NOTE ON SIMPLICITY AND

STATISTICAL EXPLANATIONS OF CORRELATIONS

Chrysovalantis Stergiou[1]

The American College of Greece
Deree

November 2016

Abstract. In this note, I discuss the simplicity of rival statistical explanations of a correlation, couched in terms of Reichenbachian Common Cause Systems. Simplicity is analyzed in two components, the so-called intrinsic and contextual simplicity. I show that if one disentangles simplicity from explanatory power then the size of the system provides an adequate for simplicity in both of its dimensions.

1. Introduction

The idea that simplicity is a virtue for a scientific theory goes back to ancient times. Comparing demonstrations, Aristotle, in *Posterior Analytics,* claimed that, "…one demonstration [is] …better than another if, other things being equal, it depends on fewer postulates or suppositions or propositions…" [1, A, 86a33] – while in the *Almagest,* Ptolemy suggested that "…we consider it a good principle to explain the phenomena by the simplest hypotheses possible, in so far as there is nothing in the observations to provide a significant objection to such a procedure" [11, III 1]. However, the most famous account of ontological

[1] cstergiou@acg.edu ; valstergiou@gmail.com

parsimony is attributed to the medieval nominalist William of Ockham. He stipulated that a theory should posit no more ontologically independent entities than necessary - providing, thus, a rational justification for the choice of the most parsimonious theory - and he shaved away anything superfluous, which would not contribute substantially to an adequate description and explanation of matters of fact, by means of his famous razor. Later on, Newton, in his *magnum opus,* placed simplicity among his Rules for the Study of Natural Philosophy: "*No more causes of natural things should be admitted than are both true and sufficient to explain the phenomena*" [9, p.794].

In contemporary philosophy, the problem of simplicity is usually broken down into three different aspects: the first, concerns the definition and measurement of simplicity; the second one, has to do with its usage in different areas of research; and, the third aspect is about its rational justification as a theoretical virtue [2]. In this note, I consider only the first aspect of the problem, namely, the problem of measuring simplicity. However, far from providing a general account, I concentrate myself on those simplicity considerations that I deem necessary in order to measure the simplicity of a particular type of scientific explanations.

Statistical explanations made their way in philosophy only in the second half of twentieth century. According to Salmon's historical account of scientific explanation, they date back to 1962, when Rescher called for an explanatory argument "…which provides a rationalization of … [a] fact from premises which render it *not necessary but merely probable*" [14, p.50] and Hempel met this need by suggesting the first explicit model of statistical explanation [4]. However, particular aspects of statistical explanation have been investigated even earlier, when Reichenbach in 1956 suggested his account of probabilistic causality and his statistical model of common cause explanations of correlations in terms of conjunctive forks

[12, p. 159]. Conjunctive forks consist of a partition of the unit element of the event space induced by a single element (and its complement), that satisfies certain probabilistic relations that are deemed explanatory with respect to a given correlation. In the 90s a generalization of conjunctive forks has been proposed by Hofer-Szabó and Rédei [6] so as to consider as explanatory, partitions that have size greater than two, the Reichenbachian Common Cause Systems.

In this note, I try to compare two *equally adequate* statistical explanations of a given correlations couched in terms of partitions of the unity of the event algebra of a probability space in terms of a plausible criterion of simplicity. A basic assumption is that simplicity considerations *do not* influence explanatory power; this is what 'equally adequate' means. Simplicity has to do with the price to be paid in exchange for explanatory power. This price is measured in terms of the number of independent assumptions to be satisfied and the minimum information content required for a theory in order to be able to provide the requested explanation. Obviously, the smaller the price is, the simpler and more desirable the explanation is.

2. Simplicity considerations

To begin with, consider the problem to some degree of generality. Let H_1, H_2 be two rival hypotheses and $\{e_j\}_{j\in J}$ a collection of facts. Assume that H_1 *explains* $\{e_j\}_{j\in J}$, H_2 *explains* $\{e_j\}_{j\in J}$ and that H_1, H_2 do not differ in any other respects relevant to explanatory power. Can one compare the two explanatory hypotheses on the basis of an intuitively plausible criterion of simplicity?

It is quite common to consider a theory as a collection of models. An explanation is realized in a model of the theory only if that model, apart from the axioms of the theory, satisfies both the explanatory hypothesis H_i and the facts to be explained, $\{e_j\}_{j\in J}$[2]. Nevertheless, not all models of the theory make true both the explanans and the explanandum. For reasons that will be soon made clear, some models may not be 'rich' enough to satisfy these two conditions, while others might make true more than one explanatory hypotheses H_1, H_2 for a given collection of facts $\{e_j\}_{j\in J}$ giving, thus, rise to rival explanations realized in a single model of the theory.

Now, suppose that the models that satisfy both the fact to be explained and the explanatory fact can be singled out in terms of a certain set of independent assumptions stipulated to hold true. In general, these assumptions need not be considered something different from the ones describing the explanans and the explanandum. Their number may provide a natural measure for the simplicity of an explanation, since the simplest explanation of a fact depends on the least number of independent assumptions, other things being equal. Let us call this measure, *intrinsic simplicity*.

Of course, as it is well known, there is no unambiguous way of counting independent assumptions, since any two assumptions can be reduced to a single one, just by considering their conjunction [5, p. 42]. However, if one decided on how to particularize and count the assumptions on the basis of some local convention, then the problem would be partially overcome. By 'local convention', I intend a case-by-case convention that does not presuppose the application of any general rule, other than that of consistency: one decides in the case at hand, how to particularize and count the assumptions satisfied by two rival explanatory hypotheses and their explanandum and does that consistently. Of course, this would never determine any global measure of simplicity; hence my claim that the problem is only partially overcome.

---

[2] The requirement of satisfiability of the explanans and the explanandum by a model of the theory expresses one of Hempel's explanatory adequacy requirements, namely, the one that stipulates that the explanans and the explanandum are true.

What happens if two explanations of a given set of facts have the same intrinsic simplicity? Do they differ in any other way relevant to simplicity? I believe that the answer should be in the positive. Among the models that satisfy the *same number* of independent assumptions for two rival hypotheses H_1, H_2 and a given collection of facts $\{e_j\}_{j\in J}$, or, even, the *same set* of independent assumptions (in which case we refer to a single hypothesis H explaining $\{e_j\}_{j\in J}$), not all of them are equally informative; some might have greater information content than others. For instance, some models might have superfluous content with respect to the needs of the one or the other explanation realized within them and a restriction of the model might be equally adequate for explaining. This is because an explanation realized in the context of a model of a theory can also be realized in any extension of that model[3]. Let us call this aspect of simplicity *contextual,* since it depends on the information content of the model in which the explanation is realized when both the explanatory hypothesis and the explanans are satisfied. An intuitively plausible idea for the evaluation of contextual simplicity of an explanation is that the more information the explanation requires the less simple it is. In other words, contextual simplicity of an explanation of a given fact is a monotonously *decreasing* function of its information content.

At this point, let me clarify something that might raise confusion. For many philosophers, simplicity of a scientific theory is a monotonously *increasing* function of its information content. There are various justifications of this point the most straightforward of which is that a simple theory must say the most about its subject matter (have the greatest information content) by assuming

---

[3] Both the notions of extension and restriction of a model can be rendered precise, in a model-theoretic context by means of the *embedding of models.* Thus, one will say that a model is embedded into another if there is an isomorphism of the first model onto a sub-model of the second model. We will then call the first model a *restriction* of the second one and, respectively, the second model an *extension* of the first one.

the least possible [16, p. 119][4]. This ideal, however, is not applicable to our problem. What would it mean for an explanatory hypothesis to say the most about its subject matter? As I understand it, it would mean an increased capacity of explaining many and disparate matters of fact. However, as I formulated the problem since from the beginning, we seek to compare explanatory hypotheses of *a given set of facts*, in terms of a simplicity criterion. In addition, I assumed that the explanatory hypotheses to be compared *do not differ in any respect relevant to explanatory power.* Thus, the line of reasoning from increased information content to explanation of different facts, increased explanatory power, and, increased simplicity does not apply here; information content does not contribute to the simplicity of an explanation. On the contrary, as I explained before, a demand for increased information content of the model that satisfies both the explanans and the explanandum leads to diminished simplicity.

To take a step further, consider the simplicity of a hypothesis H which explains a collection of facts $\{e_j\}_{j\in J}$ in a model $\mathcal{M}_H$ of a scientific theory, $S\left(H,\{e_j\}_{j\in J};\mathcal{M}_H\right)$ as the product of the intrinsic simplicity of H, $IS\left(H,\{e_j\}_{j\in J}\right)$, with the contextual simplicity of H in $\mathcal{M}_H$, $CS\left(H,\{e_j\}_{j\in J};\mathcal{M}_H\right)$. Intrinsic simplicity may be regarded as inversely proportional to the number of independent assumptions employed in its definition and contextual simplicity as inversely proportional to the information content of the model that satisfy the hypothesis and the fact to be explained. Hence,

$$S\left(H,\{e_j\}_{j\in J};\mathcal{M}_H\right)=IS\left(H,\{e_j\}_{j\in J}\right)\times CS\left(H,\{e_j\}_{j\in J};\mathcal{M}_H\right)=\frac{1}{\begin{matrix}number\ of\ independent\ defining\\ assumptions\ of\ H\end{matrix}}\times\frac{1}{information\ content\ of\ \mathcal{M}_H}\ . \quad (1)$$

[4] For Popper, "[s]imple statements, if knowledge is our object, are to be prized more highly than less simple ones *because they tell us more; because their empirical content is greater; and because they are better testable.*" [10, p.128]

Notice that the simplicity measure defined in (1), depends on the model $\mathcal{M}_H$ in which the explanation is realized. This measure can be useful when one wants to compare different formulations of the same explanatory hypothesis; as in the one provided by a model and a sequence of its extensions. However, if the aim is to compare different hypotheses then one needs a model-independent simplicity measure. To tackle the problem, I contend that one should take into account the *minimum* information requirements that the model should satisfy in order to host the explanation. Hence,

$$S\left(H,\{e_j\}_{j\in J}\right) = IS\left(H,\{e_j\}_{j\in J}\right) \times CS\left(H,\{e_j\}_{j\in J}\right) =$$

$$\frac{1}{\substack{number\ of\ independent\ defining \\ assumptions\ of\ H}} \times \frac{1}{inf\left\{ information\ content\ of\ \mathcal{M}_H \middle| \substack{\mathcal{M}_H \text{ models of } T \\ \text{that satisfy } H\ and\ \{e_j\}_{j\in J}} \right\}}$$

, (2)

where $S\left(H,\{e_j\}_{j\in J}\right)$ and $CS\left(H,\{e_j\}_{j\in J}\right)$ is the simplicity and the contextual simplicity of H, respectively; the latter being calculated by taking the *infimum* of the information content over all models $\mathcal{M}_H$ that satisfy H and $\{e_j\}_{j\in J}$.

3. Statistical Correlations: Explanation and Simplicity

In this section, I apply the aforementioned considerations to the special problem of explaining statistical correlations by partitions. To begin with, I employ the theory of probability spaces in the models of which the statistical correlations, the *explananda*, are to be defined, as well as the Reichenbachian Common Cause Systems, the explanatory hypotheses. For those who find physical intuitions clearer, I urge them to employ the theory of thermodynamic systems, since as Mackey states in Postulate [A]: "[a] thermodynamic system is equivalent to a measure space." [8, p. 2]. However, I do not see how this can be more useful or informative, so I keep on in the realm of mathematics. The models $\mathfrak{R}$ of the theory are probability spaces, i.e. triples $\mathfrak{R} = (\Omega, \mathcal{F}, p)$ that satisfy the following conditions:

a) Ω is a non-empty set – the sample space;

b) $\mathcal{F}$ is a non-empty family of subsets of X which is closed under complement and countable union – the event algebra;

c) p is function, $p: \mathcal{F} \to [0,1] \subset \mathbb{R}$ such that $p(\Omega) = 1$ and p is countably additive; for a countable family $\{X_i\}_{i \in I}$, of pairwise disjoint members of $\mathcal{F}$ ($X_i \cap X_j = \emptyset$ for all $i \neq j,\ i, j \in I$), $p(\bigcup_{i \in I} X_i) = \sum_{i \in I} p(X_i)$.

Statistical correlations as any other probabilistic concepts acquire meaning only if defined within a probability space. Hence, given a probability space $\mathfrak{R} = (\Omega, \mathcal{F}, p)$ and $A, B \in \mathcal{F}$, the quantity

$$Corr_p(A, B) \doteq p(A \cap B) - p(A)p(B), \qquad (3)$$

for two events $A, B \in \mathcal{F}$ describes a statistical correlation of A, B in p if and only if $Corr_p(A, B) \neq 0$. In addition, A, B are said to be *positively* (*negatively*) *statistically correlated* if and only if, $Corr_p(A, B) > 0$, ($Corr_p(A, B) < 0$, respectively).

Consider a positive statistical correlation as the fact to be explained,

$$e = Corr_p(A, B) > 0. \qquad (4)$$

I am interested in statistical explanations of (4) given in terms of partitions of the unit of the event algebra $\mathcal{F}$. These partitions satisfy certain conditions and are called *Reichenbachian Common Cause Systems* [7, p. 81]. Hence, an explanatory hypothesis H of the correlation e would assert the existence of a Reichenbachian Common Cause System for e in $\mathcal{F}$, i.e. the existence of a partition

$$\{C_i\}_{i \in I} \subseteq \mathcal{F},$$

where $C_i \cap C_j = \emptyset$ for all $i \neq j,\ i, j \in I$, and $\bigcup_{i \in I} C_i = \Omega$, such that

$$p(A \cap B|C_i) = p(A|C_i)p(B|C_i) \text{ for all } i \in I \; ; \qquad (5)$$

$$[p(A|C_i) - p(A|C_j)][p(B|C_i) - p(B|C_j)] > 0, \text{ for all } i \neq j, \; i,j \in I. \qquad (6)$$

To make explicit the nature of the explanatory relation between the correlation e and the hypothesis H of the existence of a Reichenbachian Common Cause System for e in $\mathcal{F}$, I refer to the following three features: (a) Screening-Off, (b) Derivability and (c) Explanatory Relevance.

(a) Screening-Off: The correlation $e = Corr_p(A,B) > 0$ is screened-off in each cell C_i of the partition $\{C_i\}_{i \in I} \subseteq \mathcal{F}$, as indicated by (5). To understand this feature, think that were an explaining factor taken into account, the statistical correlation should disappear, since there is no other reason for the correlation to exist. In Reichenbach's own words, "[w]hen we say that the common cause C explains the frequent coincidence, we refer…also to the fact that relative to the cause C the events A and B are mutually independent." [12, p. 161]

(b) Derivability: The correlation $e = Corr_p(A,B) > 0$ is *entailed* by the totality of the explaining factors in the partition which is a Reichenbachian Common Cause System for the pair A, B [7, p. 81; Prop.7.2].

(c) Explanatory Relevance: The partition is uniformly relevant to the pair of events A, B, as (6) requires. This last feature guarantees the inclusion in the explanans of factors that are explanatory relevant to the explanandum, in the statistical sense, as stipulated by Salmon in his statistical relevance account of scientific explanation. In addition, given a Reichenbachian common cause system for e in $\mathcal{F}$ there is no strictly finer or coarser partitions of the unit of $\mathcal{F}$ that is a Reichenbachian common cause system for e in $\mathcal{F}$ as well. Thus, not only the partition contains *exclusively* relevant factors that screen-off the correlation but also *all* relevant factors that do the job (for a more detailed discussion of explanatory adequacy, see [15]).

As a last comment before taking up simplicity issues, I would like to mention that in the special case of partitions of the unit of $\mathcal{F}$ of size two, $(C, C^{\perp})$, in which cells the correlation e is screened-off, condition (6) is redundant, since it is automatically satisfied and the partition is a Reichenbachian common cause system of size 2 [18, p.35; Cor.4]. Furthermore, Wrónski has shown that this is not true for partitions of size greater than 2; in his example of a partition of size 3, the correlation is screened-off in each cell without it being a Reichenbachian common cause system [18, p.35; Example 1]. Apart from the obvious consequence that if an event and its complement in an event algebra screen-off a correlation then they are always explanatory relevant to it, this fact reduces the number of independent assumptions employed in the explanation.

Consider the explanatory hypothesis H that there is a Reichenbachian common cause system $\{C_i\}_{i\in I}$ for e in $\mathcal{F}$. The number of independent assumptions that the hypothesis satisfies, hence, its inner simplicity, can be expressed in terms of the size of $\{C_i\}_{i\in I}$, the cardinal number of its index set $|I|$. For a correlation of size greater than 2, $|I| > 2$, the number of independent assumptions satisfied is $|I| + \frac{|I|(|I|-1)}{2}$. While, for $|I| = 2$, which is the case of common cause explanations, the independent assumptions satisfied are just the screening-off conditions, as already explained; hence their number equals the size of the partition, it is 2 (= $|I|$), the, Thus, the inner simplicity of H is given by the following function:

$$IS(H,e) = \begin{cases} \frac{1}{|I| + \frac{|I|(|I|-1)}{2}} & , \text{ for } |I| > 2 \\ & , \\ \frac{1}{2}, & \text{for } |I| = 2 \end{cases} \qquad (7)$$

Notice that $0 \leq IS(H,e) \leq \frac{1}{2}$ with the countably infinite explanations of correlations having the least degree of inner simplicity, zero.

The problem of assessing the contextual simplicity of an explanatory hypothesis H of a correlation fact e that claims the existence of a Reichenbachian common cause system for e in $\mathcal{F}$, $CS(H, e; \mathcal{F})$, is the same as that of assessing the information content of the Boolean algebra $\mathcal{F}$ that satisfies both e and H. If we delimit our attention to the case of finite Boolean algebras, or, more precisely to the case of finite partitions of the unit of the algebra, the problem is easily tackled. In this case, a measure of the information content of the Boolean algebra can be given in terms of the number of events that are expressible in the context of that algebra; hence, by the cardinal number of the algebra, $|\mathcal{F}|$. Hence,

$$Information\ content\ of H\ in\ \mathcal{F} = |\mathcal{F}|$$

and

$$CS(H, e; \mathcal{F}) = {}^{1}\!/_{|\mathcal{F}|}\,. \qquad (8)$$

At this point, one may object that the information content of a theory should be identified with the number of theorems that can be deductively inferred from the postulates or the axioms of the theory and not with the cardinality of the event algebra of the theory. I disagree with this idea for two reasons: firstly, the number of theorems that are deducible from the axioms of a theory depend on the number of meaningful propositions that can be formulated in the context of that theory. However, as Reichenbach, in his first principle of the probability theory of meaning, suggests, "…a proposition has meaning if it is possible to determine a weight, i.e., a degree of probability, for the proposition." [13, p.54]. Thus, the cardinality of a Boolean algebra in a probability space measure the meaningful propositions that can be formulated in the context of the theory and provides an adequate measure of its information content. Secondly, seeking deductive consequences of the postulates would complicate things unnecessarily since it would leave us open to the undesirable "circumstance that the information (logical) content of each theory is infinite in the sense that one may deductively infer an infinite number of theorem-consequences from the postulates of a theory." [16]

To construct a model-independent measure of contextual simplicity for an explanation H of a correlation e, one needs to consider the minimum, in terms of cardinality, Boolean algebra that satisfies both the *explanans* and the *explanandum*. I will construct this algebra as a Boolean subalgebra of an arbitrarily chosen $\mathcal{F}$ that satisfies both H and e.

To begin with, consider a partition $H = \{C_i\}_{i\in I}$ of the unity of the event algebra $\mathcal{F}$ which satisfies (5) and (6) in a probability space $\mathfrak{R}$. Assume that $|I| < \infty$ and $\mathcal{L}_H$ is the Boolean subalgebra of $\mathcal{F}$ generated by $H = \{C_i\}_{i\in I}$. $\mathcal{L}_H$ is a finite algebra having $2^{|I|}$ elements, since it is isomorphic to the Boolean algebra of subsets of the family $\{C_i\}_{i\in I}$ [17, p.37; Lm 1].

By construction, $\mathcal{L}_H$ is the minimal Boolean algebra containing the explanatory partition. The next step is to extend this algebra so as to contain the correlated pair $A, B \in \mathcal{F}$ as well. The extension will be constructed in two steps; first I consider $\mathcal{L}_{H,\{A\}}$ the Boolean subalgebra of $\mathcal{F}$ generated by the set $\mathcal{L}_H \cup \{A\}$. $\mathcal{L}_{H,\{A\}}$ consists of elements of the form, $(X \cap A) \cup (Y \cap A')$, where $X, Y \in \mathcal{L}_H$ [3, p.83; Lm 2]. In order to calculate the cardinality of $\mathcal{L}_{H,\{A\}}$, consider the number of 2-permutations of the $2^{|I|}$ elements of $\mathcal{L}_H$, which corresponds to the number of elements of $\mathcal{L}_{H,\{A\}}$ given by the aforementioned formula for $X \neq Y$, and add to this number, $2^{|I|}$, which are is the number of the additional elements of $\mathcal{L}_{H,\{A\}}$, for $X = Y$. Hence,

$$|\mathcal{L}_{H,\{A\}}| = \left(2^{|I|}\right)_2 + 2^{|I|} = \frac{(2^{|I|})!}{(2^{|I|}-2)!} + 2^{|I|} = \left(2^{|I|} - 1\right)2^{|I|} + 2^{|I|} = 2^{2|I|}.$$

Next, I repeat the same procedure to construct $\mathcal{L}_{H,\{A,B\}}$, the subalgebra generated by the set $\mathcal{L}_H \cup \{A, B\}$. The cardinality of this algebra, which will provide a measure of the contextual simplicity of H, is the following:

$$|\mathcal{L}_{H,\{A,B\}},| = \left(2^{2|I|}\right)_2 + 2^{2|I|} = \frac{(2^{2|I|})!}{(2^{2|I|}-2)!} + 2^{2|I|} = \left(2^{2|I|} - 1\right)2^{2|I|} + 2^{2|I|} = 2^{4|I|}.$$

From the above considerations, one infers that

$$Information\ content\ of H = 2^{4|I|},$$

and

$$CS(H,e) = {}^{1}\!/_{2^{4|I|}}\,. \qquad (9)$$

Moreover, to the contextual simplicity $CS(H,e)$ of an infinite partition $\{C_i\}_{i\in I}$ constituting an explanatory hypothesis for e, I suggest to assign the value zero, since the Boolean algebra $\mathcal{L}_H$ generated by a countably infinite partition has the cardinality of the continuum [17, p.37; Lm 1].

Consequently, the simplicity of H , $S(H,e)$ for a finite partition is,

$$S(H,e) = \begin{cases} \frac{1}{|I|+\frac{|I|(|I|-1)}{2}} \cdot \frac{1}{2^{4|I|}} & , \text{ for } |I| > 2 \\ \\ \frac{1}{2^9} & , \text{ for } |I| = 2 \end{cases} , \quad (10)$$

while for countably infinite partitions is zero. Hence,

$$0 \le S(H,e) \le \frac{1}{2^9}\,. \quad (11)$$

As a final remark, notice that to compare two explanatory hypotheses of two respective correlations using the aforementioned suggestions, one first needs to identify the two correlations by means of a Boolean isomorphism that preserves the probability.

Let $\mathfrak{R}_1 = (\Omega_1, \mathcal{F}_1, p_1)$ and $\mathfrak{R}_2 = (\Omega_2, \mathcal{F}_2, p_2)$ be two probability spaces and $e_1 = Corr_{p_1}(A_1, B_1) > 0$, $e_2 = Corr_{p_2}(A_2, B_2) > 0$, where $A_1, B_1 \in \mathcal{F}_1$ and $A_2, B_2 \in \mathcal{F}_2$, two correlations defined in $\mathfrak{R}_1$ and $\mathfrak{R}_2$, respectively. In order to consider e_1 and e_2 identical, there must be a mapping $h: \mathcal{F}_1 \to \mathcal{F}_2$, which maps isomorphically the Boolean subalgebra of $\mathcal{F}_1$ generated by the correlated pair $A_1, B_1 \in \mathcal{F}_1$,

$$\mathcal{F}_{A_1B_1} = \begin{Bmatrix} \emptyset, A_1, B_1, A_1{}', B_1{}', A_1 \cap B_1, A_1 \cap B_1{}', A_1{}' \cap B_1, A_1{}' \cap B_1{}', A_1 \cup B_1, A_1 \cup B_1{}', \\ A_1{}' \cup B_1, A_1{}' \cup B_1{}', (A_1 \cap B_1) \cup (A_1{}' \cap B_1{}'), (A_1 \cap B_1{}') \cup (A_1{}' \cap B_1), \Omega_1 \end{Bmatrix}$$

onto the Boolean subalgebra $\mathcal{F}_{A_2B_2}$ of $\mathcal{F}_2$, generated by $A_2, B_2 \in \mathcal{F}_2$. In addition, it is required that at least with respect to the elements of $\mathcal{F}_{A_1B_1}$ the mapping preserves the measure,

$$\forall X \in \mathcal{F}_{A_1B_1},\ p_2\big(h(X)\big) = p_1(X).$$

Suppose that the aforementioned condition is satisfied then the *same* correlation is satisfied by two probability spaces $\mathfrak{R}_1$ and $\mathfrak{R}_2$,

$$e = Corr_{p_1}(A_1, B_1) = Corr_{p_2}(A_2, B_2) > 0, \qquad (8)$$

Next, we assume that there are partitions of the unity of the Boolean algebras $\mathcal{F}_1$ and $\mathcal{F}_2$,

$$H_1 = \{C^1{}_i\}_{i\in I_1} \subseteq \mathcal{F}_1 \text{ and } H_2 = \{C^2{}_i\}_{i\in I_2} \subseteq \mathcal{F}_2,$$

which satisfy relations (5) and (6) for the pairs $A_1, B_1 \in \mathcal{F}_1$ and $A_2, B_2 \in \mathcal{F}_2$, respectively. H_1 and H_2 are two hypotheses which explain adequately the *same* correlation e.

4. Concluding Remarks

Simplicity is a desideratum for scientific explanations. However, there is no consensus on how to measure simplicity. I suggest that one might disentangle simplicity from explanatory power and compare in terms of simplicity, equally adequate explanations of some fact. Given that, I put forth two different criteria of simplicity: firstly, the number of independent assumptions required by the explanatory hypothesis; the simpler explanation makes less assumptions - I call this intrinsic simplicity. Secondly, the information content of the model of the theory in which the explanation is realized. Again, the simpler requires a smaller information content. This type of simplicity is dubbed contextual. Contextual simplicity of an explanation depends on the model of the theory in the context of which the explanation is couched. However, there is a minimum information content requirement that defines a model-independent measure of contextual simplicity. These ideas apply to the case of Reichenbachian

Common Cause systems for a given correlation. After providing three conditions of explanatory adequacy, I suggest a measure of intrinsic simplicity, (7), that is a function of the size of the partition of the unity of the Boolean algebra of events. Then, I argue for the understanding of contextual simplicity in terms of the cardinality of the Boolean algebra of events, (8); in doing so I delimit the discussion to finite algebras. Lastly, I provide a model-independent measure of contextual simplicity, (9), as a function of the cardinality of the partition. Here the assumption is that the Boolean algebra with the minimum information content is the one generated by the partition and the correlated events. The bottom-line of this account is that the measure of simplicity of a Reichenbachian system is a decreasing function of its size. In particular, let $|I_1|, |I_2|$ be respectively the cardinalities of the explanatory partitions H_1, H_2 for a given correlation e: if $|I_1| < |I_2|$ then $S(H_1, e) > S(H_2, e)$; if $|I_1| = |I_2|$ then $S(H_1, e) = S(H_2, e)$; and if $|I_1| > |I_2|$ then $S(H_1, e) < S(H_2, e)$.

References

1. Aristotle, *Posterior Analytics.* In Barnes, J., (ed.). *The Complete Works of Aristotle: the revised Oxford translation vol.1*, Princeton-New Jersey: Princeton University Press (2014).
2. Baker, A., "Simplicity", The Stanford Encyclopedia of Philosophy (2013), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2013/entries/simplicity/>.
3. Givant, S., Halmos, P., *Introduction to Boolean Algebras.* New York: Springer, (2009)
4. Hempel, C.G., "Deductive Nomological vs. Statistical Explanation". In Feigl, H., and Grover Maxwell (eds.), *Minnesota Studies in the Philosophy of Science III.* Minneapolis: University of Minnesota Press (1962).
5. Hempel, C.G. *Philosophy of Natural Science*. New Jersey: Prentice Hall, (1966)
6. Hofer-Szabó, G., Redei, M. Reichenbachian common cause systems. Int. J. Theor. Phys. **43**, 1819–1826 (2004)

7. Hofer-Szabó, G., Rédei, M., and Szabó L.E. *The Principle of the Common Cause.* Cambridge: Cambridge University Press, (2013)
8. Mackey, M.C. *Time's Arrow: The Origins of Thermodynamic Behavior*. New York: Springer-Verlag (1992)
9. Newton, I. *The Principia: Mathematical Principles of Natural Philosophy.* Cohen, B., Whitman, A., and Budenz, J., (tr.). Berkeley, Los Angeles and London: University of California Press (1999)
10. Popper K. *The Logic of Scientific Discover*. London and New York: Routledge (2002)
11. Ptolemy. *Almagest,* Toomer, G. J. (tr.). London: Duckworth, (1984)
12. Reichenbach, H. *The Direction of Time.* Mineola, New York: Dover, (1956)
13. Reichenbach, H. *Experience and Prediction: An Analysis of the Foundations and the Structure of Knowledge*. Notre Dame, Indiana: University of Notre Dame Press ([1938], 2002)
14. Salmon, W. *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press (1989
15. Stergiou, C. Explaining Correlations by Partitions. *Foundations of Physics.* **45** (12):1599-1612 (2015)
16. Such, J. "Are there definitely falsifying procedures in science?" In Krajewsky, W. (ed.), *Polish Essays in the Philosophy of Natural Sciences.* Dordrecht-Boston-London: D Reidel Publishing Company, pp. 113-126, (1982)
17. Vladimirov, D.A. *Boolean Algebras in Analysis.* Dordrecht: Springer Science +Business Media, (2002)
18. Wrónski, L. *Reichenbach's Paradise: Constructing the Realm of Probabilistic Common "Causes".* Warsaw-Berlin: De Gruyter Open Ltd, (2014)