



HAL
open science

Asymptotics for conformal inference

Ulysse Gazin

► **To cite this version:**

| Ulysse Gazin. Asymptotics for conformal inference. 2026. <hal-04701683v2>

HAL Id: hal-04701683

<https://hal.science/hal-04701683v2>

Preprint submitted on 18 Feb 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Asymptotics for conformal inference

Ulysse Gazin¹

¹*Université Paris Cité and Sorbonne Université, CNRS, Laboratoire de Probabilités, Statistique et Modélisation, F-75013 Paris, France, e-mail: ugazin@lpsm.paris*

Abstract: Conformal inference is a versatile tool for building prediction sets in regression or classification. We study the false coverage proportion (FCP) in a simultaneous inference setting with a calibration sample of n points and a test sample of m points. We identify the exact, distribution-free, asymptotic distribution of the FCP when both n and m tend to infinity. This shows in particular that FCP control can be achieved by using the well-known Kolmogorov distribution, and puts forward that the asymptotic variance is decreasing in the ratio n/m . We then provide a number of extensions by considering the problems of novelty detection, weighted conformal inference or distribution shift between the calibration sample and the test sample. In particular, our asymptotic results allow to accurately quantify the asymptotic behavior of the errors (a miscovering interval or declaring a false novelty) when weighted conformal inference is used.

Keywords and phrases: Conformal prediction, empirical cumulative distribution function, functional central limit theorem, functional delta method, multiple testing, weighted conformal prediction.

1. Introduction

1.1. Background

In classical statistics, producing prediction sets for outcomes often relies on strong model assumptions. Recent advances involve complex data sets and sophisticated machine learning methods, for which such an approach is not appropriate. One recent solution is given by conformal prediction (Saunders, Gammerman and Vovk, 1999; Vovk, Gammerman and Shafer, 2005; Angelopoulos and Bates, 2021) which consists in calibrating the prediction set according to an appropriate quantile of a calibration/training sample. Strikingly, this method provides a finite-sample valid coverage (i.e. valid for any size $n \geq 1$ of the calibration sample) for any underlying distribution of the data and for any underlying point-prediction machine learning algorithm. Similar techniques can be employed for the novelty detection task (Balasubramanian, Ho and Vovk, 2014; Bates et al., 2023; Marandon et al., 2024).

1.2. Aim and contributions

We consider here the multiple setting (Vovk, 2013), where it is given a training dataset (which is considered as fixed here), a calibration sample of n points $(X_1, Y_1), \dots, (X_n, Y_n)$ and a test sample of m points $(X_{n+1}, Y_{n+1}), \dots, (X_{n+m}, Y_{n+m})$. While the calibration sample is fully observed, the Y_i 's of the test points are not observed and a prediction set should be provided for each of them. The false coverage proportion (FCP) for the m conformal prediction sets $\mathcal{C}^\alpha(X_{n+1}), \dots, \mathcal{C}^\alpha(X_{n+m})$ (see below for a formal definition) is defined as the proportion of coverage errors among the test sample:

$$\text{FCP}_m^{(n)}(\alpha) := \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{Y_{n+i} \notin \mathcal{C}^\alpha(X_{n+i})}.$$

Under standard assumptions and when the conformal prediction sets come from the inductive/split conformal procedure the distribution of the process $\text{FCP}_m^{(n)}$ has been shown to be distribution-free, in the sense that it does only depend on n and m (Marques F., 2025; Huang et al., 2024; Gazin, Blanchard and Roquain, 2024). Due to the dependence between the individual coverage

errors, combinatorial formulas derived in those references are particularly complex in m and n . Nevertheless, focusing on the maximum absolute deviation

$$\|\text{FCP}_m^{(n)} - I_n\|_\infty := \sup_{\alpha \in [0,1]} |\text{FCP}_m^{(n)}(\alpha) - I_n(\alpha)|, \quad (1)$$

with $I_n(\alpha) = \lfloor (n+1)\alpha \rfloor / (n+1)$, a DKW-type concentration inequality has been derived in [Gazin, Blanchard and Roquain \(2024\)](#), which is both simple and finite-sample valid. It explicitly involves a rate $\tau_{n,m}^{1/2}$ with

$$\tau_{n,m} := \frac{nm}{n+m} \in \left[\frac{n \wedge m}{2}, n \wedge m \right]. \quad (2)$$

However, this DKW inequality is conservative in general (see [Figure 1](#) below), which makes the corresponding FCP control stringent.

The aim of this paper is to complement the above studies by analyzing $\text{FCP}_m^{(n)}$ from an asymptotic point of view, where *both* m and n tend to infinity which is equivalent to the convergence of $\tau_{n,m}$ to infinity. This asymptotic study will also cover the case of a distribution shift i.e. when the exchangeability assumption does not hold thus extending the analysis of [Gazin, Blanchard and Roquain \(2024\)](#) and more broadly the usual framework of conformal inference. We also study, in the same asymptotic regime, the asymptotic of the Benjamini-Hochberg procedure when conformal p -values are used ([Bates et al., 2023](#)), which complement the non-asymptotic work of [Bates et al. \(2023\)](#); [Gazin, Blanchard and Roquain \(2024\)](#) among others. We refer to [Section 1.3](#) for more details on the relation of this work with previous studies. More precisely, our contributions are the following ones:

1. We show that under standard assumptions $\text{FCP}_m^{(n)}$ converges uniformly to the nominal value at rate $\tau_{n,m}^{1/2}$ and that the asymptotic covariance process is a standard Brownian bridge ([Theorem 3.1](#)). Compared to the ‘‘oracle’’ case where $n = \infty$, it means that the variance is inflated by a factor asymptotically equivalent to $(n+m)/n$, for instance 2 in the case where $n \sim m$.
2. A direct corollary of this result is that $\tau_{n,m}^{1/2} \|\text{FCP}_m^{(n)} - I_n\|_\infty$ converges to the well known Kolmogorov distribution, that is, the distribution with c.d.f. $x \mapsto (1 - 2 \sum_{k \geq 1} (-1)^k e^{-2k^2 x^2}) \mathbb{1}_{x \geq 0}$. A comparison between the quantiles of this distribution, those given by empirical simulations of $\text{FCP}_m^{(n)}$ or those from the Conformal-DKW is provided in [Figure 1](#). While its validity is only asymptotic with $n, m \rightarrow +\infty$, it appears that the Kolmogorov quantile is simple and more accurate than the quantile obtained from DKW. Hence, our limit theorem allows to get simple, accurate and asymptotically-valid confidence bounds for the FCP process as well as asymptotic approximations of all quantities related to the distribution of the FCP process.
3. We then extend this result to the case where the distribution of the calibration sample is not equal to the test sample, that is, under a *distribution shift* ([Theorem 3.2](#)). As expected, the convergence of the FCP is not towards the nominal level in this case but rather towards a new term G that takes this shift into account. The asymptotic covariance process is also modified according to G and an explicit formula is given. The role of the ratio $n/(n+m)$ (number of calibration points divided by the total number of points) clearly appears in the limit and indicates the existence of three asymptotic regimes: under, proportional and over calibrated, see [Table 1](#).
4. To recover the appropriate nominal level in the limit, we adopt the weighted conformal approach ([Tibshirani et al., 2019](#); [Barber et al., 2023](#)) with specific weights that rely on the data distribution, that we refer to as *oracle weights*. The obtained central limit theorem shares similarities with the exchangeable case described above, with the essential difference that the asymptotic covariance process is not distribution free, and depends on the sample distributions ([Theorem 3.4](#)).
5. We also obtain a convergence result in case of non-oracle weights ([Theorem 3.3](#)), which is crucial to quantify the FCP asymptotic behavior in the difficult but realistic case where the user has not access to the true distribution shift. An illustration is displayed in [Figure 2](#), where the asymptotic confidence interval for the FCP is given as a function of an error Δ ,

measuring how the used weights deviate from the oracle ones. This puts forward that the FCP gets significantly away from α when Δ is above ≈ 0.133 or below ≈ -0.11 in this framework.

- Finally, we obtain similar results for the novelty detection task, by studying the asymptotic behavior of the false discovery proportion (FDP) of classical procedures (Bates et al., 2023; Jin and Candès, 2023) (Theorems 4.1 and 4.2).

The proofs are based on specific decompositions of the processes that can be found in Section 5, while further details are postponed to appendices, notably Section G with a presentation of the main results of the paper under relaxed assumptions.

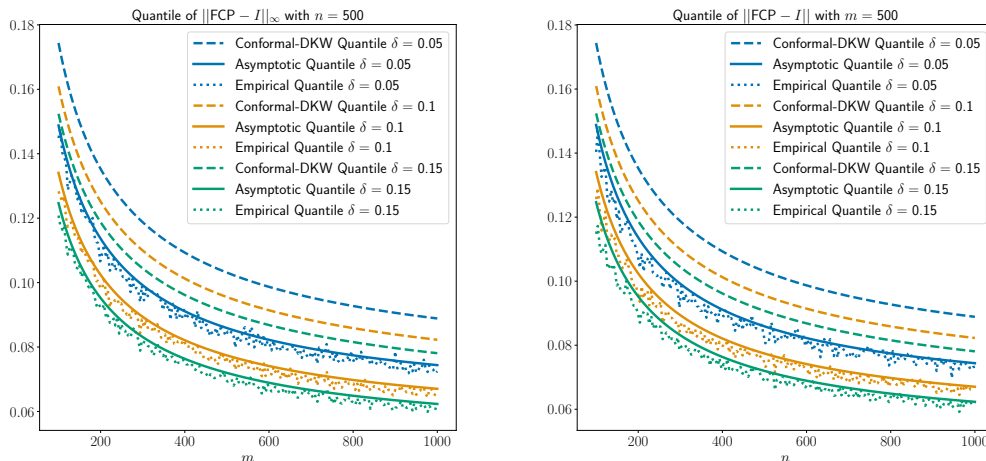


Fig 1: Comparison of the $(1 - \delta)$ -quantile of different approximations of the distribution of $\|\text{FCP}_m^{(n)} - I_n\|_\infty$ in (1) for different values of n, m and δ . The approximations include Monte-Carlo (1000 replications), Conformal-DKW (Gazin, Blanchard and Roquain, 2024) and the asymptotic exact Kolmogorov one.

1.3. Relation to previous work

Conformal prediction is a general pipeline and we refer the reader to Vovk, Gammernan and Shafer (2005) or Angelopoulos and Bates (2021) for reviews. We focus here on the inductive/split conformal inference (Papadopoulos et al., 2002), where an independent training sample is used to build the predictors, while a calibration sample is used to adjust the prediction sets. All our results can be thought of holding conditionally on the training sample, that is, they hold once the point-predictions have been computed.

The process $\text{FCP}_m^{(n)}$ coincides with the empirical cumulative distribution function (e.c.d.f.) of conformal p -values as introduced by Saunders, Gammernan and Vovk (1999). As recalled above, for a given nominal level α , the non-asymptotic distribution of $\text{FCP}_m^{(n)}(\alpha)$ has been obtained in Marques F. (2025); Huang et al. (2024) and the full distribution of the process $\text{FCP}_m^{(n)}$ has been given in Gazin, Blanchard and Roquain (2024). The main difference with the current study is that they provide non-asymptotic properties under the exchangeable setting while we provide asymptotic properties under the exchangeable and the non-exchangeable frameworks. In addition, Marques F. (2025) gives an asymptotic result for $\text{FCP}_m^{(n)}(\alpha)$ with $\alpha \in (0, 1)$ fixed in the regime where n tends to infinity only after making m tend to infinity, whereas here we make n and m grow to infinity independently leading to a double asymptotic. Furthermore all our results are uniform in α .

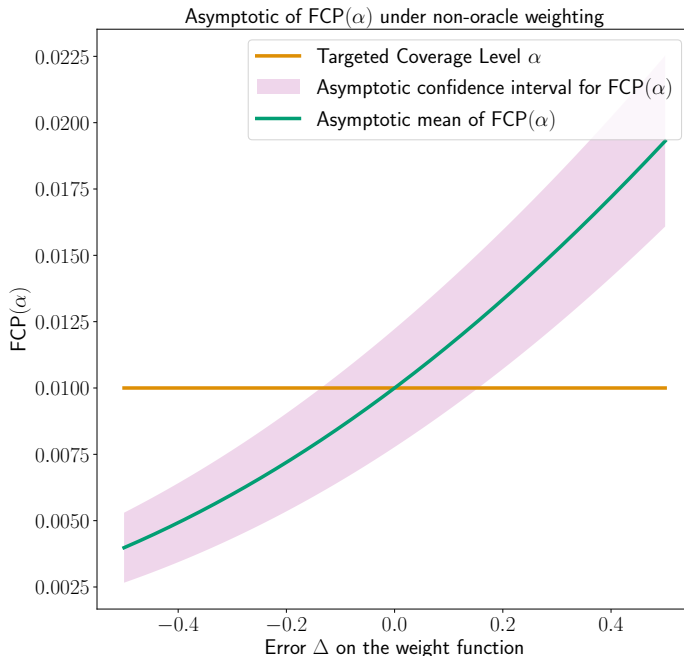


Fig 2: Plot of the asymptotic confidence interval for the FCP of the weighted conformal method (at level 80%) obtained in Theorem 3.3 versus a parameter estimation error Δ . The calibration sample and the test sample are distributed according to the exponential distribution with mean 1 and $1/3$, respectively. The weight function used in the conformal method is $w_\Delta(x) = \exp(-(2 + \Delta)x)\mathbb{1}_{x>0}$, which corresponds to the oracle choice if $\Delta = 0$ (for which the asymptotic average of $\text{FCP}(\alpha)$ is equal to α) and deviates from it if $\Delta \neq 0$.

Next, the asymptotic distribution of $\text{FCP}_m^{(n)}$ with n fixed and $m \rightarrow +\infty$ is also studied with a PAC point of view since this limit is equal to the miscoverage probability given the calibration sample $\mathbb{P}[Y_{n+1} \notin \mathcal{C}^\alpha(X_n + 1) | (X_i, Y_i)_{i \in [n]}]$: Vovk (2012) derives Hoeffding-like tail bounds for a fixed α , and other bounds was derived by Bian and Barber (2022); Sarkar and Kuchibhotla (2023) the latter being uniform in α . Angelopoulos and Bates (2021) states that $\mathbb{P}[Y_{n+1} \notin \mathcal{C}^\alpha(X_n + 1) | (X_i, Y_i)_{i \in [n]}]$ follows a Beta distribution, which is exactly the limit identified by Marques F. (2025). Similarly, Nguyen et al. (2024) obtained asymptotic results for the error as m tends to infinity while n is kept fixed but with more general risk function and not necessary the 0 – 1 loss function. In the novelty detection setting, Bates et al. (2023) also provided bounds for the type I error given the calibration sample uniformly in α .

Finally, we study here novelty detection procedures obtained by applying the Benjamini-Hochberg procedure (BH, Benjamini and Hochberg, 1995) to the conformal p -values (Mary and Roquain, 2022; Bates et al., 2023; Marandon et al., 2024) or to the weighted conformal p -values (Tibshirani et al., 2019; Barber et al., 2023; Jin and Candès, 2023). Asymptotics for the FDP of such procedures have been extensively studied in the literature, see Genovese and Wasserman (2002, 2004); Neuvial (2008) for i.i.d. uniform p -values and Delattre and Roquain (2011, 2016); Kluger and Owen (2024) for several types of dependence structures. The present work is in this line of research by seeing (weighted) conformal p -values as a particular case of dependent p -values having a specific dependency structure induced by the common calibration sample. In particular, our work shows that the ratio $n/(n + m)$ measures the dependency strength in the double asymptotic regime where both n (calibration sample size) and m (the number of test) tend to infinity. In

the novelty detection setting, we derive our results by using the techniques introduced in [Neu-vial \(2008\)](#), that combine functional central limit theorems with the functional delta method. In particular, we provide the full asymptotic FDP distribution for the procedure of [Jin and Candès \(2023\)](#) (or a variant thereof), for which only in-expectation results were known.

2. Preliminaries

2.1. Prediction setting

We consider the classical inductive/split conformal prediction ([Papadopoulos et al., 2002](#)). A calibration set $\{Z_k, k \in \llbracket n \rrbracket\}$ is observed with $Z_k = (X_k, Y_k)$ and, given a new point $Z_{n+1} = (X_{n+1}, Y_{n+1})$ for which only the covariate $X_{n+1} \in \mathcal{X}$ is observed, a prediction set should be inferred for the outcome $Y_{n+1} \in \mathcal{Y}$.¹ A training dataset $\mathcal{D}_{\text{train}}$ independent of the test point and the calibration sample is observed and used to compute a non-conformity score function $S : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$. The scalar $S(X, Y) \in \mathbb{R}$ measures the non-conformity of the response $Y \in \mathcal{Y}$ with the covariate $X \in \mathcal{X}$. All our results hold conditionally on the training dataset, hence to alleviate notation we will consider the score function S as fixed and determinist. This is in accordance with the usual split conformal model: the training/calibration split is considered as already done, with a fixed independent training sample and a growing calibration sample. A classical example in regression is the absolute value of the residual $S(X, Y) = |Y - \hat{\mu}(X)|$ where $\hat{\mu} : \mathcal{X} \rightarrow \mathcal{Y}$ is a regression function trained from an independent training sample (considered as fixed here). The (split) conformal prediction set at level $(1 - \alpha)$ for X_{n+1} , denoted by $\mathcal{C}^\alpha(X_{n+1})$, is defined as

$$\mathcal{C}^\alpha(X_{n+1}) := \{y \in \mathcal{Y} : S(X_{n+1}, y) \leq S_{([\!(n+1)(1-\alpha)\!]})\}, \quad (3)$$

where $S_{(1)} \leq S_{(2)} \leq \dots \leq S_{(n)} < S_{(n+1)} := +\infty$ correspond to the ordered calibration scores $\mathcal{D}_{\text{cal}} := \{S_k : k \in \llbracket n \rrbracket\} = \{S(X_k, Y_k) : k \in \llbracket n \rrbracket\}$. The interest of the conformal prediction set is the validity of the prediction region under weak assumption. More precisely, if the random variables $(Z_1, \dots, Z_n, Z_{n+1})$ (i.e. the n calibration points and the new test point) are exchangeable (or i.i.d.) then we have the marginal coverage property: $\mathbb{P}(Y_{n+1} \notin \mathcal{C}^\alpha(X_{n+1})) \leq \alpha$. Furthermore, if the scores $(S(X_i, Y_i))_{i \in \llbracket n+1 \rrbracket}$ are exchangeable and without ties almost surely, then $\mathbb{P}(Y_{n+1} \notin \mathcal{C}^\alpha(X_{n+1}))$ is lower bounded by $\alpha - (n+1)^{-1}$ (see [Vovk, Gammerman and Shafer, 2005](#) for more details). The set $\mathcal{C}^\alpha(X_{n+1})$ can be equivalently described by using p -values (this classical fact can be retrieved from [Lemma F.1](#)). In particular, by defining the (unobserved) conformal p -value $p_1^{(n)}$ as in (4) with the test score $T_1 = S(X_{n+1}, Y_{n+1})$ (see below), we obtain that $Y_{n+1} \notin \mathcal{C}^\alpha(X_{n+1})$ if and only if $p_1^{(n)} \leq \alpha$. By the above marginal coverage property, $p_1^{(n)}$ is a valid p -value for testing exchangeability of $(Z_1, \dots, Z_n, Z_{n+1})$ since it is super-uniformly distributed under exchangeability. In this paper, we consider a multiple/simultaneous inference setting (also called transductive by [Vovk, 2013](#); [Gazin, Blanchard and Roquain, 2024](#); [Behboodi et al., 2025](#)), where a decision should be made for a whole test sample of size m $(Z_{n+1}, \dots, Z_{n+m})$ with $Z_{n+i} = (X_{n+i}, Y_{n+i})$, for which only the covariates X_{n+1}, \dots, X_{n+m} are observed. We denote $\mathcal{D}_{\text{test}} := \{T_i : i \in \llbracket m \rrbracket\}$ with $T_i = S(X_{n+i}, Y_{n+i})$ the set of the unobserved (since only the X_{n+i} 's are observed) scores of the test sample. Considering the conformal prediction sets $\mathcal{C}^\alpha(X_{n+i})$, $i \in \llbracket m \rrbracket$ — corresponding to (3) for any test point — gives rise to a family of m conformal p -values $(p_i^{(n)})_{i \in \llbracket m \rrbracket}$ defined by

$$p_i^{(n)} := (n+1)^{-1} \left(1 + \sum_{k \in \llbracket n \rrbracket} \mathbb{1}_{S_k \geq T_i} \right), \quad i \in \llbracket m \rrbracket. \quad (4)$$

We also define the family of theoretical p -values $(p_i^{(+\infty)})_{i \in \llbracket m \rrbracket}$ when the calibration scores are i.i.d. as

$$p_i^{(+\infty)} := \mathbb{P}[S_1 \geq T_i | T_i], \quad i \in \llbracket m \rrbracket. \quad (5)$$

¹The regression and classification settings correspond to $\mathcal{Y} = \mathbb{R}$ and \mathcal{Y} finite, respectively.

By applying the law of large number when the calibration scores are i.i.d., the p -value $p_i^{(+\infty)}$ can be seen as the limit of $p_i^{(n)}$ when n tends to infinity. The false coverage proportion $\text{FCP}_m^{(n)}(\alpha)$ of the prediction set family $(\mathcal{C}^\alpha(X_{n+i}))_{i \in \llbracket m \rrbracket}$ is defined by

$$\text{FCP}_m^{(n)}(\alpha) := \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{Y_{n+i} \notin \mathcal{C}^\alpha(X_{n+i})} = \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{p_i^{(n)} \leq \alpha}. \quad (6)$$

Note that this corresponds to the empirical cumulative distribution function (e.c.d.f.) of the conformal p -values family $(p_i^{(n)})_{i \in \llbracket m \rrbracket}$. We emphasise that the set $\mathcal{D}_{\text{test}}$ of test scores is not observed, and neither are the conformal p -values (4) in the prediction setting. The FCP is not observed, and the aim is to study its behavior when the sizes of the calibration and test samples both tend to infinity. The following assumption will be considered throughout the paper for the (weighted or not) conformal prediction task.

Assumption 1. *The set of calibration scores $(S_k)_{k \geq 1}$ and the set of test scores $(T_i)_{i \geq 1}$ are two independent families of real random variables. The S_k 's ($k \geq 1$), (resp. T_i 's ($i \geq 1$)) are i.i.d. with distribution P_{cal} (resp. P_{test}) and cumulative distribution function F_{cal} (resp. F_{test}). Moreover, F_{cal} and F_{test} are continuous functions.*

When the variables $Z_i, i \in \llbracket n+m \rrbracket$, are i.i.d., Assumption 1 is true with $P_{\text{cal}} = P_{\text{test}}$ conditionally on $\mathcal{D}_{\text{train}}$, see Remark 2.1 below for more details. Since the vector of scores $\mathcal{D}_{\text{cal}} \cup \mathcal{D}_{\text{test}}$ is exchangeable in this case, the conformal p -values are marginally super-uniform (Vovk, Gammernan and Shafer, 2005; Romano and Wolf, 2005) which leads to non-asymptotically valid prediction sets. However, in case of a distribution shift between the distributions of the calibration and the test sample, that is $P_{\text{cal}} \neq P_{\text{test}}$, this property is lost. To solve this issue, Tibshirani et al. (2019); Barber et al. (2023) proposed in this case to use weighted conformal prediction. We follow this approach (with light formal variations for mathematical convenience) by introducing a nonnegative *weight function* $w : \mathbb{R} \rightarrow \mathbb{R}^+$, $w(+\infty) \in \mathbb{R}^+$ and the prediction set

$$\mathcal{C}^{w,\alpha}(X_{n+1}) := \left\{ y \in \mathcal{Y} : S(X_{n+1}, y) \leq Q_{1-\alpha} \left(\frac{w(+\infty)}{W} \delta_{+\infty} + \sum_{k \in \llbracket n \rrbracket} \frac{w(S_k)}{W} \delta_{S_k} \right) \right\},$$

where $Q_{1-\alpha}(\mu)$ denotes the $1-\alpha$ -quantile of the probability measure μ , δ_a denotes the Dirac measure in a and $W := w(+\infty) + \sum_{k \in \llbracket n \rrbracket} w(S_k)$ is a normalization constant. There are different ways to choose the weight function w . In the case of a known distribution shift in some specific model, the weight function is usually chosen as the Radon-Nikodym derivative of the shift (Tibshirani et al., 2019; Jin and Candès, 2023). Barber et al. (2023) assumes that the weight function w is deterministic and fixed by the user before looking at the data. To accommodate both points of view, we assume that w can be estimated from a (possibly infinite) independent training dataset (potentially different from $\mathcal{D}_{\text{train}}$ above) containing both data with the same distribution as the calibration points, and data with the same distribution as the test points. Our results hold conditionally on this training sample and w can be considered as deterministic. Thanks to Lemma F.1, the set $\mathcal{C}^{w,\alpha}(X_{n+1})$ can be described with *weighted conformal p -values*: $Y_{n+1} \notin \mathcal{C}^{w,\alpha}(X_{n+1})$ if and only if $p_1^{w,(n)} \leq \alpha$, with $p_1^{w,(n)}$ the (unobserved) *weighted conformal p -value* given by (7) below. Similarly, with a test sample of size m , we obtained m *weighted conformal p -values*,

$$p_i^{w,(n)} := \frac{w(+\infty) + \sum_{k=1}^n w(S_k) \mathbb{1}_{S_k \geq T_i}}{w(+\infty) + \sum_{k=1}^n w(S_k)}, \text{ for all } i \in \llbracket m \rrbracket. \quad (7)$$

The FCP of these weighted conformal prediction sets is then defined by

$$\text{FCP}_m^{w,(n)}(\alpha) := \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{Y_{n+i} \notin \mathcal{C}^{w,\alpha}(X_{n+i})} = \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{p_i^{w,(n)} \leq \alpha}, \quad (8)$$

which, similarly to (6), is the e.c.d.f. of the weighted conformal p -values family. As for the unweighted case, the weighted conformal p -values are not observed, and neither is the corresponding false coverage proportion.

Finally, under Assumption 1 and if P_{test} is absolutely continuous with respect to P_{cal} , a particularly interesting weight function, called the *oracle weight function*, is given by

$$w^* := \frac{d P_{\text{test}}}{d P_{\text{cal}}}. \quad (9)$$

Clearly, if the calibration and test sample are exchangeable ($P_{\text{cal}} = P_{\text{test}}$), the function w^* is constantly equal to 1. However, under a distribution shift ($P_{\text{cal}} \neq P_{\text{test}}$), the oracle weight function is different, and oracle-weighted conformal p -values $p_i^{w^*,(n)}$ (7) are different from the regular ones $p_i^{(n)}$ (4). They satisfy the marginally super-uniform property provided that $w^*(+\infty) \geq \sup_{u \in \mathbb{R}} w^*(u)$, see Tibshirani et al. (2019).

Remark 2.1. *In split conformal inference the score function $S(\cdot)$ and the weight function $w(\cdot)$ are constructed here with an independent training dataset which creates a dependence among the scores $(S(Z_i))_{i \in \llbracket n+m \rrbracket}$. However, our results are stated conditionally on the training phase hence the weight function $w(\cdot)$ and score function $S(\cdot)$ can be considered as deterministic and the scores $(S(Z_i))_{i \in \llbracket n+m \rrbracket}$ are independent as stated in Assumption 1 provided that the data $(Z_i)_{i \in \llbracket n+m \rrbracket}$ are independent (which is a common assumption in the conformal prediction literature).*

2.2. Novelty detection setting

In the novelty detection setting (see for instance Vovk, Gammernan and Shafer, 2005; Bates et al., 2023), we observe a calibration set $\{Z_k, k \in \llbracket n \rrbracket\}$ of size n , with values in \mathcal{Z} distributed according to an unknown “null” distribution \mathbb{P}_0 and a test sample $\{Z_{n+i}, i \in \llbracket m \rrbracket\}$ with $Z_{n+i} \in \mathcal{Z}$ either distributed as \mathbb{P}_0 or not. Formally, we introduce a subset $\mathcal{H}_0 \subset \{1, 2, \dots\}$ so that $Z_{n+i} \sim \mathbb{P}_0$ when $i \in \mathcal{H}_0$. We also denote $\mathcal{H}_1 = \{1, 2, \dots\} \setminus \mathcal{H}_0$. We assume that we have access to an independent training dataset $\mathcal{D}_{\text{train}}$ used to learn a non-conformity score function $S : \mathcal{Z} \rightarrow \mathbb{R}$ such that $S(Z) \in \mathbb{R}$ measures the non-conformity of the variables Z with respect to \mathbb{P}_0 . Our results are stated conditionally on $\mathcal{D}_{\text{train}}$ and the training phase. Therefore, as in the prediction setting, we assume that the score function $S(\cdot)$ is fixed. We denote $\mathcal{D}_{\text{cal}} = \{S_k : k \in \llbracket n \rrbracket\}$ for $S_k := S(Z_k)$, and $\mathcal{D}_{\text{test}} = \{T_i : i \in \llbracket m \rrbracket\}$ for $T_i := S(Z_{n+i})$ the set of the scores from the calibration and test samples respectively. A novelty detection procedure decides, for each $i \in \llbracket m \rrbracket$, whether Z_{n+i} is a novelty (that is, does not follow \mathbb{P}_0), or not. The procedure described by Bates et al. (2023) consists in computing the conformal p -values defined in (4) (by using the specific S_k ’s and T_i ’s of novelty detection which are observed in this case), and then applying the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995) on this p -value family $(p_i^{(n)})_{i \in \llbracket m \rrbracket}$ (see Section 5.2.1 for more formal details). This gives a rejection set $\mathcal{R}_\alpha \subset \llbracket m \rrbracket$ corresponding to the indices of the declared novelties. We introduce the two main quantities to measure the performance of the rejection set,

$$\text{FDP}_m^{(n)}(\mathcal{R}_\alpha) := \frac{|\mathcal{R}_\alpha \cap \mathcal{H}_0|}{|\mathcal{R}_\alpha| \vee 1}; \quad (10)$$

$$\text{TDP}_m^{(n)}(\mathcal{R}_\alpha) := \frac{|\mathcal{R}_\alpha \cap \mathcal{H}_1|}{|\mathcal{H}_1| \vee 1}. \quad (11)$$

The FDP corresponds to the proportion of errors among the declared novelties (related to a type I error notion), while the TDP corresponds to the proportion of correct decisions among the true novelties (related to power and type II error notions). In the novelty detection setting all the scores are available hence the conformal p -values can be computed. Nevertheless, the FDP and the TDP remain unknown since the indices of the nulls/alternatives are not observed. The following is assumed throughout the paper for (weighted) novelty detection.

Assumption 2. *The set of calibration scores $(S_k)_{k \geq 1}$ and the set of test scores $(T_i)_{i \geq 1}$ are two independent families of real random variables. The variables S_k , $k \geq 1$, are i.i.d. with distribution P_{cal} and c.d.f. F_{cal} . The variables T_i , $i \geq 1$, are independent, the variables T_i , $i \in \mathcal{H}_0$, are identically distributed as a null score distribution P_0 and c.d.f. F_0 , and the variables T_i , $i \in \mathcal{H}_1$, are identically distributed as an alternative score distribution P_{test} (potentially different from P_0) with c.d.f. F_{test} . Moreover, F_{cal} , F_0 and F_{test} are continuous.*

We also consider the case of a distribution shift between the distribution of the calibration set P_{cal} and the null distribution P_0 , in which case we propose to use a weighted p -value approach (as for the prediction task). The weights are assumed to be built with an independent training sample so they are considered as fixed here (our results hold conditionally on this sample) considering oracle weights as a particular case with perfect training. For some *weight function* $w : \mathbb{R} \rightarrow \mathbb{R}^+$ and $w(+\infty) \in \mathbb{R}^+$, the weighted conformal p -values are the ones from (7) (by using the specific S_k 's and T_i 's of novelty detection). We define the weighted version of the procedure of Bates et al. (2023) as the BH procedure applied to this weighted p -value family and we denote its rejection set by $\mathcal{R}_\alpha^w \subset \llbracket m \rrbracket$. The FDP and TDP of \mathcal{R}_α^w are defined similarly to above.

Under Assumption 2 and if P_0 is absolutely continuous with respect to P_{cal} , a particular weight is the *oracle weight function* defined by

$$w^* := \frac{dP_0}{dP_{\text{cal}}}. \quad (12)$$

The oracle weighted conformal p -values $p_i^{w^*,(n)}$, $i \in \mathcal{H}_0$, are marginally super-uniform provided that $w^*(+\infty) \geq \sup_{u \in \mathbb{R}} w^*(u)$, see Tibshirani et al. (2019).

Remark 2.2. *The score function $S(\cdot)$ and the weight function $w(\cdot)$ are trained here with an independent training dataset which induces a non-trivial dependence between the scores $(S(Z_i))_{i \in \llbracket n+m \rrbracket}$. However, our results are stated conditionally on the training phase and the score function $S(\cdot)$ and weight function $w(\cdot)$ can be considered as deterministic. Hence, the scores $(S(Z_i))_{i \in \llbracket n+m \rrbracket}$ satisfy Assumption 2 conditionally on the training phase as long as the underlying data $(Z_i)_{i \in \llbracket n+m \rrbracket}$ are independent and have the same distribution within a group (calibration, null, alternative), which is a classical assumption (Bates et al., 2023; Jin and Candès, 2023).*

2.3. Spaces for process convergence

We study the asymptotic convergence of random processes and we consider usual spaces defined in Billingsley (1999) and van der Vaart and Wellner (1996), denoted as usual $D[0, 1]$, $D(0, 1)$ and $\ell^\infty(0, 1)$, and which are briefly described below.

First, $D[0, 1]$ is the set of càdlàg functions $f : [0, 1] \rightarrow \mathbb{R}$ with the usual Skorohod topology. Second, $D(0, 1)$ is the set of càdlàg functions $f : (0, 1) \rightarrow \mathbb{R}$ with the extended Skorohod topology² defined as follows: $(x_n)_n \in D(0, 1)^\mathbb{N}$ converges to $x \in D(0, 1)$ if and only if, for all $(a, b) \in (0, 1)^2$ with $a < b$ such that x is continuous at points a and b , the sequence $(x_n)_n$ restricted to $[a, b]$ converges to x restricted to $[a, b]$ in $D[a, b]$ with the usual Skorohod topology. Finally, $\ell^\infty(0, 1)$ is the space of locally bounded functions $f : (0, 1) \rightarrow \mathbb{R}$ with the topology of the uniform convergence on all compact sets of $(0, 1)$: a sequence $(f_n)_n \in [\ell^\infty(0, 1)]^\mathbb{N}$ converges to $f \in \ell^\infty(0, 1)$ if and only if for all K compact subsets of $(0, 1)$, $\sup_{x \in K} |f_n(x) - f(x)| \rightarrow 0$. We also consider $D(\mathbb{R})$ the space of càdlàg functions from \mathbb{R} to \mathbb{R} with the extended Skorohod topology and $\ell^\infty(\mathbb{R})$ the space of locally bounded functions from \mathbb{R} to \mathbb{R} with the topology of the uniform convergence on all compact sets of \mathbb{R} .

3. Asymptotics for conformal prediction

The aim here is to study the asymptotic properties of the processes $\text{FCP}_m^{(n)}$ (6) and $\text{FCP}_m^{w, (n)}$ (8). We study each process in both exchangeable or distribution shift framework.

²This definition is similar to the convergence in $D[0, +\infty)$ considered in Billingsley (1999)

3.1. Main results

Let us introduce the two following quantities:

$$I(t) := t, \quad t \in [0, 1]; \quad (13)$$

$$G(t) := 1 - F_{\text{test}} \circ F_{\text{cal}}^{-1}(1 - t), \quad t \in (0, 1), \quad (14)$$

where $F_{\text{cal}}^{-1}(u) = \inf\{x \in \mathbb{R} : F_{\text{cal}}(x) \geq u\}$ denotes the general inverse of F_{cal} . Formally, I and G correspond to the c.d.f. of the theoretical p -values $p^{(+\infty)} = 1 - F_{\text{cal}}(T)$ when $T \sim P_{\text{cal}}$ and $T \sim P_{\text{test}}$, respectively. We denote G' the derivative of G provided that it is well defined.

Theorem 3.1. *Under Assumption 1 with $F_{\text{cal}} = F_{\text{test}}$, we have*

$$\sqrt{\tau_{n,m}} \left(\text{FCP}_m^{(n)} - I \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathbb{U} \text{ on } D[0, 1],$$

where $\tau_{n,m}$ is defined by (2) and \mathbb{U} is a standard Brownian bridge.

Remark 3.1. *Theorem 3.1 is true under slightly more general assumptions: namely, it is sufficient that all the scores are exchangeable and without ties almost surely (see Theorem 5.2 for a formal statement). Using the latter is useful since non i.i.d. scores arise naturally in various setting, like in link prediction for graph (Blanchard et al., 2024) or when using adaptive scores (Gazin, Blanchard and Roquain, 2024).*

Theorem 3.1 is proved in Section 5.1.3. As a corollary, if $n/(n+m) \rightarrow \sigma^2 \in (0, 1]$, this gives

$$\sqrt{m} \left(\text{FCP}_m^{(n)} - I \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \sigma^{-1} \mathbb{U}.$$

The latter shows that if the ratio $\tau_{n,m}m^{-1}$ or nm^{-1} is small (see Lemma F.2 for the equivalence), the FCP becomes asymptotically over-dispersed. This is markedly different from the case where we have m i.i.d. p -values uniformly distributed on $(0, 1)$ (compare with the Donsker theorem, see Theorem E.3).

Theorem 3.1 (or more precisely its extension Theorem 5.2 when the scores are exchangeable and without ties almost surely) can be formulated as a result on the asymptotic properties of the empirical c.d.f. of the random vectors (q_1, \dots, q_m) with the distribution $P_{n,m}$ defined in Gazin, Blanchard and Roquain (2024). More formally, let us assume that we are doing m successive draws in a Pólya urn model with initially 1 ball of each of the $n+1$ colors labelled $\{i/(n+1), i \in \llbracket n+1 \rrbracket\}$. We denote $(q_i^{(n)})_{i \in \llbracket m \rrbracket}$ the colors of the m successive draws. Then Theorem 3.1 yields,

$$\left(\sqrt{\tau_{n,m}} \left[m^{-1} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{q_i^{(n)} \leq t} - t \right] \right)_{t \in [0,1]} \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathbb{U} \text{ on } D[0, 1],$$

with \mathbb{U} being a standard Brownian Bridge. Since obtaining this Pólya urn consists on a first sampling of $n+m$ i.i.d. real random variables with a continuous c.d.f., and then sample without replacement n points to compute the ranks of each of the m other points, Theorem 3.1 can also be used to obtain the asymptotic of this kind of sampling without replacement procedure. Another interesting application is the convergence of the uniform distribution over the set of discrete random measure on the set $\{i/(n+1), i \in \llbracket n+1 \rrbracket\}$ with m atoms to the uniform distribution on $(0, 1)$ whenever $n \wedge m \rightarrow +\infty$. Formally, let $Q_m^{(n)}$ be a random probability measure drawn uniformly from the set of discrete random measure on the set $\{i/(n+1), i \in \llbracket n+1 \rrbracket\}$ with m atoms. Denoting for all $t \in [0, 1]$, $F_m^{(n)}(t) = Q_m^{(n)}([0, t])$ its c.d.f. Theorem 3.1 yields the following convergence in distribution

$$\sqrt{\tau_{n,m}} \left(F_m^{(n)} - I \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathbb{U} \text{ on } D[0, 1],$$

with \mathbb{U} being a standard Brownian Bridge. This gives that the random measure $(Q_m^{(n)})_{m,n}$ converges in probability to the uniform distribution on $(0, 1)$ for the Kolmogorov Smirnov divergence, with

an asymptotic quantification of its fluctuations. All those convergences are direct consequences a direct corollary of Theorem 3.1 (and Theorem A in the supplementary material of Gazin, Blanchard and Roquain (2024) giving us various properties of $P_{n,m}$). However, establishing such links between conformal inference and Pólya urn theory in the event of a distribution shift is neither straightforward nor direct.

Let us now consider the case where we have potentially a distribution shift, that is, $F_{\text{cal}} \neq F_{\text{test}}$. For this, we consider the following additional assumptions on F_{cal} and F_{test} .

Assumption 3. F_{cal} is increasing on its support $(a, b) \subset \mathbb{R}$ with $-\infty \leq a < b \leq +\infty$ and is continuously differentiable. In addition, F_{test} is continuously differentiable.

Remark 3.2. Assumptions on F_{cal} and F_{test} (Assumption 3 and the continuity assumption of Assumption 1) can be slightly relaxed by rather imposing a condition on the marginal distribution $F_{\text{test}} \circ F_{\text{cal}}^{-1}$ of the theoretical p -value (5): Section G is devoted to the formal statement of those relaxed theorems.

Theorem 3.2. Under Assumptions 1 and 3 and assuming that $n/(n+m)$ tends to $\sigma^2 \in [0, 1]$, we have

$$\sqrt{\tau_{n,m}} \left(\text{FCP}_m^{(n)} - G \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \sigma \mathbb{U} \circ G + \sqrt{1 - \sigma^2} G' \mathbb{V} \text{ on } D(0, 1),$$

where \mathbb{U} and \mathbb{V} are two independent Brownian bridges.

Theorem 3.2 is proved in Section 5.1.4. When $F_{\text{test}} = F_{\text{cal}}$ it recovers Theorem 3.1 (under the additional convergence assumption of $n/(n+m)$) since if \mathbb{U} and \mathbb{V} are two independent standard Brownian bridges and $\sigma \in [0, 1]$, then $\sigma \mathbb{U} + \sqrt{1 - \sigma^2} \mathbb{V}$ is also a standard Brownian bridge. By contrast, when $F_{\text{test}} \neq F_{\text{cal}}$, Theorem 3.2 quantifies the effect of the distribution shift. In particular, the limit of $\text{FCP}_m^{(n)}$ is G instead of simply I , thus $\text{FCP}_m^{(n)}(\alpha)$ is not provided to converge to α anymore.

The conclusions of Theorems 3.1 and 3.2 involve a double asymptotic in n and m for which $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$. The different regimes are summarized in Table 1. Whenever σ^2 is neither 0 nor 1 this is equivalent to have $m \sim_{\tau_{n,m} \rightarrow \infty} (\sigma^2 - 1)n$ and thus m and n are asymptotically proportional. For $\sigma^2 = 0$ ('under-calibrated' regime), the convergence rate is slower, of order $\sqrt{n} \ll \sqrt{m}$, for instance $m^{1/4}$ for $n = \sqrt{m}$, or $\sqrt{m(\log(m))^{-1}}$ for $n = m(\log(m))^{-1}$. This corresponds to the case where the conformal p -values are 'strongly' dependent. On the other hand, if $\sigma^2 = 1$ ('over-calibrated' regime), the rate $\sqrt{\tau_{n,m}}$ is equivalent to the classical rate in case of independence \sqrt{m} , with the same asymptotic variance. In that case, the convergence result is the same as for the independent theoretical p -values (5). This observation still apply in the weighted conformal prediction case.

	Regime of convergence	Rate of convergence: $\sqrt{\tau_{n,m}}$	Variance term
Conformal	Under-calibrated: $m/n \rightarrow +\infty$	\sqrt{n}	$\sigma^2 = 0$
	Proportional: $n \sim \gamma m$, $\gamma > 0$	$\sqrt{\frac{\gamma}{\gamma+1} m}$	$\sigma^2 = \frac{\gamma}{\gamma+1} \in (0, 1)$
	Over-calibrated: $m/n \rightarrow 0$	\sqrt{m}	$\sigma^2 = 1$
Independent	Theoretical p -values (as $n = +\infty$)	\sqrt{m}	as $\sigma^2 = 1$

TABLE 1

Three different regimes of convergence: $\sigma^2 = 0$, $\sigma^2 \in (0, 1)$ and $\sigma^2 = 1$ linked with the ratio between the number of calibration points and test points. The case of theoretical independent p -values is presented for comparison.

3.2. Weighted case

Let $w : \mathbb{R} \cup \{\infty\} \mapsto \mathbb{R}^+$ be a bounded and measurable weight function. To describe the asymptotic behavior of $\text{FCP}_m^{w,(n)}$, we need to introduce few additional quantities. First, we let

$$F_{\text{cal}}^w(t) := \frac{\int_{-\infty}^t w(u) dP_{\text{cal}}(u)}{\int_{\mathbb{R}} w(u) dP_{\text{cal}}(u)}, \quad t \in \mathbb{R}, \quad (15)$$

which corresponds to the c.d.f. of the distribution induced by weight function w on the distribution P_{cal} (i.e. proportional to $w(u) dP_{\text{cal}}(u)$). Note that when choosing the oracle weight function (9), the latter is simply the c.d.f. of P_{test} . Second, we let

$$G^w(t) := 1 - F_{\text{test}} \circ (F_{\text{cal}}^w)^{-1}(1 - t), \quad t \in (0, 1), \quad (16)$$

the c.d.f. of the theoretical p -values $p^{w,(+\infty)} = 1 - F_{\text{cal}}^w(T)$ with $T \sim P_{\text{test}}$, and we denote $(G^w)'$ its derivative when it exists. Finally, we introduce quantities involved in the asymptotic variance:

$$\rho^w := \frac{\left(\int_{\mathbb{R}} w(u)^2 dP_{\text{cal}}(u)\right)^{1/2}}{\int_{\mathbb{R}} w(u) dP_{\text{cal}}(u)} \geq 1; \quad (17)$$

$$I^w(t) := 1 - V_{\text{cal}}^w \circ (F_{\text{cal}}^w)^{-1}(1 - t), \quad t \in (0, 1), \quad (18)$$

with

$$V_{\text{cal}}^w(t) := \frac{\int_{-\infty}^t w(u)^2 dP_{\text{cal}}(u)}{\int_{\mathbb{R}} w(u)^2 dP_{\text{cal}}(u)}, \quad t \in \mathbb{R}. \quad (19)$$

Assumption 4. *The weight function $w : \mathbb{R} \cup \{\infty\} \rightarrow \mathbb{R}^+$ is uniformly bounded by a constant and is measurable. Moreover, F_{cal}^w is increasing on its support $(a^w, b^w) \subset \mathbb{R}$ with $-\infty \leq a^w < b^w \leq +\infty$ and is continuously differentiable.*

Typically, Assumption 4 holds if $\{t \in \mathbb{R}, w(t) > 0\}$ is a possibly infinite interval of \mathbb{R} on which w is bounded and continuous, and F_{cal} satisfies Assumption 3.

Theorem 3.3. *Under Assumptions 1, 3, 4 and assuming that $n/(n+m)$ tends to $\sigma^2 \in [0, 1]$, we have*

$$\sqrt{\tau_{n,m}} \left(\text{FCP}_m^{w,(n)} - G^w \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \sigma \mathbb{U} \circ G^w + \sqrt{1 - \sigma^2} \rho^w (G^w)' (\mathbb{V} \circ I^w + [I - I^w] N) \text{ on } D(0, 1),$$

where \mathbb{U}, \mathbb{V} are two independent standard Brownian bridges, N is an independent standard Gaussian random variable, ρ^w is defined by (17) and I^w is defined by (18).

Theorem 3.3 is proved in Section 5.1.5. Note that it recovers Theorem 3.2 when $w \equiv 1$. The asymptotic expectation of $\text{FCP}_m^{w,(n)}(\alpha)$ is $G^w(\alpha)$ which is different from α for a general w . Interestingly, when choosing the oracle weight function w^* given by (9), we have $G^{w^*} = I$ and the convergence in Theorem 3.3 reads as follows:

Theorem 3.4. *Under Assumptions 1 and 3, assume that P_{test} is absolutely continuous with respect to P_{cal} and that the oracle weight function w^* (9) satisfies Assumption 4. If the ratio $n/(n+m)$ tends to $\sigma^2 \in [0, 1]$, we have,*

$$\sqrt{\tau_{n,m}} \left(\text{FCP}_m^{w^*,(n)} - I \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \sigma \mathbb{U} + \sqrt{1 - \sigma^2} \rho^{w^*} \left(\mathbb{V} \circ I^{w^*} + [I - I^{w^*}] N \right) \text{ on } D(0, 1),$$

where \mathbb{U}, \mathbb{V} are two independent standard Brownian bridge and N is an independent standard Gaussian random variable.

Theorem 3.4 shows that, in the case of a perfect training of the weight function, one can recover the same asymptotic behavior as in case of no distribution shift. However, note that the asymptotic variance is no longer universal and depends on the distributions proportionnal to $w dP_{\text{cal}}$ and

$w^2 \text{d} P_{\text{cal}}$. By contrast, Theorem 3.3 quantifies the impact of the distribution shift when the weight function is mis-specified, with a limit G^w and not I . Comparing this to Theorem 3.2, Theorem 3.3 indicates that, for a level α , the FDP behavior is (asymptotically) more favorable if we consider a weight function such that $G^w(\alpha)$ is closer to α than $G(\alpha)$. This phenomenon is illustrated in Figure 2 in a particular setting.

Finally, the covariance terms of Theorems 3.3 and 3.2 can be used to build asymptotic confidence intervals for the FCP, as illustrated in Figure 2.

4. Asymptotics for novelty detection

Recall the novelty detection setting of Section 2.2. The aim here is to study the asymptotic properties of the process $\text{FDP}_m^{(n)}(\mathcal{R}_\alpha)$ (10) and of the process $\text{TDP}_m^{(n)}(\mathcal{R}_\alpha)$ (11).

4.1. Additional notation and assumptions

We denote, for all $m \geq 1$, $m_0(m) = |\mathcal{H}_0 \cap \llbracket m \rrbracket|$ and $\pi_0(m) = m_0(m)/m$ the number and the proportion of nulls (i.e. of non-novelty) among m tested points, respectively. We introduce the following quantities and recall the definition of G (14):

$$G(t) = 1 - F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-t), \quad t \in (0, 1),$$

$$G_0(t) = 1 - F_0 \circ F_{\text{cal}}^{-1}(1-t), \quad t \in (0, 1), \quad (20)$$

$$G_{\text{mixt}} = \pi_0 G_0 + (1 - \pi_0) G, \quad (21)$$

that correspond to the (limiting) c.d.f. of the theoretical p -values (5) under the alternative, the null and under the test sample mixture, respectively. For a given weight function w , we also define their weighted counterparts and recall the definition of G^w (16):

$$G^w(t) = 1 - F_{\text{test}} \circ (F_{\text{cal}}^w)^{-1}(1-t), \quad t \in (0, 1),$$

$$G_0^w(t) = 1 - F_0 \circ (F_{\text{cal}}^w)^{-1}(1-t), \quad t \in (0, 1), \quad (22)$$

$$G_{\text{mixt}}^w = \pi_0 G_0^w + (1 - \pi_0) G^w. \quad (23)$$

We denote G'_{mixt} and $(G^w_{\text{mixt}})'$ the derivatives of G_{mixt} and G^w_{mixt} , respectively. Note that Assumption 2 with $F_0 = F_{\text{cal}}$ entails $G_0 = I$. If $F_0 \neq F_{\text{cal}}$, we still have $G_0^{w^*} = I$ when w^* is the oracle weight function.

Let us introduce the two following assumptions:

Assumption 5. G_{mixt} is strictly concave on $(0, 1)$.

Assumption 6. G^w_{mixt} is strictly concave on $(0, 1)$.

Since G_{mixt} can be seen as the cdf of the theoretical p -values (5) in the test sample, Assumption 5 is classical in the multiple testing literature (Genovese and Wasserman, 2002) (see Section G for more precision), and Assumption 6 corresponds to its weighted version.

As shown in Chi (2007); Neuvial (2008) in the independent case, there is a critical value for the BH procedure given by $\alpha^* = [G'_{\text{mixt}}(0^+)]^{-1}$; the central limit theorem for the FDP/TDP of BH procedure at level α can only be obtained if $\alpha > \alpha^*$ (for $\alpha < \alpha^*$, the BH procedure at level α has asymptotically no power). This quantity plays a similar role in the (dependent) conformal setting and we prove results only if $\alpha > \alpha^*$.

Finally, as in Neuvial (2008), we consider also the following thresholds

$$\mathcal{T}_\alpha = \sup \left\{ t \in (0, 1), G_{\text{mixt}}(t) \geq \frac{t}{\alpha} \right\}; \quad (24)$$

$$\mathcal{T}_\alpha^w = \sup \left\{ t \in (0, 1), G^w_{\text{mixt}}(t) \geq \frac{t}{\alpha} \right\}, \quad (25)$$

which are well defined (and belong to $(0, 1)$) under Assumption 5 if $\alpha > [(G_{\text{mixt}})'(0^+)]^{-1}$ and Assumption 6 if $\alpha > [(G^w_{\text{mixt}})'(0^+)]^{-1}$, respectively.

4.2. Main results

When $P_0 = P_{\text{cal}}$ (no distribution shift), the following result provides the convergence behavior of the novelty detection procedure of [Bates et al. \(2023\)](#), that is, of \mathcal{R}_α given in [Section 2.2](#).

Theorem 4.1. *Under Assumption 2 with $P_0 = P_{\text{cal}}$, Assumption 3 and Assumption 5, let us consider \mathcal{R}_α the BH procedure at level α applied to the conformal p -values (4), with a level $\alpha > [G'_{\text{mixt}}(0^+)]^{-1}$. If $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0, 1)$, we have*

$$\sqrt{\tau_{n,m}} \left(\text{FDP}_m^{(n)}(\mathcal{R}_\alpha) - \pi_0 \alpha \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \alpha^2 \pi_0 [\sigma^2 + (1 - \sigma^2) \pi_0] \frac{1 - \mathcal{T}_\alpha}{\mathcal{T}_\alpha} \right); \quad (26)$$

$$\sqrt{\tau_{n,m}} \left(\text{TDP}_m^{(n)}(\mathcal{R}_\alpha) - G(\mathcal{T}_\alpha) \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{\Sigma_\alpha}{(\alpha^{-1} - G'_{\text{mixt}}(\mathcal{T}_\alpha))^2} \right), \quad (27)$$

with $\Sigma_\alpha = G'(\mathcal{T}_\alpha)^2 \mathcal{T}_\alpha (1 - \mathcal{T}_\alpha) [\pi_0 \sigma^2 + (1 - \sigma^2) \alpha^{-2}] + [\alpha^{-1} - \pi_0]^2 (1 - \pi_0)^{-1} G(\mathcal{T}_\alpha) (1 - G(\mathcal{T}_\alpha)) \sigma^2$.

Theorem 4.1, proved in [Section 5.2.3](#), relies on the pipeline introduced by [Neuviel \(2008\)](#), which consists in first deriving a functional central limit theorem for the e.c.d.f. of the p -values and then using the functional delta method ([van der Vaart, 1998](#)).

Interestingly, under the assumptions of [Theorem 4.1](#) and if $\sigma^2 > 0$, (26) yields

$$\sqrt{m} \left(\text{FDP}_m^{(n)}(\mathcal{R}_\alpha) - \pi_0 \alpha \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \alpha^2 \pi_0 [1 + (\sigma^{-2} - 1) \pi_0] \frac{1 - \mathcal{T}_\alpha}{\mathcal{T}_\alpha} \right).$$

When $\sigma = 1$ (that is $n/m \rightarrow \infty$), we note that the FDP convergence is the same as when the p -values are independent ([Neuviel, 2008](#)). However, when n/m is bounded, the asymptotic variance is affected by the dependence and gets larger when σ decreases. In addition, and maybe more importantly, the convergence rate is heavily affected in the under-calibrated setting when $\sigma^2 = 0$, with a convergence rate $\sqrt{\tau_{n,m}} \ll \sqrt{m}$. This phenomenon is illustrated in [Figure 3](#) for $m = 10^3$ and various n in a particular setting, with corresponding convergence rates reported in [Table 2](#). The three different regimes from the prediction setting displayed in [Table 1](#) also appear in this novelty detection setting.

	Calibration sample size: n	Rate of convergence: $\sqrt{\tau_{n,m}}$	Independence coefficient: σ^2
Conformal	\sqrt{m}	$m^{1/4}$	0
	$\frac{m}{\log(m)}$	$\sqrt{\frac{m}{\log(m)}}$	0
	$\frac{m}{2}$	$\sqrt{\frac{m}{3}}$	$\frac{1}{3}$
	m	$\sqrt{\frac{m}{2}}$	$\frac{1}{2}$
	$2m$	$\sqrt{\frac{2m}{3}}$	$\frac{2}{3}$
	$\gamma m, \gamma > 0$	$\sqrt{\frac{\gamma}{\gamma+1} m}$	$\frac{\gamma}{\gamma+1}$
	$m \log(m)$	\sqrt{m}	1
	$m^{3/2}$	\sqrt{m}	1
	m^2	\sqrt{m}	1
Independent	$+\infty$	\sqrt{m}	1

TABLE 2

Rates $\sqrt{\tau_{n,m}}$ and σ^2 in [Theorems 3.2, 3.3, 4.1 and 4.2](#), with different choices of n and m . Except the case $n = \gamma m$, they all correspond to the different regimes displayed in [Figure 3](#). The case with independent p -values is presented for comparison.

Under distribution shift, the following result holds.

Theorem 4.2. *Under Assumptions 2 and 6, assume that P_0 is absolutely continuous with respect to P_{cal} and that the oracle weight function w^* (12) satisfies Assumptions 3 and 4. Consider the*

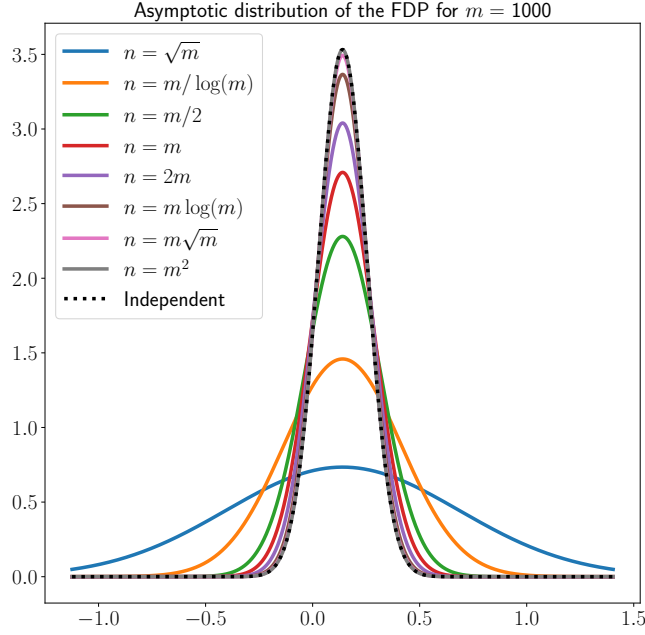


Fig 3: Asymptotic approximation of the distribution of the FDP of the Benjamini-Hochberg procedure with conformal p -values given by Theorem 4.1 with $m = 10^3$ and several choice of n (plain lines) at level $1 - \alpha = 0.8$. The case of independent theoretical p -values (dashed line) is displayed for comparison. The null distribution and the calibration distribution is $\mathcal{N}(0, 1)$ while the distribution of the alternative is $\mathcal{N}(3, 1)$, with a null proportion π_0 equal to 70%.

BH procedure at level α applied to the oracle weighted p -values $(p_i^{w^*,(n)})_{i \in \llbracket m \rrbracket}$ (7), with a level $\alpha > [(G_{mixt}^{w^*})'(0^+)]^{-1}$. Then if $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0, 1)$, we have

$$\sqrt{\tau_{n,m}} \left(\text{FDP}_m^{(n)}(\mathcal{R}_\alpha^{w^*}) - \pi_0 \alpha \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \alpha^2 \pi_0 \Xi_\alpha^{w^*} \frac{1 - \mathcal{T}_\alpha^{w^*}}{\mathcal{T}_\alpha^{w^*}} \right); \quad (28)$$

$$\sqrt{\tau_{n,m}} \left(\text{TDP}_m^{(n)}(\mathcal{R}_\alpha^{w^*}) - G^{w^*}(\mathcal{T}_\alpha^{w^*}) \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{\Sigma_\alpha^{w^*}}{(\alpha^{-1} - (G_{mixt}^{w^*})'(\mathcal{T}_\alpha^{w^*}))^2} \right), \quad (29)$$

where we denote

$$\begin{aligned} \Xi_\alpha^{w^*} &= \sigma^2 + (1 - \sigma^2) \rho^{w^*2} \pi_0 \frac{I^{w^*}(\mathcal{T}_\alpha^{w^*}) \mathcal{T}_\alpha^{w^*-1} + \mathcal{T}_\alpha^{w^*} - 2I^{w^*}(\mathcal{T}_\alpha^{w^*})}{1 - \mathcal{T}_\alpha^{w^*,*}}; \\ \Sigma_\alpha^{w^*} &= \left((G^{w^*})'(\mathcal{T}_\alpha^{w^*}) \right)^2 \mathcal{T}_\alpha^{w^*} (1 - \mathcal{T}_\alpha^{w^*}) \pi_0 \sigma^2 \\ &\quad + \left(\rho^{w^*} (G^{w^*})'(\mathcal{T}_\alpha^{w^*}) \alpha^{-1} \right)^2 \left[\mathcal{T}_\alpha^{w^*} - I^{w^*}(\mathcal{T}_\alpha^{w^*})^2 \right] (1 - \sigma^2) \\ &\quad + [\alpha^{-1} - \pi_0]^2 (1 - \pi_0)^{-1} G^{w^*}(\mathcal{T}_\alpha^{w^*}) (1 - G^{w^*}(\mathcal{T}_\alpha^{w^*})) \sigma^2. \end{aligned}$$

Theorem 4.2 is proved in Section 5.2.4. The weighted conformal BH procedure with (similar) weighted conformal p -values has been studied in Jin and Candès (2023) and they obtained a convergence of FDR and TDR (i.e. $\mathbb{E}[\text{FDP}_m^{(n)}(\mathcal{R}_\alpha^{w^*})]$ and $\mathbb{E}[\text{TDP}_m^{(n)}(\mathcal{R}_\alpha^{w^*})]$) to quantities analogue to $\pi_0 \alpha$ and $G^{w^*}(\mathcal{T}_\alpha^{w^*})$, respectively. By contrast, Theorem 4.2 provides the full asymptotic distribution.

While Theorem 4.2 focuses on the oracle weight function, the non-oracle case is deferred to Section D (see Theorem D.2).

The limiting distribution in Theorems 4.1 and 4.2 are similar to the previous ones in the literature (Neuville, 2008; Delattre and Roquain, 2011, 2016; Kluger and Owen, 2024) which is well expected since all these works study the asymptotic normality of the FDP and the TDP of the BH procedure. However, the novelty is in the setting: we do not have access to the theoretical and uniform p -values under the null, but only to the conformal p -values. Conformal p -values are by nature dependent (and PRDS for the unweighted ones (Bates et al., 2023)) even with independent scores, while the dependency structures on the theoretical p -values in Delattre and Roquain (2011, 2016); Kluger and Owen (2024) come from the dependency of the gaussian test statistics. Furthermore conformal p -values are approximations of the theoretical ones, hence the study of the double asymptotic in both the number of test m (which is the only one studied in the literature) and the size n of the calibration set is crucial and provided here. Equivalent of the rate of convergence $\sqrt{\tau_{n,m}}$ and the limit σ^2 in Theorems 3.2 and 4.1 and their weighted version Theorems 3.3 and 4.2 for different choice of n depending on m .

4.3. Applications

Theorem 4.1 has several possible applications. A first application is to obtain an asymptotic confidence region of the FDP of the BH procedure in the conformal case.

Corollary 4.3. *Let $\delta > 0$. Under Assumption 2 with $P_0 = P_{cal}$ and Assumptions 3 and 5, let us consider \mathcal{R}_α the BH procedure at level α applied to m conformal p -values (4) with a calibration sample of size n , with a level $\alpha > [G'_{mixt}(0^+)]^{-1}$. Denote $\hat{\mathcal{T}}_\alpha = \alpha |\mathcal{R}_\alpha| / m$ the threshold of \mathcal{R}_α . Define the upper bound:*

$$\hat{\lambda}_{n,m}(\delta, \alpha, \pi_0) = \alpha \sqrt{\pi_0 \left[\frac{1}{m} + \frac{1}{n} \pi_0 \right] \frac{1 - \hat{\mathcal{T}}_\alpha}{\hat{\mathcal{T}}_\alpha} \Phi^{-1}(1 - \delta)},$$

with Φ^{-1} the quantile function of the standard gaussian distribution. Then if $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0, 1)$, we have

$$\mathbb{P}\left(\text{FDP}_m^{(n)}(\mathcal{R}_\alpha) \leq \pi_0 \alpha + \hat{\lambda}_{n,m}(\delta, \alpha, \pi_0)\right) \xrightarrow{\tau_{n,m} \rightarrow +\infty} 1 - \delta.$$

In addition, if $\hat{\pi}_0 \in [0, 1]$ is an estimator of π_0 such that $\underline{\lim} \hat{\pi}_0 \geq \pi_0$ then,

$$\underline{\lim}_{\tau_{n,m} \rightarrow +\infty} \mathbb{P}\left(\text{FDP}_m^{(n)}(\mathcal{R}_\alpha) \leq \hat{\pi}_0 \alpha + \hat{\lambda}_{n,m}(\delta, \alpha, \hat{\pi}_0)\right) \geq 1 - \delta. \quad (30)$$

Corollary 4.3 is a direct application of the asymptotic result (26). It provides an asymptotic FDP control which takes into account the fluctuation of the FDP around the FDR. In (30) the estimator $\hat{\pi}_0$ should satisfy $\underline{\lim} \hat{\pi}_0 \geq \pi_0$. This includes the conservative case where $\hat{\pi}_0$ is constantly equal to 1, but also the (in general) sharper bound for which $\hat{\pi}_0$ is the Storey estimator (Storey, Taylor and Siegmund, 2004), defined as

$$\hat{\pi}_0^{\text{Storey}}(m, n) = \frac{1 + \sum_{i=1}^m \mathbf{1}_{p_i^{(m)} \geq \lambda}}{m(1 - \lambda)}, \quad (31)$$

where $\lambda \in (0, 1)$ is a fixed parameter. Indeed, Applying Theorem 5.3 (or more precisely arguments in its proof, because the convergence of all empirical processes holds almost surely by the Glivenko-Cantelli theorem and by continuity), we have

$$\hat{\pi}_0^{\text{Storey}}(n, m) \xrightarrow{\tau_{n,m} \rightarrow +\infty} \pi_0 + (1 - \pi_0) \frac{1 - G(\lambda)}{1 - \lambda} \geq \pi_0 \text{ a.s.}$$

Note that using the Storey estimator in the context of conformal p -values has been already considered in the literature (e.g., [Bates et al. \(2023\)](#); [Marandon et al. \(2024\)](#)), but not for deriving an FDP bound.

Another application of [Theorem 4.1](#) is to use the asymptotic distribution in [\(27\)](#) to design the score function $S : \mathcal{Z} \rightarrow \mathbb{R}$, see [Section 2.2](#). Recall that the latter is computed from an independent training sample $\mathcal{D}_{\text{train}}$. For simplicity, assume that $\mathcal{D}_{\text{train}} = \{(Z_i, h_i) \in \mathcal{Z} \times \{0, 1\}, i \in \llbracket n_{\text{train}} \rrbracket\}$ is a sample containing both labeled data under the null ($h_i = 0$) and under the alternative ($h_i = 1$), so that the score function can typically be trained as the best classifier of null versus alternative ([Marandon et al., 2024](#)), or as the likelihood ratio of null and alternative “conformal p -values” leading to some pruning step ([Liang, Sesia and Sun, 2024](#)). The asymptotic distribution in [\(27\)](#) delivers additional insight, and allows to use the asymptotic distribution to design a score function in a more TDP (or broadly “multiple testing”) oriented way. Namely, let us assume that the score function S lies in a parametric class $\{S_\theta, \theta \in \Theta\}$ of score functions (for example a classifier based on a neural network). The classical machine learning approach is to seek for the best two-class classifier by minimising the empirical risk $\theta \in \Theta \mapsto n_{\text{train}}^{-1} \sum_{i \in \llbracket n_{\text{train}} \rrbracket} \mathbb{1}_{S_\theta(Z_i) \neq h_i}$ which is an empirical substitute of $\mathbb{P}(S_\theta(Z_1) \neq h_1)$ computed on $\mathcal{D}_{\text{train}}$. Alternatively, let us consider the situation where the user seeks to have the “majority” of the TDP above a specific power level $\beta \in (0, 1)$, while conserving the finite-sample FDR control given by the conformal BH procedure. In this case, the task is to choose θ that maximises $\mathbb{P}(\text{TDP}_m^{(n)} > \beta) \approx 1 - \Phi(\sqrt{\tau_{n,m}} L_\alpha(\beta))$ with $L_\alpha(\beta) = (\beta - G(\mathcal{T}_\alpha))(\alpha^{-1} - G'_{\text{mixt}}(\mathcal{T}_\alpha))\Sigma_\alpha^{-1/2}$ by using [Theorem 4.1](#). A natural empirical substitution leads to minimise

$$\theta \in \Theta \mapsto (\beta - \widehat{G}_\theta)(\widehat{\mathcal{T}}_{\theta,\alpha})(\alpha^{-1} - \widehat{G}'_{\theta,\text{mixt}}(\widehat{\mathcal{T}}_{\theta,\alpha}))\widehat{\Sigma}_{\theta,\alpha}^{-1/2}, \quad (32)$$

where \widehat{G}_θ is a differentiable estimate based on $\mathcal{D}_{\text{train}}$ (e.g. a kernel estimator or a local polynomial estimator) of G [\(14\)](#) when the score function is S_θ (i.e. of $1 - F_{\theta,\text{test}} \circ F_{\theta,\text{cal}}^{-1}(1 - \cdot)$ where for all $t \in \mathbb{R}$, $F_{\theta,\text{test}}(t)$ (resp. $F_{\theta,\text{cal}}(t)$) is equal to $\mathbb{P}(S_\theta(Z_i) \leq t)$ with $i \in \mathcal{H}_0$ (resp. $i \in \mathcal{H}_1$)), its derivative is denoted \widehat{G}'_θ and all the other quantities are plug-in estimates of G_{mixt} [\(21\)](#), \mathcal{T}_α [\(46\)](#) and Σ_α defined in [\(27\)](#) with G and G' replaced by \widehat{G}_θ and \widehat{G}'_θ , the function G_0 replaced by I and π_0 replaced by $n_{\text{train}}^{-1} \sum_{i \in \llbracket n_{\text{train}} \rrbracket} (1 - h_i)$. More generally, other objective functions can be optimised by considering other TDP related quantities approximated by the asymptotic distribution [\(27\)](#) or by using versions of the empirical risk that are penalised by [\(32\)](#). Finally, a similar approach can be followed in the case of a distribution shift by applying [Theorem 4.2](#) (or its generalisation [Theorem D.2](#)). In this case, the oracle weight should also be seen as a function of θ (if the oracle weight can be known for every score function $S_\theta, \theta \in \Theta$) or must be chosen among a parameterised family of weight function $\{w_{\lambda,\theta}, (\lambda, \theta) \in \Lambda \times \Theta\}$ (when only some information on the distribution shift is known) and the training sample should contain both clean and shifted data. This heuristic way to develop score (and weight) function with a TDP focus and not only a best “marginal” classification is left for future work.

5. Proofs

In this section, we prove the main results of the paper. They rely on a particular decomposition of the FCP/FDP into two processes that further jointly converge. Applying the functional delta method ([van der Vaart, 1998](#)) then allows to conclude.

5.1. Proofs for [Section 3](#)

5.1.1. FCP decomposition

To emphasise that the FCP is the e.c.d.f. of $(p_i^{(n)})_{i \in \llbracket m \rrbracket}$, we let for all $n \geq 1$ and $m \geq 1$,

$$\widehat{G}_m^{(n)} := \frac{1}{m} \sum_{i=1}^m \mathbb{1}_{p_i^{(n)} \leq t} = \text{FCP}_m^{(n)}(t), \quad t \in [0, 1]. \quad (33)$$

We also introduce, for all $n \geq 1$ and $m \geq 1$,

$$\widehat{F}_{\text{cal}}^{(n)}(t) = \frac{1}{n+1} \sum_{k=1}^n \mathbb{1}_{S_k \leq t}, \quad t \in \mathbb{R}, \quad (34)$$

$$\widehat{F}_{m,\text{test}}(t) = \frac{1}{m} \sum_{i=1}^m \mathbb{1}_{T_i \leq t}, \quad t \in \mathbb{R}, \quad (35)$$

corresponding to the e.c.d.f. of the calibration score sample (with $+\infty$) and test score sample, respectively. Note that we have $p_i^{(n)} = 1 - \widehat{F}_{\text{cal}}^{(n)}(T_i)$ almost surely (under Assumption 1 to ensure that there are no ties almost surely).

By Lemma F.1 we have that $p_i^{(n)} \leq t$ if and only if $T_i > (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t)$ for all $t \in (0, 1)$. Recalling (14), this leads to the following decomposition for $\widehat{G}_m^{(n)}$: for all $t \in (0, 1)$,

$$\begin{aligned} \widehat{G}_m^{(n)}(t) - G(t) &= \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{p_i^{(n)} \leq t} - G(t) \\ &= \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{(\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t) < T_i} - (1 - F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-t)) \\ &= 1 - \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{(\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t) \geq T_i} - (1 - F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-t)) \\ &= F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-t) - \widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t). \end{aligned}$$

Hence, it follows

$$\begin{aligned} \widehat{G}_m^{(n)}(t) - G(t) &= F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t) - \widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t) \\ &\quad + F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-t) - F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t). \end{aligned}$$

Finally, we obtain the decomposition

$$\begin{aligned} \sqrt{\tau_{n,m}} \left(\widehat{G}_m^{(n)}(t) - G(t) \right) &= - \sqrt{\frac{\tau_{n,m}}{m}} \sqrt{m} \left(\widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t) - F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t) \right) \\ &\quad - \sqrt{\frac{\tau_{n,m}}{n}} \sqrt{n} \left(F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-t) - F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-t) \right). \quad (36) \end{aligned}$$

5.1.2. Joint convergence

Thanks to (36), in order to obtain a convergence result for $\sqrt{\tau_{n,m}} \left(\widehat{G}_m^{(n)} - G \right)$, we only need to derive the joint convergence of the two processes delineated in the decomposition (36).

Proposition 5.1. *Under Assumptions 1 and 3 we have,*

$$\left(\begin{array}{c} \sqrt{m} \left[\widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) \right] \\ \sqrt{n} \left[F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-I) \right] \end{array} \right) \xrightarrow{\mathcal{L}} \begin{pmatrix} \mathbb{U} \circ G \\ G' \mathbb{V} \end{pmatrix} \text{ on } [D(0,1)]^2, \quad (37)$$

where \mathbb{U} and \mathbb{V} are two independent Brownian bridges.

Proposition 5.1 is proved in Section B.1. The main idea of the proof is to study each coordinate separately and then to use independence to obtain a joint convergence. The process of the first coordinate (test term) is studied by using the Donsker Theorem for $\widehat{F}_{m,\text{test}}$ and the continuous mapping theorem with a random change of time (Lemma E.10). The second coordinate (calibration term) is investigated by using the Donsker theorem for $(\widehat{F}_{\text{cal}}^{(n)})^{-1}$ and then using the functional delta method with the map $\phi \mapsto F_{\text{test}} \circ \phi$ (Lemma E.5).

5.1.3. Proof of Theorem 3.1

To prove Theorem 3.1, we show the following more general result.

Theorem 5.2. *Assume that for all n, m the calibration and test scores $(S_1, \dots, S_n, T_1, \dots, T_m)$ are exchangeable without ties almost surely. Then*

$$\sqrt{\tau_{n,m}} \left(\text{FCP}_m^{(n)} - I \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathbb{U} \text{ on } D[0, 1],$$

where $\tau_{n,m}$ is defined by (2) and \mathbb{U} is a standard Brownian bridge.

The two assumptions in Theorem 5.2 are implied by Assumption 1 with $F_{\text{cal}} = F_{\text{test}}$, hence Theorem 3.1 is implied by Theorem 5.2. Let us prove the latter. Since the calibration and test scores are exchangeable without ties a.s. by Proposition 2.2 from Gazin, Blanchard and Roquain (2024) we get for all $n, m \in \mathbb{N}^2$ that the conformal p -values $(p_i^{(n)})_{i \in \llbracket m \rrbracket}$ has a distribution only depending on n and m , denoted by $P_{n,m}$. Hence, for all $T : D[0, 1] \rightarrow \mathbb{R}$ bounded continuous function we have :

$$\mathbb{E} \left[T \left(\sqrt{\tau_{n,m}} \left(\text{FCP}_m^{(n)} - I \right) \right) \right] = \mathbb{E} \left[T \left(\sqrt{\tau_{n,m}} \left(\widetilde{\text{FCP}}_m^{(n)} - I \right) \right) \right], \quad (38)$$

where $\widetilde{\text{FCP}}_m^{(n)}$ is the FCP obtained with i.i.d scores uniformly distributed on $(0, 1)$. Those specific scores satisfy Assumptions 1 and 3 with $F_{\text{cal}} = F_{\text{test}} = I$ on $[0, 1]$, and by (38) we only have to prove the convergence in this specific setting.

Up to consider a subsequence, one can assume that $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ by compacity of $[0, 1]$. By applying Proposition 5.1 with the i.i.d. uniformly distributed on $(0, 1)$ scores from the right hand side of (38), we get by Slutsky's Lemma:

$$\left(\begin{array}{c} \sqrt{m} \left[\widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1} (1 - I) - (\widehat{F}_{\text{cal}}^{(n)})^{-1} (1 - I) \right] \\ \sqrt{n} \left[(\widehat{F}_{\text{cal}}^{(n)})^{-1} (1 - I) - (1 - I) \right] \\ \frac{\tau_{n,m}}{m} \\ \frac{\tau_{n,m}}{n} \end{array} \right) \xrightarrow{\mathcal{L}} \left(\begin{array}{c} \mathbb{U} \\ \mathbb{V} \\ \sigma^2 \\ 1 - \sigma^2 \end{array} \right) \text{ on } [D(0, 1)]^2 \times [0, 1]^2,$$

with (\mathbb{U}, \mathbb{V}) two independent standard Brownian bridges. Applying the continuous mapping theorem with the decomposition (36) gives the following convergence:

$$\sqrt{\tau_{n,m}} \left(\widetilde{\text{FCP}}_m^{(n)} - I \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \sigma \mathbb{U} + \sqrt{1 - \sigma^2} \mathbb{V} \stackrel{\mathcal{L}}{=} \mathbb{W},$$

with \mathbb{W} a standard Brownian bridge. By universality of the limiting distribution and compacity of $[0, 1]$, the convergence in distribution holds without considering any subsequence.

5.1.4. Proof of Theorem 3.2

Thanks to Proposition 5.1 we have \mathbb{U} and \mathbb{V} two independent standard Brownian bridges such that (37) holds. Since $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\tau_{n,m}/n \rightarrow 1 - \sigma^2 \in [0, 1]$ by assumption, we obtain by Slutsky's Lemma

$$\left(\begin{array}{c} \sqrt{m} \left[\widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1} (1 - I) - F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1} (1 - I) \right] \\ \sqrt{n} \left[F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1} (1 - I) - F_{\text{test}} \circ F_{\text{cal}}^{-1} (1 - I) \right] \\ \frac{\tau_{n,m}}{m} \\ \frac{\tau_{n,m}}{n} \end{array} \right) \xrightarrow{\mathcal{L}} \left(\begin{array}{c} \mathbb{U} \circ G \\ G' \mathbb{V} \\ \sigma^2 \\ 1 - \sigma^2 \end{array} \right) \text{ on } [D(0, 1)]^2 \times [0, 1]^2.$$

The result follows by applying the continuous mapping theorem with the decomposition (36).

5.1.5. Proof of Theorem 3.3

For weighted conformal p -values, we introduce

$$\widehat{G}_m^{w,(n)}(t) := \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{p_i^{w,(n)} \leq t} = \text{FCP}_m^{w,(n)}(t), \quad t \in \mathbb{R}; \quad (39)$$

$$\widehat{F}_{\text{cal}}^{w,(n)}(t) := \frac{\sum_{k=1}^n w(S_k) \mathbb{1}_{S_k \leq t}}{w(+\infty) + \sum_{k=1}^n w(S_k)}, \quad t \in \mathbb{R}, \quad (40)$$

the counterparts of $\widehat{G}_m^{(n)}$ (33) and $\widehat{F}_{\text{cal}}^{(n)}$ (34) in the weighted case, respectively. Note that $p_i^{w,(n)} = 1 - \widehat{F}_{\text{cal}}^{w,(n)}(T_i)$ almost surely under Assumption 1. Hence, the following decomposition, analogue to (36), holds:

$$\begin{aligned} \sqrt{\tau_{n,m}} \left(\widehat{G}_m^{w,(n)}(t) - G^w(t) \right) &= - \sqrt{\frac{\tau_{n,m}}{m}} \sqrt{m} \left(\widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{w,(n)})^{-1}(1-t) - F_{\text{test}} \circ (F_{\text{cal}}^w)^{-1}(1-t) \right) \\ &\quad - \sqrt{\frac{\tau_{n,m}}{n}} \sqrt{n} \left(F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{w,(n)})^{-1}(1-t) - F_{\text{test}} \circ (F_{\text{cal}}^w)^{-1}(1-t) \right). \end{aligned} \quad (41)$$

Thus, the novelty of (41) with respect to (36) is only the presence of $(\widehat{F}_{\text{cal}}^{w,(n)})^{-1}$ instead of $(\widehat{F}_{\text{cal}}^{(n)})^{-1}$. By Assumption 4, we can show that the family of function $\mathcal{F} = \{w_t : x \in \mathbb{R} \mapsto w(x) \mathbb{1}_{x \leq t}; t \in \mathbb{R}\}$ is P_{cal} -Donsker and P_{cal} -Glivenko-Cantelli. Since \mathcal{F} is Glivenko-Cantelli, the convergence of the test term happens with the same argument in the unweighted case. Because this class is Donsker, using twice the functional delta method gives us that $(\sqrt{n}[(\widehat{F}_{\text{cal}}^{w,(n)})^{-1} - (F_{\text{cal}}^w)^{-1}])_n$ converges in distribution to some known random process on the set $D(0,1)$. This result is stated and then proved in Lemma C.1. This leads to the joint convergence Proposition C.2, which is the analogue of Proposition 5.1 in the weighted case. Theorem 3.3 is thus proved by applying the continuous mapping theorem to the decomposition (41).

5.2. Proofs for Section 4

5.2.1. FDP expression

We introduced, as in the works from Genovese and Wasserman (2004) and Neuvial (2008), $\widehat{G}_{m,0}^{(n)}$ and $\widehat{G}_{m,1}^{(n)}$ the two e.c.d.f.'s of conformal p -values for non-novelties and novelties, respectively:

$$\widehat{G}_{m,0}^{(n)}(t) := \frac{1}{m_0(m)} \sum_{i \in \llbracket m \rrbracket \cap \mathcal{H}_0} \mathbb{1}_{p_i^{(n)} \leq t}, \quad t \in (0,1), \quad (42)$$

$$\widehat{G}_{m,1}^{(n)}(t) := \frac{1}{m_1(m)} \sum_{i \in \llbracket m \rrbracket \cap \mathcal{H}_1} \mathbb{1}_{p_i^{(n)} \leq t}, \quad t \in (0,1). \quad (43)$$

We also introduce the mixture e.c.d.f. of the test sample

$$\widehat{G}_m^{(n)}(t) = \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{p_i^{(n)} \leq t} = \pi_0(m) \widehat{G}_{m,0}^{(n)}(t) + (1 - \pi_0(m)) \widehat{G}_{m,1}^{(n)}(t), \quad t \in (0,1). \quad (44)$$

For all $t \in (0,1)$, we denote simply $\text{FDP}_m^{(n)}(t) := \text{FDP}_m^{(n)}(\{i \in \llbracket m \rrbracket, p_i^{(n)} \leq t\})$ (resp. $\text{TDP}_m^{(n)}(t) := \text{TDP}_m^{(n)}(\{i \in \llbracket m \rrbracket, p_i^{(n)} \leq t\})$) the FDP (resp. TDP) of the procedure rejecting all the conformal p -values smaller than t , see (10) and (11). The following equalities hold:

$$\begin{aligned} \text{FDP}_m^{(n)}(t) &= \frac{\pi_0(m) \widehat{G}_{m,0}^{(n)}(t)}{\widehat{G}_m^{(n)}(t) \vee m^{-1}}, \quad t \in (0,1), \\ \text{TDP}_m^{(n)}(t) &= \widehat{G}_{m,1}^{(n)}(t), \quad t \in (0,1). \end{aligned} \quad (45)$$

Following [Neuviel \(2008\)](#), and since the novelty detection procedure from [Bates et al. \(2023\)](#) is the BH procedure applied to conformal p -values, we can be described it as a thresholding procedure with threshold $\mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{(n)})$, where the functional $\mathcal{T}^{\text{BH}\alpha}$ is defined by

$$\mathcal{T}^{\text{BH}\alpha}(F) := \sup \left\{ t \in [0, 1], F(t) \geq \frac{t}{\alpha} \right\}. \quad (46)$$

In other words, we have $\mathcal{R}_\alpha = \{i \in \llbracket m \rrbracket : p_i^{(n)} \leq \mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{(n)})\}$ and $\text{FDP}_m^{(n)}(\mathcal{R}_\alpha) = \text{FDP}_m^{(n)}(\mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}))$ with the notation above.

5.2.2. Joint convergence and application to the FDP and TDP

Following [Neuviel \(2008\)](#), $G \mapsto \mathcal{T}^{\text{BH}\alpha}(G)$ and $G \mapsto \text{FDP}_m^{(n)}(\mathcal{T}^{\text{BH}\alpha}(G))$ are Hadamard differentiable at G_{mixt} provided that G_{mixt} is concave and differentiable and that $\alpha > [G'_{\text{mixt}}(0^+)]^{-1}$. The following result complete the picture by studying the convergence of $\widehat{G}_m^{(n)}$.

Proposition 5.3. *Under Assumptions 2 and 3, assuming that $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0, 1)$, we have*

$$\sqrt{\tau_{n,m}} \begin{pmatrix} \widehat{G}_{m,0}^{(n)} - I \\ \widehat{G}_{m,1}^{(n)} - G \end{pmatrix} \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \begin{pmatrix} \frac{\sigma}{\sqrt{\pi_0}} \mathbb{U} + \sqrt{1-\sigma^2} \mathbb{W} \\ \frac{\sigma}{\sqrt{1-\pi_0}} \mathbb{V}_G + G' \sqrt{1-\sigma^2} \mathbb{W} \end{pmatrix} =: \begin{pmatrix} \mathbb{Z}_0 \\ \mathbb{Z}_1 \end{pmatrix} \text{ on } [D(0, 1)]^2,$$

with \mathbb{U} , \mathbb{V} and \mathbb{W} three independent Brownian bridges. As a result,

$$\begin{aligned} \sqrt{\tau_{n,m}} \left(\widehat{G}_m^{(n)} - G_{\text{mixt}} \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \sqrt{\pi_0 \sigma^2} \mathbb{U} + \sqrt{(1-\pi_0)\sigma^2} \mathbb{V}_G + G'_{\text{mixt}} \sqrt{1-\sigma^2} \mathbb{W} \\ &= \pi_0 \mathbb{Z}_0 + (1-\pi_0) \mathbb{Z}_1 =: \mathbb{Z} \text{ on } D(0, 1). \end{aligned}$$

Proposition 5.3 is proved in Section B.2. The proof is similar to Proposition 5.1, with the additional technicality that the decomposition (36) should be considered for the two processes $\widehat{G}_{m,0}^{(n)}$ and $\widehat{G}_{m,1}^{(n)}$. The fact that these decompositions are based on the same process $\widehat{F}_{\text{cal}}^{(n)}$ induces a dependence between the components that results in the term \mathbb{W} in the asymptotic variance.

Using Proposition 5.3 with the continuous functionals (45) gives us convergence in $D(0, 1)$ of the asymptotic processes $(\text{FDP}(t))_{t \in (0,1)}$ and $(\text{TDP}(t))_{t \in (0,1)}$ and is left to the readers. This can be compared with the results in [Genovese and Wasserman \(2004\)](#) where they use independent and uniform p -values under the null. Here we have to deal with the dependence giving us the \mathbb{W} term and the double asymptotic giving us the σ^2 term in Proposition 5.3. The same applies for the weighted case. However we are interested to understand the asymptotic when the threshold is data-driven, and specifically the BH threshold.

By using the functional delta method theorem with the BH functionals (45) and (46) described above (see [Neuviel, 2008](#), supplementary files from [Delattre and Roquain, 2016](#) and Lemma S.2.2 from [Kluger and Owen, 2024](#)), we obtain the following result.

Proposition 5.4. *Under Assumptions 2, 3 and 5, assume that BH is applied with a level $\alpha > [G'_{\text{mixt}}(0^+)]^{-1}$. If $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0, 1)$, we have the following convergences:*

$$\begin{aligned} \sqrt{\tau_{n,m}} \left(\mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}) - \mathcal{T}_\alpha \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \frac{1}{\frac{1}{\alpha} - G'_{\text{mixt}}(\mathcal{T}_\alpha)} \mathbb{Z}(\mathcal{T}_\alpha); \\ \sqrt{\tau_{n,m}} \left(\text{FDP}_m^{(n)} \left(\mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}) \right) - \pi_0 \alpha \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \frac{\pi_0}{G'_{\text{mixt}}(\mathcal{T}_\alpha)} \mathbb{Z}_0(\mathcal{T}_\alpha); \\ \sqrt{\tau_{n,m}} \left(\text{TDP}_m^{(n)} \left(\mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}) \right) - G(\mathcal{T}_\alpha) \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \frac{G'(\mathcal{T}_\alpha)}{\frac{1}{\alpha} - G'_{\text{mixt}}(\mathcal{T}_\alpha)} \mathbb{Z}(\mathcal{T}_\alpha) + \mathbb{Z}_1(\mathcal{T}_\alpha), \end{aligned}$$

where $\mathcal{T}_\alpha = \mathcal{T}^{\text{BH}\alpha}(G_{\text{mixt}})$ and \mathbb{Z}_0 , \mathbb{Z}_1 and \mathbb{Z} are the three processes defined in Proposition 5.3.

Proposition 5.4 is proved in Section B.3. As a sanity check, we see that Proposition 5.4 reduces to the result of [Neuviel \(2008\)](#) when n/m tends to infinity ($\sigma^2 = 1$).

5.2.3. Proof of Theorem 4.1

Theorem 4.1 is obtained from Proposition 5.4 by computing the different asymptotic covariance functions. For this, we use that $(\mathbb{Z}_0, \mathbb{Z}_1, \mathbb{Z})$ can be written as

$$\begin{pmatrix} \mathbb{Z}_0 \\ \mathbb{Z}_1 \\ \mathbb{Z} \end{pmatrix} = \begin{pmatrix} \frac{\sigma}{\sqrt{\pi_0}}\mathbb{U} + \sqrt{1 - \sigma^2}\mathbb{W} \\ \frac{\sigma}{\sqrt{1 - \pi_0}}\mathbb{V}_G + G'\sqrt{1 - \sigma^2}\mathbb{W} \\ \sqrt{\pi_0\sigma^2}\mathbb{U} + \sqrt{(1 - \pi_0)\sigma^2}\mathbb{V}_G + G'_{\text{mixt}}\sqrt{1 - \sigma^2}\mathbb{W} \end{pmatrix},$$

with \mathbb{U} , \mathbb{V} and \mathbb{W} being three independent standard Brownian bridges. The result follows from direct computations.

5.2.4. Proof of Theorem 4.2

Proving Theorem 4.2 is analogue to the proof of Theorem 4.1 above, but starting from weighted processes. For short, the full description of the proof is postponed to Section C.3. Therein, Propositions C.3 and C.4 are the analogues of Propositions 5.3 and 5.4, respectively.

6. Conclusion

In this paper we obtain the exact asymptotic distribution of FCP and FDP for conformal inference methods when both the sizes of the calibration sample and test sample grow simultaneously. Our theory covers both the prediction and novelty detection settings, including a potential distribution shift. Our results quantify exactly how the covariance process is affected by the dependence inherent to the conformal settings, that use the same calibration sample for all the test examples. First, we prove that the convergence rate $\sqrt{\tau_{n,m}}$ can be largely deteriorated when n/m vanishes to zero (that is, $\sigma = 0$). Otherwise, when n and m are of the same order, the convergence rate is the usual one ($m^{1/2}$), but the asymptotic covariance is affected. Nevertheless, when n/m tends to infinity, the convergence is strictly the same as in the usual independent case. Interestingly, our results can be used to calibrate easily and accurately quantiles for controlling an error amount when performing conformal inference with large n and m . We also quantify the effects of doing conformal inference while there is a distribution shift between the calibration sample and the test sample. We exhibit how this distribution shift acts on the asymptotic behaviour of the FCP, by changing the mean and the variance, and how the correction with weighted conformal p -values impacts the asymptotic variance.

While our work paves the way for studying asymptotic convergences in conformal inferences, it leaves some open directions. For instance, in case of a distribution shift, the oracle weight function is mostly unknown, and is often estimated (Jin and Candès, 2023). Finding the exact asymptotic distribution for the processes using estimated weights is a very interesting and challenging problem for future investigations. In a finite sample setting, using the asymptotic expectation and variance as a proxy to design weight and score functions can provide very compelling new ideas and can unlock new training procedures.

Acknowledgments

I am grateful to Étienne Roquain for his precious advices and reviews throughout this work. I wish to thank Gilles Blanchard and Sylvain Delattre for their comments, and Magalie Fromont for fruitful discussions. I acknowledge the Emergence project MARS of Sorbonne Université, and grant ANR-23-CE40-0018-01 (BACKUP) of the French National Research Agency ANR.

References

- ANGELOPOULOS, A. N. and BATES, S. (2021). A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv preprint arXiv:2107.07511*.
- BALASUBRAMANIAN, V., HO, S.-S. and VOVK, V. (2014). *Conformal prediction for reliable machine learning: theory, adaptations and applications*. Morgan Kaufmann books.
- BARBER, R. F., CANDÈS, E. J., RAMDAS, A. and TIBSHIRANI, R. J. (2023). Conformal prediction beyond exchangeability. *The Annals of Statistics* **51** 816–845.
- BATES, S., CANDÈS, E., LEI, L., ROMANO, Y. and SESIA, M. (2023). Testing for outliers with conformal p-values. *Ann. Statist.* **51** 149–178.
- BEHBOODI, A., CORREIA, A. H., MASSOLI, F. V. and LOUIZOS, C. (2025). Fundamental bounds on efficiency-confidence trade-off for transductive conformal prediction. *arXiv preprint arXiv:2509.04631*.
- BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* **57** 289–300. [MR1325392 \(96d:62143\)](#)
- BIAN, M. and BARBER, R. F. (2022). Training-conditional coverage for distribution-free predictive inference. *arXiv preprint arXiv:2205.03647*.
- BILLINGSLEY, P. (1999). *Convergence of probability measures*, second ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. John Wiley & Sons Inc., New York. A Wiley-Interscience Publication. [MR1700749 \(2000e:60008\)](#)
- BLANCHARD, G., DURAND, G., MARANDON-CARLHIAN, A. and PÉRIER, R. (2024). FDR control and FDP bounds for conformal link prediction. *arXiv preprint arXiv:2404.02542*.
- CHI, Z. (2007). On the performance of FDR control: constraints and a partial solution. *Ann. Statist.* **35** 1409–1431. [MR2351091](#)
- DELATTRE, S. and ROQUAIN, E. (2011). On the false discovery proportion convergence under Gaussian equi-correlation. *Statist. Probab. Lett.* **81** 111–115. [MR2740072](#)
- DELATTRE, S. and ROQUAIN, E. (2016). On empirical distribution function of high-dimensional Gaussian vector components with an application to multiple testing. *Bernoulli* **22** 302–324. [MR3449784](#)
- DONSKER, M. D. (1952). Justification and extension of Doob’s heuristic approach to the Komogorov-Smirnov theorems. *Ann. Math. Statistics* **23** 277–281. [MR0047288 \(13,853n\)](#)
- FROMONT, M., LERASLE, M. and REYNAUD-BOURET, P. (2016). Family-wise separation rates for multiple testing. *The Annals of Statistics* **44** 2533–2563.
- GAZIN, U., BLANCHARD, G. and ROQUAIN, E. (2024). Transductive conformal inference with adaptive scores. In *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics* (S. DASGUPTA, S. MANDT and Y. LI, eds.). *Proceedings of Machine Learning Research* **238** 1504–1512. PMLR.
- GENOVESE, C. and WASSERMAN, L. (2002). Operating characteristics and extensions of the false discovery rate procedure. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64** 499–517. [MR1924303 \(2003h:62027\)](#)
- GENOVESE, C. and WASSERMAN, L. (2004). A stochastic process approach to false discovery control. *Ann. Statist.* **32** 1035–1061. [MR2065197 \(2005k:62149\)](#)
- HUANG, K., JIN, Y., CANDÈS, E. and LESKOVEC, J. (2024). Uncertainty quantification over graph with conformalized graph neural networks. *Advances in Neural Information Processing Systems* **36**.
- JIN, Y. and CANDÈS, E. J. (2023). Model-free selective inference under covariate shift via weighted conformal p-values. *arXiv preprint arXiv:2307.09291*.
- KLUGER, D. M. and OWEN, A. B. (2024). A central limit theorem for the Benjamini-Hochberg false discovery proportion under a factor model. *Bernoulli* **30** 743–769.
- LIANG, Z., SESIA, M. and SUN, W. (2024). Integrative conformal p-values for out-of-distribution testing with labelled outliers. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **86** 671–693.

- MARANDON, A., LEI, L., MARY, D. and ROQUAIN, E. (2024). Adaptive novelty detection with false discovery rate guarantee. *The Annals of Statistics* **52** 157 – 183.
- MARQUES F. , P. C. (2025). Universal distribution of the empirical coverage in split conformal prediction. *Statistics & Probability Letters* **219** 110350.
- MARY, D. and ROQUAIN, E. (2022). Semi-supervised multiple testing. *Electronic Journal of Statistics* **16** 4926 – 4981.
- NEUVIAL, P. (2008). Asymptotic properties of false discovery rate controlling procedures under independence. *Electron. J. Stat.* **2** 1065–1110. [MR2460858 \(2010c:62154\)](#)
- NGUYEN, D. T., PATHAK, R., ANGELOPOULOS, A. N., BATES, S. and JORDAN, M. I. (2024). Data-Adaptive Tradeoffs among Multiple Risks in Distribution-Free Prediction. *arXiv preprint arXiv:2403.19605*.
- PAPADOPOULOS, H., PROEDROU, K., VOVK, V. and GAMMERMAN, A. (2002). Inductive confidence machines for regression. In *13th European Conference on Machine Learning (ECML 2002)* 345–356. Springer.
- ROMANO, J. P. and WOLF, M. (2005). Exact and approximate stepdown methods for multiple hypothesis testing. *J. Amer. Statist. Assoc.* **100** 94–108. [MR2156821](#)
- SARKAR, S. and KUCHIBHOTLA, A. K. (2023). Post-selection Inference for Conformal Prediction: Trading off Coverage for Precision. *arXiv preprint arXiv:2304.06158*.
- SAUNDERS, C., GAMMERMAN, A. and VOVK, V. (1999). Transduction with confidence and credibility. In *16th International Joint Conference on Artificial Intelligence (IJCAI 1999)* 722–726.
- SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical processes with applications to statistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics.* John Wiley & Sons Inc., New York. [MR838963 \(88e:60002\)](#)
- STOREY, J. D. (2002). A direct approach to false discovery rates. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64** 479–498. [MR1924302 \(2003h:62029\)](#)
- STOREY, J. D., TAYLOR, J. E. and SIEGMUND, D. (2004). Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **66** 187–205. [MR2035766 \(2004k:62056\)](#)
- TIBSHIRANI, R. J., FOYGE BARBER, R., CANDÈS, E. and RAMDAS, A. (2019). Conformal prediction under covariate shift. *Advances in neural information processing systems* **32**.
- VAN DER VAART, A. W. (1998). *Asymptotic statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge University Press, Cambridge. [MR1652247 \(2000c:62003\)](#)
- VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak convergence and empirical processes. Springer Series in Statistics.* Springer-Verlag, New York. With applications to statistics. [MR1385671 \(97g:60035\)](#)
- VOVK, V. (2012). Conditional validity of inductive conformal predictors. In *4th Asian conference on machine learning (ACML 2012)* 475–490. PMLR.
- VOVK, V. (2013). Transductive conformal predictors. In *Artificial Intelligence Applications and Innovations: 9th IFIP WG 12.5 International Conference (AIAI 2013)* 348–360. Springer.
- VOVK, V., GAMMERMAN, A. and SHAFER, G. (2005). *Algorithmic learning in a random world.* Springer.

Appendix A: Standardisation lemma

In this section, we introduce the standardisation lemma, which will be extensively used in our proofs. Let us introduce the notation

$$F_{\text{test,cal}}(t) = F_{\text{test}} \circ F_{\text{cal}}^{-1}(t), \quad t \in [0, 1]; \quad (47)$$

$$F_{0,\text{cal}}(t) = F_0 \circ F_{\text{cal}}^{-1}(t), \quad t \in [0, 1], \quad (48)$$

where by convention $F_{\text{cal}}^{-1}(0)$ denotes the infimum of the support of the distribution given by F_{cal} and $F_{\text{test,cal}}(1) = F_{0,\text{cal}}(1) = 1$. The following lemma holds.

Lemma A.1. Consider either the prediction setting with Assumption 1 (with parameters $F_{\text{cal}}, F_{\text{test}}$) or the novelty detection setting with Assumption 2 (with parameters $F_{\text{cal}}, F_{\text{test}}, F_0$). If $F_{\text{cal}}, F_{\text{test}}$ satisfy Assumption 3, for any weight function $w : \mathbb{R} \cup \{\infty\} \mapsto \mathbb{R}^+$ with $w(+\infty) > 0$, the distribution of the w -weighted conformal p -value family under the parameters $F_{\text{cal}}, F_{\text{test}}$ (and F_0) is the same as the $w \circ F_{\text{cal}}^{-1}$ -weighted conformal p -value family under the parameters $F_{\text{cal}} = I, F_{\text{test}} = F_{\text{test,cal}}$ (and $F_0 = F_{0,\text{cal}}$).

Proof of Lemma A.1. Since F_{cal} is continuous increasing on its support, we can write almost surely

$$\begin{aligned} \sum_{k \in \llbracket n \rrbracket} w(S_k) \mathbb{1}_{S_k \geq T_i} &= \sum_{k \in \llbracket n \rrbracket} (w \circ F_{\text{cal}}^{-1})(F_{\text{cal}}(S_k)) \mathbb{1}_{F_{\text{cal}}(S_k) \geq F_{\text{cal}}(T_i)} \\ &= \sum_{k \in \llbracket n \rrbracket} (w \circ F_{\text{cal}}^{-1})(S'_k) \mathbb{1}_{S'_k \geq T'_i}, \end{aligned}$$

with $S'_k = F_{\text{cal}}(S_k)$ which are i.i.d. uniform and $T'_i = F_{\text{cal}}(T_i)$ which are i.i.d. $\sim F_{\text{test,cal}}$ (or either $\sim F_{0,\text{cal}}$ under the null or $\sim F_{\text{test,cal}}$ under the alternative, in the novelty detection setting). \square

Appendix B: Proofs of auxiliary results

In this section, we prove Proposition 5.1, Propositions 5.3 and 5.4. Proofs for the weighted case are given in Section C.

B.1. Proof of Proposition 5.1

First, the Donsker theorem (Theorem E.3) provides

$$\sqrt{m} \left(\widehat{F}_{m,\text{test}} - F_{\text{test}} \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathbb{U} \circ F_{\text{test}} \text{ on } D(\mathbb{R}),$$

with \mathbb{U} being a standard Brownian bridge. Moreover, by the Glivenko-Cantelli theorem, we have that $(\widehat{F}_{\text{cal}}^{(n)})_n$ converges in probability on $\ell^\infty(\mathbb{R})$ to F_{cal} . Since by Assumption 3 the inverse map is continuous at F_{cal} we obtain that $((\widehat{F}_{\text{cal}}^{(n)})^{-1})_n$ converges in probability on $D(0, 1)$ to F_{cal}^{-1} .

Second, applying again the Donsker theorem (Theorem E.3), we have

$$\sqrt{n} \left(\widehat{F}_{\text{cal}}^{(n)} - F_{\text{cal}} \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathbb{V} \circ F_{\text{cal}} \text{ on } D(\mathbb{R}),$$

with \mathbb{V} a standard Brownian bridge independent of \mathbb{U} . Now, by using the functional delta method with the inverse map (see Lemma E.6) by Assumption 3, we obtain,

$$\sqrt{n} \left((\widehat{F}_{\text{cal}}^{(n)})^{-1} - F_{\text{cal}}^{-1} \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} (F_{\text{cal}}^{-1})' \mathbb{V} \text{ on } D(0, 1), \quad (49)$$

where $(F_{\text{cal}}^{-1})'$ denotes the derivative of F_{cal}^{-1} . Again, we use the Hadamard differentiability of the map $\varphi \mapsto F_{\text{test}} \circ \varphi$ which is true by Assumption 3 (see Lemma E.5) to obtain

$$\sqrt{n} \left(F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1} - F_{\text{test}} \circ F_{\text{cal}}^{-1} \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} F'_{\text{test}} \circ F_{\text{cal}}^{-1} (F_{\text{cal}}^{-1})' \mathbb{V} = (F_{\text{test}} \circ F_{\text{cal}}^{-1})' \mathbb{V} \text{ on } D(0, 1).$$

Since for all $t \in (0, 1)$, $G'(t) = (F_{\text{test}} \circ F_{\text{cal}}^{-1})'(1-t)$ and since \mathbb{V} is a standard Brownian bridge if and only if $(\mathbb{V}_t)_t = (\mathbb{V}_{1-t})_t$ is a standard Brownian bridge we obtain the second term in (37).

Finally, since \mathcal{D}_{cal} and $\mathcal{D}_{\text{test}}$ are independent, we obtain the joint convergence:

$$\left(\begin{array}{c} \sqrt{m} \left[\widehat{F}_{m,\text{test}} - F_{\text{test}} \right] \\ \sqrt{n} \left[F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1} (1-I) - F_{\text{test}} \circ F_{\text{cal}}^{-1} (1-I) \right] \end{array} \right) \xrightarrow{\mathcal{L}} \left(\begin{array}{c} \mathbb{U} \circ F_{\text{test}} \\ G' \mathbb{V} \end{array} \right) \text{ on } D(\mathbb{R}) \times D(0, 1),$$

with \mathbb{U} and \mathbb{V} two independent standard Brownian bridges. By Slutsky's Lemma, we obtain the following joint convergence

$$\begin{pmatrix} \sqrt{m} [\widehat{F}_{m,\text{test}} - F_{\text{test}}] \\ \sqrt{n} [F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-I)] \\ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} \mathbb{U} \circ F_{\text{test}} \\ G' \mathbb{V} \\ F_{\text{cal}}^{-1}(1-I) \end{pmatrix}$$

on $D(\mathbb{R}) \times D(0,1) \times \ell^\infty(0,1)$.

Then, using Lemma E.10, we finally obtain

$$\begin{pmatrix} \sqrt{m} [\widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I)] \\ \sqrt{n} [F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-I)] \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} \mathbb{U}(F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-I)) \\ G' \mathbb{V} \end{pmatrix}$$

on $[D(0,1)]^2$.

Since $\mathbb{U}(F_{\text{test}} \circ F_{\text{cal}}^{-1}(1-I))$ has the same distribution as $\mathbb{U}(G)$, we obtain (37).

B.2. Proof of Proposition 5.3

By using Lemma A.1, one can assume without loss of generality that $P_0 = P_{\text{cal}} = \mathcal{U}(0,1)$ and P_{test} has for c.d.f. $F_{\text{test,cal}}$. By applying the Donsker theorem (Theorem E.3) with the independent families $(S_k, k \geq 1)$, $(T_i, i \in \mathcal{H}_0)$ and $(T_i, i \in \mathcal{H}_1)$ and following the same reasoning as in the proof of Proposition 5.1, there exist \mathbb{U} , \mathbb{V} and \mathbb{W} three independent standard Brownian bridges such that

$$\begin{pmatrix} \sqrt{m_0(m)} [\widehat{F}_{m,0,\text{test}} - I] \\ \sqrt{m_1(m)} [\widehat{F}_{m,1,\text{test}} - F_{\text{test,cal}}] \\ \sqrt{n} [(\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - (1-I)] \\ \sqrt{n} [F_{\text{test,cal}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test,cal}}(1-I)] \\ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} \mathbb{U} \\ \mathbb{V}(1 - F_{\text{test,cal}}) \\ \mathbb{W} \\ G' \mathbb{W} \\ 1 - I \end{pmatrix} \text{ on } [D(0,1)]^4 \times \ell^\infty(0,1),$$

where we denoted $\widehat{F}_{m,r,\text{test}}$ the e.c.d.f. of $\{T_i, i \in [m] \cap \mathcal{H}_r\}$ for $r \in \{0,1\}$. Now, by using Lemma E.10 (or more precisely, an obvious extension of it for 5 joint processes), we obtain

$$\begin{pmatrix} \sqrt{m_0(m)} [\widehat{F}_{m,0,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I)] \\ \sqrt{m_1(m)} [\widehat{F}_{m,1,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test,cal}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I)] \\ \sqrt{n} [(\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - (1-I)] \\ \sqrt{n} [F_{\text{test,cal}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test,cal}}(1-I)] \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} \mathbb{U} \\ \mathbb{V} \circ G \\ \mathbb{W} \\ G' \mathbb{W} \end{pmatrix} \text{ on } [D(0,1)]^4.$$

Now, $\widehat{G}_{m,0}^{(n)}$ (42) and $\widehat{G}_{m,1}^{(n)}$ (43) satisfy the following decomposition (obtained similarly to (36)):

$$\begin{aligned} \sqrt{\tau_{n,m}} [\widehat{G}_{m,0}^{(n)} - I] &= -\sqrt{\frac{\tau_{n,m}}{m}} \times \frac{m}{m_0} \sqrt{m_0} [\widehat{F}_{m,0,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I)] \\ &\quad - \sqrt{\frac{\tau_{n,m}}{n}} \sqrt{n} [(\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - (1-I)], \\ \sqrt{\tau_{n,m}} [\widehat{G}_{m,1}^{(n)} - G] &= -\sqrt{\frac{\tau_{n,m}}{m}} \times \frac{m}{m_1} \sqrt{m_1} [\widehat{F}_{m,1,\text{test}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test,cal}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I)] \\ &\quad - \sqrt{\frac{\tau_{n,m}}{n}} \sqrt{n} [F_{\text{test,cal}} \circ (\widehat{F}_{\text{cal}}^{(n)})^{-1}(1-I) - F_{\text{test,cal}}(1-I)]. \end{aligned}$$

We conclude by using $m_0/m \rightarrow \pi_0 \in (0, 1)$, $m_1/m \rightarrow 1 - \pi_0 \in (0, 1)$, $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and Slutsky's lemma.

B.3. Proof of Proposition 5.4

By Assumption 5, and since $\alpha > [G'_{\text{mixt}}(0^+)]^{-1}$, the asymptotic threshold $\mathcal{T}_\alpha = \mathcal{T}^{\text{BH}\alpha}(G_{\text{mixt}})$ is well defined and belongs to $(0, 1)$. Now observe that there exists a compact interval $[a, b] \subset (0, 1)$ such that

$$\mathcal{T}_\alpha = \sup \left\{ t \in [a, b], G_{\text{mixt}}(t) \geq \frac{t}{\alpha} \right\}, \quad (50)$$

To see this, note that since $G_{\text{mixt}} \leq 1$ we have that $\mathcal{T}_\alpha \leq \alpha$, hence \mathcal{T}_α is smaller than any $b \in (\alpha, 1)$. Now, since $\mathcal{T}_\alpha > 0$, any $a \in (0, \mathcal{T}_\alpha)$ leads to $\mathcal{T}_\alpha \in (a, b)$, which leads to (50).

Now consider the functional

$$\mathcal{T}_{[a,b]}^{\text{BH}\alpha}(F) := \sup \left\{ t \in [a, b], F(t) \geq \frac{t}{\alpha} \right\}, \quad (51)$$

which is similar to (46) but with a restricted range on t . Equation (50) hence reads as $\mathcal{T}_\alpha = \mathcal{T}_{[a,b]}^{\text{BH}\alpha}(G_{\text{mixt}})$. In addition, we have

$$\left\{ \left| \widehat{G}_m^{(n)}(a) - G_{\text{mixt}}(a) \right| \leq \eta \right\} \subset \left\{ \mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}) = \mathcal{T}_{[a,b]}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}) \right\}, \quad (52)$$

where $\eta = (G_{\text{mixt}}(a) - a\alpha^{-1})/2$, which is positive by the choice of a . Indeed, if $|\widehat{G}_m^{(n)}(a) - G_{\text{mixt}}(a)| \leq \eta$, we have $\widehat{G}_m^{(n)}(a) - a\alpha^{-1} = \widehat{G}_m^{(n)}(a) - G_{\text{mixt}}(a) + G_{\text{mixt}}(a) - a\alpha^{-1} \geq \eta$, which implies that $\left\{ \left| \widehat{G}_m^{(n)}(a) - G_{\text{mixt}}(a) \right| \leq \eta \right\} \subset \left\{ \mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}) > a \right\}$. This proves (52).

Now, $\mathcal{T}_{[a,b]}^{\text{BH}\alpha}$ is Hadamard differentiable at G_{mixt} , tangentially to $\mathcal{C}[a, b]$ with a derivative coinciding with the one of $\mathcal{T}^{\text{BH}\alpha}$, that is,

$$(\dot{\mathcal{T}}_{[a,b]}^{\text{BH}\alpha})_{G_{\text{mixt}}} = (\dot{\mathcal{T}}^{\text{BH}\alpha})_{G_{\text{mixt}}}.$$

By the convergence of Proposition 5.3, which holds on $D[a, b]$, we can apply the functional delta method (see Lemma E.9 for the exact expression of derivatives) to obtain

$$\begin{aligned} \sqrt{\tau_{n,m}} \left(\mathcal{T}_{[a,b]}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}) - \mathcal{T}_\alpha \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \frac{1}{\frac{1}{\alpha} - G'_{\text{mixt}}(\mathcal{T}_\alpha)} \mathbb{Z}(\mathcal{T}_\alpha); \\ \sqrt{\tau_{n,m}} \left(\text{FDP}_m^{(n)} \left(\mathcal{T}_{[a,b]}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}) \right) - \pi_0 \alpha \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \frac{\pi_0}{G_{\text{mixt}}(\mathcal{T}_\alpha)} \mathbb{Z}_0(\mathcal{T}_\alpha); \\ \sqrt{\tau_{n,m}} \left(\text{TDP}_m^{(n)} \left(\mathcal{T}_{[a,b]}^{\text{BH}\alpha}(\widehat{G}_m^{(n)}) \right) - G(\mathcal{T}_\alpha) \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \frac{G'(\mathcal{T}_\alpha)}{\frac{1}{\alpha} - G'_{\text{mixt}}(\mathcal{T}_\alpha)} \mathbb{Z}(\mathcal{T}_\alpha) + \mathbb{Z}_1(\mathcal{T}_\alpha), \end{aligned}$$

where \mathbb{Z}_0 , \mathbb{Z}_1 and \mathbb{Z} are the three processes defined in Proposition 5.3. Now, the same convergences hold for $\mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{(n)})$ by using (52) because $\widehat{G}_m^{(n)}(a)$ converges in probability to $G_{\text{mixt}}(a)$. This concludes the proof.

Appendix C: Proofs for the weighted case

In this section, we state and prove Lemma C.1, Propositions C.2, C.3 and C.4.

C.1. Weighted versions of the Glivenko-Cantelli and Donsker theorem.

Recall that $\widehat{F}_{\text{cal}}^{w,(n)}(t)$ is given by (40). The next result applies both in the prediction and novelty detection settings and will be used to prove Propositions C.2, C.3 and C.4.

Lemma C.1. *Assume $P_{\text{cal}} = \mathcal{U}(0, 1)$. Let w a weight function satisfying Assumption 4. Then it holds*

$$(\widehat{F}_{\text{cal}}^{w, (n)})^{-1} \xrightarrow{\mathbb{P}} (F_{\text{cal}}^w)^{-1} \text{ on } \ell^\infty(0, 1),$$

and

$$\sqrt{n} \left((\widehat{F}_{\text{cal}}^{w, (n)})^{-1} - (F_{\text{cal}}^w)^{-1} \right) \xrightarrow{\mathcal{L}} \left((F_{\text{cal}}^w)^{-1} \right)' \rho^w \left(\mathbb{V} \circ V_{\text{cal}}^w \circ (F_{\text{cal}}^w)^{-1} + [I - V_{\text{cal}}^w \circ (F_{\text{cal}}^w)^{-1}] N \right) \\ \text{on } D(0, 1),$$

where \mathbb{V} is a standard Brownian bridge and N an independent standard Gaussian random variable.

Note that in the case $w \equiv 1$ (unweighted case), we recover the convergence presented in Corollary E.7, because $F_{\text{cal}}^w = V_{\text{cal}}^w = I$.

Proof of Lemma C.1. Denote for $n \geq 1$ and $t \in [0, 1]$,

$$K^{(n)}(t) := \frac{1}{n+1} \sum_{k=1}^n w(S_k) \mathbb{1}_{S_k \leq t}.$$

Since w is uniformly bounded, the family $\mathcal{F} = \{w_t : x \in \mathbb{R} \mapsto w(x) \mathbb{1}_{x \leq t}; t \in \mathbb{R}\}$ is $\mathcal{U}(0, 1)$ -Glivenko-Cantelli, and $(K^{(n)})_n$ converges uniformly on $[0, 1]$ in probability to the function K given by $K(t) := \int_0^t w(x) dx$ (Shorack and Wellner, 1986). Then, by continuity at K of the map of Lemma E.8, we get that $(K^{(n)}/K^{(n)}(1))_n$ converges uniformly (in probability) to F_{cal}^w . Since $\|\widehat{F}_{\text{cal}}^{w, (n)} - K^{(n)}/K^{(n)}(1)\|_\infty \leq w(+\infty)/((n+1)K^{(n)}(1) + w(+\infty))$ tends uniformly to 0 a.s., $(\widehat{F}_{\text{cal}}^{w, (n)})_n$ converges uniformly (in probability) to F_{cal}^w . By continuity of the inverse map at F_{cal}^w (see Lemma E.6), we obtain that $((\widehat{F}_{\text{cal}}^{w, (n)})^{-1})_n$ converges in probability to $(F_{\text{cal}}^w)^{-1}$ on $\ell^\infty(0, 1)$. This proves the first statement.

Next, we turn to prove the second statement. Since w is uniformly bounded, the family $\mathcal{F} = \{w_t : x \in \mathbb{R} \mapsto w(x) \mathbb{1}_{x \leq t}; t \in \mathbb{R}\}$ is $\mathcal{U}(0, 1)$ -Donsker (Shorack and Wellner, 1986). Hence there exist $\mathbb{K} = (\mathbb{K}(t))_{t \in [0, 1]}$ a Gaussian process such that

$$\sqrt{n} \left(K^{(n)} - K \right) \xrightarrow{\mathcal{L}} \mathbb{K} \text{ in } \ell^\infty[0, 1],$$

where the distribution of \mathbb{K} is given by $\mathbb{E}(\mathbb{K}) = 0$ and for $(s, t) \in [0, 1]^2$,

$$\text{Cov}(\mathbb{K}(s), \mathbb{K}(t)) = \int_0^{s \wedge t} w^2(x) dx - \int_0^s w(x) dx \int_0^t w(x) dx; \\ = K(1)^2 \left(\rho^{w^2} V_{\text{cal}}^w(s \wedge t) - F_{\text{cal}}^w(t) F_{\text{cal}}^w(s) \right),$$

where we used $K(1)^2 \rho^{w^2} V_{\text{cal}}^w(t) = \int_0^t w^2(x) dx$ and $F_{\text{cal}}^w = K/K(1)$ by the definition of V_{cal}^w (19), ρ^w (17) and F_{cal}^w (15). We easily check that the condition of the Kolmogorov-Čentsov theorem is satisfied, so that \mathbb{K} has a continuous version. Now, by applying the functional delta method with the map of Lemma E.8, we obtain

$$\sqrt{n} \left(\left(\frac{K^{(n)}}{K^{(n)}(1)} \right) - F_{\text{cal}}^w \right) \xrightarrow{\mathcal{L}} \frac{\mathbb{K}}{K(1)} - \frac{\mathbb{K}(1)}{K(1)} F_{\text{cal}}^w, \text{ in } \mathcal{C}[0, 1].$$

Since $\sqrt{n} \|\widehat{F}_{\text{cal}}^{w, (n)} - K^{(n)}/K^{(n)}(1)\|_\infty \leq \sqrt{nw} w(+\infty)/((n+1)K^{(n)}(1) + w(+\infty)) \rightarrow 0$ a.s., then by Slutsky's lemma, we obtain

$$\sqrt{n} \left(\widehat{F}_{\text{cal}}^{w, (n)} - F_{\text{cal}}^w \right) \xrightarrow{\mathcal{L}} \frac{\mathbb{K}}{K(1)} - \frac{\mathbb{K}(1)}{K(1)} F_{\text{cal}}^w, \text{ in } \mathcal{C}[0, 1].$$

And by applying the functional delta method with the inverse map at F_{cal}^w (see Lemma E.6) we obtain,

$$\sqrt{n} \left(\widehat{F}_{\text{cal}}^{w,(n)-1} - (F_{\text{cal}}^w)^{-1} \right) \xrightarrow{\mathcal{L}} ((F_{\text{cal}}^w)^{-1})' \left(-\frac{\mathbb{K} \circ (F_{\text{cal}}^w)^{-1}}{K(1)} + \frac{\mathbb{K}(1)}{K(1)} I \right), \text{ in } D(0, 1).$$

To conclude, we identify the covariance of the Gaussian limiting process. For, $1 > t \geq s > 0$,

$$\begin{aligned} & \text{Cov} \left(-\mathbb{K}((F_{\text{cal}}^w)^{-1}(t))/K(1) + \mathbb{K}(1)/K(1)t, -\mathbb{K}((F_{\text{cal}}^w)^{-1}(s))/K(1) + \mathbb{K}(1)/K(1)s \right) \\ &= \rho^{w^2} V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) - F_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) F_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) \\ & \quad - s \rho^{w^2} V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) + s F_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) \\ & \quad - t \rho^{w^2} V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) + t F_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) \\ & \quad + ts \rho^{w^2} - ts \\ &= \rho^{w^2} V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) - ts - s \rho^{w^2} V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) + st \\ & \quad - t \rho^{w^2} V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) + st + ts \rho^{w^2} - ts \\ &= \rho^{w^2} (V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) - s V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) - t V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) + ts). \end{aligned}$$

Therefore, the last display is equal to

$$\begin{aligned} & \rho^{w^2} \left(V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) - V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) \right) \\ & \quad + \left[t - V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) \right] \left[s - V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) \right] \\ &= \rho^{w^2} \left(V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(s)) \left[1 - V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) \right] + \left[t - V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) \right] \left[s - V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) \right] \right). \end{aligned}$$

The first term coincides with the covariance term of $\mathbb{V} \circ V_{\text{cal}}^w \circ (F_{\text{cal}}^w)^{-1}$ for a standard Brownian bridge \mathbb{V} , while the second term is the covariance of the process $\left(\left[t - V_{\text{cal}}^w((F_{\text{cal}}^w)^{-1}(t)) \right] N \right)_{t \in (0,1)}$ with $N \sim \mathcal{N}(0, 1)$. This concludes the proof for the second statement. \square

C.2. Prediction setting

The following result is the weighted version of Proposition 5.1.

Proposition C.2. *In the prediction setting with Assumptions 1, 3 and 4 and assuming that $n/(n+m)$ tends to $\sigma^2 \in [0, 1]$, we have,*

$$\begin{aligned} & \left(\begin{array}{c} \sqrt{m} \left[\widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{w,(n)})^{-1}(1-I) - F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{w,(n)})^{-1}(1-I) \right] \\ \sqrt{n} \left[F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{w,(n)})^{-1}(1-I) - F_{\text{test}} \circ (F_{\text{cal}}^w)^{-1}(1-I) \right] \end{array} \right) \xrightarrow{\mathcal{L}} \left(\begin{array}{c} \mathbb{U}(G^w) \\ \rho^w (G^w)' [\mathbb{V} \circ I^w + [I - I^w] N] \end{array} \right) \\ & \hspace{20em} \text{on } [D(0, 1)]^2, \end{aligned}$$

where \mathbb{U} and \mathbb{V} are two independent standard Brownian bridges and N is an independent standard Gaussian random variable.

Proof. By the Donsker theorem (Theorem E.3), we have

$$\sqrt{m} \left(\widehat{F}_{m,\text{test}} - F_{\text{test,cal}} \right) \xrightarrow{\mathcal{L}} \mathbb{U}(1 - F_{\text{test,cal}}) \text{ on } D(0, 1),$$

where \mathbb{U} is a standard Brownian bridge. By Lemma C.1 (which applies because we can standardize the calibration set, see Lemma A.1), we have

$$\begin{aligned} & \sqrt{n} \left((\widehat{F}_{\text{cal}}^{w,(n)})^{-1} - (F_{\text{cal}}^w)^{-1} \right) \xrightarrow{\mathcal{L}} ((F_{\text{cal}}^w)^{-1})' \rho^w \left(\mathbb{U}(V_{\text{cal}}^w \circ (F_{\text{cal}}^w)^{-1}) + [I - V_{\text{cal}}^w \circ (F_{\text{cal}}^w)^{-1}] N \right) \\ & \hspace{20em} \text{on } D(0, 1), \end{aligned}$$

where \mathbb{V} is a standard Brownian bridge and N is an independent standard Gaussian random variable. Thus, by independence between \mathcal{D}_{cal} and $\mathcal{D}_{\text{test}}$,

$$\left(\begin{array}{c} \sqrt{m} [\widehat{F}_{m,\text{test}} - F_{\text{test,cal}}] \\ \sqrt{n} [F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{w,(n)})^{-1}(1-I) - F_{\text{test}} \circ (F_{\text{cal}}^w)^{-1}(1-I)] \end{array} \right) \xrightarrow{\mathcal{L}} \left(\begin{array}{c} \mathbb{U}(1 - F_{\text{test,cal}}) \\ \rho^w(G^w)'[\mathbb{V} \circ I^w + [I - I^w]N] \end{array} \right) \\ \text{on } [D(0, 1)]^2,$$

where \mathbb{U} and \mathbb{V} two independent Brownian bridges and N an independent Gaussian r.v.. By Lemma C.1, we have that $(\widehat{F}_{\text{cal}}^{w,(n)})^{-1} \xrightarrow{\mathbb{P}} (F_{\text{cal}}^w)^{-1}$ on $\ell^\infty(0, 1)$ hence by using Lemma E.10 we get

$$\left(\begin{array}{c} \sqrt{m} [\widehat{F}_{m,\text{test}} \circ (\widehat{F}_{\text{cal}}^{w,(n)})^{-1}(1-I) - F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{w,(n)})^{-1}(1-I)] \\ \sqrt{n} [F_{\text{test}} \circ (\widehat{F}_{\text{cal}}^{w,(n)})^{-1}(1-I) - F_{\text{test}} \circ (F_{\text{cal}}^w)^{-1}(1-I)] \end{array} \right) \xrightarrow{\mathcal{L}} \left(\begin{array}{c} \mathbb{U} \circ (1 - F_{\text{test,cal}} \circ (F_{\text{cal}}^w)^{-1}(1-I)) \\ \rho^w(G^w)'[\mathbb{V} \circ I^w + [I - I^w]N] \end{array} \right) \\ \text{on } [D(0, 1)]^2,$$

which concludes the proof. \square

C.3. Novelty detection setting

We define

$$\widehat{G}_{m,0}^{w,(n)}(t) = \frac{1}{m_0(m)} \sum_{i \in \llbracket m \rrbracket \cap \mathcal{H}_0} \mathbb{1}_{p_i^{w,(n)} \leq t}, \quad t \in (0, 1);$$

$$\widehat{G}_{m,1}^{w,(n)}(t) = \frac{1}{m - m_0(m)} \sum_{i \in \llbracket m \rrbracket \cap \mathcal{H}_0^c} \mathbb{1}_{p_i^{w,(n)} \leq t}, \quad t \in (0, 1),$$

the counterparts of $\widehat{G}_{m,0}^{(n)}$ (42) and $\widehat{G}_{m,1}^{(n)}$ (43), respectively. Hence, the mixture e.c.d.f.

$$\widehat{G}_m^{w,(n)}(t) = \frac{1}{m} \sum_{i \in \llbracket m \rrbracket} \mathbb{1}_{p_i^{w,(n)} \leq t} = \pi_0(m) \widehat{G}_{m,0}^{w,(n)}(t) + (1 - \pi_0(m)) \widehat{G}_{m,1}^{w,(n)}(t), \quad t \in (0, 1),$$

is the counterpart of $\widehat{G}_m^{(n)}$ (44).

The following result is the weighted version of Proposition 5.3 (with an oracle weight function).

Proposition C.3. *In the novelty detection setting with Assumption 2, assume that P_0 is absolutely continuous with respect to P_{cal} and that the oracle weight function w^* (12) satisfies Assumption 4. Under Assumption 3, assuming that $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0, 1)$, we have*

$$\sqrt{\tau_{n,m}} \begin{pmatrix} \widehat{G}_{m,0}^{w^*,(n)} - I \\ \widehat{G}_{m,1}^{w^*,(n)} - G^{w^*} \end{pmatrix} \xrightarrow{\tau_{n,m} \rightarrow +\infty} \mathcal{L} \left(\begin{array}{c} \frac{\sigma}{\sqrt{\pi_0}} \mathbb{U} + \rho^{w^*} \sqrt{1 - \sigma^2} (\mathbb{W}_{I^{w^*}} + [I - I^{w^*}]N) \\ \frac{\sigma}{\sqrt{1 - \pi_0}} \mathbb{V}_{G^{w^*}} + (G^{w^*})' \rho^w \sqrt{1 - \sigma^2} (\mathbb{W}_{I^{w^*}} + [I - I^{w^*}]N) \end{array} \right) \\ =: \begin{pmatrix} \mathbb{Z}_0^{w^*} \\ \mathbb{Z}_1^{w^*} \end{pmatrix} \text{ on } [D(0, 1)]^2.$$

with \mathbb{U} , \mathbb{V} and \mathbb{W} three independent Brownian bridges and an independent standard Gaussian random variable N . Furthermore,

$$\sqrt{\tau_{n,m}} \left(\widehat{G}_m^{w^*,(n)} - G^{w^*} \right) \xrightarrow{\tau_{n,m} \rightarrow +\infty} \mathcal{L} \left(\begin{array}{c} \sqrt{\pi_0 \sigma^2} \mathbb{U} + \sqrt{(1 - \pi_0) \sigma^2} \mathbb{V}_{G^{w^*}} \\ + (G^{w^*})' \rho^{w^*} \sqrt{1 - \sigma^2} (\mathbb{W}_{I^{w^*}} + [I - I^{w^*}]N) \end{array} \right) \\ = \pi_0 \mathbb{Z}_0^{w^*} + (1 - \pi_0) \mathbb{Z}_1^{w^*} =: \mathbb{Z}^{w^*} \text{ on } D(0, 1).$$

Proof. We apply twice the argument of the proof of Proposition C.2 to the null and the alternative processes to obtain \mathbb{U} , \mathbb{V} and \mathbb{W} three independent standard Brownian bridge and an independent standard Gaussian r.v. N such that,

$$\begin{pmatrix} \sqrt{m_0(m)} \left(\widehat{F}_{m,0,\text{test}} \circ \left(\widehat{F}_{\text{cal}}^{w^*,(n)} \right)^{-1} (1-I) - \left(\widehat{F}_{\text{cal}}^{w^*,(n)} \right)^{-1} (1-I) \right) \\ \sqrt{m_1(m)} \left(\widehat{F}_{m,1,\text{test}} \circ \left(\widehat{F}_{\text{cal}}^{w^*,(n)} \right)^{-1} (1-I) - F_{\text{test}} \circ \left(\widehat{F}_{\text{cal}}^{w^*,(n)} \right)^{-1} (1-I) \right) \\ \sqrt{n} \left(F_{\text{cal}}^{w^*} \circ \left(\widehat{F}_{\text{cal}}^{w^*,(n)} \right)^{-1} (1-I) - (1-I) \right) \\ \sqrt{n} \left(F_{\text{test}} \circ \left(\widehat{F}_{\text{cal}}^{w^*,(n)} \right)^{-1} (1-I) - F_{\text{test}} \circ \left(F_{\text{cal}}^{w^*} \right)^{-1} (1-I) \right) \end{pmatrix}$$

converges when $\tau_{n,m} \rightarrow +\infty$ in distribution to

$$\begin{pmatrix} \mathbb{U} \\ \mathbb{V} \circ G^{w^*} \\ \rho^{w^*} (\mathbb{W}(I^{w^*}) + [I - I^{w^*}]N) \\ \rho^{w^*} (G^{w^*})' (\mathbb{W}(I^{w^*}) + [I - I^{w^*}]N) \end{pmatrix} \text{ on } [D(0,1)]^4.$$

We conclude by using the decomposition (41) on $\widehat{G}_{m,0}^{w^*,(n)}$ and $\widehat{G}_{m,1}^{w^*,(n)}$, the Slutsky lemma and the continuous mapping theorem. \square

The following result is the weighted version of Proposition 5.4.

Proposition C.4. *Under Assumption 2, assume that P_0 is absolutely continuous with respect to P_{cal} and that the oracle weight function w^* (12) satisfies Assumption 4. Under Assumptions 3, and 6, assuming that the targeted level $\alpha > [(G_{\text{mixt}}^{w^*})'(0^+)]^{-1}$, assuming that $n/(n+m) \rightarrow \sigma^2 \in [0,1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0,1)$, we have,*

$$\begin{aligned} \sqrt{\tau_{n,m}} \left(\mathcal{T}^{\text{BH}_\alpha}(\widehat{G}_m^{w^*,(n)}) - \mathcal{T}_\alpha^{w^*} \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \frac{1}{\frac{1}{\alpha} - (G_{\text{mixt}}^{w^*})'(\mathcal{T}_\alpha^{w^*})} \mathbb{Z}^{w^*}(\mathcal{T}_\alpha^{w^*}); \\ \sqrt{\tau_{n,m}} \left(\text{FDP}_m^{(n)} \left(\mathcal{T}^{\text{BH}_\alpha}(\widehat{G}_m^{w^*,(n)}) \right) - \pi_0 \alpha \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \frac{\pi_0}{G_{\text{mixt}}^{w^*}(\mathcal{T}_\alpha^{w^*})} \mathbb{Z}_0^{w^*}(\mathcal{T}_\alpha^{w^*}); \\ \sqrt{\tau_{n,m}} \left(\text{TDP}_m^{(n)} \left(\mathcal{T}^{\text{BH}_\alpha}(\widehat{G}_m^{w^*,(n)}) \right) - G^w(\mathcal{T}_\alpha^{w^*}) \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \frac{(G^{w^*})'(\mathcal{T}_\alpha^{w^*})}{\frac{1}{\alpha} - (G_{\text{mixt}}^{w^*})'(\mathcal{T}_\alpha^{w^*})} \mathbb{Z}^{w^*}(\mathcal{T}_\alpha^w) + \mathbb{Z}_1^{w^*}(\mathcal{T}_\alpha^w), \end{aligned}$$

where $\mathbb{Z}_0^{w^*}$, $\mathbb{Z}_1^{w^*}$ and \mathbb{Z}^{w^*} are the three processes defined in Proposition C.3.

Proof of Proposition C.4. It is mutatis mutandis the same proof as in Section B.3 by using the weighted processes and the weighted convergence in distribution delineated in Proposition C.3. \square

Appendix D: Weighted novelty detection for a general weight function

In this section, we extend Propositions C.3 and C.4 to the case of a general weight function. We should add the following technical assumption (which was implicitly satisfied in the oracle case):

Assumption 7. F_0 is continuously differentiable.

Let us introduce

$$\begin{aligned} \mathbf{p}(t) &= \frac{\pi_0 G_0^w(t)}{G_{\text{mixt}}^w(t)} \\ \zeta_\alpha(t) &= 1 - (1 - \pi_0) \frac{(G_0^w)' G^w - G_0^w (G^w)'}{(\alpha^{-1} - (G_{\text{mixt}}^w)'(t)) G_0^w} \end{aligned}$$

where $\mathbf{p}(t)$ corresponds to the positive false discovery rate at t (Storey, 2002). Note that in the oracle case $\zeta_\alpha(\mathcal{T}_\alpha^{w^*}) = 0$, but $\zeta_\alpha(\mathcal{T}_\alpha^w)$ is not necessarily zero for a general w .

Proposition D.1. *Under Assumptions 2, 3, 7, assuming that the weight function w satisfies Assumption 4, and assuming that $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0, 1)$, we have*

$$\sqrt{\tau_{n,m}} \begin{pmatrix} \widehat{G}_{m,0}^{w,(n)} - G_0^w \\ \widehat{G}_{m,1}^{w,(n)} - G^w \end{pmatrix} \xrightarrow{\tau_{n,m} \rightarrow +\infty} \begin{pmatrix} \frac{\sigma}{\sqrt{\pi_0}} \mathbb{U}_{G_0^w} + (G_0^w)' \rho^w \sqrt{1 - \sigma^2} (\mathbb{W}_{I^w} + [I - I^w]N) \\ \frac{\sigma}{\sqrt{1 - \pi_0}} \mathbb{V}_{G^w} + (G^w)' \rho^w \sqrt{1 - \sigma^2} (\mathbb{W}_{I^w} + [I - I^w]N) \end{pmatrix} =: \begin{pmatrix} \mathbb{Z}_0^w \\ \mathbb{Z}_1^w \end{pmatrix} \text{ on } [D(0, 1)]^2.$$

with \mathbb{U} , \mathbb{V} and \mathbb{W} three independent Brownian bridges and an independent standard Gaussian random variable N . Furthermore,

$$\begin{aligned} \sqrt{\tau_{n,m}} \left(\widehat{G}_m^{w,(n)} - G_{\text{mixt}}^w \right) &\xrightarrow{\tau_{n,m} \rightarrow +\infty} \sqrt{\pi_0 \sigma^2} \mathbb{U}_{G_0^w} + \sqrt{(1 - \pi_0) \sigma^2} \mathbb{V}_{G^w} \\ &\quad + (G_{\text{mixt}}^w)' \rho^w \sqrt{1 - \sigma^2} (\mathbb{W}_{I^w} + [I - I^w]N) \\ &= \pi_0 \mathbb{Z}_0^w + (1 - \pi_0) \mathbb{Z}_1^w =: \mathbb{Z}^w \text{ on } D(0, 1). \end{aligned}$$

The proof of Proposition D.1 is omitted because it is completely analogue to the one of Proposition C.3.

Theorem D.2. *Under Assumptions 2, 3, 6, 7, assuming that $\alpha > [(G_{\text{mixt}}^w)'(0^+)]^{-1}$ and that the weight function w satisfies Assumption 4. Then if $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0, 1)$, we have*

$$\begin{aligned} \sqrt{\tau_{n,m}} \left(\mathcal{T}^{\text{BH}\alpha}(\widehat{G}_m^{w,(n)}) - \mathcal{T}_\alpha^w \right) &\xrightarrow{\tau_{n,m} \rightarrow +\infty} \frac{1}{\frac{1}{\alpha} - (G_{\text{mixt}}^w)'(\mathcal{T}_\alpha^w)} \mathbb{Z}^w(\mathcal{T}_\alpha^w); \\ \sqrt{\tau_{n,m}} \left(\text{FDP}_m^{(n)}(\mathcal{R}_\alpha^w) - \mathbf{p}(\mathcal{T}_\alpha^w) \right) &\xrightarrow{\tau_{n,m} \rightarrow +\infty} \frac{\mathbb{Z}_0^w(\mathcal{T}_\alpha^w)}{G_0^w(\mathcal{T}_\alpha^w)} \mathbf{p}(\mathcal{T}_\alpha^w) [1 - \mathbf{p}(\mathcal{T}_\alpha^w) \zeta_\alpha(\mathcal{T}_\alpha^w)] \\ &\quad - \frac{\mathbb{Z}_1^w(\mathcal{T}_\alpha^w)}{G^w(\mathcal{T}_\alpha^w)} \mathbf{p}(\mathcal{T}_\alpha^w) [1 - \mathbf{p}(\mathcal{T}_\alpha^w)] \zeta_\alpha(\mathcal{T}_\alpha^w); \\ \sqrt{\tau_{n,m}} \left(\text{TDP}_m^{(n)}(\mathcal{R}_\alpha^w) - G^w(\mathcal{T}_\alpha^w) \right) &\xrightarrow{\tau_{n,m} \rightarrow +\infty} \frac{(G^w)'(\mathcal{T}_\alpha^w)}{\frac{1}{\alpha} - (G_{\text{mixt}}^w)'(\mathcal{T}_\alpha^w)} \mathbb{Z}^w(\mathcal{T}_\alpha^w) + \mathbb{Z}_1^w(\mathcal{T}_\alpha^w), \end{aligned}$$

where the three processes $(\mathbb{Z}_0^w, \mathbb{Z}_1^w, \mathbb{Z}^w)$ are defined in Proposition D.1.

The proof of Theorem D.2 is omitted because completely analogue to the one of Proposition C.4.

Appendix E: Additional tools for asymptotics

E.1. Donsker and Glivenko-Cantelli theorems

The results below can be found in van der Vaart (1998) and Shorack and Wellner (1986).

Theorem E.1 (Glivenko-Cantelli). *Let F_{test} be a c.d.f., and $\widehat{F}_{m,\text{test}}$ be an empirical version of F_{test} with m i.i.d. points. Then,*

$$\left\| \widehat{F}_{m,\text{test}} - F_{\text{test}} \right\|_\infty \rightarrow 0 \text{ a.s.}$$

Proposition E.2. *Let $(U_k)_{k \geq 1}$ be i.i.d. random variables uniformly distributed over $(0, 1)$. Denote $\mathbb{F}^{(n)}$ the e.c.d.f. the family $(U_k)_{k \in \llbracket n \rrbracket}$. Then,*

$$\left\| \left(\mathbb{F}^{(n)} \right)^{-1} - I \right\|_\infty \rightarrow 0 \text{ a.s.}$$

Theorem E.3 (Donsker (1952)). *Let $(X_i)_{i \geq 1}$ i.i.d. real random variables. Let F the cumulative distribution function of X_1 and \mathbb{F}_n the empirical cdf of $\{X_1, \dots, X_n\}$. Then,*

$$\sqrt{n}(\mathbb{F}_n - F) \xrightarrow{\mathcal{L}} \mathbb{U}_F,$$

with \mathbb{U} a standard Brownian bridge.

Proposition E.4. *Let w a weight function satisfying Assumption 4. Then the family of function $\mathcal{F} = \{w_t : x \in \mathbb{R} \mapsto w(x)\mathbb{1}_{x \leq t}; t \in \mathbb{R}\}$ is P_{cat} -Donsker and P_{cat} -Glivenko-Cantelli.*

E.2. Hadamard differentiability

The three first results below can be found in van der Vaart (1998).

Lemma E.5. *Let $F : [a, b] \rightarrow \mathbb{R}$ be a continuously differentiable function. The map $T : \phi \mapsto F \circ \phi$ with entries being functions $\phi : E \mapsto [a, b]$ contained in $\ell^\infty(E)$ is Hadamard differentiable tangentially to $\ell^\infty(E)$ with derivative:*

$$\dot{J}_\phi(H) = F'(\phi)H.$$

Assumption 8. *The cumulative distribution function F have a compact support $[a, b]$ in the sense that for all $x \leq a$, $F(x) = 0$, for all $x \geq b$, $F(x) = 1$ and for all $x \in (a, b)$, $0 < F(x) < 1$. Furthermore, F is continuously differentiable on (a, b) with strictly positive derivative.*

Lemma E.6. *Let F satisfying Assumption 8. Then the inverse map $G \mapsto G^{-1}$ with domain the set of cumulative distribution function of probability measure on $(a, b]$ with value in $\ell^\infty(0, 1)$ is Hadamard differentiable at F tangentially to $\mathcal{C}[a, b]$ with derivative being the map*

$$H \mapsto -H \circ F^{-1}F^{-1'} = -\frac{H \circ F^{-1}}{F' \circ F^{-1}}.$$

Corollary E.7. *Let $(X_i)_{i \geq 1}$ be i.i.d. real random variables. Let F be the cumulative distribution function of X_1 and \mathbb{F}_n be the empirical cdf of $\{X_1, \dots, X_n\}$. Denote F^{-1} and \mathbb{F}_n^{-1} the corresponding quantile functions. Assume that F satisfies Assumption 8 with derivative f . Then,*

$$\sqrt{n}(\mathbb{F}_n^{-1} - F^{-1}) \xrightarrow{\mathcal{L}} \frac{\mathbb{U}}{f \circ F^{-1}} = (F^{-1})'\mathbb{U},$$

with \mathbb{U} being a standard Brownian bridge.

Lemma E.8. *Let w a weight function satisfying Assumption 4. Denote $K : t \in [0, 1] \rightarrow \int_0^t w(x) dx$. The map $T : F \in D[0, 1] \rightarrow F/F(1) \in D[0, 1]$ is Hadamard differentiable at K tangentially to $\mathcal{C}[0, 1]$ with the following formula:*

$$\dot{T}_K(H)(u) = \frac{H}{K(1)} - \frac{H(1)}{K(1)}W,$$

with $W = T(K) = K/K(1)$.

Proof of Lemma E.8. Let $K \in D[0, 1]$ defined as in the statement. Let $(H_t)_{t \in [0, 1]}$ be a family of function in $\mathcal{C}[0, 1]$ with $H_t \rightarrow H \in \mathcal{C}[0, 1]$ uniformly on $[0, 1]$. We have, for t small enough,

$$\begin{aligned} T(K + tH_t) - T(K) &= \frac{K(1)(K + tH_t) - (K(1) + tH_t(1))K}{(K(1) + tH_t(1))K(1)} \\ &= \frac{K(1)tH_t - tH_t(1)K}{(K(1) + tH_t(1))K(1)} \\ &= t \times \left(\frac{H_t}{K(1) + tH_t(1)} - \frac{H_t(1)}{K(1) + tH_t(1)} \frac{K}{K(1)} \right). \end{aligned}$$

Hence, we obtain

$$\frac{T(K + tH_t) - T(K)}{t} \xrightarrow{t \rightarrow 0} \frac{H}{K(1)} - \frac{H(1)}{K(1)} \frac{K}{K(1)},$$

which prove the Hadamard differentiability. \square

We gather below the formulas of the Hadamard derivatives of the functionals of interest when studying the asymptotic of the FDP. They are obtained from [Neuvial \(2008\)](#).

Lemma E.9. *Let $G : [0, 1] \rightarrow [0, 1]$ be a continuously differentiable increasing strictly concave function. If $\alpha > [G'(0)]^{-1}$, then $\mathcal{T}^{\text{BH}\alpha}$ is Hadamard differentiable at G tangentially to $\mathcal{C}[0, 1]$ with the following expression for all $H \in \mathcal{C}[0, 1]$*

$$(\dot{\mathcal{T}}^{\text{BH}\alpha})_G(H) = \frac{H(\mathcal{T}^{\text{BH}\alpha}(G))}{\alpha^{-1} - G'(\mathcal{T}^{\text{BH}\alpha}(G))}.$$

Let (G_0, G_1) be two continuously differentiable c.d.f. functions from $[0, 1]$ to $[0, 1]$ and let $\pi \in (0, 1)$ such that $G = \pi G_0 + (1 - \pi)G_1$ is strictly concave. Let $\alpha > [G'(0)]^{-1}$. In addition, let us define

$$\begin{aligned} \nu : (F_0, F_1) &\rightarrow F_0(\mathcal{T}^{\text{BH}\alpha}(\pi F_0 + (1 - \pi)F_1)); \\ \Psi : (F_0, F_1) &\rightarrow \frac{\pi F_0(\mathcal{T}^{\text{BH}\alpha}(F))}{F(\mathcal{T}^{\text{BH}\alpha}(F))}; \end{aligned} \tag{53}$$

$$\Phi : (F_0, F_1) \rightarrow F_1(\mathcal{T}^{\text{BH}\alpha}(F)), \tag{54}$$

where $F = \pi F_0 + (1 - \pi)F_1$. Then, ν , Ψ and Φ are Hadamard differentiable at (G_0, G_1) tangentially to $(\mathcal{C}[0, 1])^2$ with the following derivative expressions: for all $(H_0, H_1) \in (\mathcal{C}[0, 1])^2$,

$$\begin{aligned} \dot{\nu}_{G_0, G_1}(H_0, H_1) &= G'_0(\mathcal{T}^{\text{BH}\alpha}(G)) \frac{H(\mathcal{T}^{\text{BH}\alpha}(G))}{\alpha^{-1} - G'(\mathcal{T}^{\text{BH}\alpha}(G))} + H_0(\mathcal{T}^{\text{BH}\alpha}(G)); \\ \dot{\Psi}_{G_0, G_1}(H_0, H_1) &= \frac{\pi}{G(\mathcal{T}^{\text{BH}\alpha}(G))} \left(\dot{\nu}_{G_0, G_1}(H_0, H_1) - \frac{G_0(\mathcal{T}^{\text{BH}\alpha}(G))}{G(\mathcal{T}^{\text{BH}\alpha}(G))} \left[H(\mathcal{T}^{\text{BH}\alpha}(G)) + G'(\mathcal{T}^{\text{BH}\alpha}(G))(\dot{\mathcal{T}}^{\text{BH}\alpha})_G(H) \right] \right); \\ \dot{\Phi}_{G_0, G_1}(H_0, H_1) &= G'_1(\mathcal{T}^{\text{BH}\alpha}(G)) \frac{H(\mathcal{T}^{\text{BH}\alpha}(G))}{\alpha^{-1} - G'(\mathcal{T}^{\text{BH}\alpha}(G))} + H_1(\mathcal{T}^{\text{BH}\alpha}(G)), \end{aligned}$$

where $H = \pi H_0 + (1 - \pi)H_1$.

E.3. Random change of time

We present here a version of the random change of time lemma of [Billingsley \(1999\)](#) (page 151), which is adapted to the topological spaces $D(\mathbb{R})$, $\ell^\infty(0, 1)$ and $D(0, 1)$.

Lemma E.10. *Let $(U_n, V_n)_n$ be a sequence of random processes in $D(\mathbb{R}) \times D(0, 1)$, \mathbb{U} and \mathbb{V} two random processes in $D(\mathbb{R})$ and $D(0, 1)$, respectively, which are both a.s. continuous and such that $(U_n, V_n) \xrightarrow{\mathcal{L}} (\mathbb{U}, \mathbb{V})$ in $D(\mathbb{R}) \times D(0, 1)$. Let $(F_n)_n$ be a sequence of random processes in $D(0, 1)$ and $F \in D(0, 1)$ such that $(F_n)_n$ converges in probability to F on $\ell^\infty(0, 1)$. Assume that for all $n \in \mathbb{N}$, $U_n \circ F_n \in D(0, 1)$. Then,*

$$(U_n \circ F_n, V_n) \xrightarrow{\mathcal{L}} (\mathbb{U} \circ F, \mathbb{V}) \text{ on } (D(0, 1))^2.$$

Proof of Lemma E.10. Since \mathbb{U} and \mathbb{V} are continuous, $(U_n, V_n)_n$ converges in distribution to (\mathbb{U}, \mathbb{V}) on $\ell^\infty(\mathbb{R}) \times \ell^\infty(0, 1)$ by Lemma E.12. Hence, by Slutsky's lemma, the sequence $(U_n, V_n, F_n)_n$ converges in distribution to $(\mathbb{U}, \mathbb{V}, F)$ on $\ell^\infty(\mathbb{R}) \times (\ell^\infty(0, 1))^2$. Hence by Lemma E.11 and the continuous mapping theorem we obtain,

$$(U_n \circ F_n, \mathbb{V}) \xrightarrow{\mathcal{L}} \mathbb{U} \circ F \text{ on } \ell^\infty(0, 1).$$

Since \mathbb{U} is continuous, then $\mathbb{U} \circ F$ is in $D(0, 1)$ and the previous convergence implies the convergence in $(D(0, 1))^2$. \square

Note that when using Lemma E.10 in our work, the convergence of the sequence $(U_n)_n$ is typically given by the Donsker Theorem, while the convergence of $(F_n)_n$ is given by the Glivenko-Cantelli theorem.

Lemma E.11. *Define the following map $\Psi : (F_1, F_2, F_3) \in \ell^\infty(\mathbb{R}) \times (\ell^\infty(0, 1))^2 \mapsto (F_1 \circ F_3, F_2) \in (\ell^\infty(0, 1))^2$. For all $U \in \mathcal{C}(\mathbb{R})$ and $(F, V) \in (\ell^\infty(0, 1))^2$, Ψ is continuous at (U, V, F) .*

Proof of Lemma E.11. Let $U \in \mathcal{C}(\mathbb{R})$ and $(V, F) \in (\ell^\infty(0, 1))^2$. Let $(U_n, V_n, F_n)_n$ a sequence in $\ell^\infty(\mathbb{R}) \times (\ell^\infty(0, 1))^2$ which converges to (U, V, F) . Let $K \subset (0, 1)$ be a compact set. Let $\varepsilon > 0$. The set $\{F_n(x), x \in K, n \in \mathbb{N}\}$ is bounded since $\|F_n - F\|_{\infty, K} \rightarrow 0$ and $(F_n)_n$ and F belongs in $\ell^\infty(0, 1)$, hence is included in a compact set K' of \mathbb{R} . Since U is continuous, we have $\eta > 0$ such that for all $(x, y) \in K'^2$, $|x - y| \leq \eta$ implies that $|U(x) - U(y)| \leq \varepsilon$. Let $N > 0$ such that for all $n \geq N$, $\|U_n - U\|_{\infty, K'} \leq \varepsilon$ and $\|F_n - F\|_{\infty, K} \leq \eta$. Then for all $n \geq N$,

$$\begin{aligned} \|U_n \circ F_n - U \circ F\|_{\infty, K} &\leq \|U_n \circ F_n - U \circ F_n\|_{\infty, K} + \|U \circ F_n - U \circ F\|_{\infty, K} \\ &\leq \|U_n - U\|_{\infty, K'} + \varepsilon \\ &\leq 2\varepsilon, \end{aligned}$$

where the first ε appears since for all $x \in K$, $F_n(x) \in K'$, $F(x) \in K'$ and $|F_n(x) - F(x)| \leq \eta$. Furthermore for all \tilde{K} a compact subset of $(0, 1)$, $\|V_n - V\|_{\infty, \tilde{K}} \rightarrow 0$. Hence, Ψ is continuous at (U, V, F) . \square

Finally, the following result is classical and shows how a convergence in distribution on $D[0, 1]$ can imply a uniform convergence (Billingsley, 1999), with the local topology defined in Section 2.3.

Lemma E.12 (Billingsley (1999)). *Let $(X_n)_n$ be a sequence of random processes on $D(0, 1)$. Let $X \in D(0, 1)$ be a continuous function a.s.. Then, $(X_n)_n$ converges in distribution to X on $D(0, 1)$ if and only if $(X_n)_n$ converges in distribution to X on $\ell^\infty(0, 1)$*

Appendix F: Useful Lemmas

Lemma F.1. *Let $(s_1, \dots, s_{n+1}) \in \mathbb{R} \cup \{+\infty\}$, $(w_k)_{k \in \llbracket n+1 \rrbracket}$ be a family of positive weight summing to 1 and $\mu = \sum_{k \in \llbracket n+1 \rrbracket} w_k \delta_{s_k}$ be a probability measure on $\mathbb{R} \cup \{+\infty\}$. Define for all $\alpha \in (0, 1)$, $Q_\alpha(\mu) = \inf\{x \in \mathbb{R} : \mu([-\infty, x]) \geq \alpha\}$ the α -quantile of the probability measure μ . Then, for all $t \in \mathbb{R}$ and $\alpha \in (0, 1)$, the two following assertions are equivalent:*

- (i) $\sum_{k \in \llbracket n+1 \rrbracket} w_k \mathbb{1}_{s_k \geq t} \leq \alpha$
- (ii) $t > Q_{1-\alpha}(\mu)$

Proof of Lemma F.1. First note that, since the weights are summing to 1, (i) is equivalent to $\sum_{k \in \llbracket n+1 \rrbracket} w_k \mathbb{1}_{s_k < t} \geq 1 - \alpha$. Now prove that (ii) implies (i). If $t > Q_{1-\alpha}(\mu)$, then $\sum_{k \in \llbracket n+1 \rrbracket} w_k \mathbb{1}_{s_k < t} \geq \sum_{k \in \llbracket n+1 \rrbracket} w_k \mathbb{1}_{s_k \leq Q_{1-\alpha}(\mu)} \geq 1 - \alpha$ and thus (i) holds.

Conversely, if (i) holds, we have $\sum_{k \in \llbracket n+1 \rrbracket} w_k \mathbb{1}_{s_k < t} \geq 1 - \alpha > 0$ and thus

$$S := \max_{k \in \llbracket n+1 \rrbracket} \{s_k : s_k < t\} \in \mathbb{R}$$

exists. With such a $S < t$ we have $\sum_{k \in \llbracket n+1 \rrbracket} w_k \mathbb{1}_{s_k \leq S} = \sum_{k \in \llbracket n+1 \rrbracket} w_k \mathbb{1}_{s_k < t} \geq 1 - \alpha$. This proves $S \geq Q_{1-\alpha}(\mu)$ by definition of the quantile function. Hence $t > Q_{1-\alpha}(\mu)$ and we have proved (ii). \square

Lemma F.2. $\frac{\tau_{n,m}}{m} \rightarrow \sigma^2 \in [0, 1]$ if and only if $\frac{m}{n} \rightarrow \sigma^{-2} - 1 \in [0, +\infty]$.

Appendix G: Relaxing assumptions

In this section, we show that some of our results can be stated under weaker distributional assumptions. More precisely, we aim to transfer the assumptions on the marginal distributions of the scores (e.g. Assumption 3) to assumptions on the marginal distribution of the theoretical p -value $(p_i^{(+\infty)})_{i \in \llbracket m \rrbracket}$ defined in (5).

For this, let S be a random variable with c.d.f. F_{cal} . We denote $F_{\text{cal}-}$ its left continuous version, i.e. $F_{\text{cal}-}(t) = \mathbb{P}(S < t)$ for all $t \in \mathbb{R}$. We denote its general inverse $F_{\text{cal}-}^{-1}$ as follows:

$$F_{\text{cal}-}^{-1}(u) = \sup\{t \in \mathbb{R} : F_{\text{cal}-}(t) \leq u\}, \quad u \in (0, 1),$$

with the convention that $\sup \emptyset = -\infty$ (see Lemma 13 in the supplementary material of Fromont, Lerasle and Reynaud-Bouret (2016) for an extensive study of $F_{\text{cal}-}^{-1}$). Let us introduce

$$\tilde{F}_{\text{test,cal}}(t) = F_{\text{test}} \circ F_{\text{cal}-}^{-1}(t), \quad t \in [0, 1]; \quad (55)$$

with the convention that $\tilde{F}_{\text{test,cal}}(1) = 1$ and that $F_{\text{test}}(-\infty) = 0$.

Assuming that the calibration scores $(S_k)_{k \geq 1}$ and test scores $(T_i)_{i \in \llbracket n \rrbracket}$ are independent and that the calibration scores are i.i.d. with c.d.f. F_{cal} and the test scores are i.i.d. with c.d.f. F_{test} , then for all $i \in \mathbb{N}$:

$$p_i^{(+\infty)} = 1 - F_{\text{cal}-}(T_i). \quad (56)$$

Since $F_{\text{cal}-}(T_i)$ has for c.d.f. $\tilde{F}_{\text{test,cal}}$ on $[0, 1]$, using the latter gives us that:

$$\mathbb{P}(p_i^{(+\infty)} < t) = 1 - \tilde{F}_{\text{test,cal}}(1 - t) =: \tilde{G}(t), \quad \forall t \in (0, 1) \quad (57)$$

hence making assumptions on \tilde{G} (57) or equivalently on $\tilde{F}_{\text{test,cal}}$ (55) consists on making assumptions on the marginal distribution of the theoretical p -value.

G.1. Results without weight

Assumptions 1 and 3 can be relaxed as follows.

Assumption 9. *The set of calibration scores $(S_k)_{k \geq 1}$ and the set of test scores $(T_i)_{i \geq 1}$ are two independent families of real random variables. The $S_k, k \geq 1$, (resp. $T_i, i \geq 1$) are i.i.d. with distribution P_{cal} (resp. P_{test}) and cumulative distribution function F_{cal} (resp. F_{test}). The function \tilde{G} (57) is continuously differentiable on $(0, 1)$.*

Assumption 9 is an assumption on the data structure (as in the beginning of Assumption 1 and most importantly on the distribution of the theoretical p -value and not directly on the distribution of the scores as in the last part of Assumption 1 and Assumption 3. Example G.1 shows that Assumption 9 is strictly weaker than Assumption 3.

Example G.1. *Assume that P_{cal} is the uniform distribution on $(0, 1) \cup (2, 3)$ and P_{test} has for density $f_1(x)\mathbb{1}_{x \in (0,1)} + f_2(x-2)\mathbb{1}_{x \in (2,3)}$ with f_1, f_2 two continuous function from $[0, 1]$ to \mathbb{R}^+ with $f_1(1) = f_2(0)$. We get for all $t \in [0, 3]$ that $F_{\text{cal}}(t) = 1/2(t\mathbb{1}_{t \in (0,1)} + \mathbb{1}_{t \in [1,3]} + (t-2)\mathbb{1}_{t \in (2,3)})$ which is continuous nondecreasing and constant on $(1, 2)$, while for all $u \in (0, 1)$, $F_{\text{cal}}^{-1}(u) = 2u\mathbb{1}_{u \in (0,1/2]} + (1+2u)\mathbb{1}_{u \in (1/2,1)}$ which has a discontinuity point at $1/2$. Denoting $F_1(t) = \int_0^t f_1 \, d\lambda$ and $F_2(t) = \int_0^t f_2 \, d\lambda$ for all $t \in [0, 1]$ we get that*

$$\tilde{F}_{\text{test,cal}}(u) = F_1(2u)\mathbb{1}_{u \in (0,1/2)} + (F_1(1) + F_2(2u-1))\mathbb{1}_{u \in [1/2,1)}, \quad \forall u \in (0, 1),$$

which is continuously differentiable on $(0, 1)$ with derivative equals to $2f_1(2u)$ for all $u \in (0, 1/2]$ and to $2f_2(2u-1)$ for all $u \in [1/2, 1)$ since $f_1(1) = f_2(0)$.

Theorem G.1. *Under Assumptions 9 and assuming that $n/(n+m)$ tends to $\sigma^2 \in [0, 1]$, we have*

$$\sqrt{\tau_{n,m}} \left(\text{FCP}_m^{(n)} - \tilde{G} \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \sigma \mathbb{U} \circ \tilde{G} + \sqrt{1 - \sigma^2} \tilde{G}' \mathbb{V} \text{ on } D(0, 1),$$

where \mathbb{U} and \mathbb{V} are two independent Brownian bridges and \tilde{G} is defined in (57).

Theorem G.1 is proved in Section G.2.1 and extends Theorem 3.2. The function \tilde{G} (57) is the generalisation of G (14) when F_{cal} is not continuous and is equal to G when F_{cal} is continuous.

The same kind of weakened assumption holds in the novelty detection setting. In this setting there are two different distribution for the theoretical p -values: the one under the null hypothesis (the test score has for c.d.f. F_0 under the null) and the one under the alternative (the test score has for c.d.f. F_{test} under the alternative). \tilde{G} given by (57) is the c.d.f. of the theoretical p -values under the alternative, and the c.d.f. of the theoretical p -value under the null is given by:

$$\tilde{G}_0(t) = 1 - F_0 \circ F_{\text{cal}}^{-1}(1 - t), \quad \forall t \in (0, 1). \quad (58)$$

Since we want uniform theoretical p -values under the null, we will assume that $\tilde{G}_0 = I$. This is for example true when $P_{\text{cal}} = P_0$ and P_{cal} is the distribution in Example G.1. We define the mixture c.d.f. \tilde{G}_{mixt} :

$$\tilde{G}_{\text{mixt}} = \pi_0 \tilde{G}_0 + (1 - \pi_0) \tilde{G}, \quad (59)$$

and we state our new weakened assumption in the novelty detection setting.

Assumption 10. *The set of calibration scores $(S_k)_{k \geq 1}$ and the set of test scores $(T_i)_{i \geq 1}$ are two independent families of real random variables. The variables S_k , $k \geq 1$, are i.i.d. with distribution P_{cal} and c.d.f. F_{cal} . The variables T_i , $i \geq 1$, are independent, the variables T_i , $i \in \mathcal{H}_0$, are identically distributed as a null score distribution P_{cal} and c.d.f. F_{cal} , and the variables T_i , $i \in \mathcal{H}_1$, are identically distributed as an alternative score distribution P_{test} (potentially different from P_0) with c.d.f. F_{test} . Moreover, the function \tilde{G} (55) is continuously differentiable on $(0, 1)$, the function \tilde{G}_0 (58) is equal to I on $(0, 1)$ and the function \tilde{G}_{mixt} (59) is concave.*

Assumption 10 is a generalisation of Assumptions 2, 3 and 5. More precisely, while the data share the same structure (independence and same distributions between groups), the assumptions are now on the marginal distribution of the theoretical p -values under the null and the alternative, and on the mixture distribution of the theoretical p -value in Efron's two groups model. Example G.2 shows the generalisation given by Assumption 10.

Example G.2. *Assume that $P_0 = P_{\text{cal}}$ is the uniform distribution on $(0, 1) \cup (2, 3)$. Let $\gamma > 0$, we take P_{test} a distribution with density $Z_\gamma \gamma (e^{\gamma x} \mathbf{1}_{x \in (0,1)} + e^{\gamma(x-1)} \mathbf{1}_{x \in (2,3)})$ where $Z_\gamma = (e^{2\gamma} - 1)^{-1}$ is a normalising constant. P_{cal} does not verify Assumption 3, but as shown in Example G.1 $\tilde{F}_{\text{test,cal}}$ is continuously differentiable. More precisely we get for all $u \in (0, 1)$,*

$$\tilde{F}_{\text{test,cal}}(u) = Z_\gamma (e^{2\gamma u} - 1),$$

hence for any $\pi_0 \in (0, 1)$, we get that \tilde{G}_{mixt} is concave. Furthermore, even if F_{cal} is constant on $(1, 2)$, we still get that $F_{\text{cal}} \circ F_{\text{cal}}^{-1}(u) = u$ for all $u \in (0, 1)$. So even if Assumption 3 does not hold, Assumption 10 holds hence Theorem G.2 can be applied in this case.

Theorem G.2. *Under Assumption 10, let us consider \mathcal{R}_α the BH procedure at level α applied to the conformal p -values (4), with a level $\alpha > [\tilde{G}'_{\text{mixt}}(0^+)]^{-1}$. If $n/(n+m) \rightarrow \sigma^2 \in [0, 1]$ and $\pi_0(m) \rightarrow \pi_0 \in (0, 1)$, we have*

$$\begin{aligned} \sqrt{\tau_{n,m}} \left(\text{FDP}_m^{(n)}(\mathcal{R}_\alpha) - \pi_0 \alpha \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \alpha^2 \pi_0 [\sigma^2 + (1 - \sigma^2) \pi_0] \frac{1 - \mathcal{T}_\alpha}{\mathcal{T}_\alpha} \right); \\ \sqrt{\tau_{n,m}} \left(\text{TDP}_m^{(n)}(\mathcal{R}_\alpha) - \tilde{G}(\mathcal{T}^*) \right) &\xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{\Sigma_\alpha}{(\alpha^{-1} - \tilde{G}'_{\text{mixt}}(\mathcal{T}_\alpha))^2} \right), \end{aligned}$$

with $\Sigma_\alpha = \tilde{G}'(\mathcal{T}_\alpha)^2 \mathcal{T}_\alpha (1 - \mathcal{T}_\alpha) [\pi_0 \sigma^2 + (1 - \sigma^2) \alpha^{-2}] + [\alpha^{-1} - \pi_0]^2 (1 - \pi_0)^{-1} \tilde{G}(\mathcal{T}_\alpha) (1 - \tilde{G}(\mathcal{T}_\alpha)) \sigma^2$, and $\mathcal{T}_\alpha = \sup\{t \in (0, 1), \tilde{G}_{\text{mixt}}(t) \geq t \alpha^{-1}\}$.

Theorem G.2 is proved in Section G.2.2 and is a generalisation of Theorem 4.1.

G.2. Proof of Theorems G.1 and G.1

The proofs of Theorems G.1 and G.2 are consequences of the following lemma.

Lemma G.3. *Let S and T be any independent real random variables. Let $w : \mathbb{R} \rightarrow \mathbb{R}^+$ be any weight function. Let $U \sim \mathcal{U}(0, 1)$ a random variable independent of T . We have the following equality in distribution:*

$$w(S) \mathbb{1}_{S \geq T} \stackrel{\mathcal{L}}{=} (w \circ F_{\text{cal-}}^{-1})(U) \mathbb{1}_{U \geq F_{\text{cal-}}(T)},$$

and $F_{\text{cal-}}(T)$ has for c.d.f. $\tilde{F}_{\text{test,cal}}$ (55).

Proof. First, we use that $S \stackrel{\mathcal{L}}{=} F_{\text{cal-}}^{-1}(U)$. Let $t \in \mathbb{R}$. We have that $\mathbb{P}(F_{\text{cal-}}^{-1}(U) < t) = \mathbb{P}(U < F_{\text{cal-}}(t))$ by Lemma 13 (3) from Fromont, Lerasle and Reynaud-Bouret (2016). This implies that $\mathbb{P}(F_{\text{cal-}}^{-1}(U) < t) = F_{\text{cal-}}(t)$ since $F_{\text{cal-}}(t) \in [0, 1]$ hence the c.d.f. of $F_{\text{cal-}}^{-1}(U)$ is the right continuous version of $F_{\text{cal-}} : F_{\text{cal-}}$. Since S and T are independent, and that U and T are also independent we have that $(S, T) \stackrel{\mathcal{L}}{=} (F_{\text{cal-}}^{-1}(U), T)$. Using the latter we obtain

$$\begin{aligned} w(S) \mathbb{1}_{S \geq T} &\stackrel{\mathcal{L}}{=} w(F_{\text{cal-}}^{-1}(U)) \mathbb{1}_{F_{\text{cal-}}^{-1}(U) \geq T} \\ &= (w \circ F_{\text{cal-}}^{-1})(U) \mathbb{1}_{U \geq F_{\text{cal-}}(T)}, \end{aligned}$$

by using again Lemma 13 (3) from Fromont, Lerasle and Reynaud-Bouret (2016). \square

Thanks to this general lemma, one can assume that all the calibration point are i.i.d. uniform on $(0, 1)$, and to change F_{test} by $\tilde{F}_{\text{test,cal}}$ and w by $w \circ F_{\text{cal-}}^{-1}$ (which is equal to 1 in the unweighted case).

G.2.1. Proof of Theorem G.1

Using Lemma G.3, one can assume without any loss of generality that all the calibration point are i.i.d. uniform on $(0, 1)$ and that all the test points are i.i.d. with c.d.f. $\tilde{F}_{\text{test,cal}}$. Since for all $t \in (0, 1)$, $\tilde{F}_{\text{test,cal}}(t) = 1 - \tilde{G}(1 - t)$ we have by Assumption 9 that the data, I (the new c.d.f. of the calibration points) and $\tilde{F}_{\text{test,cal}}$ (the new c.d.f. of the test points) satisfy Assumption 1 (since the scores are still i.i.d. within each groups and independent by Assumption 9) and Assumption 3. The result is then implied by Theorem 3.2 since for all $t \in (0, 1)$, $\tilde{G}(t) = 1 - \tilde{F}_{\text{test,cal}}(I^{-1}(1 - t))$.

G.2.2. Proof of Theorem G.2

Using Lemma G.3, one can assume without any loss of generality that all the calibration point are i.i.d. uniform on $(0, 1)$, that all the test points under the alternative are i.i.d. with c.d.f. $\tilde{F}_{\text{test,cal}}$ and that all the test points under the null are i.i.d. with c.d.f. \tilde{G}_0 which is equal to I on $(0, 1)$ by Assumption 10. In these setting, we get by Assumption 10 that Assumptions 2, 3 and 5 are true with $\tilde{G}_0 = I$, i.e. the test points under the null have the same distribution that the calibration points. The result is then implied by Theorem 4.1 since for all $t \in (0, 1)$, $\tilde{G}(t) = 1 - \tilde{F}_{\text{test,cal}}(I^{-1}(1 - t))$ and $\tilde{G}_{\text{mixt}}(t) = \pi_0 t + (1 - \pi_0) \tilde{G}(t)$.

G.3. Result with weight

When using a weight function, after applying Lemma G.3 one still needs to put assumptions on w and F_{cal}^w (at least on a quantity which will play a similar role) in order to use the asymptotics results necessary to the convergence of the FCP, namely Lemma C.1 and therein obtaining that the family $\tilde{\mathcal{F}} = \{\tilde{w}_t : x \in (0, 1) \mapsto w \circ F_{\text{cal}-}^{-1}(x) \mathbb{1}_{x \leq t}\}$ is $\mathcal{U}(0, 1)$ -Glivenko-Cantelli and Donsker. We introduce the new weighted calibration function:

$$\widetilde{F}_{\text{cal}}^w : t \in [0, 1] \mapsto \frac{\int_0^t w \circ F_{\text{cal}-}^{-1}(u) \, d u}{\int_0^1 w \circ F_{\text{cal}-}^{-1}(u) \, d u},$$

which corresponds to F_{cal}^w (15) with w replaced by $w \circ F_{\text{cal}-}^{-1}$ and P_{cal} replaced by the uniform distribution over $(0, 1)$. By definition, $\widetilde{F}_{\text{cal}}^w$ is differentiable on $(0, 1)$ with derivative $w \circ F_{\text{cal}-}^{-1}$ which leads us to introduce the following assumption which will replace Assumption 4 in this relaxed setting.

Assumption 11. *The weight function $w : \mathbb{R} \cup \{+\infty\} \rightarrow \mathbb{R}^+$ is measurable function. Furthermore, the function $w \circ F_{\text{cal}-}$ is bounded and continuous on $(0, 1)$ and positive on its support in the sense that there exist $0 \leq a < b \leq 1$ such that for all $x \in (a, b)$, $w \circ F_{\text{cal}-}(x) > 0$ and for all $x \in (0, 1) \setminus (a, b)$, $w \circ F_{\text{cal}-}^{-1}(x) = 0$.*

Assumption 11 is still an assumption on the distribution of an asymptotic p -values, but it is less straightforward than Assumptions 9 and 10. To understand this, assume that the weight function w is non-negative and bounded, and let $(S_k)_{k \geq 1}$ be i.i.d. real random variables with uniform distribution on $(0, 1)$ and let S be an independent real random variable with uniform distribution $(0, 1)$. Let $F_{\text{cal}-}$ be any left continuous with right limits c.d.f. and assume that $\mathbb{E}(w \circ F_{\text{cal}-}^{-1}(S_1)) > 0$. Then, by applying the law of large numbers we get, for any $w(+\infty) > 0$,

$$\frac{w(+\infty) + \sum_{k \in \llbracket n \rrbracket} w \circ F_{\text{cal}-}^{-1}(S_k) \mathbb{1}_{S_k \geq S}}{w(+\infty) + \sum_{k \in \llbracket n \rrbracket} w \circ F_{\text{cal}-}^{-1}(S_k)} \rightarrow \frac{\mathbb{E}[w \circ F_{\text{cal}-}^{-1}(S_1) \mathbb{1}_{S_1 \geq S} | S]}{\mathbb{E}[w \circ F_{\text{cal}-}^{-1}(S_1)]}.$$

Since S as an uniform distribution on $(0, 1)$, Assumption 11 is an assumption on the c.d.f. of the limiting p -values when you test if a uniform random variable has the distribution induced by $w \circ F_{\text{cal}-}^{-1}$ on the uniform distribution on $(0, 1)$. Assumption 11 states that the p -value $\mathbb{E}[w \circ F_{\text{cal}-}^{-1}(S_1) \mathbb{1}_{S_1 \geq S} | S] / \mathbb{E}[w \circ F_{\text{cal}-}^{-1}(S_1)]$ has a density which is positive on its support. Example G.3 shows the generalisation given by Assumption 11.

Example G.3. *Assume the setting of Example G.1: P_{cal} is the uniform distribution on $(0, 1) \cup (2, 3)$ and P_{test} has for density $f_1(x) \mathbb{1}_{x \in (0, 1)} + f_2(x - 2) \mathbb{1}_{x \in (2, 3)}$ with f_1, f_2 two continuous function from $[0, 1]$ to \mathbb{R}^+ with $f_1(1) = f_2(0)$. Those two distributions satisfy Assumption 9. We define the weight function $w : x \in (0, 3) \mapsto w_1(x) \mathbb{1}_{x \in (0, 1)} + w_2(x - 2) \mathbb{1}_{x \in (2, 3)}$ with w_1, w_2 two bounded continuous functions from $[0, 1]$ to \mathbb{R}^+ with $w_1(1) = w_2(0)$ and with a couple $(a, b) \in (0, 1)^2$ such that if $u \in (0, 1)$ $w_1(u) = 0$ if and only if $u \geq a$, and $w_2(u) = 0$ if and only if $u \geq b$. Denoting for all $t \in [0, 1]$ $W_1(t) = \int_0^t w_1(u) \, d u$, $W_2(t) = \int_0^t w_2(u) \, d u$ and the normalising constant $Z_w = W_1(1) + W_2(1)$ we get for all $u \in (0, 1)$,*

$$\widetilde{F}_{\text{cal}}^w(u) = Z_w^{-1} (W_1(2u) \mathbb{1}_{u \in [0, 1/2]} + [W_1(1) + W_2(2u - 1)] \mathbb{1}_{u \in [1/2, 1]}), \quad (60)$$

which is differentiable with derivative $w \circ F_{\text{cal}-}^{-1}$ equals to $2Z_w^{-1} w_1(2u)$ for all $u \in (0, 1/2)$ and to $2Z_w^{-1} w_2(2u - 1)$ for all $u \in [1/2, 1)$ which is continuous on $(0, 1)$, bounded, and is positive on its support $(a/2, (b + 1)/2)$ hence Assumption 11 is satisfied.

Now we define the new asymptotic mean,

$$\widetilde{G}^w(t) := 1 - \widetilde{F}_{\text{test, cal}} \circ (\widetilde{F}_{\text{cal}}^w)^{-1}(1 - t), \quad t \in (0, 1), \quad (61)$$

the c.d.f. of the theoretical p -values $p^{w, (+\infty)} = 1 - \widetilde{F}_{\text{cal}}^w(T)$. We introduce the new quantities involved in the variance:

$$\widetilde{V}_{\text{cal}}^w(t) := \frac{\int_0^t w(F_{\text{cal}-}^{-1}(u))^2 \, du}{\int_0^1 w(F_{\text{cal}-}^{-1}(u))^2 \, du}, \quad t \in [0, 1]; \quad (62)$$

$$\widetilde{I}^w(t) := 1 - \widetilde{V}_{\text{cal}}^w \circ (\widetilde{F}_{\text{cal}}^w)^{-1}(1-t), \quad t \in (0, 1), \quad (63)$$

which are the quantities replacing V_{cal}^w (19) and I^w (18) in Lemma C.1 and Proposition C.2 with the parameters $(F_{\text{cal}}, F_{\text{test}}, w)$ equal to $(I, \widetilde{F}_{\text{test, cal}}, w \circ F_{\text{cal}-}^{-1})$.

Theorem G.4. *Under Assumptions 9 and 11 and assuming that $n/(n+m)$ tends to $\sigma^2 \in [0, 1]$, we have*

$$\sqrt{\tau_{n,m}} \left(\text{FCP}_m^{(n)} - \widetilde{G}^w \right) \xrightarrow[\tau_{n,m} \rightarrow +\infty]{\mathcal{L}} \sigma \mathbb{U} \circ \widetilde{G}^w + \sqrt{1 - \sigma^2} \rho^w \left(\widetilde{G}^w \right)' \left(\mathbb{V} \left(\widetilde{I}^w \right) + \left[I - \widetilde{I}^w \right] N \right) \text{ on } D(0, 1),$$

where \mathbb{U}, \mathbb{V} are two independent standard Brownian bridges and N is an independent standard Gaussian random variable.

Theorem G.4 is a corollary of Theorem 3.2 after using the general standardisation Lemma G.3.