot)$ may further optimize α -wealth allocation while ensuring online FDR control. This online optimization remains an open problem and a direction for future work (Gang et al., 2023). Instead, we focus on the SAFFRON-type weight $\frac{\mathbb{I}\{p > \lambda\}}{1 - \lambda}$, which offers a theoretically grounded and practically appealing compromise with robust empirical performance.

A concrete instance of Adaptive GAIF follows the same structure as SAFFRON, augmented with the feedback correction term. For $t = 1$, $\alpha_1^{\text{Ada-GAIF}} = \min\{\gamma_1 s_0(1 - \lambda), \lambda\}$. For $t > 1$,

$$\begin{aligned} \alpha_t^{\text{Ada-GAIF}} = \min \Bigg\{ & \lambda, (1 - \lambda) \left[s_0 \gamma_{t-C_{0+}} + (\alpha - s_0) \gamma_{t-\tau_1-C_{1+}} + \alpha \sum_{j \geq 2} \gamma_{t-\tau_j-C_{j+}} \right] \\ & + \sum_{j: j \in \mathcal{I}_t} \gamma_{t-j-C_{j+}^*} \alpha_j \theta_j \mathbb{I}\{p_j > \lambda\} \Bigg\}, \end{aligned} \quad (9)$$

where $C_{j+} = C_{j+}(t) = \sum_{i=\tau_j+1}^{t-1} \mathbb{I}\{p_i \leq \lambda\}$ and $C_{j+}^* = C_{j+}^*(t) = \sum_{i=j+1}^{t-1} \mathbb{I}\{p_i \leq \lambda\}$. The first term coincides with the standard SAFFRON update; the second term recycles the wealth recovered through feedback. We denote the resulting procedures under full and instant, bandit and instant, full and delayed, and bandit and delayed feedback by SF, SF-BI, SF-FD, and SF-BD, respectively.

The Proposition 6 states that the $\alpha_t^{\text{Ada-GAIF}}$ construction in Eq. (9) satisfies Definition 4.

Proposition 6. Assume the testing levels $\alpha_t^{\text{Ada-GAIF}}$ are generated by Eq. (9), the sequence $\{\gamma_j\}_{j=1}^\infty$ is non-negative and satisfies $\sum_{j=1}^\infty \gamma_j \leq 1$, and the initial wealth parameter satisfies $s_0 \leq \alpha$. Then, the following is satisfied:

$$\widehat{\text{FDP}}_{\text{Ada-GAIF}}(t) \leq \alpha.$$

Figure 1 depicts the average testing thresholds $\{\alpha_t\}$ over time for LF, SF, LORD++, SAFFRON and LOND (Javanmard and Montanari, 2015) applied to Gaussian observations under full and instant feedback. Across all non-null proportions π_1 , both LF and SF yield strictly larger thresholds than their feedback-free counterparts, confirming that feedback

enables more effective α -wealth utilization. When π_1 is small, revealed non-nulls are rare and feedback-recovered wealth is limited; adaptive α -wealth allocation then dominates, and SF outperforms LF—at $\pi_1 = 0.2$. As π_1 increases, feedback becomes more informative and wealth recycling grows, allowing LF to surpass SF.

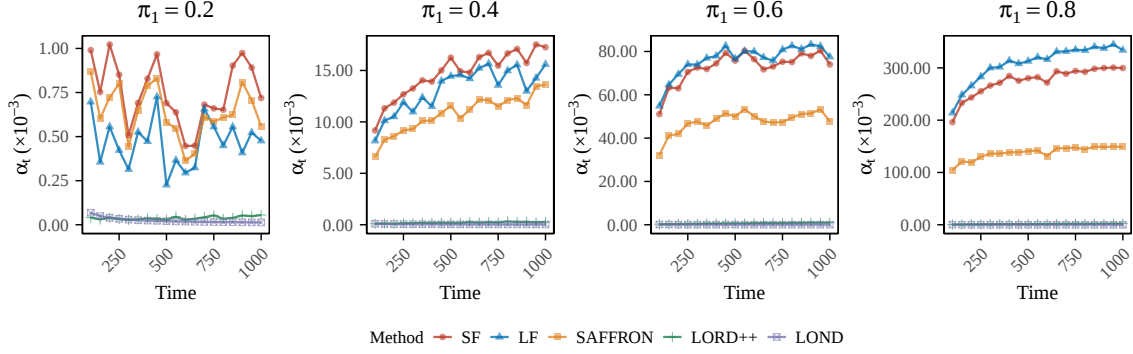


Figure 1: Average testing thresholds α_t over time under various procedures, based on 500 replications across non-null proportions $\pi_1 \in \{0.2, 0.4, 0.6, 0.8\}$. Data are generated under Scenario I of Section 5.1 with full and instant feedback and signal strength $\mu = 2$.

We now establish the theoretical results for Adaptive GAIF. Similar to (6), the conditional super-uniformity condition on the null p -values is

$$\Pr(p_t \leq \alpha_t \mid \mathcal{J}_{t-1}) \leq \alpha_t \quad \text{for all } t \in \mathcal{H}_0, \quad (10)$$

where $\mathcal{J}_{t-1} := \sigma(\delta_{1:t-1}; C_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})$ is the enlarged filtration, where $C_j = \mathbb{I}\{p_j \leq \lambda\}$. And we extend the monotonicity definition in Theorem 3 by additionally requiring α_t to be coordinate-wise non-decreasing in the indicators $\{C_j : j < t\}$.

Theorem 7 (Online mFDR and FDR control for Adaptive GAIF). *Let $\{\alpha_t\}_{t \in \mathbb{N}}$ be a sequence of test levels satisfying $\widehat{\text{FDP}}_{\text{Ada-GAIF}}(t) \leq \alpha$. Then:*

- (a) *If the null p -values are conditionally super-uniform (10), the procedure guarantees*

$$\text{mFDR}(t) \leq \alpha \quad \text{for all } t \in \mathbb{N}.$$

- (b) *If the null p -values are mutually independent and independent of the non-null p -values, and if the test levels $\{\alpha_t\}$ form a monotone sequence of functions of the past for all t , then the procedure satisfies*

$$\text{FDR}(t) \leq \alpha \quad \text{for all } t \in \mathbb{N}.$$

The test levels of Adaptive GAIF in (9) inherit the monotonicity structure from SAFFRON, and therefore satisfy the monotonicity condition needed for Theorem 7 (b). The proof of Theorem 7 is provided in Appendix F.3.

3.2 GAIF under local dependence

This subsection focuses on extending the GAIF and Adaptive GAIF framework to settings with local dependence (as defined in Definition 8). This extension is useful in applications where nearby tests may be dependent, while long-range dependence is absent or weak. For a strategy that addresses more general dependence structures via e-values, the reader may refer to Appendix B.2.

Definition 8 (Local dependence; Zrnic et al. (2021)). *We say that p -values $p_1, p_2, \dots, p_t, \dots$ are locally dependent if*

$$\text{for all } t > 0, \text{ there exists } L_t \in \mathbb{N} \text{ such that } p_t \perp p_{t-L_t-1}, p_{t-L_t-2}, \dots, p_1, \quad (11)$$

where $\{L_t\}_{t \in \mathbb{N}}$ is a fixed sequence of parameters which we refer to as lags.

Zrnic et al. (2021) pioneered a strategy to handle this local dependence in online multiple testing, leading to the development of the LORD_{dep} and $\text{SAFFRON}_{\text{dep}}$ methods. We extend the idea to the feedback-enhanced setting. Under local dependence, the dependence-adjusted FDP estimates for GAIF and Adaptive GAIF are, respectively:

$$\widehat{\text{FDP}}_{\text{GAIF}_{\text{dep}}}(t) = \frac{\sum_{j \in \mathcal{I}_t} (1 - \theta_j) \alpha_j + \sum_{j \in \bar{\mathcal{I}}_t} \alpha_j}{(\sum_{j \leq t, j \notin \{t-L_t, \dots, t-1\}} \delta_j) \vee 1},$$

and

$$\widehat{\text{FDP}}_{\text{Ada-GAIF}_{\text{dep}}}(t) = \frac{\sum_{j < t-L_t} \frac{\alpha_j \mathbb{I}\{p_j > \lambda\}}{1-\lambda} \mathcal{W}_j + \sum_{t-L_t \leq j \leq t} \frac{\alpha_j}{1-\lambda} \mathcal{W}_j}{(\sum_{j \leq t, j \notin \{t-L_t, \dots, t-1\}} \delta_j) \vee 1},$$

where $\mathcal{W}_j = (1 - \theta_j) \mathbb{I}\{j \in \mathcal{I}_t\} + \mathbb{I}\{j \in \bar{\mathcal{I}}_t\}$. The corresponding testing levels are chosen so that $\widehat{\text{FDP}}_{\text{GAIF}_{\text{dep}}}(t) \leq \alpha$ or $\widehat{\text{FDP}}_{\text{Ada-GAIF}_{\text{dep}}}(t) \leq \alpha$. The resulting testing levels are provided in Appendix A.2. Theorem 9 shows that GAIF_{dep} and $\text{Ada-GAIF}_{\text{dep}}$ methods control mFDR under local dependence.

Theorem 9 (Online mFDR control under local dependence). *Suppose that the null p -values are locally dependent as defined in (11). Then if the test levels $\{\alpha_t\}_{t \in \mathbb{N}}$ satisfy $\widehat{\text{FDP}}_{\text{GAIF}_{\text{dep}}}(t) \leq \alpha$ or $\widehat{\text{FDP}}_{\text{Ada-GAIF}_{\text{dep}}}(t) \leq \alpha$, we have $\text{mFDR}(t) \leq \alpha$ for all $t \in \mathbb{N}$.*

4. Applications of GAIF on Online Conformal Selection

Online conformal selection (Huo et al., 2024; Xu and Ramdas, 2024) provides a canonical and intuitive instantiation of the GAIF framework. Here, decisions must be made in real time about whether an incoming observation satisfies a pre-specified requirement, while the true label is typically observed after the decision and thus serves as feedback. Throughout this section, we focus on the full and instant feedback setting.

Consider a data pair $(\mathbf{X}, Y) \in \mathcal{X} \times \mathcal{Y}$, with a historical calibration dataset $\mathcal{D}_{\mathcal{C}} = \{(\mathbf{X}_i, Y_i)\}_{i=-n+1}^0$ of size n , with index set $\mathcal{C}$. Test samples $(\mathbf{X}_t, Y_t)$ arrive sequentially for $t = 1, 2, \dots$. At each time t , the covariate $\mathbf{X}_t$ is observed but responses Y_t remain hidden until a real-time decision $\delta_t \in \{0, 1\}$ is made. The goal is to determine whether Y_t lies in a target region $\mathcal{A} \subseteq \mathcal{Y}$, for example, $\mathcal{A} = [a, b]$ or $\mathcal{A} = [b, \infty)$ in regression tasks. This can be

framed as an online testing problem with $\theta_t = \mathbb{I}\{Y_t \in \mathcal{A}\}$. The GAIF framework developed in the preceding sections offers a principled basis for analyzing this setting and enables rigorous theoretical guarantees.

We apply the GAIF framework to online conformal selection as follows. Since no off-the-shelf p -values are available, we first construct feedback-assisted online conformal p -values in Section 4.1. We then incorporate them into a modified GAIF (or Adaptive GAIF) rule to obtain finite-sample mFDR control in Section 4.2. We further use feedback to adaptively select the score model for conformal testing in Section 4.3, and analyze its optimality in Section 4.4.

4.1 Construction of online conformal p -values

We begin by constructing suitable online p -values for this online conformal selection setting. Let $V(\mathbf{X})$ be a non-conformity score, where larger values indicate a higher likelihood that $\theta_i = 0$. Typically, $V(\mathbf{X})$ is a monotone transformation of the prediction $\hat{\mu}(\mathbf{X})$, assumed pre-trained to estimate Y_t . For example, if $\theta = \mathbb{I}\{Y_t \geq b\}$, one can take $V(X) = b - \hat{\mu}(\mathbf{X})$. For simplicity, we write $V_i = V(\mathbf{X}_i)$.

A natural approach to achieving online FDR control is to compute conformal p -values with the given calibration set $\mathcal{C}$ (Bates et al., 2023; Jin and Candès, 2023) and then apply GAIF or existing GAI rules. However, standard conformal p -values are not suitable here, since they do not satisfy the conditional super-uniformity in (6) and independence assumptions. To circumvent these issues, we adopt *online* conformal p -values, which are constructed by sequentially updating the calibration set. Specifically, at time t , the current point $(\mathbf{X}_t, Y_t)$ serves as the test point, but from time $t + 1$ onward, it is added to the calibration set. Let $\mathcal{C}_{0t}$ denote the dynamically updated index set of calibration samples with

$$\mathcal{C}_{0t} = \{-n < i < t : \theta_i = 0\}.$$

For test data $\mathbf{X}_t$ at time t , the online conformal p -value p_t is defined as

$$p_t = \frac{\sum_{i \in \mathcal{C}_{0t}} \mathbb{I}\{V_i < V_t\} + \xi_t \cdot [1 + \sum_{i \in \mathcal{C}_{0t}} \mathbb{I}\{V_i = V_t\}]}{1 + |\mathcal{C}_{0t}|}, \quad (12)$$

where $\xi_t \stackrel{\text{i.i.d.}}{\sim} \text{Unif}[0, 1]$ is an independent random variable for tie-breaking. Similar forms of online conformal p -values have been studied for testing exchangeability (Vovk et al., 2003; Vovk, 2021) and constructing online conformal prediction intervals (Angelopoulos et al., 2024).

Under exchangeability, this construction yields null p -values that are independent, thereby avoiding the complex dependence induced by structures inherent in offline conformal methods, which rely on shared calibration sets and produce p -values that are PRDS (Bates et al., 2023). We impose the following assumption.

Assumption 1 (Exchangeability in conformal setting). *The null data $\{(\mathbf{X}_i, Y_i) : i > -n, \theta_i = 0\}$ are exchangeable conditional on the non-null data $\{(\mathbf{X}_i, Y_i) : i > -n, \theta_i = 1\}$.*

This assumption is common in conformal inference (Marandon et al., 2024), and is weaker than requiring that the full data $\{(\mathbf{X}_i, Y_i)\}_{i > -n}$ are independent and identically distributed (i.i.d.). We formally state the validity and independence under the null in Proposition 10.

Proposition 10 (Validity and independence of online conformal p -values). *Suppose Assumption 1 holds. Then under the null, the online conformal p -value p_t defined in (12) is uniformly distributed on $[0, 1]$, and the null p -values $\{p_t : t \in \mathbb{N}, \theta_t = 0\}$ are mutually independent.*

The mutual independence above indeed relies on the full and instant feedback setting, since only in this case does $\mathcal{C}_{0t}$ contain the complete set of past null samples, allowing the sequential rank-based independence argument to hold. For delayed feedback, additional adjustments are needed; see the finite-sample analysis in Appendix C.1. We also allow the score function to vary over time and be data-dependent, provided that it remains symmetric on the null samples; see Appendix F.5. This flexibility enables leveraging non-null feedback to improve score quality, facilitating score selection and adaptation to distribution shift.

4.2 Online conformal testing with feedback

Having established the properties of online conformal p -values, we now introduce the corresponding testing procedures. Although applying GAIF directly to those p -values yields satisfactory FDR control empirically, a finite-sample theoretical guarantee remains lacking. The key challenge is that null online conformal p -values may still be dependent on non-null decisions through the adaptive testing levels.

To resolve this, we construct safe variants of LF and SF in which the rejection rewards are triggered only by confirmed null rejections. Specifically, we replace the usual rejection times τ_j with $\tilde{\tau}_j$, the time of the j -th null rejection:

$$\tilde{\tau}_j = \inf \left\{ t \in \mathbb{N} : \sum_{i \leq t} \delta_i (1 - \theta_i) \geq j \right\}.$$

Since $\tilde{\tau}_j$ is determined solely by null rejections, the resulting test levels do not depend on non-null decisions. We denote the resulting procedures by LFS and SFS, where “S” stands for “safe.” The corresponding *Online Conformal Testing with Feedback* (OCTF) procedure is given in Algorithm 2, and the following theorem establishes its finite-sample mFDR control.

Algorithm 2 Online conformal testing with feedback (OCTF)

Input: Initial calibration data $\mathcal{D}_C = \{(\mathbf{X}_i, Y_i)\}_{i=-n+1}^0$, target region $\mathcal{A}$, FDR target level $\alpha \in (0, 1)$, non-conformity score function $V(\cdot)$, parameter s_0 , parameter sequence $\{\gamma_t\}_{t \in \mathbb{N}}$, constant $\lambda \in (0, 1)$.

- 1: Initialize $\mathcal{C}_{01} = \{i \in \mathcal{C} : \theta_i = 0\}$.
 - 2: **for** $t = 1, 2, \dots$ **do**
 - 3: Observe test covariate $\mathbf{X}_t$.
 - 4: Compute conformity scores $\{V_i\}_{i \in \mathcal{C}_{0t} \cup \{t\}}$.
 - 5: Compute online conformal p -value p_t via (12).
 - 6: Update $\alpha_t = \alpha_t^{\text{LFS}}$ in (18) (or $\alpha_t = \alpha_t^{\text{SFS}}$ in (19)).
 - 7: **if** $p_t \leq \alpha_t$ **then** $\delta_t = 1$, **else** $\delta_t = 0$.
 - 8: Obtain the revealed feedback Y_t , and thus obtain θ_t .
 - 9: Update the calibration dataset $\mathcal{C}_{0t}$.
 - 10: **end for**
- Output:** Rejection set $\mathcal{R} = \{t : \delta_t = 1\}$.
-

Theorem 11 (Finite-sample online mFDR control for OCTF). *Suppose Assumption 1 holds. The output $\mathcal{R}$ of Algorithm 2 satisfies $\text{mFDR}(t) \leq \alpha$ for all $t \in \mathbb{N}$.*

Theorem 11 is established under a dependence regime that naturally arises from online conformal p -values: the null p -values are mutually independent, whereas their dependence with non-null p -values is generally unspecified. This regime is not directly covered by the standard assumptions used in classical online testing procedures such as LOND, LORD, and SAFFRON. In this setting, LFS and SFS trade some potential power for a rigorous theoretical guarantee. LF and SF retain the more operationally natural reward structure and serve as more aggressive alternatives when empirical power is the primary consideration. Although they maintain empirical FDR control in our simulations (Section 5), their finite-sample validity under this dependence regime remains open.

Remark 12. *Theorem 11 controls mFDR rather than FDR. Although the safe construction removes past non-null rejections from the threshold update, they remain in the random FDP denominator and may be dependent on a current null p -value. The resulting dependence makes extending mFDR control to FDR nontrivial (Zrnic et al., 2021). Finite-sample FDR control under the present assumptions therefore remains open.*

4.3 Optimized online conformal testing with feedback

The performance of online conformal testing depends on the choice of the non-conformity scores $V(\mathbf{X})$ and the prediction model, e.g., random forests, neural networks, support vector machines, or regularized linear models (Bai and Jin, 2024; Gasparin and Ramdas, 2024). In practice, the most effective score may also change over time, especially when the non-null distribution drifts (Gang et al., 2023; Huo et al., 2024). This motivates a feedback-driven score-selection step built on top of the OCTF framework.

Suppose there are K pre-trained candidate score functions $\{V(\cdot; k) : \mathcal{X} \rightarrow \mathbb{R}\}_{k=1}^K$. At each time t , a score is chosen as

$$\hat{k}_t = \arg \min_{k \in [K]} \mathcal{M}(\mathcal{D}_t; k),$$

where $\mathcal{M}(\mathcal{D}_t; k)$ is a given criterion evaluated by the currently observed data $\mathcal{D}_t = \{(\mathbf{X}_i, Y_i) : -n < i < t\} \cup \{\mathbf{X}_t\}$. Our goal is to choose the score that is most favorable for detecting non-nulls, which motivates the design of a reliable criterion $\mathcal{M}(\mathcal{D}_t; k)$. Since the testing threshold α_t depends recursively on the past rejection path, directly optimizing rejection probability is difficult. We therefore use the proxy

$$\mathbb{E}[p_{k,t} \mid \theta_t = 1],$$

where $p_{k,t}$ is the t -th conformal p -value obtained from the score function $V(\cdot; k)$. Smaller non-null p -values generally correspond to higher rejection probability $\Pr(p_{k,t} \leq \alpha_t \mid \theta_t = 1)$, so this provides a natural threshold-free criterion.

To estimate $\mathbb{E}[p_{k,t} \mid \theta_t = 1]$ under shifted non-null distribution while preserving symmetry, we employ an exponentially weighted moving average (EWMA) of past auxiliary non-null

p -values. Specifically, at time t , for each score function k , we construct *auxiliary non-null p -values* as

$$\tilde{p}_{k,j} = \frac{\sum_{s \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{X}_s; k) \leq V(\mathbf{X}_j; k)\}}{1 + |\mathcal{C}_{0t}|}, \quad j \in \mathcal{C}_{1t}, \quad (13)$$

where $\mathcal{C}_{1t} = \{-n < i < t : \theta_i = 1\}$ denotes the set of online non-null samples observed prior to time t . When the non-null distribution varies slowly over time, these auxiliary non-null p -values are approximately distributed like the current non-null p -value $p_{k,t}$. The construction in (13) is carefully designed so that the resulting auxiliary non-null p -values are invariant under permutations of $\mathcal{C}_{0t} \cup \{t\}$, which is crucial for ensuring valid inference after score selection. We then define the EWMA criterion as

$$\mathcal{M}_t^{\text{EWMA}}(\mathcal{D}_t; k) = \frac{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j} \tilde{p}_{k,j}}{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j}},$$

where $\rho \in (0, 1)$ is the user-specified decay parameter that downweights past observations, enabling dynamic adaptation to the recent non-null distribution and precise estimation of $\mathbb{E}[p_{k,t} \mid \theta_t = 1]$. The selected score function is accordingly as $\hat{k}_t = \arg \min_{k \in [K]} \mathcal{M}_t^{\text{EWMA}}(\mathcal{D}_t; k)$.

The online conformal p -value after score selection, denoted as p_t^{opt} , is then computed by (12) using the selected score $V(\mathbf{X}_t; \hat{k}_t)$. Then running the OCTF procedures in Section 4.2 with these optimized p -values yields the whole optimized procedure, which we refer to as Opt-OCTF algorithm.

Corollary 13 (Finite-sample online mFDR control for Opt-OCTF). *Under the assumptions in Theorem 11, the Opt-OCTF procedure by $\mathcal{M}_t^{\text{EWMA}}$ satisfies $\text{mFDR}(t) \leq \alpha$ for all $t \in \mathbb{N}$.*

Although Corollary 13 is stated for the EWMA criterion, the underlying framework for optimized online conformal testing can be extended more broadly. In particular, we allow more general selection criteria $\mathcal{M}$ by using a full permutation strategy over $\mathcal{C}_{0t} \cup \{t\}$; see Appendix C.2. In addition, data-driven score functions obtained by online updating are also compatible with the framework under a suitable symmetry-preserving training scheme. However, we focus on selection among K pre-trained candidates in the main text for computational reasons; see Appendix C.3 for details.

Remark 14. *In (13), we include the current time point t in the null calibration set to preserve symmetry over $\mathcal{C}_{0t} \cup \{t\}$, which is essential for the validity of the optimized conformal p -value after score selection. Although this may slightly affect calibration when t is non-null, the effect is typically minor in practice. A simple truncated variant can further reduce this influence while preserving validity; see Appendix E.6 for details.*

4.4 Optimality analysis of EWMA criterion

We now give a supporting analysis for the EWMA criterion by establishing certain optimality of the online score selection step. To facilitate analysis, we work under the standard two-group model, which is widely used to study the performance of online testing procedures (Gang et al., 2023; Humbert et al., 2025).

Assumption 2 (Two-group model). *Suppose the score functions follow the two-group model for each $k \in [K]$:*

$$V(\mathbf{X}_t; k) \stackrel{\text{indep}}{\sim} \pi_t F_k^{(0)} + (1 - \pi_t) F_{k,t}^{(1)}, \quad \theta_t \stackrel{\text{indep}}{\sim} \text{Bernoulli}(1 - \pi_t),$$

where $F_k^{(0)}$ and $F_{k,t}^{(1)}$ are the null and (time-varying) non-null score distributions for k -th model.

Under this model, the oracle conformal p -value associated with score $V(\cdot; k)$ is $F_k^{(0)}(V(\mathbf{X}_t; k))$. Given $\theta_t = 1$, a smaller value of $F_k^{(0)}(V(\mathbf{X}_t; k))$ indicates higher detection power. And we define the optimal score function at time t as

$$k_t^* = \arg \min_{k \in [K]} \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_t; k)) \mid \theta_t = 1].$$

Note that $\mathbb{E}[F_k^{(0)}(V(\mathbf{X}_t; k)) \mid \theta_t = 1]$ is the population analogue of the empirical proxy $\mathbb{E}[p_{k,t} \mid \theta_t = 1]$. To show that the EWMA criterion tracks this oracle target and selects k_t^* reliably, we impose standard regularity conditions.

Assumption 3 (Slowly varying distribution). *For each $k \in [K]$ and $t \in \mathbb{N}$, $F_{k,t}^{(1)}$ has a slowly varying distribution $\|F_{k,t}^{(1)} - F_{k,t-1}^{(1)}\|_\infty \leq \gamma$ such that $\gamma \rightarrow 0$.*

Assumption 4 (Non-degeneracy). *There exist constants $0 < \underline{\pi} \leq \bar{\pi} < 1$ such that null proportion satisfies $\pi_t \in [\underline{\pi}, \bar{\pi}]$.*

Assumption 5 (Margin). *For any t and a constant $c > 0$, $\min_{i \neq k_t^*} |\mathbb{E}[F_i^{(0)}(V(\mathbf{X}_t; i)) \mid \theta_t = 1] - \mathbb{E}[F_{k_t^*}^{(0)}(V(\mathbf{X}_t; k_t^*)) \mid \theta_t = 1]| \geq c$.*

Assumption 3 allows the non-null distribution $F_{k,t}^{(1)}$ to drift over time. Assumption 4 ensures that neither nulls nor non-nulls become too rare over time. Assumption 5 guarantees the optimal score is identifiable.

Theorem 15 (Optimality of online score selection). *Suppose Assumptions 2-4 hold.*

(i) *If $t + n \geq 1 + 1/(1 - \rho)$, then for any $\varepsilon_1 \in (0, \sqrt{\underline{\pi}/2})$ and $\varepsilon_2 \in (0, 1 - \rho)$,*

$$\begin{aligned} & \Pr \left(\left| \mathcal{M}^{\text{EWMA}}(k, \mathcal{D}_t) - \mathbb{E}[F_0^k(V(\mathbf{X}_t; k)) \mid \theta_t = 1] \right| \leq \varepsilon_1 + \varepsilon_2 + \frac{4\gamma}{(1 - \bar{\pi})(1 - \rho)} + \frac{2}{(t + n - 1)\underline{\pi}} \right) \\ & \geq 1 - 3 \exp \left\{ -(t + n - 1)\underline{\pi}\varepsilon_1^2 \right\} - 4 \exp \left\{ -\frac{\varepsilon_2^2(1 - \bar{\pi})^2}{6(1 - \rho)} \right\}. \end{aligned}$$

Further assume that Assumption 5 holds. Then the following results in (ii)-(iii) hold.

(ii) *If $t + n \geq 1 + 16/(c\underline{\pi})$ and $\gamma/(1 - \rho) \leq (1 - \bar{\pi})c/32$, we have the finite-sample error bound for the score selection*

$$\Pr(\hat{k}_t \neq k_t^* \mid \theta_t = 1) \leq 3K \exp \left\{ -(t + n - 1)\underline{\pi} \frac{c^2}{64} \right\} + 4K \exp \left\{ -\frac{c^2(1 - \bar{\pi})^2}{384(1 - \rho)} \right\}.$$

(iii) Moreover, the score selection consistency for all non-null points holds as

$$\Pr(\exists i \leq t : \theta_i = 1, \hat{k}_i \neq k_i^*) \leq \frac{3K(1-\pi)\exp\{-n\pi c^2/64\}}{1-\exp\{-\pi c^2/64\}} + 4K(1-\pi)t \exp\left\{-\frac{c^2(1-\pi)^2}{384(1-\rho)}\right\}.$$

In particular, if $(1-\rho)\log(t) \rightarrow 0$, $\gamma/(1-\rho) \rightarrow 0$ and $n \rightarrow \infty$, $\Pr(\exists i \leq t : \theta_i = 1, \hat{k}_i \neq k_i^*) \rightarrow 0$.

Theorem 15 (i) shows that the EWMA criterion concentrates around the oracle power proxy. The bound decomposes naturally into: an estimation term scaling with $t+n$, a non-null effective size term in EWMA operation by $(1-\rho)^{-1}$ and a drift bias of order $\gamma/(1-\rho)$ reflecting how well the EWMA estimator can track a drifting non-null distribution. Parts (ii)–(iii) translate this concentration into optimal score selection guarantees. The rate conditions reflect the classic EWMA trade-off: ρ must be sufficiently close to 1 to ensure an adequate effective sample size, yet small enough to allow the procedure to adapt to distribution drift.

5. Numerical Experiments

In this section, we present extensive synthetic experiments to demonstrate the validity and efficiency of our proposed methods. First, Section 5.1 studies GAIF and Adaptive GAIF in the non-conformal setting. Second, Sections 5.2 and 5.3 present the results for OCTF and Opt-OCTF under the conformal setting, respectively. Following the setup in Robertson et al. (2023), we fix $\lambda = 0.5$, $\gamma_j \propto j^{-1.6}$ for all $j \in \mathbb{N}$, and $s_0 = \alpha/2$ throughout. Code for implementing our methods and reproducing the experiments and figures in our paper is available at <https://github.com/lulin2023/GAIF>.

5.1 Synthetic experiments under non-conformal settings

We begin by presenting experiments on online multiple testing problems under three scenarios. Throughout, let π_1 denote the non-null proportion parameter in the data-generating model; its value varies across experiments.

- **Scenario I (Testing with Gaussian observations)** We simulate T independent test statistics $Z_t \sim N(\mu_t, 1)$ and test hypotheses $\mathbb{H}_{0t} : \mu_t = 0$ using one-sided p -values $p_t = \Phi(-Z_t)$, where Φ is the standard Gaussian CDF. The signal strengths μ_t follow a mixture model:

$$\mu_t = \begin{cases} 0 & \text{with probability } 1 - \pi_1 \\ F_1 & \text{with probability } \pi_1, \end{cases} \quad (14)$$

where the random variable $F_1 \sim N(2.5, 1)$.

- **Scenario II (Testing with Beta alternatives)** We generate p -values according to the following model:

$$p_t \sim \begin{cases} \text{Unif}[0, 1] & \text{with probability } 1 - \pi_1 \\ \text{Beta}(0.5, 4) & \text{with probability } \pi_1. \end{cases} \quad (15)$$

- **Scenario III (Testing under local dependence)** We simulate correlated test statistics $(Z_1, \dots, Z_T)^\top \sim N(\boldsymbol{\mu}, \Sigma)$ where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_T)^\top$ with $\mu_t = 2$ for a randomly chosen fraction π_1 of indices and $\mu_t = 0$ otherwise. The covariance matrix Σ has a block-diagonal structure: coordinates are grouped into blocks of size $n_{\text{block}} = 10$, with within-block correlation $\rho = 0.8$ and diagonal elements as 1. We test hypotheses $\mathbb{H}_{0t} : \mu_t = 0$ using one-sided p -values $p_t = \Phi(-Z_t)$.

We set the stopping time as $T = 1000$ and the FDR level at $\alpha = 0.1$. We evaluate the performance via empirical FDR and power averaged over 500 independent replications. Appendix E confirms that mFDR and FDR show similar trends.

Firstly, we consider the full and instant feedback setting, and compare our methods SF and LF with state-of-the-art algorithms for online FDR control, namely LOND (Javanmard and Montanari, 2015), LORD++, and SAFFRON, using their default parameters ($\lambda = 0.5$ chosen for SAFFRON).

Figure 2 shows results for varying non-null proportion $\pi_1 \in [0.1, 0.8]$ under Scenarios I and II. All benchmark methods maintain valid empirical FDR control across different π_1 under both Scenarios. Our feedback-based SF and LF methods enhance detection power while preserving valid empirical online FDR control: SF consistently outperforms SAFFRON, while LF yields higher power than LORD++. When π_1 is small, SF benefits from adaptive wealth allocation and tends to outperform LF. When π_1 is large, LF gains more from frequent non-null feedback and can recycle more released α -wealth and thereby surpassing SF in power. In contrast, LORD++ and LOND remain conservative across different values of π_1 . Numerical values together with Monte Carlo standard errors are reported in Appendix E (Table 3).

Figure 3 presents results under a positive local dependence structure in Scenario III. The baseline methods—LOND, LORD++, and SAFFRON—are included for comparison. The proposed procedures, SF_{dep} and LF_{dep} , correspond to Ada-GAIF_{dep} and GAIF_{dep} under full and instant feedback, respectively. In addition, we include the existing methods $\text{SAFFRON}_{\text{dep}}$ and LORD_{dep} from Zrnic et al. (2021). We find that SF, LF, and SAFFRON fail to control the FDR under dependence. In contrast, the dependence-aware procedures SF_{dep} and LF_{dep} successfully achieve valid FDR control and attain higher power than their feedback-ignoring counterparts, $\text{SAFFRON}_{\text{dep}}$ and LORD_{dep} . See more detailed numerical values in Appendix E (Table 4).

Next, we evaluate the performance of our proposed methods under other feedback settings: bandit and instant, and full and delayed. Taking the classical methods LOND, LORD++, and SAFFRON as benchmarks, we implement the proposed LF-BI and SF-BI under bandit and instant feedback setting. The results in Scenario I are shown in Figure 4. All methods maintain valid empirical FDR control. Under bandit feedback, LF-BI remains more powerful than LORD++ by recycling wealth from rejected non-nulls. SF-BI coincides with SAFFRON because every rejected hypothesis satisfies $p_t \leq \alpha_t \leq \lambda$, so the feedback term $\mathbf{1}\{p_t > \lambda\}$ is zero. Numerical results are reported in Table 5.

For the full and delayed feedback setting, we consider delays $d \in \{0, 10, 100\}$ and compare the proposed SF-FD and LF-FD against feedback-free baselines. Figure 5 displays the results under Scenario I. As the delay d increases, the power of our methods decreases gradually, yet they continue to outperform the baselines by a substantial margin while maintaining valid empirical FDR control. When $d = 0$ and the non-null proportion is large, LF achieves

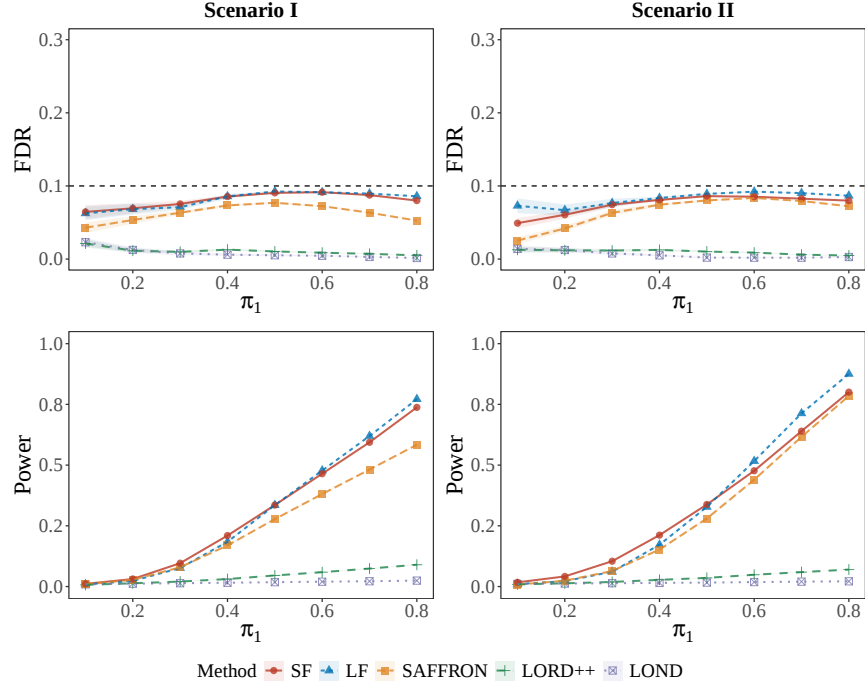


Figure 2: Results for Scenario I and Scenario II under full and instant feedback: FDR and Power at stopping time with non-null proportion π_1 ranging from 0.1 to 0.8 across 500 replications; The black dashed lines denote the FDR level $\alpha = 0.1$. Shaded areas show ± 1 standard error.

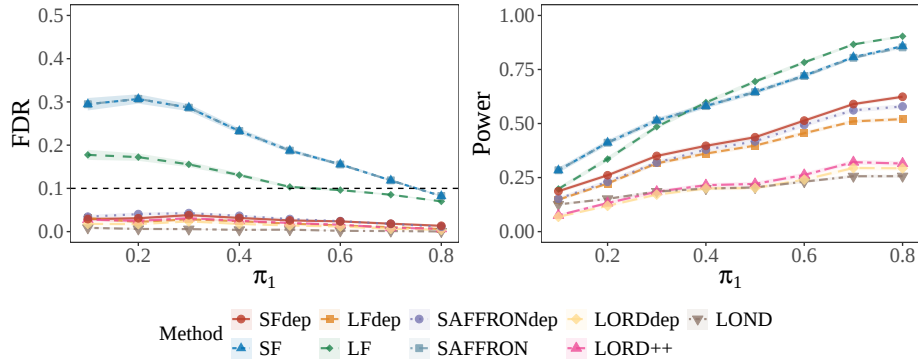


Figure 3: Results for Scenario III under full and instant feedback: FDR and Power at stopping time with non-null proportion π_1 ranging from 0.1 to 0.8 across 500 replications. The black dashed line indicates the target FDR level $\alpha = 0.1$. Shaded areas show ± 1 standard error.

the highest power, as immediate feedback enables aggressive α -wealth recycling from the abundant revealed non-nulls. Numerical results are reported in Table 6.

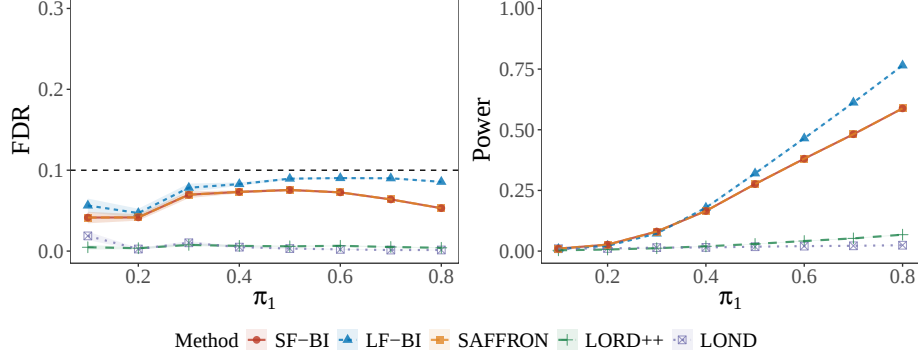


Figure 4: Results for Scenario I under bandit and instant feedback: FDR and Power at stopping time with non-null proportion π_1 ranging from 0.1 to 0.8 across 500 replications. The black dashed line indicates the target FDR level $\alpha = 0.1$. Shaded areas show ± 1 standard error.

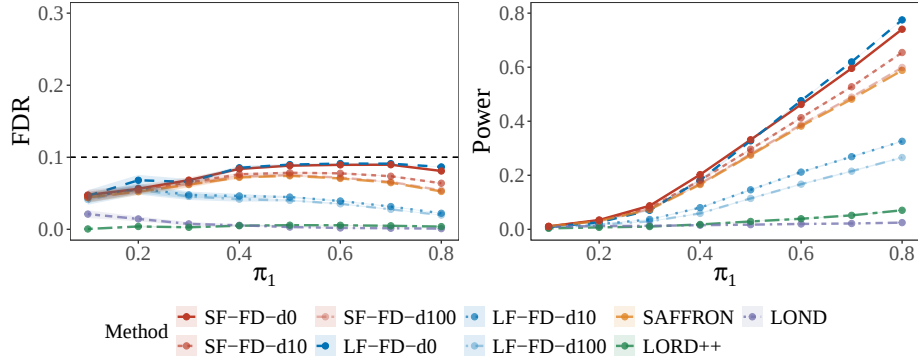


Figure 5: Results for Scenario I under full and delayed feedback: FDR and power at stopping time with non-null proportion π_1 ranging from 0.1 to 0.8 across 500 replications. The black dashed line indicates the target FDR level $\alpha = 0.1$. Shaded areas show ± 1 standard error.

5.2 Results for online conformal testing with feedback

We next evaluate OCTF with a fixed score in a real-time binary classification task.

- **Scenario IV:** Data are from $\mathbf{X} \mid Y = 0 \sim \mathcal{N}_4(\boldsymbol{\mu}_1, \mathbf{I}_4)$, and $\mathbf{X} \mid Y = 1 \sim \mathcal{N}_4(\boldsymbol{\mu}_2, \mathbf{I}_4)$, where $\boldsymbol{\mu}_1 = (2, 0, 0, 0)^\top$, $\boldsymbol{\mu}_2 = (0, 0, -1, -1)^\top$. The target region is $\mathcal{A} = \{1\}$.

We set the stopping time to $T = 1000$ and the target FDR level to $\alpha = 0.2$. The initial calibration size is $n_{\text{cal}} = 1000$. We consider the score as $V(\mathbf{X}) = 1 - \hat{\mu}(\mathbf{X})$, where the predictive model $\hat{\mu}(\mathbf{X})$ is a support vector machine classifier trained on $n_{\text{tr}} = 1000$. Appendix E reports additional results for varying n_{cal} , alternative training algorithms, and a regression setting.

We evaluate the proposed LF and SF against the benchmark procedures LOND, LORD++, and SAFFRON, all implemented within the OCTF workflow using their own testing levels $\{\alpha_t\}$. Figure 6 reports the empirical FDR and power as the non-null proportion π_1 varies

over $[0.1, 0.8]$. All five procedures control the empirical FDR below the nominal level α across the entire range. As expected, the feedback-enhanced procedures deliver substantial power gains over their non-feedback counterparts: SF improves consistently over SAFFRON across all π_1 , while LF trails SF for small π_1 but overtakes it once π_1 exceeds roughly 0.5, ultimately attaining the highest power among all methods considered. We also report results for the safe variants SFS and LFS in Appendix E.1; these procedures are more conservative, consistent with their design of prioritizing finite-sample validity via a more restrictive wealth update.

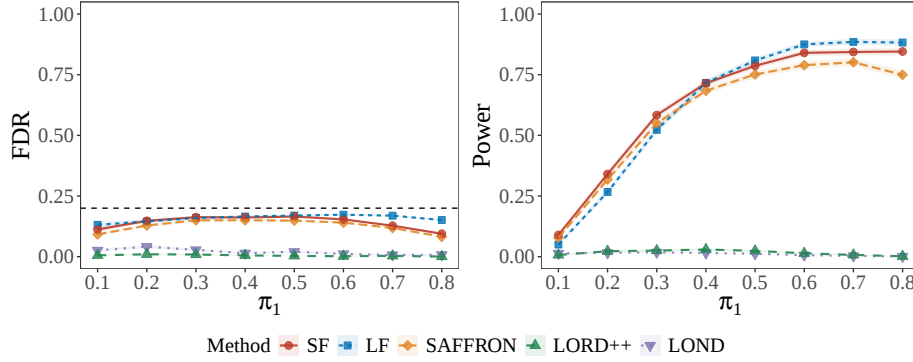


Figure 6: Results for Scenario IV: FDR and power at stopping time with non-null proportion π_1 ranging from 0.1 to 0.8 across 500 replications. The black dashed lines denote the FDR level $\alpha = 0.2$. Shaded areas show ± 1 standard error.

5.3 Results for optimized online conformal testing with feedback

We now demonstrate the effectiveness of the proposed score selection strategy from Section 4.3. The goal here is to examine whether feedback can be used not only for valid testing, but also for adaptive score selection when non-null distributions shift.

We employ the binary classification example in Scenario IV with $\boldsymbol{\mu}_1 = (3, 0, 0, 0)^\top$, $\boldsymbol{\mu}_2 = (0, 0, -1, -1)^\top$. For the historical labeled dataset, the non-null proportion is fixed as 0.2. For the testing data, we consider a sine pattern shift, where the non-null proportion varies as $\pi_t = \Pr(Y_t = 1) = \{\sin(100\pi t/T) + 1\}/4$, oscillating between 0 and 0.5.

For the implementation of the score selection, we consider $K = 3$ candidate models: random forest (RF), neural network (NN), and support vector machine (SVM) and employ the EWMA criterion $\mathcal{M}_t^{\text{EWMA}}(k, \mathcal{D}_t)$ with exponential weight parameter $\rho = 0.95$. To reduce computational cost, the auxiliary non-null p -value is computed by a sliding window of fixed length $L = 100$: for $t > L$, we define $\mathcal{C}_{1t} = \{i : t - L \leq i < t, \theta_i = 1\}$; for $t \leq L$, we use the full set $\mathcal{C}_{1t} = \{i < t : \theta_i = 1\}$.

To evaluate the score-selection step, we compare the proposed optimized rule with random score selection under different testing procedures. Here, “Opt-” means that the score model is selected adaptively by the proposed EWMA criterion, while “Ran-” means that the score model is chosen uniformly at random from the candidate library. For example, Opt-LF

denotes the optimized version of LF, whereas Ran-LF uses the same testing rule with random score selection.

Figure 7 shows results across time t under distribution shift; results for the SFS- and LFS-based variants are deferred to Appendix E. All methods successfully control empirical FDR at or below the nominal level $\alpha = 0.05$. In terms of power, the optimized (Opt) variants consistently outperform their randomly selected (Ran) counterparts, confirming that optimized score selection is beneficial under distribution shift. Additional results for scenarios without distribution shifts are presented in Appendix E, and similarly highlight the advantage of optimized score selection.

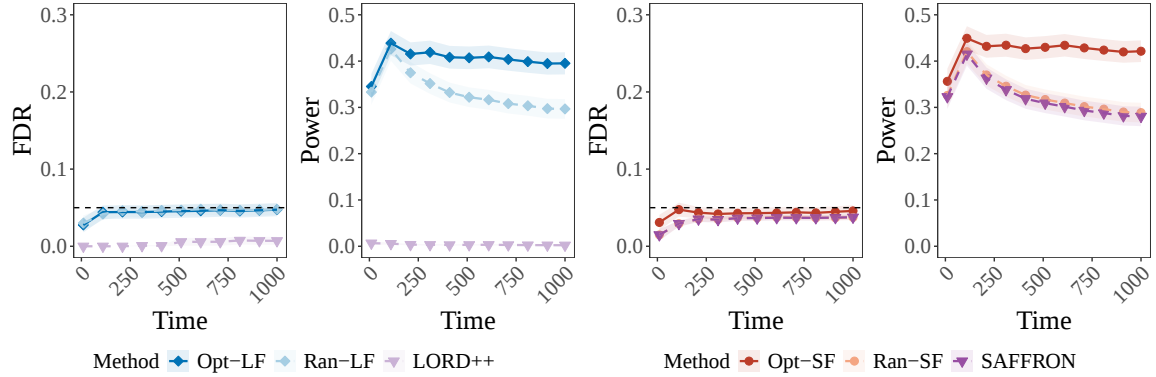


Figure 7: Results for Scenario IV (sine pattern shifts): Empirical $\text{FDR}(t)$ and $\text{Power}(t)$ across different time t . The black dashed lines denote the FDR level $\alpha = 0.05$. Shaded areas show ± 1 standard error.

6. Real Data Applications

In this section, we evaluate our proposed methods on four real-world datasets to illustrate their practical benefits in diverse online decision-making tasks.

- Task 1: Online Candidate Screening.** The first task is real-time candidate screening for selecting the candidates who can get into the interview process. We use the recruitment dataset (Kaggle, 2020), which contains $N = 45,372$ candidates with 11 attributes, including education status, handicapped or not, and gender. The target binary variable is whether the candidate passes the job interview. In this task, the non-null proportion is $\pi_1 = 0.69$.
- Task 2: High-Risk Diabetes Identification.** The second task focuses on health screening, where selecting individuals at high risk of diabetes is critical for early intervention. We use the diabetes health indicators dataset (Kaggle, 2021), which contains $N = 70,692$ samples and 22 covariates, including demographic attributes (e.g., sex, age, BMI), lifestyle-related features, and several binary health indicators. The target binary variable is whether an individual has diabetes. The non-null proportion is $\pi_1 = 0.50$ in the dataset.

- **Task 3: High-Income Individual Selection.** The third task involves using the 1994 Census Bureau dataset (Becker and Kohavi, 1996) to identify a subset of individuals with high incomes (i.e., income $> 50K$) for precision marketing. This dataset contains $N = 32,561$ records with 14 attributes, including gender, race, marital status, education level, and more. The non-null proportion is $\pi_1 = 0.25$.
- **Task 4: Airfoil Noise Detection.** The final task involves using the airfoil dataset (Brooks et al., 2014) from the UCI Machine Learning Repository to identify samples with high sound pressure. This dataset contains $N = 1,503$ observations with five physical covariates (log frequency, angle of attack, chord length, free-stream velocity, and suction side log displacement thickness). The response variable Y represents the scaled sound pressure, and we aim to test $\mathbb{H}_{0t} : Y_t \in [-\infty, c)$, where c is the $(1 - \pi_1)$ -quantile of Y , with $\pi_1 = 0.60$.

For Tasks 1-3, we randomly sample three subsets from the full dataset: $n_{\text{tr}} = 1,000$ training, $n_{\text{cal}} = 1,000$ calibration, and $n_{\text{te}} = 1,000$ test samples. For Task 4, we set $n_{\text{tr}} = 300$, $n_{\text{cal}} = 300$, and $n_{\text{te}} = 1,000$. We compare Opt-SF and Opt-LF with three feedback-free benchmarks, SAFFRON, LORD++, and LOND. At each time step, $K = 3$ candidate training algorithms (RF, NN, and SVM) are employed, and the score selection criterion is $\mathcal{M}_t^{\text{EWMA}}(k, \mathcal{D}_t)$ with $\rho = 0.9$. Our methods use the optimized conformity score, while the LORD++ and SAFFRON use a randomly selected score at each time t . All results are averaged over 500 replications.

Figure 8 summarizes the empirical FDR and power over time t , and Table 1 reports the corresponding values at the stopping time $T = 1,000$ at the target FDR level $\alpha = 0.3$. Across all four tasks, all methods maintain empirical FDR below the nominal level.

The relative performance of Opt-LF and Opt-SF is consistent with our earlier simulation findings. For Tasks 1, 2, and 4, where the non-null proportion is relatively large ($\pi_1 = 0.69$, 0.50, and 0.60 respectively), Opt-LF achieves the highest power (0.308, 0.163, and 0.684), as abundant non-null feedback enables more aggressive α -wealth recycling. For Task 3 ($\pi_1 = 0.25$), Opt-SF instead achieves the best power (0.151), consistent with the observation that adaptive α -wealth allocation becomes more consequential when non-nulls are sparse and feedback-recovered wealth is limited. SAFFRON and LORD++ exhibit substantially lower power across all tasks, with LORD++ and LOND achieving near-zero power in most settings. Results for the safe variants Opt-SFS and Opt-LFS are deferred to Appendix E (Table 9); both achieve empirical FDR control and have conservative power across all tasks.

Table 1: FDR(T) and Power(T) for different tasks across four datasets (Candidate, Diabetes, Income, Airfoil). The target FDR level is $\alpha = 0.3$. Bold numbers represent the best results.

Method	Task 1		Task 2		Task 3		Task 4	
	FDR	Power	FDR	Power	FDR	Power	FDR	Power
Opt-SF	.118(.014)	.181(.018)	.135(.016)	.121(.015)	.166(.014)	.151(.012)	.166(.008)	.499(.028)
Opt-LF	.152(.014)	.308(.026)	.138(.016)	.163(.019)	.181(.017)	.126(.012)	.228(.007)	.684(.024)
SAFFRON	.034(.010)	.001(.000)	.102(.016)	.022(.005)	.095(.013)	.031(.004)	.050(.007)	.136(.015)
LORD++	.000(.000)	.000(.000)	.010(.007)	.000(.000)	.011(.007)	.001(.000)	.002(.001)	.012(.004)
LOND	.015(.009)	.000(.000)	.045(.014)	.000(.000)	.035(.011)	.004(.000)	.007(.004)	.003(.000)

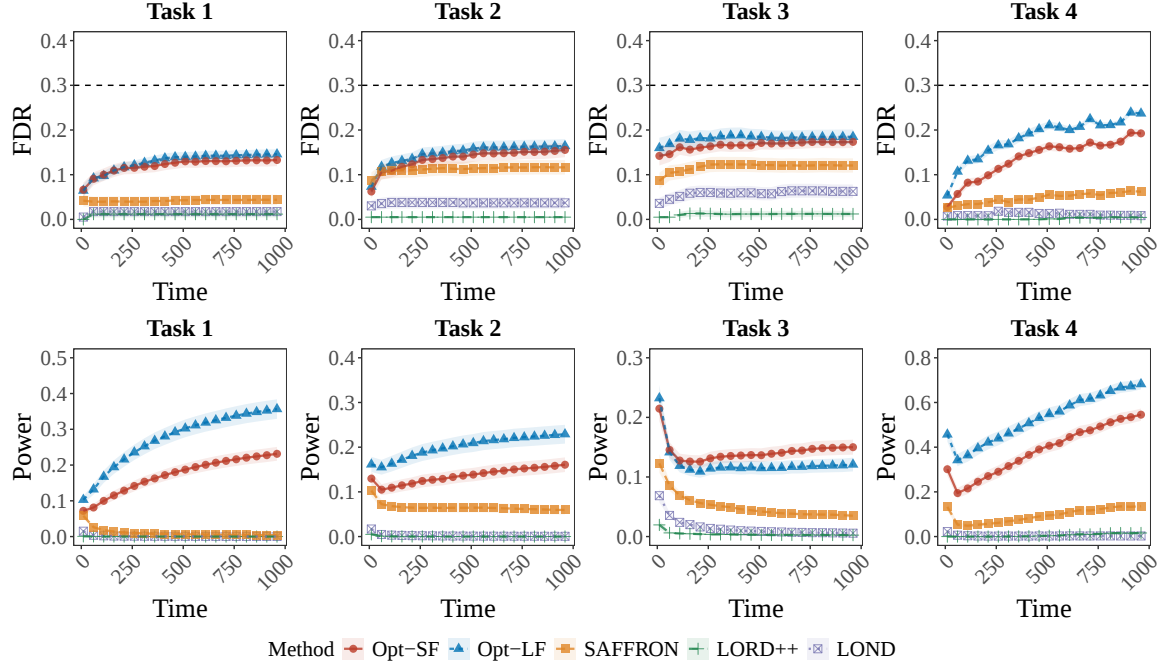


Figure 8: Results for real-data applications: the values of $\text{FDR}(\delta^t)$ and $\text{Power}(\delta^t)$ over time t for five methods across four tasks (Task 1: candidate screening, $\pi_1 = 0.69$; Task 2: diabetes identification, $\pi_1 = 0.50$; Task 3: income selection, $\pi_1 = 0.25$; Task 4: airfoil noise detection, $\pi_1 = 0.60$). The black dashed lines indicate the FDR level $\alpha = 0.3$. Shaded areas show ± 1 standard error.

7. Concluding Remarks

We study online multiple testing with feedback, aiming to develop reliable machine learning methods for real-time decision-making with rigorous FDR and mFDR control. Our key contribution is GAIF and its adaptive variants, a feedback-enhanced Generalized Alpha-Investing framework that guarantees online FDR control under the standard assumption of independence between null and all other p -values. Building on GAIF, we further develop OCTF for online conformal testing by constructing independent online conformal p -values for null hypotheses, which achieves finite-sample mFDR control. To further enhance performance and address distribution shifts among non-nulls, we propose a feedback-driven score selection strategy and provide theoretical analysis characterizing its optimality properties.

We highlight several potential directions for future work. First, the current framework mainly focuses on distribution shifts in the non-null data. To address alpha-death and piggybacking under such shifts, one can consider controlling weighted FDR using forgetting factors or decaying memory mechanism (Ramdas et al., 2017). Second, extending our framework to accommodate more general forms of distribution shift is another promising direction. One potential approach is to incorporate adaptive strategies that exploit local information, in the spirit of the adaptive conformal inference (Gibbs and Candès, 2021). Finally, to establish guarantees under weaker assumptions, one could relax the definition of valid online FDR control and develop new error-rate notions, analogous to the average-coverage criterion studied in online conformal inference (Gibbs and Candès, 2021; Humbert et al., 2025).

Acknowledgments

We would like to acknowledge the action editor and three anonymous referees for their valuable comments and suggestions, which have improved the manuscript greatly. Lu and Huo contributed equally to this work. Zou and Ren are the corresponding authors. Ren was supported by the National Key R&D Program of China(No. 2024YFA1012200), and the National Natural Science Foundation of China (Nos. 12471262,12522115), and Shanghai Jiao Tong University 2030 Initiative. Wang and Zou were supported by the National Key R&D Program of China (No. 2022YFA1003703) and the National Natural Science Foundation of China (Nos. 12231011, 12531011). Huo was supported by the National Natural Science Foundation of China (No. 124B2016).

Supplementary Material for “Feedback-Enhanced Online Multiple Testing with Applications to Conformal Selection”

This supplementary material includes:

- Notation and Algorithmic Details (Appendix A)
- Additional Discussions on GAIF and Adaptive GAIF (Appendix B)
- Additional Discussions on OCTF (Appendix C)
- Applications on Real-time LLM Alignment (Appendix D)
- Additional Experimental Results (Appendix E)
- Proofs of all the theoretical results (Appendix F)

Appendix A. Notation and Algorithmic Details

This appendix gathers the notation and additional algorithmic details used throughout the paper. For completeness, we first review several preliminary terms and benchmark procedures, and then provide the concrete forms of the proposed testing levels and related implementation details.

A.1 Preliminary terms for self-containment

Here, we list the preliminary terms we use in the paper for the sake of clarity and self-containment.

- FDR (Benjamini and Hochberg, 1995), false discovery rate, a widely-adopted error rate notion in the field of multiple testing, is defined as the expected proportion of incorrectly rejected null hypotheses as follows:

$$\text{FDP}(t) = \frac{V_t}{1 \vee R(t)} := \frac{\sum_{j=1}^t \delta_j (1 - \theta_j)}{1 \vee \sum_{j=1}^t \delta_j} \quad \text{and} \quad \text{FDR}(t) := \mathbb{E}[\text{FDP}(t)],$$

where $R(t)$ represents the number of rejections until time t and $V(t)$ is the number of false discoveries.

- mFDR, modified false discovery rate, or marginal false discovery rate, is defined as:

$$\text{mFDR}(t) := \frac{\mathbb{E}[V(t)]}{\mathbb{E}[1 \vee R(t)]} = \frac{\mathbb{E}[\sum_{j=1}^t \delta_j (1 - \theta_j)]}{\mathbb{E}[1 \vee \sum_{j=1}^t \delta_j]}.$$

- **Testing levels of existing online multiple testing methods.** For the online methods, denote the decision rule as $\delta_t = \{p_t \leq \alpha_t\}$, where p_t is the corresponding conformal p -value at time t for our problem. The test levels $\{\alpha_t\}$ for LOND (Javanmard and Montanari, 2015), LORD++ (Ramdas et al., 2017), SAFFRON (Ramdas et al., 2018) and ADDIS (Tian and Ramdas, 2019), LORD_{dep} (Zrnic et al., 2021), SAFFRON_{dep} (Zrnic et al., 2021) are listed as follows:

1. LOND: $\alpha_t = \gamma_t(R(t-1) + 1)$, where $\{\gamma_t\}_{t=1}^\infty$ is a given infinite non-increasing sequence of positive constants that sums to α and $R(n) = \sum_{t=1}^n \delta_t$ denotes the number of discoveries in the first n hypotheses tested.

2. LORD++:

$$\alpha_t^{\text{LORD++}} = \gamma_t s_0 + (\alpha - s_0) \gamma_{t-\tau_1} \mathbb{I}\{\tau_1 < t\} + \alpha \sum_{j: \tau_j < t, \tau_j \neq \tau_1} \gamma_{t-\tau_j},$$

where $\{\gamma_t\}_{t=1}^\infty$ is a given infinite non-increasing sequence of positive constants that sums to one; τ_j is the time of the j -th rejection.

3. SAFFRON: At each time t , define $C_{j+} = C_{j+}(t) = \sum_{i=\tau_j+1}^{t-1} C_i$, where $C_t = \mathbb{I}\{p_t \leq \lambda\}$. For $t = 1$, $\alpha_1 = \min\{\gamma_1(1-\lambda)s_0, \lambda\}$; For $t = 2, 3, \dots$, $\alpha_t := \min\{\lambda, \tilde{\alpha}_t\}$, where

$$\tilde{\alpha}_t = (1-\lambda)s_0\gamma_{t-C_{0+}} + ((1-\lambda)(\alpha-s_0))\gamma_{t-\tau_1-C_{1+}} + (1-\lambda)\alpha \sum_{j \geq 2} \gamma_{t-\tau_j-C_{j+}}.$$

4. ADDIS: The testing levels for ADDIS are given by $\alpha_t = \min\{\lambda, \hat{\alpha}_t\}$, where

$$\hat{\alpha}_t = (\eta - \lambda)[\omega_0 \gamma_{S^t-C_{0+}} + (\alpha - \omega_0) \gamma_{S^t-\tau_1^*-C_{1+}} + \alpha \sum_{j \geq 2} \gamma_{S^t-\tau_j^*-C_{j+}}]$$

and $S^t = \sum_{i < t} \mathbb{I}\{p_i \leq \eta\}$, $\tau_j^* = \sum_{i \leq \tau_j} \mathbb{I}\{p_i \leq \eta\}$.

5. LORD_{dep} and SAFFRON_{dep}: Define r_k under local dependence as:

$$r_k = \min\{i \in [t] : \sum_{j=1}^{i-L_{i+1}} \delta_j \geq k\}.$$

The corresponding test levels for LORD_{dep} and SAFFRON_{dep} are as follows:

$$\alpha_t^{\text{LORD}_{\text{dep}}} = \gamma_t s_0 + (\alpha - s_0) \gamma_{t-r_1} \mathbb{I}\{r_1 < t\} + \alpha \sum_{j=2}^{\infty} \gamma_{t-r_j}.$$

$$\alpha_t^{\text{SAFFRON}_{\text{dep}}} := \min \left\{ \lambda, (1-\lambda) \left(s_0 \gamma_{t-C_{0+}} + (\alpha - s_0) \gamma_{t-r_1-C_{1+}} + \alpha \left(\sum_{j \geq 2} \gamma_{t-r_j-C_{j+}} \right) \right) \right\},$$

where $C_{j+} = \sum_{i=r_j+1}^{t-L_{j+1}} C_i$.

- **Conformal p -values.** The notion of conformal p -values was originally introduced by Vovk et al. (2005) for constructing prediction intervals. A conformal p -value quantifies how well a new observation conforms to a reference set, based on a chosen nonconformity score function. More recently, several works have applied conformal p -values to sample selection from a multiple testing perspective (Jin and Candès, 2025; Bates et al., 2023; Wang et al., 2024). The conformal p -values are defined as

$$p_t = \frac{\sum_{i \in \mathcal{C}_0} \mathbb{I}\{V_i < V_t\} + \xi_t [1 + \sum_{i \in \mathcal{C}_0} \mathbb{I}\{V_i = V_t\}]}{1 + |\mathcal{C}_0|}, \quad (16)$$

where $\mathcal{C}_0$ is a hold-out calibration dataset and $V(\cdot)$ is a nonconformity score function. However, in these approaches, the conformal p -values are constructed using a fixed offline calibration dataset, which limits their flexibility in online or adaptive settings.

A.2 Testing levels of the proposed approaches

1. Testing levels for GAIF_{dep} and Ada-GAIF_{dep}:

Define r_k under local dependence as:

$$r_k = \min\{i \in [t] : \sum_{j=1}^{i-L_{i+1}} \delta_j \geq k\}.$$

The corresponding test levels for LF_{dep} and SF_{dep} are constructed as follows:

$$\begin{aligned} \alpha_t^{\text{GAIF}_{\text{dep}}} &= \gamma_t s_0 + (\alpha - s_0) \gamma_{t-r_1} \mathbb{I}\{r_1 < t\} + \alpha \sum_{j=2}^{\infty} \gamma_{t-r_j} + \sum_{j: j \in \mathcal{I}_t, j < t-L_t} \gamma_{t-j} \alpha_j \theta_j, \\ \alpha_t^{\text{Ada-GAIF}_{\text{dep}}} &:= \min \left\{ \lambda, (1-\lambda) \left(s_0 \gamma_{t-C_{0+}} + (\alpha - s_0) \gamma_{t-r_1-C_{1+}} + \alpha \sum_{j \geq 2} \gamma_{t-r_j-C_{j+}} \right) \right. \\ &\quad \left. + \sum_{j: j \in \mathcal{I}_t, j < t-L_t} \gamma_{t-j-C_{j+}^{\#}} \alpha_j \theta_j \mathbb{I}\{p_j > \lambda\} \right\}, \end{aligned}$$

where $C_{j+} = \sum_{i=r_j+1}^{t-L_t} C_i$, $C_{j+}^{\#} = \sum_{i=j+1}^{t-L_t} C_i$.

2. Testing levels for LFS and SFS: Recall that the proposed test levels of GAIF under full and instant feedback, i.e., α_t^{LF} , are given by:

$$\alpha_t^{\text{LF}} = \gamma_t s_0 + (\alpha - s_0) \gamma_{t-\tau_1} \mathbb{I}\{\tau_1 < t\} + \alpha \sum_{j: \tau_j < t, \tau_j \neq \tau_1} \gamma_{t-\tau_j} + \sum_{j: j < t} \gamma_{t-j} \alpha_j \theta_j, \quad (17)$$

where τ_j denotes the time of the j -th rejection. We revise this rule to:

$$\alpha_t^{\text{LFS}} = \gamma_t s_0 + (\alpha - s_0) \gamma_{t-\tilde{\tau}_1} \mathbb{I}\{\tilde{\tau}_1 < t\} + \alpha \sum_{j: \tilde{\tau}_j < t, \tilde{\tau}_j \neq \tilde{\tau}_1} \gamma_{t-\tilde{\tau}_j} + \sum_{j: j < t} \gamma_{t-j} \alpha_j \theta_j, \quad (18)$$

where $\tilde{\tau}_j$ denotes the time of the j -th rejection under the null, defined as

$$\tilde{\tau}_j = \inf \left\{ t \in \mathbb{N} : \sum_{i \leq t} \delta_i (1 - \theta_i) \geq j \right\}.$$

And similarly, the test levels of SFS are:

$$\alpha_t^{\text{SFS}} := \min\{\lambda, \tilde{\alpha}_t^{\text{SFS}}\}, \quad (19)$$

where

$$\tilde{\alpha}_t^{\text{SFS}} = (1-\lambda) \left[s_0 \gamma_{t-C_{0+}} + (\alpha - s_0) \gamma_{t-\tilde{\tau}_1-C_{1+}} + \alpha \sum_{j \geq 2} \gamma_{t-\tilde{\tau}_j-C_{j+}} \right] + \sum_{j: j < t} \gamma_{t-j-C_{j+}^*} \alpha_j \theta_j \mathbb{I}\{p_j > \lambda\}.$$

A.3 Choice of non-conformity score function V

In terms of the non-conformity score function, denote $W_t = \hat{\mu}_t(\mathbf{X}_t)$, in classification settings, we set $V(W_t) = 1 - W_t$. In regression settings, if $\mathcal{A} = [b, +\infty)$, we can use $V(W_j) = b - W_j$. If $\mathcal{A} = (-\infty, a] \cup [b, +\infty)$, then we can choose $V(W_t) = \max\{W_t - a, b - W_t\}$.

Appendix B. Additional Discussions on GAIF and Adaptive GAIF

In this section, we provide additional details of our GAIF and Adaptive GAIF algorithms in Sections 2 and 3.

B.1 Explanation about utilizing feedback in SAFFRON

From another perspective, one may propose using feedback to estimate the null proportion and operate in a manner similar to SAFFRON, but in a form different from ours in Section 3.1, where the construction of α_t is

$$\alpha_t^{\text{SF-variant}} = s_0 \gamma_{t-D_{0+}} + (\alpha - s_0) \gamma_{t-\tau_1-D_{1+}} + \alpha \sum_{j \geq 2} \gamma_{t-\tau_j-D_{j+}}, \quad (20)$$

here $D_{j+} = \sum_{i=\tau_j+1}^{t-1} \mathbb{I}\{\theta_i = 1\}$. This approach constructs α_t using $\{\theta_i\}_{i=1}^{t-1}$ only, and can provide the same finite-sample online FDR guarantee.

Here, we explain why this form is not adopted in the proposed GAIF. In practice, we typically have $D_j^+ < C_j^+$, since the set filtered by p -values is larger than the true non-null set. Consequently, due to the exponential decay design of γ_t , we have

$$\gamma_{t-\tau_j-D_{j+}} \ll \gamma_{t-\tau_j-C_{j+}}.$$

Although the original SAFFRON includes an adjustment factor $1 - \lambda$ which reduces the α_t , this term becomes negligible compared to the exponential decay of $\gamma_{t-\tau_1-C_{1+}}$. As a result, the original SAFFRON achieves higher power than this naive feedback-based variant, especially when λ is large.

Our empirical results in Figure 9 further confirm this phenomenon. In the Gaussian setting described in Section 5.1 with $\mu = 2$ and $\pi_1 = 0.3$, we observe that when $\lambda > 0.3$, the original SAFFRON outperforms the variant in terms of power. Moreover, its highest power occurs at $\lambda = 0.8$, which is relatively large. This highlights the trade-off between $\gamma_{t-\tau_j-C_{j+}}$ and $1 - \lambda$, where the rapidly decaying $\gamma_{t-\tau_j-C_{j+}}$ dominates, rendering $1 - \lambda$ less influential.

In conclusion, directly using feedback to estimate the null proportion is not advisable, as it leads to a poor construction of α_t . By contrast, our SF strategy leverages the p -values to achieve adaptive α -wealth allocation and demonstrates superior performance.

B.2 Extensions of GAIF based on e -values

Although feedback cannot be directly used to improve e -LOND (Xu and Ramdas, 2024), it can enhance e -LORD and e -SAFFRON (Zhang et al., 2025) through a feedback-driven approach analogous to the extension from GAI to GAIF. Denote $\delta_j = \mathbb{I}\{e_j \geq 1/\alpha_j\}$. Denote

$$\text{FDP}^*(t) = \sum_{j \leq t, j \in \mathcal{H}_0} \frac{\alpha_j}{1 + R(j-1)} \leq \sum_{j \leq t} \frac{\alpha_j}{1 + R(j-1)} = \widehat{\text{FDP}}_{e\text{-LORD}}.$$

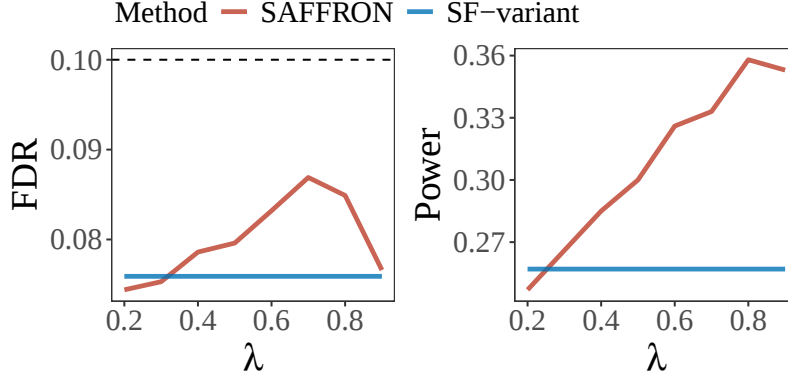


Figure 9: The FDR and Power for SAFFRON at stopping time 600 under different λ value for target FDR level $\alpha = 0.1$.

Similar to the GAIF framework, we propose the estimates $\widehat{\text{FDP}}_{\text{e-GAIF}}$ and $\widehat{\text{FDP}}_{\text{e-Ada-GAIF}}$ as follows:

$$\text{FDP}^*(t) = \frac{\sum_{j \leq t, j \in \mathcal{H}_0} \alpha_j}{1 \vee \sum_{j \leq t} \delta_j} \leq \widehat{\text{FDP}}_{\text{e-GAIF}} = \sum_{j \in \mathcal{I}_t} \frac{\alpha_j(1 - \theta_j)}{1 + R(j - 1)} + \sum_{j \in \bar{\mathcal{I}}_t} \frac{\alpha_j}{1 + R(j - 1)}.$$

$$\widehat{\text{FDP}}(t)_{\text{e-Ada-GAIF}} = \sum_{j \in \mathcal{I}_t} \frac{\alpha_j(1 - \theta_j)}{1 + R(j - 1)} \frac{\mathbb{I}\{e_j \leq 1/\lambda\}}{1 - \lambda} + \sum_{j \in \bar{\mathcal{I}}_t} \frac{\alpha_j}{1 + R(j - 1)} \frac{\mathbb{I}\{e_j \leq 1/\lambda\}}{1 - \lambda}.$$

Then we require $\widehat{\text{FDP}}_{\text{e-GAIF}} \leq \alpha$ or $\widehat{\text{FDP}}_{\text{e-Ada-GAIF}} \leq \alpha$ when constructing the testing levels. The corresponding testing levels are as follows:

$$\alpha_t^{\text{e-GAIF}} = \omega_t \left(\alpha - \sum_{j \in \mathcal{I}_t} \frac{\alpha_j(1 - \theta_j)}{1 + R(j - 1)} - \sum_{j \in \bar{\mathcal{I}}_t} \frac{\alpha_j}{1 + R(j - 1)} \right) (R(t - 1) + 1), \quad (21)$$

$$\alpha_t^{\text{e-Ada-GAIF}} = \omega_t \left(\alpha(1 - \lambda) - \sum_{j \in \mathcal{I}_t} \frac{\alpha_j(1 - \theta_j) \mathbb{I}\{e_j < 1/\lambda\}}{1 + R(j - 1)} - \sum_{j \in \bar{\mathcal{I}}_t} \frac{\alpha_j \mathbb{I}\{e_j < 1/\lambda\}}{1 + R(j - 1)} \right) (R(t - 1) + 1), \quad (22)$$

where $\omega_t \in (0, 1)$ is updated by

$$\omega_{t+1} = \omega_t + \omega_1 \varphi^{t-R(t)} (1 - \delta_t) - \omega_1 \psi^{R(t)} \delta_t \quad (23)$$

with a user-defined initial allocation coefficient $\omega_1 \in (0, 1)$, and user-defined parameters $\varphi > 0, \psi > 0$.

The Generalized Alpha-Investing with Feedback procedure based on e -values is summarized in Algorithm 3. The e -GAIF and e -Ada-GAIF can control online FDR validly if the null e -values are conditional valid.

Algorithm 3 e -GAIF (e -GAIF and e -Ada-GAIF)

Input: Target FDR level α , parameters λ, φ and $\psi \in (0, 1)$, initial allocation coefficient $\omega_1 \in (0, 1)$.

```

1: for  $t = 1, 2, \dots$  do
2:   Observe  $e$ -value  $e_t$ 
3:   Update  $\alpha_t = \alpha_t^{\text{e-GAIF}}$  in Eq.(21) (or  $\alpha_t = \alpha_t^{\text{e-Ada-GAIF}}$  in Eq.(22))
4:   if  $e_t \geq 1/\alpha_t$  then  $\delta_t = 1$ , else  $\delta_t = 0$ 
5:   Update  $R(t) = R(t-1) + \delta_t$  and  $\omega_{t+1}$  by (23)
6:   Obtain the revealed feedback  $\theta_t$ 
7: end for
    
```

Output: Rejection set $\{t : \delta_t = 1\}$.

Theorem 16 (Online FDR control for e -GAIF and e -Ada-GAIF). *(a.) If the null e -values are conditional valid, i.e.,*

$$\mathbb{E}[e_t \mid \mathcal{F}_{t-1}] \leq 1 \text{ for all } t \in \mathcal{H}_0, \quad (24)$$

where $\mathcal{F}_{t-1} = \sigma(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})$ is the sigma field generated from past rejections and feedback, then if the parameters $\{\alpha_t\}_{t \in \mathbb{N}}$ are selected such that $\widehat{\text{FDP}}_{\text{e-GAIF}}(t) \leq \alpha$, then we have

$$\text{FDR}(t) \leq \alpha \text{ for all } t \in \mathbb{N}.$$

(b.) Denote $C_t = \mathbb{I}\{e_t < 1/\lambda\}$. If the null e -values are conditional valid, i.e.,

$$\mathbb{E}[e_t \mid \mathcal{J}_{t-1}] \leq 1 \text{ for all } t \in \mathcal{H}_0, \quad (25)$$

where $\mathcal{J}_{t-1} = \sigma(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t}; C_{1:t-1})$ is the enlarged filtration, then if the parameters $\{\alpha_t\}_{t \in \mathbb{N}}$ are selected such that $\widehat{\text{FDP}}_{\text{e-Ada-GAIF}}(t) \leq \alpha$, then we have

$$\text{FDR}(t) \leq \alpha \text{ for all } t \in \mathbb{N}.$$

B.3 Discussion about improving LOND/ e -LOND with feedback

In this subsection, we discuss the feasibility of improving LOND or e -LOND through our feedback-enhanced framework.

Recall the setup of the LOND or e -LOND algorithm. Given a non-negative sequence $\{\gamma_j\}_{j=1}^\infty$ such that $\sum_{j=1}^\infty \gamma_j = 1$, the test levels are set as

$$\alpha_t^{\text{LOND}} = \alpha \gamma_t \left(\sum_{j=1}^{t-1} \delta_j \vee 1 \right) = \alpha \gamma_t (R(t-1) \vee 1).$$

Zrnic et al. (2021) proved that LOND controls the FDR under PRDS.

Improve LOND or e -LOND with feedback? (negative result) Here we prove that the feedback cannot be used to improve LOND and e -LOND.

Denote $a_j = \sum_{i=1}^{j-1} \delta_i \theta_i$ and $b_j = \sum_{i=1}^{j-1} \delta_i (1 - \theta_i)$, so that $R(j-1) = a_j + b_j$. For LOND or e -LOND, the significance level is given by

$$\alpha_j = \alpha \gamma_j (a_j + b_j + 1).$$

If $\theta_1, \dots, \theta_{j-1}$ are known, we consider the modified significance level

$$\tilde{\alpha}_j = \alpha \gamma_j (2w_j a_j + 2(1 - w_j)b_j + 1),$$

where $w_j \in [0, 1]$ is a weight that differentiates between true and false discoveries. Setting $w_j = 1/2$ recovers $\tilde{\alpha}_j = \alpha_j$. To control $\text{FDR}(t) \leq \alpha$, we require

$$\sum_{j \in \mathcal{H}_0 \cap [t]} \frac{\tilde{\alpha}_j}{R(j-1) + 1} \leq \alpha.$$

This implies

$$2w_j a_j + 2(1 - w_j)b_j + 1 \leq a_j + b_j + 1.$$

To improve power, we require

$$\tilde{\alpha}_j > \alpha_j,$$

which translates to

$$2w_j a_j + 2(1 - w_j)b_j + 1 > a_j + b_j + 1.$$

However, this contradicts the FDR control condition, making the approach infeasible.

Appendix C. Additional Discussions on OCTF

This section collects several supplementary discussions and extensions of OCTF that complement Section 4. These include delayed feedback (Appendix C.1), a general framework for score selection (Appendix C.2), extensions to online-updated score functions (Appendix C.3), and the discussion of OCTF under null distribution shift (Appendix C.4).

C.1 Online conformal selection with delayed Feedback

In this section, we study how to adapt OCTF to settings with delayed feedback with rigorous guarantees. We focus on the case of a fixed delay d , where at time t only the feedback $\{\theta_j\}_{j=1}^{t-d-1}$ is available. To maintain finite-sample mFDR control under delayed feedback, both the construction of the test levels α_t and the calibration data sets must be modified accordingly.

Specifically, note that the feedback associated with test t is revealed at time $t + d + 1$. To accommodate this delay, we partition the time indices into $(d + 1)$ disjoint sub-streams:

$$\begin{aligned} S_1 &= \{1, 2 + d, 3 + 2d, 4 + 3d, \dots\}, \\ S_2 &= \{2, 3 + d, 4 + 2d, 5 + 3d, \dots\}, \\ &\vdots \\ S_{d+1} &= \{1 + d, 2 + 2d, 3 + 3d, 4 + 4d, \dots\}. \end{aligned}$$

Within each sub-stream, the feedback for a test is revealed at the next time point belonging to the same sub-stream. For each sub-stream $j \in \{1, 2, \dots, d + 1\}$, we maintain an independent online multiple testing procedure based on safe GAIF with an adjusted target level $\alpha/(d + 1)$,

which is used to compute and update the test levels for that sub-stream. In parallel, we maintain a separate calibration data set

$$\mathcal{C}_{0t}^j = \{i < t : i \in S_j, \theta_i = 0\}.$$

For a test observation $\mathbf{X}_t$ arriving at time $t \in S_j$, the test level α_t is determined by the current state of the GAIF procedure associated with sub-stream j . The corresponding online conformal p -value is defined as

$$p_t = \frac{\sum_{i \in \mathcal{C}_{0t}^j} \mathbb{I}\{V_i < V_t\} + \xi_t \left(1 + \sum_{i \in \mathcal{C}_{0t}^j} \mathbb{I}\{V_i = V_t\}\right)}{1 + |\mathcal{C}_{0t}^j|}, \quad (26)$$

where $\xi_t \stackrel{\text{i.i.d.}}{\sim} \text{Unif}[0, 1]$ is an independent random variable used for tie-breaking.

The resulting online conformal testing procedure with delayed feedback, referred to as OCTF-delay, is summarized in Algorithm 4. Finite-sample mFDR control for OCTF-delay is established in Corollary 17.

Corollary 17 (Finite-sample online mFDR control for OCTF-delay). *Suppose that Assumption 1 holds. Then the rejection set $\mathcal{R}$ produced by Algorithm 4 satisfies*

$$\text{mFDR}(t) \leq \alpha \quad \text{for all } t \in \mathbb{N}.$$

The proof relies on a sub-stream decomposition argument. Under fixed delay d , the time indices are partitioned into $(d + 1)$ disjoint sub-streams. Within each sub-stream, feedback for a test is revealed before the next test in the same sub-stream, so the corresponding online conformal p -values are super-uniform and independent under the null conditional on the sub-stream filtration. Applying safe GAIF separately to each sub-stream therefore guarantees finite-sample mFDR control at level $\alpha/(d + 1)$ within every sub-stream. Since the sub-streams are disjoint, summing the corresponding bounds and using $\sum_{r=1}^{d+1} (1 \vee R_r(t)) \leq (d + 1)(1 \vee R(t))$ yields $\mathbb{E}[V(t)] \leq \alpha \mathbb{E}[1 \vee R(t)]$, and therefore the combined procedure controls the overall mFDR at level α .

We evaluate the proposed procedures in Scenario IV under delayed feedback. We compare the sub-stream-based procedures with direct delayed-feedback implementations. At time t , these implementations form the calibration set from the null samples whose feedback has already been revealed, compute the corresponding online conformal p -value, and then apply LF-FD or SF-FD without the sub-stream modification.

Figure 10 reports the results. All sub-stream-based procedures (SF-sub, LF-sub, SFS-sub, and LFS-sub) maintain empirical FDR below the target level under delayed feedback. Notably, SF-sub and LF-sub retain substantial power despite the sub-stream partition. Although the direct SF-FD and LF-FD implementations do not admit a finite-sample guarantee, they also maintain empirical FDR control in these simulations and attain higher power than their sub-stream-based counterparts. Thus, the direct implementations may be attractive when empirical power is the main concern and the lack of guarantee is acceptable, whereas the sub-stream construction provides a principled alternative when rigorous finite-sample error control is required.

Algorithm 4 Online conformal testing with delayed feedback (OCTF-delay)

Input: Initial calibration data $\mathcal{D}_C = \{(\mathbf{X}_i, Y_i)\}_{i=-n+1}^0$, target region $\mathcal{A}$, FDR target level $\alpha \in (0, 1)$, non-conformity score function $V(\cdot)$, initial wealth s_0 , sequence $\{\gamma_t\}_{t \in \mathbb{N}}$, delay d .

- 1: Partition $\mathbb{N}$ into $(d+1)$ sub-streams $\{S_j\}_{j=1}^{d+1}$.
- 2: For each sub-stream j , initialize an independent GAIF procedure with initial wealth $s_0/(d+1)$ and target mFDR level $\alpha/(d+1)$.
- 3: Initialize calibration sets $\mathcal{C}_{01}^j = \{i \leq 0 : \theta_i = 0\}$ for $j = 1, \dots, d+1$.
- 4: **for** $t = 1, 2, \dots$ **do**
- 5: Observe test data $\mathbf{X}_t$
- 6: Let j be the index such that $t \in S_j$
- 7: Compute conformity score V_t
- 8: Compute conformal p -value p_t using $\mathcal{C}_{0t}^j$ via (26)
- 9: Compute test level α_t using the SFS (or LFS) update rule of sub-stream j at level $\alpha/(d+1)$
- 10: **if** $p_t \leq \alpha_t$ **then** $\delta_t = 1$ **else** $\delta_t = 0$
- 11: Record decision δ_t
- 12: **if** feedback for some $i < t$ is revealed at time t :
- 13: update θ_i and the GAIF state of the corresponding sub-stream
- 14: **if** $\theta_i = 0$, add i to $\mathcal{C}_{0t}^j$
- 15: **end for**

Output: Rejection set $\mathcal{R} = \{t : \delta_t = 1\}$.

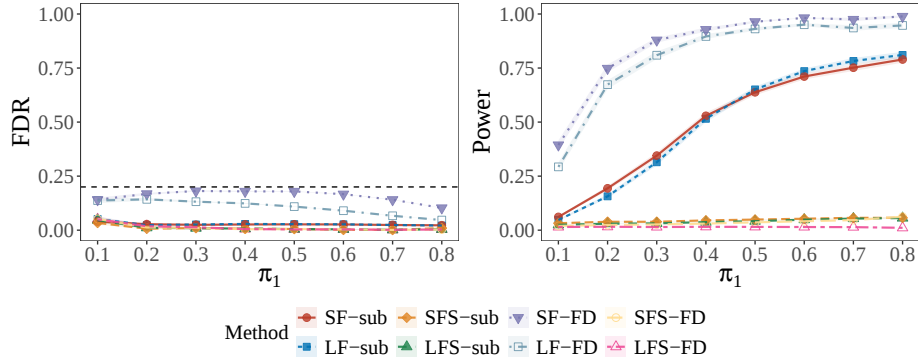


Figure 10: Results for Scenario IV with delayed feedback: Line charts of FDR and Power with varying non-null proportion π_1 from 0.1 to 0.8 after 500 replications; The black dashed lines denote the target FDR level $\alpha = 0.2$. Delay $d = 5$.

C.2 A general validity framework and practical principles for optimized OCTF

We present a general validity-preserving framework for online conformal testing after score selection. The key idea is that, even when the score selection criterion is arbitrary, one can still restore the symmetry required for conformal validity by applying the same score selection rule under all permutations of the current null calibration set and test point.

At time t , let Ω_t denote the set of all permutations of $\mathcal{C}_{0t} \cup \{t\}$ that leave indices outside $\mathcal{C}_{0t} \cup \{t\}$ unchanged. For each $\sigma \in \Omega_t$, we apply the same score selection rule to the permuted data $(\mathcal{D}_t)_\sigma$ and obtain

$$\hat{k}_t^\sigma = \arg \min_{k \in [K]} \mathcal{M}((\mathcal{D}_t)_\sigma; k).$$

We then define the permutation-based conformal p -value

$$p_t^{\text{opt}} = \frac{1}{|\Omega_t|} \left[\sum_{\sigma \in \Omega_t} \mathbb{I}\{V(\mathbf{X}_{\sigma(t)}; \hat{k}_t^\sigma) < V(\mathbf{X}_t; \hat{k}_t)\} + \xi_t \sum_{\sigma \in \Omega_t} \mathbb{I}\{V(\mathbf{X}_{\sigma(t)}; \hat{k}_t^\sigma) = V(\mathbf{X}_t; \hat{k}_t)\} \right]. \quad (27)$$

Intuitively, under each permutation, the candidate test point $\sigma(t)$ is evaluated using the score function that would have been selected under that same permutation. In this way, the score selection step and the conformal ranking step are coupled symmetrically across all elements of $\mathcal{C}_{0t} \cup \{t\}$, restoring the exchangeability needed for conformal validity.

With this construction, arbitrary score selection criteria can be incorporated without sacrificing finite-sample error control.

Theorem 18 (Finite-sample online mFDR control based on full permutations). *Under the same assumptions as in Theorem 11, for any score selection criterion $\mathcal{M}$, the procedure in Algorithm 2 with $\{p_t\}$ replaced by $\{p_t^{\text{opt}}\}$ satisfies*

$$\text{mFDR}(t) \leq \alpha, \quad \forall t \in \mathbb{N}.$$

Practical principles for score selection Theorem 18 shows that online conformal testing with score selection is possible in full generality. However, this result is mainly a theoretical benchmark: the full-permutation correction can be computationally infeasible in online settings and may also degrade power when the selected score function varies substantially across permutations. This observation motivates the search for score selection criteria that are not only valid in principle, but also practically effective.

Guided by this perspective, we highlight three principles that a practically useful criterion $\mathcal{M}(\mathcal{D}_t; \cdot)$ should satisfy.

- **Power alignment.** At time t , a natural oracle objective for power is to choose the score function k maximizing

$$\Pr(p_{k,t} \leq \alpha_t \mid \theta_t = 1).$$

A useful criterion should therefore be aligned with this target, in the sense that models with stronger non-null rejection ability receive better scores.

- **Permutation robustness.** The criterion should not vary drastically under permutation of $\mathcal{C}_{0t} \cup \{t\}$. Otherwise, the selected score function $\hat{k}_t^\sigma$ may differ substantially across σ , so that p_t^{opt} compares scores computed under different score models. Such cross-model comparisons can inject extra variability, weaken the non-null signal, and reduce power. A sufficient condition avoiding this issue is that $\mathcal{M}(\mathcal{D}_t; \cdot)$ is symmetric with respect to $\mathcal{C}_{0t} \cup \{t\}$, in which case $\hat{k}_t^\sigma \equiv \hat{k}_t$ for all σ .
- **Computational efficiency.** In principle, full permutation requires enumerating $|\Omega_t| = (|\mathcal{C}_{0t}| + 1)!$, which is prohibitive in online applications. Criteria with sufficient symmetry, especially permutation-invariant ones, allow one to avoid explicit enumeration and reduce computation essentially to the standard conformal ranking over $|\mathcal{C}_{0t}| + 1$ scores.

Alternative criteria These principles clarify why some seemingly natural criteria are unattractive in practice.

First, threshold-dependent criteria involving α_t typically violate both permutation robustness and computational efficiency. Since α_t is generated by the online testing rule and depends on the past rejection trajectory, such criteria are generally not symmetric with respect to $\mathcal{C}_{0t} \cup \{t\}$. Moreover, a full-permutation correction would require recomputing the threshold process under each permutation, making the procedure both computationally burdensome and statistically unstable.

A second illustrative example is the criterion $\mathcal{M}(\mathcal{D}_t; k) = p_{k,t}$. At first glance, this choice appears attractive, since smaller p -values are more likely to lead to rejection. However, if used directly for score selection, it leads to severe validity issues: under the null, the procedure selects the minimum of K candidate p -values, which is generally not super-uniform.

Although Theorem 18 implies that validity can still be restored through full permutation, this criterion remains practically problematic. Indeed, under this choice, for each permutation $\sigma \in \Omega_t$,

$$\mathcal{M}((\mathcal{D}_t)_\sigma; k) = \frac{\sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{X}_i; k) < V(\mathbf{X}_{\sigma(t)}; k)\} + \xi_t \cdot \sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{X}_i; k) = V(\mathbf{X}_{\sigma(t)}; k)\}}{1 + |\mathcal{C}_{0t}|}.$$

Hence, when $\sigma(t) \neq t$ and denote $i = \sigma(t)$, the quantity $\mathcal{M}((\mathcal{D}_t)_\sigma; k)$ is simply a null conformal p -value for $V(\mathbf{X}_i; k)$ based on $\{V(\mathbf{X}_s; k)\}_{s \in \mathcal{C}_{0t} \cup \{t\}}$. Score selection is then driven by null p -values, which contain little useful information for identifying a powerful model. As a result, $\hat{k}_t^\sigma$ may fluctuate nearly randomly across permutations, and the final p_t^{opt} in (27) is formed by ranking scores obtained under different selected score functions, possibly on different scales. This can substantially degrade power; empirically, we find that such instability may even make the procedure perform worse than random score selection.

In contrast, our EWMA criterion is designed to satisfy the above three principles: it is aligned with power through a threshold-free proxy, robust to the required permutation symmetry, and computationally efficient in online implementation. For this reason, although Theorem 18 provides a general validity result for arbitrary criteria, we recommend the EWMA criterion for practical use.

Our framework also allows other power-aligned surrogates. For example, one may consider oracle-aligned objectives of the form $\mathbb{E}[\phi(p_{k,t}) \mid \theta_t = 1]$, where $\phi(\cdot)$ is a deterministic nondecreasing function that emphasizes small p -values, such as $\phi(p) = \min(p, c)$ for some constant c . Such choices could move the objective closer to the tail probability $\Pr(p_{k,t} \leq \alpha_t \mid \theta_t = 1)$. However, selecting c appropriately, especially when α_t varies over time, would require additional tuning and theoretical analysis. In this paper, we therefore focus on $\mathcal{M}_t^{\text{EWMA}}$ as a simple choice.

C.3 Optimized OCTF with online-updated score functions

Beyond selecting from a pool of pre-trained score functions, our framework also accommodates *data-driven score functions* that are updated online. The key requirement is the same as in Proposition 22: at time t , the score function used to construct the conformal p -value must be symmetric with respect to the samples in $\mathcal{C}_{0t} \cup \{t\}$. Equivalently, when fitting the score model for the t -th test point, the current sample must be inserted as a provisional null sample and treated exchangeably with the historical null samples in $\mathcal{C}_{0t}$.

We focus on the binary classification setting with $Y \in \{0, 1\}$. The regression case can be handled similarly by replacing the response with the binary label $\mathbb{I}\{Y \in \mathcal{A}\}$. At each time t , for each candidate setup $k \in [K]$, we fit a one-step-ahead predictive model $\hat{\mu}_k^{(t)}$ by distinguishing the non-null samples $\{(\mathbf{X}_i, Y_i)\}_{i \in \mathcal{C}_{1t}}$ from the null samples $\{(\mathbf{X}_i, Y_i)\}_{i \in \mathcal{C}_{0t}} \cup \{(\mathbf{X}_t, 0)\}$. The associated score function is then given by $V(\cdot; \hat{\mu}_k^{(t)})$.

This construction is valid as long as the training algorithm is symmetric to its input data (or more precisely, symmetric to the treated null samples $\{(\mathbf{X}_i, Y_i)\}_{i \in \mathcal{C}_{0t}} \cup \{(\mathbf{X}_t, 0)\}$), which is satisfied by most batch learning procedures, such as random forests and neural networks. In contrast, order-dependent updating rules, such as standard online gradient descent, generally violate this symmetry requirement and therefore cannot be directly used here.

In this perspective, we can extend our framework of score selection. The K candidate setups may correspond to different score function classes or to the same score function with different hyper-parameters. When $K = 1$, the procedure reduces to pure online model/score updating. When $K > 1$, we update all K score functions in parallel and then select the score function used for inference at time t . Specifically, we choose

$$\hat{k}_t = \arg \min_{k \in [K]} \mathcal{M}_t^{\text{EWMA}}(\mathcal{D}_t; k),$$

where $\mathcal{M}_t^{\text{EWMA}}(\mathcal{D}_t; k)$ is the same evaluation criterion as in Section 4.3. The optimized conformal p -value is then constructed using the selected score function based on $\hat{\mu}_{\hat{k}_t}^{(t)}$. Since each candidate score function is symmetric in $\mathcal{C}_{0t} \cup \{t\}$, and the selection criterion depends only on permutation-invariant quantities, the selected score function remains symmetric as well. Therefore, the same argument as in Corollary 13 yields finite-sample mFDR control.

The detailed pseudo-code is given in Algorithm 5.

Although online updating of score functions improves data utilization and adaptivity, it is substantially more expensive than selecting from a fixed pool of pre-trained models. To preserve the required symmetry, each candidate model typically needs to be refit from scratch after adding the provisional null sample $(\mathbf{X}_t, 0)$. For this reason, in the main text we focus on score selection among pre-trained score functions, and view online updating as a valid but computationally heavier extension.

Experimental Evaluations We illustrate this through an additional experiment. In particular, we compare our approach with an alternative procedure that retrain the score function (predictive model) at every time step using all previously observed data. We adopt the same data-generating process as Scenario IV with sine pattern shift in Section 5.3. We set the target level $\alpha = 0.2$.

Consider three strategies for constructing the predictive score function:

- **Fixed score function.** A score function based on Random Forest is trained once using historical data and then kept fixed throughout the online testing process.
- **Score selection (ours).** We pre-train $K = 3$ candidate score functions (Neural network, SVM, and Random Forest) using historical data, which induce K candidate score functions. During the online testing process, the algorithm dynamically selects one candidate at each time step according to the feedback-based EWMA criterion described in Section 4.3.

Algorithm 5 Optimized OCTF with online-updated score functions

Input: Initial data $\mathcal{D}_1 = \{(\mathbf{X}_i, Y_i)\}_{i=-N}^0$, target region $\mathcal{A}$, FDR target level $\alpha \in (0, 1)$, K candidate model setups, evaluation criterion $\mathcal{M}$, parameter s_0 , parameter sequence $\{\gamma_t\}$, stopping time T .

- 1: **for** $t = 1, \dots, T$ **do**
- 2: Observe test covariate $\mathbf{X}_t$.
- 3: **for** $k = 1, \dots, K$ **do**
- 4: Fit the one-step-ahead model $\hat{\mu}_k^{(t)}$ by distinguishing $\{(\mathbf{X}_i, Y_i)\}_{i \in \mathcal{C}_{1t}}$ from $\{(\mathbf{X}_i, Y_i)\}_{i \in \mathcal{C}_{0t}} \cup \{(\mathbf{X}_t, 0)\}$.
- 5: Construct the score function $V(\cdot; \hat{\mu}_k^{(t)})$.
- 6: **end for**
- 7: Select the predictive model by

$$\hat{k}_t = \arg \min_{k \in [K]} \mathcal{M}_t^{\text{EWMA}}(\mathcal{D}_t; k).$$

- 8: Construct the optimized conformal p -value

$$p_t^{\text{opt}} = \frac{\sum_{i \in \mathcal{C}_{0t}} \mathbb{I}\{V(\mathbf{X}_i; \hat{\mu}_{\hat{k}_t}^{(t)}) < V(\mathbf{X}_t; \hat{\mu}_{\hat{k}_t}^{(t)})\} + \xi_t \left[1 + \sum_{i \in \mathcal{C}_{0t}} \mathbb{I}\{V(\mathbf{X}_i; \hat{\mu}_{\hat{k}_t}^{(t)}) = V(\mathbf{X}_t; \hat{\mu}_{\hat{k}_t}^{(t)})\}\right]}{1 + |\mathcal{C}_{0t}|}.$$

- 9: Update $\alpha_t = \alpha_t^{\text{LFS}}$ in (18) or $\alpha_t = \alpha_t^{\text{SFS}}$ in (19).
 - 10: Make a decision $\delta_t = \mathbb{I}\{p_t^{\text{opt}} \leq \alpha_t\}$.
 - 11: Observe the revealed feedback Y_t (and hence θ_t).
 - 12: Update the calibration sets $\mathcal{C}_{0t}$ and $\mathcal{C}_{1t}$.
 - 13: **end for**
- Output:** Rejection set $\mathcal{R}_{\text{opt}} = \{i : \delta_i = 1, \delta_i \in \delta^T\}$.
-

- **Online retraining/updating.** At each time step t , the score function based on Random Forest is retrained using all previously observed data. To maintain the validity of conformal p -values, samples with feedback $\theta_t = 0$ are incorporated into the null calibration set in the same way as other null observations.

For each strategy, we evaluate the empirical FDR and statistical power across the testing horizon. In addition, we record the total running time required to process a sequence of $T = 300$ test samples, which measures the computational cost of the corresponding score construction method.

Table 2 reports the empirical FDR and power for the three strategies. All methods control the empirical FDR at the target level. The online retraining (score updating) strategy achieves slightly higher power than the fixed model, but its improvement over the proposed score-selection approach is marginal. From a computational perspective, however, the difference is substantial. Online retraining requires fitting a predictive model at every time step, leading to a computational cost that grows linearly with the length of the testing sequence. In contrast, the proposed method only evaluates $K = 3$ pre-trained candidate score functions and updates the EWMA statistics, which incurs negligible additional overhead.

Table 2: Comparison of score construction strategies for SF.

Strategy	FDR	Power	Runtime (s)
Fixed score function	0.162(0.011)	0.829(0.035)	0.621(0.018)
Score selection (ours)	0.161(0.009)	0.837(0.031)	0.644(0.028)
Online retraining	0.179(0.010)	0.866(0.026)	122.000(1.650)

C.4 OCTF under null distribution shifts

In this subsection, we examine how shifts in the null distribution affect OCTF and discuss possible remedies.

Performance evaluation under null distribution shift To assess robustness under non-stationarity, we consider a binary classification setting similar to Scenario IV in Section 5.2, but introduce a time-varying drift in the null distribution. The initial training and calibration sets are sampled from the distribution at time $t = 0$, while the test stream evolves over time:

- Null ($Y_t = 0$): $\mathbf{X}_t \sim \mathcal{N}(\boldsymbol{\mu}_t, \mathbf{I}_4)$, where

$$\boldsymbol{\mu}_t = \left(2 - \Delta \frac{t}{T}, 0, 0, 0\right)^\top,$$

and $\Delta \in \mathbb{R}$ controls the drift magnitude.

- Alternative ($Y_t = 1$): $\mathbf{X}_t \sim \mathcal{N}(\boldsymbol{\mu}_1, \mathbf{I}_4)$ with fixed mean

$$\boldsymbol{\mu}_1 = (0, 0, 0, 0)^\top.$$

We set $T = 1000$ and $\pi_1 = 0.5$. To adapt to distribution shift, a natural idea is to use a sliding calibration window. We therefore compare:

- **Growing window:** $\mathcal{C}_{0t} = \{i < t : \theta_i = 0\}$.
- **Sliding window:** $\mathcal{C}_{0t} = \{i : t - W \leq i < t, \theta_i = 0\}$, with $W = 200$.

Under exchangeability, the growing window preserves the mutual independence of null online conformal p -values in Proposition 10, whereas the sliding window introduces local dependence among consecutive conformal p -values.

We consider two drift directions, $\Delta \in \{-2, 2\}$. The results are shown in Figures 11–12. Our main findings are:

- **Dual Effects of Drift:** Null drift does not necessarily invalidate OCTF. When the null moves away from the alternative, the procedure remains valid but becomes conservative. When the null moves toward the alternative, FDR inflation may occur.
- **Sliding windows may hurt validity.** Although sliding windows are intended to improve adaptation, in our experiments they exacerbate FDR inflation relative to the growing-window baseline.

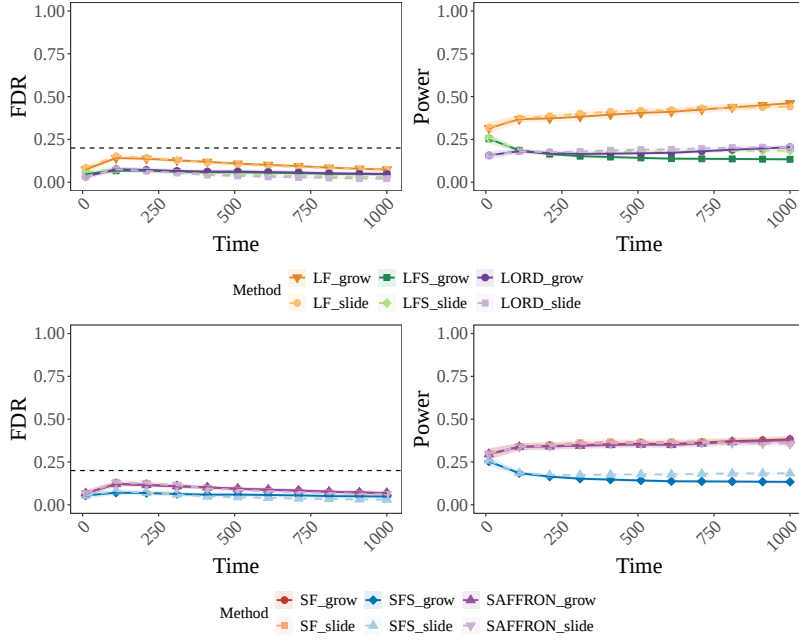


Figure 11: Results under null distribution shift: Line charts of FDR and Power across time t ; The black dashed lines denote the target FDR level $\alpha = 0.2$. We set the drift parameter $\Delta = -2$ so that the null shifts away from the non-null.

We attribute this to the loss of independence: the sliding window creates strong local dependence among nearby conformal p -values, and standard OCTF rules applied to such dependent p -values can inflate FDR. This suggests that simply discarding old data is not enough; the gain in adaptation may be outweighed by the loss of error-rate control. In contrast, the dependence-corrected LF_{dep} variant restores control, albeit at a substantial power cost.

Online weighted conformal p -values for handling null covariate shift These observations suggest that handling null distribution shift requires more principled corrections, such as *online weighted conformal p -values*. However, obtaining exactly valid online p -values becomes highly nontrivial when the null distribution changes over time.

To illustrate this difficulty, consider a stylized setting in which all observations up to time t are null, but follow different distributions:

$$V_s \sim F_s, \quad s = 1, \dots, t,$$

with $F_s \neq F_{s'}$ in general. Let Ω_t denote the set of all permutations of $\{1, \dots, t\}$. Given an unordered realization $\{V_1, \dots, V_t\} = \{v_1, \dots, v_t\}$, an oracle posterior over permutations can be defined as

$$w(\{v_i\}_{i=1}^t; \sigma) \propto \prod_{s=1}^t f_s(v_{\sigma(s)}), \quad \sigma \in \Omega_t,$$

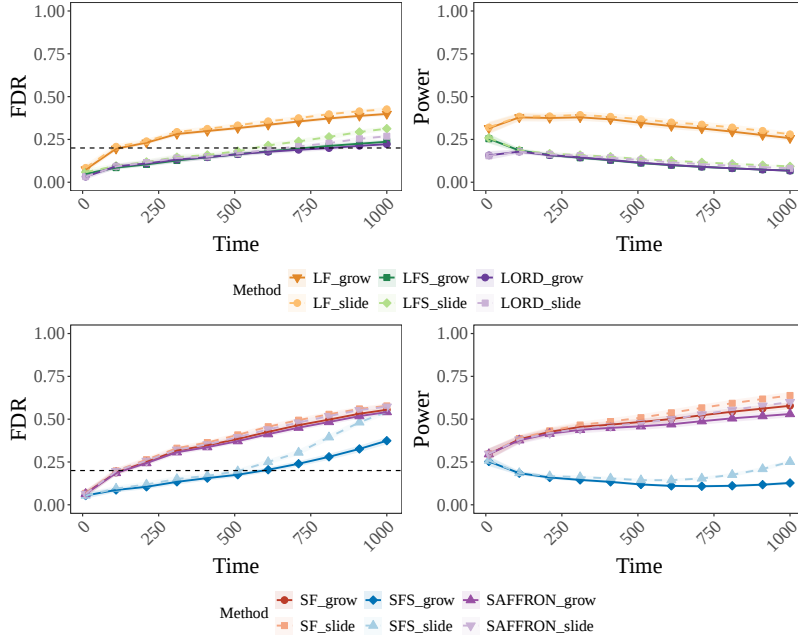


Figure 12: Results under null distribution shift: Line charts of FDR and Power across time t ; The black dashed lines denote the target FDR level $\alpha = 0.2$. We set the drift magnitude $\Delta = 2$ so that the null drifts towards the alternative/non-null region.

where f_s is the null density of F_s and the weights are normalized so that $\sum_{\sigma \in \Omega_t} w(\{v_i\}_{i=1}^t; \sigma) = 1$. Intuitively, $w(\{v_i\}; \sigma)$ quantifies how plausible it is that the time labels $1, \dots, t$ are assigned to the observed $\{v_i\}_{i=1}^t$ according to σ .

Using this oracle weighting, an online conformal p -value could be constructed by averaging the randomized rank over permutations:

$$p_t^w = \sum_{\sigma \in \Omega_t} w(\{V_i\}_{i=1}^t; \sigma) (\mathbb{I}\{V_{\sigma(t)} < V_t\} + \xi_t \mathbb{I}\{V_{\sigma(t)} = V_t\}).$$

In principle, such oracle weighting can restore super-uniformity and mutual independence by integrating over the uncertainty induced by non-identical null distributions. In offline covariate-shift settings, this weighting may simplify substantially (Jin and Candès, 2025); for example, if only the test point has a different null distribution while the calibration points remain i.i.d., the weight depends only on the identity of the test point. In the online setting, however, when all $F_1, \dots, F_t$ may differ, the weight generally depends on the entire permutation, leading to factorial computational complexity. Moreover, estimating the full sequence of drifting null distributions is itself difficult, making practical implementations unstable.

An empirical investigation based on weighted conformal p -values A practical workaround is to maintain a fixed calibration set drawn from a stable null distribution and compute weighted conformal p -values relative to this reference set, similar to the offline covariate-shift approach in Prinster et al. (2025), although a rigorous theoretical guarantee

remains open. We conduct a real-data application to verify the performance of adopting online weighted conformal p -values, where oracle weights are used.

We use the airfoil dataset (Brooks et al., 2014) from the UCI Machine Learning Repository. The task is to identify samples with high sound pressure. The dataset contains $n = 1503$ observations, where the response Y is the scaled sound pressure and the covariates $\mathbf{X} \in \mathbb{R}^5$ include log frequency, angle of attack, chord length, free-stream velocity, and suction-side log displacement thickness. We test

$$\mathbb{H}_{0t} : Y_t \in (-\infty, c),$$

where c is the $(1 - \pi_1)$ -quantile of Y , and $\pi_1 \in \{0.2, \dots, 0.8\}$ controls the non-null proportion.

To simulate covariate shift, we construct a shifted test set D_{shift} by resampling 25% of the points from D_{test} with probabilities proportional to

$$\omega(x) = \exp(x^\top \beta), \quad \beta = (2, 0, 0, 0, -3).$$

Following Prinster et al. (2025), we use a fixed calibration set and compare weighted conformal p -values with their unweighted counterparts. Figure 13 reports the results for weighted methods (wSFS, wLFS, wSF, and wLF) and unweighted methods (SFS, LFS, SF, and LF). The weighted safe procedures wSFS and wLFS consistently maintain valid empirical FDR control under covariate shift, whereas SF, LF, and SFS suffer from FDR inflation due to their reliance on unweighted conformal p -values.

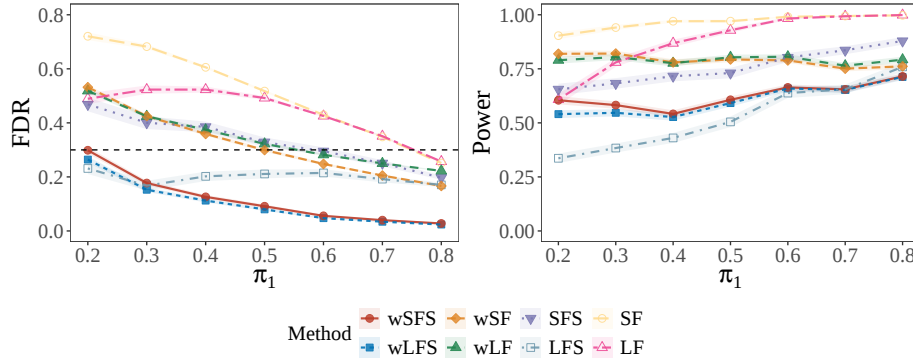


Figure 13: Results for airfoil data: the values of $\text{FDR}(T)$ and $\text{Power}(T)$ for all methods. The training algorithm is random forest. The black dashed lines denote the FDR level $\alpha = 0.3$. The non-null proportion is $\pi_1 \in \{0.2, \dots, 0.8\}$.

Appendix D. Applications on Real-time LLM Alignment

In this section, we introduce the potential application of our proposed OCTF procedure on the task of real-time LLM alignment. For example, in medical report generation tasks, we may need to sequentially select radiology images $t \in \{1, \dots, T\}$ for which the generated reports align with expert standards. Similarly, in question-answering tasks, our goal is to identify the generated answer that best matches the true reference answer in an online

fashion. Specifically, let $f : \mathcal{X} \rightarrow \mathcal{Y}$ be a pre-trained foundation model that maps a prompt to an output. A holdout set $\mathcal{D} = (\mathbf{X}_i, E_i)_{i=-n}^0$ is available, where $\mathbf{X}_i \in \mathcal{X}$ represents an input prompt, and $E_i \in \mathcal{E}$ serves as a reference for assessing alignment. The alignment score function $\mathcal{A} : \mathcal{Y} \times \mathcal{E} \rightarrow \mathbb{R}$ maps the generated output $f(\mathbf{X})$ and reference E to an alignment score $A = \mathcal{A}(f(\mathbf{X}), E)$. For example, A may represent the similarity score between the machine-generated report $f(\mathbf{X})$ and a human expert report E . The test data $\{\mathbf{X}_t\}_{t=1}^T$ arrive sequentially, forming the following online multiple hypothesis testing problem at time t :

$$\mathbb{H}_{0t} : A_t \leq c \quad \text{versus} \quad \mathbb{H}_{1t} : A_t > c,$$

where $c \in \mathbb{R}$ is a pre-specified threshold. After making a decision δ_t at time t , the corresponding human expert report is revealed either immediately or with a delay of d time steps.

Our goal is to control the online FDR:

$$\text{FDR}(t) = \mathbb{E} \left[\frac{\sum_{j \leq t} \mathbb{I}\{A_j \leq c, \delta_j = 1\}}{1 \vee \sum_{j \leq t} \delta_j} \right] \leq \alpha. \quad (28)$$

Following the conformal alignment framework Gui et al. (2024), we randomly split $\mathcal{D}$ into two subsets: a training set $\mathcal{D}_{\mathcal{T}}$ and an initial calibration set $\mathcal{D}_{\mathcal{C}}$. Using $\mathcal{D}_{\mathcal{T}}$, we train an alignment predictor $g(\mathbf{X})$ to estimate the alignment score based on features of LLM outputs $\mathbf{X}_t$ and compute the predicted alignment scores $\hat{A}_t = g(\mathbf{X}_t)$ for each $t \in [T]$. Then applying OCTF with $\mathcal{D}_{\mathcal{C}_{0t}}$ to select new images whose generated reports are aligned with expert standards ensures finite-sample online mFDR control according to Theorem 11, where the calibration set $\mathcal{C}_{0t}$ is updated online as the feedback is revealed. The conformal p -value for $t \in [T]$ is:

$$p_t = \frac{1 + \sum_{i \in \mathcal{C}_{0t}} \mathbb{I}\{\hat{A}_i \geq \hat{A}_t\}}{1 + |\mathcal{C}_{0t}|}. \quad (29)$$

The real-time LLM alignment procedure is summarized in Algorithm 6, which, similar to Theorem 11, guarantees finite-sample mFDR control under the same assumptions.

Appendix E. Additional Experimental Results

In this section, we provide additional experimental results to further demonstrate the superior performance of our proposed algorithms, with a focus on online conformal testing.

Specifically, Appendix E.1 reports numerical results for the experiments in Sections 5 and 6; Appendix E.2 reports additional results for a regression example (Scenario V); Appendix E.3 reports results under different training algorithms for $\hat{\mu}$; Appendix E.4 reports additional results after score selection; Appendix E.5 reports empirical mFDR results; and Appendix E.6 reports results on auxiliary calibration for score selection; Appendix E.7 reports the results for comparison with alternative finite-sample-valid baselines in online conformal testing.

E.1 Additional Numerical Results for Sections 5 and 6

This section provides numerical summaries complementing the simulation experiments in Section 5 and the real-data analysis in Section 6. Tables 3–7 report the FDR and Power

Algorithm 6 Real-time LLM conformal Alignment with feedback

Input: Pre-trained foundation model f ; alignment score function $\mathcal{A}$; reference dataset $\mathcal{D} = (\mathbf{X}_i, E_i)_{i=-n}^0$; algorithm for fitting alignment predictor $\mathcal{G}$; alignment level c ; target FDR level α .

- 1: Compute the alignment score $A_i = \mathcal{A}(f(\mathbf{X}_i), E_i)$, $\forall i \in \mathcal{D}$.
- 2: Randomly split $\mathcal{D}$ into two disjoint sets: the training set $\mathcal{D}_{\mathcal{T}}$ and the calibration set $\mathcal{D}_{\mathcal{C}}$.
- 3: Fit the alignment score predictor with $\mathcal{D}_{\mathcal{T}}$: $g \leftarrow \mathcal{G}(\mathcal{D}_{\mathcal{T}})$.
- 4: Initialize $\mathcal{C}_{0t} = \{i \in \mathcal{C} : A_i \leq c\}$
- 5: **for** $t \in [T]$ **do**
- 6: Observe test data $\mathbf{X}_{n+t}$
- 7: Compute the predicted alignment score: $\hat{A}_i \leftarrow g(\mathbf{X}_i)$, $\forall i \in \mathcal{C}$ and $\hat{A}_t \leftarrow g(\mathbf{X}_t)$.
- 8: Compute the conformal p-value p_t according to Equation (29).
- 9: Update $\alpha_t = \alpha_t^{\text{LFS}}$ in Equation (18) (or $\alpha_t = \alpha_t^{\text{SFS}}$ in (19))
- 10: Obtain the revealed feedback θ_t .
- 11: Update the calibration dataset $\mathcal{C}_{0t}$.
- 12: **end for**

Output: The selection set $\mathcal{R} = \{t : \delta_t = 1, t \in [T]\}$.

values underlying Figures 2–6, respectively. Table 8 reports the results for Scenario IV with sine pattern shift at the stopping time ($T = 1000$). Table 9 further reports the performance of the safe variants Opt-SFS and Opt-LFS on the real-data tasks of Section 6.

E.2 Results on synthetic data for a regression task

The corresponding data generation process for the regression task is detailed below:

- **Scenario V (Regression example):** $Y = -0.5X_1^2 + \exp X_2 + (X_3 + X_4)^2 + \varepsilon$, with $\mathbf{X} \sim \mathcal{N}_4(\mathbf{0}, \mathbf{I}_4)$ and $\varepsilon \sim \mathcal{N}(0, 2)$. The target region is $\mathcal{A} = [c, \infty)$, where c is the $1 - \pi_1$ quantile of Y .

In terms of the non-conformity score function, denote $W_t = \hat{\mu}_t(\mathbf{X}_t)$, in classification settings, we set $V(W_t) = 1 - W_t$. In regression settings, if $\mathcal{A} = [b, +\infty)$, we can use $V(W_j) = b - W_j$. If $\mathcal{A} = (-\infty, a] \cup [b, +\infty)$, then we can choose $V(W_t) = \max\{W_t - a, b - W_t\}$.

The results for the regression example (Scenario V) using a fixed training algorithm (random forest) are shown in Figure 14. The performance trends are similar to those observed in the classification case (Scenario IV).

Table 3: FDR and Power of competing methods under Scenario I and Scenario II. Results are averaged over 500 simulation runs. Standard errors are given in parentheses.

Scenario	Method	π_1							
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
I	<i>FDR</i>								
	LOND	.023 _(.005)	.012 _(.004)	.008 _(.003)	.006 _(.001)	.005 _(.001)	.005 _(.001)	.003 _(.001)	.001 _(.000)
	LORD++	.021 _(.006)	.011 _(.003)	.010 _(.002)	.013 _(.001)	.010 _(.001)	.009 _(.001)	.007 _(.001)	.005 _(.000)
	SAFFRON	.043 _(.007)	.053 _(.005)	.063 _(.003)	.073 _(.002)	.077 _(.001)	.072 _(.001)	.063 _(.001)	.053 _(.001)
	LF	.063 _(.009)	.068 _(.007)	.071 _(.004)	.086 _(.003)	.092 _(.001)	.091 _(.001)	.090 _(.001)	.086 _(.000)
	SF	.065 _(.009)	.069 _(.007)	.075 _(.004)	.086 _(.002)	.091 _(.001)	.091 _(.001)	.088 _(.001)	.080 _(.001)
	<i>Power</i>								
	LOND	.009 _(.001)	.012 _(.000)	.014 _(.000)	.016 _(.000)	.018 _(.000)	.020 _(.000)	.022 _(.000)	.024 _(.000)
	LORD++	.007 _(.001)	.013 _(.001)	.021 _(.001)	.031 _(.001)	.046 _(.001)	.059 _(.001)	.074 _(.001)	.090 _(.002)
	SAFFRON	.010 _(.001)	.028 _(.002)	.080 _(.003)	.169 _(.004)	.278 _(.003)	.381 _(.003)	.482 _(.003)	.584 _(.002)
	LF	.009 _(.001)	.023 _(.001)	.076 _(.003)	.184 _(.004)	.335 _(.003)	.477 _(.002)	.620 _(.002)	.771 _(.001)
	SF	.011 _(.001)	.031 _(.002)	.096 _(.003)	.210 _(.004)	.335 _(.003)	.465 _(.002)	.594 _(.002)	.738 _(.002)
II	<i>FDR</i>								
	LOND	.014 _(.004)	.012 _(.003)	.008 _(.002)	.005 _(.001)	.002 _(.001)	.002 _(.001)	.002 _(.001)	.003 _(.001)
	LORD++	.013 _(.004)	.012 _(.003)	.012 _(.002)	.013 _(.001)	.010 _(.001)	.009 _(.001)	.006 _(.001)	.005 _(.000)
	SAFFRON	.025 _(.005)	.042 _(.004)	.063 _(.004)	.074 _(.002)	.080 _(.001)	.084 _(.001)	.079 _(.001)	.072 _(.001)
	LF	.073 _(.010)	.067 _(.008)	.077 _(.006)	.084 _(.002)	.089 _(.001)	.092 _(.001)	.090 _(.001)	.087 _(.000)
	SF	.049 _(.007)	.061 _(.005)	.074 _(.003)	.081 _(.002)	.086 _(.001)	.085 _(.001)	.083 _(.001)	.080 _(.001)
	<i>Power</i>								
	LOND	.011 _(.001)	.012 _(.000)	.014 _(.000)	.015 _(.000)	.016 _(.000)	.018 _(.000)	.020 _(.000)	.022 _(.000)
	LORD++	.009 _(.001)	.013 _(.001)	.019 _(.001)	.028 _(.001)	.036 _(.001)	.049 _(.001)	.059 _(.001)	.070 _(.001)
	SAFFRON	.008 _(.001)	.024 _(.001)	.064 _(.003)	.152 _(.004)	.280 _(.005)	.440 _(.005)	.616 _(.004)	.785 _(.002)
	LF	.010 _(.001)	.022 _(.001)	.061 _(.002)	.172 _(.004)	.328 _(.004)	.517 _(.003)	.713 _(.002)	.875 _(.001)
	SF	.017 _(.001)	.042 _(.002)	.104 _(.003)	.212 _(.003)	.338 _(.003)	.477 _(.003)	.640 _(.003)	.800 _(.002)

Table 4: FDR and Power of competing methods under dependent setting (Scenario III). Results are averaged over 500 simulation runs. Standard errors are given in parentheses.

Method	π_1							
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
<i>FDR</i>								
LOND	.010 _(.003)	.007 _(.002)	.006 _(.002)	.004 _(.001)	.002 _(.001)	.002 _(.000)	.002 _(.001)	.001 _(.000)
LORD++	.034 _(.006)	.033 _(.005)	.030 _(.003)	.025 _(.003)	.013 _(.002)	.015 _(.002)	.013 _(.001)	.007 _(.001)
LORDdep	.023 _(.004)	.022 _(.003)	.020 _(.003)	.018 _(.002)	.009 _(.001)	.009 _(.001)	.009 _(.001)	.004 _(.001)
SAFFRON	.270 _(.014)	.312 _(.011)	.280 _(.008)	.236 _(.006)	.191 _(.005)	.160 _(.003)	.117 _(.003)	.085 _(.002)
SAFFRONdep	.036 _(.005)	.041 _(.004)	.038 _(.003)	.037 _(.003)	.025 _(.002)	.025 _(.002)	.019 _(.001)	.012 _(.001)
LF	.156 _(.011)	.177 _(.009)	.155 _(.006)	.126 _(.005)	.105 _(.004)	.097 _(.002)	.083 _(.002)	.071 _(.001)
LFdep	.030 _(.005)	.029 _(.004)	.026 _(.003)	.024 _(.002)	.013 _(.001)	.014 _(.001)	.012 _(.001)	.006 _(.001)
SF	.270 _(.014)	.312 _(.011)	.281 _(.008)	.236 _(.006)	.191 _(.005)	.160 _(.003)	.117 _(.003)	.085 _(.002)
SFdep	.031 _(.005)	.035 _(.004)	.032 _(.003)	.033 _(.003)	.022 _(.002)	.023 _(.002)	.019 _(.001)	.014 _(.001)
<i>Power</i>								
LOND	.126 _(.006)	.155 _(.006)	.186 _(.006)	.209 _(.007)	.210 _(.007)	.244 _(.007)	.248 _(.007)	.255 _(.008)
LORD++	.085 _(.006)	.128 _(.008)	.193 _(.009)	.225 _(.010)	.237 _(.010)	.295 _(.011)	.311 _(.011)	.310 _(.012)
LORDdep	.079 _(.006)	.117 _(.007)	.176 _(.009)	.205 _(.010)	.217 _(.010)	.274 _(.011)	.288 _(.011)	.285 _(.012)
SAFFRON	.282 _(.012)	.416 _(.012)	.515 _(.012)	.588 _(.012)	.648 _(.011)	.748 _(.009)	.788 _(.009)	.857 _(.007)
SAFFRONdep	.150 _(.008)	.236 _(.009)	.323 _(.010)	.387 _(.012)	.438 _(.011)	.519 _(.011)	.542 _(.011)	.580 _(.010)
LF	.188 _(.009)	.347 _(.011)	.490 _(.011)	.604 _(.011)	.698 _(.010)	.808 _(.007)	.858 _(.006)	.905 _(.005)
LFdep	.145 _(.007)	.234 _(.009)	.308 _(.009)	.372 _(.010)	.417 _(.010)	.478 _(.009)	.490 _(.010)	.520 _(.009)
SF	.283 _(.012)	.416 _(.012)	.515 _(.012)	.589 _(.012)	.650 _(.011)	.748 _(.009)	.791 _(.008)	.862 _(.007)
SFdep	.182 _(.009)	.270 _(.010)	.351 _(.010)	.412 _(.011)	.458 _(.011)	.542 _(.010)	.572 _(.010)	.628 _(.009)

Table 5: FDR and Power of competing methods under bandit and instant feedback setting. Results are averaged over 500 simulation runs. Standard errors are given in parentheses.

Method	π_1							
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
<i>FDR</i>								
LOND	.014 _(.004)	.012 _(.003)	.009 _(.002)	.008 _(.002)	.004 _(.001)	.003 _(.001)	.003 _(.001)	.002 _(.000)
LORD++	.004 _(.002)	.002 _(.001)	.006 _(.001)	.006 _(.001)	.008 _(.001)	.006 _(.001)	.004 _(.000)	.003 _(.000)
SAFFRON	.049 _(.008)	.050 _(.006)	.065 _(.004)	.071 _(.002)	.074 _(.001)	.071 _(.001)	.064 _(.001)	.053 _(.000)
LF-BI	.045 _(.007)	.058 _(.007)	.070 _(.005)	.084 _(.003)	.089 _(.002)	.091 _(.001)	.090 _(.001)	.086 _(.000)
SF-BI	.049 _(.008)	.050 _(.006)	.065 _(.004)	.071 _(.002)	.074 _(.001)	.071 _(.001)	.064 _(.001)	.053 _(.000)
<i>Power</i>								
LOND	.009 _(.000)	.013 _(.000)	.014 _(.000)	.016 _(.000)	.018 _(.000)	.021 _(.000)	.022 _(.000)	.026 _(.001)
LORD++	.004 _(.000)	.007 _(.001)	.012 _(.001)	.020 _(.001)	.028 _(.001)	.039 _(.002)	.050 _(.002)	.066 _(.002)
SAFFRON	.011 _(.001)	.028 _(.002)	.082 _(.003)	.171 _(.004)	.271 _(.004)	.377 _(.003)	.483 _(.002)	.588 _(.002)
LF-BI	.010 _(.001)	.024 _(.001)	.074 _(.003)	.183 _(.004)	.313 _(.004)	.462 _(.004)	.606 _(.003)	.765 _(.002)
SF-BI	.011 _(.001)	.028 _(.002)	.082 _(.003)	.171 _(.004)	.271 _(.004)	.377 _(.003)	.483 _(.002)	.588 _(.002)

Table 6: FDR and Power of competing methods under full and delayed feedback setting with delay $d \in \{0, 10, 100\}$ (Scenario I). Results are averaged over 500 simulation runs. Standard errors are given in parentheses.

Method	π_1							
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
<i>FDR</i>								
LOND	.021 _(.005)	.014 _(.003)	.008 _(.003)	.006 _(.001)	.003 _(.001)	.002 _(.001)	.002 _(.000)	.001 _(.000)
LORD++	.001 _(.001)	.004 _(.001)	.003 _(.001)	.005 _(.001)	.006 _(.001)	.006 _(.001)	.005 _(.001)	.004 _(.000)
SAFFRON	.044 _(.007)	.052 _(.005)	.062 _(.004)	.072 _(.002)	.074 _(.001)	.071 _(.001)	.065 _(.001)	.053 _(.000)
LF-FD $d = 0$	.047 _(.008)	.068 _(.007)	.066 _(.004)	.085 _(.003)	.090 _(.001)	.091 _(.001)	.091 _(.001)	.087 _(.000)
LF-FD $d = 10$	.044 _(.008)	.057 _(.007)	.048 _(.005)	.046 _(.003)	.045 _(.001)	.039 _(.001)	.032 _(.001)	.022 _(.000)
LF-FD $d = 100$	.041 _(.007)	.055 _(.007)	.046 _(.005)	.042 _(.003)	.040 _(.001)	.035 _(.001)	.028 _(.001)	.020 _(.000)
SF-FD $d = 0$	.048 _(.007)	.056 _(.005)	.068 _(.004)	.084 _(.002)	.088 _(.001)	.090 _(.001)	.090 _(.001)	.081 _(.001)
SF-FD $d = 10$	.045 _(.007)	.055 _(.005)	.065 _(.004)	.076 _(.002)	.079 _(.001)	.078 _(.001)	.074 _(.001)	.064 _(.001)
SF-FD $d = 100$	.044 _(.007)	.054 _(.005)	.063 _(.004)	.073 _(.002)	.075 _(.001)	.072 _(.001)	.066 _(.001)	.054 _(.000)
<i>Power</i>								
LOND	.009 _(.000)	.012 _(.000)	.013 _(.000)	.015 _(.000)	.017 _(.000)	.020 _(.000)	.022 _(.000)	.025 _(.000)
LORD++	.004 _(.000)	.007 _(.001)	.010 _(.001)	.018 _(.001)	.029 _(.001)	.039 _(.002)	.052 _(.002)	.071 _(.002)
SAFFRON	.010 _(.001)	.031 _(.002)	.073 _(.003)	.166 _(.004)	.275 _(.004)	.382 _(.003)	.482 _(.002)	.588 _(.002)
LF-FD $d = 0$	.008 _(.001)	.026 _(.002)	.070 _(.003)	.182 _(.004)	.328 _(.004)	.476 _(.002)	.620 _(.002)	.775 _(.001)
LF-FD $d = 10$	.007 _(.001)	.019 _(.001)	.036 _(.002)	.080 _(.003)	.146 _(.003)	.212 _(.003)	.269 _(.003)	.326 _(.002)
LF-FD $d = 100$	.007 _(.001)	.017 _(.001)	.029 _(.002)	.059 _(.002)	.114 _(.003)	.167 _(.003)	.215 _(.003)	.266 _(.003)
SF-FD $d = 0$	.011 _(.001)	.035 _(.002)	.087 _(.003)	.202 _(.004)	.331 _(.003)	.462 _(.002)	.596 _(.002)	.741 _(.002)
SF-FD $d = 10$	.011 _(.001)	.033 _(.002)	.079 _(.003)	.181 _(.004)	.296 _(.004)	.413 _(.002)	.528 _(.002)	.654 _(.002)
SF-FD $d = 100$	.010 _(.001)	.032 _(.002)	.075 _(.003)	.169 _(.004)	.279 _(.004)	.387 _(.003)	.489 _(.002)	.599 _(.002)

Table 7: FDR and Power of competing methods in the classification setting. Results are averaged over 500 simulation runs. Standard errors are given in parentheses.

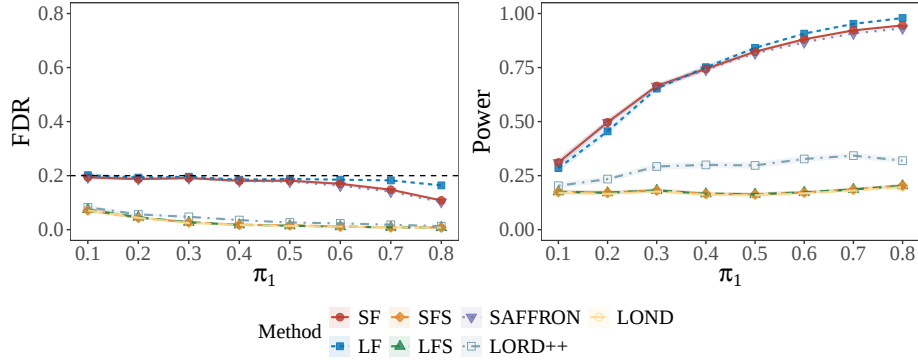
Method	π_1							
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
<i>FDR</i>								
LOND	.034 _(.007)	.023 _(.004)	.027 _(.005)	.019 _(.004)	.020 _(.005)	.015 _(.004)	.011 _(.003)	.012 _(.004)
LORD++	.011 _(.004)	.008 _(.002)	.005 _(.001)	.007 _(.002)	.005 _(.002)	.002 _(.001)	.004 _(.002)	.004 _(.002)
SAFFRON	.090 _(.007)	.138 _(.006)	.137 _(.004)	.148 _(.004)	.147 _(.003)	.142 _(.003)	.126 _(.004)	.084 _(.003)
LF	.132 _(.011)	.151 _(.008)	.152 _(.006)	.168 _(.004)	.169 _(.004)	.172 _(.003)	.175 _(.004)	.147 _(.003)
SF	.117 _(.009)	.153 _(.007)	.151 _(.005)	.165 _(.004)	.162 _(.003)	.154 _(.002)	.137 _(.004)	.094 _(.003)
LFS	.111 _(.012)	.083 _(.009)	.052 _(.007)	.044 _(.005)	.041 _(.005)	.033 _(.004)	.036 _(.006)	.031 _(.004)
SFS	.088 _(.010)	.071 _(.008)	.038 _(.005)	.031 _(.004)	.033 _(.004)	.024 _(.003)	.032 _(.005)	.024 _(.004)
<i>Power</i>								
LOND	.014 _(.001)	.018 _(.001)	.017 _(.001)	.013 _(.001)	.012 _(.001)	.007 _(.001)	.003 _(.000)	.001 _(.000)
LORD++	.007 _(.002)	.020 _(.003)	.035 _(.005)	.028 _(.004)	.023 _(.004)	.016 _(.003)	.006 _(.002)	.000 _(.000)
SAFFRON	.077 _(.006)	.340 _(.013)	.533 _(.015)	.660 _(.015)	.735 _(.015)	.779 _(.015)	.794 _(.016)	.729 _(.018)
LF	.049 _(.005)	.294 _(.012)	.510 _(.015)	.719 _(.013)	.792 _(.014)	.861 _(.012)	.881 _(.013)	.860 _(.014)
SF	.085 _(.007)	.372 _(.013)	.560 _(.014)	.717 _(.013)	.784 _(.013)	.847 _(.012)	.856 _(.013)	.820 _(.015)
LFS	.011 _(.000)	.011 _(.000)	.011 _(.000)	.011 _(.000)	.011 _(.000)	.010 _(.000)	.011 _(.000)	.009 _(.000)
SFS	.013 _(.001)	.014 _(.000)	.013 _(.000)	.012 _(.000)	.012 _(.000)	.011 _(.000)	.011 _(.000)	.009 _(.000)

 Table 8: Results for Scenario IV (sine pattern shifts): Performance at the stopping time ($T = 1000$). The target level $\alpha = 0.05$.

Framework	Method	FDR	Power
Adaptive GAIF	SAFFRON	.038 _(.004)	.280 _(.021)
	Ran-SF	.037 _(.003)	.289 _(.021)
	Opt-SF	.046 _(.007)	.421 _(.023)
	Ran-SFS	.005 _(.003)	.007 _(.000)
	Opt-SFS	.019 _(.008)	.008 _(.001)
GAIF	LORD++	.007 _(.005)	.002 _(.001)
	Ran-LF	.048 _(.007)	.297 _(.021)
	Opt-LF	.047 _(.009)	.395 _(.024)
	Ran-LFS	.015 _(.008)	.006 _(.000)
	Opt-LFS	.021 _(.009)	.006 _(.000)

Table 9: FDR(T) and Power(T) for the safe variants (Opt-SFS and Opt-LFS) across four datasets (Candidate, Diabetes, Income, Airfoil). The target FDR level is $\alpha = 0.3$.

Method	Task 1		Task 2		Task 3		Task 4	
	FDR	Power	FDR	Power	FDR	Power	FDR	Power
Opt-SFS	.071(.014)	.002(.000)	.090(.017)	.002(.000)	.077(.014)	.008(.000)	.002(.001)	.007(.003)
Opt-LFS	.070(.014)	.002(.000)	.088(.016)	.002(.000)	.105(.017)	.007(.000)	.001(.001)	.005(.001)


 Figure 14: Results for Scenario V: values of FDR(T) and Power(T) at stopping time T across different non-null proportions π_1 . The black dashed line denotes the FDR level $\alpha = 0.2$.

E.3 Additional experiments for different training algorithms

The results for Scenarios IV and V under different training algorithms—RF, SVM, and NN—with varying initial calibration sizes are presented in Figure 15-16. Thanks to the online updating of the calibration dataset, even a small initial calibration size does not significantly impact performance. While the choice of predictive model $\hat{\mu}$ does affect power, all methods maintain valid FDR control. Notably, our SF and LF methods consistently outperform the baselines across all score functions, benefiting from the distribution-free and model-agnostic nature of online conformal p -values. The variation in performance across different algorithms further underscores the importance of careful score selection in practice.

E.4 Additional experiments on score selection

The results with score selection for Scenarios IV and V are shown below Figure 17-Figure 18. The target level is $\alpha = 0.05$. In both settings, the performance gap between the Opt methods and their randomly selected counterparts is also pronounced.

E.5 Additional experiments results on empirical mFDR

To illustrate the similarity between mFDR and FDR, we present results under various settings below in Figure 19-Figure 20. We estimate mFDR by computing the ratio of the

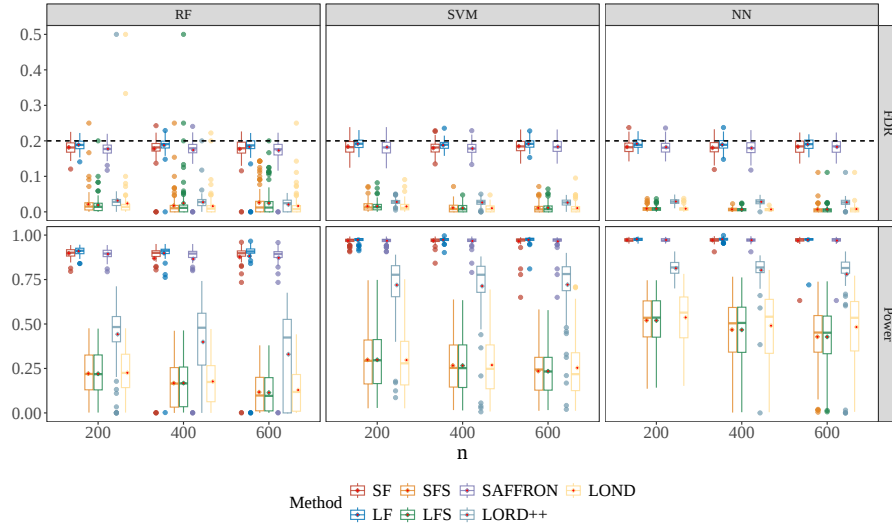


Figure 15: Results for Scenario IV: FDR(T) and Power(T) vs. initial calibration size n ($\pi_1 = 0.5$, $\alpha = 0.2$).

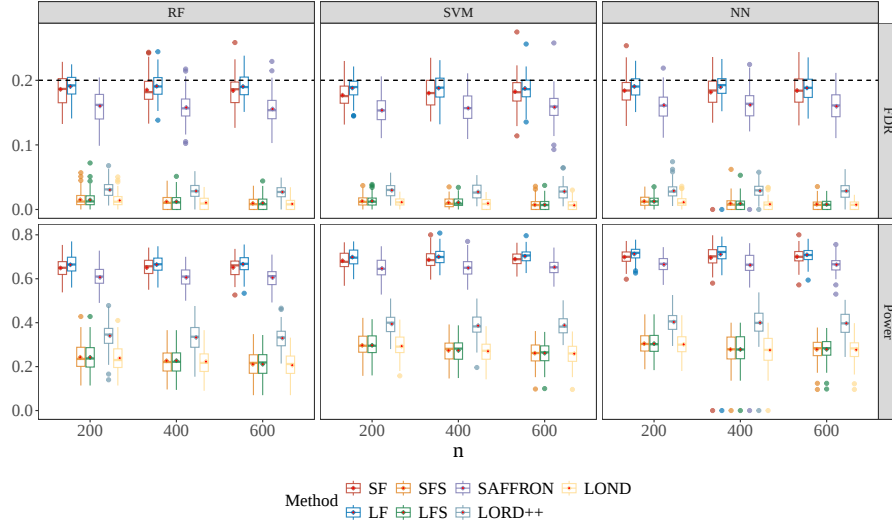


Figure 16: Results for Scenario V: FDR(T) and Power(T) vs. initial calibration size n ($\pi_1 = 0.5$, $\alpha = 0.2$).

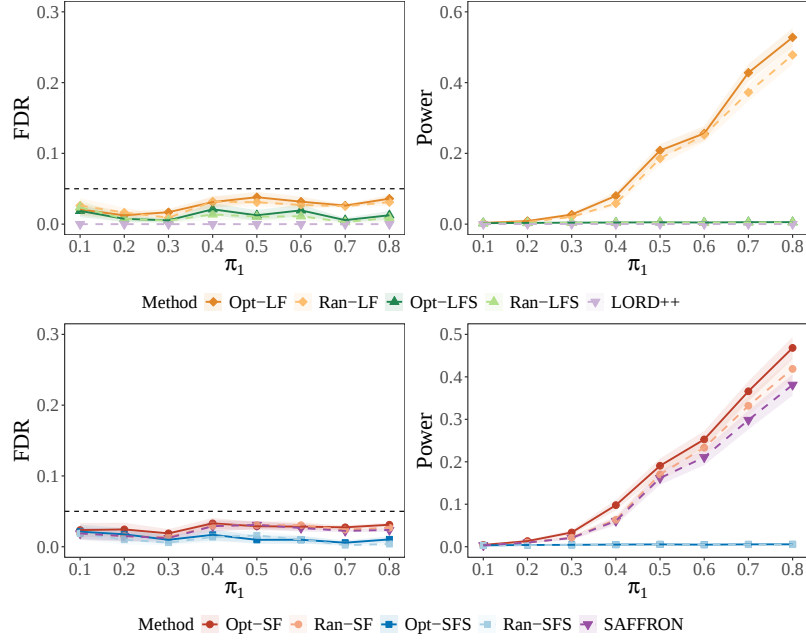


Figure 17: Results for Scenario IV: the values of $FDR(T)$ and $Power(T)$ at stopping time T across different non-null proportions π_1 . The black dashed lines denote the FDR level $\alpha = 0.05$.

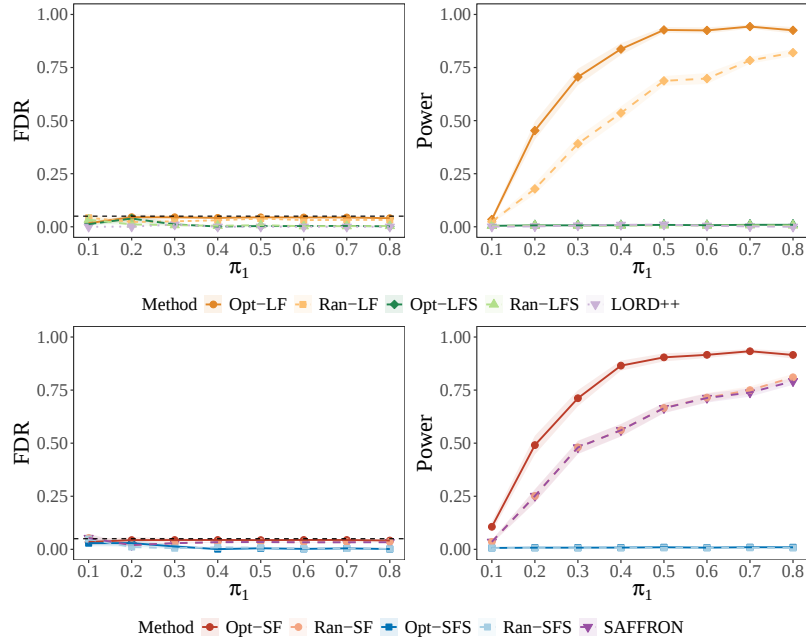


Figure 18: Results for Scenario V: the values of $FDR(T)$ and $Power(T)$ at stopping time T across different non-null proportions π_1 . The black dashed lines denote the FDR level $\alpha = 0.05$.

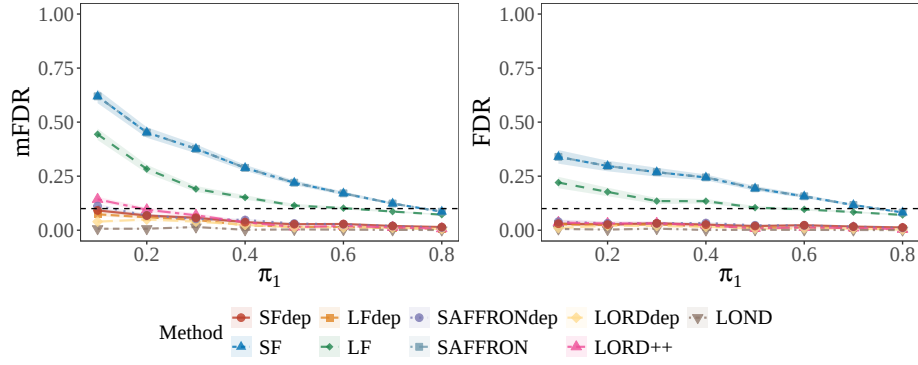


Figure 19: Results for Scenario III (local dependence): Line charts of mFDR and FDR at stopping time with varying non-null proportion π_1 from 0.1 to 0.8. The black dashed lines denote the target FDR level $\alpha = 0.1$.

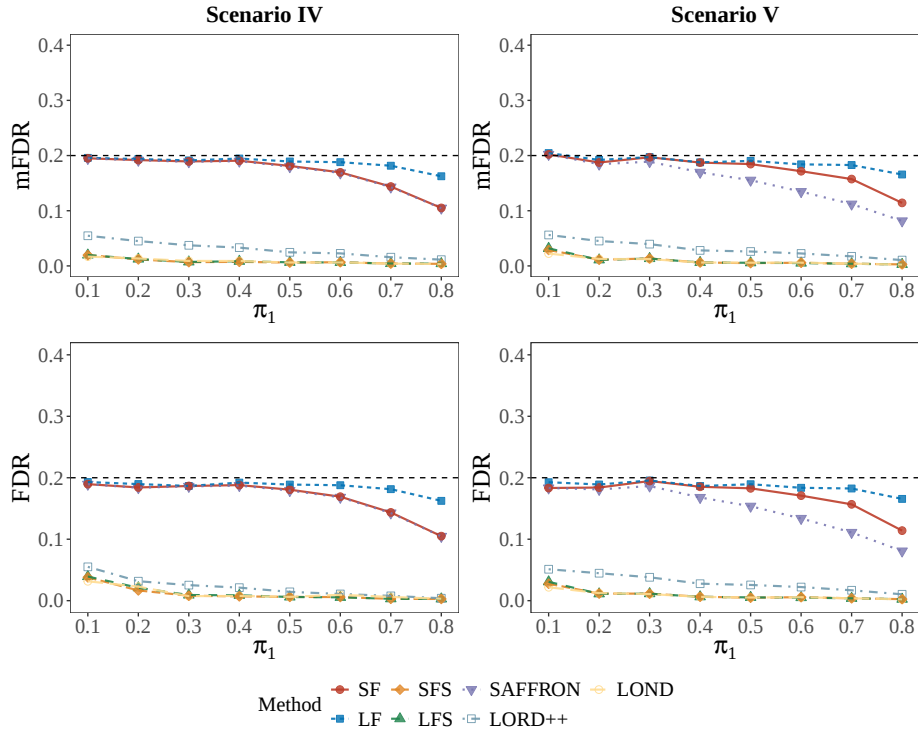


Figure 20: Results for Scenario IV and Scenario V : Line charts of mFDR and FDR at stopping time with varying non-null proportion π_1 from 0.1 to 0.8 after 500 replications; The black dashed lines denote the target FDR level $\alpha = 0.2$.

average number of false discoveries and the average total number of discoveries. Empirical mFDR closely tracks empirical FDR, and both are well controlled by our proposed methods.

E.6 Additional experiments on auxiliary calibration for score selection

In our EWMA-based model-selection procedure, we construct auxiliary non-null p -values in (13) based on $\mathcal{C}_{0t} \cup \{t\}$, where the current index t is included. The main purpose of this design is to preserve *finite-sample validity under data-adaptive score selection*. In particular, including t ensures the permutation symmetry of $\mathcal{C}_{0t} \cup \{t\}$ required by our validity argument.

From a power perspective, one may worry that, when t is non-null and t is small, including the current point could slightly perturb the auxiliary calibration distribution. In practice, however, this effect is limited in our setting. Our procedure starts from a historical calibration dataset $\mathcal{C}$ of size n , so even at early times the inclusion of a single point has only a minor impact, since the effective null calibration size is already $n + t$.

To further reduce the possible adverse effect when t is non-null, we also consider a simple robustified version based on a truncated calibration set:

$$\tilde{p}_{k,j} = \frac{\sum_{s \in \mathcal{C}_t^{\text{trun},k}} \mathbb{I}\{V(\mathbf{X}_s; k) \leq V(\mathbf{X}_j; k)\}}{|\mathcal{C}_t^{\text{trun},k}|}, \quad j \in \mathcal{C}_{1t}, \quad (30)$$

where

$$\mathcal{C}_t^{\text{trun},k} = (\mathcal{C}_{0t} \cup \{t\}) \setminus \left\{ \arg \min_{i \in \mathcal{C}_{0t} \cup \{t\}} V(\mathbf{X}_i; k) \right\}.$$

This operation remains permutation-invariant on $\mathcal{C}_{0t} \cup \{t\}$, so the symmetry needed for validity is preserved. Intuitively, when t is non-null, its score $V(\mathbf{X}_t; k)$ is more likely to be unusually small and hence removed by the truncation, reducing its impact.

To investigate the practical effect of including t , we compare the following four variants:

- **Opt-SFS:** Our proposed feedback-enhanced approach using EWMA of past non-null auxiliary p -values based on calibration set $\mathcal{C}_{0t} \cup \{t\}$
- **OptEx-SFS:** Similar to Opt-SFS but excluding t from the calibration data, i.e. constructing auxiliary p -values based on $\mathcal{C}_{0t}$.
- **OptTr-SFS:** Optimized selection using the truncated calibration set $\mathcal{C}_t^{\text{trun},k}$ defined in Eq. (30).
- **Ran-SFS:** A baseline that randomly selects a model at each time step to quantify the gain from feedback-driven selection.

We generated data streams of length $T = 1000$ characterized by smoothly varying distribution shifts according to Scenario IV in Section 5.3. Conformal p -values were computed online using the above proposed EWMA-based selection strategies, and rejections were determined via standard SFS procedures in the main text. For each method, we tracked the empirical online FDR and empirical power across all time points across 200 replications.

As shown in Figure 21, all three optimized methods (Opt-SFS, OptEx-SFS, and OptTr-SFS) maintain empirical FDR control below the nominal level ($\alpha = 0.1$) while consistently achieving higher power than the random baseline.

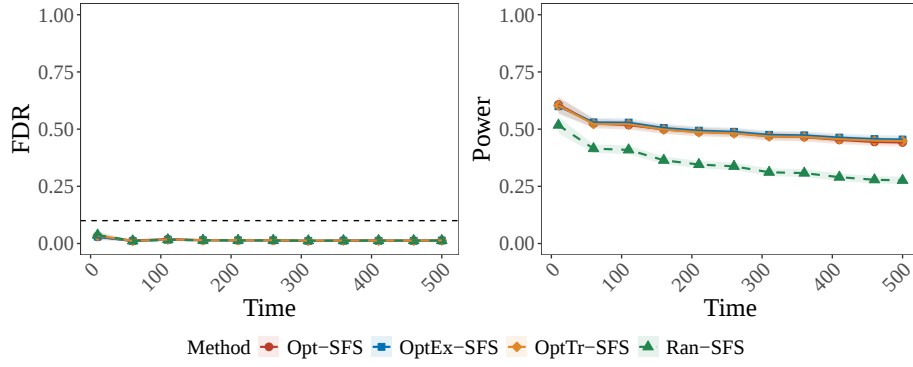


Figure 21: Online FDR and power for various Opt-SFS variants versus the Ran-SFS baseline over time.

A comparison between Opt-SFS and OptEx-SFS reveals that at early time points (t small), differences in power among these strategies are relatively minor. As more observations accumulate, the differences among the three Opt methods diminish. These results empirically verify that the inclusion of t has negligible influence on the power performance.

E.7 Comparison with alternative finite-sample-valid baselines

In this section, we compare LFS and SFS against two natural alternatives that are also finite-sample valid in online conformal testing.

Fixed-calibration conformal p -values and LOND. The first alternative keeps the calibration set fixed throughout the entire test stream, exactly as in standard split-conformal inference: for each test point t with conformity score V_t ,

$$p_t = \frac{\sum_{i \in \mathcal{C}_0} \mathbb{I}\{V_i < V_t\} + \xi_t[1 + \sum_{i \in \mathcal{C}_0} \mathbb{I}\{V_i = V_t\}]}{1 + |\mathcal{C}_0|}, \quad (31)$$

where $\mathcal{C}_0$ is a hold-out calibration dataset. Under the standard exchangeability assumption, the null conformal p -values constructed from a shared fixed calibration set $\mathcal{C}_0$ are PRDS, which is sufficient for the LOND rule $\alpha_t^{\text{LOND}} := \alpha \gamma_t(R^{\text{LOND}}(t-1) + 1)$ to control the FDR (Zrnic et al., 2021).

Online conformal e -values and e-LOND. The second alternative retains our online calibration update: each confirmed null is added to the calibration pool as soon as it is observed. It then uses e -values in place of p -values and applies e-LOND rule of (Xu and Ramdas, 2024). Because e-LOND controls the FDR under arbitrary, possibly unknown dependence among e -values, it accommodates the dependence induced by an evolving shared calibration set. We use the rank-based conformal e -value

$$e_t = \frac{n_t + 1}{k_t} \cdot \mathbf{1}\{\text{rank}_t \leq k_t\}, \quad k_t = \max\{1, \lfloor \rho(n_t + 1) \rfloor\}, \quad (32)$$

where n_t is the calibration set size immediately before test point t is processed, rank_t is the (randomized, tie-broken) rank of the test score among the n_t calibration scores and itself, and

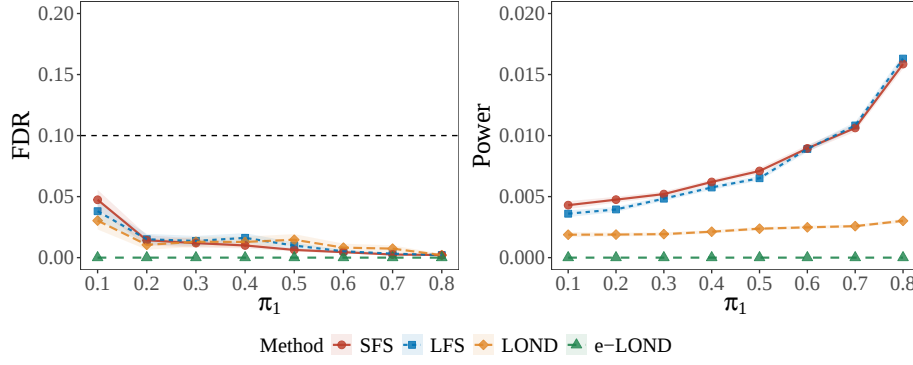


Figure 22: FDR (left) and power (right) across different non-null proportion π_1 , for LOND with classic (fixed) conformal p -values, e-LOND with online conformal e -values, and our proposed LFS/SFS with online conformal p -values. The dashed line marks the nominal level $\alpha = 0.1$.

$\rho \in (0, 1)$ is a fixed proportion (we use $\rho = 0.05$). This is a generalization of the single-test conformal e -value of Vovk and Wang (2021) and the construction underlying Bashari et al. (2023): since k_t depends only on the calibration history n_t and not on the test score itself, it is predictable, and exchangeability of the test score with the calibration pool under the null gives $\mathbb{E}[e_t] = 1$. Discoveries are then made using

$$\alpha_t^{\text{e-LOND}} := \alpha \gamma_t (R^{\text{e-LOND}}(t-1) + 1), \quad \text{reject } \mathbb{H}_{0t} \iff e_t \geq 1/\alpha_t^{\text{e-LOND}}. \quad (33)$$

Experimental setup. We employ the binary classification setting of Scenario IV with $\mu_1 = (5, 0, 0, 0)^\top$ and $\mu_2 = (0, 0, -4, -4)^\top$, sweeping the non-null proportion $\pi_1 \in \{0.1, 0.2, \dots, 0.8\}$. FDR and power are averaged over 500 independent replications at nominal level $\alpha = 0.1$, with an initial null calibration set of size $n_{\text{cal}} = 50$. The training algorithm is support vector machine. We compare four procedures: LOND applied to classic conformal p -values (31), e-LOND applied to online conformal e -values (32) via (33), and our proposed LFS and SFS applied to online conformal p -values.

Results. Figure 22 reports empirical FDR and power across different π_1 for all four methods. LOND with a fixed calibration set and e-LOND both control the empirical FDR at the nominal level across all values of π_1 , confirming their finite-sample validity. While all four methods maintain valid empirical error control, LFS and SFS consistently achieve power gains relative to LOND. In contrast, e-LOND remains highly conservative. This indicates that incorporating feedback through the proposed safe strategy in OCTF can improve detection performance while preserving finite-sample guarantees.

Appendix F. Technical Details

F.1 Auxiliary lemmas and proof of lemmas

Our proof of Theorem 3 and Theorem 7 relies on the following lemmas, which are modified versions of Lemma 1 in Ramdas et al. (2017) and Lemma 1 in Ramdas et al. (2018),

respectively. The modifications arise because, in our setting, the feedback information $\{\theta_j\}_{j \in \mathcal{I}_t}$ is available at time t , and the test levels α_t in GAIF and Adaptive GAIF procedures depend on both past rejections δ_j and feedback $\{\theta_j\}_{j \in \mathcal{I}_t}$.

For simplicity, denote $\delta_{1:t} = (\delta_1, \dots, \delta_t)$. Given a sequence $p_1, p_2, \dots$ of independent p -values (i.e., the null p -values are independent of all other p -values), we define a filtration via the sigma-field of decisions and feedback $\mathcal{F}_{t-1} = \sigma(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})$, where $\delta_t = \mathbb{I}\{p_t \leq f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})\}$ for some coordinate-wise non-decreasing function $f_t : \{0, 1\}^{t-1+|\mathcal{I}_t|} \rightarrow \mathbb{R}$. With this set-up, we have the following guarantees in Lemma 19:

Lemma 19. *Let $g : \{0, 1\}^T \rightarrow \mathbb{R}$ be any coordinate-wise non-decreasing function such that $g(\mathbf{x}) > 0$ for any vector $\mathbf{x} \neq (0, \dots, 0)$. Then for any index $t \leq T$ such that $H_t \in \mathcal{H}_0$, we have*

$$\mathbb{E} \left[\frac{\mathbb{I}\{p_t \leq f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})\}}{g(\delta_{1:T})} \mid \mathcal{F}_{t-1} \right] \leq \mathbb{E} \left[\frac{f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})}{g(\delta_{1:T})} \mid \mathcal{F}_{t-1} \right].$$

Proof Denote $p_{1:T} = (p_1, \dots, p_T)$, and $\tilde{p}_{1:T}^{-t} = (\tilde{p}_1, \dots, \tilde{p}_T)$, where

$$\tilde{p}_i = \begin{cases} 0 & \text{if } i = t \\ p_i & \text{if } i \neq t, \end{cases}$$

i.e., the t -th component of $\tilde{p}_{1:T}^{-t}$ is zero and the other components of $\tilde{p}_{1:T}^{-t}$ equals to that of $p_{1:T}$. For all i , define $\tilde{\delta}_i = \mathbb{I}\{\tilde{p}_i \leq f_i(\tilde{\delta}_{1:i-1}, \{\theta_j\}_{j \in \mathcal{I}_i})\}$. And let the decision vectors using $p_{1:T}$ and $\tilde{p}_{1:T}^{-t}$ be $\delta_{1:T} = (\delta_1, \dots, \delta_T)$ and $\tilde{\delta}_{1:T}^{-t} = (\tilde{\delta}_1, \dots, \tilde{\delta}_T)$. By construction, we have

$$\begin{cases} \tilde{\delta}_i = \delta_i & \text{for all } i < t \\ \tilde{\delta}_i \geq \delta_i & \text{for all } i \geq t, \end{cases}$$

from which we obtain $f_i(\delta_{1:i-1}; \{\theta_j\}_{j \in \mathcal{I}_i}) = f_i(\tilde{\delta}_{1:i-1}; \{\theta_j\}_{j \in \mathcal{I}_i})$ for all $i \leq t$. Also, by noting that $\tilde{\delta}_t = 1$ by construction and by definition of $g(\cdot)$, we have $g(\tilde{\delta}_{1:T}^{-t}) > 0$. Hence, on the event $\{p_t \leq f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})\}$, we have $\delta_t = \tilde{\delta}_t = 1$ and hence also $\delta_{1:T} = \tilde{\delta}_{1:T}^{-t}$, which allows us to conclude that

$$\frac{\mathbb{I}\{p_t \leq f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})\}}{g(\delta_{1:T})} = \frac{\mathbb{I}\{p_t \leq f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})\}}{g(\tilde{\delta}_{1:T}^{-t})}.$$

Since $\tilde{\delta}_{1:T}^{-t}$ is independent of p_t , we take conditional expectations to obtain

$$\begin{aligned} \mathbb{E} \left[\frac{\mathbb{I}\{p_t \leq f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})\}}{g(\delta_{1:T})} \mid \mathcal{F}_{t-1} \right] &= \mathbb{E} \left[\frac{\mathbb{I}\{p_t \leq f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})\}}{g(\tilde{\delta}_{1:T}^{-t})} \mid \mathcal{F}_{t-1} \right] \\ &\stackrel{(i)}{\leq} \mathbb{E} \left[\frac{f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})}{g(\tilde{\delta}_{1:T}^{-t})} \mid \mathcal{F}_{t-1} \right] \\ &\stackrel{(ii)}{\leq} \mathbb{E} \left[\frac{f_t(\delta_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})}{g(\delta_{1:T})} \mid \mathcal{F}_{t-1} \right], \end{aligned}$$

where inequality (i) follows by taking expectation only with respect to p_t by the conditional super-uniformity property (6); and inequality (ii) follows because $g(\delta_{1:T}) \leq g(\tilde{\delta}_{1:T}^{-t})$ since

$\delta_i \leq \tilde{\delta}_i$ for all i by monotonicity of the online FDR rule. This concludes the proof of the Lemma 19. ■

Denote $C_{1:t} := (C_1, \dots, C_t)$, where $C_t = \mathbb{I}\{p_t \leq \lambda\}$, with a fixed constant $\lambda \in (0, 1)$. Furthermore, assume $\alpha_t = h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t})$, for some coordinate-wise non-decreasing function $h_t : \{0, 1\}^{2(t-1)+|\mathcal{I}_t|} \rightarrow [0, \lambda]$. Define the sigma-fields $\mathcal{J}_{t-1} = \sigma(\delta_{1:t-1}; C_{1:t-1}; \{\theta_j\}_{j \in \mathcal{I}_t})$. For independent p -values, we have the following guarantees.

Lemma 20. *Let $g : \{0, 1\}^T \rightarrow \mathbb{R}$ be any coordinate-wise non-decreasing function. Then, for any index $t \leq T$ such that $H_t \in \mathcal{H}_0$, we have*

$$\begin{aligned} & \mathbb{E} \left[\frac{h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t}) \mathbb{I}\{p_t > \lambda\}}{(1 - \lambda)g(\delta_{1:T})} \mid \mathcal{J}_{t-1} \right] \\ & \geq \mathbb{E} \left[\frac{h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t})}{g(\delta_{1:T})} \mid \mathcal{J}_{t-1} \right] \\ & \geq \mathbb{E} \left[\frac{\mathbb{I}\{p_t \leq h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t})\}}{g(\delta_{1:T})} \mid \mathcal{J}_{t-1} \right] \end{aligned}$$

Proof The second inequality is a consequence of Lemma 19, so we only prove the first inequality. Denote $p_{1:T} = (p_1, \dots, p_T)$, and $\bar{p}_{1:T}^{t \rightarrow 1} = (\bar{p}_1, \dots, \bar{p}_T)$, where

$$\bar{p}_i = \begin{cases} 1 & \text{if } i = t \\ p_i & \text{if } i \neq t, \end{cases}$$

i.e., the t -th component of $\bar{p}_{1:T}^{t \rightarrow 1}$ is set to one and the other components of $\bar{p}_{1:T}^{t \rightarrow 1}$ equals to that of $p_{1:T}$. Define $\bar{C}_i = \mathbb{I}\{\bar{p}_i \leq \lambda\}$ and $\bar{\delta}_i = \mathbb{I}\{\bar{p}_i \leq h_i(\bar{\delta}_{1:i-1}, \bar{C}_{1:i-1}, \{\theta_j\}_{j \in \mathcal{I}_i})\}$ respectively. Let $\delta_{1:T} = (\delta_1, \dots, \delta_T)$ and $\bar{\delta}_{1:T}^{t \rightarrow 1} = (\bar{\delta}_1, \dots, \bar{\delta}_T)$ denote the decision vectors using $p_{1:T}$ and $\bar{p}_{1:T}$, respectively. Similarly, $C_{1:T} = (C_1, \dots, C_T)$ and $\bar{C}_{1:T} = (\bar{C}_1, \dots, \bar{C}_T)$. By construction, we have

1. $\bar{\delta}_i = \delta_i$ and $\bar{C}_i = C_i$ for all $i \leq t$, hence $h_i(\delta_{1:i-1}, C_{1:i-1}, \{\theta_j\}_{j \in \mathcal{I}_i}) = h_i(\bar{\delta}_{1:i-1}, \bar{C}_{1:i-1}, \{\theta_j\}_{j \in \mathcal{I}_i})$ for all $i \leq t$.
2. $\bar{\delta}_t = \bar{C}_t = 0$, and hence $\bar{\delta}_i \leq \delta_i$ for all $i \geq t$, due to monotonicity of the function h_i .

Hence, on the event $\{p_t > \lambda\}$, we have $\delta_t = \bar{\delta}_t = 0$ and $C_t = \bar{C}_t = 0$, and hence also $\delta_{1:T} = \bar{\delta}_{1:T}^{t \rightarrow 1}$. Therefore, we obtain

$$\frac{h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t}) \mathbb{I}\{p_t > \lambda\}}{(1 - \lambda)g(\delta_{1:T})} = \frac{h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t}) \mathbb{I}\{p_t > \lambda\}}{(1 - \lambda)g(\bar{\delta}_{1:T})}.$$

Since $\bar{\delta}_{1:T}^{t \rightarrow 1}$ is independent of p_t , we may take conditional expectations to obtain:

$$\begin{aligned}
 & \mathbb{E} \left[\frac{h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t}) \mathbb{I}\{p_t > \lambda\}}{(1 - \lambda)g(\delta_{1:T})} \mid \mathcal{J}_{t-1} \right] \\
 &= \mathbb{E} \left[\frac{h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t}) \mathbb{I}\{p_t > \lambda\}}{(1 - \lambda)g(\bar{\delta}_{1:T}^{t \rightarrow 1})} \mid \mathcal{J}_{t-1} \right] \\
 &\stackrel{(i)}{\geq} \mathbb{E} \left[\frac{h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t})}{g(\bar{\delta}_{1:T}^{t \rightarrow 1})} \mid \mathcal{J}_{t-1} \right] \\
 &\stackrel{(ii)}{\geq} \mathbb{E} \left[\frac{h_t(\delta_{1:t-1}, C_{1:t-1}, \{\theta_j\}_{j \in \mathcal{I}_t})}{g(\delta_{1:T})} \mid \mathcal{J}_{t-1} \right],
 \end{aligned}$$

where inequality (i) follows by taking an expectation only with respect to p_t by invoking the conditional super-uniformity property (10); and inequality (ii) follows because $g(\delta_{1:T}) \geq g(\bar{\delta}_{1:T}^{t \rightarrow 1})$ since $\delta_i \geq \bar{\delta}_i$ for all i by monotonicity of the online FDR rule. This concludes the proof of Lemma 20. $\blacksquare$

Lemma 21. *Let $\mathcal{F}$ be a sigma-field and let $\mathcal{E}, \mathcal{A} \in \mathcal{F}$ be the events and $\mathcal{E} \perp \mathcal{A}$. Let Z be a random variable and $b \in \mathbb{R}$. Assume that $\Pr(\mathcal{E}) \geq 1 - \delta'$ and on $\mathcal{E}$, $\Pr(Z \leq b \mid \mathcal{F}) \geq 1 - \delta$. Then*

$$\Pr(Z \leq b \mid \mathcal{A}) \geq 1 - \delta - \delta'.$$

Proof Since $\mathcal{E} \in \mathcal{F}$,

$$\begin{aligned}
 & \Pr(Z \leq b \mid \mathcal{A}) \geq \Pr(Z \leq b, \mathcal{E} \mid \mathcal{A}) \\
 &= \mathbb{E}[\mathbb{I}_{\mathcal{E}} \Pr(Z \leq b \mid \mathcal{F}) \mid \mathcal{A}] \geq (1 - \delta) \Pr(\mathcal{E} \mid \mathcal{A}) \geq 1 - \delta - \delta'.
 \end{aligned}$$

The last inequality comes from the independence of $\mathcal{E}$ and $\mathcal{A}$ such that $\Pr(\mathcal{E} \mid \mathcal{A}) = \Pr(\mathcal{E}) \geq 1 - \delta'$. $\blacksquare$

F.2 Proof of Theorem 3

Proof

Proof of Theorem 3 (a). (mFDR control of GAIF procedures.) For the GAIF procedures,

$$\begin{aligned}
 \mathbb{E}[V(t)] &= \mathbb{E} \left[\sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{I}\{p_j \leq \alpha_j\} \right] = \sum_{j \leq t} \mathbb{E}[(1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\}] \\
 &\stackrel{(i)}{=} \sum_{j \leq t} \mathbb{E}[\mathbb{E}[(1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\} \mid \mathcal{F}_{j-1}]] \\
 &\stackrel{(ii)}{\leq} \mathbb{E} \left[\sum_{j \leq t} \alpha_j (1 - \theta_j) \right] \\
 &\stackrel{(iii)}{\leq} \mathbb{E} \left[\sum_{j \in \mathcal{I}_t} \alpha_j (1 - \theta_j) + \sum_{j \in \bar{\mathcal{I}}_t} \alpha_j \right] \\
 &\stackrel{(iv)}{\leq} \alpha \mathbb{E} \left[1 \vee \sum_{j \leq t} \delta_j \right],
 \end{aligned}$$

where equality (i) is derived by conditioning on $\mathcal{F}_{j-1}$ and applying the law of iterated expectations. Inequality (ii) is a consequence of the conditional super-uniformity property stated in (6). Subsequent inequalities (iii) follows from the fact that $1 - \theta_j \leq 1$ for all $j \in \bar{\mathcal{I}}_t$ and (iv) follows from the definition of GAIF procedures. Therefore, we obtain the conclusion that $\text{mFDR}(t) \leq \alpha$ for all GAIF procedures.

Proof of Theorem 3 (b). (FDR control of GAIF procedures.) We first prove online FDR control for GAIF procedures. Under the independence and the monotonicity assumptions, we have

$$\begin{aligned}
 \text{FDR}(t) &= \mathbb{E} \left[\frac{\sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{I}\{p_j \leq \alpha_j\}}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &= \sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{E} \left[\frac{\mathbb{I}\{p_j \leq \alpha_j\}}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &\stackrel{(i)}{=} \sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{E} \left[\mathbb{E} \left[\frac{\mathbb{I}\{p_j \leq \alpha_j\}}{1 \vee \sum_{j \leq t} \delta_j} \mid \mathcal{F}_{j-1} \right] \right] \\
 &\stackrel{(ii)}{\leq} \sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{E} \left[\frac{\alpha_j}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &= \mathbb{E} \left[\frac{\sum_{j \leq t} (1 - \theta_j) \alpha_j}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &\stackrel{(iii)}{\leq} \mathbb{E} \left[\frac{\sum_{j \in \mathcal{I}_t} (1 - \theta_j) \alpha_j + \sum_{j \in \bar{\mathcal{I}}_t} \alpha_j}{1 \vee \sum_{j=1}^t \delta_j} \right] \\
 &= \mathbb{E} \left[\widehat{\text{FDP}}_{\text{GAIF}}(t) \right] \\
 &\stackrel{(iv)}{\leq} \alpha,
 \end{aligned} \tag{34}$$

where the equality (i) follows from the law of iterated expectations by conditioning on $\mathcal{F}_{j-1}$ and inequality (ii) results from Lemma 19 by setting $g(\delta_{1:t}) = (1 \vee \sum_{j \leq t} \delta_j)$, and the inequality (iii) holds because $1 - \theta_j \leq 1$ for all $j \in \bar{\mathcal{I}}_t$, the inequality (iv) follows from definition of GAIF, which completes the proof of FDR control for monotone GAIF procedures. $\blacksquare$

F.3 Proof of Theorem 7

Proof

Proof of Theorem 7 (a). (mFDR control of Adaptive GAIF procedures.)

Now we prove mFDR control for the Adaptive GAIF procedures. For any $c \in (0, 1)$ that is predictable with respect to $\mathcal{J}_{t-1}$, the conditional super-uniformity of null p-values implies that, for $t \in \mathcal{H}_0$,

$$\mathbb{E} \left[\frac{\mathbb{I}\{p_t > c\}}{1 - c} \mid \mathcal{J}_{t-1} \right] \geq 1 \geq \mathbb{E} \left[\frac{\mathbb{I}\{p_t \leq c\}}{c} \mid \mathcal{J}_{t-1} \right]. \quad (35)$$

Note that for any time $t \in [T]$, we have

$$\begin{aligned} \mathbb{E}[V(t)] &= \mathbb{E} \left[\sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{I}\{p_j \leq \alpha_j\} \right] = \sum_{j \leq t} \mathbb{E}[(1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\}] \\ &\stackrel{(i)}{=} \sum_{j \leq t} \mathbb{E}[\mathbb{E}[(1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\} \mid \mathcal{J}_{j-1}]] \\ &\stackrel{(ii)}{\leq} \mathbb{E} \left[\sum_{j \leq t} \alpha_j (1 - \theta_j) \right] \\ &\stackrel{(iii)}{\leq} \mathbb{E} \left[\sum_{j \leq t} \alpha_j (1 - \theta_j) \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)} \right] \\ &\stackrel{(iv)}{\leq} \mathbb{E} \left[\sum_{j \in \mathcal{I}_t} \alpha_j (1 - \theta_j) \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)} + \sum_{j \in \bar{\mathcal{I}}_t} \alpha_j \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)} \right] \\ &\stackrel{(v)}{\leq} \alpha \mathbb{E} \left[1 \vee \sum_{j \leq t} \delta_j \right], \end{aligned} \quad (36)$$

where the equality (i) follows from the law of iterated expectations by conditioning on $\mathcal{J}_{j-1}$ and inequality (ii) applying the conditional super-uniformity property in (10), and the inequality (iii) also follows by the law of iterated expectations by conditioning on $\mathcal{J}_{j-1}$ and then applying the conditional super-uniformity property in (35) with $c = \lambda$, and the inequality (iv) holds since the fact that $1 - \theta_j \leq 1$ for all $j \in \bar{\mathcal{I}}_t$, and the inequality (v) follows from the construction such that $\overline{\text{FDP}}_{\text{Adaptive-GAIF}} \leq \alpha$. Therefore, we obtain the conclusion that $\text{mFDR}(t) \leq \alpha$ for Adaptive GAIF procedures.

Proof of Theorem 7 (b). (FDR control of Adaptive GAIF procedures.)

We next prove the online FDR control for Adaptive GAIF procedures. Under the independence and the monotonicity assumptions, we have

$$\begin{aligned}
 \text{FDR}(t) &= \mathbb{E} \left[\frac{\sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{I}\{p_j \leq \alpha_j\}}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &= \sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{E} \left[\frac{\mathbb{I}\{p_j \leq \alpha_j\}}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &\stackrel{(i)}{=} \sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{E} \left[\mathbb{E} \left[\frac{\mathbb{I}\{p_j \leq \alpha_j\}}{1 \vee \sum_{j \leq t} \delta_j} \mid \mathcal{J}_{j-1} \right] \right] \\
 &\stackrel{(ii)}{\leq} \sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{E} \left[\frac{\alpha_j}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &\stackrel{(iii)}{\leq} \sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{E} \left[\frac{\alpha_j}{1 \vee \sum_{j \leq t} \delta_j} \cdot \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)} \right] \\
 &= \mathbb{E} \left[\frac{\sum_{j=1}^t \alpha_j (1 - \theta_j) \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)}}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &\stackrel{(iv)}{\leq} \mathbb{E} \left[\frac{\sum_{j \in \mathcal{I}_t} \alpha_j (1 - \theta_j) \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)} + \sum_{j \in \bar{\mathcal{I}}_t} \alpha_j \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)}}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &= \mathbb{E} \left[\widehat{\text{FDP}}_{\text{Ada-GAIF}}(t) \right] \\
 &\stackrel{(v)}{\leq} \alpha,
 \end{aligned}$$

where (i) and (ii) follows from the law of iterated expectations by conditioning on $\mathcal{J}_{j-1}$ and applying Lemma 20 by setting $g(\delta_{1:t}) = (1 \vee \sum_{j \leq t} \delta_j)$, the inequality (iii) also follows from the law of iterated expectations by conditioning on $\mathcal{J}_{j-1}$ and applying Lemma 20 by setting $g(\delta_{1:t}) = (1 \vee \sum_{j \leq t} \delta_j)$, the inequality (iv) holds since $1 - \theta_j \leq 1$ for all $j \in \bar{\mathcal{I}}_t$, and the inequality (v) follows from the definition of Adaptive GAIF, which completes the proof of online FDR control for monotone Adaptive GAIF. ■

F.4 Proof of Theorem 9

Proof Define locally conditional super-uniformity as follows: if the null hypothesis H_t is true, then for all $\alpha_t \in [0, 1]$,

$$\Pr \left(p_t \leq \alpha_t \mid \mathcal{F}_{\text{dep}}^{-\mathcal{X}^t} \right) \leq \alpha_t,$$

where $\mathcal{X}^t := \{t - L_t, \dots, t - 1\}$, and $\mathcal{F}_{\text{dep}}^{-\mathcal{X}^t} := \sigma(\delta_{1:t-L_t-1}; \{\theta_j\}_{j \in \mathcal{I}_{t-L_t-1}, 1 \leq j \leq t-L_t-1})$.

This condition is immediately true by local dependence (Zrnic et al., 2021). Note that for any time $t \in [T]$, we have

$$\begin{aligned}
 \mathbb{E}[V(t)] &= \mathbb{E} \left[\sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{I}\{p_j \leq \alpha_j\} \right] = \sum_{j \leq t} \mathbb{E}[(1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\}] \\
 &= \sum_{j \leq t} \mathbb{E} \left[\mathbb{E} \left[(1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\} \mid \mathcal{F}_{\text{dep}}^{-\mathcal{X}^j} \right] \right] \\
 &\leq \sum_{j \leq t} \mathbb{E}[(1 - \theta_j) \alpha_j] = \mathbb{E} \left[\sum_{j \leq t} \alpha_j (1 - \theta_j) \right] \\
 &\leq \mathbb{E} \left[\sum_{j \in \mathcal{I}_t} \alpha_j (1 - \theta_j) + \sum_{j \in \bar{\mathcal{I}}_t} \alpha_j \right] \\
 &\leq \alpha \mathbb{E} \left[1 \vee \sum_{j \leq t, j \notin \{t-L_t, \dots, t-1\}} \delta_j \right], \\
 &\leq \alpha \mathbb{E} \left[1 \vee \sum_{j \leq t} \delta_j \right], \tag{37}
 \end{aligned}$$

where the first inequality follows from the law of iterated expectations by conditioning on $\mathcal{F}_{\text{dep}}^{-\mathcal{X}^t}$ and then applying the conditional super-uniformity property and by noticing that the measurability of α_j with respect to $\mathcal{F}_{\text{dep}}^{-\mathcal{X}^t}$, and the second inequality follows by the fact that $1 - \theta_t \leq 1$, and the third inequality follows from the construction such that $\widehat{\text{FDP}}_{\text{GAIF}_{\text{dep}}} \leq \alpha$. Therefore, we obtain the conclusion that $\text{mFDR}(t) \leq \alpha$ for GAIF_{dep} .

Define $\mathcal{J}_{\text{dep}}^{-\mathcal{X}^t} := \sigma(\delta_{1:t-L_t-1}; C_{1:t-L_t-1}; \{\theta_j\}_{j \in \mathcal{I}_{t-L_t-1}, 1 \leq j \leq t-L_t-1})$. Under local dependence, we have the conditional super-uniformity for null p -values:

$$\mathbb{E} \left[\frac{\mathbb{I}\{p_t > c\}}{1 - c} \mid \mathcal{J}_{\text{dep}}^{-\mathcal{X}^t} \right] \geq 1 \geq \mathbb{E} \left[\frac{\mathbb{I}\{p_t \leq c\}}{c} \mid \mathcal{J}_{\text{dep}}^{-\mathcal{X}^t} \right],$$

where $c \in (0, 1)$ is predictable with respect to $\mathcal{J}_{\text{dep}}^{-\mathcal{X}^t}$. The mFDR control can be obtained for Ada-GAIF_{dep} as follows.

$$\begin{aligned}
 \mathbb{E}[V(t)] &= \mathbb{E} \left[\sum_{j \leq t, j \in \mathcal{H}_0} \mathbb{I}\{p_j \leq \alpha_j\} \right] = \sum_{j \leq t} \mathbb{E}[(1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\}] \\
 &\stackrel{(i)}{=} \sum_{j \leq t} \mathbb{E} \left[\mathbb{E} \left[(1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\} \mid \mathcal{J}_{\text{dep}}^{-\mathcal{X}^j} \right] \right] \\
 &\stackrel{(ii)}{\leq} \sum_{j \leq t} \mathbb{E}[(1 - \theta_j) \alpha_j] = \mathbb{E} \left[\sum_{j \leq t} \alpha_j (1 - \theta_j) \right] \\
 &\stackrel{(iii)}{\leq} \mathbb{E} \left[\sum_{j \in \mathcal{I}_t} \alpha_j (1 - \theta_j) \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)} + \sum_{j \in \bar{\mathcal{I}}_t} \alpha_j \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)} \right] \\
 &\stackrel{(iv)}{\leq} \mathbb{E} \left[\sum_{j < t - L_t} \frac{\alpha_j \mathbb{I}\{p_j > \lambda\}}{1 - \lambda} \mathcal{W}_j + \sum_{t - L_t \leq j \leq t} \frac{\alpha_j}{1 - \lambda} \mathcal{W}_j \right] \\
 &\stackrel{(v)}{\leq} \alpha \mathbb{E} \left[1 \vee \sum_{j \leq t, j \notin \{t - L_t, \dots, t - 1\}} \delta_j \right] \\
 &\stackrel{(vi)}{\leq} \alpha \mathbb{E} \left[1 \vee \sum_{j \leq t} \delta_j \right], \tag{38}
 \end{aligned}$$

where $\mathcal{W}_j = (1 - \theta_j) \mathbb{I}\{j \in \mathcal{I}_t\} + \mathbb{I}\{j \in \bar{\mathcal{I}}_t\}$ and the equality (i) follows from the law of iterated expectations by conditioning on $\mathcal{J}_{\text{dep}}^{-\mathcal{X}^j}$, and the inequality (ii) applying the conditional super-uniformity, and the inequality (iii) also follows by the law of iterated expectations by conditioning on $\mathcal{J}_{\text{dep}}^{-\mathcal{X}^t}$ and then applying the conditional super-uniformity. The inequality (iv) follows from the fact that $\mathbb{I}\{p_j > \lambda\} \leq 1$ for all $j \in \{t - L_t, \dots, t\}$ and $1 - \theta_t \leq 1$. The inequality (v) follows from the construction such that $\widehat{\text{FDP}}_{\text{Ada-GAIF}_{\text{dep}}} \leq \alpha$, which concludes the proof. $\blacksquare$

F.5 Proof of Proposition 10

We first restate Proposition 10 as follows and finish the proof. The proof of Proposition 22 essentially follows the argument of Theorem 8.2 in Angelopoulos et al. (2024), with an extension to our online conformal testing setting.

Proposition 22 (Validity and Mutual Independence of Online Conformal p -values under Exchangeability and Symmetric Scores). *Suppose at each time t , the score function $V(\cdot; \mathcal{D}_t)$ is constructed through the current data $\mathcal{D}_t = ((\mathbf{X}_i, Y_i) : -n \leq i \leq t)$. The p -value of each time t is constructed as*

$$p_t = \frac{\sum_{i \in \mathcal{C}_{0t}} \mathbb{I}\{V(\mathbf{X}_i; \mathcal{D}_t) < V(\mathbf{X}_t; \mathcal{D}_t)\} + \xi_t \cdot (1 + \sum_{i \in \mathcal{C}_{0t}} \mathbb{I}\{V(\mathbf{X}_i; \mathcal{D}_t) = V(\mathbf{X}_t; \mathcal{D}_t)\})}{1 + |\mathcal{C}_{0t}|}. \tag{39}$$

Suppose Assumption 1 holds and the score function $V(\cdot; \mathcal{D}_t)$ is symmetric to $\{(\mathbf{X}_i, Y_i) : -n \leq i \leq t, \theta_i = 0\}$. Then under the null hypothesis,

1. Each null p -value p_t with $\theta_t = 0$ is marginally uniformly distributed on $[0, 1]$.
2. The null p -values $\{p_t : t \in \mathbb{N}, \theta_t = 0\}$ are mutually independent.

Proof We adapt the standard argument for the validity of conformal p -values to our online setting with exchangeable data. For each time t , define

$$\Phi_t = \left(\{(\mathbf{X}_i, Y_i) : i \in \mathcal{C}_{0t} \cup \{t\}\}, (\theta_i : i < t), ((\mathbf{X}_i, Y_i) : i \in \mathcal{C}_{1t}) \right),$$

where Φ_t contains the unordered set of conformity scores for indices in $\mathcal{C}_{0t} \cup \{t\}$, the true state θ_i , and the data corresponding to indices not used in the calibration set (denoted here by $\mathcal{C}_{1t}$). To prove the mutual independence, consider any time indices $t \leq T$ and let $x_t, x_{t+1}, \dots, x_T \in [0, 1]$ be arbitrary. For ease of notation, we first consider the case in which all p -values $\{p_s\}_{s=t}^T$ are associated with null hypotheses. The argument extends directly to the more general setting in which some non-null hypotheses appear between them.

Since $\{p_k\}_{k=t+1}^T$ are determined by $\Phi_t \cup \{(\mathbf{X}_i, Y_i)\}_{i=t+1}^T$ and p_t is independent of $\{(\mathbf{X}_i, Y_i)\}_{i=t+1}^T$, it holds that

$$\begin{aligned} & \Pr(p_t \leq x_t, p_{t+1} \leq x_{t+1}, \dots, p_T \leq x_T) \\ &= \mathbb{E} \left[\mathbb{I}\{p_t \leq x_t\} \mathbb{I}\{p_{t+1} \leq x_{t+1}\} \cdots \mathbb{I}\{p_T \leq x_T\} \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\mathbb{I}\{p_t \leq x_t\} \mid \Phi_t \cup \{(\mathbf{X}_i, Y_i)\}_{i=t+1}^T \right] \cdot \mathbb{I}\{p_{t+1} \leq x_{t+1}\} \cdots \mathbb{I}\{p_T \leq x_T\} \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\mathbb{I}\{p_t \leq x_t\} \mid \Phi_t \right] \cdot \mathbb{I}\{p_{t+1} \leq x_{t+1}\} \cdots \mathbb{I}\{p_T \leq x_T\} \right]. \end{aligned}$$

The key observation is that, by exchangeability, the conditional distribution of p_t given Φ_t is uniform on $[0, 1]$, so that

$$\mathbb{E} \left[\mathbb{I}\{p_t \leq x_t\} \mid \Phi_t \right] = x_t.$$

Thus,

$$\Pr(p_t \leq x_t, p_{t+1} \leq x_{t+1}, \dots, p_T \leq x_T) = x_t \mathbb{E} \left[\mathbb{I}\{p_{t+1} \leq x_{t+1}\} \cdots \mathbb{I}\{p_T \leq x_T\} \right].$$

Next, we note the following two key facts:

(a) Uniformity of single p -value p_t :

Fix any $t \in [T]$. Let Ω_t be the set of all permutations of $\mathcal{C}_0 \cup [T]$ that fix indices outside of $\mathcal{C}_{0t} \cup \{t\}$. Note that given Φ_t , the only randomness for p_t is the order of $\{(\mathbf{X}_i, Y_i) : i \in \mathcal{C}_{0t} \cup \{t\}\}$ and ξ_t .

For any $\sigma \in \Omega_t$ and given $\{\mathbf{X}_i : i \in \mathcal{C}_{0t} \cup \{t\}\} = \{\mathbf{x}_i : i \in \mathcal{C}_{0t} \cup \{t\}\}$ as the realizations, define

$$\begin{aligned}
 & p_t((\mathcal{D}_T)_\sigma) \\
 &= \frac{\sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_{\sigma(i)}; (\mathcal{D}_t)_\sigma) < V(\mathbf{x}_{\sigma(t)}; (\mathcal{D}_t)_\sigma)\} + \xi_t \cdot \sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_{\sigma(i)}; (\mathcal{D}_t)_\sigma) = V(\mathbf{x}_{\sigma(t)}; (\mathcal{D}_t)_\sigma)\}}{1 + |\mathcal{C}_{0t}|} \\
 &\stackrel{(i)}{=} \frac{\sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_{\sigma(i)}; \mathcal{D}_t) < V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t)\} + \xi_t \cdot \sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_{\sigma(i)}; \mathcal{D}_t) = V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t)\}}{1 + |\mathcal{C}_{0t}|} \\
 &\stackrel{(ii)}{=} \frac{\sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) < V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t)\} + \xi_t \cdot \sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) = V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t)\}}{1 + |\mathcal{C}_{0t}|}.
 \end{aligned}$$

Equality (i) holds since $V(\cdot; \sigma(\mathcal{D}_t)) = V(\cdot; \mathcal{D}_t)$ from the symmetry of score function V . Equality (ii) is true by the property that σ only permutes indices in $\mathcal{C}_{0t} \cup \{t\}$ such that $\{V(\mathbf{X}_i; \mathcal{D}_t) : i \in \mathcal{C}_{0t}\} = \{V(\mathbf{X}_{\sigma(i)}; \mathcal{D}_t) : i \in \mathcal{C}_{0t}\}$.

Denote $Q_\alpha(S_i : i \in \mathcal{I})$ as the α -th quantile of the set $\{S_i\}_{i \in \mathcal{I}}$. Let $\{\mathbf{X}_i : i \in \mathcal{C}_{0t} \cup \{t\}\} = \{x_i : i \in \mathcal{C}_{0t} \cup \{t\}\}$ as a set of realizations and

$$q = Q_\alpha(V(x_i; \mathcal{D}_t) : i \in \mathcal{C}_{0t} \cup \{t\}).$$

For any $\sigma \in \Omega_t$ such that $V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t) > q$, we have

$$p_t((\mathcal{D}_T)_\sigma) \geq \frac{\sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) < V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t)\}}{1 + |\mathcal{C}_{0t}|} \geq \frac{N_+ + N_-}{1 + |\mathcal{C}_{0t}|} \geq \alpha;$$

Here

$$N_- = \sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{(V(x_i; \mathcal{D}_t) < q), \quad N_+ = \sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{(V(x_i; \mathcal{D}_t) = q).$$

The last inequality holds by the property of quantile function such that $N_+ + N_- \geq 1 + \alpha(|\mathcal{C}_{0t} + 1|)$.

For any $\sigma \in \Omega_t$ such that $V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t) < q$, we have

$$p_t((\mathcal{D}_T)_\sigma) \leq \frac{\sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) \leq V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t)\}}{1 + |\mathcal{C}_{0t}|} \leq \frac{N_-}{1 + |\mathcal{C}_{0t}|} \leq \alpha,$$

which is from the property that $N_- \leq \alpha(|\mathcal{C}_{0t} + 1|)$.

And for any $\sigma \in \Omega_t$ such that $V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t) = q$, we have

$$\begin{aligned}
 p_t((\mathcal{D}_T)_\sigma) &= \frac{\sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) < V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t)\} + \xi_t \cdot \sum_{i \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) = V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t)\}}{1 + |\mathcal{C}_{0t}|} \\
 &\leq \frac{N_- + \xi_t N_+}{1 + |\mathcal{C}_{0t}|}.
 \end{aligned}$$

Hence,

$$\begin{aligned}
 \Pr(p_t \leq \alpha \mid \Phi_t) &= \frac{1}{(|\mathcal{C}_{0t}| + 1)!} \sum_{\sigma \in \Omega_t} \Pr(p_t((\mathcal{D}_T)_\sigma) \leq \alpha) \\
 &\stackrel{(i)}{=} \frac{1}{(|\mathcal{C}_{0t}| + 1)!} \sum_{\sigma \in \Omega_t} \left(\mathbb{I}\{V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t) < q\} + \mathbb{I}\{V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t) = q\} \frac{\alpha(|\mathcal{C}_{0t}| + 1) - N_-}{N_-} \right) \\
 &\stackrel{(ii)}{=} \frac{1}{|\mathcal{C}_{0t}| + 1} \sum_{i \in \mathcal{C}_{0t} \cup \{j\}} \left(\mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) < q\} + \mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) = q\} \frac{\alpha(|\mathcal{C}_{0t}| + 1) - N_-}{N_-} \right) \\
 &\stackrel{(iii)}{=} \frac{1}{|\mathcal{C}_{0t}| + 1} \left(N_- + N_+ \frac{\alpha(|\mathcal{C}_{0t}| + 1) - N_-}{N_-} \right) = \alpha,
 \end{aligned}$$

where equality (i) holds since ξ_t is uniformly distributed and independent of everything else, making

$$\mathbb{E}[\mathbb{I}\{V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t) = q, p_t((\mathcal{D}_T)_\sigma) \leq \alpha\} \mid (\mathcal{D}_T)_\sigma] = \mathbb{I}\{V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t) = q\} \frac{\alpha(|\mathcal{C}_{0t}| + 1) - N_-}{N_-}.$$

Equality (ii) is from the fact that $\sum_{\sigma \in \Omega_t, \sigma(t)=i} \mathbb{I}\{V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t) < q\} = |\mathcal{C}_{0t}| \mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) < q\}$ and $\sum_{\sigma \in \Omega_t, \sigma(t)=i} \mathbb{I}\{V(\mathbf{x}_{\sigma(t)}; \mathcal{D}_t) = q\} = |\mathcal{C}_{0t}| \mathbb{I}\{V(\mathbf{x}_i; \mathcal{D}_t) = q\}$. And equality (iii) is direct by the definition of N_- and N_+ . Marginalizing over the Φ_t implies $\Pr(p_t \leq \alpha) = \alpha$ for all $t \in [T]$.

(b) Independence of future p -values from Φ_t :

Define the data set $\mathcal{D}_T = ((\mathbf{X}_{-n+1}, Y_{-n+1}), \dots, (\mathbf{X}_T, Y_T))$. By exchangeability of the data, it holds that $\mathcal{D}_T \stackrel{d}{=} (\mathcal{D}_T)_\sigma$, where $(\mathcal{D}_T)_\sigma$ is obtained from $\mathcal{D}_T$ by permuting the data points according to σ . A key observation is that for any $\sigma \in \Omega_t$, defining σ as above we have

$$p_{t'}(\mathcal{D}_T) = p_{t'}((\mathcal{D}_T)_\sigma)$$

for all $t' \in \{t+1, \dots, T\}$. Intuitively, this indicates that permuting the data according to σ does not change p -values after time t .

This is because $p_{t'}((\mathcal{D}_T)_\sigma)$ is

$$\begin{aligned}
 &\frac{\sum_{i \in \mathcal{C}_{0t'}} \mathbb{I}\{V(\mathbf{X}_{\sigma(i)}; \sigma(\mathcal{D}_{t'})) < V(\mathbf{X}_{\sigma(t')}; \sigma(\mathcal{D}_{t'})) + \xi_{t'}(1 + \sum_{i \in \mathcal{C}_{0t'}} \mathbb{I}\{V(\mathbf{X}_{\sigma(i)}; \sigma(\mathcal{D}_{t'})) = V(\mathbf{X}_{\sigma(t')}; \sigma(\mathcal{D}_{t'}))\})}{1 + |\mathcal{C}_{0t'}|} \\
 &\stackrel{(i)}{=} \frac{\sum_{i \in \mathcal{C}_{0t'}} \mathbb{I}\{V(\mathbf{X}_{\sigma(i)}; \mathcal{D}_{t'}) < V(\mathbf{X}_{t'}; \mathcal{D}_{t'})\} + \xi_{t'}(1 + \sum_{i \in \mathcal{C}_{0t'}} \mathbb{I}\{V(\mathbf{X}_{\sigma(i)}; \mathcal{D}_{t'}) = V(\mathbf{X}_{t'}; \mathcal{D}_{t'})\})}{1 + |\mathcal{C}_{0t'}|} \\
 &\stackrel{(ii)}{=} \frac{\sum_{i \in \mathcal{C}_{0t'}} \mathbb{I}\{V(\mathbf{X}_i; \mathcal{D}_{t'}) < V(\mathbf{X}_{t'}; \mathcal{D}_{t'})\} + \xi_{t'}(1 + \sum_{i \in \mathcal{C}_{0t'}} \mathbb{I}\{V(\mathbf{X}_i; \mathcal{D}_{t'}) = V(\mathbf{X}_{t'}; \mathcal{D}_{t'})\})}{1 + |\mathcal{C}_{0t'}|} \\
 &= p_{t'}(\mathcal{D}_T),
 \end{aligned}$$

where equality (i) holds since $\sigma(t') = t'$ by definition and $V(\cdot; \sigma(\mathcal{D}_{t'})) = V(\cdot; \mathcal{D}_{t'})$ from the symmetry of score function V . And equality (ii) is true as $\{V(\mathbf{X}_i; \mathcal{D}_{t'}) : i \in \mathcal{C}_{0t'}\} = \{V(\mathbf{X}_{\sigma(i)}; \mathcal{D}_{t'}) : i \in \mathcal{C}_{0t'}\}$.

Therefore,

$$\begin{aligned}
 & \Pr(p_t \leq x_t, \dots, p_T \leq x_T) \\
 &= \frac{1}{(|\mathcal{C}_{0t}| + 1)!} \mathbb{E} \left[\sum_{\sigma \in \Omega_t} \mathbb{I}\{p_t((\mathcal{D}_T)_\sigma) \leq x_t, \dots, p_T((\mathcal{D}_T)_\sigma) \leq x_T\} \right] \\
 &= \frac{1}{(|\mathcal{C}_{0t}| + 1)!} \mathbb{E} \left[\sum_{\sigma \in \Omega_t} \mathbb{I}\{p_t(\mathcal{D}_T) \leq x_t, \dots, p_T(\mathcal{D}_T) \leq x_T\} \right] \\
 &= \frac{1}{(|\mathcal{C}_{0t}| + 1)!} \mathbb{E} \left[\left(\sum_{\sigma \in \Omega_t} \mathbb{I}\{p_t(\mathcal{D}_T) \leq x_t\} \right) \mathbb{I}\{p_{t+1}(\mathcal{D}_T) \dots, p_T(\mathcal{D}_T) \leq x_T\} \right] \\
 &= \mathbb{E} [\Pr(p_t \leq x_t \mid \Phi_t \cup \{(\mathbf{X}_i, Y_i)\}_{i=t+1}^T) \mathbb{I}\{p_{t+1}(\mathcal{D}_T) \dots, p_T(\mathcal{D}_T) \leq x_T\}] \\
 &= x_t \Pr(p_{t+1}(\mathcal{D}_T) \dots, p_T(\mathcal{D}_T) \leq x_T).
 \end{aligned}$$

Repeating the above argument recursively for $s = t, t+1, \dots, T$, and conditioning at step s on $\Phi_s \cup \{(\mathbf{X}_i, Y_i)\}_{i=s}^T$, we obtain

$$\Pr(p_t \leq x_t, p_{t+1} \leq x_{t+1}, \dots, p_T \leq x_T) = x_t x_{t+1} \dots x_T.$$

Since the joint cumulative distribution function factors as the product of the marginals, it follows that the sequence $\{p_t, \dots, p_T\}$ is mutually independent, with each p_t marginally distributed as Uniform(0, 1).

Thus, we have demonstrated that under the exchangeability assumption and the online updating scheme, the online conformal p -values defined in (39) are mutually independent. ■

F.6 Proof of Theorem 11: finite sample mFDR control

Proof mFDR control for Algorithm 2. We first prove the results for LFS. Denote $\Psi_t = ((p_k : k \in \mathcal{C}_{0t}), (\theta_k : k < t), ((X_k, Y_k) : k \in \mathcal{C}_{1t}))$. We need to verify two facts, for any time $t \in [T]$,

(i): α_t^{LFS} is fixed given Ψ_t . Since Ψ_t contains all non-null information available up to time t , we can regard the p -values $\{p_i\}_{i \leq t}$ as being constructed without using any non-null data; the general case then follows directly by conditioning on Ψ_t as non-null data are all fixed. Note that α_t^{LFS} is fixed given all past null decisions $(\delta_i : 1 \leq i < t)$ and a null decision δ_i is decided by $((p_k, \alpha_k) : \theta_k = 0, 1 \leq k \leq i)$. By iterated discussion, α_t is decided by null p -values $(p_k : k \in \mathcal{C}_{0t})$, thereby determined fully by Ψ_t .

(ii): p_t is super-uniform given Ψ_t and $\theta_t = 0$. This is direct as long as p_t is independent of past null p -values given non-null data, which is verified by Proposition 10.

Then we have

$$\begin{aligned}
 \mathbb{E}[V(t)] &= \mathbb{E} \left[\sum_{j \leq t} (1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\} \right] \\
 &= \sum_{j \leq t} \mathbb{E} [(1 - \theta_j) \mathbb{E} [\mathbb{I}\{p_j \leq \alpha_j\} \mid \Psi_j, \theta_j = 0]] \\
 &\stackrel{(i)}{\leq} \mathbb{E} \left[\sum_{j \leq t} \alpha_j (1 - \theta_j) \right] \\
 &\stackrel{(ii)}{\leq} \mathbb{E} \left[\sum_{j=1}^{t-1} \alpha_j (1 - \theta_j) + \alpha_t \right] \\
 &\stackrel{(iii)}{\leq} \alpha \mathbb{E} \left[1 \vee \sum_{j \leq t} \delta_j (1 - \theta_j) \right] \\
 &\stackrel{(iv)}{\leq} \alpha \cdot \mathbb{E}[1 \vee R(t)], \tag{40}
 \end{aligned}$$

where the inequality (i) follows from the above two facts, thereby $\Pr(p_j \leq \alpha_j \mid \Psi_j, \theta_j = 0) \leq \alpha_j$. And the inequality (ii) follows by the fact that $1 - \theta_t \leq 1$. For inequality (iii) follows from the construction of α_t^{LFS} , and the inequality (iv) holds since $1 - \theta_j \leq 1$ for all $j \leq t$. Therefore, we conclude that $\text{mFDR}(t) \leq \alpha$ for LFS.

Similar results for SFS can also be proved as follows:

$$\begin{aligned}
 \mathbb{E}[V(t)] &= \mathbb{E} \left[\sum_{j \leq t} (1 - \theta_j) \mathbb{I}\{p_j \leq \alpha_j\} \right] \\
 &= \sum_{j \leq t} \mathbb{E} [(1 - \theta_j) \mathbb{E} [\mathbb{I}\{p_j \leq \alpha_j\} \mid \Psi_j, \theta_j = 0]] \\
 &\leq \mathbb{E} \left[\sum_{j \leq t} \alpha_j (1 - \theta_j) \right] \\
 &\leq \mathbb{E} \left[\sum_{j=1}^{t-1} \alpha_j (1 - \theta_j) \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)} + \alpha_t (1 - \theta_t) \frac{\mathbb{I}\{p_t > \lambda\}}{(1 - \lambda)} \right] \\
 &\leq \mathbb{E} \left[\sum_{j=1}^{t-1} \alpha_j (1 - \theta_j) \frac{\mathbb{I}\{p_j > \lambda\}}{(1 - \lambda)} + \alpha_t \frac{\mathbb{I}\{p_t > \lambda\}}{(1 - \lambda)} \right] \\
 &\leq \alpha \mathbb{E} \left[1 \vee \sum_{j \leq t} \delta_j (1 - \theta_j) \right] \\
 &\leq \alpha \cdot \mathbb{E}[1 \vee R(t)].
 \end{aligned}$$

Therefore, we obtain the conclusion that $\text{mFDR}(t) \leq \alpha$ for the proposed OCTF procedures in Algorithm 2 with $\alpha_t = \alpha_t^{\text{LFS}}$ or $\alpha_t = \alpha_t^{\text{SFS}}$. $\blacksquare$

F.7 Proof of Corollary 13

Proof It suffices to verify that the optimized score function $V(\cdot; \hat{k}_t)$ is symmetric with respect to $\mathcal{C}_{0t} \cup \{t\}$, so that Proposition 22 and Theorem 11 can be directly applied to establish the validity of mFDR control at the target level for Algorithm Opt-OCTF.

For any permutation $\sigma \in \Omega_t$ that only permutes the indices in $\mathcal{C}_{0t} \cup \{t\}$, we have

$$\hat{k}_t^\sigma = \arg \min_{k \in [K]} \mathcal{M}_t^{\text{EWMA}}((\mathcal{D}_T)_\sigma; k) = \hat{k}_t.$$

To see why, the auxiliary p-value for $j \in \mathcal{C}_{1t}$ after permutation σ is

$$\begin{aligned} (\tilde{p}_{k,j})_\sigma &= \frac{\sum_{s \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{X}_{\sigma(s)}; k) \leq V(\mathbf{X}_j; k)\}}{1 + |\mathcal{C}_{0t}|} \\ &= \frac{\sum_{s \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{X}_s; k) \leq V(\mathbf{X}_j; k)\}}{1 + |\mathcal{C}_{0t}|} = \tilde{p}_{k,j}. \end{aligned}$$

This means $\{\tilde{p}_{k,j}\}_{j \in \mathcal{C}_{1t}}$ is permutation invariant to σ . Applying this, we have

$$\mathcal{M}_t^{\text{EWMA}}((\mathcal{D}_t)_\sigma; k) = \frac{\sum_{j=1}^{t-1} \rho^{t-1-j} (\tilde{p}_{k,j})_\sigma \theta_j}{\sum_{j=1}^{t-1} \rho^{t-1-j} \cdot \theta_j} = \frac{\sum_{j=1}^{t-1} \rho^{t-1-j} \tilde{p}_{k,j} \theta_j}{\sum_{j=1}^{t-1} \rho^{t-1-j} \cdot \theta_j} = \mathcal{M}_t^{\text{EWMA}}(\mathcal{D}_t; k),$$

which keeps invariant to the permutation σ . Combining together, we have $\hat{k}_t$ is symmetric to $\mathcal{C}_{0t} \cup \{t\}$. And it indicates that $V(\cdot; \hat{k}_t)$ is symmetric to $\mathcal{C}_{0t} \cup \{t\}$. $\blacksquare$

F.8 Proof of Theorem 16

Proof We first prove the results for e -GAIF. Since $\delta_j = \mathbb{I}\{e_j \geq 1/\alpha_j\}$, by definition, we have

$$\begin{aligned}
 \text{FDR}(t) &= \mathbb{E} \left[\frac{\sum_{j \leq t} \mathbb{I}\{e_j \geq 1/\alpha_j\} (1 - \theta_j)}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &\leq \mathbb{E} \left[\sum_{j \leq t} \frac{(1 - \theta_j) \mathbb{I}\{1/e_j \leq \alpha_j\}}{R(j-1) + 1} \right] \\
 &\leq \mathbb{E} \left[\sum_{j \leq t} \frac{(1 - \theta_j) e_j \alpha_j}{R(j-1) + 1} \right] \\
 &= \mathbb{E} \left[\sum_{j \leq t} \frac{(1 - \theta_j) \mathbb{E}[e_j \mid \mathcal{F}_{j-1}] \alpha_j}{R(j-1) + 1} \right] \\
 &\leq \mathbb{E} \left[\sum_{j \leq t} \frac{(1 - \theta_j) \alpha_j}{R(j-1) + 1} \right] \\
 &\leq \mathbb{E} \left[\sum_{j \in \mathcal{I}_t} \frac{\alpha_j (1 - \theta_j)}{1 + R(j-1)} + \sum_{j \in \bar{\mathcal{I}}_t} \frac{\alpha_j}{1 + R(j-1)} \right] \\
 &= \mathbb{E} \left[\widehat{\text{FDP}}_{e\text{-GAIF}}(t) \right] \leq \alpha,
 \end{aligned}$$

where the first inequality holds since $R(j-1)+1 \leq (R(t) \vee 1)$ for every $j \in \{j \leq t : \delta_j = 1\}$ by definition, the second inequality holds since $\mathbb{I}\{e_j \alpha_j \geq 1\} \leq e_j \alpha_j$, the third inequality uses the law of iterated expectations by conditioning on $\mathcal{F}_{j-1}$ and then applies the property of e -values, and the fourth inequality holds since $1 - \theta_t \leq 1$, and the last inequality follows from the construction of e -GAIF, which completes the proof of FDR control for e -GAIF.

We now proceed to establish the FDR control guarantee for the e -Ada-GAIF procedure. Specifically, we show that:

$$\begin{aligned}
 \text{FDR}(t) &= \mathbb{E} \left[\frac{\sum_{j \leq t} \mathbb{I}\{e_j \geq 1/\alpha_j\}(1 - \theta_j)}{1 \vee \sum_{j \leq t} \delta_j} \right] \\
 &\leq \mathbb{E} \left[\sum_{j \leq t} \frac{(1 - \theta_j) \mathbb{I}\{1/e_j \leq \alpha_j\}}{R(j-1) + 1} \right] \\
 &\leq \mathbb{E} \left[\sum_{j \leq t} \frac{(1 - \theta_j) e_j \alpha_j}{R(j-1) + 1} \right] \\
 &= \mathbb{E} \left[\sum_{j \leq t} \frac{(1 - \theta_j) \mathbb{E}[e_j \mid \mathcal{J}_{j-1}] \alpha_j}{R(j-1) + 1} \right] \\
 &\leq \mathbb{E} \left[\sum_{j \leq t} \frac{(1 - \theta_j) \alpha_j}{R(j-1) + 1} \right] \\
 &\leq \sum_{j \leq t} \mathbb{E} \left[\frac{(1 - \theta_j) \alpha_j}{R(j-1) + 1} \cdot \frac{\mathbb{E}[\mathbb{I}\{e_j < 1/\lambda\} \mid \mathcal{J}_{j-1}]}{1 - \lambda} \right] \\
 &= \sum_{j \leq t} \mathbb{E} \left[\frac{(1 - \theta_j) \alpha_j}{R(j-1) + 1} \cdot \frac{\mathbb{I}\{e_j < 1/\lambda\}}{1 - \lambda} \right] \\
 &\leq \mathbb{E} \left[\sum_{j \in \mathcal{I}_t} \frac{\alpha_j (1 - \theta_j)}{R(j-1) + 1} \cdot \frac{\mathbb{I}\{e_j \leq 1/\lambda\}}{1 - \lambda} + \sum_{j \in \bar{\mathcal{I}}_t} \frac{\alpha_j}{R(j-1) + 1} \cdot \frac{\mathbb{I}\{e_j \leq 1/\lambda\}}{1 - \lambda} \right] \\
 &= \mathbb{E} \left[\widehat{\text{FDP}}_{e\text{-Ada-GAIF}}(t) \right] \leq \alpha,
 \end{aligned}$$

where the first inequality holds since $R(j-1) + 1 \leq (R(t) \vee 1)$ for every $j \in \{j \leq t : \delta_j = 1\}$ by definition, the second inequality holds since $\mathbb{I}\{e_j \alpha_j \geq 1\} \leq e_j \alpha_j$, the third inequality uses the law of iterated expectations by conditioning on $\mathcal{F}_{j-1}$ and then applies the property of e -values, and the fourth inequality holds since $\mathbb{E}[\mathbb{I}\{e_j < 1/\lambda\} \mid \mathcal{J}_{j-1}] \geq 1 - \lambda$ by the property of e -values, the fifth inequality follows from $1 - \theta_t \leq 1$, and the last inequality follows from the construction of e -Ada-GAIF, which completes the proof of FDR control for e -Ada-GAIF. ■

F.9 Proof of Corollary 17

Proof We prove the result in two steps.

Step 1: mFDR control within each sub-stream. Fix a sub-stream index $j \in \{1, 2, \dots, d+1\}$ and consider the sequence of testing times $\{t_k\}_{k \geq 1}$ such that $t_k \in S_j$. By construction of the sub-streams, the feedback associated with test t_k is revealed at time t_{k+1} , which belongs to the same sub-stream. Therefore, when conditioning on the sub-stream

index j , the testing procedure restricted to $\{t_k\}_{k \geq 1}$, using an adjusted level $\alpha/(d+1)$, has fully observed feedback and coincides with the standard OCTF procedure (without delay).

Under Assumption 1, the conformal p -values $\{p_{t_k}\}_{k \geq 1}$ computed using the calibration sets $\{\mathcal{C}_{0t}^j\}$ are super-uniform under the null, conditional on the past within the same sub-stream. Moreover, the test levels $\{\alpha_{t_k}\}_{k \geq 1}$ are updated based only on information available prior to t_k within the sub-stream. Hence, by the proof of finite-sample mFDR guarantee for OCTF in Appendix F.6, the rejection set restricted to sub-stream S_j , denoted by $\mathcal{R}^{(j)} = \mathcal{R} \cap S_j$, satisfies

$$\mathbb{E}[V^{(j)}(t)] \leq \frac{\alpha}{d+1} \mathbb{E}[1 \vee R^{(j)}(t)].$$

and thus

$$\text{mFDR}^{(j)}(t) \leq \alpha/(d+1) \leq \alpha \quad \text{for all } t \in S_j.$$

Step 2: Aggregation across sub-streams. Let $V(t)$ and $R(t)$ denote the numbers of false discoveries and total rejections up to time t , respectively. Since the sub-streams $\{S_j\}_{j=1}^{d+1}$ form a partition of $\mathbb{N}$,

$$V(t) = \sum_{j=1}^{d+1} V^{(j)}(t), \quad R(t) = \sum_{j=1}^{d+1} R^{(j)}(t).$$

By Step 1,

$$\mathbb{E}[V(t)] = \sum_{j=1}^{d+1} \mathbb{E}[V^{(j)}(t)] \leq \frac{\alpha}{d+1} \sum_{j=1}^{d+1} \mathbb{E}[1 \vee R^{(j)}(t)].$$

Since $R^{(j)}(t) \leq R(t)$ pointwise and $1 \vee (\cdot)$ is monotone,

$$1 \vee R^{(j)}(t) \leq 1 \vee R(t),$$

for every j . Hence,

$$\sum_{j=1}^{d+1} \mathbb{E}[1 \vee R^{(j)}(t)] \leq (d+1) \mathbb{E}[1 \vee R(t)].$$

Therefore,

$$\mathbb{E}[V(t)] \leq \alpha \mathbb{E}[1 \vee R(t)],$$

which implies

$$\text{mFDR}(t) = \frac{\mathbb{E}[V(t)]}{\mathbb{E}[R(t) \vee 1]} \leq \alpha.$$

■

F.10 Proof of Theorem 15

Proof For simplicity, we define that when $\theta_t = 1$,

$$\hat{M}_{k,t} := \mathcal{M}_t^{\text{EWMA}}(\mathcal{D}_t; k) = \frac{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j} \tilde{p}_{k,j}}{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j}} \quad \text{and} \quad M_{k,t} := \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_t; k)) \mid \theta_t = 1].$$

And we set historical data $\mathcal{D}_C$ as empty set for notational convenience. The final guarantee will remain valid by replacing t with $n + t$ when considering historical samples.

We decompose the error by

$$\begin{aligned} \hat{M}_{k,t} - M_{k,t} &= \underbrace{\frac{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j} (\tilde{p}_{k,j} - F_k^{(0)}(V(\mathbf{X}_j; k)))}{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j}}}_{\text{(I) null CDF estimation}} \\ &+ \underbrace{\frac{\sum_{j < t} \rho^{t-1-j} \theta_t (F_k^{(0)}(V(\mathbf{X}_j; k)) - \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_j; k)) \mid \theta_j = 1])}{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j}}}_{\text{(II) gap of EWMA estimation}} \\ &+ \underbrace{\frac{\sum_{j < t} \rho^{t-1-j} \theta_j (\mathbb{E}[F_k^{(0)}(V(\mathbf{X}_j; k)) \mid \theta_j = 1] - \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_t; k)) \mid \theta_t = 1])}{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j}}}_{\text{(III) drift bias}}. \end{aligned} \tag{41}$$

Then we check the three term one by one.

(I) Null CDF estimation. Firstly, we condition on the past states $(\theta_s)_{s < t}$ and $\{\theta_t = 1\}$ to give results. In this case, $\mathcal{C}_{0t}$ and $\mathcal{C}_{1t}$ are fixed. Denote

$$\hat{F}_{k,t}^{(0)}(v) = \frac{\sum_{s \in \mathcal{C}_{0t} \cup \{t\}} \mathbb{I}\{V(\mathbf{X}_s; k) \leq v\}}{1 + |\mathcal{C}_{0t}|}.$$

On the event $\{\theta_t = 1\}$, the score $V(\mathbf{X}_t; k)$ is drawn from the non-null distribution and thus is not a null sample for estimating F_0^k . Let

$$\hat{F}_{k,t}^{(0),\text{null}}(v) := \frac{1}{|\mathcal{C}_{0t}|} \sum_{s \in \mathcal{C}_{0t}} \mathbb{I}\{V(\mathbf{X}_s; k) \leq v\}.$$

Then we can decompose $\hat{F}_{k,t}^{(0)}$ as

$$\hat{F}_{k,t}^{(0)}(v) = \frac{|\mathcal{C}_{0t}|}{|\mathcal{C}_{0t}| + 1} \hat{F}_{k,t}^{(0),\text{null}}(v) + \frac{1}{|\mathcal{C}_{0t}| + 1} \mathbb{I}\{V(\mathbf{X}_t; k) \leq v\}.$$

Therefore, for all v ,

$$\begin{aligned} |\hat{F}_{k,t}^{(0)}(v) - F_k^{(0)}(v)| &\leq \frac{|\mathcal{C}_{0t}|}{|\mathcal{C}_{0t}| + 1} |\hat{F}_{k,t}^{(0),\text{null}}(v) - F_k^{(0)}(v)| + \frac{1}{|\mathcal{C}_{0t}| + 1} |\mathbb{I}\{V(\mathbf{X}_t; k) \leq v\} - F_k^{(0)}(v)| \\ &\leq \sup_u |\hat{F}_{k,t}^{(0),\text{null}}(u) - F_k^{(0)}(u)| + \frac{1}{|\mathcal{C}_{0t}| + 1}, \end{aligned}$$

since $|\mathbb{I}\{V(\mathbf{X}_t; k) \leq v\} - F_k^{(0)}(v)| \leq 1$.

Conditional on $(\theta_s)_{s < t}$, the collection $\{V(\mathbf{X}_s; k)\}_{s \in \mathcal{C}_{0t}}$ are i.i.d. from $F_k^{(0)}$. By the DKW inequality, for any $\varepsilon_1 > 0$,

$$\Pr\left(\sup_v |\hat{F}_{k,t}^{(0),\text{null}}(v) - F_k^{(0)}(v)| \leq \varepsilon_1 \mid (\theta_s)_{s < t}, \theta_t = 1\right) \geq 1 - 2 \exp\{-2|\mathcal{C}_{0t}|\varepsilon_1^2\}.$$

Combining the two displays yields

$$\Pr\left(\sup_v |\hat{F}_{k,t}^{(0)}(v) - F_k^{(0)}(v)| \leq \varepsilon_1 + \frac{1}{|\mathcal{C}_{0t}| + 1} \mid (\theta_s)_{s < t}, \theta_t = 1\right) \geq 1 - 2 \exp\{-2|\mathcal{C}_{0t}|\varepsilon_1^2\}.$$

As $\tilde{p}_{k,j} = \hat{F}_{k,t}^{(0)}(V(\mathbf{X}_j; k))$, we further have

$$\Pr\left(\sup_{j \in \mathcal{C}_{1t}} |\tilde{p}_{k,j} - F_k^{(0)}(V(\mathbf{X}_j; k))| \leq \varepsilon_1 + \frac{1}{|\mathcal{C}_{0t}| + 1} \mid (\theta_s)_{s < t}, \theta_t = 1\right) \geq 1 - 2 \exp\{-2|\mathcal{C}_{0t}|\varepsilon_1^2\}.$$

Since $|\mathcal{C}_{0t}| = \sum_{s=1}^{t-1} \mathbb{I}\{\theta_s = 0\}$ and $\Pr(\theta_s = 0) \geq \pi$, by Hoeffding's inequality,

$$\Pr\left(|\mathcal{C}_{0t}| \geq (t-1)\pi/2\right) = \Pr\left(|\mathcal{C}_{0t}| - (t-1)\pi \geq -(t-1)\pi/2\right) \geq 1 - \exp\{-(t-1)\pi^2/2\}.$$

Therefore, on the event $\mathbb{P}_t = \{|\mathcal{C}_{0t}| \geq (t-1)\pi/2\}$, we have

$$\Pr\left(\sup_{j \in \mathcal{C}_{1t}} |\tilde{p}_{k,j} - F_k^{(0)}(V(\mathbf{X}_j; k))| \leq \varepsilon_1 + \frac{2}{(t-1)\pi} \mid \mathbb{P}_t, \theta_t = 1\right) \geq 1 - 2 \exp\{-\{(t-1)\pi\}\varepsilon_1^2\}.$$

Then by the independence of $\mathbb{P}_t$ and $\{\theta_t = 1\}$, invoking Lemma 21 leads to

$$\begin{aligned} & \Pr\left(\left|\frac{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j} (\tilde{p}_{k,j} - F_k^{(0)}(V(\mathbf{X}_j; k)))}{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j}}\right| \leq \varepsilon_1 + \frac{2}{(t-1)\pi} \mid \theta_t = 1\right) \\ & \geq 1 - 2 \exp\{-\{(t-1)\pi\}\varepsilon_1^2\} - \exp\{-(t-1)\pi^2/2\}. \end{aligned} \quad (42)$$

(II) Gap of EWMA estimation. Define $\Omega_t = \sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j} = \sum_{j=1}^{t-1} \rho^{t-1-j} \theta_j$. It follows that $\mathbb{E}[\Omega_t] = \sum_{j=1}^{t-1} \rho^{t-1-j} \mathbb{E}[\theta_j] \geq (1 - \pi)(1 - \rho^{t-1})/(1 - \rho)$. By Hoeffding inequality, we have

$$\Pr\left(\Omega_t \geq \frac{1 - \pi}{2} \frac{1 - \rho^{t-1}}{1 - \rho}\right) \geq 1 - \exp\left\{-\frac{(1 - \pi)^2(1 - \rho^2)}{2(1 - \rho^{2(t-1)})}\right\}. \quad (43)$$

Then we can apply the Hoeffding inequality to the weighted sum as

$$\left\{\theta_j(F_k^{(0)}(V(\mathbf{X}_j; k)) - \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_j; k)) \mid \theta_j = 1])\right\}_{j < t}$$

are bounded between $[0, 1]$ and independent:

$$\begin{aligned}
 & \Pr \left(\left| \sum_{j < t} \rho^{t-1-j} \theta_j (F_k^{(0)}(V(\mathbf{X}_j; k)) - \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_j; k)) \mid \theta_j = 1]) \right| \leq \varepsilon_2 \Omega_t \right) \\
 & \geq \Pr \left(\left| \sum_{j < t} \rho^{t-1-j} \theta_j (F_k^{(0)}(V(\mathbf{X}_j; k)) - \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_j; k)) \mid \theta_j = 1]) \right| \leq \varepsilon_2 \Omega_t, \Omega_t \geq \frac{1 - \bar{\pi}}{2} \frac{1 - \rho^{t-1}}{1 - \rho} \right) \\
 & \geq 1 - \Pr \left(\left| \sum_{j < t} \rho^{t-1-j} \theta_j (F_k^{(0)}(V(\mathbf{X}_j; k)) - \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_j; k)) \mid \theta_j = 1]) \right| \geq \varepsilon_2 \left\{ \frac{1 - \bar{\pi}}{2} \frac{1 - \rho^{t-1}}{1 - \rho} \right\} \right) \\
 & \quad - \Pr \left(\Omega_t \leq \frac{1 - \bar{\pi}}{2} \frac{1 - \rho^{t-1}}{1 - \rho} \right) \\
 & \geq 1 - 2 \exp \left\{ - \frac{2\varepsilon_2^2(1 - \rho^2) \left\{ \frac{1 - \bar{\pi}}{2} \frac{1 - \rho^{t-1}}{1 - \rho} \right\}^2}{1 - \rho^{2(t-1)}} \right\} - \exp \left\{ - \frac{(1 - \bar{\pi})^2(1 - \rho^2)}{2(1 - \rho^{2(t-1)})} \right\}. \tag{44}
 \end{aligned}$$

(III) Drift bias. By Stieltjes integrals,

$$\begin{aligned}
 \left| \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_j; k)) \mid \theta_j = 1] - \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_t; k)) \mid \theta_t = 1] \right| &= \left| \int F_k^{(0)}(v) dF_{k,j}^{(1)}(v) - \int F_k^{(0)}(v) dF_{k,t}^{(1)}(v) \right| \\
 &= \left| \int (F_{k,j}^{(1)}(v) - F_{k,t}^{(1)}(v)) dF_k^{(0)}(v) \right| \\
 &\leq \|F_{k,j}^{(1)} - F_{k,t}^{(1)}\|_{\infty}.
 \end{aligned}$$

So relating it with our slowly varying distribution condition leads to

$$\begin{aligned}
 & \left| \frac{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j} (\mathbb{E}[F_k^{(0)}(V(\mathbf{X}_j; k)) \mid \theta_j = 1] - \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_t; k)) \mid \theta_t = 1])}{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j}} \right| \\
 & \leq \gamma \frac{\sum_{j=1}^{t-1} \rho^{t-1-j} (t-j)}{\Omega_t} \leq \gamma \frac{1}{\Omega_t(1 - \rho)^2}.
 \end{aligned}$$

The last inequality comes from the fact that $\sum_{j=1}^{t-1} (t-j) \rho^{t-1-j} = \{1 - t\rho^{t-1} + (t-1)\rho^t\} / (1 - \rho)^2 \leq 1 / (1 - \rho)^2$.

Since we have already analyzed the high-probability lower bound of Ω_t in (43), it follows that

$$\Pr(|\bar{\Delta}_t| \leq 2\gamma \{(1 - \bar{\pi})(1 - \rho^{t-1})(1 - \rho)\}^{-1}) \geq 1 - \exp \left\{ - \frac{(1 - \bar{\pi})^2(1 - \rho^2)}{2(1 - \rho^{2(t-1)})} \right\}, \tag{45}$$

where

$$\bar{\Delta}_t := \frac{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j} (\mathbb{E}[F_k^{(0)}(V(\mathbf{X}_j; k)) \mid \theta_j = 1] - \mathbb{E}[F_k^{(0)}(V(\mathbf{X}_t; k)) \mid \theta_t = 1])}{\sum_{j \in \mathcal{C}_{1t}} \rho^{t-1-j}}.$$

Combining the results in (42), (44) and (45) together into (41), we have for any $\varepsilon_1 \in (0, \sqrt{\pi/2})$ and $\varepsilon_2 \in (0, 1 - \rho)$

$$|\hat{M}_{k,t} - M_{k,t}| \leq \varepsilon_1 + \varepsilon_2 + 2\gamma\{(1 - \bar{\pi})(1 - \rho^{t-1})(1 - \rho)\}^{-1} + 2\{(t - 1)\underline{\pi}\}^{-1}$$

with probability

$$1 - 3 \exp\{-\{(t - 1)\underline{\pi}\varepsilon_1^2\}\} - 4 \exp\left\{-\frac{\varepsilon_2^2(1 + \rho)(1 - \bar{\pi})^2(1 - \rho^{t-1})}{2(1 + \rho^{t-1})(1 - \rho)}\right\}.$$

Suppose $t > 1 + 1/(1 - \rho)$ such that $\rho^{t-1} < 1/2$, then

$$\frac{(1 + \rho)(1 - \rho^{t-1})}{2(1 + \rho^{t-1})(1 - \rho)} \geq \frac{1}{3} \frac{1 + \rho}{1 - \rho}.$$

Then the part (i) is verified.

For part (ii),

$$\max_{k \in [K]} |\hat{M}_t^k - M_t^k| \leq \varepsilon_1 + \varepsilon_2 + 4\gamma\{(1 - \bar{\pi})(1 - \rho)\}^{-1} + 2\{(t - 1)\underline{\pi}\}^{-1}$$

holds with probability

$$1 - 3K \exp\{-\{(t - 1)\underline{\pi}\varepsilon_1^2\}\} - 4K \exp\left\{-\frac{\varepsilon_2^2(1 - \bar{\pi})^2}{6(1 - \rho)}\right\}.$$

By Assumption 5,

$$\Pr(\hat{k}_t \neq k_t^* \mid \theta_t = 1) \leq \Pr(\max_{k \in [K]} |\hat{M}_{k,t} - M_{k,t}| \geq c/2).$$

To enforce $\max_k |\hat{M}_{k,t} - M_{k,t}| \leq c/2$, it suffices to let

$$\varepsilon_1 = \varepsilon_2 \leq \frac{c}{8}, \quad \frac{4\gamma}{(1 - \bar{\pi})(1 - \rho)} \leq \frac{c}{8}, \quad \frac{2}{(t - 1)\underline{\pi}} \leq \frac{c}{8}.$$

So it is required that

$$t \geq 1 + \frac{16}{c\underline{\pi}}, \quad \frac{\gamma}{1 - \rho} \leq \frac{(1 - \bar{\pi})c}{32}.$$

And we have the explicit bound

$$\Pr(\hat{k}_t \neq k_t^*) \leq 3K \exp\left\{-(t - 1)\underline{\pi}\frac{c^2}{64}\right\} + 4K \exp\left\{-\frac{c^2(1 - \bar{\pi})^2}{384(1 - \rho)}\right\}.$$

And this bound can be directly generalized when the historical data are not empty by replacing t with $n + t$.

Finally, we can jointly guarantee the score selection consistency for all non-null points:

$$\begin{aligned} \Pr(\exists i \leq t : \theta_i = 1, \hat{k}_i \neq k_i^*) &\leq \sum_{i=1}^t \Pr(\theta_i = 1) \Pr(\hat{k}_i \neq k_i^* \mid \theta_i = 1) \\ &\leq 3K(1 - \underline{\pi}) \sum_{i=1}^t \exp\left\{-(n + i - 1)\underline{\pi}\frac{c^2}{64}\right\} + 4K(1 - \underline{\pi})t \exp\left\{-\frac{c^2(1 - \bar{\pi})^2}{384(1 - \rho)}\right\} \\ &\leq \frac{3K(1 - \underline{\pi}) \exp\{-n\underline{\pi}c^2/64\}}{1 - \exp\{-\underline{\pi}c^2/64\}} + 4K(1 - \underline{\pi})t \exp\left\{-\frac{c^2(1 - \bar{\pi})^2}{384(1 - \rho)}\right\}. \end{aligned}$$

■

F.11 Proof of Theorem 18

Proof It suffices to prove Proposition 22 for the permuted p -value. Naturally, we can view score selection as part of the score construction such that $V(\cdot; \hat{k}_t)$ can be defined as $V(\cdot; \mathcal{D}_t)$. Here we do not require $V(\cdot; \mathcal{D}_t)$ is symmetric to past null data and allow it perform non-symmetrically. For convenience, we set $\theta_t = \theta_{t+1} = \dots = \theta_T = 0$.

For any $\sigma' \in \Omega_t$ and given $\{\mathbf{X}_i : i \in \mathcal{C}_{0t} \cup \{t\}\} = \{\mathbf{x}_i : i \in \mathcal{C}_{0t} \cup \{t\}\}$ as the realizations, define $\bar{\sigma} = \sigma' \cdot \sigma$ and

$$\begin{aligned}
 & p_t^{\text{opt}}((\mathcal{D}_T)_{\sigma'}) \\
 &= \frac{1}{|\Omega_t|} \sum_{\sigma \in \Omega_t} \mathbb{I}\{V(\mathbf{x}_{\sigma'(\sigma(i))}; ((\mathcal{D}_t)_{\sigma})_{\sigma'}) < V(\mathbf{x}_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'})\} \\
 & \quad + \frac{1}{|\Omega_t|} \xi_t \cdot \sum_{\sigma \in \Omega_t} \mathbb{I}\{V(\mathbf{x}_{\sigma'(\sigma(i))}; ((\mathcal{D}_t)_{\sigma})_{\sigma'}) = V(\mathbf{x}_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'})\} \\
 &= \frac{\sum_{\bar{\sigma} \in \Omega_t} \mathbb{I}\{V(\mathbf{x}_{\bar{\sigma}(i)}; (\mathcal{D}_t)_{\bar{\sigma}}) < V(\mathbf{x}_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'})\} + \xi_t \cdot \sum_{\bar{\sigma} \in \Omega_t} \mathbb{I}\{V(\mathbf{x}_{\bar{\sigma}(i)}; (\mathcal{D}_t)_{\bar{\sigma}}) = V(\mathbf{x}_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'})\}}{|\Omega_t|} \\
 &= \frac{\sum_{\sigma \in \Omega_t} \mathbb{I}\{V(\mathbf{x}_{\sigma(i)}; (\mathcal{D}_t)_{\sigma}) < V(\mathbf{x}_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'})\} + \xi_t \cdot \sum_{\sigma \in \Omega_t} \mathbb{I}\{V(\mathbf{x}_{\sigma(i)}; (\mathcal{D}_t)_{\sigma}) = V(\mathbf{x}_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'})\}}{|\Omega_t|}.
 \end{aligned}$$

The last equality comes from the fact that $\sigma' \in \Omega_t$ is a bijection such that $\{\sigma : \sigma \in \Omega_t\} = \{\sigma' \cdot \sigma : \sigma \in \Omega_t\}$.

Follow the same discussion in the proof of Proposition 22. Let $\{X_i : i \in \mathcal{C}_{0t} \cup \{t\}\} = \{x_i : i \in \mathcal{C}_{0t} \cup \{t\}\}$ as a set of realizations and

$$q = Q_{\alpha}(V(x_{\sigma(t)}; (\mathcal{D}_t)_{\sigma}) : \sigma \in \Omega_t).$$

For any $\sigma' \in \Omega_t$ such that $V(\mathbf{x}_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'}) > q$, we have

$$p_t((\mathcal{D}_T)_{\sigma'}) \geq \frac{N_+ + N_-}{|\Omega_t|} \geq \alpha;$$

Here

$$N_- = \sum_{\sigma' \in \Omega_t} \mathbb{I}\{V(x_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'}) < q\}, \quad N_+ = \sum_{\sigma' \in \Omega_t} \mathbb{I}\{V(x_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'}) = q\}.$$

For any $\sigma' \in \Omega_t$ such that $V(\mathbf{x}_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'}) < q$, we have

$$p_t((\mathcal{D}_T)_{\sigma'}) \leq \frac{N_-}{|\Omega_t|} \leq \alpha.$$

Hence

$$\begin{aligned}
 \Pr(p_t \leq \alpha \mid \Phi_t) &= \frac{1}{|\Omega_t|} \sum_{\sigma' \in \Omega_t} \Pr(p_t^{\text{opt}}((\mathcal{D}_T)_{\sigma'}) \leq \alpha) \\
 &= \frac{1}{|\Omega_t|} \sum_{\sigma' \in \Omega_t} \left(\mathbb{I}\{V(\mathbf{x}_{\sigma'(t)}; (\mathcal{D}_t)_{\sigma'}) < q\} + \mathbb{I}\{V(\mathbf{x}_{\sigma'(t)}; \mathcal{D}_t) = q\} \frac{\alpha|\Omega_t| - N_-}{N_-} \right) \\
 &= \frac{1}{|\Omega_t|} \left(N_- + N_- \frac{\alpha|\Omega_t| - N_-}{N_-} \right) = \alpha,
 \end{aligned}$$

As for the independence of null p -values, by a similar discussion in Proposition 22, we can still have

$$p_{t'}(\mathcal{D}_T) = p_{t'}((\mathcal{D}_T)_{\sigma'})$$

for all $t' \in \{t+1, \dots, T\}$ and any $\sigma' \in \Omega_t$. Then by the same iterate conditional expectation without relying on additional symmetry discussion, we can provide the independence results. $\blacksquare$

F.12 Proof of Proposition 2

Proof The inequality $\widehat{\text{FDP}}_{\text{GAIF}}(t) \leq \alpha$ is equivalent to

$$\sum_{m=1}^t \alpha_m \leq \alpha(1 \vee R(t)) + \sum_{j \in \mathcal{I}_t} \theta_j \alpha_j, \quad (46)$$

where $R(t) = \sum_{m=1}^t \delta_m$. Since

$$\alpha_m = \gamma_m s_0 + (\alpha - s_0) \gamma_{m-\tau_1} \mathbb{I}\{\tau_1 < m\} + \alpha \sum_{k \geq 2: \tau_k < m} \gamma_{m-\tau_k} + \sum_{j \in \mathcal{I}_m} \theta_j \alpha_j \gamma_{m-j},$$

summing over $m = 1, \dots, t$ and exchanging the order of summation, we bound each inner sum separately.

Base wealth terms. For any fixed origin $a \in \{0\} \cup \{\tau_k : k \geq 1\}$, define

$$S_{\text{base}}(a) = \sum_{m > a}^t \gamma_{m-a}.$$

Since the index $m - a$ increments by exactly 1 at each step, and $\gamma_i \geq 0$ for all i ,

$$S_{\text{base}}(a) = \sum_{i=1}^{t-a} \gamma_i \leq \sum_{i=1}^{\infty} \gamma_i \leq 1.$$

Since $R(t-1) = \#\{k : \tau_k < t\}$, applying this bound with $\alpha^{(1)} = \alpha - s_0$ and $\alpha^{(k)} = \alpha$ for $k \geq 2$ yields:

$$\sum_{m=1}^t \left[s_0 \gamma_m + \sum_{k \geq 1: \tau_k < m} \alpha^{(k)} \gamma_{m-\tau_k} \right] \leq s_0 + (\alpha - s_0) \mathbb{I}\{R(t-1) \geq 1\} + \alpha (R(t-1) - 1)_+.$$

Since

$$s_0 + (\alpha - s_0) \mathbb{I}\{R(t-1) \geq 1\} + \alpha (R(t-1) - 1)_+ = \begin{cases} s_0, & R(t-1) = 0, \\ \alpha R(t-1), & R(t-1) \geq 1, \end{cases}$$

and $s_0 \leq \alpha$, we obtain

$$s_0 + (\alpha - s_0) \mathbb{I}\{R(t-1) \geq 1\} + \alpha (R(t-1) - 1)_+ \leq \alpha (1 \vee R(t)).$$

Therefore,

$$\sum_{m=1}^t \left[s_0 \gamma_m + \sum_{k \geq 1: \tau_k < m} \alpha^{(k)} \gamma_{m-\tau_k} \right] \leq \alpha (1 \vee R(t)).$$

Feedback wealth terms. By definition, $\mathcal{I}_t = \{j \in [t] : \theta_j \text{ is revealed by time } t\}$; since once revealed a label stays revealed, $\{\mathcal{I}_t\}_{t \in \mathbb{N}}$ is automatically nondecreasing, i.e., $\mathcal{I}_m \subseteq \mathcal{I}_t$ for all $m \leq t$. Consequently, in the double sum $\sum_{m=1}^t \sum_{j \in \mathcal{I}_m} \theta_j \alpha_j \gamma_{m-j}$, every index j that appears for some $m \leq t$ already belongs to $\mathcal{I}_t$. Hence, after exchanging the order of summation, it suffices to sum over $j \in \mathcal{I}_t$. For each $j \in \mathcal{I}_t$, define

$$S_{\text{fb}}(j) = \sum_{m=j+1}^t \mathbb{I}\{j \in \mathcal{I}_m\} \gamma_{m-j}.$$

Since a revealed label stays revealed, if $j \in \mathcal{I}_t$ then $\mathbb{I}\{j \in \mathcal{I}_m\} = 1$ for all $m \geq m_0$ where m_0 is the first time j enters $\mathcal{I}_m$. Hence, using $\gamma_i \geq 0$,

$$S_{\text{fb}}(j) \leq \sum_{m=j+1}^t \gamma_{m-j} = \sum_{i=1}^{t-j} \gamma_i \leq 1.$$

The total feedback contribution therefore satisfies:

$$\sum_{j \in \mathcal{I}_t} \theta_j \alpha_j \cdot S_{\text{fb}}(j) \leq \sum_{j \in \mathcal{I}_t} \theta_j \alpha_j.$$

Combining. Adding the two contributions:

$$\sum_{m=1}^t \alpha_m \leq \alpha (1 \vee R(t)) + \sum_{j \in \mathcal{I}_t} \theta_j \alpha_j,$$

which is exactly (46). Rearranging gives $\widehat{\text{FDP}}_{\text{GAIF}}(t) \leq \alpha$. ■

F.13 Proof of Proposition 6

Proof The inequality $\widehat{\text{FDP}}_{\text{Ada-GAIF}}(t) \leq \alpha$ is equivalent to

$$\sum_{m=1}^t \alpha_m \kappa(p_m) \leq \alpha(1 \vee R(t)) + \sum_{j \in \mathcal{I}_t} \theta_j \alpha_j \kappa(p_j), \quad (47)$$

where $R(t) = \sum_{m=1}^t \delta_m$ and $\kappa(p) = \frac{\mathbb{I}\{p > \lambda\}}{1-\lambda}$.

Denote

$$\begin{aligned} \tilde{\alpha}_m &= (1-\lambda) \left[s_0 \gamma_{m-C_{0+}} + (\alpha - s_0) \gamma_{m-\tau_1-C_{1+}} + \alpha \sum_{j \geq 2} \gamma_{m-\tau_j-C_{j+}} \right] \\ &\quad + \sum_{j: j \in \mathcal{I}_m} \gamma_{m-j-C_{j+}^*} \alpha_j \theta_j \mathbb{I}\{p_j > \lambda\}. \end{aligned}$$

Since $\alpha_m = \min\{\lambda, \tilde{\alpha}_m\} \leq \tilde{\alpha}_m$, and $\kappa(p_m) = \frac{\mathbb{I}\{p_m > \lambda\}}{1-\lambda}$, the factor $(1-\lambda)$ in the denominator of $\kappa(p_m)$ cancels the leading $(1-\lambda)$ in $\tilde{\alpha}_m$. Hence, for each m :

$$\begin{aligned} \alpha_m \kappa(p_m) &\leq \mathbb{I}\{p_m > \lambda\} \left[s_0 \gamma_{m-C_{0+}(m)} + (\alpha - s_0) \gamma_{m-\tau_1-C_{1+}(m)} \mathbb{I}\{\tau_1 < m\} \right. \\ &\quad \left. + \alpha \sum_{k \geq 2: \tau_k < m} \gamma_{m-\tau_k-C_{k+}(m)} \right. \\ &\quad \left. + \sum_{j \in \mathcal{I}_m} \theta_j \alpha_j \kappa(p_j) \gamma_{m-j-C_{j+}^*(m)} \right]. \end{aligned} \quad (48)$$

Summing (48) over $m = 1, \dots, t$ and exchanging the order of summation, we bound each resulting inner sum separately.

Base wealth terms. For any fixed origin $a \in \{0\} \cup \{\tau_k : k \geq 1\}$, define

$$S_{\text{base}}(a) = \sum_{m > a}^t \mathbb{I}\{p_m > \lambda\} \gamma_{m-a-C_{a+}(m)}.$$

Since $C_{a+}(m)$ counts the number of indices $i \in (a, m)$ with $p_i \leq \lambda$, on the event $\{p_m > \lambda\}$ the effective index satisfies

$$m - a - C_{a+}(m) = \sum_{i=a+1}^m \mathbb{I}\{p_i > \lambda\},$$

which increments by exactly 1 relative to its value at the previous occurrence of $\{p_m > \lambda\}$, and stays unchanged whenever $p_m \leq \lambda$ (in which case the term is annihilated by the indicator $\mathbb{I}\{p_m > \lambda\}$ in $S_{\text{base}}(a)$). Consequently, the non-zero terms in $S_{\text{base}}(a)$ map bijectively onto a contiguous prefix of $\{\gamma_i\}_{i \geq 1}$, that is, they correspond exactly to $\gamma_1, \gamma_2, \dots, \gamma_{\Delta_{a,t}}$ without gaps. Therefore, using $\gamma_i \geq 0$,

$$S_{\text{base}}(a) = \sum_{i=1}^{\Delta_{a,t}} \gamma_i \leq \sum_{i=1}^{\infty} \gamma_i \leq 1,$$

where $\Delta_{a,t} = \sum_{i=a+1}^t \mathbb{I}\{p_i > \lambda\}$.

Applying this bound with $\alpha^{(1)} = \alpha - s_0$ and $\alpha^{(k)} = \alpha$ for $k \geq 2$, we obtain

$$\sum_{m=1}^t \mathbb{I}\{p_m > \lambda\} \left[s_0 \gamma_{m-C_{0+}(m)} + \sum_{k \geq 1: \tau_k < m} \alpha^{(k)} \gamma_{m-\tau_k-C_{k+}(m)} \right] \leq s_0 + (\alpha - s_0) \mathbb{I}\{R(t-1) \geq 1\} + \alpha(R(t-1)-1)_+.$$

Since

$$s_0 + (\alpha - s_0) \mathbb{I}\{R(t-1) \geq 1\} + \alpha(R(t-1)-1)_+ = \begin{cases} s_0, & R(t-1) = 0, \\ \alpha R(t-1), & R(t-1) \geq 1, \end{cases}$$

and $s_0 \leq \alpha$, it follows that

$$s_0 + (\alpha - s_0) \mathbb{I}\{R(t-1) \geq 1\} + \alpha(R(t-1)-1)_+ \leq \alpha(1 \vee R(t)).$$

Therefore,

$$\sum_{m=1}^t \mathbb{I}\{p_m > \lambda\} \left[s_0 \gamma_{m-C_{0+}(m)} + \sum_{k \geq 1: \tau_k < m} \alpha^{(k)} \gamma_{m-\tau_k-C_{k+}(m)} \right] \leq \alpha(1 \vee R(t)).$$

Feedback wealth terms. As in the proof of Proposition 2, $\{\mathcal{I}_t\}_{t \in \mathbb{N}}$ is nondecreasing by definition, so terms with $j \notin \mathcal{I}_t$ do not occur in the double sum, and after exchanging the order of summation it suffices to sum over $j \in \mathcal{I}_t$. For each $j \in \mathcal{I}_t$, define

$$S_{\text{fb}}(j) = \sum_{m=j+1}^t \mathbb{I}\{p_m > \lambda\} \mathbb{I}\{j \in \mathcal{I}_m\} \gamma_{m-j-C_{j+}^*(m)}.$$

As m ranges over the non-candidate times after j , on the event $\{p_m > \lambda\}$ this effective index $m - j - C_{j+}^*(m) = \sum_{i=j+1}^m \mathbb{I}\{p_i > \lambda\}$ increases by one each time. Hence the corresponding γ -indices are all distinct. In the full and instant feedback case, these indices form a consecutive prefix of $\{\gamma_i\}_{i \geq 1}$. Under general feedback, the indicator $\mathbb{I}\{j \in \mathcal{I}_m\}$ may remove some terms before the feedback for j becomes available, but it cannot duplicate any γ -index. Consequently, using $\gamma_i \geq 0$,

$$S_{\text{fb}}(j) \leq \sum_{i=1}^{\infty} \gamma_i \leq 1.$$

Hence the total feedback contribution satisfies

$$\sum_{j \in \mathcal{I}_t} \theta_j \alpha_j \kappa(p_j) \cdot S_{\text{fb}}(j) \leq \sum_{j \in \mathcal{I}_t} \theta_j \alpha_j \kappa(p_j).$$

Combining. Adding the two contributions:

$$\sum_{m=1}^t \alpha_m \kappa(p_m) \leq \alpha(1 \vee R(t)) + \sum_{j \in \mathcal{I}_t} \theta_j \alpha_j \kappa(p_j),$$

which is exactly (47). Rearranging gives $\widehat{\text{FDP}}_{\text{Ada-GAIF}}(t) \leq \alpha$. ■

References

- Ehud Aharoni and Saharon Rosset. Generalized α -investing: definitions, optimality results and application to public databases. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(4):771–794, 2014.
- Anastasios N Angelopoulos, Rina Foygel Barber, and Stephen Bates. Theoretical foundations of conformal prediction. *arXiv preprint arXiv:2411.11824*, 2024.
- Tian Bai and Ying Jin. Optimized conformal selection: Powerful selective inference after conformity score optimization. *arXiv preprint arXiv:2411.17983*, 2024.
- Meshi Bashari, Amir Epstein, Yaniv Romano, and Matteo Sesia. Derandomized novelty detection with FDR control via conformal e-values. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.
- Stephen Bates, Emmanuel Candès, Lihua Lei, Yaniv Romano, and Matteo Sesia. Testing for outliers with conformal p-values. *The Annals of Statistics*, 51(1):149–178, 2023.
- Barry Becker and Ronny Kohavi. Adult income investigation. UCI Machine Learning Repository <https://archive-beta.ics.uci.edu/dataset/2/adult>, 1996.
- Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 57(1):289–300, 1995.
- Thomas Brooks, D Pope, and Michael Marcolini. Airfoil Self-Noise. UCI Machine Learning Repository <https://archive.ics.uci.edu/dataset/291/airfoil+self+noise>, 2014.
- Sebastian Döhler and Iqraa Meah. A unified class of null proportion estimators with plug-in fdr control. *arXiv preprint arXiv:2307.13557*, 2023.
- Evanthia Faliagka, Lazaros Iliadis, Ioannis Karydis, Maria Rigou, Spyros Sioutas, Athanasios Tsakalidis, and Giannis Tzimas. On-line consistent ranking on e-recruitment: seeking the truth behind a well-formed cv. *Artificial Intelligence Review*, 42(3):515–528, 2014.
- Lasse Fischer, Ziyu Xu, and Aaditya Ramdas. An online generalizations of the e-BH and BH procedure. *arXiv preprint arXiv:2407.20683*, 2024.
- Aaron Fisher. Online false discovery rate control for LORD++ and SAFFRON under positive, local dependence. *Biometrical Journal*, 66(1):2300177, 2024.
- Dean P Foster and Robert A Stine. α -investing: a procedure for sequential control of expected false discoveries. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(2):429–444, 2008.
- Bowen Gang, Wenguang Sun, and Weinan Wang. Structure-adaptive sequential testing for online false discovery rate control. *Journal of the American Statistical Association*, 118(541):732–745, 2023.

- Matteo Gasparin and Aaditya Ramdas. Conformal online model aggregation. *arXiv preprint arXiv:2403.15527*, 2024.
- Isaac Gibbs and Emmanuel Candès. Adaptive conformal inference under distribution shift. *Advances in Neural Information Processing Systems*, 34:1660–1672, 2021.
- Isaac Gibbs and Emmanuel J Candès. Conformal inference for online prediction with arbitrary distribution shifts. *Journal of Machine Learning Research*, 25(162):1–36, 2024.
- Yu Gui, Ying Jin, and Zhimei Ren. Conformal alignment: Knowing when to trust foundation models with guarantees. *Advances in Neural Information Processing Systems*, 37:73884–73919, 2024.
- Yu Gui, Ying Jin, Yash Nair, and Zhimei Ren. ACS: An interactive framework for conformal selection. *arXiv preprint arXiv:2507.15825*, 2025.
- Lei Huang, Weijiang Yu, Weitao Ma, Weihong Zhong, Zhangyin Feng, Haotian Wang, Qianglong Chen, Weihua Peng, Xiaocheng Feng, Bing Qin, et al. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, 43(2):1–55, 2025.
- Pierre Humbert, Ulysse Gazin, Ruth Heller, and Etienne Roquain. Online selective conformal inference: Adaptive scores, convergence rate and optimality. *arXiv preprint arXiv:2508.10336*, 2025.
- Yuyang Huo, Lin Lu, Haojie Ren, and Changliang Zou. Real-time selection under general constraints via predictive inference. *Advances in Neural Information Processing Systems*, 37:61267–61305, 2024.
- Adel Javanmard and Andrea Montanari. On online control of false discovery rate. *arXiv preprint arXiv:1502.06197*, 2015.
- Adel Javanmard and Andrea Montanari. Online rules for control of false discovery rate and false discovery exceedance. *The Annals of Statistics*, 46(2):526–554, 2018.
- Ying Jin and Emmanuel J Candès. Selection by prediction with conformal p-values. *Journal of Machine Learning Research*, 24(244):1–41, 2023.
- Ying Jin and Emmanuel J Candès. Model-free selective inference under covariate shift via weighted conformal p-values. *Biometrika*, page asaf066, 2025.
- Kaggle. Candidate selection dataset. <https://www.kaggle.com/datasets/tarunchilkur/client>, 2020.
- Kaggle. Diabetes health indicators dataset. <https://www.kaggle.com/datasets/alexteboul/diabetes-health-indicators-dataset>, 2021.
- Etienne Krönert, Alain Célisse, and Dalila Hattab. Fdr control for online anomaly detection. *arXiv preprint arXiv:2312.01969*, 2023.

- Kangdao Liu, Huajun Xi, Chi Man Vong, and Hongxin Wei. Online conformal selection with accept-to-reject changes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pages 23765–23773, 2026.
- Ariane Marandon, Lihua Lei, David Mary, and Etienne Roquain. Adaptive novelty detection with false discovery rate guarantee. *The Annals of Statistics*, 52(1):157 – 183, 2024.
- Yash Nair, Ying Jin, James Yang, and Emmanuel Candes. Diversifying conformal selections. *arXiv preprint arXiv:2506.16229*, 2025.
- Gergely Neu, Andras Antos, András György, and Csaba Szepesvári. Online markov decision processes under bandit feedback. *Advances in Neural Information Processing Systems*, 23, 2010.
- Drew Prinster, Xing Han, Anqi Liu, and Suchi Saria. WATCH: Adaptive monitoring for AI deployments via weighted-conformal martingales. In *Forty-second International Conference on Machine Learning*, 2025.
- Aaditya Ramdas, Fanny Yang, Martin J Wainwright, and Michael I Jordan. Online control of the false discovery rate with decaying memory. *Advances in Neural Information Processing Systems*, 30:5655–5664, 2017.
- Aaditya Ramdas, Tijana Zrnic, Martin Wainwright, and Michael Jordan. SAFFRON: an adaptive algorithm for online control of the false discovery rate. In *International Conference on Machine Learning*, pages 4286–4294. PMLR, 2018.
- Quentin Rebjock, Baris Kurt, Tim Januschowski, and Laurent Callot. Online false discovery rate control for anomaly detection in time series. *Advances in Neural Information Processing Systems*, 34:26487–26498, 2021.
- David S Robertson, James MS Wason, and Aaditya Ramdas. Online multiple hypothesis testing. *Statistical Science*, 38(4):557, 2023.
- John D Storey, Jonathan E Taylor, and David 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.
- Jinjin Tian and Aaditya Ramdas. ADDIS: an adaptive discarding algorithm for online FDR control with conservative nulls. *Advances in Neural Information Processing Systems*, 32: 9388–9396, 2019.
- Vladimir Vovk. Testing randomness online. *Statistical Science*, 36(4):595–611, 2021.
- Vladimir Vovk and Ruodu Wang. E-values: Calibration, combination and applications. *The Annals of Statistics*, 49(3):1736–1754, 2021.
- Vladimir Vovk, Ilia Nouretdinov, and Alexander Gammerman. Testing exchangeability on-line. In *Proceedings of the 20th International Conference on Machine Learning*, pages 768–775, 2003.

- Vladimir Vovk, Alexander Gammernan, and Glenn Shafer. *Algorithmic learning in a random world*. New York: Springer, 2005.
- Xiaoning Wang, Yuyang Huo, Liuhua Peng, and Changliang Zou. Conformalized multiple testing after data-dependent selection. *Advances in Neural Information Processing Systems*, 37:58574–58609, 2024.
- Xiaoyang Wu, Yuyang Huo, Haojie Ren, and Changliang Zou. Optimal subsampling via predictive inference. *Journal of the American Statistical Association*, 119(548):2844–2856, 2024.
- Xiaoyang Wu, Lin Lu, Zhaojun Wang, and Changliang Zou. Conditional testing based on localized conformal p-values. In *The Thirteenth International Conference on Learning Representations*, 2025.
- HuaJun Xi, Kangdao Liu, Hao Zeng, Wenguang Sun, and Hongxin Wei. Exploring the noise robustness of online conformal prediction. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.
- Ziyu Xu and Aaditya Ramdas. Online multiple testing with e-values. In *International Conference on Artificial Intelligence and Statistics*, pages 3997–4005. PMLR, 2024.
- Yifan Zhang, Zijian Wei, Haojie Ren, and Changliang Zou. e-GAI: e-value-based generalized alpha-investing for online false discovery rate control. In *Forty-second International Conference on Machine Learning*, pages 76308–76331, 2025.
- Tijana Zrnic, Aaditya Ramdas, and Michael I. Jordan. Asynchronous online testing of multiple hypotheses. *Journal of Machine Learning Research*, 22(33):1–39, 2021.