The Logical Consistency of Simultaneous Agnostic Hypothesis Tests

Luís G. Esteves [1], Rafael Izbicki [2], Julio M. Stern [1] and Rafael B. Stern [2,*]

[1] Institute of Mathematics and Statistics, University of São Paulo, São Paulo 13565-905, Brazil; lesteves@ime.usp.br (L.G.E.); jstern@ime.usp.br (J.M.S.)

[2] Department of Statistics, Federal University of São Carlos, São Carlos 05508-090, Brazil; rafaelizbicki@gmail.com

* Correspondence: rbstern@gmail.com; Tel.: +55-16-3351-8241

Abstract: Simultaneous hypothesis tests can fail to provide results that meet logical requirements. For example, if A and B are two statements such that A implies B, there exist tests that, based on the same data, reject B but not A. Such outcomes are generally inconvenient to statisticians (who want to communicate the results to practitioners in a simple fashion) and non-statisticians (confused by conflicting pieces of information). Based on this inconvenience, one might want to use tests that satisfy logical requirements. However, Izbicki and Esteves shows that the only tests that are in accordance with three logical requirements (monotonicity, invertibility and consonance) are trivial tests based on point estimation, which generally lack statistical optimality. As a possible solution to this dilemma, this paper adapts the above logical requirements to agnostic tests, in which one can accept, reject or remain agnostic with respect to a given hypothesis. Each of the logical requirements is characterized in terms of a Bayesian decision theoretic perspective. Contrary to the results obtained for regular hypothesis tests, there exist agnostic tests that satisfy all logical requirements and also perform well statistically. In particular, agnostic tests that fulfill all logical requirements are characterized as region estimator-based tests. Examples of such tests are provided.

Keywords: agnostic tests; multiple hypothesis testing; logical consistency; decision theory; loss functions

1. Introduction

One of the practical shortcomings of simultaneous test procedures is that they can lack logical consistency [1,2]. As a result, recent papers have discussed minimum logical requirements and methods that achieve these requirements [3–7]. For example, it has been argued that simultaneous tests ought to be in agreement with the following criterion: if hypothesis A implies hypothesis B, a procedure that rejects B should also reject A.

In particular, Izbicki and Esteves [3] and da Silva et al. [7] examine classical and bayesian simultaneous tests with respect to four consistency properties:

- Monotonicity: if A implies B, then a test that does not reject A should not reject B.
- Invertibility: A test should reject A if and only if it does not reject not-A.
- Union consonance: If a test rejects A and B, then it should reject $A \cup B$.
- Intersection consonance: If a test does not reject A and does not reject B, then it should not reject $A \cap B$.

Izbicki and Esteves [3] prove that the only tests that are fully coherent are trivial tests based on point estimation, which are generally void of statistical optimality. This finding suggests that alternatives to the standard "reject versus accept" tests should be explored.

Such an alternative are agnostic tests [8], which can take the following decisions: (i) accept an hypothesis (decision 0); (ii) reject it (decision 1); or (iii) noncommittally neither accept or reject it; thus abstaining or remaining agnostic about the other two actions (decision $\frac{1}{2}$). Decision (iii) is also called a *no-decision classification*. The set of samples, $x \in \mathcal{X}$, for which one abstains from making a decision about a given hypothesis is called a no-decision region [8]. An agnostic test enables one to explicitly deal with the difference between "accepting a hypothesis H" and "not rejecting H (remaining agnostic)". This distinction will be made clearer in Section 5, which derives agnostic tests under a Bayesian decision-theoretic standpoint by means of specific penalties for false rejection, false acceptance and excessive abstinence.

We use the above framework to revisit the logical consistency of simultaneous hypothesis tests. Section 2 defines *agnostic testing scheme* (ATS), a transformation that assigns to each statistical hypothesis an agnostic test function. This definition is illustrated with bayesian and frequentist examples, using both existing and novel agnostic tests. Section 3 generalizes the logical requirements in [3] to agnostic testing schemes. Section 4 presents tests that satisfy all of these logical requirements. Section 5 obtains, under the Bayesian decision-theoretic paradigm, necessary and sufficient conditions on loss functions to ensure that Bayes tests meet each of the logical requirements. All theorems are proved in the Appendix.

2. Agnostic Testing Schemes

This section describes the mathematical setup for agnostic testing schemes. Let $\mathcal{X}$ denote the sample space, Θ the parameter space and $L_x(\theta)$ the likelihood function at the point $\theta \in \Theta$ generated by the data $x \in \mathcal{X}$. We denote by $\mathcal{D} = \{0, \frac{1}{2}, 1\}$ the set of all decisions that can be taken when testing a hypothesis: accept (0), remain agnostic ($\frac{1}{2}$) and reject (1). By an agnostic hypothesis test (or simply agnostic test) we mean a decision function from $\mathcal{X}$ to $\mathcal{D}$ [8,9]. Similar tests are commonly used in machine learning in the context of classification [1,2]. Moreover, let $\Phi = \{\phi : \phi : \mathcal{X} \longrightarrow \mathcal{D}\}$ be the set of all (agnostic) hypothesis tests. The following definition adapts testing schemes [3] to agnostic tests.

Definition 1 (**Agnostic Testing Scheme; ATS**). Let $\sigma(\Theta)$, a σ-field of subsets of the parameter space Θ, be the set of hypotheses to be tested. An *ATS* is a function $\mathcal{L} : \sigma(\Theta) \to \Phi$ that, for each hypothesis $A \in \sigma(\Theta)$, assigns the test $\mathcal{L}(A) \in \Phi$ for testing A.

A way of creating an agnostic testing scheme is to find a collection of statistics and to compare them to thresholds:

Example 1. For every $A \in \sigma(\Theta)$, let $s_A : \mathcal{X} \longrightarrow \mathbb{R}$ be a statistic. Let $c_1, c_2 \in \mathbb{R}$, with $c_1 \geqslant c_2$, be fixed thresholds. For each $A \in \sigma(\Theta)$, one can define $\mathcal{L}(A) : \mathcal{X} \to \mathcal{D}$ by

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } s_A(x) > c_1 \\ \frac{1}{2} & \text{if } c_1 \geqslant s_A(x) > c_2 \\ 1 & \text{if } c_2 \geqslant s_A(x) \end{cases}$$

The ATS in Example 1 rejects a hypothesis if the value of the statistic s_A is small, accepts it if this value is large, and remains agnostic otherwise. If $s_A(x)$ is a measure of how much evidence that x brings about A, then this ATS rejects a hypothesis if the evidence brought by the data is small, accepts it if this evidence is large, and remains agnostic otherwise. The next examples present particular cases of this ATS These examples will be explored in the following sections.

Example 2 (ATS based on posterior probabilities). Let $\Theta = \mathbb{R}^d$ and $\sigma(\Theta) = \mathcal{B}(\Theta)$, the Borelians of $\mathbb{R}^d$. Assume that a prior probability $\mathbb{P}$ in $\sigma(\Theta)$ is fixed, and let $c_1, c_2 \in (0, 1)$, with $c_1 \geqslant c_2$, be fixed thresholds. For each $A \in \sigma(\Theta)$, let $\mathcal{L}(A) : \mathcal{X} \to \mathcal{D}$ be defined by

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } \mathbb{P}(A|x) > c_1 \\ \frac{1}{2} & \text{if } c_1 \geqslant \mathbb{P}(A|x) > c_2 \\ 1 & \text{if } c_2 \geqslant \mathbb{P}(A|x) \end{cases}$$

where $\mathbb{P}(.|x)$ is the posterior distribution of θ, given x. This is essentially the test that Ripley [10] proposed in the context of classification, which was also investigated by Babb et al. [9]. When $c_1 = c_2$, this ATS is a standard (non-agnostic) Bayesian testing scheme.

Example 3 (Likelihood Ratio Tests with fixed threshold). Let $\Theta = \mathbb{R}^d$ and $\sigma(\Theta) = \mathcal{P}(\Theta)$, the set of the parts of $\mathbb{R}^d$. Let $c_1, c_2 \in (0, 1)$, with $c_1 \geqslant c_2$, be fixed thresholds. For each $A \in \sigma(\Theta)$, let $\lambda_x(A) = \frac{\sup_{\theta \in A} L_x(\theta)}{\sup_{\theta \in \Theta} L_x(\theta)}$ be the likelihood ratio statistic for sample $x \in \mathcal{X}$. Define $\mathcal{L}$ by

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } \lambda_x(A) > c_1 \\ \frac{1}{2} & \text{if } c_1 \geqslant \lambda_x(A) > c_2 \\ 1 & \text{if } c_2 \geqslant \lambda_x(A) \end{cases}$$

When $c_1 = c_2$, this is the standard likelihood ratio with fixed threshold (non-agnostic) testing scheme [3].

A similar test to that of Example 3 is developed by Berg [8]; however, the values of the cutoffs c_1 and c_2 are allowed to change with the hypothesis of interest, and they are chosen so as to control the level of significance and the power of each of the tests.

Example 4 (FBST ATS). Let $\Theta = \mathbb{R}^d$, $\sigma(\Theta) = \mathcal{B}(\mathbb{R}^d)$, and $f(\theta)$ be the prior probability density function (p.d.f.) for θ. Suppose that, for each $x \in \mathcal{X}$, there exists $f(\theta|x)$, the p.d.f. of the posterior distribution of θ, given x. For each hypothesis $A \in \sigma(\Theta)$, let

$$T_x^A = \left\{ \boldsymbol{\theta} \in \Theta : f(\boldsymbol{\theta}|\boldsymbol{x}) > \sup_{\theta \in A} f(\boldsymbol{\theta}|\boldsymbol{x}) \right\}$$

be the set tangent to the null hypothesis and let $ev_x(A) = 1 - P(\theta \in T_x^A | x)$ be the Pereira–Stern evidence value for A [11]. Let $c_1, c_2 \in (0, 1)$, with $c_1 \geqslant c_2$, be fixed thresholds. One can define an ATS $\mathcal{L}$ by

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } ev_x(A) > c_1 \\ \frac{1}{2} & \text{if } c_1 \geqslant ev_x(A) > c_2 \\ 1 & \text{if } c_2 \geqslant ev_x(A) \end{cases}$$

When $c_1 = c_2$, this ATS reduces to the standard (non-agnostic) FBST testing scheme [3].

The following example presents a novel ATS based on region estimators.

Example 5 (Region Estimator-based ATS). Let $R : \mathcal{X} \longrightarrow \mathcal{P}(\Theta)$ be a region estimator of θ. For every $A \in \sigma(\Theta)$ and $x \in \mathcal{X}$, one can define an ATS $\mathcal{L}$ via

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } R(x) \subseteq A \\ 1 & \text{if } R(x) \subseteq A^c \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

Hence, $\mathcal{L}(A)(x) = \frac{\mathbb{I}(R(x) \subseteq A^c) + \mathbb{I}(R(x) \nsubseteq A)}{2}$. See Figure 1 for an illustration of this procedure.
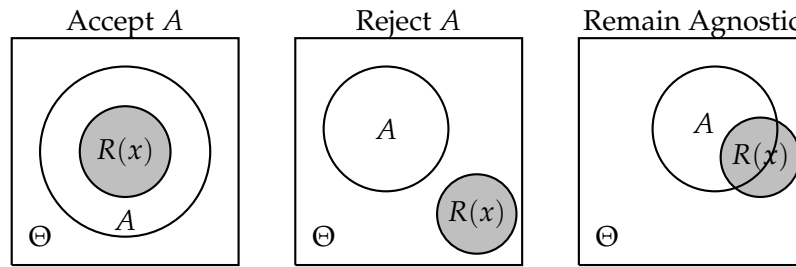

Figure 1. Agnostic test based on the region estimate $R(x)$ from Example 5.

Notice that for continuous Θ, Example 5 does not accept precise (i.e., null Lebesgue measure) hypotheses, yielding either rejection or abstinence (unless region estimates are themselves precise). Therefore, the performance of region estimator-based ATS's is in agreement with the prevailing position among both bayesian and frequentist statisticians: to accept a precise hypothesis is inappropriate. From a Bayesian perspective, precise null hypothesis usually have zero posterior probabilities, and thus should not be accepted. From a Frequentist perspective, not rejecting a hypothesis is not the same as accepting it. See Berger and Delampady [12] and references therein for a detailed account on the controversial problem of testing precise hypotheses.

In principle, R can be any region estimator. However, some choices of R lead to better statistical performance. For example, from a frequentist, one might choose R to be a confidence region. This choice is explored in the next example.

Example 6. From a frequentist perspective, one might choose R in Example 5 to be a confidence region: if the region estimator has confidence at least $1 - \alpha$, then type I error probability, $\sup_{\theta \in A} \mathbb{P}(\mathcal{L}(A)(X) = 1 | \theta)$, is smaller than α for each of the hypothesis tests. Indeed,

$$\sup_{\theta \in A} \mathbb{P}(\mathcal{L}(A)(X) = 1 | \theta) = \sup_{\theta \in A} \mathbb{P}(\theta' \notin R(X) \text{ for every } \theta' \in A | \theta) \leqslant \sup_{\theta \in A} \mathbb{P}(\theta \notin R(X) | \theta) \leqslant \alpha.$$

If R is a confidence region, then this ATS also controls the Family Wise Error Rate (FWER, [13]), as shown in Section 3.1.

Consider $X_1, \ldots, X_{20} | \mu \overset{i.i.d.}{\sim}$ Normal$(\mu, 1)$. In Figure 2, we illustrate how the probability of each decision, $\mathbb{P}(\mathcal{L}(A)(X) = d | \mu)$ for $d \in \{0, \frac{1}{2}, 1\}$, varies as a function of μ for three hypotheses: (i) $\mu < 0$; (ii) $\mu = 0$; and (iii) $0 < \mu < 1$. We consider the standard region estimator for μ, $R(X) = [\bar{X} - z_{1-\alpha/2} \frac{1}{\sqrt{n}}; \bar{X} + z_{1-\alpha/2} \frac{1}{\sqrt{n}}]$ with $\alpha = 5\%$. These curves represent the generalization of the standard power function to agnostic hypothesis tests. Notice that $\mu = 0$ is never accepted, and that, under the null hypothesis, all tests have at most 5% of probability of rejecting H.
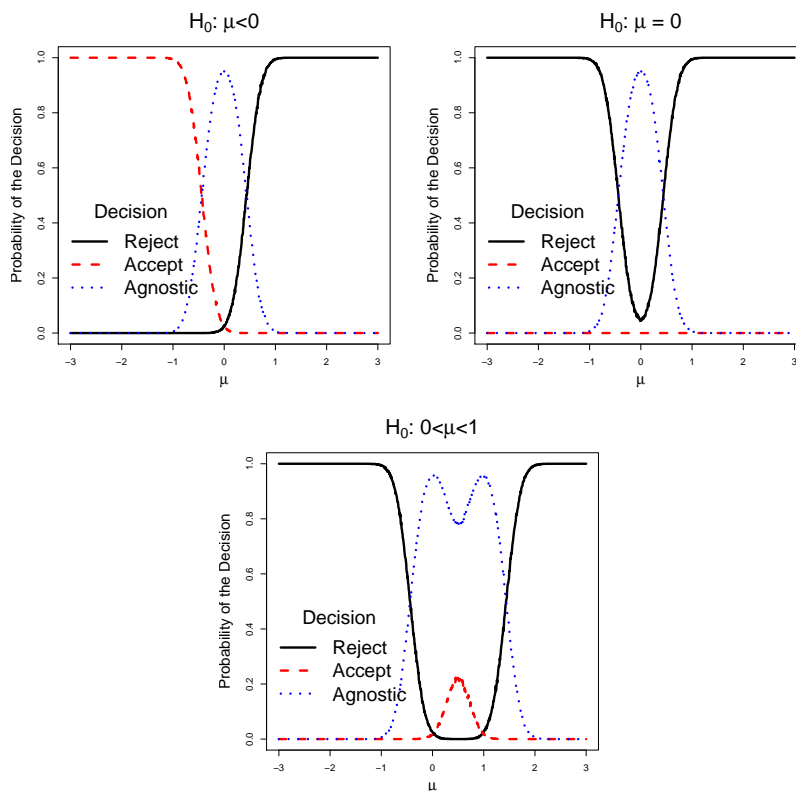
Figure 2. Illustrations of the performance of the agnostic region testing scheme (Example 5) for three different hypotheses (specified on the top of each picture). The pictures present the probability of each decision, $\mathbb{P}(\mathcal{L}(A)(X) = d|\mu)$ for $d \in \{0, \frac{1}{2}, 1\}$, as a function of the mean, μ.

The next two examples show other cases of the ATS in Example 5 that use region estimators based on the measures of evidence in Examples 3 and 4.

Example 7 (Region Likelihood Ratio ATS). For a fixed value $c \in (0, 1)$, define the region estimate $R_c(x) = \{\theta \in \Theta : \lambda_x(\{\theta\}) \geqslant c\}$, where λ_x is the likelihood ratio statistics from Example 3. For every $A \in \sigma(\Theta)$ and $x \in \mathcal{X}$, the ATS based on this region estimator (Example 5) satisfies $\mathcal{L}(A)(x) = 1 \iff A \bigcap R_c(x) = \varnothing$, and $\mathcal{L}(A)(x) = 0 \iff A^c \bigcap R_c(x) = \varnothing$. It follows that this ATS can be written as

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } \lambda_x(A^c) < c \\ 1 & \text{if } \lambda_x(A) < c \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

Example 8 (Region FBST ATS). For a fixed value of $c \in (0, 1)$, let HPD_c^x be the Highest Posterior Probability Density region with probability $1 - c$, based on observation x [3,14]. For every $A \in \sigma(\Theta)$ and $x \in \mathcal{X}$, the ATS based on this region estimator (Example 5) satisfies $\mathcal{L}(A)(x) = 1 \iff A \bigcap \text{HPD}_c^x = \varnothing$, and $\mathcal{L}(A)(x) = 0 \iff A^c \bigcap \text{HPD}_c^x = \varnothing$. It follows that this ATS can be written as

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } ev_x(A^c) < c \\ 1 & \text{if } ev_x(A) < c \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

In the sequence, we introduce four logical coherence properties for agnostic testing schemes and investigate which tests satisfy them.

3. Coherence Properties

3.1. Monotonicity

Monotonicity restricts the decisions that are available for nested hypotheses. If hypothesis A *implies* hypothesis B (i.e., $A \subseteq B$), then a testing scheme that rejects B should also reject A. Monotonicity has received a lot of attention in the literature (e.g., [5,6,15–19]). It can be extended to ATS's in the following way.

Definition 2 (Monotonicity). $\mathcal{L} : \sigma(\Theta) \to \Phi$ is monotonic if, for every $A, B \in \sigma(\Theta)$, $A \subset B$ implies that $\mathcal{L}(A) \geqslant \mathcal{L}(B)$.

$\mathcal{L}$ is monotonic if, for every hypotheses $A \subset B$,

1.　if $\mathcal{L}$ accepts A, then it also accepts B.
2.　if $\mathcal{L}$ remains agnostic about A, then it either remains agnostic about B or accepts B.

Next, we illustrate some monotonic agnostic testing schemes.

Example 9 (Tests based on posterior probabilities). The ATS from Example 2 is monotonic. Indeed, $A \subset B$ implies that $\mathbb{P}(A|x) \leqslant \mathbb{P}(B|x) \ \forall x \in \mathcal{X}$, and hence $\mathbb{P}(A|x) > c_i$ implies that $\mathbb{P}(B|x) > c_i$ for $i = 1, 2$.

Example 10 (Likelihood Ratio Tests with fixed threshold). The ATS from Example 3 is monotonic. This is because if $A, B \in \sigma(\Theta)$ are such that $A \subset B$, then $\sup_{\theta \in A} L_x(\theta) \leqslant \sup_{\theta \in B} L_x(\theta)$, $\forall x \in \mathcal{X}$, which implies that $\lambda_x(A) \leqslant \lambda_x(B)$. It follows that $\lambda_x(A) > c_i$ implies that $\lambda_x(B) > c_i$ for $i = 1, 2$.

Example 11 (FBST). The ATS from Example 4 is monotonic. In fact, let $A, B \in \sigma(\Theta)$ be such that $A \subset B$. We have $\sup_B f(\theta|x) \geqslant \sup_A f(\theta|x) \ \forall x \in \mathcal{X}$. Hence, $T_x^B \subseteq T_x^A$, and, therefore, $ev_x(A) \leqslant ev_x(B)$. It follows that $ev_x(A) > c_i$ implies that $ev_x(B) > c_i$ for $i = 1, 2$.

Notice that *p*-values and Bayes factors are *not* (coherent) measures of support for hypotheses [19,20], and therefore using them in a similar fashion as in Examples 2–4 would not lead to monotonic agnostic testing schemes. On the other hand, any monotonic statistic s_A does, however, provide a monotonic ATS, because, if $A \subseteq B$, $s_A(x) > c_i$ implies that $s_B(x) > c_i$ for $i = 1, 2$. Another example of such statistic is the s-value defined by Patriota [6]. As a matter of fact, every ATS is, in a sense, associated to monotonic statistics as shown in the next theorem.

Theorem 1. *Let $\mathcal{L}$ be an agnostic testing scheme. $\mathcal{L}$ is monotonic if, and only if, there exist a sequence of test statistics $(s_A)_{A \in \sigma(\Theta)}$, $s_A : \mathcal{X} \longrightarrow I \subseteq \mathbb{R}$, with $s_A \leqslant s_B$ whenever $A \subset B$, $A, B \in \sigma(\Theta)$, and cutoffs $c_1, c_2 \in I$, $c_1 \geqslant c_2$, such that for every $A \in \sigma(\Theta)$ and $x \in \mathcal{X}$,*

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } s_A(x) > c_1 \\ \frac{1}{2} & \text{if } c_1 \geqslant s_A(x) > c_2 \\ 1 & \text{if } c_2 \geqslant s_A(x) \end{cases} \tag{1}$$

Example 12 (Region Estimator). The ATS from Example 5 is monotonic, because if $A \subseteq B$, $A, B \in \sigma(\Theta)$, then

$$\mathcal{L}(A)(x) = \frac{\mathbb{I}(R(x) \subseteq A^c) + \mathbb{I}(R(x) \nsubseteq A)}{2} \geqslant \frac{\mathbb{I}(R(x) \subseteq B^c) + \mathbb{I}(R(x) \nsubseteq B)}{2} = \mathcal{L}(B)(x),$$

as $\mathbb{I}(R(x) \subseteq B^c) \leqslant \mathbb{I}(R(x) \subseteq A^c)$ and $\mathbb{I}(R(x) \nsubseteq B) \leqslant \mathbb{I}(R(x) \nsubseteq A)$. Because this ATS is monotonic, it also controls Family Wise Error Rate [21].

3.2. Union Consonance

Finner and Strassburger [4] and Izbicki and Esteves [3] investigated the following logical property, named union consonance: if a (non-agnostic) testing scheme rejects each of the hypotheses A and B, it should also reject their union $A \cup B$. In other words, a TS cannot accept the union while rejecting its components. In this section, we adapt the concept of union consonance to the framework of agnostic testing schemes by considering two extensions for such desideratum: the weak and the strong union consonance.

Definition 3 (**Weak Union Consonance**). An ATS $\mathcal{L} : \sigma(\Theta) \to \Phi$ is weakly consonant with the union if, for every $A, B \in \sigma(\Theta)$, and for every $x \in \mathcal{X}$,

$$\mathcal{L}(A)(x) = 1 \text{ and } \mathcal{L}(B)(x) = 1 \text{ implies } \mathcal{L}(A \cup B)(x) \neq 0$$

This is exactly the definition of union consonance for non-agnostic testing schemes. Notice that, according to such definition, it is possible to remain agnostic about $A \cup B$ while rejecting A and B.

Remark 1. Izbicki and Esteves [3] show that if a non agnostic testing scheme $\mathcal{L}$ satisfies union consonance, then for every *finite* set of indices I and for every $\{A_i\}_{i \in I} \subseteq \sigma(\Theta), \min\{\mathcal{L}(A_i)\}_{i \in I} = 1$ implies that $\mathcal{L}(\cup_{i \in I} A_i) \neq 0$. This is not the case for weak union consonant agnostic testing schemes; we leave further details to Section 4.3.

The second definition of union consonance is more stringent than the first one:

Definition 4 (**Strong Union Consonance**). An $\mathcal{L} : \sigma(\Theta) \to \Phi$ is strongly consonant with the union if, for every arbitrary set of indices I and for every $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ such that $\cup_{i \in I} A_i \in \sigma(\Theta)$, and for every $x \in \mathcal{X}$,

$$\min\{\mathcal{L}(A_i)(x)\}_{i \in I} = 1 \text{ implies } \mathcal{L}(\cup_{i \in I} A_i)(x) = 1$$

Definition 3 is less stringent than Definition 4 in two senses: (i) the latter imposes the (strict) rejection of a union of hypotheses whenever each of them is rejected while the former imposes just non-acceptance (rejection or abstention) of the union is such circumstances; and (ii) in Definition 4 consonance is required to hold for every set (possibly infinite) of hypotheses as opposed to Definition 3 which applies only to pairs of hypotheses. Notice that if an ATS is strongly consonant with union, it is also weakly consonant with union, and that both definitions are indeed extensions of the concept presented by Izbicki and Esteves [3].

The following examples show ATSs that are consonant with union.

Example 13 (**Tests based on posterior probabilities**). Consider again Example 2 with the restriction $c_1 \geqslant 2c_2$. If A and B are rejected after observing $x \in \mathcal{X}$, then

$$\mathbb{P}(A \cup B | x) \leqslant \mathbb{P}(A | x) + \mathbb{P}(B | x) \leqslant 2c_2 \leqslant c_1,$$

and therefore $A \cup B$ cannot be accepted. Thus, with this restrictions, that ATS is *weakly* consonant with union. The restriction $c_1 \geqslant 2c_2$ is not only sufficient to ensure weak union consonance, but it is actually necessary to ensure it holds for every prior distribution (see Theorem 2). Notice, however, that this ATS is not strongly consonant with union in general.

Example 14 (**Likelihood Ratio Tests with fixed threshold**). The ATS of Example 3 is strongly consonant with union. Indeed, let I be an arbitrary set of indices and $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ be such that $\cup_{i \in I} A_i \in \sigma(\Theta)$. For every $x \in \mathcal{X}$, $\lambda_x(\cup_{i \in I} A_i) = \sup_{i \in I}\{\lambda_x(A_i)\}$ [3]. It follows that if $\lambda_x(A_i) \leqslant c_2$ for

every $i \in I$, then $\lambda_x(\cup_{i \in I} A_i) \leqslant c_2$. Thus, if $\mathcal{L}$ rejects all hypotheses A_i after x is observed, it also rejects $\cup_{i \in I} A_i$. In addition, $\mathcal{L}$ is also weakly consonant with union.

Example 15 (FBST). The ATS from Example 4 is also strongly consonant with union. Indeed, let I be an arbitrary set of indices and $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ be such that $\cup_{i \in I} A_i \in \sigma(\Theta)$. For every $x \in \mathcal{X}$, $ev_x(\cup_{i \in I} A_i) = \sup_{i \in I}\{ev_x(A_i)\}$ [22]. Strong union consonance holds due to the same argument from Example 14. It follows that $\mathcal{L}$ is also weakly consonant with union.

Example 16 (Region Estimator). The TS from Example 5 satisfies strong union consonance. Indeed, let I be an arbitrary set of indices and $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ be such that $\cup_{i \in I} A_i \in \sigma(\Theta)$. If $\mathcal{L}(A_i)(x) = 1$, then $R(x) \subseteq A_i^c$. Hence, if $\mathcal{L}(A_i)(x) = 1$ for every $i \in I$, $R(x) \subseteq \bigcap_{i \in I} A_i^c = (\bigcup_{i \in I} A_i)^c$, and, therefore, $\bigcup_{i \in I} A_i$ is rejected. It follows that $\mathcal{L}$ is also weakly consonant with union.

3.3. Intersection Consonance

The third property we investigate, named intersection consonance [3], states that if a (non agnostic) testing scheme cannot accept hypotheses A and B while rejecting its intersection. We consider two extensions of such definition to agnostic testing schemes.

Definition 5 (Weak Intersection Consonance). An ATS $\mathcal{L} : \sigma(\Theta) \to \Phi$ is consonant with the intersection if, for every $A, B \in \sigma(\Theta)$ and $x \in \mathcal{X}$,

$$\mathcal{L}(A)(x) = 0 \text{ and } \mathcal{L}(B)(x) = 0 \text{ implies } \mathcal{L}(A \cap B)(x) \neq 1.$$

This is exactly the definition of intersection consonance for non-agnostic testing schemes. Notice that it is possible to accept A and B while being agnostic about $A \cap B$.

The second definition of intersection consonance is more stringent:

Definition 6 (Strong Intersection Consonance). An ATS $\mathcal{L} : \sigma(\Theta) \to \Phi$ is strongly consonant with the intersection if, for every arbitrary set of indices I and for every $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ such that $\cap_{i \in I} A_i \in \sigma(\Theta)$, and for every $x \in \mathcal{X}$,
$$\max\{\mathcal{L}(A_i)(x)\}_{i \in I} = 0 \text{ implies } \mathcal{L}(\cap_{i \in I} A_i)(x) = 0.$$

As in the case of union consonance, Definition 5 is less stringent than Definition 6 in two senses: (i) the latter imposes the (strict) acceptance of an intersection of hypotheses whenever each of them is accepted while the former imposes just non-rejection (acceptance or abstention) of the intersection is such circumstances; and (ii) in Definition 6 consonance is required to hold for every set (possibly infinite) of hypotheses as opposed to Definition 5 which applies only to pairs of hypotheses. Notice that if an ATS is strongly consonant with intersection, it is also weakly consonant with intersection, and that both definitions are indeed extensions of the concept presented by Izbicki and Esteves [3].

Example 17 (Tests based on posterior probabilities). Consider Example 2 with the restriction $c_2 \leqslant 2c_1 - 1$. If A and B are accepted when $x \in \mathcal{X}$ is sampled, then $\mathbb{P}(A|x) > c_1$ and $\mathbb{P}(B|x) > c_1$. By Fréchet inequality, it follows that

$$\mathbb{P}(A \cap B|x) \geqslant \mathbb{P}(A|x) + \mathbb{P}(B|x) - 1 > 2c_1 - 1 \geqslant c_2$$

and, therefore, $A \cap B$ cannot be rejected. It follows that weak intersection consonance holds. The restriction $c_2 \leqslant 2c_1 - 1$ is not only sufficient to ensure weak intersection consonance, but it is actually necessary to ensure this property holds for every prior distribution; see Theorem 2.

Notice, however, that this ATS is not strongly consonant with intersection in general (Take, for example, $\Theta = [0,1]$, $\mathbb{P} = \lambda$ (the Lebesgue measure), $A = [0, 2/3]$, $B = [1/3, 1]$, and $c_1 = 3/5$).

The ATS based on the likelihood ratio statistic from Example 3 does not satisfy intersection consonance, because there are examples in which $\lambda_x(A \cap B) = 0$, while $\lambda_x(A) > 0$ and $\lambda_x(B) > 0$ (Consider, for example, that every $\theta \in \Theta$ has the same likelihood and $A \cap B = \varnothing$). Similarly, the ATS based on FBST from Example 4 is not consonant with intersection, because there are examples such that $ev_x(A \cap B) = 0$, while $ev_x(A) > 0$ and $ev_x(B) > 0$. ATSs based on region estimators are consonant with intersection.

Example 18 (**Region Estimator**). The TS from Example 5 satisfies both strong and weak intersection consonance. Indeed, let I be an arbitrary set of indices and $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ be such that $\cap_{i \in I} A_i \in \sigma(\Theta)$. If $\mathcal{L}(A_i)(x) = 0$ for every $i \in I$, then $R(x) \subseteq A_i$ for every $i \in I$. It follows that $R(x) \subseteq \cap_{i \in I} A_i$, and hence $\cap_{i \in I} A_i$ is accepted.

It follows that the ATSs from Examples 7 and 8 are also consonant with intersection. Hence, it is possible to use e-values and likelihood ratio statistics to define ATS that are consonant with intersection.

Example 19 (**ANOVA**). In [3], the authors present an example which we now revisit. Suppose that $X_1, \ldots, X_{20}$ are i.i.d. $N(\mu_1, \sigma^2)$; $X_{21}, \ldots, X_{40}$ are i.i.d. $N(\mu_2, \sigma^2)$ and $X_{41}, \ldots, X_{60}$ are i.i.d. $N(\mu_3, \sigma^2)$. Consider the following hypotheses:

$$H_0^{(1,2,3)} : \mu_1 = \mu_2 = \mu_3 \qquad H_0^{(1,2)} : \mu_1 = \mu_2 \qquad H_0^{(1,3)} : \mu_1 = \mu_3$$

and suppose that we observe the following means and standard-deviations on the data: $\overline{X}_1 = 0.15$; $S_1 = 1.09$; $\overline{X}_2 = -0.13$; $S_2 = 0.5$ $\overline{X}_3 = -0.38$; $S_3 = 0.79$. Using the likelihood ratio statistics, we have the following *p*-values for these hypotheses:

$$p_{H_0^{(1,2,3)}} = 0.0498 \qquad p_{H_0^{(1,2)}} = 0.2564 \qquad p_{H_0^{(1,3)}} = 0.0920$$

Therefore, the testing scheme given by the likelihood ratio tests with common level of significance $\alpha = 5\%$ rejects $H_0^{(1,2,3)}$ but does not reject either $H_0^{(1,2)}$ or $H_0^{(1,3)}$. It follows intersection consonance does not hold. Now, consider the region estimator ATS based on the region estimate given by [23] for this setting,

$$R(x) = \left\{ (\mu_1, \mu_2, \mu_3) \in \mathbb{R}^3 : \mu_1 - \mu_2 \in [-1.65, 2.21], \mu_2 - \mu_3 \in [-1.68, 2.18], \mu_1 - \mu_3 \in [-1.40, 2.46] \right\}$$

All hypotheses $H_0^{(1,2,3)}$, $H_0^{(1,2)}$, and $H_0^{(1,3)}$ intercept both $R(x)$ and its complement, so that one remains agnostic about all of them. As expected, intersection consonance holds using this ATS.

3.4. Invertibility

Invertibility formalizes the notion of simultaneous tests free from the labels "null" and "alternative" for the hypotheses of interest and has been suggested by several authors, specially under a Bayesian perspective [3,24,25].

Definition 7 (**Invertibility**). An ATS $\mathcal{L} : \sigma(\Theta) \to \Phi$ is invertible if, for every $A \in \sigma(\Theta)$,

$$\mathcal{L}(A^c) = 1 - \mathcal{L}(A)$$

Example 20 (**Tests based on posterior probabilities**). The ATS from Example 2 is invertible for every prior distribution if and only if $c_2 = 1 - c_1$.

Example 21 (**Region Estimator**). The ATS from Example 5 is invertible. Indeed,

$$\mathcal{L}(A)(x) = \frac{\mathbb{I}(R(x) \subseteq A^c) + \mathbb{I}(R(x) \nsubseteq A))}{2} = \frac{(1 - \mathbb{I}(R(x) \nsubseteq A^c)) + (1 - \mathbb{I}(R(x) \subseteq A))}{2} = 1 - \mathcal{L}(A^c)(x)$$

It follows that the ATS from Examples 7 and 8 are also invertible.

4. Satisfying All Properties

Is it possible to construct non-trivial agnostic testing schemes that satisfy all consistency properties simultaneously? Contrary to the case of non agnostic testing schemes [3], the answer is yes. We next examine this question considering three desiderata: the weak desiderata (Section 4.1), the strong desiderata (Section 4.2), and the *n*-weak desiderata (Section 4.3).

4.1. Weak Desiderata

Definition 8 (**Weakly Consistent ATS**). An ATS, $\mathcal{L}$, is said to be weakly consistent if $\mathcal{L}$ is monotonic (Definition 2), invertible (Definition 7), weakly consonant with the union (Definition 3), and weakly consonant with the intersection (Definition 5).

Example 22 (**Region Estimator**). The ATS from Example 5 was already shown to satisfy all consistency properties from Definition 8 (Examples 12, 16, 18 and 21). Thus, it is a weakly consistent ATS.

It follows that the ATSs from Examples 7 and 8, based on measures of support (likelihood ratio statistics and e-values), are weakly consistent ATSs.

Example 23 (**Tests based on posterior probabilities**). Consider Example 2. We have seen that the following restrictions are sufficient to guarantee union weak consonance (Example 13), weak intersection consonance (Example 17) and invertibility (Example 20), respectively: $c_1 \geqslant 2c_2$, $2c_1 - 1 \geqslant c_2$ and $c_2 = 1 - c_1$. It follows from these relations and the fact that this ATS is monotonic (Example 9) that if $c_1 > 2/3$ and $c_2 = 1 - c_1$, then it is weakly consistent, whatever the prior distribution for θ is.

The next theorem shows necessary and sufficient conditions for agnostic tests based on posterior distribution (with possibly different thresholds c_1 and c_2 for each hypothesis of interest) to satisfy each of the coherence properties.

Theorem 2. *Let $\Theta = \mathbb{R}^d$ and $\sigma(\Theta) = \mathcal{B}(\Theta)$, the Borelians of $\mathbb{R}^d$. Let $\mathbb{P}$ be a prior probability measure in $\sigma(\Theta)$. For each $A \in \sigma(\Theta)$, let $\mathcal{L}(A) : \mathcal{X} \to \mathcal{D}$ be defined by*

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } \mathbb{P}(A|x) > c_1^A \\ \frac{1}{2} & \text{if } c_1^A \geqslant \mathbb{P}(A|x) > c_2^A \\ 1 & \text{if } c_2^A \geqslant \mathbb{P}(A|x) \end{cases}$$

where $\mathbb{P}(.|x)$ is the posterior distribution of θ, given x, and $0 \leqslant c_2^A \leqslant c_1^A \leqslant 1$. This is a generalization of the ATS of Example 2. Assume that the likelihood function is positive for every $x \in \mathcal{X}$ and $\theta \in \Theta$. Such ATS satisfies:

1. *Monotonicity for every prior distribution if, and only if, for every $A, B \in \sigma(\Theta)$ with $A \subseteq B$, $c_2^A \geqslant c_2^B$ and $c_1^A \geqslant c_1^B$*
2. *Weak union consonance for every prior distribution if, and only if, for every $A, B \in \sigma(\Theta)$ such that $A \neq B$, $c_2^A + c_2^B \leqslant c_1^{A \cup B}$*

3. *Weak intersection consonance for every prior distribution if, and only if, for every $A, B \in \sigma(\Theta)$ such that $A \neq B$, $c_1^A + c_1^B - 1 \geqslant c_2^{A \cap B}$*
4. *Invertibility for every prior distribution if, and only if, for every $A \in \sigma(\Theta)$, $c_1^A = 1 - c_2^{A^c}$*

It follows from Theorem 2 that if the cutoffs used in each of the tests (c_1 and c_2) are required to be the same for all hypothesis of interest, then the conditions in Example 23 are not only sufficient, but they are also necessary to ensure that all (weak) consistency properties hold for every prior distribution for θ.

4.2. Strong Desiderata

Definition 9 (**Fully Consistent ATS**). An ATS, $\mathcal{L}$, is said to be fully consistent if $\mathcal{L}$ is monotonic (Definition 2), invertible (Definition 7), strongly consonant with the union (Definition 4), and strongly consonant with the intersection (Definition 6).

The following theorem shows that, under mild assumptions, the only ATSs that are fully consistent are those based on region estimators.

Theorem 3. *Assume that for every $\theta \in \Theta$, $\{\theta\} \in \sigma(\Theta)$. An ATS is fully consistent if, and only if, it is a region estimator-based ATS (Example 5).*

Hence, the only way to create a fully consistent ATS is by designing an appropriate region estimator and using Example 5. In particular, ATSs based on posterior probabilities (Example 2) are typically not fully consistent. It should be emphasized that when the region estimator that characterizes a fully consistent ATS $\mathcal{L}$ maps $\mathcal{X}$ to singletons of Θ, no sample point will lead to abstention, as either $R(x) \subseteq A$ or $R(x) \subseteq A^c$, for every $A \in \sigma(\Theta)$. In such situations, region estimators reduce to point estimator which charaterize full consistent *non-agnostic* TSs [3].

In the next section, we consider a desiderata for simultaneous tests which is not as strong as that of Definition 9, but which is more stringent that that of Definition 8.

4.3. n-Weak Desiderata

In Sections 3.2 and 3.3, weak consonance was defined for two hypotheses only. It is however possible to define it for $n < \infty$ hypotheses:

Definition 10 (**Weak n-union Consonance**). An A-TS $\mathcal{L} : \sigma(\Theta) \to \Phi$ satisfies weak n-union consonant if, for every finite set of indices I, with $|I| \leqslant n$, for every $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$, and for every $x \in \mathcal{X}$

$$\min\{\mathcal{L}(A_i)(x)\}_{i \in I} = 1 \text{ implies } \mathcal{L}(\cup_{i \in I} A_i)(x) \neq 0.$$

Definition 11 (**Weak n-intersection Consonance**). An ATS $\mathcal{L} : \sigma(\Theta) \to \Phi$ is weak n-intersection consonant if, for every finite set of indices I, with $|I| \leqslant n$, for every $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$, and for every $x \in \mathcal{X}$

$$\max\{\mathcal{L}(A_i)(x)\}_{i \in I} = 0 \text{ implies } \mathcal{L}(\cap_{i \in I} A_i)(x) \neq 1.$$

Although in the context of *non agnostic* testing schemes (union or intersection) consonance holds for $n = 2$ if, and only if, it holds for every $n \in \mathbb{N}$ [3], this is not the case in the agnostic setting. We hence define

Definition 12 (*n-Weakly Consistent ATS*). An ATS, $\mathcal{L}$, is said to be n-weakly consistent if $\mathcal{L}$ is monotonic (Definition 2), invertible (Definition 7), n-weakly consonant with the union (Definition 10), and n-weakly consonant with the intersection (Definition 11).

Example 24 (**Region Estimator**). The ATS from Example 5 satisfies weak *n*-union and weak *n*-intersection consonance. The argument is the same as that presented in Examples 16 and 18. It follows that this is a *n*-weakly consistent ATS.

Example 25 (**Tests based on posterior probabilities**). Consider Example 2. In order to guarantee weak *n*-union consonance for every prior, it is necessary and sufficient to have $c_1 \geqslant nc_2$. Moreover, to guarantee weak *n*-intersection consonance for every prior, it is necessary and sufficient to have $c_2 \leqslant nc_1 - (n-1)$. It follows from these conditions and Example 20 that the following restrictions are necessary and sufficient to guarantee monotonicity, *n*-union consonance, *n*-intersection consonance and invertibility: $c_1 > n/(n+1)$ and $c_2 = 1 - c_1$. Hence, these conditions are sufficient to guarantee this ATS is *n*-weakly consistent. Now, because these conditions are also necessary, it follows that this ATS is *n*-weakly consistent *for every n* > 1 if, and only if, it remains agnostic about every hypothesis which has probability in $(0, 1)$.

5. Decision-Theoretic Perspective

In this section, we investigate agnostic testing schemes from a Bayesian decision-theoretic perspective. First, we define an ATS generated by a family of loss functions. Note that, in the context of agnostic tests, a loss function is a function $L : \mathcal{D} \times \Theta \to \mathbb{R}$ that assigns to each $\theta \in \Theta$ the loss $L(d, \theta)$ for making the decision $d \in \{0, \frac{1}{2}, 1\}$.

Definition 13 (**ATS generated by a family of loss functions**). Let $(\mathcal{X} \times \Theta, \sigma(\mathcal{X} \times \Theta), \mathbb{P})$ be a Bayesian statistical model. Let $(L_A)_{A \in \sigma(\Theta)}$ be a family of loss functions, where $L_A : \mathcal{D} \times \Theta \to \mathbb{R}$ is the loss function to be used to test $A \in \sigma(\Theta)$. An ATS generated by the family of loss functions $(L_A)_{A \in \sigma(\Theta)}$ is any ATS $\mathcal{L}$ defined over the elements of $\sigma(\Theta)$ such that, $\forall A \in \sigma(\Theta)$, $\mathcal{L}(A)$ is a Bayes test for hypothesis A against $\mathbb{P}$.

Example 26 (**Bayesian ATS generated by a family of error-wise constant loss functions**). For $A \in \sigma(\Theta)$, consider the loss function L_A of the form of Table 1, where all entries are assumed to be non negative. This is a generalization of standard $0 - 1 - c$ loss functions to agnostic tests in the sense that it penalizes not only false acceptance and false rejection with constant losses b_A and d_A, respectively, but also an eventual abstention from deciding between accepting and rejecting A with the values a_A and c_A. If $b_A d_A > a_A b_A + c_A d_A$, then the Bayes test against L_A consists in rejecting A if $\mathbb{P}(A|x) < \frac{c_A}{d_A + c_A - a_A}$, accept A if $\mathbb{P}(A|x) > \frac{b_A - c_A}{a_A + b_A - c_A}$, and remain agnostic otherwise. It follows that the following ATS is generated by the family of loss functions $(L_A)_{A \in \sigma(\Theta)}$:

$$\mathcal{L}(A)(x) = \begin{cases} 0 & \text{if } \mathbb{P}(A|x) > \frac{b_A - c_A}{a_A + b_A - c_A} \\ \frac{1}{2} & \text{if } \frac{b_A - c_A}{a_A + b_A - c_A} \geqslant \mathbb{P}(A|x) > \frac{c_A}{d_A + c_A - a_A} \\ 1 & \text{if } \frac{c_A}{d_A + c_A - a_A} \geqslant \mathbb{P}(A|x) \end{cases}$$

Notice that if, for every $A, B \in \sigma(\Theta)$, $a_A = a_B$, $b_A = b_B$, $c_A = c_B$, and $d_A = d_B$, this ATS matches that from Example 2 for a particular value of c_1 and c_2.

Table 1. The loss function for the hypothesis $\theta \in A$ used in Example 26.

Decision	State of Nature	
	$\theta \in A$	$\theta \notin A$
0 (accept A)	0	b_A
$\frac{1}{2}$ (remain agnostic about A)	a_A	c_A
1 (reject A)	d_A	0

We restrict out attention to ATSs generated by *proper* losses, a concept we adapt from [3] to agnostic tests:

Definition 14 (Proper losses). A family of loss functions $(L_A)_{A \in \sigma(\Theta)}$ has *proper losses* if

$$\begin{cases} L_A(0, \theta) < L_A(\frac{1}{2}, \theta) < L_A(1, \theta) & \text{, if } \theta \in A \\ L_A(0, \theta) > L_A(\frac{1}{2}, \theta) > L_A(1, \theta) & \text{, if } \theta \notin A \\ L_A(\frac{1}{2}, \theta) < \frac{L_A(0, \theta) + L_A(1, \theta)}{2} & \text{, for all } \theta \end{cases}$$

Definition 14 states that (i) by taking a correct decision we lose less than by taking a wrong decision; (ii) by remaining agnostic we do not lose as much as when taking a wrong decision, but we lose more than by taking a correct decision; and (iii) it is better to remain agnostic about *A* than to flip a coin to decide if we reject or accept this hypothesis.

Example 27 (Bayesian ATS generated by a family of error-wise constant loss functions). In order to ensure that the loss in Example 26 is proper, the following restrictions must be satisfied: $0 < a_A < d_A/2$ and $0 < c_A < b_A/2$. In particular, these conditions imply those stated in Example 26.

5.1. Monotonicity

We now turn our attention towards characterizing Bayesian monotonic ATS using a decision-theoretic framework. In order to do this, we first adapt the concept of relative losses [3] to the context of agnostic testing schemes.

Definition 15 (Relative Loss). Let L_A be a loss function for testing hypothesis *A*. The relative losses $r_A^{(1, \frac{1}{2})} : \Theta \to \mathbb{R}$ and $r_A^{(\frac{1}{2}, 0)} : \Theta \to \mathbb{R}$ are defined by

$$\begin{cases} r_A^{(1, \frac{1}{2})}(\theta) & = L_A(1, \theta) - L_A(\frac{1}{2}, \theta) \\ r_A^{(\frac{1}{2}, 0)}(\theta) & = L_A(\frac{1}{2}, \theta) - L_A(0, \theta) \end{cases}$$

The relative losses thus measure the difference between the losses of rejecting a given hypothesis and remaining agnostic about it, as well as the difference between the losses of remaining agnostic and accepting it. In order to guarantee that a Bayesian ATS is monotonic, certain constraints on the relative losses must be imposed. The next definition presents one of such assumptions, which we interpret in the sequence.

Definition 16 (Monotonic Relative Loss). Let $\mathcal{D}_>^2 = \{(1, \frac{1}{2}), (\frac{1}{2}, 0)\}$. $(L_A)_{A \in \sigma(\Theta)}$ has *monotonic relative losses* if the family $(L_A)_{A \in \sigma(\Theta)}$ is proper and, for all $A, B \in \sigma(\Theta)$ such that $A \subset B$ and for all $(i, j) \in \mathcal{D}_>^2$,

$$r_B^{(i,j)}(\theta) \geqslant r_A^{(i,j)}(\theta) \ \forall \theta \in \Theta$$

Let $A, B \in \sigma(\Theta)$ with $A \subseteq B$. If $\theta \in A$, both *A* and *B* are true, so $(L_A)_{A \in \sigma(\Theta)}$ having monotonic relative losses reflects the situation in which the rougher error of rejecting B compared to rejecting A (with respect to remaining agnostic about these hypotheses) should be assigned a larger relative loss. Similarly, the rougher error of remaining agnostic about *B* should be assigned a larger relative loss than remaining agnostic about *A* (with respect to correctly accepting these hypotheses). If $\theta \in B$ but $\theta \notin A$, these conditions are a consequence of the assumption that the family $(L_A)_{A \in \sigma(\Theta)}$ is proper. The case $\theta \notin B$ can be interpreted in a similar fashion as the case $\theta \in A$.

The following example presents necessary and sufficient conditions to ensure that the loss functions from Example 26 yield monotonic relative losses.

Example 28. Consider the losses presented in Example 26. Assuming the losses are proper (see Example 27), the conditions required to ensure $(L_A)_{A \in \sigma(\Theta)}$ has monotonic relative losses are

$$a_A \leqslant a_B, \ c_B \leqslant c_A, \ c_B - b_B \geqslant c_A - b_A \text{ and } d_B - a_B \geqslant d_A - a_A$$

Notice that these restrictions imply that $b_A \geqslant b_B$.

As a particular example, let $k > 2$ and λ be a finite measure in $\sigma(\Theta)$ with $\lambda(\Theta) > 0$. The following assignments yield a proper and monotonic loss: for every $A \in \sigma(\Theta)$, $b_A = \lambda(A^c)$, $a_A = \lambda(A)/k$, $c_A = \lambda(A^c)/k$, and $d_A = \lambda(A)$. Another particular case is when $a_A = a_B$, $b_A = b_B$, $c_A = c_B$, and $d_A = d_B$ for every $A, B \in \sigma(\Theta)$.

Another concept that helps us characterizing the Bayesian monotonic agnostic testing schemes is that of balanced relative losses, which we adapt from [7].

Definition 17 (Balanced Relative Loss). $(L_A)_{A \in \sigma(\Theta)}$ has *balanced relative losses* if, for all $A, B \in \sigma(\Theta)$ such that $A \subset B$, for all $\theta_1 \in A$ and $\theta_2 \in B^c$, and for all $(i, j) \in D_>^2$,

$$\frac{r_A^{(i,j)}(\theta_1)}{r_A^{(i,j)}(\theta_2)} \geqslant \frac{r_B^{(i,j)}(\theta_1)}{r_B^{(i,j)}(\theta_2)}$$

Lemma 1. *If $(L_A)_{A \in \sigma(\Theta)}$ has monotonic relative losses, then $(L_A)_{A \in \sigma(\Theta)}$ has balanced relative losses.*

The following result shows that balanced relative losses characterize Bayesian monotonic ATS.

Theorem 4. *Let $(L_A)_{A \in \sigma(\Theta)}$ be a family of proper loss functions. Assume that for every $\theta \in \Theta$ and $x \in \mathcal{X}$, $L_x(\theta) > 0$. For every prior π for θ, let $\mathcal{L}^\pi$ denote a Bayesian ATS generated by $(L_A)_{A \in \sigma(\Theta)}$. There exists a monotonic $\mathcal{L}^\pi$ for every prior π if, and only if, $(L_A)_{A \in \sigma(\Theta)}$ has balanced relative losses.*

Example 29. In Example 28, we obtained conditions on the loss functions $(L_A)_{A \in \sigma(\Theta)}$ from Example 26 in order to guarantee that family to have monotonic relative losses. From Lemma 1 and Theorem 4, it follows that such family of loss functions yield monotonic Bayesian ATSs whatever the prior for θ is. In other words, there are family of loss functions that induce monotonic tests based on posterior probabilities.

5.2. Union Consonance

We now turn our attention towards characterizing union consonant Bayesian ATS using a decision theoretic framework.

Definition 18. $(L_A)_{A \in \sigma(\Theta)}$ is *compatible with weak union consonance* if there exists **no** $A, B \in \sigma(\Theta)$, $\theta_1, \theta_2, \theta_3 \in \Theta$ and $p_1, p_2, p_3 \geqslant 0$ such that $p_1 + p_2 + p_3 = 1$ and

$$\begin{cases} p_1 \cdot r_A^{(1, \frac{1}{2})}(\theta_1) + p_2 \cdot r_A^{(1, \frac{1}{2})}(\theta_2) + p_3 \cdot r_A^{(1, \frac{1}{2})}(\theta_3) & < 0 \\ p_1 \cdot r_B^{(1, \frac{1}{2})}(\theta_1) + p_2 \cdot r_B^{(1, \frac{1}{2})}(\theta_2) + p_3 \cdot r_B^{(1, \frac{1}{2})}(\theta_3) & < 0 \\ p_1 \cdot r_{A \cup B}^{(\frac{1}{2}, 0)}(\theta_1) + p_2 \cdot r_{A \cup B}^{(\frac{1}{2}, 0)}(\theta_2) + p_3 \cdot r_{A \cup B}^{(\frac{1}{2}, 0)}(\theta_3) & > 0 \end{cases}$$

Definition 18 states that the family of loss functions $(L_A)_{A \in \sigma(\Theta)}$ being compatible with weak union consonance cannot induce any Bayesian ATS on the basis of which one may prefer rejecting both hypotheses A and B over remaining agnostic about them while accepting $A \cup B$ rather than abstaining.

As we will see in the next theorem, proper loss functions compatible with weak union consonance characterize Bayesian ATSs that are weakly consonant with the union.

Theorem 5. *Let $(L_A)_{A \in \sigma(\Theta)}$ be a family of proper loss functions. Assume that for every $\theta \in \Theta$ and $x \in \mathcal{X}$, $L_x(\theta) > 0$. For every prior π for θ, let $\mathcal{L}^\pi$ denote a Bayesian ATS generated by $(L_A)_{A \in \sigma(\Theta)}$. There exists an ATS $\mathcal{L}^\pi$ that is weakly consonant with the union for every priori π if, and only if, $(L_A)_{A \in \sigma(\Theta)}$ is compatible with weak union consonance.*

Example 30. We saw that the ATS from Example 2 is a Bayes test against a particular proper loss (Examples 26 and 27) and that it is weakly consonant with the union (Example 13). It follows from Theorem 5 that the family of loss functions that lead to this ATS are compatible with weak union consonance.

Definition 19 (Union consonance-balanced relative losses [7]). $(L_A)_{A \in \sigma(\Theta)}$ has union consonance-balanced relative losses if, for every $A, B \in \sigma(\Theta)$, $\theta_1 \in A \cup B$ and $\theta_2 \in (A \cup B)^c$,

$$\frac{r_A^{(1,\frac{1}{2})}(\theta_1)}{r_A^{(1,\frac{1}{2})}(\theta_2)} \leqslant \frac{r_{A \cup B}^{(\frac{1}{2},0)}(\theta_1)}{r_{A \cup B}^{(\frac{1}{2},0)}(\theta_2)}, \textbf{ or}$$

$$\frac{r_B^{(1,\frac{1}{2})}(\theta_1)}{r_B^{(1,\frac{1}{2})}(\theta_2)} \leqslant \frac{r_{A \cup B}^{(\frac{1}{2},0)}(\theta_1)}{r_{A \cup B}^{(\frac{1}{2},0)}(\theta_2)}$$

Corollary 1. *Let $(L_A)_{A \in \sigma(\Theta)}$ be a family of proper loss functions. Assume that for every $\theta \in \Theta$ and $x \in \mathcal{X}$, $L_x(\theta) > 0$. If $(L_A)_{A \in \sigma(\Theta)}$ does not have union consonance-balanced relative losses, then there exists a prior π such that every Bayesian ATS, $\mathcal{L}^\pi$, is not weakly consonant with the union.*

5.3. Intersection Consonance

Next, we characterize intersection consonant Bayesian ATS under a Bayesian perspective.

Definition 20. $(L_A)_{A \in \sigma(\Theta)}$ is *compatible with weak intersection consonance* if there exists **no** $A, B \in \sigma(\Theta)$, $\theta_1, \theta_2, \theta_3 \in \Theta$ and $p_1, p_2, p_3 \geqslant 0$ such that $p_1 + p_2 + p_3 = 1$ and

$$\begin{cases} p_1 \cdot r_A^{(\frac{1}{2},0)}(\theta_1) + p_2 \cdot r_A^{(\frac{1}{2},0)}(\theta_2) + p_3 \cdot r_A^{(\frac{1}{2},0)}(\theta_3) &> 0 \\ p_1 \cdot r_B^{(\frac{1}{2},0)}(\theta_1) + p_2 \cdot r_B^{(\frac{1}{2},0)}(\theta_2) + p_3 \cdot r_B^{(\frac{1}{2},0)}(\theta_3) &> 0 \\ p_1 \cdot r_{A \cap B}^{(1,\frac{1}{2})}(\theta_1) + p_2 \cdot r_{A \cap B}^{(1,\frac{1}{2})}(\theta_2) + p_3 \cdot r_{A \cap B}^{(1,\frac{1}{2})}(\theta_3) &< 0 \end{cases}$$

Definition 20 states that the family of loss functions $(L_A)_{A \in \sigma(\Theta)}$ being compatible with weak intersection consonance cannot induce any Bayesian ATS on the basis of which one may prefer accepting both hypotheses A and B to remaining agnostic about them while rejecting $A \cap B$ rather than abstaining.

As we will see in the next theorem, proper loss functions compatible with weak intersection consonance characterize Bayesian ATSs that are weakly consonant with the intersection .

Theorem 6. *Let $(L_A)_{A \in \sigma(\Theta)}$ be a family of proper loss functions. Assume that for every $\theta \in \Theta$ and $x \in \mathcal{X}$, $L_x(\theta) > 0$. For every prior π for θ, let $\mathcal{L}^\pi$ denote a Bayesian ATS generated by $(L_A)_{A \in \sigma(\Theta)}$. There exists an ATS $\mathcal{L}^\pi$ that is weakly consonant with the intersection for every prior π if, and only if, $(L_A)_{A \in \sigma(\Theta)}$ is compatible with weak intersection consonance.*

Example 31. We saw that the ATS from Example 2 is a Bayes test against a particular proper loss (Examples 26 and 27) and that it is weakly consonant with the intersection (Example 17). It follows from Theorem 6 that the family of loss functions that lead to this ATS are compatible with weak intersection consonance.

Definition 21 (Intersection consonance-balanced relative losses [7]). $(L_A)_{A \in \sigma(\Theta)}$ has intersection consonance-balanced relative losses if, for every $A, B \in \sigma(\Theta)$, $\theta_1 \in A \cap B$ and $\theta_2 \in (A \cap B)^c$,

$$\frac{r_{A \cap B}^{(1,\frac{1}{2})}(\theta_1)}{r_{A \cap B}^{(1,\frac{1}{2})}(\theta_2)} \leqslant \frac{r_A^{(\frac{1}{2},0)}(\theta_1)}{r_A^{(\frac{1}{2},0)}(\theta_2)}\text{, or}$$

$$\frac{r_{A \cap B}^{(1,\frac{1}{2})}(\theta_1)}{r_{A \cap B}^{(1,\frac{1}{2})}(\theta_2)} \leqslant \frac{r_B^{(\frac{1}{2},0)}(\theta_1)}{r_B^{(\frac{1}{2},0)}(\theta_2)}$$

Corollary 2. *Let $(L_A)_{A \in \sigma(\Theta)}$ be a family of proper loss functions. Assume that for every $\theta \in \Theta$ and $x \in \mathcal{X}$, $L_x(\theta) > 0$. If $(L_A)_{A \in \sigma(\Theta)}$ does not have intersection consonance-balanced relative losses, then there exists a prior π such that every Bayesian ATS, $\mathcal{L}^\pi$, is not weakly consonant with the intersection.*

We end this section by noting that although we focused our results on weak consonance, they can be extended to strong consonance using the same techniques presented in the Appendix.

5.4. Invertibility

Finally, we examine invertible Bayesian ATSs from a decision-theoretic standpoint.

Definition 22 (Invertible Relative Losses). $(L_A)_{A \in \sigma(\Theta)}$ has invertible relative losses if, for every $A \in \sigma(\Theta)$, for all $\theta_1 \in A$, $\theta_2 \in A^c$ and $(i,j) \in D_>^2$,

$$\frac{r_A^{(i,j)}(\theta_1)}{r_A^{(i,j)}(\theta_2)} = \frac{r_{A^c}^{(i,j)}(\theta_1)}{r_{A^c}^{(i,j)}(\theta_2)}$$

We end this section by showing that invertible Bayesian ATSs are determined by family of loss functions that fulfill the conditions of Definition 22.

Theorem 7. *Let $(L_A)_{A \in \sigma(\Theta)}$ be a family of proper loss functions. Assume that for every $\theta \in \Theta$ and $x \in \mathcal{X}$, $L_x(\theta) > 0$. For every prior π for θ, let $\mathcal{L}^\pi$ denote a Bayesian ATS generated by $(L_A)_{A \in \sigma(\Theta)}$. There exists an ATS $\mathcal{L}^\pi$ that is invertible for every prior π if, and only if, $(L_A)_{A \in \sigma(\Theta)}$ has invertible relative losses.*

Example 32. For every $A \in \sigma(\Theta)$, let $(L_A)_{A \in \sigma(\Theta)}$ be such that $L_A(1, \theta) = L_{A^c}(0, \theta)$ and $L_A(\frac{1}{2}, \theta) = L_{A^c}(\frac{1}{2}, \theta)$. It is easily seen that the conditions from Definition 22 hold. Theorem 7 then implies that any Bayesian ATS generated by $(L_A)_{A \in \sigma(\Theta)}$ is invertible.

6. Final Remarks

Agnostic tests allow one to explicitly capture the difference between "not rejecting" and "accepting" a null hypothesis. When the agnostic decision is chosen, the null hypothesis is neither rejected or accepted. This possibility aligns with the idea that although precise null hypotheses can be tested, they shouldn't be accepted. This idea is followed by the region based agnostic tests derived in this paper, which can either remain agnostic or reject precise null hypotheses.

This distinction provides a solution to the problem raised by Izbicki and Esteves [3], in which all (non-agnostic) logically coherent tests were shown to be based on point estimators which lack statistical optimality. We show that agnostic tests based on region estimators satisfy logical consistency and also allow statistical optimality. For example, agnostic tests based on frequentist confidence intervals control family wise error. Similarly, agnostic tests based on posterior density regions are shown to be an extension of the Full Bayesian Significance Test [11].

Future research includes investigating the consequences and generalizations of the logical requirements in this paper. For example, one could study what kinds of trivariate logic derive from

the different definition of logical consistency studied in this paper. One could also generalize these logical requirements to generalized agnostic tests, in which one can decide among different degrees of agnosticism. The scale of such degrees can be either discrete or continuous. One could also investigate region estimator-based ATSs with respect to other optimality criteria such as statistical power.

The results of this paper can also be tied to the philosophical literature that studies the consequences and importance of precise hypothesis. Agnostic tests can be used to revisit the role of testing precise hypotheses in science. Agnostic tests also provide a framework to interpret the scientific meaning of measures of possibility or significance of precise hypotheses.

Author Contributions: The manuscript has come to fruition by the substantial contributions of all authors. All authors have also been involved in either writing the article or carefully revising it. All authors have read and approved the submitted version of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Theorem 1. The sufficiency is immediate. Let $I = \{0, \frac{1}{2}, 1\} \subseteq \mathbb{R}$, $c_1 = \frac{1}{2}$ and $c_2 = 0$. For $A \in \sigma(\Theta)$, let $s_A = 1 - \mathcal{L}(A)$, $(s_A)_{A \in \sigma(\Theta)}$ is such that $s_A(x) \leqslant s_B(x)$, $\forall x \in \mathcal{X}$, if $A \subset B$ and it is straightforward to verify Equation (1).

Now, let $A, B \in \sigma(\Theta)$, with $A \subset B$. If $\mathcal{L}$ is given by Equation (1), it follows that:

1. $\mathcal{L}_A(x) = 0 \implies s_A(x) > c_1 \implies s_B(x) > c_1 \implies \mathcal{L}_B(x) = 0$.
2. $\mathcal{L}_A(x) = \frac{1}{2} \implies c_1 \geqslant s_A(x) > c_2 \implies s_B(x) > c_2 \implies L_B(x) \in \{0, \frac{1}{2}\}$.
3. $\mathcal{L}_A(x) = 1 \implies \mathcal{L}_B(x) \leqslant \mathcal{L}_A(x) = 1$.

From (1), (2), (3) it follows that $\mathcal{L}_A(x) \geqslant \mathcal{L}_B(x)$, thus $\mathcal{L}$ is monotonic. □

Proof of Theorem 2. We start by proving the sufficiency of these conditions.

1. If for every $A, B \in \sigma(\Theta)$ with $A \subseteq B$, $c_2^A \geqslant c_2^B$ and $c_1^A \geqslant c_1^B$, then for every $x \in \mathcal{X}$, $\mathbb{P}(A|x) > c_2^A \implies \mathbb{P}(B|x) > c_2^B$, and $\mathbb{P}(A|x) > c_1^A \implies \mathbb{P}(B|x) > c_1^B$. It follows that monotonicity holds.
2. If for every $A, B \in \sigma(\Theta)$ such that $A \neq B$, $c_2^A + c_2^B \leqslant c_1^{A \cup B}$, then for every $x \in \mathcal{X}$, $\mathbb{P}(A|x) \leqslant c_2^A$ and $\mathbb{P}(B|x) \leqslant c_2^B$ implies that $\mathbb{P}(A \cup B|x) \leqslant \mathbb{P}(A|x) + \mathbb{P}(B|x) = c_2^A + c_2^B \leqslant c_1^{A \cup B}$. It follows that union consonance holds.
3. If for every $A, B \in \sigma(\Theta)$ such that $A \neq B$, $c_1^A + c_1^B - 1 \geqslant c_2^{A \cap B}$, then for every $x \in \mathcal{X}$, $\mathbb{P}(A|x) > c_1^A$ and $\mathbb{P}(B|x) > c_1^B$ implies that $\mathbb{P}(A \cap B|x) \geqslant \mathbb{P}(A|x) + \mathbb{P}(B|x) - 1 > c_1^A + c_1^B - 1 \geqslant c_2^{A \cap B}$. It follows that intersection consonance holds.
4. If for every $A \in \sigma(\Theta)$, $c_1^A = 1 - c_2^{A^c}$, then for every $x \in \mathcal{X}$, $\mathbb{P}(A|x) \leqslant c_1^A$ if, and only if, $\mathbb{P}(A^c|x) \geqslant 1 - c_1^A = c_2^{A^c}$. Similarly, $\mathbb{P}(A|x) \leqslant c_2^A$ if, and only if, $\mathbb{P}(A^c|x) \geqslant 1 - c_2^A = c_1^{A^c}$. It follows that invertibility holds.

We prove the necessary condition only for union consonance; the other statements have a similar proof. Suppose there are $A, B \in \sigma(\Theta)$, $A \neq B$, such that $c_2^A + c_2^B > c_1^{A \cup B}$. Let $\theta_1 \in A \cap B^c$ and $\theta_2 \in B \cap A^c$.

First, assume $c_2^A + c_2^B \leqslant 1$. Let $\epsilon > 0$ be such that $\epsilon \leqslant (c_2^A + c_2^B - c_1^{A \cup B})/2$ and $\epsilon < \min\{c_2^A, c_2^B\}$. Assume that the posterior distribution on Θ given x is such that

$$\mathbb{P}(A|x) = \mathbb{P}(\{\theta_1\}|x) = c_2^A - \epsilon \text{ and } \mathbb{P}(B|x) = \mathbb{P}(\{\theta_2\}|x) = c_2^B - \epsilon$$

(see the Appendix of [3] for a prior distribution that leads to such posterior). It follows that $\mathbb{P}(A|x) = \mathbb{P}(\{\theta_1\}|x) \leqslant c_2^A$, $\mathbb{P}(B|x) = \mathbb{P}(\{\theta_2\}|x) \leqslant c_2^B$, and

$$\mathbb{P}(A \cup B|x) = \mathbb{P}(\{\theta_1\}|x) + \mathbb{P}(\{\theta_2\}|x) = c_2^A + c_2^B - 2\epsilon > c_2^A + c_2^B - (c_2^A + c_2^B - c_1^{A \cup B}) = c_1^{A \cup B}$$

Hence A and B are rejected, but $A \cup B$ is accepted.

Now, assume $c_2^A + c_2^B > 1$. Let $b_2^A < c_2^A$ and $b_2^B < c_2^B$ be such that $b_2^A + b_2^B = 1$. Assume that the posterior distribution on Θ is such that

$$\mathbb{P}(A|x) = \mathbb{P}(\{\theta_1\}|x) = b_2^A \text{ and } \mathbb{P}(B|x) = \mathbb{P}(\{\theta_2\}|x) = b_2^B$$

It follows that $\mathbb{P}(A|x) = \mathbb{P}(\{\theta_1\}|x) < b_2^A < c_2^A$, $\mathbb{P}(B|x) = \mathbb{P}(\{\theta_2\}|x) < b_2^B < c_2^B$, and

$$\mathbb{P}(A \cup B|x) = \mathbb{P}(\{\theta_1\}, \{\theta_2\}|x) = b_2^A + b_2^B = 1 \geqslant c_1^{A \cup B}$$

Hence A and B are rejected, but $A \cup B$ is not. $\quad\square$

Lemma A1. *Let $\mathcal{L}$ be an invertible ATS. If, for every x, there exists $R(x) \subset \Theta$ such that $\forall A \in \sigma(\Theta)$ $\mathcal{L}(A)(x) = 0$ if and only if $R(x) \subset A$, then $\mathcal{L}$ is a region estimator-based ATS (Example 5).*

Proof of Lemma A1. It follows from definition that, for every $A \in \sigma(\Theta)$ such that $R(x) \subset A$, $\mathcal{L}(A)(x) = 0$. Furthermore, for every $A \in \sigma(\Theta)$ such that $R(x) \subset A^c$, $\mathcal{L}(A^c)(x) = 0$. Therefore, it follows from invertibility that $\mathcal{L}(A)(x) = 1$. Finally, let $A \in \sigma(\Theta)$ be such that $A \cap R(x) \neq \varnothing$ and $A^C \cap R(x) \neq \varnothing$. Since $A \cap R(x) \neq R(x)$ and $A^c \cap R(x) \neq R(x)$ it follows that $\mathcal{L}(A)(x) \geqslant \frac{1}{2}$ and $\mathcal{L}(A^c)(x) \geqslant \frac{1}{2}$. Conclude from invertibility that $\mathcal{L}(A)(x) = \frac{1}{2}$. $\quad\square$

Proof of Theorem 3. It follows from definition that a region estimator ATS is fully consistent. In order to prove the reverse implication, consider the following notation. For every $x \in \mathcal{X}$ and $\theta \in \Theta$, let $A_\theta = \Theta - \{\theta\}$. Let $R(x) = \bigcap\{A_\theta : \mathcal{L}(A_\theta)(x) = 0\}$.

Next, we prove that, for every $B \in \sigma(\Theta)$, $\mathcal{L}(B)(x) = 0$ if and only if $R(x) \subset B$. Let $B \in \sigma(\Theta)$ be such that $R(x) \subset B$. Therefore, it follows from the definition of $R(x)$ that, for every $\theta \in B^c$, $\mathcal{L}(A_\theta) = 0$. Since $B = \bigcap\{A_\theta : \theta \in B^c\}$, it follows from strong intersection consonance (Definition 6) that $\mathcal{L}(B)(x) = 0$. Let $B \in \sigma(\Theta)$ be such that $\mathcal{L}(B)(x) = 0$. It follows from the monotonicity of $\mathcal{L}$ (Definition 2) that, for every $\theta \in B^c$, $\mathcal{L}(A_\theta)(x) = 0$ as $B \subseteq A_\theta$. Therefore,

$$R(x) = \bigcap\{A_\theta : \mathcal{L}(A_\theta)(x) = 0\} \subset \bigcap\{A_\theta : \theta \in B^c\} = B$$

Conclude that for every $B \in \sigma(\Theta)$, $\mathcal{L}(B)(x) = 0$ if and only if $R(x) \subset B$.

Since $\mathcal{L}$ is invertible (Definition 7), it follows from the above conclusion and Lemma A1 that $\mathcal{L}$ is a region estimator-based ATS. $\quad\square$

Lemma A2. *Let, for $i \in \Theta$, $\{i\} \in \sigma(\Theta)$ and $f_1, \ldots, f_m$ be $\sigma(\Theta)/\mathbb{R}$-measurable bounded functions. If there exists a probability $\mathbb{P}$ on $\sigma(\Theta)$ such that, for all $1 \leqslant i \leqslant m$, $\int f_i d\mathbb{P} > 0$, then are $A \in \sigma(\Theta)$, A finite, and a probability $\mathbb{P}^*$ with a finite support such that $\mathbb{P}^*(A) = 1$ and such that, for all $1 \leqslant i \leqslant m$, $\int f_i d\mathbb{P}^* > 0$.*

Proof. Let $\epsilon_i > 0$ be such that,

$$\int f_i d\mathbb{P} > \epsilon_i \tag{A1}$$

Since f_i are bounded, there exist simple measurable functions, g_i, such that

$$\sup_{x \in \Theta} |g_i(x) - f_i(x)| < \frac{\min_i \epsilon_i}{2} \tag{A2}$$

Therefore,

$$\int g_i d\mathbb{P} > \frac{\epsilon_i}{2} \tag{A3}$$

Let $\mathcal{G}_i = \{g_i^{-1}(\{j\}) : j \in g_i(\Theta)\}$. Observe that $\mathcal{G}_i$ is a finite partition of Θ. Let $\mathcal{G}^*$ be the coarsest partition that is finer than every $\mathcal{G}_i$. Since every $\mathcal{G}_i$ is finite, $\mathcal{G}^*$ is finite. Let $h : \mathcal{G}^* \to \Theta$ be such that $h(G) \in G$. Define $\mathbb{P}^* : \sigma(\Theta) \to \mathbb{R}_*$ by $\mathbb{P}^*(A) = \sum_{G \in \mathcal{G}*} \mathbb{P}(G)\mathbb{I}_A(h(G))$. $\mathbb{P}^*$ is such that $\mathbb{P}^*(\{h(G)\}) = \mathbb{P}(G), \forall G \in \mathcal{G}^*$, and that $\mathbb{P}^*(h(\mathcal{G}^*)) = 1$, where $h(\mathcal{G}^*)$ is a finite subset of Θ. Also, conclude from the definition of $\mathcal{G}^*$ and Equation (A3) that

$$\int g_i d\mathbb{P}^* = \int g_i d\mathbb{P} > \frac{\epsilon_i}{2} \tag{A4}$$

Conclude from Equations (A2) and (A4) that

$$\int f_i d\mathbb{P}^* > 0, \ i = 1, \ldots, m$$

□

Lemma A3. *Let, for $i \in \Theta$, $\{i\} \in \sigma(\Theta)$ and $f_1, \ldots, f_m$ be $\sigma(\Theta)/\mathbb{R}$-measurable bounded functions. If there exists a probability $\mathbb{P}$ on $\sigma(\Theta)$ such that $\mathbb{P}$ has a finite support and, for all $1 \leqslant i \leqslant m$, $\int f_i d\mathbb{P} > 0$, then there exists a probability $\mathbb{P}^*$ with a support of size smaller or equal than m such that, for all $1 \leqslant i \leqslant m$, $\int f_i d\mathbb{P}^* > 0$.*

Proof. Let $\epsilon_i > 0$ be such that

$$\int f_i d\mathbb{P} \geqslant \epsilon_i$$

Let $\Theta_{\mathbb{P}}$ denote the support of $\mathbb{P}$. Let $\theta_1, \ldots, \theta_{|\Theta_{\mathbb{P}}|}$ be an ordering of the elements of $\Theta_{\mathbb{P}}$. Let F be a $m \times |\Theta_{\mathbb{P}}|$ matrix such that $F_{i,j} = f_i(\theta_j)$. Let $p \in \mathbb{R}^{|\Theta_{\mathbb{P}}|}$ be such that $p_j = \mathbb{P}(\{\theta_j\})$, $j = 1, \ldots, |\Theta_{\mathbb{P}}|$. Observe that

$$Fp \geqslant \epsilon; p \geqslant 0$$

Therefore, the set $C = \{p^* \in \mathbb{R}^{|\Theta_{\mathbb{P}}|} : p \geqslant 0, Fp \geqslant \epsilon\}$ is a non-empty polyhedron. Conclude that there exists a vertex $p^* \in C$ such that $|\{i : p_i^* = 0\}| \geqslant |\Theta_{\mathbb{P}}| - m$. Define $\mathbb{P}^*(\{\theta_i\}) = \frac{p_i^*}{\|p^*\|_1}$. □

Theorem A1. *Let, for $i \in \Theta$, $\{i\} \in \sigma(\Theta)$ and $f_1, \ldots, f_m$ be $\sigma(\Theta)/\mathbb{R}$-measurable bounded functions. There exists a probability $\mathbb{P}$ on $\sigma(\Theta)$ such that, for all $1 \leqslant i \leqslant m$, $\int f_i d\mathbb{P} > 0$, if and only if there exists a probability $\mathbb{P}^*$ with a support of size smaller or equal to m such that, for all $1 \leqslant i \leqslant m$, $\int f_i d\mathbb{P}^* > 0$.*

Proof. Follows directly from Lemmas A2 and A3. □

Lemma A4. *Let $(L_A)_{A \in \sigma(\Theta)}$ have proper losses. For every $x \in \mathcal{X}$,*

- *If $E[L_A(1,\theta)|x] < E[L_A(\frac{1}{2},\theta)|x]$, then $E[L_A(1,\theta)|x] < E[L_A(0,\theta)|x]$.*
- *If $E[L_A(0,\theta)|x] < E[L_A(\frac{1}{2},\theta)|x]$, then $E[L_A(0,\theta)|x] < E[L_A(1,\theta)|x]$.*

Proof of Lemma A4. The proof follows directly from the monotonicity of conditional expectation. □

Proof of Lemma 1. Let $A \subset B$, $\theta_1 \in A$, $\theta_2 \in B^c$ and $(i,j) \in D_>^2$. Since $(L_A)_{A \in \sigma(\Theta)}$ has *proper and monotonic relative losses*,

$$r_B^{(i,j)}(\theta_1) \geqslant r_A^{(i,j)}(\theta_1) > 0$$
$$r_A^{(i,j)}(\theta_2) \leqslant r_B^{(i,j)}(\theta_2) < 0$$

Conclude that $(L_A)_{A \in \sigma(\Theta)}$ has *balanced relative losses*. □

Lemma A5. *Let* $(L_A)_{A \in \sigma(\Theta)}$ *have proper losses,* L_A *be bounded for every* $A \in \sigma(\Theta)$ *and* $L_x(\theta) > 0$ *for every* $\theta \in \Theta$ *and* $x \in \mathcal{X}$. *There exists a prior for* θ *such that, for some* $A \subset B$ *and* $(i, j) \in D_>$ *and some* $x \in \mathcal{X}$, $E[L_B(i, \theta)|x] < E[L_B(j, \theta)|x]$ *and* $E[L_A(i, \theta)|x] > E[L_A(j, \theta)|x]$ *if and only if* $(L_A)_{A \in \sigma(\Theta)}$ *does not have balanced relative losses.*

Proof of Lemma A5. Since $L_x(\theta) > 0$, the space of posteriors is exactly the space of priors over $\sigma(\Theta)$ [3]. Therefore, there exists a prior such that $E[L_B(i, \theta)|x] < E[L_B(j, \theta)|x]$ and $E[L_A(i, \theta)|x] > E[L_A(j, \theta)|x]$ if and only if there exists $\mathbb{P}$ such that

$$\int -r_B^{(i,j)} d\mathbb{P} > 0 \text{ and}$$

$$\int r_A^{(i,j)} d\mathbb{P} > 0$$

It follows from Theorem A1 that there exists such a $\mathbb{P}$ if and only if there exists $\theta_1, \theta_2 \in \Theta$ and $p \in [0, 1]$ such that

$$\begin{cases} p \cdot r_B^{(i,j)}(\theta_1) + (1-p) \cdot r_B^{(i,j)}(\theta_2) & < 0 \\ p \cdot r_A^{(i,j)}(\theta_1) + (1-p) \cdot r_A^{(i,j)}(\theta_2) & > 0 \end{cases}$$

Since $(L_A)_{A \in \sigma(\Theta)}$ has *proper losses*, the above condition is satisfied if and only if $p \in (0, 1)$, that is, if and only if $(L_A)_{A \in \sigma(\Theta)}$ doesn't have *balanced relative losses*. $\square$

Proof of Theorem 4. Assume that $(L_A)_{A \in \sigma(\Theta)}$ has *balanced relative losses*. Let P_θ be an arbitrary prior and A, B be arbitrary sets such that $A \subset B$. It follows from Lemma A5 that, for every $(i, j) \in \mathcal{D}_>^2$, it cannot be the case that $E[L_B(i, \theta)|x] < E[L_B(j, \theta)|x]$ and $E[L_A(i, \theta)|x] > E[L_A(j, \theta)|x]$. Conclude from Lemma A4 that there exists a monotonic Bayesian ATS.

Assume that $(L_A)_{A \in \sigma(\Theta)}$ does not have *balanced relative losses*. It follows from Lemma A5 that there exists a prior P_θ, $A \subset B$ and $(i, j) \in \mathcal{D}_>^2$ and $x \in \chi$ such that $E[L_B(i, \theta)|x] < E[L_B(j, \theta)|x]$ and $E[L_A(i, \theta)|x] > E[L_A(j, \theta)|x]$. Conclude from Lemma A4 that, for every Bayesian ATS, $\mathcal{L}^{P_\theta}(A)(x) \leqslant j < i \leqslant \mathcal{L}^{P_\theta}(B)(x)$. Therefore there exists no monotonic Bayesian ATS against P_θ. $\square$

Proof of Theorem 5. The proof follows directly from Theorem A1 and Lemma A4. $\square$

Proof of Corollary 1. Assume that $(L_A)_{A \in \sigma(\Theta)}$ doesn't satisfy Definition 19. Therefore, there exist $A, B \in \sigma(\Theta)$, $\theta_1 \in A \cup B$ and $\theta_2 \in (A \cup B)^c$ such that

$$\begin{cases} \dfrac{r_A^{(1,\frac{1}{2})}(\theta_1)}{r_A^{(1,\frac{1}{2})}(\theta_2)} & > \dfrac{r_{A \cup B}^{(\frac{1}{2},0)}(\theta_1)}{r_{A \cup B}^{(\frac{1}{2},0)}(\theta_2)} \\ \dfrac{r_B^{(1,\frac{1}{2})}(\theta_1)}{r_B^{(1,\frac{1}{2})}(\theta_2)} & > \dfrac{r_{A \cup B}^{(\frac{1}{2},0)}(\theta_1)}{r_{A \cup B}^{(\frac{1}{2},0)}(\theta_2)} \end{cases}$$

Since $(L_A)_{A \in \sigma(\Theta)}$ has *proper losses*, there exist $q_1, q_2 \in (0, 1)$ such that

$$\begin{cases} q_1 \cdot r_A^{(1,\frac{1}{2})}(\theta_1) + (1-q_1) \cdot r_A^{(1,\frac{1}{2})}(\theta_2) & < 0 \\ q_1 \cdot r_{A \cup B}^{(\frac{1}{2},0)}(\theta_1) + (1-q_1) \cdot r_{A \cup B}^{(\frac{1}{2},0)}(\theta_2) & > 0 \\ q_2 \cdot r_B^{(1,\frac{1}{2})}(\theta_1) + (1-q_2) \cdot r_B^{(1,\frac{1}{2})}(\theta_2) & < 0 \\ q_2 \cdot r_{A \cup B}^{(\frac{1}{2},0)}(\theta_1) + (1-q_2) \cdot r_{A \cup B}^{(\frac{1}{2},0)}(\theta_2) & > 0 \end{cases} \tag{A5}$$

Let $p_1 = \min(q_1, q_2)$, $p_2 = 1 - p_1$, $p_3 = 0$ and $\theta_3 \in \Theta$. Since $(L_A)_{A \in \sigma(\Theta)}$ has *proper losses*, it follows directly from Equation (A5) that

$$\begin{cases} p_1 \cdot r_A^{(1,\frac{1}{2})}(\theta_1) + p_2 \cdot r_A^{(1,\frac{1}{2})}(\theta_2) + p_3 \cdot r_A^{(1,\frac{1}{2})}(\theta_3) & < 0 \\ p_1 \cdot r_B^{(1,\frac{1}{2})}(\theta_1) + p_2 \cdot r_B^{(1,\frac{1}{2})}(\theta_2) + p_3 \cdot r_B^{(1,\frac{1}{2})}(\theta_3) & < 0 \\ p_1 \cdot r_{A \cup B}^{(\frac{1}{2},0)}(\theta_1) + p_2 \cdot r_{A \cup B}^{(\frac{1}{2},0)}(\theta_2) + p_3 \cdot r_{A \cup B}^{(\frac{1}{2},0)}(\theta_3) & > 0 \end{cases} \tag{A6}$$

Equation (A6) shows that $(L_A)_{A \in \sigma(\Theta)}$ is not compatible with finite union consonance. Therefore, since $(L_A)_{A \in \sigma(\Theta)}$ has *proper losses*, it follows from Theorem 5 that there exists a prior P_θ, such that, against P_θ, no Bayesian ATS is consonant with pairwise union. □

Proof of Theorem 6. The proof follows directly from Theorem A1 and Lemma A4. □

Proof of Corollary 2. The proof follows the same steps as in Corollary 1. □

Proof of Theorem 7. It follows from Theorem A1 and $(L_A)_{A \in \sigma(\Theta)}$ being proper that $(L_A)_{A \in \sigma(\Theta)}$ has invertible relative losses (Definition 22) if and only if there exists no $A \in \sigma(\Theta)$, $(i,j) \in \mathcal{D}_>$, $\theta_1 \in A$, $\theta_2 \in A^c$ and $p \in [0,1]$ such that

$$\begin{cases} p \cdot r_A^{(i,j)}(\theta_1) + (1-p)r_A^{(i,j)}(\theta_2) & > 0 \\ p \cdot r_{A^c}^{(i,j)}(\theta_1) + (1-p)r_{A^c}^{(i,j)}(\theta_2) & > 0 \end{cases} \text{ or } \begin{cases} p \cdot r_A^{(i,j)}(\theta_1) + (1-p)r_A^{(i,j)}(\theta_2) & < 0 \\ p \cdot r_{A^c}^{(i,j)}(\theta_1) + (1-p)r_{A^c}^{(i,j)}(\theta_2) & < 0 \end{cases} \tag{A7}$$

Furthermore, Equation (A7) is equivalent to there existing no $k > 0$ such that

$$\begin{cases} k & > -\dfrac{r_A^{(i,j)}(\theta_2)}{r_A^{(i,j)}(\theta_1)} \\ k^{-1} & < -\dfrac{r_{A^c}^{(i,j)}(\theta_1)}{r_{A^c}^{(i,j)}(\theta_2)} \end{cases} \text{ or } \begin{cases} k & < -\dfrac{r_A^{(i,j)}(\theta_2)}{r_A^{(i,j)}(\theta_1)} \\ k^{-1} & > -\dfrac{r_{A^c}^{(i,j)}(\theta_1)}{r_{A^c}^{(i,j)}(\theta_2)} \end{cases}$$

Conclude that $(L_A)_{A \in \sigma(\Theta)}$ is compatible with invertibility if and only if, for every for every $(i,j) \in D_>^2$, $A \in \sigma(\Theta)$, $\theta_1 \in A$ and $\theta_2 \in A^c$,

$$\frac{r_A^{(i,j)}(\theta_2)}{r_A^{(i,j)}(\theta_1)} = \frac{r_{A^c}^{(i,j)}(\theta_1)}{r_{A^c}^{(i,j)}(\theta_2)}$$

□

References

1. Wiener, Y.; El-Yaniv, R. Agnostic selective classification. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2011; pp. 1665–1673.
2. Balsubramani, A. Learning to abstain from binary prediction. **2016**, arXiv:1602.08151.
3. Izbicki, R.; Esteves, L.G. Logical consistency in simultaneous statistical test procedures. *Logic J. IGPL* **2015**, *23*, 732–758.
4. Finner, H.; Strassburger, K. The partitioning principle: A powerful tool in multiple decision theory. *Ann. Stat.* **2002**, *30*, 1194–1213.
5. Sonnemann, E. General solutions to multiple testing problems. *Biom. J.* **2008**, *50*, 641–656.
6. Patriota, A.G. S-value: An alternative measure of evidence for testing general null hypotheses. *Cienc. Nat.* **2014**, *36*, 14–22.
7. Da Silva, G.M.; Esteves, L.G.; Fossaluza, V.; Izbicki, R.; Wechsler, S. A bayesian decision-theoretic approach to logically-consistent hypothesis testing. *Entropy* **2015**, *17*, 6534–6559.
8. Berg, N. No-decision classification: An alternative to testing for statistical significance. *J. Socio-Econ.* **2004**, *33*, 631–650.

9. Babb, J.; Rogatko, A.; Zacks, S. Bayesian sequential and fixed sample testing of multihypothesis. In *Asymptotic Methods in Probability and Statistics*; Elsevier: Amsterdam, The Netherlands, 1998; pp. 801–809.

10. Ripley, B.D. *Pattern Recognition and Neural Networks*; Cambridge University Press: Cambridge, UK, 1996.

11. De Bragança Pereira, C.A.; Stern, J.M. Evidence and credibility: Full bayesian significance test for precise hypotheses. *Entropy* **1999**, *1*, 99–110.

12. Berger, J.O.; Delampady, M. Testing precise hypotheses. *Stat. Sci.* **1987**, *2*, 317–335.

13. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **1995**, *57*, 289–300.

14. Jaynes, E.T. Confidence intervals vs. Bayesian intervals. In *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*; Springer: Dodrecht, The Netherlands, 1976.

15. Gabriel, K.R. Simultaneous test procedures—Some theory of multiple comparisons. *Ann. Math. Stat.* **1969**, *41*, 224–250.

16. Fossaluza, V.; Izbicki, R.; da Silva, G.M.; Esteves, L.G. Coherent hypothesis testing. *Am. Stat.* **2016**, submitted for publication.

17. Sonnemann, E.; Finner, H. Vollständigkeitssätze für multiple testprobleme. In *Multiple Hypothesenprüfung*; Bauer, P., Hommel, G., Sonnemann, E., Eds.; Springer: Berlin, Germany, 1988; pp. 121–135. (In German)

18. Lavine, M.; Schervish, M. Bayes factors: What they are and what they are not. *Am. Stat.* **1999**, *53*, 119–122.

19. Izbicki, R.; Fossaluza, V.; Hounie, A.G.; Nakano, E.Y.; Pereira, C.A.B. Testing allele homogeneity: The problem of nested hypotheses. *BMC Genet.* **2012**, *13*, doi:10.1186/1471-2156-13-103.

20. Schervish, M.J. *p* values: What they are and what they are not. *Am. Stat.* **1996**, *50*, 203–206.

21. Hochberg, Y.; Tamhane, A.C. *Multiple Comparison Procedures*; John Wiley & Sons: New York, NY, USA, 1987.

22. Borges, W.; Stern, J.M. The rules of logic composition for the bayesian epistemic e-values. *Logic J. IGPL* **2007**, *15*, 401–420.

23. Johnson, R.A.; Wichern, D.W. *Applied Multivariate Statistical Analysis*, 6th ed.; Pearson: Upper Saddle River, NJ, USA, 2007.

24. Schervish, M.J. *Theory of Statistics*; Springer: New York, NY, USA, 1997.

25. Robert, C. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, 2nd ed.; Springer: New York, NY, USA, 2007.