HAL open science

Interfacing biology, category theory and mathematical statistics

Erwan Beurier, Dominique Pastor, Andrée Ehresmann

▶ To cite this version:

Erwan Beurier, Dominique Pastor, Andrée Ehresmann. Interfacing biology, category theory and mathematical statistics. [Research Report] IMTA-RR-2019-04-SC, IMT-Atlantique. 2019. hal-02280488v1

HAL Id: hal-02280488

https://imt-atlantique.hal.science/hal-02280488v1

Submitted on 6 Sep 2019 (v1), last revised 19 Dec 2019 (v2)

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

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

IMT Atlantique
Dépt. Signal & Communications
Technopôle de Brest-Iroise - CS 83818
29238 Brest Cedex 3
Téléphone: +33 (0)2 29 00 13 04
Télécopie: +33 (0)2 29 00 10 12
URL: **www.imt-atlantique.fr**

Collection des rapports de recherche d'IMT Atlantique
IMTA-RR-2019-04-SC

Interfacing biology, category theory and mathematical statistics

Erwan Beurier
IMT Atlantique

Dominique Pastor
IMT Atlantique

Andrée Ehresmann
University of Picardie

Date d'édition : 10 July 2019
Version : 1.0

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Contents

1. Introduction **2**

2. Multiplicity Principle **4**
2.1. General case 4
2.2. Application to preorders 5

3. Statistical detection of a signal in noise **5**
3.1. Problem statement 5
3.2. Decision with level $\gamma \in (0,1)$ and oracles 5
3.3. Observations 6

4. Selectivity, landscapes of tests and preordering **6**

5. The Neyman-Pearson (NP) solution **8**

6. The RDT solution **8**
6.1. An elementary RDT problem 8
6.2. Application to Detection 10

7. Multiplicity Principle **11**

8. Conclusions and Perspectives **12**

Abstract: Motivated by the concept of degeneracy in biology [1], we establish a first connection between the Multiplicity Principle [2, 3] and mathematical statistics. Specifically, we exhibit two families of tests that satisfy this principle to achieve the detection of a signal in noise.

The contents of this paper were presented during ACT 2019 (Applied Category Theory 2019), held in Oxford (England) between the 15 and 19 July (see http://www.cs.ox.ac.uk/ACT2019/*).*

1. Introduction

In [1], Edelman & Gally pointed out degeneracy as the fundamental property allowing for living systems to evolve through natural selection towards more complexity in fluctuating environments. Degeneracy is defined [1] as "... *the ability of elements that are structurally different to perform the same function or yield the same output*". Degeneracy is a crucial feature of immune systems and neural networks, at all organization levels.

The Multiplicity Principle (MP) [2, 3], introduced by Ehresmann & Vanbremeersch, is a mathematical formalization of degeneracy in Categorical terms. The consequences of this principle, as treated in [2, 3], underpin Edelman & Gally's conjecture according to which "*complexity and degeneracy go hand in hand*" [1].

Another property of many biological and social systems is their resilience: (i) they can perform in degraded mode, with some performance loss, but without collapsing; (ii) they can recover their initial performance level when nominal conditions are satisfied again; (iii) they can perform corrections and auto-adaption so as to maintain essential tasks for their survival. In addition, resilience of social or biological systems is achieved via agents with different skills. For instance, cells are simply reactive organisms, whereas social agents have some cognitive properties. Thence the idea that resilience may derive from fundamental properties satisfied by agents, interactions and organizations. Could this fundamental property be a possible consequence of degeneracy [3, Section 3.1, p. 15]?

The notion of resilience remains, however, somewhat elusive, mathematically speaking. In contrast, the notion of robustness has a long history and track record in mathematical statistics [4]. By and large, a statistical method is robust if its performance is not unduly altered in case of outliers or fluctuations around the model for which it is designed. Can we fathom the links between resilience and robustness?

As an attempt to embrace the questions raised above from a comprehensive outlook, the original question addressed in this work-in-progress is the possible connection between MP and robustness to account for emergence of resilience in complex systems. As a first step in our study aimed at casting the notions of robustness, resilience and degeneracy within the same theoretical framework based on MP, we hereafter establish that statistical tests do satisfy MP. The task to perform by the tests is the fundamental problem of detecting a signal in noise. However, to ease the reading of a paper at the interface between category theory and mathematical statistics, we consider a simplified version of this problem.

The paper is organized as follows. We begin by specifying notation and notions in mathematical statistics. In Section 2, we state MP in categorical words on the basis of [2] and consider the particular case of preorders, which will be sufficient at the present time to establish that statistical tests satisfy MP for detecting signals in noise. In Section 3, we set out the statistical detection problem. We will then introduce, in Section 4, a preorder that makes it possible to exhibit two types of "structurally different" tests, namely, the Neyman-Pearson tests (Section 5) and the RDT tests (Section 6). Section 7 concludes the paper by establishing that these two types of tests achieve the MP for the detection problem under consideration. For space considerations, we limit proofs to the minimum making it possible to follow the approach without too much undue effort.

Summary of main results

Because this paper lies at the interface between different mathematical specialties, the present section summarizes its contents in straight text. To begin with, the MP is a property of a category that structurally different diagrams that share the same cocones. To state our main results, it will not be necessary to consider the general MP though. In fact, the particular case of preordered sets will suffice, in which case MP reduces

to Proposition 1.

Second, in statistical hypothesis testing, a hypothesis can be seen as a predicate, the truth value of which we aim at determining using statistical decisions, some of which being potentially better than others. There exist many optimality criteria to devise a decision to test a given hypothesis. In non-Bayesian approaches, which will be our focus below, such criteria are encoded through the notions of size and power.

The size is the least upper bound for the probability of rejecting the hypothesis when this one is actually true. We generally want this size to remain below a certain value called level, because the hypothesis to test mostly represents the standard situation. For instance, planes in the sky are rare events, after all, and the standard hypothesis is "there is no plane", which represents the standard situation. A too large level may result in an almost intractable cluttering of a radar screen.

We do not want to be bothered by too many alarms. In contrast, when the hypothesis is false, we want to reject it with highest confidence. The probability of a test rejecting the hypothesis when this one is actually false is called the powerof the test. For a given testing problem, we thus look for decisions with maximal power within the set of decisions whose size does not exceed the specified level. This defines a preorder. A maximal element in this preorder is said to be optimal.

Different hypotheses to test may thus require different criteria, specified through different notions of size and two different notions of power. This is what we exploit below to exhibit two sets of "structurally different" decisions that satisfy the MP.

To carry out this construction, we consider the detection of a signal in independent standard gaussian noise, a classical problem in many applications. This is an hypothesis testing problem for which there exists an optimality criterion where the size is the so-called probability of false alarm and the power is the so-called probability of detection. This criterion has a solution, the Neyman-Pearson (NP) decision, which is thus the maximal element of a certain preorder. We can consider a second class of tests, namely, the RDT tests. These tests are aimed at detecting deviations of a signal with respect to a known deterministic model in presence of independent standard gaussian noise. This problem is rotationally invariant and the RDT tests are optimal with respect to a specific criterion defined through suitable notions of size and power. They are maximal elements of another preordered set. Although not dedicated to signal detection, these tests can be used as surrogates to NP tests to perform signal detection. It turns out that the family of RDT tests and that of NP tests satisfy the MP as stated in Theorem 4. This is because the more data we have, the closer to perfection both tests are.

Notation

Random variables. Given two measurable spaces $\mathcal{E}$ and $\mathcal{F}$, $\mathcal{M}(\mathcal{E},\mathcal{F})$ denotes the set of all measurable functions defined on $\mathcal{E}$ and valued in $\mathcal{F}$. The two σ-algebra involved are omitted in the notation because, in the sequel, they will always be obvious from the context. In particular, we will throughout consider a probability space $(\Omega,\mathcal{B},\mathbb{P})$ and systematically endow $\mathbb{R}$ with the Borel σ-algebra, which will not be recalled. Therefore, $\mathcal{M}(\Omega,\mathbb{R})$ designates the set of all real random variables. Given $q \in [0,\infty[$, $\mathcal{B}_\infty(q)$ is the set of all real random variables $\Delta \in \mathcal{M}(\Omega,\mathbb{R})$ such that $|\Delta|_\infty \leqslant q$. As usual, we write $X \sim \mathcal{N}(0,1)$ to mean that $X \in \mathcal{M}(\Omega,\mathbb{R})$ is standard normal. The cumulative distributive function (cdf) of any $X \sim \mathcal{N}(0,1)$ is denoted by Φ. Given a sequence $(X_n)_{n\in\mathbb{N}} \in \mathcal{M}(\Omega,\mathbb{R})^{\mathbb{N}}$ of real random variables, we write $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$ to mean that the random variables $X_1, X_2, \ldots$ are independent and identically distributed with common distribution $\mathcal{N}(0,1)$. For any $n \in \mathbb{N}$, the Generalized Marcum function $Q_{n/2}$ is defined for all $\rho \geqslant 0$ and all $\lambda \geqslant 0$ by setting: $Q_{n/2}(\rho,\lambda) := 1 - \mathbb{F}_{\chi^2_n(\rho^2)}(\lambda^2)$, where $\mathbb{F}_{\chi^2_n(\rho^2)}$ is the cdf of the χ^2 law with n degrees of freedom and non-centrality parameter ρ^2.

Decisions et Observations. Throughout, $\mathcal{M}\left(\{0,1\}\times\Omega,\{0,1\}\right)$ designates the set of all measurable functions $D : \{0,1\}\times\Omega \to \{0,1\}$. Any element of $\mathcal{M}\left(\{0,1\}\times\Omega,\{0,1\}\right)$ is called a decision for obvious reasons given below. If $D \in \mathcal{M}\left(\{0,1\}\times\Omega,\{0,1\}\right)$ then, for any $\varepsilon \in \{0,1\}$, $D(\varepsilon)$ denotes the Bernoulli-distributed random variable $D(\varepsilon) : \Omega \to \{0,1\}$ defined for any given $\omega \in \Omega$ by $D(\varepsilon)(\omega) = D(\varepsilon,\omega)$. An n-dimensional test is hereafter any measurable function $f : \mathbb{R}^n \to \{0,1\}$ and $\mathcal{M}(\mathbb{R}^n,\{0,1\})$ stands for the set of all n-dimensional tests. A measurable function $\boldsymbol{X} : \{0,1\}\times\Omega \to \mathbb{R}^n$ is hereafter called

an observation and $\mathcal{M}(\{0,1\}\times\Omega,\mathbb{R}^n)$ denotes the set of all these observations. Given a test $f \in \mathcal{M}(\mathbb{R}^n,\{0,1\})$ and $\boldsymbol{X} \in \mathcal{M}(\{0,1\}\times\Omega,\mathbb{R}^n)$, $D = f(\boldsymbol{X})$ is trivially a decision: $D \in \mathcal{M}\left(\{0,1\}\times\Omega,\{0,1\}\right)$. If $\boldsymbol{X} \in \mathcal{M}(\{0,1\}\times\Omega,\mathbb{R}^n)$ then, for any $\varepsilon \in \{0,1\}$, $\boldsymbol{X}(\varepsilon) = \boldsymbol{X}(\varepsilon,\cdot) \in \boldsymbol{X}(\varepsilon) \in \mathcal{M}(\Omega,\mathbb{R}^n)$ is defined for every $\omega \in \Omega$ by $\boldsymbol{X}(\varepsilon)(\omega) = \boldsymbol{X}(\varepsilon,\omega)$.

Empirical means. We define the empirical mean of a given sequence $y = (y_n)_{n\in\mathbb{N}}$ of real values as the sequence $(\langle y\rangle_n)_{n\in\mathbb{N}}$ of real values such that, $\forall n \in \mathbb{N}, \langle y\rangle_n := \frac{1}{n}\sum_{i=1}^n y_i$. By extension, the empirical mean of a sequence $Y = (Y_n)_{n\in\mathbb{N}}$ of random variables where each $Y_n \in \mathcal{M}(\Omega,\mathbb{R})$ is the sequence $(\langle Y\rangle_n)_{n\in\mathbb{N}}$ of random variables where, for each $n \in \mathbb{N}$, $\langle Y\rangle_n \in \mathcal{M}(\Omega,\mathbb{R})$ is defined by $\langle Y\rangle_n := \frac{1}{n}\sum_{i=1}^n Y_i$. Therefore, for any $\omega \in \Omega$, $\langle Y\rangle_n(\omega) := \langle Y(\omega)\rangle_n$ with $Y(\omega) = (Y_n(\omega))_{n\in\mathbb{N}}$. If $Y = (Y_n)_{n\in\mathbb{N}}$ is a sequence of observations ($\forall n \in \mathbb{N}$, $Y_n \in \mathcal{M}(\{0,1\}\times\Omega,\mathbb{R})$), we define the empirical mean of Y as the sequence $(\langle Y\rangle_n)_{n\in\mathbb{N}}$ of observations such that, for $\varepsilon \in \{0,1\}$, $\langle Y\rangle_n \in \mathcal{M}(\{0,1\}\times\Omega,\mathbb{R})$ with$\langle Y\rangle_n(\varepsilon) = \langle Y(\varepsilon)\rangle_n$ and $Y(\varepsilon) = (Y_n(\varepsilon))_{n\in\mathbb{N}}$.

2. Multiplicity Principle

2.1. General case

The multiplicity principle (MP) comes from [2]. It proposes a categorical approach to the biological degeneracy principle which ensures a kind of flexible redundancy. Roughly, MP, in a category $\mathscr{C}$, ensures the existence of *structurally non isomorphic* diagrams which have the same colimit. A formal definition relies on the notion of a cluster between diagrams in a category $\mathscr{C}$.

Definition 1 (Cluster)**.** *Let $D : \mathscr{D} \to \mathscr{C}$ and $E : \mathscr{E} \to \mathscr{C}$ be two (small) diagrams. A cluster $G : D \to E$ is a set $G = \{f : D(d) \to E(e) \mid d \in \mathscr{D}, e \in \mathscr{E}, f \in \mathscr{C}\}$ such that:*

(i) for all $d \in \mathscr{D}$ there exist $e \in \mathscr{E}$ and $g : D(d) \to E(e)$ such that $g \in G$
(ii) if $g : D(d) \to E(e)$ and $g' : D(d) \to E(e')$ are in the cluster, then e and $e' \in \mathscr{E}$ are in the same connected component of $\mathscr{E}$
(iii) if $g : D(d) \to E(e) \in G$ and $\varepsilon : e \to e' \in \mathscr{E}$, then $E(\varepsilon) \circ g \in G$
(iv) if $\delta : d' \to d \in \mathscr{D}$ and $g : D(d) \to E(e)$, then $g \circ D(\delta) \in G$

For instance, a cone from c to D can be seen as a cluster from the constant functor $\Delta(c)$ to D; and a cocone from E to c is a cluster $E \to \Delta(c)$.

Remark 1. *Adjacent clusters can be composed: a cluster $G : D \to E_0$ and a cluster $G_0 : E_0 \to E$ can be composed to a cluster $G \circ G_0$. We can then consider a category of clusters of $\mathscr{C}$, whose objects are the (small) diagrams $\mathscr{D} \to \mathscr{C}$, and an arrow $D \to E$ is a cluster. This category is isomorphic to the free cocompletion of $\mathscr{C}$ [2].*

Definition 2 (Multiplicity principle (MP))**.** *A category $\mathscr{C}$* satisfies the multiplicity principle (MP) *if there exist two diagrams $D : \mathscr{D} \to \mathscr{C}$ and $E : \mathscr{E} \to \mathscr{C}$ such that:*

(i) **Cocones** $(D) \cong$ **Cocones** (E)*;*
(ii) There is no cluster $D \to E$ nor $E \to D$ that is an isomorphism in the free cocompletion of $\mathscr{C}$ (cf. Remark 1).

D and E having the same cocones translates the property of both systems to accomplish the same function. The absence of isomorphic clusters between D and E reflects the structural difference between them, which is key to robustness and adaptability: if the system described by E fails, then D may replace it.

2.2. Application to preorders

The main purpose of this paper is to find a meaningful instance of the multiplicity principle in some preorder. In the following, we do not distinguish between a preorder and its associated category.

Proposition 1 (Multiplicity principle in a preorder)**.** *Let $(E, \leqslant)$ be a preorder. If there are two disjoint subsets $A, B \subset E$ such that the following conditions hold, then E verifies the multiplicity principle:*

(i) A and B have the same sets of upper bounds

(ii) There is an $a \in A$ with no upper bounds in B

(iii) There is a $b \in B$ with no upper bounds in A

Proof. Condition (i) ensures that A and B have isomorphic categories of cocones. Conditions (ii) and (iii) respectively ensure that there is no cluster $i_A \to i_B$ nor $i_B \to i_A$ where $i_A : A \hookrightarrow E$ and $i_B : B \hookrightarrow E$ are the inclusion functors. □

Albeit trivial, the following lemma will be helpful.

Lemma 1. *Given a preordered set $(E, \preceq)$, if A and B are two subsets of E such that* $\sup(A, (E, \preceq)) = \sup(B, (E, \preceq))$ *and if $A \times B \cap \preceq = \emptyset$, then E satisfies MP.*

3. Statistical detection of a signal in noise

3.1. Problem statement

Let $\varepsilon \in \{0, 1\}$ be the unknown indicator value on whether a certain physical phenomenon has occurred ($\varepsilon = 1$) or not ($\varepsilon = 0$). We aim at determining this value. It is desirable to resort to something more evolved than tossing a coin to estimate ε. However, whatever D, the decision is erroneous for any $\omega \in \Omega$ such that $D(\varepsilon, \omega) \neq \varepsilon$. We thus have two distinct cases.

False alarm probability: If $\varepsilon = 0$ and $D(0, \omega) = 1$, we commit an error of the 1st kind false or false alarm, since we have erroneously decided that the phenomenon has occurred while nothing actually happened. We thus define the false alarm probability (aka size, aka error probability of the 1st kind) of D as:

$$\mathbb{P}_{\text{FA}}[D] \stackrel{\text{def}}{=} \mathbb{P}\big[D(0) = 1\big] \tag{1}$$

Detection probability: If $\varepsilon = 1$ and $D(1, \omega) = 0$, we commit an error of the 2nd kind, also called missed detection since, in this case, we have missed the occurrence of the phenomenon. As often in the literature on the topic, we prefer using the probability of correctly detecting the phenomenon by defining the detection probability:

$$\mathbb{P}_{\text{DET}}[D] \stackrel{\text{def}}{=} \mathbb{P}\big[D(1) = 1\big] \tag{2}$$

3.2. Decision with level $\gamma \in (0, 1)$ and oracles

Among all the possible decisions, the omniscient oracle $D^* \in \mathcal{M}\left(\{0, 1\} \times \Omega, \{0, 1\}\right)$ is defined for any pair $(\varepsilon, \omega) \in \{0, 1\} \times \Omega$ by setting $D^*(\varepsilon, \omega) = \varepsilon$. Its probability of false alarm is 0 and its probability of detection is 1: $\mathbb{P}_{\text{FA}}[D^*] = 0$ et $\mathbb{P}_{\text{DET}}[D^*] = 1$. This omniscient oracle has no practical interest since it knows ε. That's not really fair! Since it is not possible in practice to guarantee a null false alarm probability, we focus on decisions whose false alarm probabilities are upper-bounded by a real number $\gamma \in (0, 1)$ called level. We state the following definition.

Definition 3 (Level)**.** *We say that $D \in \mathcal{M}\left(\{0, 1\} \times \Omega, \{0, 1\}\right)$ has level $\gamma \in]0, 1[$ if $\mathbb{P}_{FA}[D] \leqslant \gamma$. The set of all decisions with level $\gamma \in]0, 1[$ is denoted by C_γ.*

We can easily prove the existence of an infinite number of elements in C_γ that all have a detection probability equal to 1. Whence the following definition.

Definition 4. *Given $\gamma \in (0,1)$, an oracle with level γ is any decision $D \in \mathcal{C}_\gamma$ such that $\mathbb{P}_{DET}[D] = 1$. The set of all the oracles with level γ is denoted by $\mathfrak{D}^*_\gamma$.*

Oracles with level γ have no practical interest either since they require prior knowledge of ε! Therefore, we restrict our attention to decisions in $\mathcal{C}_\gamma$ that "approximate" at best the oracles with level γ, without prior knowledge of ε, of course. To this end, we must preorder decisions.

Lemma-Definition 1 (Total preorder $(\mathcal{C}_\gamma, \preceq)$)**.** *For any given $\gamma \in (0,1)$ and any pair $(D, D') \in \mathcal{C}_\gamma \times \mathcal{C}_\gamma$, we define a preorder $(\mathcal{C}_\gamma, \preceq)$ by setting:*

$$D \preceq D' \quad \text{if} \quad \mathbb{P}_{DET}[D] \leqslant \mathbb{P}_{DET}[D']\,. \tag{3}$$

We write $D \cong D'$ if $D \preceq D'$ and $D' \preceq D$.

3.3. Observations

In practice, observations help us decide whether the phenomenon has occurred or not. By collecting a certain number of them, we can expect to make a decision. Hereafter, observations are assumed to be elements of $\mathcal{M}(\{0,1\} \times \Omega, \mathbb{R})$ and corrupted versions of ε. We suppose to have a sequence $(Y_n)_{n\in\mathbb{N}}$ of such random variables. As a first standard model, we could assume that, for any $n \in \mathbb{N}$ and any $(\varepsilon, \omega) \in \{0,1\} \times \Omega$, $Y_n(\varepsilon, \omega) = \varepsilon + X_n(\omega)$ with $X_1, X_2, \ldots, X_n, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$. In this additive model, X_n models noise on the nth observation. We could make this model more complicated and realistic by considering random vectors instead of variables. However, with respect to our purpose, the significant improvement we can bring to the model is elsewhere. Indeed, we have assumed above that the signal, regardless of noise, is ε. However, from a practical point of view, it is more realistic to assume that the nth observation Y_n captures ε in presence of some interference Δ_n, independent of X_n. In practice, the probability distribution of Δ_n will hardly be known and, as a means to compensate for this lack of knowledge, we assume the existence of a uniform bound on the amplitude of all possible interferences. Therefore, we assume that, for all $(\varepsilon, \omega) \in \{0,1\} \times \Omega$, $Y_n(\varepsilon, \omega) = \varepsilon + X_n(\omega) + \Delta_n(\omega)$ and the existence of $q \in [0, \infty)$ such that $\Delta_n \in \mathcal{B}_\infty(q)$. After all, this model is standard in time series analysis: ε plays the role of a trend, Δ_n is the seasonal variation and X_n is the measurement noise.

For each $q \in [0, \infty)$, $\mathrm{Seq}^{[q]}$ is hereafter the set of all the sequences $Y^{[q]} = \left(Y_n^{[q]}\right)_{n\in\mathbb{N}}$ where, $\forall n \in \mathbb{N}$ and $\forall(\varepsilon, \omega) \in \{0,1\} \times \Omega$, $Y_n^{[q]}(\varepsilon, \omega) = \varepsilon + \Delta_n(\omega) + X_n(\omega)$, where $\Delta_n \in \mathcal{B}_\infty(q)$ and $X_n \sim \mathcal{N}(0,1)$ are independant. Therefore, for all $n \in \mathbb{N}$ and all $\varepsilon \in \{0,1\}$, $Y_n^{[q]}(\varepsilon) = \varepsilon + \Delta_n + X_n$, with $X_1, X_2, \ldots, X_n, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$.

4. Selectivity, landscapes of tests and preordering

For any $q \in [0, 1/2[$, any $n \in \mathbb{N}$ and any $Y^{[q]} = \left(Y_n^{[q]}\right)_{n\in\mathbb{N}} \in \mathrm{Seq}^{[q]}$, we set:

$$\boldsymbol{Y}_n^{[q]} = \left(Y_1^{[q]}, Y_2^{[q]}, \ldots, Y_n^{[q]}\right) \tag{4}$$

Definition 5 (Selectivity of a test)**.** *Given any $n \in \mathbb{N}$ and any test $f \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$, the selectivity of f at given level $\gamma \in (0,1)$ is defined as the set:*

$$\mathrm{Sel}_\gamma(f) \;:=\; \left\{ q \in [0, 1/2) : \forall Y^{[q]} \in \mathrm{Seq}^{[q]}, f\left(\boldsymbol{Y}_n^{[q]}\right) \in \mathcal{C}_\gamma \right\}$$

Definition 6 (Landscape of a test)**.** *Given any $n \in \mathbb{N}$ and any test $f \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$, the landscape of f at given level $\gamma \in (0,1)$ is the subset of $\mathcal{C}_\gamma$ defined by:*

$$\mathfrak{D}_\gamma(f) \;:=\; \bigcup_{q \in \mathrm{Sel}_\gamma(f)} \left\{ f\left(\boldsymbol{Y}_n^{[q]}\right) : Y^{[q]} \in \mathrm{Seq}^{[q]} \right\} \tag{5}$$

This notion of landscape makes it possible to compare tests via landscapes via the following preorder. The proofs that the following definition is consistent and that the next lemma holds true are left to the reader.

Definition 7 (Preorder $(2^{C_\gamma}, \preceq^*)$)**.** *Given any level $\gamma \in (0,1)$, we define the preorder $(2^{C_\gamma}, \preceq^*)$ via the three following properties:*

(P1) $\forall n \in \mathbb{N}$, $\forall (f,g) \in \mathcal{M}(\mathbb{R}^n, \{0,1\}) \times \mathcal{M}(\mathbb{R}^n, \{0,1\})$, $\mathfrak{D}_\gamma(f) \preceq^* \mathfrak{D}_\gamma(g)$ *if:*

$$\mathrm{Sel}_\gamma(f) = \mathrm{Sel}_\gamma(g) \quad \textit{and} \quad \forall q \in \mathrm{Sel}_\gamma(f), \forall Y^{[q]} \in \mathrm{Seq}^{[q]}, f\left(\boldsymbol{Y}_n^{[q]}\right) \preceq g\left(\boldsymbol{Y}_n^{[q]}\right)$$

(P2) $\forall (\mathfrak{D}, \mathfrak{D}') \in \left(\mathbb{D}_\gamma \cup \mathbb{D}_\gamma^*\right) \times \mathbb{D}_\gamma^*$, $\mathfrak{D} \preceq^* \mathfrak{D}'$ *with:*

$$\mathbb{D}_\gamma := \bigcup_{n \in \mathbb{N}} \left\{ \mathfrak{D}_\gamma(f) : f \in \mathcal{M}(\mathbb{R}^n, \{0,1\}) \right\} \quad \textit{and} \quad \mathbb{D}_\gamma^* := 2^{\mathfrak{D}_\gamma^*}$$

(P3) $\forall \mathfrak{D} \in 2^{C_\gamma} \setminus \left(\mathbb{D}_\gamma \cup \mathbb{D}_\gamma^*\right)$, $\mathfrak{D} \preceq^* \mathfrak{D}$

Lemma 2. $\forall (\mathfrak{D}, \mathfrak{D}') \in \left(\mathbb{D}_\gamma \cup \mathbb{D}_\gamma^*\right) \times \left(\mathbb{D}_\gamma \cup \mathbb{D}_\gamma^*\right)$, $\mathfrak{D} \preceq^* \mathfrak{D}' \Rightarrow \mathfrak{D} \times \mathfrak{D}' \subset \preceq$.

Henceforth, given a preordered set $(E, \preceq)$ and $A \subset E$, the set of maximal elements of A is denoted by $\max(A, (E, \preceq))$, the set of upper bounds is denoted by $\mathrm{upper}(A, (E, \preceq))$ and the set of least upper bounds of A in $(E, \preceq)$ is denoted by $\sup(A, (E, \preceq))$.

Theorem 1 (Approximation of oracles in $(2^{C_\gamma}, \preceq^*)$)**.** *Given $\gamma \in (0,1)$, let a set $\mathfrak{X}_\gamma$ and a family of tests $\left(f_{\xi,n}\right)_{\xi \in \mathfrak{X}_\gamma, n \in \mathbb{N}}$ such that:*

(i) $\forall (\xi, n) \in \mathfrak{X}_\gamma \times \mathbb{N}$, $f_{\xi,n} \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$*;*

(ii) $\exists \mathcal{Q}_\gamma \subset [0, \infty)$, $\forall (\xi, n) \in \mathfrak{X}_\gamma \times \mathbb{N}$, $\mathrm{Sel}_\gamma\left(f_{\xi,n}\right) = \mathcal{Q}_\gamma$*;*

(iii) $\forall (\xi, q) \in \mathfrak{X}_\gamma \times \mathcal{Q}_\gamma$, $\forall Y^{[q]} \in \mathrm{Seq}^{[q]}$, $\lim_{n \to \infty} \mathbb{P}_{DET}\left[f_{\xi,n}\left(\boldsymbol{Y}_n^{[q]}\right)\right] = 1$*;*

then, by setting $\mathbb{D}_\gamma' = \left\{ \mathfrak{D}_\gamma\left(f_{\xi,n}\right) : n \in \mathbb{N}, \xi \in \mathfrak{X}_\gamma \right\}$*, we have:*

$$\mathbb{D}_\gamma^* = \mathrm{upper}\left(\mathbb{D}_\gamma', \left(2^{C_\gamma}, \preceq^*\right)\right) = \sup\left(\mathbb{D}_\gamma', \left(2^{C_\gamma}, \preceq^*\right)\right) \tag{6}$$

Proof. For any $(\xi, n) \in \mathfrak{X}_\gamma \times \mathbb{N}$ and any $\mathfrak{D} \in \mathbb{D}_\gamma^*$, (P2) in Definition 7 straightforwardly implies that $\mathfrak{D}_\gamma\left(f_{\xi,n}\right) \preceq^* \mathfrak{D}$. As a consequence:

$$\mathbb{D}_\gamma^* \subset \mathrm{upper}\left(\mathbb{D}_\gamma', \left(2^{C_\gamma}, \preceq^*\right)\right) \tag{7}$$

To prove the converse inclusion, consider some $\mathfrak{D} \in \mathrm{upper}\left(\mathbb{D}_\gamma', \left(2^{C_\gamma}, \preceq^*\right)\right)$. We thus have $\forall (\xi, n) \in \mathbb{N} \times \mathfrak{X}_\gamma$, $\mathfrak{D}_\gamma\left(f_{\xi,n}\right) \preceq^* \mathfrak{D}$. According to Lemma 2, we have $\forall (\xi, n) \in \mathfrak{X}_\gamma \times \mathbb{N}$, $\mathfrak{D}_\gamma\left(f_{\xi,n}\right) \times \mathfrak{D} \subset \preceq$. Therefore, $\forall (\xi, n) \in \mathbb{N} \times \mathfrak{X}_\gamma$, $\forall q \in \mathrm{Sel}_\gamma\left(f_{\xi,n}\right)$, $\forall Y^{[q]} \in \mathrm{Seq}^{[q]}$ and $\forall D \in \mathfrak{D}$, $f_{\xi,n}\left(\boldsymbol{Y}_n^{[q]}\right) \preceq \mathfrak{D}$. It follows from the definition of $\preceq$ and assumption (ii) above that:

$$\forall (\xi, n) \in \mathbb{N} \times \mathfrak{X}_\gamma, \forall q \in \mathcal{Q}_\gamma, \forall Y^{[q]} \in \mathrm{Seq}^{[q]}, \forall D \in \mathfrak{D}, \mathbb{P}_{\mathrm{DET}}\left[f_{\xi,n}\left(\boldsymbol{Y}_n^{[q]}\right)\right] \leqslant \mathbb{P}_{\mathrm{DET}}[D]$$

We derive from assumption (iii) that $\mathbb{P}_{\mathrm{DET}}[D] = 1$ and thus that $D \in \mathfrak{D}_\gamma^*$. It follows that $\mathfrak{D} \in \mathbb{D}_\gamma^*$. We thus obtain that $\mathrm{upper}\left(\mathbb{D}_\gamma', \left(2^{C_\gamma}, \preceq^*\right)\right) \subset \mathbb{D}_\gamma^*$ and therefore, from (7), $\mathbb{D}_\gamma^* = \mathrm{upper}\left(\mathbb{D}_\gamma', \left(2^{C_\gamma}, \preceq^*\right)\right)$. The second inequality in (6) is straightforward since the elements of $\mathbb{D}_\gamma^*$ are isomorphic in the sense of $\preceq^*$. □

For later use, given $\mathcal{J} \subset [0, \infty)$, $n \in \mathbb{N}$ and $\mathfrak{F} \subset \mathcal{M}(\mathbb{R}^n, \{0,1\})$, we hereafter set:

$$\mathbb{D}_\gamma^{\mathcal{J}}(\mathfrak{F}) := \left\{ \mathfrak{D}_\gamma(f) \in \mathbb{D}_\gamma : f \in \mathfrak{F}, \mathrm{Sel}_\gamma(f) = \mathcal{J} \right\} \tag{8}$$

5. The Neyman-Pearson (NP) solution

When n spans $\mathbb{N}$, the Neyman-Pearson (NP) Lemma makes it possible to pinpoint a maximal element in each $(\mathbb{D}_\gamma^{\{0\}}(\mathfrak{F}),\ \preceq\ )$ with $\mathfrak{F} = \mathcal{M}(\mathbb{R}^n, \{0,1\})$. These maximal elements are hereafter called NP decisions. Specifically, we have the following result.

Lemma 3 (Maximality of the NP decisions)**.** *For any $\gamma \in (0,1)$ and any $n \in \mathbb{N}$,*

$$\mathfrak{D}_\gamma\left(f_n^{NP(\gamma)}\right) = \max\left(\mathbb{D}_\gamma^{\{0\}}\left(\mathcal{M}\left(\mathbb{R}^n, \{0,1\}\right)\right),\ \preceq^*\ \right) \tag{9}$$

where $f_n^{NP(\gamma)} \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$ is the n-dimensional NP test with size γ defined by:

$$\forall (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n, f_n^{NP(\gamma)}(y_1, y_2, \ldots, y_n) = \begin{cases} 1 & \text{if} \quad \sum_{i=1}^n y_i > \sqrt{n}\,\Phi^{-1}(1-\gamma) \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

and satisfies, $\forall Y^{[0]} \in \mathrm{Seq}^{[0]}$,

$$\begin{cases} \mathbb{P}_{FA}\left[f_n^{NP(\gamma)}\left(\boldsymbol{Y}_n^{[0]}\right)\right] = \gamma \\ \mathbb{P}_{DET}\left[f_n^{NP(\gamma)}\left(\boldsymbol{Y}_n^{[0]}\right)\right] = 1 - \Phi\left(\Phi^{-1}(1-\gamma) - \sqrt{n}\right) \end{cases}$$

Proof. A direct application of the Neyman-Pearson Lemma [5, Theorem 3.2.1, page 60], followed by some standard algebra to obtain $\mathbb{P}_{\mathrm{DET}}\left[f_n^{\mathrm{NP}(\gamma)}\boldsymbol{Y}_n^{[0]}\right]$. □

The next result states that it suffices to increase the number of observations to approximate oracles with level γ by NP decisions.

Theorem 2 (Approximation of oracles with level γ by NP decisions in $(2^{C_\gamma},\ \preceq^*\ )$)**.**

Setting $\mathbb{D}^{\mathrm{NP}(\gamma)} := \left\{\mathfrak{D}_\gamma\left(f_n^{NP(\gamma)}\right) : n \in \mathbb{N}\right\}$ *for any $\gamma \in (0,1)$, we have:*

$$\mathbb{D}_\gamma^* = \mathrm{upper}\left(\mathbb{D}^{\mathrm{NP}(\gamma)},\ \left(2^{C_\gamma},\ \preceq^*\ \right)\right) = \sup\left(\mathbb{D}^{\mathrm{NP}(\gamma)},\ \left(2^{C_\gamma},\ \preceq^*\ \right)\right)$$

Proof. According to Lemma 3, $\lim_{n\to\infty} \mathbb{P}_{\mathrm{DET}}\left[f_n^{\mathrm{NP}(\gamma)}\boldsymbol{Y}_n^{[0]}\right] = 1$. The set $\mathbb{D}^{\mathrm{NP}(\gamma)} \subset \mathbb{D}_\gamma$ satisfies the conditions of Theorem 1 for any $\gamma \in (0,1)$ and $\mathfrak{X}_\gamma = \emptyset$. Thence the result. □

6. The RDT solution

6.1. An elementary RDT problem

Problem statement. The RDT theoretical framework is exposed in full details in [6, 7]. To ease the reading of the present paper, we directly focus on the particular RDT problem that can be used in connection with the detection problem at stake.

In this respect, suppose that $\boldsymbol{Y} = \boldsymbol{\Theta} + \boldsymbol{X} \in \mathcal{M}(\Omega, \mathbb{R}^n)$, where $\boldsymbol{\Theta}$ and $\boldsymbol{X}$ are independent elements of $\mathcal{M}(\Omega, \mathbb{R}^n)$. In the sequel, we assume that $\boldsymbol{X} \sim \mathcal{N}(0, \mathbf{I}_n)$, $\mathbf{I}_n$ being the $n \times n$ identity matrix, and consider the mean testing problem of deciding on whether $|\langle\boldsymbol{\Theta}\rangle_n(\omega)| \leqslant \tau$ (null hypothesis $\mathcal{H}_0$) or $|\langle\boldsymbol{\Theta}\rangle_n(\omega)| > \tau$ (alternative hypothesis $\mathcal{H}_1$), when we are given $\boldsymbol{Y}(\omega) = \boldsymbol{\Theta}(\omega) + \boldsymbol{X}(\omega)$, for $\omega \in \Omega$. The idea is that $\boldsymbol{\Theta}$ oscillates uncontrollably around 0 and that only sufficient large deviations of the norm should be detected. This is a particular Block-RDT problem, following the terminology and definition given in [7]. This problem is summarized by discarding ω and writing:

$$\begin{cases} \textbf{Observation: } \boldsymbol{Y} = \boldsymbol{\Theta} + \boldsymbol{X} \in \mathcal{M}(\Omega, \mathbb{R}^n) \text{ with } \begin{cases} \boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^n), \boldsymbol{X} \sim \mathcal{N}(0, \mathbf{I}_N), \\ \boldsymbol{\Theta} \text{ and } \boldsymbol{X} \text{ are independent,} \end{cases} \\ \mathcal{H}_0 : \ |\langle\boldsymbol{\Theta}\rangle_n| \leqslant \tau, \\ \mathcal{H}_1 : \ |\langle\boldsymbol{\Theta}\rangle_n| > \tau. \end{cases} \tag{11}$$

Standard likelihood theory [5, 8, 9] does not make it possible to solve this problem. Fortunately, this problem can be solved as follows via the Random Distortion Testing (RDT) framework.

Size and power of tests for mean testing. We seek tests with guaranteed size and optimal power, in the sense specified below.

Definition 8 (Size for the mean testing problem)**.** *The size of $f \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$ for testing the empirical mean of the signals $\boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^n)$ such that $\mathbb{P}\big[|\langle \boldsymbol{\Theta} \rangle_n| \leqslant \tau\big] \neq 0$, given the processes $\boldsymbol{\Theta} + \boldsymbol{X} \in \mathcal{M}(\Omega, \mathbb{R}^n)$ with $\boldsymbol{X}$ independent of $\boldsymbol{\Theta}$, is defined by:*

$$\alpha^{[n]}(f) = \sup_{\boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^n) : \mathbb{P}[\,|\langle \boldsymbol{\Theta} \rangle_n| \leqslant \tau\,] \neq 0} \mathbb{P}\big[f(\boldsymbol{\Theta} + \boldsymbol{X}) = 1 \,\big|\, |\langle \boldsymbol{\Theta} \rangle_n| \leqslant \tau\big] \tag{12}$$

We say that $f \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$ has level (resp. size) γ if $\alpha^{[n]}(f) \leqslant \gamma$ (resp. $\alpha^{[n]}(f) = \gamma$). The class of all the tests with level γ is denoted by $\mathcal{K}_\gamma^{[n]}$:

$$\mathcal{K}_\gamma^{[n]} = \Big\{ f \in \mathcal{M}(\mathbb{R}^n, \{0,1\}) : \alpha^{[n]}(f) \leqslant \gamma \Big\}$$

Definition 9 (Power for the mean testing problem)**.** *The power of $f \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$ for testing the empirical mean of $\boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^n)$ such that $\mathbb{P}\big[|\langle \boldsymbol{\Theta} \rangle_n| > \tau\big] \neq 0$ when we are given $\boldsymbol{Y} = \boldsymbol{\Theta} + \boldsymbol{X} \in \mathcal{M}(\Omega, \mathbb{R}^n)$, with $\boldsymbol{X}$ independent of $\boldsymbol{\Theta}$, is defined by:*

$$\beta_{\boldsymbol{\Theta}}^{[n]}(f) = \mathbb{P}\big[f(\boldsymbol{\Theta} + \boldsymbol{X}) = 1 \,\big|\, |\langle \boldsymbol{\Theta} \rangle_n| > \tau\big] \tag{13}$$

The RDT solution. We can easily construct a preorder $\Big(\mathcal{K}_\gamma^{[n]}, \preceq^\diamond\Big)$ by setting:

$$\forall (f, f') \in \mathcal{K}_\gamma^{[n]} \times \mathcal{K}_\gamma^{[n]}, f \preceq^\diamond f' \text{ if } \forall \boldsymbol{\Theta} \in \mathcal{M}(\Omega, \mathbb{R}^n), \mathbb{P}\big[|\langle \boldsymbol{\Theta} \rangle_n| > \tau\big] \neq 0 \Rightarrow \beta_{\boldsymbol{\Theta}}^{[n]}(f) \leqslant \beta_{\boldsymbol{\Theta}}^{[n]}(f')$$

No maximal element exists in $\Big(\mathcal{K}_\gamma^{[n]}, \preceq^\diamond\Big)$. However, we can exhibit $\mathfrak{C}_\gamma^{[n]} \subset \mathcal{K}_\gamma^{[n]}$ whose elements satisfy suitable invariance properties with respect to the mean testing problem and prove the existence of a maximal element in $\Big(\mathfrak{C}_\gamma^{[n]}, \preceq^\diamond\Big)$.

Set $\mathcal{S} = \big\{\mathrm{id}, -\mathrm{id}\big\}$ where id is the identity of $\mathbb{R}$. Endowed with the usual composition law $\circ$ of functions, $(\mathcal{S}, \circ)$ is a group. Let $\mathcal{A}$ be the group action that associates to each given $s \in \mathcal{S}$ the map $\mathcal{A}_s : \mathbb{R}^n \to \mathbb{R}^n$ defined for every $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ by $\mathcal{A}_s(x) = (s(x_1), s(x_2), \ldots, s(x_n))$. Readily, the mean testing problem is invariant under the action of $\mathcal{A}$ in that $\mathcal{A}_s(\boldsymbol{Y}) = \mathcal{A}_s(\boldsymbol{\Theta}) + \boldsymbol{X}'$ where $\boldsymbol{X}' = (\boldsymbol{X}_1', \boldsymbol{X}_2', \ldots, \boldsymbol{X}_n') \sim \mathcal{N}(0, \mathbf{I}_n)$ is independent of $\mathcal{A}_s(\boldsymbol{\Theta})$. Therefore, $\mathcal{A}_s(\boldsymbol{Y})$ satisfies the same hypotheses as $\boldsymbol{Y}$. We also have $|\langle \mathcal{A}_s(\boldsymbol{\Theta}) \rangle_n| = |\langle \boldsymbol{\Theta} \rangle_n|$. Hence, the mean testing problem remains unchanged by substituting $\mathcal{A}_s(\boldsymbol{\Theta})$ for $\boldsymbol{\Theta}$ and $\boldsymbol{X}'$ for $\boldsymbol{X}$. It is thus natural to seek $\mathcal{A}$-invariant tests, that is, tests $f \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$ such that $f(\mathcal{A}_s(\boldsymbol{x})) = f(\boldsymbol{x})$ for any $s \in \mathcal{S}$ and any $\boldsymbol{x} \in \mathbb{R}^n$.

On the other hand, since we can reduce the noise variance by averaging observations, we consider $\mathcal{A}$-invariant integrator tests, that is, $\mathcal{A}$-invariant tests $f \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$ for which exists $\overline{f} \in \mathcal{M}\big(\mathbb{R}^1, \{0,1\}\big)$, henceforth called the reduced form of f, such that $f(\boldsymbol{x}) = \overline{f}(\langle \boldsymbol{x} \rangle_n)$ for any $\boldsymbol{x} \in \mathbb{R}^n$. Reduced forms of $\mathcal{A}$-invariant integrator tests are also $\mathcal{A}$-invariant: $\forall x \in \mathbb{R}$, $\forall s \in \mathcal{A}$, $\overline{f}(s(x)) = \overline{f}(x)$. We thus define $\mathfrak{C}_\gamma^{[n]} \subset \mathcal{K}_\gamma^{[n]}$ of all $\mathcal{A}$-invariant integrator tests. We thus have $f \in \mathfrak{C}_\gamma^{[n]}$ if:

[Size]: $\alpha^{[n]}(f) \leqslant \gamma$;

[$\mathcal{A}$-invariance]: $\forall (s, \boldsymbol{x}) \in \mathcal{S} \times \mathbb{R}^n$, $f(\mathcal{A}_s(\boldsymbol{x})) = f(\boldsymbol{x})$;

[Integration]: $\exists \overline{f} \in \mathcal{M}\big(\mathbb{R}^1, \{0,1\}\big)$, $\forall \boldsymbol{x} \in \mathbb{R}^n$, $f(\boldsymbol{x}) = \overline{f}(\langle \boldsymbol{x} \rangle_n)$.

The following result derives from the foregoing and [6, 7].

Proposition 2 (Maximal element of $\left(\mathfrak{C}_\gamma^{[n]}, \preceq^\diamond\right)$)**.** *For any $\gamma \in (0,1)$ and any $n \in \mathbb{N}$,*

$$\left\{f_{n,\tau}^{RDT(\gamma)}\right\} = \max\left(\mathfrak{C}_\gamma^{[n]}, \preceq^\diamond\right) \tag{14}$$

where $f_{n,\tau}^{RDT(\gamma)} \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$ is defined by setting:

$$\forall (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n, f_{n,\tau}^{RDT(\gamma)}(y_1, y_2, \ldots, y_n) = \begin{cases} 1 & \text{if} \quad \left|\sum_{i=1}^n y_i\right| \leqslant \sqrt{n}\lambda_\gamma(\tau\sqrt{n}) \\ 0 & \text{otherwise} \end{cases}$$

and $\lambda_\gamma(\tau\sqrt{n})$ is the unique solution in x to the equation $Q_{1/2}(\tau\sqrt{n}, x) = \gamma$

RDT and NP tests are structurally different because dedicated to two different testing problems and optimal with respect to two different criteria. This structural difference will be enhanced by coming back to our initial detection problem.

6.2. Application to Detection

Consider again the problem of estimating $\varepsilon \in \{0,1\}$, when we have a sequence $Y^{[q]} \in \mathrm{Seq}^{[q]}$ of observations. We thus have:

$$\forall n \in \mathbb{N}, \forall (\varepsilon, \omega) \in \left\{0,1\right\} \times \Omega, Y_n^{[q]}(\varepsilon, \omega) = \varepsilon + \Delta_n(\omega) + X_n(\omega) \tag{15}$$

where $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$ and $\forall n \in \mathbb{N}$, $\Delta_n \in \mathcal{B}_\infty(q)$ with $q \in [0,\infty)$. The empirical mean of $Y^{[q]}$ satisfies: $\forall n \in \mathbb{N}, \langle Y^{[q]}\rangle_n(\varepsilon) = \langle Y^{[q]}(\varepsilon)\rangle_n = \varepsilon + \langle\Delta\rangle_n + \langle X\rangle_n$. We thus have $|\langle\Delta\rangle_n| \leqslant q$ (a-s). Set $\mathbf{\Theta}_n = \varepsilon + \Delta_n$ for every $n \in \mathbb{N}$. In the sequel, we assume $q < 1/2$ because, in this case, we straightforwardly verify that

$$\begin{cases} \varepsilon = 0 & \Leftrightarrow & |\langle\mathbf{\Theta}\rangle_n| \leqslant q \\ \varepsilon = 1 & \Leftrightarrow & |\langle\mathbf{\Theta}\rangle_n| \geqslant 1 - q \end{cases} \tag{16}$$

Therefore, when $q \in [0, 1/2)$, deciding on whether ε is zero or not when we are given $\mathbf{Y}_n^{[q]}(\omega)$ amounts to testing whether $|\langle\mathbf{\Theta}\rangle_n(\omega)| \leqslant \tau$ or not for $\tau \in [q, 1-q]$. We thus can use the decision $f_{n,\tau}^{\mathrm{RDT}(\gamma)}\left(\mathbf{Y}_n^{[q]}\right)$, where $f_{n,\tau}^{\mathrm{RDT}(\gamma)}$ is given by Proposition 2.

We can calculate the false alarm probability of $f_{n,\tau}^{\mathrm{RDT}(\gamma)}\left(\mathbf{Y}_n^{[q]}\right)$ defined by (1) and the theoretical results in [6] yield that $\forall \tau \in [q, 1-q], \mathbb{P}_{\mathrm{FA}}\left[f_{n,\tau}^{\mathrm{RDT}(\gamma)}\left(\mathbf{Y}_n^{[q]}\right)\right] \leqslant \gamma$. In the sequel, for the sake of simplifying notation, we assume that both τ and q are in $[0, 1/2)$. In this case, we have:

$$\forall \tau \in [0, 1/2), \begin{cases} \mathrm{Sel}_\gamma\left(f_{n,\tau}^{\mathrm{RDT}(\gamma)}\right) = [0, \tau] \\ \mathfrak{D}_\gamma\left(f_{n,\tau}^{\mathrm{RDT}(\gamma)}\right) = \bigcup_{q \in [0,\tau]} \left\{f_{n,\tau}^{\mathrm{RDT}(\gamma)}\left(\mathbf{Y}_n^{[q]}\right) : Y^{[q]} \in \mathrm{Seq}^{[q]}\right\} \end{cases} \tag{17}$$

We can then state the following lemma, which is the counterpart to Lemma 3.

Theorem 3 (Maximality of RDT decisions)**.** *For any $\gamma \in (0,1)$, any $n \in \mathbb{N}$ and any $0 \leqslant q \leqslant \tau < 1/2$, $\mathfrak{D}_\gamma\left(f_{n,\tau}^{RDT(\gamma)}\right) = \max\left(\mathbb{D}_\gamma^{[0,\tau]}\left(\mathfrak{C}_\gamma^{[n]}\right), \preceq^*\right)$.*

Proof. It results from Definition 6 that $\mathfrak{D}_\gamma(f) := \left\{f\left(\mathbf{Y}_n^{[q]}\right) : Y^{[q]} \in \mathrm{Seq}^{[q]}, q \in [0,\tau]\right\}$. According to (8), we also have $\mathbb{D}_\gamma^{[0,\tau]}\left(\mathfrak{C}_\gamma^{[n]}\right) = \left\{\mathfrak{D}_\gamma(f) \in \mathbb{D}_\gamma : f \in \mathfrak{C}_\gamma^{[n]}, \mathrm{Sel}_\gamma(f) = [0,\tau]\right\}$. Given $q \in [0,\tau]$ and $Y^{[q]} \in \mathrm{Seq}^{[q]}$, set

$$\begin{cases} \mathbf{Y} & = & \mathbf{Y}_n^{[q]} = \left(Y_1^{[q]}, Y_2^{[q]}, \ldots, Y_n^{[q]}\right) \text{(see (4))} \\ \mathbf{X} & = & (X_1, X_2, \ldots, X_n) \sim \mathcal{N}(0, \mathbf{I}_n) \\ \mathbf{\Theta} & = & (1 + \Delta_1, 1 + \Delta_2, \ldots, 1 + \Delta_n) \end{cases}$$

We basically have $\boldsymbol{Y} = \boldsymbol{\Theta} + \boldsymbol{X}$. Consider now the mean testing problem (11) with $\boldsymbol{\Theta}$, $\boldsymbol{X}$ and $\boldsymbol{Y}$ defined as above. For any $f \in \mathcal{M}(\mathbb{R}^n, \{0,1\})$, it follows from Eqs. (15) , (16), (2) and (12) that:

$$\beta_{\boldsymbol{\Theta}}^{[n]}(f) = \mathbb{P}_{\text{DET}}\left[f\left(\boldsymbol{Y}_n^{[q]}\right)\right] \tag{18}$$

Suppose now that $f \in \mathfrak{C}_\gamma^{[n]}$ with $\text{Sel}_\gamma(f) = [0,\tau]$. We derive from Proposition 2, (18) and its application to $f_{n,\tau}^{\text{RDT}(\gamma)}$, that $\mathbb{P}_{\text{DET}}\left[f\left(\boldsymbol{Y}_n^{[q]}\right)\right] \leqslant \mathbb{P}_{\text{DET}}\left[f_{n,\tau}^{\text{RDT}(\gamma)}\left(\boldsymbol{Y}_n^{[q]}\right)\right]$. Since $q \leqslant \tau < 1/2$ implies that $q \in \text{Sel}_\gamma(f)$ and since $\text{Sel}_\gamma(f) = \text{Sel}_\gamma\left(f_{n,\tau}^{\text{RDT}(\gamma)}\right) = [0,\tau]$, we can rewrite the foregoing equality as $f\left(\boldsymbol{Y}_n^{[q]}\right) \preceq f_{n,\tau}^{\text{RDT}(\gamma)}\left(\boldsymbol{Y}_n^{[q]}\right)$. This holding true for any $q \in \text{Sel}_\gamma(f)$, any $Y^{[q]} \in \text{Seq}^{[q]}$ and since f and $f_{q,n}^{\text{RDT}(\gamma)}$ have same selectivity $[0,\tau]$, we derive from the foregoing and Definition 7 that $\mathfrak{D}_\gamma(f) \preceq^* \mathfrak{D}_\gamma\left(f_{n,\tau}^{\text{RDT}(\gamma)}\right)$. □

We now prove that the oracles with level γ are approximated by RDT decisions.

Lemma 4 (Approximation of oracles with γ by RDT decisions in $(2^{C_\gamma}, \preceq^*)$)**.**
Setting $\mathbb{D}_\tau^{\text{RDT}(\gamma)} := \left\{\mathfrak{D}_\gamma\left(f_{n,\tau}^{RDT(\gamma)}\right) : n \in \mathbb{N}\right\}$ *for any given* $\gamma \in (0,1)$*, we have:*

$$\mathbb{D}_\gamma^* = \text{upper}\left(\mathbb{D}_\tau^{\text{RDT}(\gamma)}, (2^{C_\gamma}, \preceq^*)\right) = \sup\left(\mathbb{D}_\tau^{\text{RDT}(\gamma)}, (2^{C_\gamma}, \preceq^*)\right)$$

Proof. According to (2) and [6, Theorem 2], we obtain:

$$\forall (q,\tau) \in [0,1/2) \times [0,1/2), \forall n \in \mathbb{N}, \mathbb{P}_{\text{DET}}\left[f_{n,\tau}^{\text{RDT}(\gamma)}\left(\boldsymbol{Y}_n^{[q]}\right)\right] \geqslant Q_{1/2}\big((1-q)\sqrt{n}, \lambda_\gamma(\tau\sqrt{n})\big)$$

In [10], we proved that $\lim_{\sigma\to 0} Q_{1/2}(\rho/\sigma, \lambda_\gamma(\tau/\sigma)) = \mathbb{1}_{]\tau,\infty[}(\rho)$. Therefore, since $\tau < 1-q$, we have $\lim_{n\to\infty} Q_{1/2}\big(\sqrt{n}(1-q), \lambda_\gamma(\tau\sqrt{n})\big) = 1$. Thus, $\lim_{n\to\infty} \mathbb{P}_{\text{DET}}\left[f_{n,\tau}^{\text{RDT}(\gamma)}\left(\boldsymbol{Y}_n^{[q]}\right)\right] = 1$. The set $\mathbb{D}_\tau^{\text{RDT}(\gamma)} \subset \mathbb{D}_\gamma$ satisfies Theorem 1 conditions for any $\gamma \in (0,1)$ and $\mathfrak{X}_\gamma = \{\tau\}$. □

7. Multiplicity Principle in $(2^{C_\gamma}, \preceq^*)$

To state the MP in $(2^{C_\gamma}, \preceq^*)$, we need the following lemma.

Lemma 5 (Selectivity of NP tests)**.** $\forall n \in \mathbb{N}$, $\text{Sel}_\gamma\left(f_n^{NP(\gamma)}\right) = \{0\}$

Proof. A consequence of [6, Section B., p. 6.]. □

We have now all the material to state the main result.

Theorem 4 (Multiplicity Principle in $(2^{C_\gamma}, \preceq^*)$)**.** *For any given* $\tau \in [0,1/2)$*, the MP is satisfied in* $(2^{C_\gamma}, \preceq^*)$ *by the pair* $\left(\mathbb{D}^{\text{NP}(\gamma)}, \mathbb{D}_\tau^{\text{RDT}(\gamma)}\right)$.

Proof. According to Theorems 2 and 3, the subsets $\mathbb{D}^{\text{NP}(\gamma)}$ and $\mathbb{D}_\tau^{\text{RDT}(\gamma)}$ of 2^{C_γ} are such that

$$\sup\left(\mathbb{D}^{\text{NP}(\gamma)}, (2^{C_\gamma}, \preceq^*)\right) = \sup\left(\mathbb{D}^{\text{RDT}(\gamma)}, (2^{C_\gamma}, \preceq^*)\right) = \mathbb{D}_\gamma^*$$

In addition, Lemma 5 implies that $\mathbb{D}^{\text{NP}(\gamma)} \times \mathbb{D}_\tau^{\text{RDT}(\gamma)} \cap \preceq^* = \emptyset$. The conclusion follows from Lemma 1. □

8. Conclusions and Perspectives

In this paper, via the framework provided by the Multiple Principle (MP), which is motivated by the concept of degeneracy in biology, and by introducing the notions of test landscapes and selectivity, we have established that this principle is satisfied when we consider the standard NP tests and the RDT tests applied to a detection problem. One interest of this result is that it opens prospects on the construction of Memory Evolutive Systems [2, 3] via tests.

More elaborated statistical decision problems should be considered beyond this preliminary work. Sequential tests are particularly appealing because they collect information till they can decide with guaranteed performance bounds. On the one hand, the Sequential Probability Ratio Test (SPRT) established in [11] is proved to be optimal; on the other hand, in [10], we have exhibited non-optimal tests with performance guarantees in presence of interferences. In the same way as NP and RDT tests satisfy PM, we conjecture that these two types of sequential tests satisfy MP as well.

References

[1] G. M. Edelman and J. A. Gally, "Degeneracy and complexity in biological systems," *Proceedings of the National Academy of Sciences*, vol. 98, no. 24, pp. 13 763–13 768, 2001. [Online]. Available: https://www.pnas.org/content/98/24/13763

[2] A. Ehresmann and J.-P. Vanbremeersch, *Memory Evolutive Systems; Hierarchy, Emergence, Cognition*, 1st ed., ser. Studies in multidisciplinarity. Elsevier, 2007, vol. 4.

[3] ——, "MES: A mathematical model for the revival of natural philosophy," *Philosophies*, vol. 4, no. 1, p. 9, Feb 2019. [Online]. Available: http://dx.doi.org/10.3390/philosophies4010009

[4] F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel, *Robust Statistics: the Approach based on Influence Functions*. John Wiley and Sons, New York, 1986.

[5] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*, 3rd ed. Springer, 2005.

[6] D. Pastor and Q.-T. Nguyen, "Random Distortion Testing and Optimality of Thresholding Tests," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4161 – 4171, Aug. 2013.

[7] D. Pastor and F.-X. Socheleau, "Random distortion testing with linear measurements," *Signal Processing*, vol. 145, pp. 116 – 126, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168417304127

[8] A. A. Borovkov, *Mathematical statistics*. Gordon and Breach Science Publishers, 1998.

[9] M. L. Eaton, *Multivariate statistics. A vector space approach*. Wiley, 1983.

[10] P. Khanduri, D. Pastor, V. Sharma, and P. K. Varshney, "Sequential Random Distortion Testing of Non-Stationary Processes," *IEEE Transactions on Signal Processing (accepted with minor revisions)*, 2019.

[11] A. Wald, "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, 06 1945.

OUR WORLDWIDE PARTNERS UNIVERSITIES - DOUBLE DEGREE AGREEMENTS

3 CAMPUS, 1 SITE

IMT Atlantique Bretagne–Pays de la Loire – **http://www.imt-atlantique.fr/**

Campus de Brest
Technopôle Brest-Iroise
CS 83818
29238 Brest Cedex 3
France
T +33 (0)2 29 00 11 11
F +33 (0)2 29 00 10 00

Campus de Nantes
4, rue Alfred Kastler
CS 20722
44307 Nantes Cedex 3
France
T +33 (0)2 51 85 81 00
F +33 (0)2 99 12 70 08

Campus de Rennes
2, rue de la Châtaigneraie
CS 17607
35576 Cesson Sévigné Cedex
France
T +33 (0)2 99 12 70 00
F +33 (0)2 51 85 81 99

Site de Toulouse
10, avenue Édouard Belin
BP 44004
31028 Toulouse Cedex 04
France
T +33 (0)5 61 33 83 65

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

© IMT Atlantique, 2019

Imprimé à IMT Atlantique
Dépôt légal : Juillet 2019
ISSN : 2556-5060