Topological Criteria for Hypothesis Testing with Finite-Precision Measurements

Philip Boeken[*] Eduardo Skapinakis[†] Konstantin Genin[‡] Joris M. Mooij[§]

January 20, 2026

Abstract

We establish topological necessary and sufficient conditions under which a pair of statistical hypotheses can be consistently distinguished when i.i.d. observations are recorded only to finite precision. Requiring the test's decision regions to be open in the sample-space topology to accommodate finite-precision data, we show that a pair of null- and alternative hypotheses H_0 and H_1 admits a consistent test if and only if they are F_σ in the weak topology on the space of probability measures $W := H_0 \cup H_1$. Additionally, the hypotheses admit uniform error control under H_0 and/or H_1 if and only if H_0 and/or H_1 are closed in W. Under compactness assumptions, uniform consistency is characterised by H_0 and H_1 having disjoint closures in the ambient space of probability measures. These criteria imply that — without regularity assumptions — conditional independence is not consistently testable. We introduce a Lipschitz-continuity assumption on the family of conditional distributions under which we recover testability of conditional independence with uniform error control under the null, with testable smoothness constraints.

1 Introduction

Whether a statistical hypothesis is testable remains an important question across the natural and social sciences. For example, judgements of conditional independence underlie many scientific inferences and are particularly fundamental in structural causal discovery (Spirtes et al., 1993). Recent results demonstrate that, unless regularity assumptions are made, conditional independence is not testable with finite-sample bounds on the probability of errors of the first type (Shah and Peters, 2020). Nevertheless, it was believed that conditional independence is consistently testable (Györfi and Walk, 2012). The true situation turns out to be more complicated (Neykov et al., 2021). These recent results in conditional independence testing are arrived at via ingenious ad hoc arguments. Despite recent developments, there remains no simple, unified criterion for characterizing all and only the testable statistical hypotheses. In this respect, statistics is in sharp contrast to the theory of computation, where the testable hypotheses (co-semidecidable sets) receive elegant complexity-theoretic characterizations (Kelly, 1996). In this paper, we attempt to remedy this situation by laying out general topological conditions for statistical testability, generalizing results from Dembo and Peres (1994), Ermakov (2017), Genin and Kelly (2017) and Kleijn (2022).

We demonstrate our results on the following elementary parametric examples since, contrary to many nonparametric hypotheses like conditional independence, their topological properties are evident.

[*]Department of Mathematics, VU Amsterdam, p.a.boeken@vu.nl

[†]Carl Friedrich von Weizsäcker Zentrum, Universität Tübingen, Center for Mathematics and Applications (NOVA Math), NOVA School of Science and Technology (NOVA FCT), eduardo.skapinakis@uni-tuebingen.de

[‡]Deptartment of Philosophy, University of Utah, konstantin.genin@utah.edu

[§]Korteweg-de Vries Institute for Mathematics, University of Amsterdam, j.m.mooij@uva.nl

Let $X_1, ..., X_n \sim \text{Bernoulli}(p)$ i.i.d. and consider for given $\varepsilon > 0$ the pairs of hypotheses

$$H_0 : p \in [0,1] \cap \mathbb{Q} \qquad H_1 : p \in [0,1] \setminus \mathbb{Q}. \tag{1}$$

$$H_0 : p \in [0,1] \cap \mathbb{Q} \qquad H_1 : p \in [0,1] \cap (\mathbb{Q} + \sqrt{2}) \tag{2}$$

$$H_0 : p \in [0,1/2] \qquad H_1 : p \in (1/2,1] \tag{3}$$

$$H_0 : p \in [0,1/2) \qquad H_1 : p \in (1/2,1] \tag{4}$$

$$H_0 : p \in [0,1/2-\varepsilon) \qquad H_1 : p \in (1/2+\varepsilon,1] \tag{5}$$

As might be intuitively clear, for these pairs of hypotheses one can achieve different consistency properties of the test, ranging from uniform consistency for (5) to mere consistency for (2), and (1) is not consistently testable. Our main results, presented in Section 2, state that these modes of testability are characterised by the topological properties of the hypotheses: whether they are F_σ (i.e. a countable union of closed sets), closed or open, or clopen with respect to the subspace topology on $W := H_0 \cup H_1$, or metrically separated. We don't restrict to parametric hypotheses, but we consider arbitrary nonparametric hypotheses H_0 and H_1 as subsets of the space of Borel probability measures $\mathcal{P}(\mathcal{X})$ on a separable metric space $\mathcal{X}$. The topological characterisations are with respect to the weak topology on $\mathcal{P}(\mathcal{X})$. The data is assumed to be i.i.d. from a single $\mathbb{P}$ in either H_0 or H_1. Notably, our results do not require any regularity conditions on the probability measures under consideration. However, it turns out that the weak topology characterises testability with *some* type of regularity: that the critical regions of the tests are open — a property that is useful when considering finite-precision measurements, to be further motivated in Section 1.1 below.

Our main application of these topological characterisations is to analyse the feasibility of conditional independence testing. For given sample spaces $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ and set of probability measures $W \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ we consider

$$H_0 := \{\mathbb{P} \in W : X \underset{\mathbb{P}}{\perp\!\!\!\perp} Y \mid Z\} \qquad H_1 := \{\mathbb{P} \in W : X \underset{\mathbb{P}}{\not\!\perp\!\!\!\perp} Y \mid Z\}, \tag{6}$$

where conditional independence, denoted with $X \perp\!\!\!\perp_{\mathbb{P}} Y \mid Z$, means that for all measurable A and B the factorisation $\mathbb{P}(X \in A, Y \in B \mid Z) = \mathbb{P}(X \in A \mid Z)\mathbb{P}(Y \in B \mid Z)$ holds $\mathbb{P}(Z)$-almost surely. Conditional dependence is denoted with $X \not\!\perp\!\!\!\perp_{\mathbb{P}} Y \mid Z$. In Section 3, we show that conditional independence and conditional dependence are both dense in the weak topology on $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$. It follows from the Baire category theorem that there does not exist a consistent FP-test for conditional independence when $W = \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$. Hence, conditional independence testing requires assumptions, that is, a restriction of the set W. We introduce the space of distributions with Lipschitz conditional distributions $\mathbb{P}(X \mid Z)$ or $\mathbb{P}(Y \mid Z)$ and investigate some properties in Section 4. We show that if the Lipschitz constant is bounded then this space is closed in the weak topology, and conditional independence is closed in this space. In Section 5 we discuss the consequences for various modes of testability of conditional independence.

1.1 Accommodating finite-precision measurements

Given a test $\varphi_n : \mathcal{X}^n \to \{0,1\}$ and a sample $x_1, ..., x_n$, the verdict of the test is given by the evaluation $\varphi_n(x_1, ..., x_n)$. From now on, we write $\{\varphi_n = i\}$ for the region $\{x \in \mathcal{X}^n : \varphi_n(x) = i\}$. In standard presentations of hypothesis testing, the acceptance and rejection regions of a test are taken to be arbitrary *measurable* sets. But is it really natural to consider a test that rejects if the sample point is rational-valued? Or if it is precisely π? If measurements of real-valued quantities can be made with only finite precision, such tests cannot be implemented.

In the following, we will consider tests whose decision regions $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$ are open in the topology on the sample space $\mathcal{X}^n$. If this condition is satisfied, some amount of finite precision is sufficient to determine the verdict of the test, no matter where in $\{\varphi_n = 0\}$ or $\{\varphi_n = 1\}$ the sample lands. If $\mathcal{X}^n$ is connected, $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$ cannot both be open and cover the entire sample space. Therefore, it is necessary to introduce the region $\{\varphi_n = 2\}$ that recommends neither acceptance nor rejection of H_0, but suspension of judgement.

Definition 1. A *finite-precision test (FP-test)* is a measurable map $\varphi_n : \mathcal{X}^n \to \{0,1,2\}$ such that $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$ are open.

We sometimes abuse terminology by referring to a *testing sequence* $\varphi = (\varphi_n)_{n\in\mathbb{N}}$ as a *test*. Note that if the sample lands in the boundary of $\{\varphi_n = 2\}$, no finite-precision measurement will be precise enough to verify that it is outside of the decision regions $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$. That situation is not so vicious: it will be impossible to discern the verdict of the test, which is practically equivalent to suspension of judgement. Every FP-test can be transformed into a binary test by merging $\{\varphi_n = 2\}$ with $\{\varphi_n = 0\}$ or $\{\varphi_n = 1\}$. For a consistent FP-test we have $\mathbb{P}^n(\varphi_n = 2) \to 0$, so consistency is maintained by transforming it into a binary test; for maintaining uniform consistency, once should be more careful. More details are given in Section 2.

Why is it reasonable to restrict our attention to finite-precision tests? After all, measurements are made with bounded, not arbitrary finite, precision. If a sample landing in $\{\varphi_n = 0\}$ is sufficiently close to its boundary, it might still be impossible to determine the verdict of the test if the precision required exceeds that of our measurement device. The first thing to notice is that this requirement is already more realistic than the usual one in which we impose no conditions on the complexity of the zones. Furthermore, our positive results are strengthened: if it is informative to learn that a hypothesis is testable with arbitrary zones, it is even more informative to learn that it is testable with arbitrary but finite precision. Negative results are still informative: if you cannot test a hypothesis with arbitrary but finite precision, then you certainly cannot test it with bounded precision.

A more explicit justification of the restriction to open regions requires a model of measurement. Let $\mathcal{B}$ be a basis for a topology on the sample space $\mathcal{X}$. We interpret $\mathcal{B}(x)$, the local basis at x, in the following way: the elements of $\mathcal{B}(x)$ are the possible outcome of measurements performed on x. In other words: we assume that the set of feasible measurements of x satisfies the properties of a local basis. Moreover, if $E \in \mathcal{B}(x)$, we assume that a sufficiently diligent empiricist measuring x will eventually produce a measurement $F \subseteq E$ in $\mathcal{B}(x)$ that is at least as precise as E. On this interpretation, it is always possible, with sufficient measurement effort, to determine whether a sample has landed in $\{\varphi = 0\}$ or $\{\varphi = 1\}$. The appropriate basis is determined by the available measurement protocol. If we are observing the outcomes of a coin toss, the appropriate basis is $\{\{H\}, \{T\}\}$, since it is possible to directly observe the outcome of a coin toss. For real-valued measurements, the usual basis of open intervals with rational endpoints is a natural candidate. For more on topological models of measurement, see Vickers (1990); Kelly (1996); Genin and Kelly (2017); Resende (2021).

1.2 Consistency, uniform consistency, and error control

Given a pair of hypotheses $H_0, H_1 \subseteq \mathcal{P}(\mathcal{X})$, a sequence $(\varphi_n)_{n\in\mathbb{N}}$ of tests $\varphi_n : \mathcal{X}^n \to \{0,1,2\}$ is *weakly consistent* if for $i \in \{0,1\}$ and all $\mathbb{P} \in H_i$ we have

$$\mathbb{P}^n(\varphi_n = i) \to 1,$$

and *strongly consistent* if for $i \in \{0,1\}$ and all $\mathbb{P} \in H_i$ we have

$$\mathbb{P}^\infty(\varphi_n \neq i \text{ for finitely many } n) = 1.$$

This is equivalent to the condition $\mathbb{P}^\infty(\liminf_n \{\varphi_n = i\}) = 1$ and to $\lim_{n\to\infty} \mathbb{P}^\infty(\exists m \geq n : \varphi_m \neq i) = 0$. We say that a pair of hypotheses is weakly (strongly) consistently testable if there exists a weakly (strongly) consistent testing sequence. Hypotheses which are strongly consistently testable are sometimes referred to as *discernible* (Dembo and Peres, 1994).

For binary tests, it has been shown by Nobel (2006) that the existence of a weakly consistent test is equivalent to the existence of strongly consistent test. This also applies to FP-tests, as is shown in the following result. A proof of the equivalence of weak and strong consistency of binary tests is achieved by changing one of the strict inequalities to an inequality in the definition of ψ_n below.

Theorem 1. *Given a pair of hypotheses $H_0, H_1 \subseteq \mathcal{P}(\mathcal{X})$, there exists a weakly consistent FP-test if and only if there exists a strongly consistent FP-test.*

Proof. Any strongly consistent test is also weakly consistent. Conversely, let $\varphi_n : \mathcal{X}^n \to \{0,1,2\}$ be weakly consistent, let $k_n := \lfloor \log(n) \rfloor$ and $m_n := \lfloor n/k_n \rfloor$. For $i \in \{0,1\}$ define the random variable $Y^i_{n,j} := \mathbb{1}\{\varphi_{k_n}(X_{k_n j+1}, ..., X_{k_n j+k_n}) = i\}$ such that $Y^i_{n,0}, ..., Y^i_{n,m_n-1}$ are i.i.d., and let

$$\psi_n(X_1, ..., X_n) := \begin{cases} 0 & \text{if } \frac{1}{m_n}\sum_{j=0}^{m_n-1} Y^0_{n,j} > 1/2; \\ 1 & \text{if } \frac{1}{m_n}\sum_{j=0}^{m_n-1} Y^1_{n,j} > 1/2; \\ 2 & \text{otherwise.} \end{cases}$$

Fix $i \in \{0,1\}$. For $\mathbb{P} \in H_i$ we have $\mu^i_n := \mathbb{E}_{\mathbb{P}}[Y^i_{n,j}] = \mathbb{P}^n(\varphi_{k_n}(X_{k_n j+1}, ..., X_{k_n j+k_n}) = i) \to 1$ by weak consistency of φ_n, and hence for every $\varepsilon \in (0, 1/2)$ there is an $N_{\mathbb{P}}$ such that $1/2 \le \mu^i_n - \varepsilon$ for all $n \ge N_{\mathbb{P}}$, so using Hoeffding's inequality this gives

$$\begin{aligned} \mathbb{P}^n(\psi_n \neq i) = \mathbb{P}^n\left(\frac{1}{m_n}\sum_{j=0}^{m_n-1} Y^i_{n,j} \le 1/2\right) &\le \mathbb{P}^n\left(\frac{1}{m_n}\sum_{j=0}^{m_n-1} (Y^i_{n,j} - \mu^i_n) \le -\varepsilon\right) \\ &\le \mathbb{P}^n\left(\left|\frac{1}{m_n}\sum_{j=0}^{m_n-1} (Y^i_{n,j} - \mu^i_n)\right| > \varepsilon\right) \le 2e^{-2m_n\varepsilon^2}, \end{aligned}$$

we get $\sum_{n \ge N_{\mathbb{P}}} e^{-2m_n\varepsilon^2} < \infty$, hence $\sum_{n=1}^{\infty} \mathbb{P}^n(\psi_n \neq i) < \infty$, so the result follows from the Borel-Cantelli lemma. ■

On top of consistency, one can consider various notions of error control: bounds on error probabilities, asymptotically or for finite samples, and uniform consistency, all under H_0 and/or H_1. For binary tests, Pfanzagl (1968) has shown that the existence of a uniformly weakly consistent test (property (8) below) is equivalent to the existence of a uniformly strongly consistent test (property (9)). The following result shows that testability with various asymptotic notions of uniform error control are equivalent, both for one-sided and two-sided error control. If for only one of the hypotheses the error is uniformly controlled asymptotically, then it can also be controlled at finite sample sizes. The following theorem is stated in terms of FP-tests, but the result also holds for binary tests, following the same alteration as suggested right before Theorem 1.

Theorem 2. *Let a pair of hypotheses $H_0, H_1 \subseteq \mathcal{P}(\mathcal{X})$ be given, and let $K \subseteq \{H_0, H_1\}$ be a set of hypotheses. For every $\alpha > 0$ there exists a consistent FP-testing sequence $(\varphi_n)_{n\in\mathbb{N}}$ with for all $H_i \in K$*

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_i} \mathbb{P}^n(\varphi_n \neq i) < \alpha \tag{7}$$

if and only if there exists a consistent FP-testing sequence $(\varphi_n)_{n\in\mathbb{N}}$ with any of the following uniform consistency properties for all $H_i \in K$:

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_i} \mathbb{P}^n(\varphi_n \neq i) = 0; \tag{8}$$

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_i} \mathbb{P}^\infty(\exists m \ge n : \varphi_m \neq i) = 0; \tag{9}$$

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_i} \sum_{m\ge n}^{\infty} \mathbb{P}^m(\varphi_m \neq i) = 0. \tag{10}$$

If for only a single hypothesis the errors are controlled, say $K = \{H_i\}$, the above is equivalent to the statement that for every $\alpha > 0$ there exists a consistent FP-testing sequence $(\varphi_n)_{n\in\mathbb{N}}$ with any of the following properties for all $n \in \mathbb{N}$:

$$\sup_{\mathbb{P}\in H_i} \mathbb{P}^n(\varphi_n \neq i) \le \alpha; \tag{11}$$

$$\sup_{\mathbb{P}\in H_i} \mathbb{P}^\infty(\exists n : \varphi_n \neq i) \le \alpha; \tag{12}$$

$$\sup_{\mathbb{P}\in H_i} \sum_{n=1}^{\infty} \mathbb{P}^n(\varphi_n \neq i) \le \alpha. \tag{13}$$

Throughout, it is equivalent to consider weak or strong consistency of the tests.

Proof. Every claim with strong consistency implies the corresponding claim with weak consistency.

Let $K \subseteq \{H_0, H_1\}$.

(7) with weak consistency $\implies$ (10) with strong consistency: Let $\varepsilon \in (0, 1/2)$ and let $\varphi_n : \mathcal{X}^n \to \{0, 1, 2\}$ be such that there is a N such that $\sup_{\mathbb{P} \in H_i} \mathbb{P}^n(\varphi_n \neq i) \leq 1/2 - \varepsilon$ for all $n \geq N$ for all $H_i \in K$. Then equivalently $\inf_{\mathbb{P} \in H_i} \mathbb{P}^n(\varphi_n = i) \geq 1/2 + \varepsilon$. With ψ_n and $N_{\mathbb{P}}$ as defined in the proof of Theorem 1, we have that $N_{\mathbb{P}} = N$ is independent of $\mathbb{P}$, and we obtain $\mathbb{P}^n(\psi_n \neq i) \leq 2e^{-2m_n\varepsilon^2}$ for all $n \geq N$, and hence $\sup_{\mathbb{P} \in H_i} \sum_{n \geq N} \mathbb{P}^n(\psi_n \neq i) \leq \sum_{n \geq N} e^{-2m_n\varepsilon^2} < \infty$. As the tail of a convergent sequence, we get $\sup_{\mathbb{P} \in H_i} \sum_{m \geq n} \mathbb{P}^m(\psi_m \neq i) \to 0$.

(10) $\implies$ (9) $\implies$ (8) $\implies$ (7) If for every $\alpha > 0$ there is an N such that $\sup_{\mathbb{P} \in H_i} \sum_{m \geq n} \mathbb{P}^m(\varphi_m \neq i) \leq \alpha$ for all $n \geq N$, then

$$\mathbb{P}^n(\varphi_n \neq i) \leq \mathbb{P}^\infty(\exists m \geq n : \varphi_m \neq i) \leq \sum_{m \geq n} \mathbb{P}^m(\varphi_m \neq i) \leq \alpha$$

for all $n \geq N$, so we obtain the implications. This holds both with weak and strong consistency. Further, without loss of generality, let $K = \{H_0\}$.

(10) $\implies$ (13) For every $\alpha > 0$ there is an N such that $\sup_{\mathbb{P} \in H_0} \sum_{m \geq n} \mathbb{P}^n(\varphi_n \neq 0) \leq \alpha$ for all $n \geq N$. Then $\psi_n := \mathbb{1}\{n \geq N\}\varphi_n$ satisfies (13), both for weak and strong consistency.

(13) $\implies$ (12) $\implies$ (11) This follows from the union bound

$$\mathbb{P}^n(\varphi_n = 1) \leq \mathbb{P}^\infty(\exists n : \varphi_n = 1) \leq \sum_{n=1}^{\infty} \mathbb{P}^n(\varphi_n = 1) \leq \alpha,$$

both with weak and strong consistency.

(11) $\implies$ (7) Follows immediately, both with weak and strong consistency. ∎

We will loosely refer to these conditions as *uniform error control under H_i*. However, interpreting the conditions of the FP-tests in Theorem 2 as 'error control' is a slight misnomer, since the outcome $\varphi_n = 2$ should be interpreted as suspension of judgement and not as an error. For one-sided error control on say $K = \{H_i\}$ this is a futile discussion: Theorem 4 below shows that the existence of a consistent FP-test which satisfies (8) is equivalent to the existence of a consistent FP-test which satisfies the condition $\lim_n \sup_{\mathbb{P} \in H_i} \mathbb{P}^n(\varphi_n = 1 - i) = 0$. A different interpretation of this apparent distinction is whether one wants the measure of the suspension zone $\{\varphi_n = 2\}$ to converge to 0 pointwise, or uniformly over all $\mathbb{P} \in H_i$.

For two-sided error control this distinction *is* important. Pointwise or uniform convergence of the measure of the suspension zone are essentially different, and they correspond to different topological properties of the hypotheses, as shown below in Section 2.

1.3 Related literature

In a sufficiently regular parametric setting, Berger and Wald (1949) provide necessary and sufficient conditions for the existence of a test with Type-I and Type-II error control with $\alpha = \beta = 1/2$. A rather technical characterisation for the existence of uniformly consistent tests has been given by Berger (1951). A topological characterisation has been given by LeCam and Schwartz (1960) in terms of a topology of setwise convergence on the space $\cup_{n=1}^{\infty} \mathcal{P}(\mathcal{X})^n$ of all n-fold products of probability measures on $\mathcal{X}$. The downside of this condition is that it is hard to verify since this topology is not metrisable, not first countable and hence "not very easily accessible" (LeCam and Schwartz, 1960) — see also Kleijn (2022). In contrast, later work (and this paper as well) considers topological properties of H_0, H_1 as subsets of $\mathcal{P}(\mathcal{X})$, instead of the space of all n-fold products. Cover (1973) considered testing whether the bias of a coin is rational or irrational (example (1)), and found that there exists a consistent test for testing the null of rational bias versus the alternative of irrational bias, if a subset of Lebesgue measure zero is removed from the alternative. Dembo and Peres (1994) show that if H_0 and H_1 are disjoint sets which are F_σ in the weak topology on $W = H_0 \cup H_1$, then there exists a strongly consistent test. For the converse direction, they note that the hypotheses $H_0 := \{\delta_x : x \in [0,1] \cap \mathbb{Q}\}$ and $H_1 := \{\delta_x : x \in [0,1] \setminus \mathbb{Q}\}$ are discernible by the test $\varphi_n(x_1, ..., x_n) := \mathbb{1}_{[0,1]\setminus\mathbb{Q}}(x_1)$, but H_1 is not F_σ[1] so *some* regularity condition has to be imposed for such a topological characterisation of discernibility. Dembo and Peres (1994) prove this characterisation under the assumption that every measure in W has a $p > 1$-integrable density, and Kleijn (2022) (Corollary 9.4.23) weakens this assumption to uniformly integrable densities. We show without any assumptions on the hypotheses H_0 and H_1 that consistent FP-testability is characterised by hypotheses that are F_σ in the weak topology, so requiring the regions $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$ to be open is *the* regularity condition which resolves the issue raised by Dembo and Peres (1994).

Ermakov (2017) (Theorem 4.4) characterises the existence of consistent tests in terms of the topology of setwise convergence under the assumption of σ-compactness of W (which is satisfied if e.g., one assumes $p > 1$-integrable densities), in which case the topology of setwise convergence is equivalent to the weak topology (Gänssler, 1971, Lemma 2.3). For a comprehensive overview of the literature on consistent hypothesis testing, we refer to Kleijn (2022), Chapter 9.

Larsson et al. (2026) show that a test — as general $[0,1]$-valued random variable on the underlying measurable space, so without i.i.d. assumption — that satisfies the condition $\sup_{\mathbb{P}\in H_0} \mathbb{E}_\mathbb{P}[\varphi] < \inf_{\mathbb{P}\in H_1} \mathbb{E}_\mathbb{P}[\varphi]$ exists if and only if the weak* closures (in the space of bounded finitely additive measures) of the convex hulls of H_0 and H_1 are disjoint.

Our results in Section 2 are largely based on the work by Genin and Kelly (2017), who work under the assumption that the sample space $\mathcal{X}$ has a subbasis $\mathcal{O}$ such that $\mathbb{P}(\partial A) = 0$ for all $\mathbb{P} \in W$ and $A \in \mathcal{O}$ — this assumption is satisfied if e.g., all measures in W have a density with respect to some dominating measure (Bogachev, 2007, Proposition 8.2.8). The results of Genin and Kelly (2017) are applied to the study of causal discovery in Genin and Mayo-Wilson (2020); Genin (2021); Genin and Mayo-Wilson (2024).

Regarding conditional independence testing, Shah and Peters (2020) and Neykov et al. (2021) show, in the setting where the variables are real-valued and the probability measures have densities, that no conditional independence test with Type-I error control and consistency under the alternative exists. Lundborg et al. (2022) generalise this to a specific setting where the samples are $L^2([0,1], \mathbb{R})$ functions, e.g. continuous-time stochastic processes. Györfi and Walk (2012) provide a conditional independence test and prove that it is strongly consistent, but Neykov et al. (2021) point out a mistake in their proof, so it remains an open question whether conditional independence is consistently testable. We answer this question for FP-testability by showing that if X, Y, Z take values in arbitrary Polish spaces $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ where $\mathcal{Z}$ has no isolated points, there exists no consistent FP-test for conditional independence $X \perp\!\!\!\perp Y \mid Z$.

In Section 4 we show that conditional independence is weakly closed under similar conditions as considered by Barbie and Gupta (2014); we provide a more direct, alternative proof. See Section 4.3 for further discussion of related literature In Section 5, we provide sufficient conditions for the consistent testability of conditional independence with and without with uniform error control under H_0. Our conditions for testability with error control are similar to those considered by Warren (2021) and Neykov et al. (2021). To the best of our knowledge, our conditions for consistent testability are novel: we require that only one of the maps

$$z \mapsto \mathbb{P}(X \mid Z = z),$$
$$z \mapsto \mathbb{P}(Y \mid Z = z)$$

satisfies a regularity condition. For a more in-depth comparison with existing literature, see Section 5.1.

1.4 The weak topology

We will characterise FP-testability in terms of topological properties of the hypotheses $H_0, H_1 \subseteq \mathcal{P}(\mathcal{X})$, where $\mathcal{X}$ is a separable metric space, and $\mathcal{P}(\mathcal{X})$ denotes the set of probability measures on the Borel σ-algebra $\mathcal{B}(\mathcal{X})$ on $\mathcal{X}$. Recall that a sequence of probability measures $\mathbb{P}_1(X), \mathbb{P}_2(X), ...$ converges

[1]This holds since $[0,1] \setminus \mathbb{Q}$ is not F_σ and $x \mapsto \delta_x$ is a homeomorphism (Bogachev, 2007, Lemma 8.9.2).

weakly to another probability measure $\mathbb{P}(X)$, denoted with $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$, if $\int f(x)\mathrm{d}\mathbb{P}_n(x) \to \int f(x)\mathrm{d}\mathbb{P}(x)$ for all continuous functions $f : \mathcal{X} \to \mathbb{R}$. The *weak topology* on the space of Borel probability measures $\mathcal{P}(\mathcal{X})$ is the smallest topology that makes the maps $\mathbb{P} \mapsto \int f(x)\mathrm{d}\mathbb{P}$ continuous for all bounded continuous f. Because $\mathcal{X}$ is separable, the weak topology is separable and metrisable, for example by the *bounded Lipschitz metric*[2] d_{BL}, defined by

$$d_{BL}(\mathbb{P}_0(X), \mathbb{P}_1(X)) := \sup\left\{\left|\int f\mathrm{d}(\mathbb{P}_0 - \mathbb{P}_1)\right| : f \in \mathrm{BL}(\mathcal{X};\mathbb{R})\right\},$$

where $\mathrm{BL}(\mathcal{X};\mathbb{R}) := \left\{f : \mathcal{X} \to \mathbb{R} \;\middle|\; \sup_{x\neq x'} \frac{|f(x)-f(x')|}{d(x,x')} \le 1 \text{ and } \|f\|_\infty \le 1\right\}$. If $\mathcal{X}$ is complete, then d_{BL} is complete as well. See also Bogachev (2007), Theorem 8.3.2. Since the weak topology is sequential, convergence in the weak topology coincides with weak convergence. Weak convergence $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ is equivalent to the condition that $\liminf_n \mathbb{P}_n(A) \ge \mathbb{P}(A)$ for all $A \subseteq \mathcal{X}$ open, and to the condition that $\limsup_n \mathbb{P}_n(B) \le \mathbb{P}(B)$ for all $B \subseteq \mathcal{X}$ closed. This is also known as the 'Portmanteau theorem' (Bogachev, 2007, Theorem 8.2.3). A subbasis for the weak topology is given by sets $\{\mathbb{P} : \mathbb{P}(A) > q\}$, with $A \subseteq \mathcal{X}$ open and $q \in [0, 1]$; see for example Bogachev (2007), Section 8.2. This means that for any weakly open set $H \subseteq \mathcal{P}(\mathcal{X})$, there exist open sets $A_{ij} \subseteq \mathcal{X}$ and $q_{ij} \in [0, 1]$ such that $H = \bigcup_{i\in I} \bigcap_{j=1}^{m_i} \{\mathbb{P} : \mathbb{P}(A_{ij}) > q_{ij}\}$ for some index set I. Because $\mathcal{X}$ and hence $\mathcal{P}(\mathcal{X})$ are assumed to be separable, the index set I can be taken to be countable.

2 Main results

Let $\mathcal{X}$ be a separable metric space, and let H_0, H_1 be disjoint sets of Borel probability measures on $\mathcal{X}$.

Theorem 3. *The following are equivalent:*

1. *there exists a consistent FP-test;*
2. *H_0 and H_1 are F_σ in the weak topology on $W := H_0 \cup H_1$.*

Theorem 4. *The following are equivalent:*

1. *there exists a consistent FP-test φ_n with*

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_0} \mathbb{P}^n(\varphi_n = 1) = 0;$$

2. *there exists a consistent FP-test φ_n with*

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_0} \mathbb{P}^n(\varphi_n \neq 0) = 0;$$

3. *H_0 is closed in the weak topology on $W := H_0 \cup H_1$;*

Theorem 5. *The following are equivalent:*

1. *there exists a consistent FP-test φ_n with for $i \in \{0, 1\}$:*

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_i} \mathbb{P}^n(\varphi_n = 1 - i) = 0;$$

2. *H_0 and H_1 are clopen in the weak topology on $W := H_0 \cup H_1$.*

Theorem 6. *Let H_0 and H_1 be contained in some compact set in the weak topology, then the following are equivalent:*

[2] The bounded Lipschitz metric is strongly equivalent to the *Kantorovic-Rubinstein metric*, and weaker than the *p-Wasserstein metric* W_p (we have $d_{BL} \le W_1 \le W_p$, with equality $d_{BL} = W_1$ if the measures under consideration have bounded support (Bogachev, 2007, Theorem 8.10.45)) and the total variation metric d_{TV}.

1. *there exists an FP-test φ_n with $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$ having disjoint closures, and for $i \in \{0,1\}$:*

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_i} \mathbb{P}^n(\varphi_n \neq i) = 0;$$

2. *H_0 and H_1 have disjoint closures in $\mathcal{P}(\mathcal{X})$.*

The proofs are given below in Section 2.1. Throughout one can equivalently consider strong consistency and weak consistency, by Theorem 1. Theorem 2 shows that the uniform consistency of Theorem 4 can equivalently be stated as Type-I error control, either asymptotically, or for finite samples. The uniform consistency of Theorem 6 can equivalently be stated as asymptotic Type-I and Type-II error control. Similarly, it can be shown that the uniform consistency of Theorem 5 can equivalently be stated as Type-I and Type-II error control, either asymptotically or for finite samples.

By swapping the roles of H_0 and H_1 in Theorem 4, we obtain that H_1 is closed if and only if the pair is consistently FP-testable with uniform error control under H_1.

The condition in Theorem 6 that H_0 and H_1 have disjoint closures is (under the assumption of relative compactness) equivalent to being metrically separated, that is, $d_{BL}(H_0, H_1) > 0$. Note that by Prokhorov's theorem, the relative compactness condition is equivalent to uniform tightness of the set of probability measures in W. This compactness assumption is not *necessary* for uniformly consistent FP-testability: one can uniformly consistently test whether the mean of a Gaussian with fixed variance is smaller than 0 or larger than 1, and this space of hypotheses is not tight.

The topological condition of each theorem implies the topological condition of the previous theorem: metric separation implies the hypotheses to be clopen in the subspace topology, which implies H_0 to be closed, which implies H_0 and H_1 to be F_σ (since the weak topology is a separable metric topology).

These results are readily applicable to the example from the introduction, since the parameter space $[0,1]$ with the Euclidean topology is homeomorphic to the set of Bernoulli distributions equipped with the weak topology. The hypothesis H_1 in (1) is not F_σ, so there does not exist a consistent FP-test for this problem. The hypotheses from (2) are both F_σ so there exists a consistent test, but no form of uniform error control is possible since neither H_0 nor H_1 is closed. In (3) H_0 is closed, so there exists a consistent FP-test with uniform error control under H_0. Both hypotheses of (4) are clopen in the relative topology on W, so there exists a consistent FP-test with the type of error control as in Theorem 5, but uniform consistency as in Theorem 6 is not feasible. In (5) the two hypotheses are relatively compact and have disjoint closures in the ambient space, so there exists a uniformly consistent test.

In general, these sufficient conditions for various types of testability are not constructive: the tests of Theorems 3, 4 and 5 are constructed by expressing the hypotheses in terms of the subbasis elements $\{\mathbb{P} \in W : \mathbb{P}(A) > q\}$ of the weak topology. For example, in Theorem 4 one might know that H_1 is open without knowing an explicit representation in terms of the subbasis elements. However, for the examples (1)–(4) these subbasis representations are easily derived, allowing us to give explicit tests, following the constructions in the proofs. We leave this as an exercise to the reader. The test of Theorem 6 requires the computation of the BL-distance between the empirical measure and the hypotheses, which might be computationally intractable as well.

These sufficient conditions for FP-testability also imply binary testability with corresponding modes of error control, by joining the suspension zone $\{\varphi_n = 2\}$ with either $\{\varphi_n = 0\}$ or $\{\varphi_n = 1\}$. In Theorem 3, merging the suspension zone with $\{\varphi_n = 0\}$ or $\{\varphi_n = 1\}$ both give a consistent test. Actually, the topological condition of H_0 and H_1 being F_σ is equivalent to the existence of two binary testing sequences: one with $\{\varphi_n = 0\}$ open, and one with $\{\varphi_n = 1\}$ open. This can rather straightforwardly be deduced from the proof of Theorem 3. For the test of clause 2 of Theorem 4 it does not matter for maintaining error control whether $\{\varphi_n = 2\}$ is joined with $\{\varphi_n = 0\}$ or $\{\varphi_n = 1\}$, but for the test of clause 1 the suspension zone must be joined with $\{\varphi_n = 0\}$ to maintain the error control under H_0. It can be shown that H_0 being closed is equivalent to the existence of a binary testing sequence with $\{\varphi_n = 1\}$ open and uniform error control under H_0. For Theorem 5 the FP-test cannot be converted to a binary test while maintaining the same error control. Finally, for Theorem 6 the binary test can be constructed both ways while maintaining uniform consistency.

Results similar to Theorems 3 – 6 exist in the literature. Theorem 3 generalises Dembo and Peres (1994) (Theorem 2) who provide the similar characterisation that, under the assumption that all measures in W have uniformly integrable densities, there exists a consistent binary test if and only if the hypotheses are F_σ in the weak topology. Genin and Kelly (2017) (Theorem 4.3) drop the uniform integrability assumption (but the assumption of having densities remains), for which they show that the F_σ condition is equivalent to the existence of a ternary testing sequence $(\varphi_n)_{n\in\mathbb{N}}$ with $\mathbb{P}(\partial\{\varphi_n = 0\}) = \mathbb{P}(\partial\{\varphi_n = 1\}) = 0$ for all $\mathbb{P} \in W$. Ermakov (2017) (Theorem 4.4) considers the topology of setwise convergence instead of the weak topology for which he shows – under the assumption that W is contained in a σ-compact set – that the F_σ condition is equivalent to the existence of a binary testing sequence, without any regularity of the critical regions.

An analogue of Theorem 4 is also shown by Genin and Kelly (2017) (Theorem 4.1) under the assumption of having densities, for binary tests with $\mathbb{P}(\partial\{\varphi_n = 0\}) = \mathbb{P}(\partial\{\varphi_n = 1\}) = 0$ for all $\mathbb{P} \in W$. Ermakov (2017) (Theorem 4.3) shows that when H_0 and H_1 are contained in respectively a compact set and a σ-compact set, consistent binary testability with uniform consistency under H_0 is equivalent to H_0 being closed and H_1 being F_σ in the topology of setwise convergence.

Theorem 5 is also given by Genin (2018) (Theorem 3.2.3) under the assumption of having densities, for ternary tests with $\mathbb{P}(\partial\{\varphi_n = 0\}) = \mathbb{P}(\partial\{\varphi_n = 1\}) = 0$ for all $\mathbb{P} \in W$.

Theorem 6 is similar to Ermakov (2017) (Theorem 4.1), who shows that if W is contained in a compact set in the topology of setwise convergence, then H_0 and H_1 having disjoint closures is equivalent to the existence of a consistent binary test. Note that compactness in the topology of setwise convergence implies that the weak topology and the topology of setwise convergence coincide, so we obtain weaker sufficient conditions for the existence of uniformly consistent tests.

2.1 Proofs

The proofs that the existence of certain tests imply certain topological conditions on the hypotheses are stand-alone. There is however a dependency between the proofs that certain topological conditions imply the existence of FP-tests with the desired properties. We first prove Theorem 4. From that one, we can build under F_σ conditions consistent tests in Theorem 3, and under clopen conditions we can build tests with the two-sided error control as required in Theorem 5. To prove Theorem 6, we use an entirely different proof strategy using convergence of empirical measures.

2.1.1 Proof of Theorem 4

The main argument for the existence of a test with uniform error control is that if H_1 is open, we can write it as a countable union of finite intersections of subbasis elements $\{\mathbb{P} \in W : \mathbb{P}(A) > q\}$. For each disjoint pair of subbasis elements there exists a test with the required error control. The proofs largely follow the structure of the proof of Genin and Kelly (2017), Theorem 4.1.

Lemma 1. *Let $W \subseteq \mathcal{P}(\mathcal{X})$ be given, let $A \subseteq \mathcal{X}$ be open and let $q \in [0, 1]$. For the hypotheses $H_0 := \{\mathbb{P} \in W : \mathbb{P}(A) \leq q\}$ and $H_1 = \{\mathbb{P} \in W : \mathbb{P}(A) > q\}$, for every $\alpha > 0$ there exists a strongly consistent FP-testing sequence $(\varphi_n)_{n\in\mathbb{N}}$ with $\sup_{\mathbb{P}\in H_0} \sum_{n=1}^{\infty} \mathbb{P}^n(\varphi_n \neq 0) \leq \alpha$.*

Proof. Let $\alpha > 0$ be given, let $A^c_{1/n} := \bigcup_{a\in A^c} B(a, 1/n)$ be the $1/n$-neighborhood of A^c, and define

$$\varphi_n(X_1, ..., X_n) := \begin{cases} 0 & \text{if } \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{A^c_{1/n}}(X_i) > 1 - q - t_n; \\ 1 & \text{if } \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\text{ext}(A^c_{1/n})}(X_i) \geq q + t_n; \\ 2 & \text{otherwise,} \end{cases}$$

where $t_n := \sqrt{\frac{1}{2n}\ln(\pi^2 n^2/6\alpha)}$.

For $\mathbb{P} \in H_0$ we have $\mathbb{P}(A) \leq q$, so the fact that $A^c_{1/n} \supseteq A^c$ and Hoeffding's inequality gives

$$\mathbb{P}^n(\varphi_n \neq 0) = \mathbb{P}^n\left(\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{A^c_{1/n}}(X_i) \leq 1 - q - t_n\right) \leq \mathbb{P}^n\left(\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{A^c}(X_i) \leq 1 - q - t_n\right)$$
$$= \mathbb{P}^n\left(\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{A}(X_i) \geq q + t_n\right) \leq \mathbb{P}^n\left(\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{A}(X_i) \geq \mathbb{P}(A) + t_n\right) \leq e^{-2nt_n^2} = \frac{6\alpha}{\pi^2 n^2},$$

hence $\sum_{n=1}^\infty \mathbb{P}^n(\varphi_n \neq 0) \leq \alpha$. Strong consistency under the null then follows from the Borel-Cantelli lemma.

If $\mathbb{P} \in H_1$ then $\mathbb{P}(A) > q$. Since $A = \bigcup_n \operatorname{ext}(A^c_{1/n})$, we have that $\mathbb{P}(\operatorname{ext}(A^c_{1/n})) \uparrow \mathbb{P}(A)$. Therefore, there is an $M \in \mathbb{N}$ such that $\mathbb{P}(\operatorname{ext}(A^c_{1/M})) > q$. Since $\operatorname{ext}(A^c_{1/n}) \supseteq \operatorname{ext}(A^c_{1/M})$ for all $n \geq M$ we have for all these n that $\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\operatorname{ext}(A^c_{1/n})}(X_i) \geq \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\operatorname{ext}(A^c_{1/M})}(X_i)$, which by the strong law of large numbers converges almost surely to $\mathbb{P}(\operatorname{ext}(A^c_{1/M})) > q$. Combined with the fact that $t_n \downarrow 0$ we almost surely have that $\frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\operatorname{ext}(A^c_{1/n})}(X_i) \geq q + t_n$ for all sufficiently large n, so $\varphi_n \to 1$ a.s.

To prove that $\{\varphi_n = 0\}$ is open, let for $\gamma \in \{0,1\}^n$ the set $(A^c_{1/n})^\gamma \subseteq \mathcal{X}^n$ be the Cartesian product of n sets, where the i-th set is $A^c_{1/n}$ if $\gamma_i = 1$, and $\mathcal{X}$ if $\gamma_i = 0$. For example, $\gamma = (1, ..., 1, 0)$ gives $(A^c_{1/n})^\gamma = A^c_{1/n} \times ... \times A^c_{1/n} \times \mathcal{X}$. Since $A^c_{1/n}$ is open, so is $(A^c_{1/n})^\gamma$, hence $\{\varphi_n = 0\} = \bigcup\{(A^c_{1/n})^\gamma : \gamma \in \{0,1\}^n, |\gamma| > n(1 - q - t_n)\}$ is open as well. Since $\operatorname{ext}(A^c_{1/n})$ is open, we have analogously that $\{\varphi_n = 1\}$ is open as well. ∎

The following lemma shows that for pairs of hypotheses which are consistently testable with uniform error control under the null, the alternative hypotheses enjoy a topological structure: they are closed under countable unions and finite intersections. For convenience in proving Theorem 4, the result is stated in terms of strong consistency and control of the sum over n of all errors, but also holds for weak consistency and error control pointwise in n.

Lemma 2. *Let $\{(H_0^{ij}, H_1^{ij}) : (i,j) \in \mathbb{N}^2\}$ be pairs of disjoint hypotheses such that for every pair (H_0^{ij}, H_1^{ij}) and every $\alpha^{ij} > 0$ there exists a strongly consistent FP-test φ_n^{ij} with $\sup_{\mathbb{P}\in H_0}\sum_{n=1}^\infty \mathbb{P}^n(\varphi_n^{ij} \neq 0) \leq \alpha^{ij}$, then for the hypotheses $H_0 := \bigcap_{i=1}^\infty \bigcup_{j=1}^{m_i} H_0^{ij}$ and $H_1 := \bigcup_{i=1}^\infty \bigcap_{j=1}^{m_i} H_1^{ij}$ and any $\alpha > 0$ there exists a strongly consistent FP-test φ_n with $\sup_{\mathbb{P}\in H_0}\sum_{n=1}^\infty \mathbb{P}^n(\varphi_n \neq 0) \leq \alpha$.*

Proof. Let $\alpha > 0$ be given. For the hypotheses (H_0^{ij}, H_1^{ij}), let φ_n^{ij} be a strongly consistent FP-test with $\sum_{n=1}^\infty \mathbb{P}^n(\varphi_n^{ij} \neq 0) \leq \alpha/2^i$. Let φ_n be the FP-test defined by

$$\{\varphi_n = 0\} := \bigcap_{i=1}^n \bigcup_{j=1}^{m_i} \{\varphi_n^{ij} = 0\}$$
$$\{\varphi_n = 1\} := \bigcup_{i=1}^n \bigcap_{j=1}^{m_i} \{\varphi_n^{ij} = 1\}$$
$$\{\varphi_n = 2\} := \mathcal{X}^n \setminus (\{\varphi_n = 0\} \cup \{\varphi_n = 1\}),$$

then φ_n has uniform error control under the null: if $\mathbb{P} \in H_0$, then for every $i \in \mathbb{N}$ there is a $j_i \leq m_i$ such that $\mathbb{P} \in H_0^{ij_i}$, and since $\{\varphi_n \neq 0\} = \bigcup_{i=1}^n \bigcap_{j=1}^{m_i} \{\varphi_n^{ij} \neq 0\} \subseteq \bigcup_{i=1}^\infty \{\varphi_n^{ij_i} \neq 0\}$ we have

$$\sum_{n=1}^\infty \mathbb{P}^n(\varphi_n \neq 0) \leq \sum_{n=1}^\infty \mathbb{P}^n\left(\cup_{i=1}^\infty \{\varphi_n^{ij_i} \neq 0\}\right) \leq \sum_{n=1}^\infty \sum_{i=1}^\infty \mathbb{P}^n(\varphi_n^{ij_i} \neq 0) \leq \alpha.$$

The Borel-Cantelli lemma then gives strong consistency under H_0.

For every $\mathbb{P} \in H_1$ there is an $i \in \mathbb{N}$ such that $\mathbb{P} \in \bigcap_{j=1}^{m_i} H_1^{ij}$. Because of the inclusion $\cap_{j=1}^{m_i}\{\varphi_n^{ij} = 1\} \subseteq \{\varphi_n = 1\}$ and the Fréchet inequality[3] we have for $n \geq i$ that

$$\begin{aligned}\mathbb{P}^\infty(\liminf_n\{\varphi_n = 1\}) &= \mathbb{P}^\infty(\liminf_{n\geq i}\{\varphi_n = 1\}) \\ &\geq \mathbb{P}^\infty(\liminf_n \cap_{j=1}^{m_i}\{\varphi_n^{ij} = 1\}) \geq \sum_{j=1}^{m_i} \mathbb{P}^\infty(\liminf_n\{\varphi_n^{ij} = 1\}) - m_i + 1 = 1,\end{aligned}$$

so φ_n is strongly consistent. ■

We are now able to prove Theorem 4:

Proof.

1 $\implies$ 3 By Theorem 2, 1 is equivalent to the existence (for every $\alpha > 0$) of a weakly consistent FP-test with $\sup_{\mathbb{P}\in H_0} \mathbb{P}^n(\varphi_n = 1) \leq \alpha$ for all n. If H_0 is not closed, then there is a $\mathbb{Q} \in H_1$ and a sequence $\{\mathbb{P}_m\}_{m\in\mathbb{N}}$ in H_0 such that $\mathbb{P}_m \xrightarrow{w} \mathbb{Q}$. The map $\mathbb{P} \mapsto \mathbb{P}^n$ is continuous (Billingsley, 1999, Theorem 2.8) hence $\mathbb{P}_m^n \xrightarrow{w} \mathbb{Q}^n$. Let φ_n be an FP-test with $\sup_{\mathbb{P}\in H_0} \mathbb{P}^n(\varphi_n = 1) \leq \alpha$ for all n, then $\{\varphi_n = 1\}$ is open so by the Portmanteau theorem we have $\mathbb{Q}^n(\varphi_n = 1) \leq \liminf_m \mathbb{P}_m^n(\varphi_n = 1) \leq \alpha$, and hence φ_n cannot be weakly consistent at $\mathbb{Q}$.

3 $\implies$ 2 If H_1 is open, then we can write H_0 and H_1 as

$$H_0 = \bigcap_{i=1}^{\infty}\bigcup_{j=1}^{m_i}\{\mathbb{P} \in W : \mathbb{P}(A_{ij}) \leq q_{ij}\} \qquad H_1 = \bigcup_{i=1}^{\infty}\bigcap_{j=1}^{m_i}\{\mathbb{P} \in W : \mathbb{P}(A_{ij}) > q_{ij}\}$$

for $A_{ij} \subseteq \mathcal{X}$ open and $q_{ij} \in [0, 1]$. By Lemma 1, for each of these pairs of hypotheses $H_0^{ij} := \{\mathbb{P} \in W : \mathbb{P}(A_{ij}) \leq q_{ij}\}$ and $H_1^{ij} := \{\mathbb{P} \in W : \mathbb{P}(A_{ij}) > q_{ij}\}$, for each $\alpha^{ij} > 0$ there exists a consistent FP-test φ_n^{ij} with $\sup_{\mathbb{P}\in H_0} \sum_{n=1}^{\infty} \mathbb{P}^n(\varphi_n^{ij} \neq 0) < \alpha^{ij}$, and by Lemma 2 this implies that for H_0, H_1 there exists a strongly consistent FP-test φ_n with $\sup_{\mathbb{P}\in H_0} \sum_{n=1}^{\infty} \mathbb{P}^n(\varphi_n \neq 0) < \alpha$. The result now follows from Theorem 2.

2 $\implies$ 1 Immediate. ■

2.1.2 Proof of Theorem 3

We now show that consistent FP-testability is equivalent to the hypotheses being F_σ in W.

Proof.

1 $\implies$ 2 Let φ_n be a weakly consistent FP-test. For $i \in \{0, 1\}$ we have $\mathbb{P} \in H_i$ if and only if there is an $m \in \mathbb{N}$ such that $\mathbb{P}^n(\varphi_n = i) > 2/3$ for all $n \geq m$, hence we can write

$$\begin{aligned} H_0 &= \bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\{\mathbb{P} \in W : \mathbb{P}^n(\varphi_n = 1) \leq 1/3\} \\ H_1 &= \bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\{\mathbb{P} \in W : \mathbb{P}^n(\varphi_n = 0) \leq 1/3\}. \end{aligned}$$

Since sets of the form $\{\mathbb{P} \in W : \mathbb{P}(A) > q\}$ with $A \subseteq \mathcal{X}$ open are open in the weak topology on W, and both critical regions $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$ are open, the set $\{\mathbb{P} \in W : \mathbb{P}^n(\varphi_n = i) \leq 1/3\}$ is closed, so H_0 and H_1 are F_σ sets in the weak topology.

2 $\implies$ 1 If H_0 and H_1 are disjoint F_σ sets, then we can write $H_0 = \bigcup_{m=1}^{\infty} H_0^m$ and $H_1 = \bigcup_{m=1}^{\infty} H_1^m$ for closed sets H_0^m and H_1^m.

By Theorem 4, there exists for every $m \in \mathbb{N}$, for the pair of hypotheses $(H_0^m, W \setminus H_0^m)$ a strongly consistent FP-test $\varphi_{0,n}^m$, where $\varphi_{0,n}^m = 0$ corresponds to H_0^m and where $\varphi_{0,n}^m = 1$ corresponds to $W \setminus H_0^m$.

[3] For measurable sets $A_1, ..., A_n$ we have $\mathbb{P}(\cap_{i=1}^n A_i) \geq \sum_{i=1}^n \mathbb{P}(A_i) - n + 1$.

Similarly, for the pair of hypotheses $(W \setminus H_1^m, H_1^m)$ there exists a strongly consistent FP-test $\varphi_{1,n}^m$ where $\varphi_{1,n}^m = 0$ corresponds to $W \setminus H_1^m$ and $\varphi_{1,n}^m = 1$ corresponds to H_1^m. Define the FP-test φ_n such that

$$\begin{aligned}
\{\varphi_n = 0\} &:= \bigcup_{m=1}^{n} \{\varphi_{0,n}^m = 0\} \bigcap_{k=1}^{m} \{\varphi_{1,n}^k = 0\} \\
\{\varphi_n = 1\} &:= \bigcup_{m=1}^{n} \{\varphi_{1,n}^m = 1\} \bigcap_{k=1}^{m} \{\varphi_{0,n}^k = 1\} \\
\{\varphi_n = 2\} &:= \mathcal{X}^n \setminus (\{\varphi_n = 0\} \cup \{\varphi_n = 1\}).
\end{aligned}$$

For $\mathbb{P} \in H_i$ there is a $m \in \mathbb{N}$ such that $\mathbb{P} \in H_i^m$, and note that $\mathbb{P} \notin H_{1-i}^k$ for all $k \in \mathbb{N}$, so we obtain

$$\begin{aligned}
\mathbb{P}^\infty(\liminf_n \{\varphi_n = i\}) = \mathbb{P}^\infty(\liminf_{n \geq m} \{\varphi_n = i\}) \\
&\geq \mathbb{P}^\infty(\liminf_n \{\varphi_{i,n}^m = i\} \cap_{k=1}^m \{\varphi_{1-i,n}^k = i\}) \\
&\qquad \geq \sum_{k=1}^{m} \mathbb{P}^\infty(\liminf_n \{\varphi_{i,n}^m = i\} \cap \{\varphi_{1-i,n}^k = i\}) - m + 1 \geq 1
\end{aligned}$$

by the Fréchet inequality and the consistency of all $\varphi_{i,n}^m$ and $\varphi_{1-i,n}^k$. ■

2.1.3 Proof of Theorem 5

Having Theorem 4 at our disposal, we obtain the following characterisation of testability in terms of both hypotheses being clopen in W.

Proof.

$1 \implies 2$ By Theorem 4, if H_0 is not closed in W then there is no test with

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_0} \mathbb{P}^n(\varphi_n = 1) = 0$$

and if H_1 is not closed in W then there is no test with

$$\lim_{n\to\infty} \sup_{\mathbb{P}\in H_1} \mathbb{P}^n(\varphi_n = 0) = 0.$$

$2 \implies 1$ By Theorem 4 there exists for each $i \in \{0, 1\}$ a strongly consistent FP-tests φ_n^i such that $\lim_n \sup_{\mathbb{P}\in H_i} \mathbb{P}^n(\varphi_n^i = 1 - i) = 0$. By defining the FP-test

$$\begin{aligned}
\{\varphi_n = 0\} &:= \{\varphi_n^0 = 0\} \cap \{\varphi_n^1 = 0\} \\
\{\varphi_n = 1\} &:= \{\varphi_n^0 = 1\} \cap \{\varphi_n^1 = 1\} \\
\{\varphi_n = 2\} &:= \mathcal{X}^n \setminus (\{\varphi_n = 0\} \cup \{\varphi_n = 1\}),
\end{aligned}$$

we have for both $i \in \{0, 1\}$ that $\lim_n \sup_{\mathbb{P}\in H_i} \mathbb{P}^n(\varphi_n = 1 - i) \leq \lim_n \sup_{\mathbb{P}\in H_i} \mathbb{P}^n(\varphi_n^i = 1 - i) = 0$. ■

2.1.4 Proof of Theorem 6

Proof.

$1 \implies 2$ Let φ_n be a uniformly consistent FP-test with $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$ having disjoint closures, so for every $\varepsilon > 0$ there is a $N > 0$ such that $\inf_{\mathbb{P}\in H_i} \mathbb{P}(\varphi_n = i) \geq 1 - \varepsilon$ for all $n \geq N$, for $i = 0$

and $i = 1$. Then for any limit point $\mathbb{Q}$ of H_0 and H_1 there are sequences $\mathbb{P}_m^0$ and $\mathbb{P}_m^1$ in H_0 and H_1 respectively converging to $\mathbb{Q}$. Since $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$ have disjoint closures, we have

$$\begin{aligned}
1 &\geq \mathbb{Q}\left(\overline{\{\varphi_n = 0\}}\right) + \mathbb{Q}\left(\overline{\{\varphi_n = 0\}}\right) \\
&\geq \limsup_m P_m^0\left(\overline{\{\varphi_n = 0\}}\right) + \limsup_m P_m^1\left(\overline{\{\varphi_n = 1\}}\right) \\
&\geq \limsup_m P_m^0\left(\{\varphi_n = 0\}\right) + \limsup_m P_m^1\left(\{\varphi_n = 1\}\right) \\
&\geq \inf_{\mathbb{P}\in H_0} \mathbb{P}\left(\{\varphi_n = 0\}\right) + \inf_{\mathbb{P}\in H_1} \mathbb{P}\left(\{\varphi_n = 1\}\right) \\
&\geq 2(1-\varepsilon),
\end{aligned}$$

so we reach a contradiction.

$2 \implies 1$ Let H_0, H_1 have disjoint closures, then by relative compactness, we have $d_{BL}(H_0, H_1) > 0$. Let $\gamma < d_{BL}(H_0, H_1)/2$, let $\mathbb{P}_n^x$ denote the empirical measure at $x \in \mathcal{X}^n$, and let

$$\varphi_n(x) := \begin{cases} 0 & \text{if } d_{BL}(\mathbb{P}_n^x, H_0) < \gamma \\ 1 & \text{if } d_{BL}(\mathbb{P}_n^x, H_1) < \gamma \\ 2 & \text{otherwise,} \end{cases}$$

then by weak continuity of $x \mapsto \mathbb{P}_n^x$, the sets $\{\varphi_n = 0\}$ and $\{\varphi_n = 1\}$ are open with disjoint closures. By tightness of H_0, for every $\varepsilon > 0$ there is a compact $K \subseteq \mathcal{X}$ such that $\mathbb{P}(\mathcal{X} \setminus K) \leq \varepsilon$. This gives

$$\begin{aligned}
d_{BL}(\mathbb{P}_n^x, \mathbb{P}) &= \sup\left\{\left|\int f \mathrm{d}\mathbb{P}_n^x - \int f \mathrm{d}\mathbb{P}\right| : f \in \mathrm{BL}\right\} \\
&\leq \sup\left\{\left|\int_K f \mathrm{d}\mathbb{P}_n^x - \int_K f \mathrm{d}\mathbb{P}\right| : f \in \mathrm{BL}\right\} + \mathbb{P}_n^x(\mathcal{X} \setminus K) + \mathbb{P}(\mathcal{X} \setminus K),
\end{aligned}$$

so $\mathbb{E}[d_{BL}(\mathbb{P}_n^x, \mathbb{P})] \leq \mathbb{E}[\sup\{|\int_K f \mathrm{d}\mathbb{P}_n^x - \int_K f \mathrm{d}\mathbb{P}| : f \in \mathrm{BL}\}] + 2\varepsilon$. Since BL is equicontinuous and bounded, by Arzela-Ascoli it is precompact in the topology of uniform convergence on compacta. Restricted to K, this gives precompactness of $\mathrm{BL}(K;\mathbb{R})$ in the sup-norm, hence for every $\eta > 0$ there is a N_η and functions $f_1, ..., f_{N_\eta} \in \mathrm{BL}(K;\mathbb{R})$ such that for every $f \in \mathrm{BL}(K;\mathbb{R})$, $\|f - f_i\| \leq \eta$ for some i. The triangle inequality then gives

$$\sup\left\{\left|\int_K f \mathrm{d}\mathbb{P}_n^x - \int_K f \mathrm{d}\mathbb{P}\right| : f \in \mathrm{BL}\right\} \leq \max\left\{\left|\int_K f_i \mathrm{d}\mathbb{P}_n^x - \int_K f_i \mathrm{d}\mathbb{P}\right| : i = 1, ..., N_\eta\right\} + 2\eta.$$

By Jensen's inequality applied to the square, we obtain

$$\begin{aligned}
\mathbb{E}\left[\left|\int_K f_i \mathrm{d}\mathbb{P}_n^x - \int_K f_i \mathrm{d}\mathbb{P}\right|\right] &= \mathbb{E}\left[\left|\frac{1}{n}\sum_{j=1}^n \mathbb{1}_K(X_j) f_i(X_j) - \mathbb{E}_{\mathbb{P}}[\mathbb{1}_K f_i]\right|\right] \\
&\leq \sqrt{\mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^n \mathbb{1}_K(X_j) f_i(X_j)\right)} \leq \frac{1}{\sqrt{n}},
\end{aligned}$$

so by combing these intermediate result we obtain

$$\mathbb{E}_{\mathbb{P}}\left[d_{BL}(\mathbb{P}_n^X, \mathbb{P})\right] \leq \frac{N_\eta}{\sqrt{n}} + 2\eta + 2\varepsilon.$$

This implies uniform consistency of the test under H_0 as follows: we have $d_{BL}(\mathbb{P}_n^x, H_0) \leq d_{BL}(\mathbb{P}_n^x, \mathbb{P}) + d_{BL}(\mathbb{P}, H_0) = d_{BL}(\mathbb{P}_n^x, \mathbb{P})$, and we can pick ε, η and $N > 0$ such that $\mathbb{E}_{\mathbb{P}}\left[d_{BL}(\mathbb{P}_n^X, \mathbb{P})\right] \leq \gamma/2$ for all $n \geq N$, and so

$$\begin{aligned}
\mathbb{P}(\varphi_n \neq 0) &= \mathbb{P}(d_{BL}(\mathbb{P}_n^x, H_0) \geq \gamma) \leq \mathbb{P}(d_{BL}(\mathbb{P}_n^x, \mathbb{P}) \geq \gamma) \\
&= \mathbb{P}(d_{BL}(\mathbb{P}_n^x, \mathbb{P}) - \mathbb{E}_{\mathbb{P}}\left[d_{BL}(\mathbb{P}_n^X, \mathbb{P})\right] \geq \gamma - \mathbb{E}_{\mathbb{P}}\left[d_{BL}(\mathbb{P}_n^X, \mathbb{P})\right]) \\
&\leq \mathbb{P}(d_{BL}(\mathbb{P}_n^x, \mathbb{P}) - \mathbb{E}_{\mathbb{P}}\left[d_{BL}(\mathbb{P}_n^X, \mathbb{P})\right] \geq \gamma/2).
\end{aligned}$$

Since for $x = (x_1, ..., x_n)$ and any i and $x'_i \in \mathcal{X}$ we have with $x' = (x_1, ..., x'_i, ..., x_n)$ the bound $\left|d_{BL}(\mathbb{P}_n^x, \mathbb{P}) - d_{BL}(\mathbb{P}_n^{x'}, \mathbb{P})\right| \leq 2/n$, McDiarmids inequality (Van Der Vaart and Wellner, 2023, Proposition 2.15.3) gives

$$\sup_{\mathbb{P} \in H_0} \mathbb{P}(\varphi_n \neq 0) \leq \sup_{\mathbb{P} \in H_0} \mathbb{P}(d_{BL}(\mathbb{P}_n^x, \mathbb{P}) - \mathbb{E}_{\mathbb{P}}\left[d_{BL}(\mathbb{P}_n^X, \mathbb{P})\right] \geq \gamma/2) \leq \exp(-n^2\gamma^2/8).$$

■

3 The hardness of conditional independence testing

We now turn to applications of the preceding results to conditional independence testing. Lauritzen (2024) shows that conditional independence is closed under limits in the total variation metric. However, in general, weak convergence does not preserve conditional independence: for a weakly convergent sequence $\mathbb{P}_n(X, Y, Z) \xrightarrow{w} \mathbb{P}(X, Y, Z)$ with $X \perp\!\!\!\perp_{\mathbb{P}_n} Y \mid Z$ for all $n \in \mathbb{N}$, we might have $X \not\!\perp\!\!\!\perp_{\mathbb{P}} Y \mid Z$.

Example 1 (Lauritzen, 1996, Example 3.11). Consider a trivariate Gaussian $\mathbb{P}_n(X, Y, Z)$ with mean zero, and (conditional) covariance matrices given by

$$\Sigma^n = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{\sqrt{n}} \\ \frac{1}{2} & 1 & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{2}{n} \end{pmatrix} \quad \text{and} \quad \Sigma^n_{XY|Z} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix},$$

so we have conditional independence $X \perp\!\!\!\perp_{\mathbb{P}_n} Y \mid Z$. Since $\mathbb{P}_n(X, Y, Z) \xrightarrow{w} \mathbb{P}(X, Y, Z)$ with $\mathbb{P}(X, Y, Z)$ a degenerate multivariate Gaussian distribution with (conditional) covariance matrices

$$\Sigma = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \Sigma_{XY|Z} = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix},$$

we have $X \not\!\perp\!\!\!\perp_{\mathbb{P}} Y \mid Z$, so conditional independence is not maintained under weak limits. Other examples can be found in Barbie and Gupta (2014) and Saldi and Yüksel (2022).

Combining this example with Theorem 4, we immediately obtain that for real-valued random variables, conditional independence is not consistently FP-testable with uniform error control under the null. We strengthen this result by allowing for rather general sample spaces, and showing that consistent FP-tests don't exist for conditional independence. We use the following lemmas:

Lemma 3. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be complete separable metric spaces with $\mathcal{Z}$ perfect*[4] *then $H_0 := \{\mathbb{P} : X \perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$ is dense in $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$.*

Proof. Probability measures of the form $\mathbb{P} = \sum_{i=1}^n a_i \delta_{k_i}$ with $k_i = (x_i, y_i, z_i) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ are dense in $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ (Bogachev, 2007, Example 8.1.6). If there are $i \neq j$ such that $z_i = z_j$, then for every $\varepsilon > 0$ there is a $z'_j \in \mathcal{Z}$ such that $z'_j \neq z_i$ for all $i = 1, ..., n$ and $d(z_j, z'_j) < \varepsilon/a_j$. Defining $k'_j := (x_j, y_j, z'_j)$ and $\mathbb{P}' := \sum_{i=1}^n a_i \delta_{k_i} + a_j(\delta_{k'_j} - \delta_{k_j})$ we have that $d_{BL}(\mathbb{P}, \mathbb{P}') < \varepsilon$. Hence, convex combinations of point-masses with distinct Z-coordinates are dense in $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$. For any such $\mathbb{P}'$ and any z_i in the support of $\mathbb{P}'(Z)$ we have $\mathbb{P}'(X, Y \mid Z = z_i) = \delta_{(x_i, y_i)} = \delta_{x_i}\delta_{y_i}$, so $X \perp\!\!\!\perp_{\mathbb{P}'} Y \mid Z$. ■

Lemma 4. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be complete separable metric spaces, then $H_1 := \{\mathbb{P} : X \not\!\perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$ is dense in $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$.*

Proof. In Boeken et al. (2025), Corollary 1 it is shown that $H_1 := \{\mathbb{P} : X \not\!\perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$ is dense in $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ with respect to the total variation metric. Since $d_{BL} \leq d_{TV}$, the assertion holds. ■

Theorem 7. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be complete separable metric spaces with $\mathcal{Z}$ perfect, then the hypotheses $H_0 := \{\mathbb{P} : X \perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$ and $H_1 := \{\mathbb{P} : X \not\!\perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$ are not consistently FP-testable.*

[4] That is, $\mathcal{Z}$ has no isolated points. Since $\mathcal{Z}$ is also Polish, it is uncountable (Kechris, 1995, Corollary 6.3).

Proof. Since $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ equipped with the weak topology is complete, the Baire category theorem implies that it is not meager in itself. By Lemmas 3 and 4 we have that H_0 and H_1 are dense. If H_0 and H_1 are both F_σ, then they are both G_δ (i.e. a countable union of open sets). The complement of a dense G_δ set is meager, implying that H_0 and H_1 are meager. The space $W = H_0 \cup H_1$ would then be meager as well giving a contradiction, so H_0 and H_1 cannot both be F_σ. From Theorem 3 we conclude that H_0, H_1 are not consistently FP-testable. ■

The preceding result is for example applicable when X, Y, Z take values in (complete separable metric) function spaces such as $C([0,1], \mathbb{R}^d)$ (Manten et al., 2024), $L^p([0,1], \mathbb{R}^d)$ (Lundborg et al., 2022) or the Skorohod space $\mathbb{D}([0,1], \mathbb{R}^d)$ (Boeken and Mooij, 2024), for example when they represent measurements of continuous-time stochastic processes.

4 Weak closedness of conditional independence

Having established that conditional independence is generally not FP-testable, we will now search for conditions under which conditional independence is closed or F_σ, implying the existence of a consistent FP-test. In particular, we will show that if $\mathbb{P}_n(X, Y, Z) \xrightarrow{w} \mathbb{P}(X, Y, Z)$ and the conditional distributions $z \mapsto \mathbb{P}_n(X \mid Z = z)$ are uniformly Lipschitz then conditional independence $X \perp\!\!\!\perp Y \mid Z$ is maintained in the limit. First, we formalise what is meant with a conditional distribution being Lipschitz.

A *Markov kernel* $\mathbb{P}(X \mid Z)$ is a measurable map $\mathcal{Z} \to \mathcal{P}(\mathcal{X})$.[5] For Markov kernels $\mathbb{P}(X \mid Y), \mathbb{P}(Y \mid Z)$, their *product* is defined as the Markov kernel

$$\mathbb{P}(X \mid Y) \otimes \mathbb{P}(Y \mid Z) : \mathcal{Z} \to \mathcal{P}(\mathcal{X} \times \mathcal{Y}),\ z \mapsto \left(D \mapsto \int_D \mathrm{d}\mathbb{P}(x \mid y) \mathrm{d}\mathbb{P}(y \mid z) \right)$$

where $D \in \mathcal{B}(\mathcal{X} \times \mathcal{Y})$. If $\mathcal{X}$ is a separable complete metric space, then for a given joint distribution $\mathbb{P}(X, Z)$ there exists a Markov kernel (a version of the *conditional distribution*) $\mathbb{P}(X \mid Z)$ such that $\mathbb{P}(X, Z) = \mathbb{P}(X \mid Z) \otimes \mathbb{P}(Z)$. This Markov kernel is $\mathbb{P}(Z)$-almost everywhere uniquely defined.

Our main focus is on distributions $\mathbb{P}(X, Y, Z)$ for which there exists a version of the conditional distribution $\mathbb{P}(X \mid Z)$ such that

$$\mathbb{P}(X \mid Z) : (\mathcal{Z}, d_{\mathcal{Z}}) \to (\mathcal{P}(\mathcal{X}), d_{BL}), \quad z \mapsto \mathbb{P}(X \mid Z = z)$$

is an L-Lipschitz map, i.e.

$$d_{BL}(\mathbb{P}(X \mid Z = z), \mathbb{P}(X \mid Z = z')) \leq L \cdot d_{\mathcal{Z}}(z, z')$$

for all $z, z' \in \mathcal{Z}$, where $d_{\mathcal{Z}}$ denotes the metric on $\mathcal{Z}$ – we will always take this version of the conditional distribution. It follows from the definitions that L-Lipschitz continuity of the Markov kernel $z \mapsto \mathbb{P}(X \mid Z = z)$ is equivalent to L-Lipschitz continuity of the conditional expectation $z \mapsto \mathbb{E}[f(X) \mid Z = z]$ for all $f \in \mathrm{BL}(\mathcal{X}; \mathbb{R})$.

First, we have that the Lipschitz assumption is closed in the weak topology:

Theorem 8. *Let $\mathcal{X}, \mathcal{Z}$ be separable metric spaces with $\mathcal{Z}$ complete. The set $\{\mathbb{P} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z}) : \mathbb{P}(X \mid Z)$ is L-Lipschitz$\}$ is closed in the weak topology.*

Further, this Lipschitz assumption on the conditional $\mathbb{P}(X \mid Z)$ or $\mathbb{P}(Y \mid Z)$ implies that conditional independence is closed in the weak topology.

Theorem 9. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be separable metric spaces with $\mathcal{Z}$ complete. If $\mathbb{P}_n(X, Y, Z) \xrightarrow{w} \mathbb{P}(X, Y, Z)$ where $X \perp\!\!\!\perp_{\mathbb{P}_n} Y \mid Z$ and $\mathbb{P}_n(X \mid Z)$ is L-Lipschitz for all $n \in \mathbb{N}$, then $X \perp\!\!\!\perp_{\mathbb{P}} Y \mid Z$.*

The proofs of Theorems 8 and 9 are given in Appendix A.

[5] Here, $\mathcal{P}(\mathcal{X})$ is equipped with the Borel σ-algebra generated by the weak topology, which coincides with the smallest σ-algebra that makes for all $D \in \mathcal{B}(\mathcal{X})$ the evaluation map $\mathbb{P} \mapsto \mathbb{P}(X \in D)$ measurable (Ghosal and Van Der Vaart, 2017, Proposition A.5). This definition of a Markov kernel is equivalent to the common definition that for all $D \in \mathcal{B}(\mathcal{X})$ the map $z \mapsto \mathbb{P}(X \in D \mid Z = z)$ is measurable, and for every $z \in \mathcal{Z}$ the map $D \mapsto \mathbb{P}(X \in D \mid Z = z)$ is a probability measure.

4.1 Sufficient conditions for Lipschitz Markov kernels

In Section 5 we will use this Lipschitz assumption for conditional independence testing. To make it better accessible, we investigate sufficient conditions for this Lipschitz assumption to hold.

Scheffé's theorem implies that if a conditional density is continuous in its conditioning variable, so $p(x \mid z_n) \to p(x \mid z)$ for some sequence $z_n \to z$, then $\mathbb{P}(X \mid Z = z_n) \xrightarrow{w} \mathbb{P}(X \mid Z = z)$. The following proposition contains a Lipschitz-continuity analogue of this result. We also consider the composition of Markov kernels $\mathbb{P}(X \mid Y) \circ \mathbb{P}(Y \mid Z) : \mathcal{Z} \to \mathcal{P}(\mathcal{X})$, defined as $z \mapsto \pi_{\mathcal{X}}(\mathbb{P}(X \mid Y) \otimes \mathbb{P}(Y \mid Z = z))$ (where $\pi_{\mathcal{X}}$ denotes the projection $\pi_{\mathcal{X}}(\mathbb{P}(X, Y \mid Z = z)) = \mathbb{P}(X \mid Z = z)$), which is of particular importance in the context of graphical models like Bayesian networks. We show that the composition of Lipschitz Markov kernels behaves similar to the composition of Lipschitz functions.

Proposition 1. *Lipschitz Markov kernels have the following properties:*

a) If $\mathbb{P}(X \mid Z)$ has a density $p(x \mid z)$ with respect to a finite measure $\mathbb{Q}(X \in \mathcal{X}) \leq M < \infty$ such that $z \mapsto p(x \mid z)$ is L-Lipschitz, then $\mathbb{P}(X \mid Z)$ is LM-Lipschitz.

b) If $\mathcal{Z}$ is discrete then $\mathbb{P}(X \mid Z)$ is 2-Lipschitz.

c) If $y \mapsto \mathbb{P}(X \mid Y = y)$ and $z \mapsto \mathbb{P}(Y \mid Z = z)$ are L-Lipschitz and M-Lipschitz respectively, then $z \mapsto \mathbb{P}(X \mid Z = z) := \mathbb{P}(X \mid Y) \circ \mathbb{P}(Y \mid Z = z)$ is $\max\{1, L\} \cdot M$-Lipschitz.

d) If $z \mapsto \mathbb{P}(X \in D \mid Z = z)$ is L-Lipschitz for all $D \in \mathcal{B}(\mathcal{X})$, then $z \mapsto \mathbb{P}(X \mid Z = z)$ is L-Lipschitz as well.

e) If $\mathbb{P}(X, Z) \sim \mathcal{N}(\mu, \Sigma)$ is multivariate Gaussian, then $\mathbb{P}(X \mid Z)$ has Lipschitz constant $\|\Sigma_{XZ}\Sigma_{ZZ}^{-1}\|_{\mathrm{op}}$.[6]

Proof.

a) For any $f \in \mathrm{BL}(\mathcal{X}; \mathbb{R})$ and $z, z' \in \mathcal{Z}$ we have $|\mathbb{E}[f(X) \mid z] - \mathbb{E}[f(X) \mid z']| \leq \int_{\mathcal{X}} |p(x \mid z) - p(x \mid z')| \mathrm{d}\mathbb{Q}(x) \leq L d_{\mathcal{Z}}(z, z') \mathbb{Q}(\mathcal{X}) \leq LM d_{\mathcal{Z}}(z, z')$.

b) For any $f \in \mathrm{BL}(\mathcal{X}; \mathbb{R})$ and $z, z' \in \mathcal{Z}$ we have $\left|\int f(x) \mathrm{d}(\mathbb{P}(x \mid z) - \mathbb{P}(x \mid z'))\right| \leq 2\mathbb{1}\{z = z'\} = 2 d_{\mathcal{Z}}(z, z')$.

c) For any $f \in \mathrm{BL}(\mathcal{X}; \mathbb{R})$ we have that $y \mapsto \frac{1}{\max\{1,L\}} \mathbb{E}[f(X) | Y = y]$ is in $\mathrm{BL}(\mathcal{Y}; \mathbb{R})$, hence $\left|\int \mathbb{E}[f(X) \mid Y = y] \mathrm{d}(\mathbb{P}(y \mid z) - \mathbb{P}(y \mid z'))\right| \leq \max\{1, L\} d_{BL}(\mathbb{P}(Y \mid z), \mathbb{P}(Y \mid z')) \leq \max\{1, L\} M d_{\mathcal{Z}}(z, z')$.

d) This follows directly from the definition of the total variation distance, which upper-bounds the bounded Lipschitz metric.

e) Given $\mathbb{P}(X, Z) \sim \mathcal{N}(\mu, \Sigma)$, the conditional $\mathbb{P}(X \mid Z = z)$ is Gaussian with mean $\mu_{X \mid Z}(z) := \mu_X + \Sigma_{XZ}\Sigma_{ZZ}^{-1}(z - \mu_Z)$ and covariance matrix $\Sigma_{X|Z} := \Sigma_{XX} - \Sigma_{XZ}\Sigma_{ZZ}^{-1}\Sigma_{ZX}$. Let $U \sim \mathcal{N}(0, \Sigma_{X|Z})$. For any $f \in \mathrm{BL}(\mathcal{X}; \mathbb{R})$ we have

$$
\begin{aligned}
|\mathbb{E}[f(X) \mid Z = z] - \mathbb{E}[f(X) \mid Z = z']| &= |\mathbb{E}[f(\mu_{X \mid Z}(z) + U)] - \mathbb{E}[f(\mu_{X \mid Z}(z') + U)]| \\
&\leq \mathbb{E}[|f(\mu_{X \mid Z}(z) + U) - f(\mu_{X \mid Z}(z') + U)|] \\
&\leq \|\mu_{X \mid Z}(z) - \mu_{X \mid Z}(z')\| \\
&\leq \|\Sigma_{XZ}\Sigma_{ZZ}^{-1}\|_{\mathrm{op}} \|z - z'\|.
\end{aligned}
$$

■

One readily verifies that the conditional distributions $\mathbb{P}_n(X \mid Z)$ and $\mathbb{P}_n(Y \mid Z)$ from Example 1 are $\sqrt{n}/2$-Lipschitz. In particular, no L-Lipschitz assumption holds for these sequences, so the fact that conditional independence is not maintained in the limit hints at the sharpness of this condition in Theorem 9.

[6]Here, $\|\cdot\|_{\mathrm{op}}$ denotes the operator norm.

4.2 Weak closedness of conditional independence via total variation

We can also take a completely different approach to finding sufficient conditions under which conditional independence is weakly closed, which does not use uniform continuity of the Markov kernels. Namely, by Lauritzen (2024), conditional independence is closed in total variation. For sets of probability measures for which the total variation topology and weak topology coincide, we then immediately obtain that conditional independence is weakly closed. To this end, we consider classes of measures with well-behaved densities. Let μ be a σ-finite Borel measure on standard Borel spaces $\mathcal{X}_V$ and let W_μ be any class of distributions with a density with respect to μ which are uniformly equicontinuous and uniformly bounded by some envelope function, that is, for every $\varepsilon > 0$ there is a $\delta > 0$ such that $|x - y| < \delta$ implies $|p(x) - p(y)| < \varepsilon$ for all densities p, and there is a $M : \mathcal{X} \to \mathbb{R}$ such that $p(x) \leq M(x)$ for all p. By generalising Boos (1985) (Lemma 1) to arbitrary metric spaces we have that the total variation topology and weak topology coincide on W_μ.

Theorem 10. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be metric spaces. If $\mathbb{P}_n(X,Y,Z) \xrightarrow{w} \mathbb{P}(X,Y,Z)$ with $\mathbb{P}_n \in W_\mu$ for all $n \in \mathbb{N}$, then $\mathbb{P}_n(X,Y,Z) \xrightarrow{tv} \mathbb{P}(X,Y,Z)$.*

The proof is found in Appendix A. By Lauritzen (2024), this implies the following result.

Theorem 11. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be metric spaces. If $\mathbb{P}_n(X,Y,Z) \xrightarrow{w} \mathbb{P}(X,Y,Z)$ with $X \perp\!\!\!\perp_{\mathbb{P}_n} Y \mid Z$ and $\mathbb{P}_n \in W_\mu$ for all $n \in \mathbb{N}$, then $X \perp\!\!\!\perp_{\mathbb{P}} Y \mid Z$.*

Another approach would be to consider sufficiently regular exponential family, for which by Barndorff-Nielsen (2014), Section 8.1, Theorem 8.3, the parameter space with the Euclidean topology is homeomorphic to the set of probability measures with the weak topology. By Scheffé's theorem, this topology then coincides with the total variation topology.

Note that weak convergence plus the Lipschitz assumption does not imply total variation convergence (for example, let $\mathbb{P}_n(X) = \delta_{1/n}$ and $\mathbb{P}_n(Y|X = x) = \delta_x$), so Theorem 9 cannot be proven by invoking Lauritzen (2024).

4.3 Related literature

As sufficient condition for conditional independence to be closed in the weak topology, Jordan (1977) and Hellwig (1996) considered probability measures $\mathbb{P}$ with $z \mapsto \mathbb{P}(X \mid Z = z)$ continuous, but Barbie and Gupta (2014) gave a counterexample. They show that conditional independence is closed in the *topology of information* (see also Backhoff-Veraguas et al. (2020)), and show that this topology coincides with the weak topology under similar uniform equicontinuity conditions as Theorem 9.

5 Sufficient conditions for conditional independence testing

Having the topological sufficient conditions for testability from Section 2 and the sufficient conditions for conditional independence to be weakly closed from Section 4 at our disposal, we can relatively easily prove the existence of conditional independence tests under various settings.

To this end, let $W_L \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ be the set of measures such that $\mathbb{P}(X \mid Z)$ is L-Lipschitz and let $W_\infty := \cup_{L \in \mathbb{N}} W_L$, i.e. the set of measures such that $\mathbb{P}(X \mid Z)$ is Lipschitz. Also, recall the definition of W_μ from Section 4.2.

Theorem 12. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be complete separable metric spaces, let $W \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ and consider the hypotheses $H_0 := \{\mathbb{P} \in W : X \perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$ and $H_1 := \{\mathbb{P} \in W : X \not\perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$.*

a) If $W = W_L$ then there exists a strongly consistent FP-test with uniform error control under H_0.

b) If $\mathcal{Z}$ is discrete and $W = \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$, then there exists a strongly consistent FP-test with uniform error control under H_0.

c) If $W = W_\mu$ then there exists a strongly consistent FP-test with uniform error control under H_0.

d) If $W = W_L$ and we consider the alternative hypothesis

$$H_1^\varepsilon := \{\mathbb{P} \in W_L : d_{BL}(\mathbb{P}(X \mid Z) \otimes \mathbb{P}(Y, Z), \mathbb{P}(X, Y, Z)) \geq \varepsilon\}$$

for given $\varepsilon > 0$, then there exists a uniformly consistent FP-test as specified in Theorem 6.

e) If $W = W_\infty$, then there exists a strongly consistent FP-test.

Proof.

a) By Theorem 9 we have that H_0 is closed in $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$, hence also in W_L. The result follows from Theorem 4.

b) If $\mathcal{Z}$ is discrete, then by Proposition 1.b) we have that $W_2 = \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$, so the result follows from part a).

c) By Theorem 11, H_0 is closed in $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$, hence also in W_μ, so the result follows from Theorem 4.

d) Another way of phrasing Theorem 9 is that the map $\mathbb{P}(X, Y, Z) \mapsto \mathbb{P}(X \mid Z) \otimes \mathbb{P}(Y, Z)$ from W_L to $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ is continuous. Hence, the map

$$f : W_L \mapsto \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}), \quad \mathbb{P}(X, Y, Z) \mapsto d_{BL}(\mathbb{P}(X \mid Z) \otimes \mathbb{P}(Y, Z), \mathbb{P}(X, Y, Z))$$

is continuous as well, so $H_0 = f^{-1}(\{0\})$ and $H_1^\varepsilon = f^{-1}([\varepsilon, 1])$ are closed in W_L. The result follows from Theorem 5.

e) The set W_L is closed in $\mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ hence also in W_∞, and writing $H_0' = \{\mathbb{P} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}) : X \perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$ we have that $H_0' \cap W_L$ is closed in W_L hence also in W_∞, which gives that $H_0 = \bigcup_{L \in \mathbb{N}} H_0' \cap W_L$ is F_σ in W_∞. Similarly, writing $H_1' = \{\mathbb{P} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}) : X \not\!\perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$ we have that $H_1' \cap W_L$ is open in W_L, hence F_σ in W_L, hence also F_σ in W_∞, hence $H_1 = \bigcup_{L \in \mathbb{N}} H_1' \cap W_L$ is F_σ in W_∞, so the result follows from Theorem 3. ■

Note that by symmetry, in the definition of W_L as considered in Theorem 12, one can also let all $\mathbb{P}(Y \mid Z)$ be L-Lipschitz.

Remark 1. Note that the hypotheses $H_0 := W_L$ and $H_1 := \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}) \setminus W_L$ are also FP-testable, since W_L is closed by Theorem 8. Hence, these regularity assumptions for conditional independence testing *are themselves testable*. An assumption-free consistent FP-test with uniform error control under H_0 therefore exists for the hypotheses $H_0 := \{\mathbb{P} \in W_L : X \perp\!\!\!\perp_{\mathbb{P}} Y \mid Z\}$ and $H_1 := \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}) \setminus H_0$.

5.1 Related literature

Warren (2021) proposes a conditional independence test based on binning the space $\mathcal{Z}$, and proves that it is pointwise asymptotically valid under the null and consistent under the alternative, under the assumption that the distributions $\mathbb{P}(X, Y \mid Z), \mathbb{P}(X \mid Z)$ and $\mathbb{P}(Y \mid Z)$ are L-Lipschitz maps from a compact space $\mathcal{Z}$ to the space of probability measures equipped with the p-Wasserstein distance W_p. This assumption is close to the assumption that we make, since we have $d_{BL}(\mathbb{P}_0, \mathbb{P}_1) \leq W_p(\mathbb{P}_0, \mathbb{P}_1)$, with equality $d_{BL}(\mathbb{P}_0, \mathbb{P}_1) = W_1(\mathbb{P}_0, \mathbb{P}_1)$ if $\mathbb{P}_0$ and $\mathbb{P}_1$ have bounded support (Bogachev, 2007, Theorem 8.10.45). Theorem 12.a) shows that, under the assumptions of Warren (2021), there even exists a test which is valid at every sample size $n \in \mathbb{N}$, and one requires only $\mathbb{P}(X \mid Z)$ or $\mathbb{P}(Y \mid Z)$ to be L-Lipschitz.

Neykov et al. (2021) propose a conditional independence test which obtains a minimax-optimal rate, under the assumption that X, Y, Z are compactly supported, have densities, and $\mathbb{P}(X \mid Z)$ and $\mathbb{P}(Y \mid Z)$ are L-Lipschitz maps from $\mathcal{Z}$ to the space of probability measures equipped with the total variation metric.

Györfi and Walk (2012) propose a conditional independence test and aim at proving its strong consistency without any regularity conditions on H_0 and H_1, but Neykov et al. (2021) point out a mistake in their proof. Dai and Song (2025) propose a consistent test, under the assumption of a smooth parametric model. For data beyond real-valued random variables, there is some work about conditional independence testing where X, Y, Z take values in function spaces, representing measurements of continuous time stochastic processes. Lundborg et al. (2022) propose an asymptotically valid test under the assumption that $\mathbb{E}[X \mid Z]$ and $\mathbb{E}[Y \mid Z]$ can be estimated sufficiently well, and Manten et al. (2024) propose a weakly consistent test under the assumption that $\mathbb{E}[X \mid Z], \mathbb{E}[Y \mid Z]$ and $\mathbb{E}[(X, Y) \mid Z]$ can be estimated with certain kernel methods.

None of these tests take any regularity of the critical region into account. For example, the minimax rate of Neykov et al. (2021) is computed over all tests with *measurable* critical regions. Hence, tests are considered which can never be implemented when one has finite-precision measurements.

6 Discussion

This work establishes a precise topological framework for understanding the testability of statistical hypotheses under the constraint of finite-precision measurements. Our results reveal that the feasibility of constructing consistent tests with controlled error — particularly in nonparametric settings — can be characterized entirely in terms of the topological properties (F_σ, closed, clopen, or disjoint closures) of the null and alternative hyptheses in the weak topology on the space of probability measures. An important implication is the non-testability of conditional independence hypotheses in general: because both the null and alternative hypotheses are dense and their union is a Polish space, no consistent FP-test exists without additional assumptions. This generalizes and strengthens prior results in the literature, and underscores that conditional independence — while foundational in causal inference and graphical models — is not a testable property without smoothness or structural constraints. By imposing a Lipschitz continuity assumption on the conditional distributions, we recover testability. We show that Lipschitz conditions ensure closedness in the weak topology, enabling the construction of tests with uniform error control under H_0. This yields sufficient conditions under which conditional independence becomes statistically testable. In particular, we prove that this regularity assumption of having Lipschitz conditional distributions is itself testable, allowing for a recursive framework for assessing the testability of higher-order statistical claims.

Despite these contributions, several important questions remain open. While the established topological conditions are necessary and sufficient, they are not always constructive: given an open H_1, one must find open $A_{ij} \subseteq \mathcal{X}$ and $q_{ij} \in [0, 1]$ such that $H_1 = \bigcup_{i=1}^{\infty} \bigcap_{j=1}^{m_i} \{\mathbb{P} : \mathbb{P}(A_{ij}) > q_{ij}\}$ in order to construct a test, and this representation may be difficult to identify in practice. It would be interesting to find such an explicit representation for conditional independence. An interesting avenue for future research would be to allow for dependent data sampling: by Theorem 2, uniform error control under H_0 can be interpreted as $\sup_{\mathbb{P} \in H_0} \mathbb{P}^\infty(\exists n : \varphi_n \neq 0) \leq \alpha$, which is closely related to anytime-valid p-values and e-values (Grünwald et al., 2024), and topological characterisations of the existence of p/e-values would be of significant interest. It is plausible that analogous topological criteria govern the possibility of valid online or interactive hypothesis testing in complex models. Also, Theorem 6 curretly employs a compactness assumptions which can be loosened. It remains an open question whether there exists a consistent FP-test for conditional independence if one merely assumes that the distribution has a density. Finally, from Example 1 it seems that some kind of equicontinuity of $\mathbb{P}(X \mid Z)$ or $\mathbb{P}(Y \mid Z)$ is necessary for conditional independence testing with uniform error control under H_0. Finding these necessary conditions would be of great interest, for example for constraint-based causal discovery.

In summary, our results offer a unified topological perspective on the fundamental limits of statistical inference under measurement constraints. They resolve open problems around the testability of nonparametric hypotheses, while opening avenues for further exploration in both theoretical and applied statistics.

A Proofs of results in Section 4

Lemma 5. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be separable metric spaces. Let $\mathbb{P}_0(Y, Z)$ and $\mathbb{P}_1(Y, Z)$ be given, and let $\mathbb{P}(X \mid Z)$ be an L-Lipschitz Markov kernel, then*

$$d_{BL}(\mathbb{P}(X \mid Z) \otimes \mathbb{P}_0(Y, Z), \mathbb{P}(X \mid Z) \otimes \mathbb{P}_1(Y, Z)) \leq \max\{1, L\} d_{BL}(\mathbb{P}_0(Y, Z), \mathbb{P}_1(Y, Z)).$$

Proof. Considering the product metric $d_{\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}} = d_{\mathcal{X}} + d_{\mathcal{Y}} + d_{\mathcal{Z}}$, for any $f(x,y,z) \in \mathrm{BL}(\mathcal{X}\times\mathcal{Y}\times\mathcal{Z};\mathbb{R})$ we have that $g_f(y,z) := \frac{1}{\max\{1,L\}}\int_{\mathcal{X}} f(x,y,z)\mathrm{d}\mathbb{P}(x\,|\,z)$ is in $\mathrm{BL}(\mathcal{Y}\times\mathcal{Z};\mathbb{R})$, so we get

$$\begin{aligned}\left|\int f(x,y,z)\mathrm{d}\mathbb{P}(x\,|\,z)\mathrm{d}\mathbb{P}_0(y,z) - \int f(x,y,z)\mathrm{d}\mathbb{P}(x\,|\,z)\mathrm{d}\mathbb{P}_1(y,z)\right| \\ = \max\{1,L\}\left|\int g_f(y,z)\mathrm{d}\mathbb{P}_0(y,z) - \int g_f(y,z)\mathrm{d}\mathbb{P}_1(y,z)\right| \\ \le \max\{1,L\} d_{BL}(\mathbb{P}_0(Y,Z),\mathbb{P}_1(Y,Z)).\end{aligned}$$

■

Lemma 6. *Let $\mathcal{X},\mathcal{Y},\mathcal{Z}$ be separable metric spaces, with $\mathcal{Z}$ complete. If $\mathbb{P}_n(X\,|\,Z) \to \mathbb{P}(X\,|\,Z)$ uniformly on compacta, then for any given $\mathbb{P}(Y,Z)$ we have*

$$d_{BL}(\mathbb{P}_n(X\,|\,Z)\otimes\mathbb{P}(Y,Z), \mathbb{P}(X\,|\,Z)\otimes\mathbb{P}(Y,Z)) \to 0.$$

Proof. For any $f(x,y,z) \in \mathrm{BL}(\mathcal{X}\times\mathcal{Y}\times\mathcal{Z};\mathbb{R})$ we have that $x \mapsto f(x,y,z) \in \mathrm{BL}(\mathcal{X};\mathbb{R})$, so

$$\left|\int f(x,y,z)\mathrm{d}\mathbb{P}_n(x\,|\,z) - \int f(x,y,z)\mathrm{d}\mathbb{P}(x\,|\,z)\right| \le d_{BL}(\mathbb{P}_n(X\,|\,z),\mathbb{P}(X\,|\,z)).$$

Since $\mathcal{Z}$ is separable and complete, the measure $\mathbb{P}(Z)$ is tight, so for every $\varepsilon > 0$ there exists a compact $K_\varepsilon \subseteq \mathcal{Z}$ such that $\mathbb{P}(Z \notin K_\varepsilon) \le \varepsilon$. Since $d_{BL} \le 2$, we have

$$\begin{aligned}\left|\int f(x,y,z)\mathrm{d}\mathbb{P}_n(x\,|\,z)\mathrm{d}\mathbb{P}(y,z) - \int f(x,y,z)\mathrm{d}\mathbb{P}(x\,|\,z)\mathrm{d}\mathbb{P}(y,z)\right| \\ \le \int d_{BL}(\mathbb{P}_n(X\,|\,z),\mathbb{P}(X\,|\,z))\mathrm{d}\mathbb{P}(z) \\ \le \sup_{z\in K_\varepsilon} d_{BL}(\mathbb{P}_n(X\,|\,z),\mathbb{P}(X\,|\,z)) + 2\varepsilon.\end{aligned}$$

Since $\mathbb{P}_n(X\,|\,Z) \to \mathbb{P}(X\,|\,Z)$ uniformly on compacta we get the result. ■

Lemma 7. *Let $\mathcal{X},\mathcal{Z}$ be separable metric spaces with $\mathcal{Z}$ complete. If $\mathbb{P}_n(X,Z) \xrightarrow{w} \mathbb{P}(X,Z)$ and $\mathbb{P}_n(X\,|\,Z)$ is L-Lipschitz for all $n \in \mathbb{N}$, there is a subsequence n_k such that $\mathbb{P}_{n_k}(X\,|\,Z) \to \mathbb{P}(X\,|\,Z)$ uniformly on compacta.*

Proof. By Ascoli's theorem (Munkres, 2014, Theorem 47.1) there exists a subsequence n_k and a continuous Markov kernel $\mathbb{Q}(X\,|\,Z)$ such that $\mathbb{P}_{n_k}(X\,|\,Z) \to \mathbb{Q}(X\,|\,Z)$ uniformly on compacta. Then we have

$$\begin{aligned}d_{BL}(\mathbb{P}(X\,|\,Z)\otimes\mathbb{P}(Z), \mathbb{Q}(X\,|\,Z)\otimes\mathbb{P}(Z)) &\le d_{BL}(\mathbb{P}(X\,|\,Z)\otimes\mathbb{P}(Z), \mathbb{P}_{n_k}(X\,|\,Z)\otimes\mathbb{P}_{n_k}(Z)) \\ &\quad + d_{BL}(\mathbb{P}_{n_k}(X\,|\,Z)\otimes\mathbb{P}_{n_k}(Z), \mathbb{P}_{n_k}(X\,|\,Z)\otimes\mathbb{P}(Z)) \\ &\quad + d_{BL}(\mathbb{P}_{n_k}(X\,|\,Z)\otimes\mathbb{P}(Z), \mathbb{Q}(X\,|\,Z)\otimes\mathbb{P}(Z)).\end{aligned}$$

The first term is equal to $d_{BL}(\mathbb{P}(X,Z),\mathbb{P}_{n_k}(X,Z))$ and by Lemma 5 the second term is bounded by $\max\{1,L\}d_{BL}(\mathbb{P}(Z),\mathbb{P}_{n_k}(Z))$, which both go to zero. The last term converges to zero by Lemma 6, hence we obtain $\mathbb{P}(X\,|\,Z) = \mathbb{Q}(X\,|\,Z)$, $\mathbb{P}(Z)$-almost surely. ■

Theorem 8. *Let $\mathcal{X},\mathcal{Z}$ be separable metric spaces with $\mathcal{Z}$ complete. The set $\{\mathbb{P} \in \mathcal{P}(\mathcal{X}\times\mathcal{Z}) : \mathbb{P}(X\,|\,Z)$ is L-Lipschitz$\}$ is closed in the weak topology.*

Proof. Note that if $\mathbb{P}_n(X\,|\,Z)$ is L-Lipschitz for all $n \in \mathbb{N}$ and $\mathbb{P}_{n_k}(X\,|\,Z=z) \xrightarrow{w} \mathbb{P}(X\,|\,Z=z)$ for all $z \in \mathcal{Z}$, then $\mathbb{P}(X\,|\,Z)$ is L-Lipschitz as well. By Lemma 7 we get that $\{\mathbb{P} \in \mathcal{P}(\mathcal{X}\times\mathcal{Z}) : \mathbb{P}(X\,|\,Z)$ is L-Lipschitz$\}$ is closed in the weak topology. ■

Theorem 9. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be separable metric spaces with $\mathcal{Z}$ complete. If $\mathbb{P}_n(X,Y,Z) \xrightarrow{w} \mathbb{P}(X,Y,Z)$ where $X \perp\!\!\!\perp_{\mathbb{P}_n} Y \mid Z$ and $\mathbb{P}_n(X \mid Z)$ is L-Lipschitz for all $n \in \mathbb{N}$, then $X \perp\!\!\!\perp_{\mathbb{P}} Y \mid Z$.*

Proof. By Lemma 7, there is a subsequence n_k such that $\mathbb{P}_{n_k}(X \mid Z) \to \mathbb{P}(X \mid Z)$ uniformly on compacts. When writing

$$\begin{aligned} d_{BL}(\mathbb{P}_{n_k}(X \mid Z) \otimes \mathbb{P}_{n_k}(Y,Z),\ & \mathbb{P}(X \mid Z) \otimes \mathbb{P}(Y,Z)) \\ & \leq d_{BL}(\mathbb{P}_{n_k}(X \mid Z) \otimes \mathbb{P}_{n_k}(Y,Z), \mathbb{P}_{n_k}(X \mid Z) \otimes \mathbb{P}(Y,Z)) \\ & + d_{BL}(\mathbb{P}_{n_k}(X \mid Z) \otimes \mathbb{P}(Y,Z), \mathbb{P}(X \mid Z) \otimes \mathbb{P}(Y,Z)) \end{aligned}$$

the first term is bounded by $\max\{1, L\} d_{BL}(\mathbb{P}_{n_k}(Y,Z), \mathbb{P}(Y,Z))$ by Lemma 5 and hence converges to zero, and the second term converges to zero by Lemma 6, which gives $\mathbb{P}(X,Y,Z) = \lim_k \mathbb{P}_{n_k}(X,Y,Z) = \lim_k \mathbb{P}_{n_k}(X \mid Z) \otimes \mathbb{P}_{n_k}(Y,Z) = \mathbb{P}(X \mid Z) \otimes \mathbb{P}(Y,Z)$, which is the desired result. ■

Theorem 10. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be metric spaces. If $\mathbb{P}_n(X,Y,Z) \xrightarrow{w} \mathbb{P}(X,Y,Z)$ with $\mathbb{P}_n \in W_\mu$ for all $n \in \mathbb{N}$, then $\mathbb{P}_n(X,Y,Z) \xrightarrow{tv} \mathbb{P}(X,Y,Z)$.*

Proof. Let $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ weakly with $\mathbb{P}_n \in W_\mu$ for all $n \in \mathbb{N}$. By Ascoli's theorem (Munkres, 2014, Theorem 47.1), the class of uniformly bounded and uniformly equicontinuous densities is relatively compact in the topology of uniform convergence on compacta. In particular, for any subsequence n' there is a further subsequence n'' and a $p^* : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \to [0, \infty)$ such that $p_{n''} \to p^*$ uniformly on compacta. This implies that p^* integrates to 1, has modulus of continuity ω and is uniformly bounded by M, and hence $\mathbb{P}^* \in W_\mu$. By Scheffe (1947) we then have weak convergence $\mathbb{P}_{n''} \xrightarrow{w} \mathbb{P}^*$. The weak convergence $\mathbb{P}_n \xrightarrow{w} \mathbb{P}$ also implies convergence of the subsequence $\mathbb{P}_{n''} \to \mathbb{P}$, and thus $p = p^*$ μ-a.e. The μ-a.e. convergence $p_{n''} \to p$ implies convergence $p_n \to p$ as well (otherwise there exists a subsequence n' with $|p_{n'} - p| > \varepsilon$ on some set with positive μ-measure, which contradicts the existence of a convergent further subsequence), which implies total variation convergence $\mathbb{P}_n \xrightarrow{tv} \mathbb{P}$, again by Scheffé's Theorem. ■

Acknowledgements

Philip Boeken was supported by Booking.com. Eduardo Skapinakis was funded by national funds through the FCT – Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UIDB/00297/2020 (`https://doi.org/10.54499/UIDB/00297/2020`) and UIDP/00297/2020 (`https://doi.org/10.54499/UIDP/00297/2020`) (Center for Mathematics and Applications) and the FCT scholarship, reference 2022.10596.BD (`https://doi.org/10.54499/2022.10596.BD`). The authors thank Bas Kleijn and Sourbh Bhadane for helpful discussions.

References

Backhoff-Veraguas, J., Bartl, D., Beiglböck, M., and Eder, M. (2020). All adapted topologies are equal. *Probability Theory and Related Fields*, 178(3):1125–1172.

Barbie, M. and Gupta, A. (2014). The topology of information on the space of probability measures over polish spaces. *Journal of Mathematical Economics*, 52:98–111.

Barndorff-Nielsen, O. E. (2014). *Information and Exponential Families: In Statistical Theory*. John Wiley & Sons, Chichester [U.K.] New York.

Berger, A. (1951). On Uniformly Consistent Tests. *The Annals of Mathematical Statistics*, 22(2):289–293.

Berger, A. and Wald, A. (1949). On Distinct Hypotheses. *The Annals of Mathematical Statistics*, 20(1):104–109.

Billingsley, P. (1999). *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. Probability and Statistics Section. Wiley, New York, 2nd ed edition.

Boeken, P., Forré, P., and Mooij, J. M. (2025). Are Bayesian networks typically faithful?

Boeken, P. and Mooij, J. M. (2024). Dynamic Structural Causal Models.

Bogachev, V. I. (2007). *Measure theory*, volume 2. Springer.

Boos, D. D. (1985). A Converse to Scheffe's Theorem. *The Annals of Statistics*, 13(1).

Cover, T. M. (1973). On Determining the Irrationality of the Mean of a Random Variable. *The Annals of Statistics*, 1(5):862–871.

Dai, S. and Song, X. (2025). Consistent tests for semiparametric conditional independence. *Statistics & Probability Letters*, 216:110253.

Dembo, A. and Peres, Y. (1994). A Topological Criterion for Hypothesis Testing. *The Annals of Statistics*, 22(1).

Ermakov, M. (2017). On Consistent Hypothesis Testing. *Journal of Mathematical Sciences*, 225(5):751–769.

Gänssler, P. (1971). Compactness and sequential compactness in spaces of measures. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 17(2):124–146.

Genin, K. (2018). *The Topology of Statistical Inquiry.* PhD thesis, Carnegie Mellon University, Pittsburgh, PA, USA.

Genin, K. (2021). Statistical undecidability in linear, non-gaussian causal models in the presence of latent confounders. *Advances in neural information processing systems*, 34:13564–13574.

Genin, K. and Kelly, K. T. (2017). The Topology of Statistical Verifiability. *Electronic Proceedings in Theoretical Computer Science*, 251:236–250.

Genin, K. and Mayo-Wilson, C. (2020). Statistical decidability in linear, non-gaussian causal models. In *Proceedings of the Causal Discovery & Causality-Inspired Machine Learning Workshop, NeurIPS, Virtual*, pages 11–12.

Genin, K. and Mayo-Wilson, C. (2024). Success concepts for causal discovery: The topology of success in LiNGAM models with and without hidden variables. *Behaviormetrika*, 51(1):515–538.

Ghosal, S. and Van Der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference.* Cambridge University Press, 1 edition.

Grünwald, P., de Heide, R., and Koolen, W. (2024). Safe testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(5):1091–1128.

Györfi, L. and Walk, H. (2012). Strongly consistent nonparametric tests of conditional independence. *Statistics & Probability Letters*, 82(6):1145–1150.

Hellwig, M. F. (1996). Sequential decisions under uncertainty and the maximum theorem. *Journal of Mathematical Economics*, 25(4):443–464.

Jordan, J. S. (1977). The Continuity of Optimal Dynamic Decision Rules. *Econometrica*, 45(6):1365–1376.

Kechris, A. (1995). *Classical descriptive set theory.* Springer.

Kelly, K. T. (1996). *The Logic of Reliable Inquiry.* Logic and Computation in Philosophy. Oxford University Press, Oxford.

Kleijn, B. J. K. (2022). *The Frequentist Theory of Bayesian Statistics.* Springer. (in preparation to be published).

Larsson, M., Ruf, J., and Ramdas, A. (2026). A complete characterization of testable hypotheses.

Lauritzen, S. (2024). Total variation convergence preserves conditional independence. *Statistics & Probability Letters*, 214:110200.

Lauritzen, S. L. (1996). *Graphical models*, volume 17. Clarendon Press.

LeCam, L. and Schwartz, L. (1960). A Necessary and Sufficient Condition for the Existence of Consistent Estimates. *The Annals of Mathematical Statistics*, 31(1):140–150.

Lundborg, A. R., Shah, R. D., and Peters, J. (2022). Conditional Independence Testing in Hilbert Spaces with Applications to Functional Data Analysis. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(5):1821–1850.

Manten, G., Casolo, C., Ferrucci, E., Mogensen, S. W., Salvi, C., and Kilbertus, N. (2024). Signature Kernel Conditional Independence Tests in Causal Discovery for Stochastic Processes.

Munkres, J. R. (2014). *Topology*. Pearson, Harlow, 2. ed., pearson new internat. ed edition.

Neykov, M., Balakrishnan, S., and Wasserman, L. (2021). Minimax optimal conditional independence testing. *The Annals of Statistics*, 49(4):2151–2177.

Nobel, A. B. (2006). Hypothesis Testing for Families of Ergodic Processes. *Bernoulli*, 12(2):251–269.

Pfanzagl, J. (1968). On the existence of consistent estimates and tests. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 10(1):43–62.

Resende, P. (2021). An Abstract Theory of Physical Measurements. 51(6):108.

Saldi, N. and Yüksel, S. (2022). Geometry of information structures, strategic measures and associated stochastic control topologies. *Probability Surveys*, 19(none):450–532.

Scheffe, H. (1947). A Useful Convergence Theorem for Probability Distributions. *The Annals of Mathematical Statistics*, 18(3):434–438.

Shah, R. D. and Peters, J. (2020). The hardness of conditional independence testing and the generalised covariance measure. *The Annals of Statistics*, 48(3):1514–1538.

Spirtes, P., Glymour, C., Scheines, N., et al. (1993). Causation, prediction, and search.

Van Der Vaart, A. W. and Wellner, J. A. (2023). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Series in Statistics. Springer International Publishing, Cham.

Vickers, S. (1990). *Topology via Logic*. Number 5 in Cambridge Tracts in Theoretical Computer Science. Cambridge university press, Cambridge.

Warren, A. (2021). Wasserstein Conditional Independence Testing.