

A philosophical foundation of nonadditive measure and probability

Sebastian Maaß

Email: sebastian.maass@math.ethz.ch
ETH Zürich, Departement Mathematik
Rämistrasse 101, 8092 Zürich, Switzerland

ABSTRACT. In this paper, nonadditivity of a set function is interpreted as a method to express relations between sets which are not modeled in a set theoretic way. Drawing upon a concept called “quasi-analysis” of the philosopher Rudolf Carnap, we introduce a transform for sets, functions, and set functions. The transformed of a set can be interpreted as a class of components or properties representing this set. Since the transformed of a nonadditive set function is σ -additive, probabilistic concepts like independence and conditioning can be canonically introduced to the nonadditive theory via this transform.

KEY WORDS. nonadditive measure, conditioning, independence, products, Möbius transform, quasi-analysis

1 Introduction

The starting point of our research is that we want to interpret nonadditivity of a set function μ , i.e. $\mu(A \cup B) \neq \mu(A) + \mu(B)$ for disjoint sets A and B , in such a way that a) there is a relation between the sets A and B not being modelled in a set theoretic way since this would imply $A \cap B \neq \emptyset$ and that b) $\mu(A) + \mu(B) - \mu(A \cup B)$ is a measure for the strength of this relation. To formalize this idea, we draw upon a concept called “quasi-analysis” introduced by the philosopher Rudolf Carnap in “The logical construction of the world” in 1928 being a generalization of some abstraction principles of Frege, Russell and Whitehead.

Carnap’s concept of quasi-analysis can be outlined as follows. Suppose there is given a set of basis elements and a system of logical connections over this set, either in the form of binary relations or in the form of a set system over the set of basis elements. In the latter case, every element of the set system is interpreted to represent a component or property being shared by each of its elements. Since the basis elements are supposed to be indivisible unities, these sets are called “quasi-components” or “quasi-properties” of the basis elements. This method of analyzing basis elements using their quasi-components is called quasi-analysis. If the set system is large enough then every basis element can be represented by the set of its quasi-components.

The next step is to apply the quasi-analysis to our problem. Starting with a nonempty and for the moment finite set Ω and an algebra \mathcal{A} over Ω , we interpret sets $A \in \mathcal{A}$ in two different ways, either just as a set containing its elements or as a quasi-component of its elements. In the latter case A stands for a quasi-property or quasi-component being shared by all basis elements of A and not being shared by all basis-elements outside A . For every set $A \in \mathcal{A}$ of basis elements, we call a set $B \in \mathcal{A}$ a “quasi-component of A ”, if B is a quasi-component of some element of A , i.e. if $B \cap A \neq \emptyset$. Therefore, any set A of basis elements can be represented by the set A of its quasi-components, $A := \{B \in \mathcal{A} \mid B \cap A \neq \emptyset\}$. We then show that any set function $\mu : \mathcal{A} \rightarrow \mathbb{R}$ can be transformed into a (signed) measure μ on the algebra $\mathcal{A} := \sigma\{A \mid A \in \mathcal{A}\}$ generated by the sets of quasi-components of sets in \mathcal{A} such that μ preserves the values of μ , i.e. $\mu(A) = \mu(A)$ holds for all $A \in \mathcal{A}$. Furthermore, every μ -(Choquet-)integrable function $f : \Omega \rightarrow \mathbb{R}$ can be transformed into a function f on the set of quasi-components such that the value of the μ -integral is preserved, $\int f d\mu = \int f d\mu$. As one simple result, we obtain $\mu(A \cap B) = \mu(A) + \mu(B) - \mu(A \cup B)$ solving our problem introduced at the beginning. Thus, nonadditivity of a set function μ can be interpreted in such a way that μ

assigns to a set A the value of the set of its quasi-components A . Moreover, μ can be decomposed into the measure μ and the transform which maps to every set A the set of its quasi-components A .

With a view to this interpretation of nonadditive measure and integral we show how this transform can be applied to different fields of mathematical economics like cooperative game theory, multi-criteria decision making and expected utility theory by showing that the quasi-components have a meaningful interpretation in the respective theories. Moreover, we show that this transform enables us to transfer concepts like products, conditioning and independence from standard to nonadditive measure and probability theory in a very natural way. Our definitions of products, conditioning and independence are different from already existing ones but they are well-founded as they are based on a very comprehensive interpretation of nonadditive measures. Finally, we extend our mathematical results to the general, nondiscrete case. Mathematically, the presented transform is somehow dually to the Möbius transform and closely related to the representation of plausibility functions in terms of basic probability assignments.

2 Notations

Throughout this paper, let Ω denote a nonempty set, \mathcal{A} a σ -algebra over Ω , and $\mu : \mathcal{A} \rightarrow [0, \infty[$ a set function on \mathcal{A} with $\mu(\emptyset) = 0$. To every set function μ its **dual** set function $\bar{\mu}$ is defined by $\bar{\mu}(A) := \mu(\Omega) - \mu(A^c)$. A set function μ is called **monotone** if $A \subset B$ implies $\mu(A) \leq \mu(B)$. It is called **k-monotone** if

$$\mu(\cup_{i=1}^k A_i) + \sum_{\substack{I \subset \{1, \dots, k\} \\ I \neq \emptyset}} (-1)^{|I|} \mu(\cap_{i \in I} A_i) \geq 0 \quad \text{for } k \geq 2 \text{ and } A_1, \dots, A_k \in \mathcal{A},$$

k-alternating if $\bar{\mu}$ is k-monotone or, equivalently, if

$$\mu(\cap_{i=1}^k A_i) + \sum_{\substack{I \subset \{1, \dots, k\} \\ I \neq \emptyset}} (-1)^{|I|} \mu(\cup_{i \in I} A_i) \leq 0 \quad \text{for } k \geq 2 \text{ and } A_1, \dots, A_k \in \mathcal{A},$$

totally monotone if μ is monotone and k-monotone for every $k \geq 2$, and **totally alternating** if $\bar{\mu}$ is totally monotone or, equivalently, if μ is monotone and k-alternating for every $k \geq 2$. A **belief**

In the subsequent proposition we collect elementary results on the transformed sets. We omit the proof since it consists of simple calculations.

Proposition 3.1 For $A, B \in \mathcal{A}$

$$\begin{aligned} (a) \quad \emptyset = \emptyset, \quad \Omega = \mathcal{A} \setminus \{\emptyset\}, & \quad (c) \quad \widehat{A \cup B} = A \cup B, & \quad (e) \quad \widehat{A \setminus B} \supset A \setminus B, \\ (b) \quad A \subset B \Leftrightarrow \widehat{A} \subset \widehat{B}, & \quad (d) \quad \widehat{A \cap B} \subset A \cap B, & \quad (f) \quad A^c \subset \widehat{A}^c. \end{aligned}$$

Remark 3.2 The following statements on the relation between the $\widehat{\cdot}$ -operator and monotone sequences can easily be proved and will be of interest later on.

$$(A_n)_{n \in \mathbb{N}} \nearrow A \Rightarrow \bigcap_{n \in \mathbb{N}} A_n = \widehat{\bigcup_{n \in \mathbb{N}} A_n}, \quad (3)$$

$$(A_n)_{n \in \mathbb{N}} \searrow A, A_n \neq A \forall n \in \mathbb{N} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \not\supseteq \widehat{\bigcup_{n \in \mathbb{N}} A_n}. \quad (4)$$

Especially, in the latter case, $A^c \in \bigcap_{n \in \mathbb{N}} A_n \setminus \widehat{\bigcup_{n \in \mathbb{N}} A_n}$.

We now address the problem if every set-function $\mu : \mathcal{A} \rightarrow \mathbb{R}$ can be transformed into a measure $\mu : \mathcal{A} \rightarrow \mathbb{R}$ on the set of properties, satisfying

$$\mu(A) = \mu(\widehat{A}) \quad \text{for all } A \in \mathcal{A}, \quad (5)$$

i.e. the measure of the set A should equal the measure of all of its quasi-components. Such a transformed set function would solve our task for interpreting nonadditivity as a method of expressing a connection between the sets in \mathcal{A} which is not modeled in a set theoretic way. For two disjoint sets A and B , we interpret $\mu(A) + \mu(B) - \mu(A \cup B)$ as a measure of their connection. This interpretation perfectly fits with that of $A \cap B$ being the set of all common quasi-components of A and B and of $\mu(A \cap B)$ being a measure of their weight because

$$\begin{aligned} \mu(A \cap B) &= \mu(A) + \mu(B) - \mu(A \cup B) \\ &= \mu(A) + \mu(B) - \mu(\widehat{A \cup B}) \\ &= \mu(A) + \mu(B) - \mu(A \cup B). \end{aligned}$$

Furthermore, equation (5) can be interpreted in such a way that μ accumulates the weights of the quasi-components of any set $A \in \mathcal{A}$ and assigns this weight to the set A .

According to Remark 3.2, the transform μ of μ won't be σ -additive in general since whenever $\lim_{n \rightarrow \infty} \mu(A_n) < \mu(A)$ for a monotone increasing sequence $(A_n)_{n \in \mathbb{N}}$ we also would have that $\lim_{n \rightarrow \infty} \mu(A_n) < \mu(A)$ by equations (3) and (5), i.e. μ wouldn't be continuous from below. In order to make μ σ -additive, we have to enlarge the set of quasi-components to that effect that equation (3) changes to

$$(A_n)_{n \in \mathbb{N}} \nearrow A, A_n \neq A \forall n \in \mathbb{N} \Rightarrow \bigcap_{n \in \mathbb{N}} A_n \subsetneq \widehat{\bigcup_{n \in \mathbb{N}} A_n}, \quad (6)$$

while maintaining equation (4). A natural approach is very similar to the completion of the rational numbers via Cauchy sequences: The set of quasi-component (more precisely, the generator of \mathcal{A}) is then the set of equivalence classes of monotonously increasing sequences in \mathcal{A} , under the equivalence relation $(A_n) \sim (B_n)$ if there exists a monotonously increasing interleave sequence, i.e. if

$$(A_n) \sim (B_n) \quad :\Leftrightarrow \quad \exists (C_n) \text{ increasing with } \{C_n \mid n \in \mathbb{N}\} = \{A_n \mid n \in \mathbb{N}\} \cup \{B_n \mid n \in \mathbb{N}\}.$$

We state without a proof that every such equivalence class (A_n) can uniquely be represented by a plausibility function:

$$\nu_{[(A_n)]}(A) := \begin{cases} 1 & \text{if } \exists (B_n) \in (A_n) : A = B_n \text{ for some } n \in \mathbb{N} \\ 0 & \text{else} \end{cases}.$$

We therefore redefine

$$\mathcal{A} := \{ \nu \text{ plausibility function} \mid \nu(A) = 1 \text{ for all } A \in \mathcal{A} \}, \quad (7)$$

$$\mathcal{A} := \{ \sigma \mid A \in \mathcal{A} \}. \quad (8)$$

We call the elements of any set A the set of quasi-components of A . It is easy to prove that Proposition 3.1 remains valid and that inequalities (4) and (6) hold under the redefined terms. Moreover, it can be shown that in the discrete case any plausibility function is a dual unanimity game which again can be represented by a set in \mathcal{A} such that the definitions (1) and (7) as well as (2) and (8) coincide.

Furthermore implicitly understanding that there exists a transformation between nonadditive set function and signed measures satisfying equation (5), we now show how (μ -integrable) real-valued functions f

Theorem 3.3

- (a) For any set function μ on an algebra $\mathcal{A} \subset 2^\Omega$ there exists a unique (signed) measure μ on \mathcal{A} satisfying $\mu(A) = \mu(A)$ for all $A \in \mathcal{A}$.
- (b) The measure μ is nonnegative if and only if μ is totally monotone.
- (c) For any μ -integrable function f holds $f d\mu = \tilde{f} d\mu$.

The unique transformed μ of μ is then called the **Möbius transformed** of μ .

The existence part of Theorem 3.3 (a) can be shown algebraically (cf. Denneberg 1997 [6]) or using a version of Choquet's Theorem using that Ω consists of all extreme points of the set of belief functions (cf. Choquet 1953 [4]).

We now “dualize” the preceding results by switching from belief functions to their dual, the plausibility functions. Mathematically, this is merely a reformulation of Theorem 3.3 but in contrast to the Möbius transform, our results have nice interpretation.

Additionally to the transforms for sets, σ -algebras and functions introduced on the preceding page, we define $T : \Omega \rightarrow \Omega$ by $T(\nu) := \bar{\nu}$. Simple calculations yield $T^{-1}(A) = \Omega \setminus A^c$, i.e. T is measurable and transforms the generating system of \mathcal{A} into that of \mathcal{A} and thus, the domain of the image measure μ^T of μ under T is \mathcal{A} . Moreover,

$$\mu^T(A) = \mu(\Omega \setminus A^c) = \mu(\Omega) - \mu(A^c) = \mu(\Omega) - \mu(A^c) = \bar{\mu}(A).$$

This is already enough to obtain the “dual version” of Theorem 3.3.

Theorem 3.4

- (a) For any set function μ on an algebra \mathcal{A} there exists a unique (signed) measure μ on \mathcal{A} satisfying $\mu(A) = \mu(A)$ for all $A \in \mathcal{A}$.
- (b) The measure μ is nonnegative if and only if μ is totally alternating.
- (c) For any μ -integrable function f holds $f d\mu = f d\mu$.

Proof.

(a) Existence and uniqueness follows from Theorem 3.3 (a) for $\mu := \bar{\mu}^T$.

(b) Using Theorem 3.3 (b), we obtain

$$\begin{aligned} \mu \text{ is totally alternating} &\iff \bar{\mu} \text{ is totally monotone} \\ &\iff \bar{\mu} \text{ is nonnegative} \\ &\iff \mu \text{ is nonnegative.} \end{aligned}$$

(c) By Theorem 3.3 (c),

$$f d\mu = - \quad -f d\bar{\mu} = - \quad -f d\bar{\mu} = \quad f \circ T d\bar{\mu} = \quad f d\bar{\mu}^T = \quad f d\mu. \quad \square$$

In the discrete case, $\mu(A) = \mu(A)$ can be rewritten to

$$\mu(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

for all $A \in \mathcal{A}$ with $m : \mathcal{A} \setminus \{\emptyset\} \rightarrow \mathbb{R}$, $m(B) := \mu(\{B\})$ for all $B \in \mathcal{A}$. This result is well-known in the Dempster-Shafer theory of evidence as the representation of a plausibility functions by a *basic probability assignment* m (cf. Shafer [16]).

4 Examples and applications

In this section, we provide some examples and applications showing that the dual Möbius transform is nearly all-purpose in those theories making use of nonadditive set functions and their corresponding Choquet integrals. We show how the dual Möbius transform can be interpreted in multicriteria decision theory and in cooperative game theory.

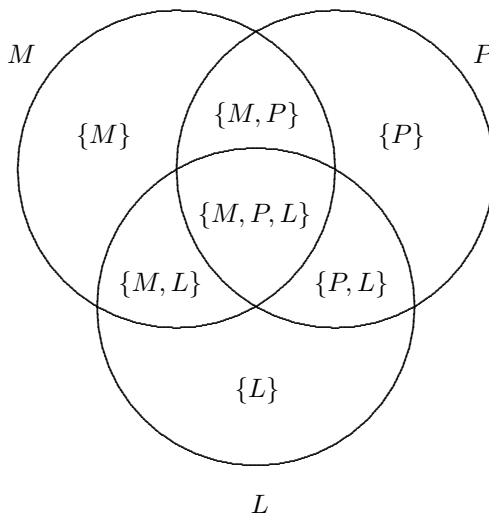
4.1 Multi-criteria decision making

Usually, a multicriteria decision problem consists of at least two different alternatives and at least two different criterias or attributes of the alternatives. A standard method to solve such a problem is the usage of aggregation: each alternative is represented by function on the set of criterias such that its values represent how “good” this criteria is distinct in this alternative. One has to introduce a weight function (measure) on the set of criterias and the rank problem can then be solved just by calculating the integral of each alternative representing function and rank them by their values. To proceed in this way, one has to impose that the criterias are unrelated since otherwise the common parts of different criterias are multiple weighted. Naturally, this presupposition is not fulfilled in all multicriteria decision problems like e.g. in the following evaluation problem (cf. Grabisch 1995 [10, p. 295] and Grabisch 1996 [11, p. 451]).

Ranking students (alternatives) on the basis of their marks in different subjects (criterias) poses the problem that the different subjects do not measure mutually disjoint skills. For instance, mathematical skills are not only tested in mathematics but also in physics and chemistry. How can the students be ranked in a reasonable way incorporating these redundancies? The simplest approach, i.e. calculating weighted sums (i.e. expected values) of the marks is not an appropriate method as mentioned above.

Instead of an expected value as an evaluation functional, Grabisch proposes the usage of the Choquet integral. The measurable space then consists of the set of all subjects together with an appropriate σ -algebra (usually the power set). A nonadditive set function μ is able to represent correlations between subjects. For example, a redundancy of mathematics and physics can be modeled by $\mu(\{\text{mathematics, physics}\}) < \mu(\{\text{mathematics}\}) + \mu(\{\text{physics}\})$. Roughly speaking, the more correlations between the subjects exists the more has the set function μ deviate from (σ -)additivity. In this method of resolution, the missing richness of the set of subjects to model correlations in a set theoretic way is compensated by the use of nonadditive set functions.

Now, let us consider how the problem can be solved using the new transform. Let $\Omega := \{M, P, L\}$ be the set of subjects (mathematics, physics, literature). The transformed subjects together with their quasi-components are mapped in the following diagram.



The subject mathematics then consists of four quasi-components: the component $\{M\}$ of skills exclusively necessary for mathematics, the component $\{M, P\}$ of skills necessary for mathematics and physics but not necessary for literature and so on. To model redundancy between mathematics and physics, a positive value must be assigned to the set $M \cap P = \{M, P\}, \{M, P, L\}$.

4.2 Cooperative game theory

We now show how to apply our new transform in cooperative game theory. Let us briefly recall the setting in cooperative game theory. Denote by

- S a nonempty set of players,
- $\mathcal{A} \subset 2^S$ an algebra over S interpreted as the set of all possible coalitions of players in S ,
- $v : \mathcal{A} \rightarrow \mathbb{R}_+$ the characteristic function of the game satisfying $v(\emptyset) = 0$ interpreted as the maximum utility/payoff the coalition A can get without correlating strategies with the other $S \setminus A$ players. This interpretation justifies the often imposed condition of v being superadditive.

In most cases, the characteristic function v is nonadditive. But we can interpret the nonadditivity in an overlap of the worth generating abilities of the coalitions. E.g. $v(C_1) + v(C_2) - v(C_1 \cup C_2) > 0$ or, equivalently, $v(C_1 \cap C_2) > 0$, means that the coalitions C_1 and C_2 share some of their worth gaining skills. Conversely, $v(C_1) + v(C_2) - v(C_1 \cup C_2) < 0$ resp. $v(C_1 \cap C_2) < 0$ could be interpreted in terms of that the coalitions C_1 and C_2 share some properties barring them from gaining worth. Thus, the dual Möbius transform of a characteristic function in cooperative game theory is a signed measure on the worth generating abilities of the players and coalitions.

5 Mathematical consequences

Since the transformed of a set is interpreted as the class of quasi-components representing this set, we can directly introduce probabilistic concepts via this transform. For example, two sets will be called independent if the two classes of quasi-components representing these sets are independent w.r.t. the transformed measure.

Definition 5.1 *Let (Ω, \mathcal{A}) be a measurable space and $\mu : \mathcal{A} \rightarrow [0, 1]$ be a set function interpreted as a not necessarily σ -additive probability measure. We call a family $(\mathcal{A}_i)_{i \in I}$, I nonempty index-set, of sub- σ -algebras of \mathcal{A} **independent** if for any finite nonempty subset J of I and any sets $A_j \in \mathcal{A}_j$, $j \in J$,*

$$\mu \left(\bigcap_{j \in J} A_j \right) = \prod_{j \in J} \mu(A_j).$$

Since, by the inclusion exclusion principle, $\mu \left(\bigcap_{j \in J} A_j \right) = \sum_{\emptyset \neq K \subset J} (-1)^{|K|+1} \mu \left(\bigcap_{k \in K} A_k \right)$, we obtain that independence of the family $(\mathcal{A}_i)_{i \in I}$ is equivalent to

$$\sum_{\emptyset \neq K \subset J} (-1)^{|K|+1} \mu \left(\bigcap_{k \in K} A_k \right) = \prod_{j \in J} \mu(A_j) \quad (13)$$

for any nonempty finite subset J of I .

*A family $(X_i)_{i \in I}$ of random variables is called **independent** if the family $(\sigma(X_i))_{i \in I}$ is independent, i.e. if*

$$\sum_{\emptyset \neq K \subset J} (-1)^{|K|+1} \mu \left(\bigcap_{k \in K} \{X_k \in B_k\} \right) = \prod_{j \in J} \mu(X_j \in B_j) \quad (14)$$

and any sets $B_k \in \mathcal{B}(\mathbb{R})$.

Definition 5.2 Given a nonadditive probability μ and two events A and B with $\mu(B) \neq 0$ the **conditional nonadditive probability** of A given B is defined as the conditional probability of A given B

$$\begin{aligned} \nu(A | B) &:= \mu(A | B) \\ &= \frac{\mu(A \cap B)}{\mu(B)} \\ &= \frac{\mu(A) + \mu(B) - \mu(A \cup B)}{\mu(B)}. \end{aligned} \tag{15}$$

For the definition of products of nonadditive set function we are confronted with the problem that we can not define $\mu_1 \otimes \mu_2$ via $\mu_1 \otimes \mu_2$ since the domains are not compatible. In the discrete case, we can give a natural definition.

Definition 5.3 Let μ_1 and μ_2 be two nonadditive set functions on a discrete measurable space Ω, \mathcal{A} . Then their product $\mu_1 \otimes \mu_2$ is defined by

$$\mu_1 \otimes \mu_2(A) := \sum_{\emptyset \neq K_i \in \mathcal{A}_i} \mu_1(K_1) \cdot \mu_2(K_2) \cdot \overline{u_{K_1 \times K_2}}(A). \tag{16}$$

For the general case, there exist some results for the Möbius transform (cf. e.g. Brüning [1]) which should be able to be transported to our setting similar to Theorem 3.4.

6 Conclusions and Outlook

We have introduced a transform for sets, functions, and set functions. Since the transformed of a set can be interpreted as the class of the properties or components representing this set, any result on the transformed space can directly interpreted in terms of the original space. This has enabled us to define independence and conditioning for nonadditive probability measures via the transform.

The first results arising from this new approach presented in this paper give reason to hope that the presented transform can establish a basis to build up a nonadditive probability theory in a very natural way.

References

- [1] Brüning, M.: (2003), *Products of Monotone Measures, Möbius Transform and k-monotonicity*. Dissertation thesis, Universität Bremen.
- [2] Carnap, R.: (1923), *Die Quasizerlegung. Ein Verfahren zur Ordnung nichthomogener Mengen mit den Mitteln der Beziehungslehre*. Unpublished manuscript RC 081-04-01, University of Pittsburgh.
- [3] Carnap, R.: (1928), *Der logische Aufbau der Welt*. Meiner, Hamburg.
- [4] Choquet, G.: (1953/54), Theory of Capacities, *Annales de l'Institut Fourier*, 5, 131-295.
- [5] Denneberg, D.: (1994), *Non-additive Measure and Integral*. Kluwer, Dordrecht.
- [6] Denneberg, D.: (1997), Representation of the Choquet integral with the σ -additive Möbius transform, *Fuzzy Sets and Systems*, 92, 139-156.
- [7] Denneberg, D.: (2000), Totally monotone core and products of monotone measures, *International Journal of Approximate Reasoning*, 24, 273-281.
- [8] Denneberg, D.: (2002), Conditional expectation for monotone measures, the discrete case, *Journal of Mathematical Economics*, 37, 105-121.

- [9] Gilboa, I. and Schmeidler, D.: (1989), Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics*, 18, 141-153.
- [10] Grabisch, M.: (1995), Fuzzy integral in multicriteria decision making, *Fuzzy Sets and Systems*, 69, 279-298.
- [11] Grabisch, M.: (1996), The application of fuzzy integrals in multicriteria decision making, *European Journal of Operational Research*, 89, 445-456.
- [12] Hendon, E., Jacobsen, H. J., Sloth, B. and Tranæs, T.: (1996), The product of capacities and belief functions, *Mathematical Social Sciences*, 32, 95-108.
- [13] Koshevoy, G. A.: (1998), Distributive lattices and products of capacities, *Journal of Mathematical Analysis and Applications*, 219, 427-441.
- [14] Murofushi, T. and Sugeno, M.: (1991), A theory for fuzzy measures: representations, the Choquet integral, and null sets, *Journal of Mathematical Analysis and Applications*, 159, 532-549.
- [15] Murofushi, T. and Sugeno, M.: (1989), An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure, *Fuzzy Sets and Systems*, 29, 201-227.
- [16] Shafer, G.: (1976), *A Mathematical Theory of Evidence*. Princeton Univ. Press, Princeton, NJ.
- [17] Walley, P. and Fine, T. L.: (1982), Towards a frequentist theory of upper and lower probability, *Annals of Statistics*, 10, 741-761.