Macsum: A new interval-valued linear operator

Olivier Strauss, Agnès Rico, Yassine Hmidy

Macsum: a new way to approximate linear operations

Olivier Strauss[a], Agnès Rico[b], Yassine Hmidy[a]

[a] *LIRMM, Université de Montpellier, CNRS, France*
[b] *LIRIS, Université Claude Bernard Lyon 1, CNRS, France*

Abstract

Many applications in fields as diverse as chemistry, mechanics, medicine, economics, robotics, environment, ecology, meteorology, etc. are based on the notion of system modeling. A system is a real process associating in a deterministic way an output value to one or more input values. A model is a mathematical object that allows the analysis of real phenomena and the prediction of results at a given level of approximation. One of the difficulties of modeling is the choice of the model and how to control the level of approximation. The linear model, where the output is obtained by a weighted sum of the inputs, is a simple model, based on a reduced number of parameters, but describing the functioning of a system in a very approximate way, without the level of approximation being known. Non-linear models are much more specific but much more difficult to use, the level of approximation being even more difficult to control. What we propose in this article is an imprecise linear model, so the simplicity of representation and use is quite comparable to a linear model. This model is imprecise in the sense that the output is imprecise, although the inputs are precise, thus potentially reflecting how close the model is to the system behavior: the more imprecise the output, the less likely the model is to describe the system correctly. This imprecise linear model can be seen as a convex set of conventional linear models, the imprecise output of this model being the convex set of outputs that would have been obtained by each linear model individually. This modeling is based on non-monotonic real-valued concave set measures.

Key words: Imprecise linear system, Choquet integral, non-additive aggregation, non-monotonic set functions

1. Introduction

Linear relationships between entities occupy a prominent place in a plethora of subjects, as diverse as chemistry, mechanics, medicine, economics, robotics, environment, ecology, meteorology, etc. Whether it is expressing the voltage as a function of the current in an electronic circuit [1] or the budget constraint of

Email addresses: Olivier.Strauss@lirmm.fr (Olivier Strauss), Agnes.Rico@univ-lyon1.fr (Agnès Rico), Yassine.Hmidy@lirmm.fr (Yassine Hmidy)

a homo-economicus in a model of microeconomics [6], these questions require knowledge of a precise proportionality relating the evolution of one variable to another. Modeling a system, whether it is mechanical, economic, robotic, chemical, biological, medical, etc., makes it possible to predict how this system will behave when knowing the system inputs. For example, in medicine, this allows for diagnostic assistance [16], in mechanics to predict the performance of an assembly and thus optimize the manufacturing of objects [18], in chemistry to develop new products without real tests thanks to simulation techniques [32], in econometrics to improve predictive models [8], etc. In many of these domains, linear models have a preponderant place because of their simplicity of implementation, their efficiency and their predictive power. This preeminence of linear models has increased with the popularization of computers.

A linear model can be seen as a linear aggregation operation involving a set of weights, the output of the model being a weighted sum of the inputs, the weights being symptomatic of the system that we want to describe. In the field of systems and signal processing, this set of weights is called a convolution kernel or impulse response. One of the difficulties is to choose the weights that best represent the system that we want to model. Several methods exist under the generic term of linear regression, potentially including a prior information on the input-output relations of the system and the ability of the model to describe the system. This fitness is usually characterized by the distance between the output predicted by the model and the actual output of the system on a benchmark data-set used to identify (or learn) the system.

Of course, a linear model is only an approximation of the system behavior to be described, linearity being rare in our world. Using a linear model means to use an approximation of the real system that we want to control, predict, ... One of the problems we are often confronted with is that, although we are aware that the linear model is an approximation of a real system, it is often not easy to know how close the output predicted by the system is to the real output. The identification of the weights characterizing the linear system is usually based on an optimality criterion (a quadratic distance for example) but nothing really allows to characterize the adequacy of this model to the real system: no robust method is available to make use of this distance to predict how close is the output of the model to the output of the system. More complex models are proposed (non-linear models for example) which are less easy to use: a slight gain in accuracy is generally made at the expense of the simplicity of use of the model.

Several approaches are proposed in the literature to describe a system in an approximate way. Some authors focus on dealing with parametric models whose imprecise parameters are specified by numerical intervals. On a theoretical level, for example, Shary [36] proposes to consider solving ill-conditioned systems of linear algebraic equations by *intervalizing* the parameters of the sough after system. They mention in a very relevant way the ambivalence of the meaning of an interval in this context, which is known in the fuzzy literature under the name of *epistemic* and *ontic* interpretation of the intervals [7]. On a more applied level, Rinner and Weiss [34] propose using this interval-valued representation

to deal with incomplete knowledge about a system to be supervised. In [2], Boukezzoula et al. propose to refine this imprecise representation by using a gradual number-based representation of the parameters. They propose different regression methods for identifying this sophisticate representation [3]. In a completely different manner, in [21], Jaulin et al. propose to approximate a function by an interval-based decomposition of the domain of the function to be approximated. This representation leads to very interesting tools for approximately represent imprecise non-linear input-output systems [22] with applications to robotics [20]. In a recent article [4], an attempt to fuse the two previous approaches is proposed. In the field of decision theory, it has been proposed to approximate a system using imprecision on the model itself. Walley [43] first and then many other authors in multi-criteria decision making [5, 24, 44, 35] or random sets approximation [15] have proposed the imprecise probability theory to represent the fact that the probabilistic model is poorly known, it could not be summarized by a distribution of weights. They propose to move to a more complex model inducing an imprecision on the prediction generated by the model. This can be applied in many fields such as artificial intelligence [23]. However, most of the work carried out to date to represent both a model and its approximate power leads to complex representations that are not easy to use.

In recent work, Loquin et al, inspired by a particular case of imprecise probability theory, possibility measures, have proposed an imprecise model of a convolution kernel under the name of maxitive kernel [25]. A maxitive kernel can be seen as a convex set of convolution kernels. The extension of the aggregation operation (also called convolution) proposed by Loquin et al. allows to compute the (convex) set of outputs that would have been obtained using this convex set of convolution kernels [33]. One of the main advantages of this representation is that its complexity is low enough to be comparable with that of classical linear models. A disadvantage of this representation is that it only allows to represent linear systems whose kernel is positive and normalized, i.e. linear systems whose weights are positive and sum to one.

What we propose in this paper is to extend the work of Loquin et al. to any convolution kernel, thus to any linear system. We end up with what we call the **macsum** representation of a system, which can be interpreted as an imprecise linear representation of a system whose imprecision can be controlled.

2. Theoretical background

2.1. Notations

- $\Omega = \{1, \ldots, N\} \subset \mathbb{N}$.
- $x : \Omega \to \mathbb{R}^N$, is a function defined by a discrete subset of $\mathbb{R}^N$: $x = (x_1, \cdots, x_N) \in \mathbb{R}^N$.
- $\lceil . \rceil$ is the permutation that sorts the x_i's in increasing order: $x_{\lceil 1 \rceil} \leq x_{\lceil 2 \rceil} \leq \cdots \leq x_{\lceil N \rceil}$.

- $\lfloor . \rfloor$ is the permutation that sorts the x_i's in decreasing order: $x_{\lfloor 1 \rfloor} \geq x_{\lfloor 2 \rfloor} \geq \cdots \geq x_{\lfloor N \rfloor}$.
- $A_{\lceil i \rceil}$ $(i \in \Omega)$ is the coalition of Ω defined by $A_{\lceil i \rceil} = \{\lceil i \rceil, \ldots, \lceil N \rceil\}$.
- $A_{\lfloor i \rfloor}$ $(i \in \Omega)$ is the coalition of Ω defined by $A_{\lfloor i \rfloor} = \{\lfloor i \rfloor, \ldots, \lfloor N \rfloor\}$

2.2. Kernels and capacities

The aim of this section is to define the fundamental notions that will be used throughout this article.

- A **kernel** of Ω is a discrete function $\varphi : \Omega \to \mathbb{R}$ defined by $\varphi = (\varphi_1, \cdots, \varphi_N)$.
- The set of kernels of Ω is denoted $\mathcal{K}(\Omega)$.
- A set function is a function $\vartheta : 2^\Omega \to \mathbb{R}$ that associates a real value to any subset of Ω.
- A set function ϑ of Ω is said to be **concave** if $\forall A, B \subseteq \Omega$, $\vartheta(A \cup B) + \vartheta(A \cap B) \geq \vartheta(A) + \vartheta(B)$.
- A set function ϑ of Ω is said to be **convex** if $\forall A, B \subseteq \Omega$, $\vartheta(A \cup B) + \vartheta(A \cap B) \leq \vartheta(A) + \vartheta(B)$.
- A set function ϑ of Ω is said to be **additive** if $\forall A, B \subseteq \Omega$, $\vartheta(A \cup B) + \vartheta(A \cap B) = \vartheta(A) + \vartheta(B)$.
- To a set function ϑ of Ω can be associated a complementary set function ϑ^c defined by $\forall A \subseteq \Omega$, $\vartheta^c(A) = \vartheta(\Omega) - \vartheta(A^c)$, A^c being the complementary set of A in Ω.
- If a set function ϑ is concave (rsp. convex) then ϑ^c is convex (rsp. concave).
- A capacity is a normalized increasing set function $\upsilon : 2^\Omega \to \mathbb{R}^+$ with $\upsilon(\varnothing) = 0$. Normalized means $\upsilon(\Omega) = 1$ and increasing means that $\forall A \subseteq B \subseteq \Omega$, $\upsilon(A) \leq \upsilon(B)$. To a capacity υ is associated its complementary capacity υ^c: $\forall A \in \Omega$, $\upsilon^c(A) = 1 - \upsilon(A^c)$.
- A **maxitive kernel** of Ω is a discrete function $\pi : \Omega \to [0, 1]^N$ defined by $\pi = (\pi_1, \cdots, \pi_N)$ such that $\max_{i \in \Omega} \pi_i = 1$.
- The set of maxitive kernels of Ω is denoted $\mathcal{K}_m(\Omega) \subseteq \mathcal{K}(\Omega)$.
- To a maxitive kernel can be associated a capacity Π_π defined by $\forall A \subseteq \Omega$ $\Pi_\pi(A) = \max_{i \in A} \pi_i$. This function is called a **possibility** measure when π is interpreted as a possibility distribution [11]. Its complementary set function $\Pi_\pi^c(A) = 1 - \Pi_\pi(A^c) = 1 - \max_{i \in A^c} \pi_i$ is called a **necessity** measure in the context of confidence measures (possibility theory).

- A **summative kernel** of Ω is a discrete function $\rho : \Omega \to (\mathbb{R}^N)^+$ defined by $\rho = (\rho_1, \cdots, \rho_N)$ such that $\sum_{i\in\Omega} \rho_i = 1$.
- The set of summative kernels of Ω is denoted $\mathcal{K}_s(\Omega) \subseteq \mathcal{K}(\Omega)$.
- To a summative kernel can be associated an additive set function P_ρ defined by $\forall A \subseteq \Omega$, $P_\rho(A) = \sum_{i\in A} \rho_i$. P_ρ is called a **probability** measure in the context of confidence measure. The complementary set function P_ρ^c of P_ρ is P_ρ itself since
  $P_\rho^c(A) = P_\rho(\Omega) - P_\rho(A^c) = \sum_{i\in\Omega} \rho_i - \sum_{i\in A^c} \rho_i = \sum_{i\in A} \rho_i$.
- This can be generalized to any kernel: let $\varphi = (\varphi_1, \cdots, \varphi_N)$ be an unnormalized ($\sum_{i\in\Omega} \varphi_i \neq 1$) and signed ($\exists i \in \Omega$ such that $\varphi_i < 0$) kernel. The additive set function associated to φ is $P_\varphi(A) = \sum_{i\in A} \varphi_i$.
- A maxitive kernel $\pi \in \mathcal{K}_m(\Omega)$ is said to dominate a summative kernel $\rho \in \mathcal{K}_s(\Omega)$ if $\forall A \subseteq \Omega$, $\Pi_\pi(A) \geq P_\rho(A)$. [10]
- The set of summative kernels dominated by a maxitive kernel π, denoted as $\mathcal{M}(\pi)$, is defined by:
  $\mathcal{M}(\pi) = \{\rho \in \mathcal{K}_s(\Omega) : \forall A \subseteq \Omega\ \Pi_\pi^c(A) \leq P_\rho(A) \leq \Pi_\pi(A)\}$. This definition refers to the core of a capacity. The core of a capacity υ denoted as $\mathcal{M}(\upsilon)$ is the set of probability measures that it dominates. When υ is concave, this can be written $\mathcal{M}(\upsilon) = \{P \in \mathcal{P}(\Omega) : \forall A \subseteq \Omega,\ \upsilon^c(A) \leq P(A) \leq \upsilon(A)\}$, where $\mathcal{P}(\Omega)$ is the set of probability measures defined on Ω.

2.3. Choquet-based aggregation

The Choquet integral is a way to aggregate real values with respect to a capacity. The basic Choquet integral has been defined to extend the notion of expectation to non-additive confidence measures (also called capacity) [9].

Let $x \in \mathbb{R}^N$. The literature generally reports two ways for computing the discrete Choquet integral of x with respect to the capacity υ: $y = \mathbb{C}_\upsilon(x)$ [30].

$$y = \mathbb{C}_\upsilon(x) = \sum_{k=1}^{N} x_{\lceil k \rceil}.(\upsilon(A_{\lceil k \rceil}) - \upsilon(A_{\lceil k+1 \rceil})), \text{with } A_{\lceil N+1 \rceil} = \varnothing. \tag{1}$$

$$y = \mathbb{C}_\upsilon(x) = \sum_{k=1}^{N} (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\upsilon(A_{\lceil k \rceil}), \text{with } x_{\lceil 0 \rceil} = 0. \tag{2}$$

Those formulas need the values of x to be positive. The asymmetric Choquet integral, denoted $\check{\mathbb{C}}_\upsilon(x)$, has been defined to generalize the Choquet integral for signed real values: $\check{\mathbb{C}}_\upsilon(x) = \mathbb{C}_\upsilon(x^+) - \mathbb{C}_{\upsilon^c}(x^-)$, with $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$.

Proposition 2.1. *Equations* (1) *and* (2) *can be used to compute the asymetric Choquet integral.*

Proof. Remember that the asymetric Choquet integral coincides with the Choquet integral for positive values. Moreover, its translative invariant, i.e. $\forall a \in \mathbb{R}$, $\check{\mathbb{C}}_v(x+a) = \check{\mathbb{C}}_v(x) + a$.

Thus, let us define $x' = x - a$, with $a = \min_{k\in\Omega} x_k$. By construction $\forall k \in \Omega$, $x'_k \in \mathbb{R}^+$. Moreover, the permutation that sorts the vector x sorts also the vector x'. Thus:

$$\check{\mathbb{C}}_v(x') = \mathbb{C}_v(x') = \sum_{k=1}^{N} x'_{\lceil k \rceil}.(v(A_{\lceil k \rceil}) - v(A_{\lceil k+1 \rceil})) = \check{\mathbb{C}}_v(x) - a.$$

$$y = \check{\mathbb{C}}_v(x) = \check{\mathbb{C}}_v(x')+a = \sum_{k=1}^{N}(x'_{\lceil k \rceil}+a).(v(A_{\lceil k \rceil})-v(A_{\lceil k+1 \rceil})) = \sum_{k=1}^{N} x_{\lceil k \rceil}.(v(A_{\lceil k \rceil})-v(A_{\lceil k+1 \rceil})).$$

In the same way:

$$\begin{aligned} y = \check{\mathbb{C}}_v(x) = \mathbb{C}_v(x') + a &= \sum_{k=1}^{N}(x'_{\lceil k \rceil} - x'_{\lceil k-1 \rceil}).v(A_{\lceil k \rceil}) + a \\ &= (x'_{\lceil k \rceil} - 0).v(A_{\lceil 1 \rceil}) + a + \sum_{k=2}^{N}(x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).v(A_{\lceil k \rceil}), \end{aligned}$$

because $\forall i, j \in \Omega$, $x_i - x_j = x'_i - x'_j$. Thus:

$$y = (x'_{\lceil k \rceil}+a-0).v(A_{\lceil 1 \rceil})+\sum_{k=2}^{N}(x_{\lceil k \rceil}-x_{\lceil k-1 \rceil}).v(A_{\lceil k \rceil}) = \sum_{k=1}^{N}(x_{\lceil k \rceil}-x_{\lceil k-1 \rceil}).v(A_{\lceil k \rceil})$$

□

The Choquet integral with respect to v^c, the conjugate operator of v, can be computed easily by remembering that $\check{\mathbb{C}}_{v^c}(x) = -\check{\mathbb{C}}_v(-x)$.

2.4. Recent advances in non-monotonic set functions and integrals

Choquet capacities are increasing and normalized, i.e. monotonic. Choquet integral has been defined to extend expectation to non-additive confidence measures. However, as shown by the seminal work of Murofushi et al. [40], except when set functions have to be interpreted as confidence measures, monotonicity is inessential. A non-monotonic set function (also called fuzzy measure) is a function $\mu : 2^\Omega \to \mathbb{R}$ such that $\mu(\varnothing) = 0$.

Several work, including [14, 30, 41, 31] proposed to extend Choquet integral to non-monotonic set functions. In particular, in [41], Waegenære and Wakker shown that Expressions (1) and (2) can be used to compute the Choquet integral w.r.t. a non-monotonic set function. Among the properties mentioned in [40] we have that the non-monotonic Choquet integral is homogeneous positive that is to say $\check{\mathbb{C}}_\mu(\lambda.x) = \lambda.\check{\mathbb{C}}_\mu(x)$ with $\lambda \geq 0$. Moreover $\check{\mathbb{C}}_\mu$ is comonotically additive, which means that if f and g are two comonotic measurable functions , we have $\check{\mathbb{C}}_\mu(f+g) = \check{\mathbb{C}}_\mu(f) + \check{\mathbb{C}}_\mu(g)$.

2.5. Representing a convex set of linear aggregations by a maxitive aggregation

A linear aggregation, also called a weighted mean, is a function that associates to each vector $x \in \mathbb{R}^N$ a real value y computed as: $y = \mathbb{E}_\rho(x) = \sum_{k\in\Omega} \rho_k.x_k$, where $\rho \in \mathcal{K}_s(\Omega)$ is a summative kernel of Ω. When ρ is interpreted as a probability distribution, then the value y is called the expectation of x. Let P_ρ be the additive set function associated to ρ, then $\mathbb{E}_\rho(x) = \check{\mathbb{C}}_{P_\rho}(x)$.

In [26], Loquin et al. propose a new aggregation, under the name of **maxitive expectation**, denoted $\overline{\underline{\mathbb{E}}}_\pi$, where $\pi \in \mathcal{K}_m(\Omega)$ is a maxitive kernel of Ω. The goal of this representation is to deal with the idea that the appropriate summative kernel to be used to aggregate the information of x is imprecisely known. Contrarily to the linear aggregation , $\overline{\underline{\mathbb{E}}}_\pi$ leads to an imprecise expectation $\overline{\underline{\mathbb{E}}}_\pi(x) = [\underline{y}, \overline{y}] = [\check{\mathbb{C}}_{\Pi^c_\pi}(x), \check{\mathbb{C}}_{\Pi_\pi}(x)]$. It has been shown that:

Proposition 2.2. $\forall \pi \in \mathcal{K}_m(\Omega)$, $\rho \in \mathcal{M}(\pi) \Leftrightarrow \mathbb{E}_\rho(x) \in \overline{\underline{\mathbb{E}}}_\pi(x)$.

Therefore, the maxitive aggregation $\overline{\underline{\mathbb{E}}}_\pi(x)$ of x w.r.t. the maxitive kernel π can be seen as the convex set of all additive aggregations $\mathbb{E}_\rho(x)$ of x w.r.t. a summative kernel $\rho \in \mathcal{M}(\pi)$. This has many potential applications in signal processing [28], image processing [27, 12, 13], statistics [25], etc.

However, in most domains except statistics, the fact that this modeling can only represent convex sets of weighted sums w.r.t. a summative kernel can be perceived as very restrictive. For example, in signal processing, high-pass filters cannot be represented by using a summative kernel based aggregation since the weights are signed. The same acts, in image processing, with interpolation operations, since, except for the nearest neighbor and linear interpolations, most interpolation kernels are signed [38].

To circumvent this problem, in [39] Rico and Strauss have proposed a signed extension to the concept of maxitive kernels. This method is based on considering separately the positive and negative part of a summative kernel. A *signed maxitive kernel* can be seen as a pair of two maxitive kernels (π^+, π^-), one representing a convex set of positive parts and the other one a convex set of negative parts of summative kernels. To a maxitive kernel is associated a particular non-monotonic set function μ_{π^+,π^-} that is still normalized (i.e. $\mu_{\pi^+,\pi^-}(\Omega) = 1$) but that is not increasing w.r.t. union in Ω. Aggregating x w.r.t. μ_{π^+,π^-} requires the use of an extension of the Choquet integral as proposed in [42]. This extension is relevant in the context of signal processing but insufficient to deal with any application since only kernels summing to one can be represented. For example, in image processing, kernels summing to 0 are used to estimate the gradient of an image, which is one of the fundamental building blocks in image processing. The gradient of an image can be used for edge detection, interest points localization, image editing, seamless image stitching, etc.

What we propose in this article is a new way for representing a convex set of linear kernel-based aggregations that can work with any kind of kernels.

3. Generalizing the maxitive domination

In this section, we propose to extend the work of Loquin et al. to any kernel-based linear application.

We define an operator as being a concave kernel-based set function $\mu_\varphi : 2^\Omega \to \mathbb{R}$, with $\varphi \in \mathcal{K}(\Omega)$ ($\mu_\varphi(\varnothing) = 0$) where, $\forall A \in \Omega$, the value of $\mu_\varphi(A)$ only depends on the N values of the kernel φ. As for capacities, we can associate to μ_φ a complementary operator μ^c_φ: $\mu^c_\varphi(A) = \mu_\varphi(\Omega) - \mu_\varphi(A^c)$, A^c being the complementary set of A in Ω [30].

The additive set function presented in Section 2.2 is a good example of an operator: a set function P_φ defined by: $\forall A \subseteq \Omega$, $P_\varphi(A) = \sum_{k\in A} \varphi_k$ can be defined for any kernel $\varphi \in \mathcal{K}(\Omega)$. Moreover, because P is additive, we have $P^c_\varphi = P_\varphi$.

The maxitive set function proposed in [26] does not comply with our definition of an operator since it is only defined for maxitive kernels.

3.1. The macsum operator

What we propose here is an extension of the work of [39] i.e. define a new set function that can be associated to any kernel of $\varphi \in \mathcal{K}(\Omega)$. The aim of this extension is to represent a convex set of kernels to account for imprecise knowledge in a linear aggregation. The extension we propose under the name of *macsum operator* is an operator denoted ν_φ, defined by:

$$\forall A \subseteq \Omega,\ \nu_\varphi(A) = \max_{i\in A} \varphi_i^+ + \min_{i\in\Omega} \varphi_i^- - \min_{i\in A^c} \varphi_i^-, \tag{3}$$

where A^c is the complementary set of A in Ω, $\varphi^+ = \max(0, \varphi)$ and $\varphi^- = \min(0, \varphi)$.

As a consequence: $\nu_\varphi(\Omega) = \max_{i\in\Omega} \varphi_i^+ + \min_{i\in\Omega} \varphi_i^- = \overline{\alpha} + \underline{\alpha}$,
with $\overline{\alpha} = \max_{i\in\Omega} \varphi_i^+$ and $\underline{\alpha} = \min_{i\in\Omega} \varphi_i^-$.

The term *macsum* comes from the fact that ν_φ can be expressed as a sum of two maxitive operators: $\nu_\varphi(A) = \max_{i\in A} \varphi_i^+ + \underline{\alpha} - \min_{i\in A^c} \varphi_i^- = \max_{i\in A} \varphi_i^+ + \max_{i\in A^c} \left(\underline{\alpha} - \varphi_i^-\right)$.

The complementary set function of ν_φ, denoted ν^c_φ, is given by:

$$\begin{aligned} \forall A \subseteq \Omega,\ \ \nu^c_\varphi(A) &= \nu_\varphi(\Omega) - \nu_\varphi(A^c) \\ &= \overline{\alpha} + \underline{\alpha} - \left(\max_{i\in A^c} \varphi_i^+ - \min_{i\in A} \varphi_i^- + \underline{\alpha} \right), \\ &= \min_{i\in A} \varphi_i^- + \overline{\alpha} - \max_{i\in A^c} \varphi_i^+ = \min_{i\in A} \varphi_i^- + \min_{i\in A^c} \left(\overline{\alpha} - \varphi_i^+\right). \end{aligned} \tag{4}$$

Thus $\nu_\varphi(\Omega) = \nu^c_\varphi(\Omega) = \underline{\alpha} + \overline{\alpha}$. Moreover, $\nu_\varphi(\varnothing) = \max_{i\in\varnothing} \varphi_i^+ - \min_{i\in\Omega} \varphi_i^- + \min_{i\in\Omega} \varphi_i = 0$ and $\nu^c_\varphi(\varnothing) = \overline{\alpha} - \max_{i\in\Omega} \varphi_i^+ + \min_{i\in\varnothing} \varphi_i^- = 0$.

Lemma 3.1. *$\forall A, B \subseteq \Omega$, $\max_{i\in A\cup B} \varphi_i + \max_{i\in A\cap B} \varphi_i \leq \max_{i\in A} \varphi_i + \max_{i\in B} \varphi_i$.*

Proof. Let $\tau = \max_{i\in A\cup B}\varphi_i + \max_{i\in A\cap B}\varphi_i - \max_{i\in A}\varphi_i - \max_{i\in B}\varphi_i$, and let us prove $\tau \leq 0$.
Let $a = \max_{i\in A}\varphi_i$, $b = \max_{i\in B}\varphi_i$ and $c = \max_{i\in A\cap B}\varphi_i$.
By construction $\max_{i\in A\cup B}\varphi_i = \max(a,b)$ and $c \leq a, b$.
We can consider without any loss of generality that $a \leq b$, therefore $\max_{i\in A\cup B}\varphi_i = b$. Thus $c \leq a \leq b$ which implies $\tau = b + c - a - b = c - a \leq 0$. □

Lemma 3.2. $\forall A, B \subseteq \Omega$, $\min_{i\in A\cup B}\varphi_i + \min_{i\in A\cap B}\varphi_i \geq \min_{i\in A}\varphi_i + \min_{i\in B}\varphi_i$.

Proof. Let $\tau = \min_{i\in A\cup B}\varphi_i + \min_{i\in A\cap B}\varphi_i - \min_{i\in A}\varphi_i - \min_{i\in B}\varphi_i$, and let us prove $\tau \geq 0$.
Let $a = \min_{i\in A}\varphi_i$, $b = \min_{i\in B}\varphi_i$ and $c = \min_{i\in A\cap B}\varphi_i$.
By construction $\min_{i\in A\cup B}\varphi_i = \min(a,b)$ and $a, b \leq c$.
We can consider without any loss of generality that $a \leq b$, therefore $\min_{i\in A\cup B}\varphi_i = a$. Thus $a \leq b \leq c$, then $\tau = a + c - a - b = c - b \geq 0$. □

Proposition 3.3. *ν_φ is concave and thus ν_φ^c is convex.*

Proof. The proof is trivial considering Lemmas 3.1 and 3.2. Let $A, B \subseteq \Omega$.
Let $\tau = \nu_\varphi(A\cup B) + \nu_\varphi(A\cap B) - \nu_\varphi(A) + \nu_\varphi(B)$.
$\tau = \tau_1 - \tau_2$, with $\tau_1 = \max_{i\in A\cup B}\varphi_i^+ + \max_{i\in A\cap B}\varphi_i^+ - \max_{i\in A}\varphi_i^+ - \max_{i\in B}\varphi_i^+$, and $\tau_2 = \min_{i\in(A\cup B)^c}\varphi_i^- + \min_{i\in(A\cap B)^c}\varphi_i^- - \min_{i\in A^c}\varphi_i^- - \min_{i\in B^c}\varphi_i^-$.
$\tau_2 = \min_{i\in A^c\cap B^c}\varphi_i^- + \min_{i\in A^c\cup B^c}\varphi_i^- - \min_{i\in A^c}\varphi_i^- - \min_{i\in B^c}\varphi_i^-$.
Due to Lemma 3.1, $\tau_1 \leq 0$ and due to Lemma 3.2, $\tau_2 \geq 0$. Thus $\tau = \tau_1 - \tau_2 \leq 0$ and therefore ν_φ is concave. Proving ν_φ^c being convex can be done in the same way. □

Remark 1. *The maxitive operator defined in Expression* (3) *is an extension of the one proposed by Loquin et al. [26] since if $\pi \in \mathcal{K}_s$ is a maxitive kernel, then $\nu_\pi = \Pi_\pi$ is a maxitive aggregation fonction defined by: $\forall A \in \Omega$, $\nu_\pi(A) = \Pi_\pi(A) = \max_{i\in A}\pi_i$.*

Proposition 3.4. *ν_φ is a weak-maxitive set function*
i.e. $\forall A, B \subseteq \Omega$, $\nu_\varphi(A\cup B) \geq \max(\nu_\varphi(A), \nu_\varphi(B))$.

Proof. Let $A, B \subseteq \Omega$,

$$\begin{aligned}
\nu_\varphi(A\cup B) &= \max_{i\in A\cup B}\varphi_i^+ + \min_{i\in\Omega}\varphi_i^- - \min_{i\in(A\cup B)^c}\varphi_i^- \\
&= \max(\max_{i\in A}\varphi_i^+, \max_{i\in B}\varphi_i^+) + \min_{i\in\Omega}\varphi_i^- - \min_{i\in(A\cup B)^c}\varphi_i^- \\
&\geq \max(\max_{i\in A}\varphi_i^+, \max_{i\in B}\varphi_i^+) + \min_{i\in\Omega}\varphi_i^- - \min(\min_{i\in A^c}\varphi_i^-, \min_{i\in B^c}\varphi_i^-) \\
&\geq \max(\nu_\varphi(A), \nu_\varphi(B))
\end{aligned}$$

□

3.2. Domination of the macsum operator over additive operators

First, let us define $\mathcal{M}(\varphi)$ as being the core of a kernel $\varphi \in \mathcal{K}(\Omega)$ by extending the notion of core of a maxitive kernel proposed by Loquin et al. [26]:

$$\mathcal{M}(\varphi) = \left\{\psi \in \mathcal{K}(\Omega) \,/\, \forall A \subseteq \Omega, \nu_\varphi^c(A) \leq P_\psi(A) \leq \nu_\varphi(A)\right\}. \tag{5}$$

Remark 2. *This definition coincides with the one of [26] when φ is a maxitive kernel.*

Remark 3. *To be inline with the work of [26], let $\psi, \varphi \in \mathcal{K}(\Omega)$, if $\psi \in \mathcal{M}(\varphi)$ we say that φ **dominates** ψ because the macsum operator based on φ dominates the additive operator base on ψ. ψ belongs to the convex set represented by φ.*

Remark 4. *Since $\nu_\varphi^c(\Omega) = \nu_\varphi(\Omega)$, if $\psi \in \mathcal{M}(\varphi)$,*
then $P_\psi(\Omega) = \sum_{i\in\Omega} \psi_i = \max_{i\in\Omega} \max(0, \varphi_i) + \min_{i\in\Omega} \min(0, \varphi_i)$.

Now, two questions of instrumental interest arise, especially in the context of signal processing:

- let $\varphi \in \mathcal{K}(\Omega)$ be a kernel of Ω, is there a simple way to check whether a kernel $\psi \in \mathcal{K}(\Omega)$ belongs or not to $\mathcal{M}(\varphi)$?
- let $\psi \in \mathcal{K}(\Omega)$ be a kernel of Ω, is there a simple way to define a kernel $\varphi \in \mathcal{K}(\Omega)$ such that $\psi \in \mathcal{M}(\varphi)$?

In [26], Loquin et al. have used some known properties defined in the context of possibility theory [11] to partially answer to those questions by defining particular relations between maxitive and summative kernels. Let us briefly recall some useful Lemma in this paper.

Lemma 3.5. *(weak domination) Let $\rho \in \mathcal{K}_s(\Omega)$ be a summative kernel of Ω, then the maxitive kernel $\breve{\pi} \in \mathcal{K}_m(\Omega)$, defined by $\forall i \in \Omega$, $\breve{\pi}_i = \sum_{j\in\Omega} \min(\rho_i, \rho_j)$, dominates ρ, i.e. $\rho \in \mathcal{M}(\breve{\pi})$.*

Lemma 3.6. *(strong domination) Let $\rho \in \mathcal{K}_s(\Omega)$ be a summative kernel of Ω, then the maxitive kernel $\mathring{\pi} \in \mathcal{K}_m(\Omega)$, defined by $\forall i \in \Omega$, $\mathring{\pi}_i = \sum_{j\in A_i} \rho_j$, where $A_i = \{j \in \Omega / \rho_j \leq \rho_i\}$, dominates ρ i.e. $\rho \in \mathcal{M}(\mathring{\pi})$.*

As shown in [26], $\breve{\pi}$ is said to weaker dominate ρ than $\mathring{\pi}$ since $\mathcal{M}(\mathring{\pi}) \subseteq \mathcal{M}(\breve{\pi})$.

What we propose here is to use Lemma 3.5 and 3.6 to define also weak and strong domination relations between two kernels of $\mathcal{K}(\Omega)$.

Proposition 3.7. *(general weak domination) Let $\psi \in \mathcal{K}(\Omega)$ be a kernel of Ω, then the kernel $\breve{\varphi} \in \mathcal{K}(\Omega)$ defined by $\forall i \in \Omega$, $\breve{\varphi}_i = \sum_{j\in\Omega} \min(\psi_i^+, \psi_j^+) + \sum_{j\in\Omega} \max(\psi_i^-, \psi_j^-)$, where $\psi^+ = \max(0, \psi)$ and $\psi^- = \min(0, \psi)$ dominates ψ i.e. $\psi \in \mathcal{M}(\breve{\varphi})$.*

Proof. Let us define $\alpha^+ = \sum_{i\in\Omega} \psi_i^+$ and $\alpha^- = \sum_{i\in\Omega} \psi_i^-$, then the kernels ρ^+ and ρ^- defined by: $\forall i \in \Omega$, $\rho_i^+ = \psi_i^+/\alpha^+$ and $\rho_i^- = \psi_i^-/\alpha^-$ are summative by construction (i.e. positive and normalized). Thus according to Lemma 3.5 we can define two maxitive kernels $\breve{\pi}^+$ and $\breve{\pi}^-$ by:

$$\forall i \in \Omega,\ \breve{\pi}_i^+ = \sum_{j\in\Omega} \min(\rho_i^+, \rho_j^+) \text{ and } \breve{\pi}_i^- = \sum_{j\in\Omega} \max(\rho_i^-, \rho_j^-).$$

Let us define two kernels $\breve{\varphi}^+$ and $\breve{\varphi}^-$, such that

$$\forall i \in \Omega,\ \breve{\varphi}_i^+ = \alpha^+.\breve{\pi}_i^+ = \sum_{j\in\Omega} \min(\psi_i^+, \psi_j^+) \text{ and}$$
$$\breve{\varphi}_i^- = \alpha^-.\breve{\pi}_i^- = \sum_{j\in\Omega} \max(\psi_i^-, \psi_j^-).$$

By construction $\forall i \in \Omega,\ \breve{\varphi}_i^+ \geq 0$ and $\breve{\varphi}_i^- \leq 0$. Let us define $\breve{\varphi} = \breve{\varphi}^+ + \breve{\varphi}^-$.

According to Lemma 3.5 we have, $\forall A \subseteq \Omega,\ \Pi_{\breve{\pi}^+}(A) = \max_{i\in A} \breve{\pi}_i^+ \geq P_{\rho^+}(A) = \sum_{i\in A} \rho_i^+$ and $\Pi_{\breve{\pi}^-}(A) = \max_{i\in A} \breve{\pi}_i^- \geq P_{\rho^-}(A) = \sum_{i\in A} \rho_i^-$.

Let us consider the two set functions $\mu_{\breve{\varphi}^+}$ and $\mu_{\breve{\varphi}^-}$, defined by: $\forall A \subseteq \Omega$, $\mu_{\breve{\varphi}^+}(A) = \max_{i\in A} \breve{\varphi}_i^+$ and $\mu_{\breve{\varphi}^-}(A) = \min_{i\in A} \breve{\varphi}_i^-$.

We have, $\forall A \subseteq \Omega$, $\mu_{\breve{\varphi}^+}(A) = \max_{i\in A} \breve{\varphi}_i^+ = \alpha^+ \max_{i\in A} \breve{\pi}_i^+ \geq \alpha^+.P_{\rho^+}(A) = \sum_{i\in A} \alpha^+.\rho_i^+ = P_{\psi^+}(A)$ and $\mu_{\breve{\varphi}^-}(A) = \min_{i\in A} \breve{\varphi}_i^- = \alpha^- \max_{i\in A} \breve{\pi}_i^- \leq \alpha^-.P_{\rho^-}(A) = \sum_{i\in A} \alpha^-.\rho_i^- = P_{\psi^-}(A)$. Remarking that $\min_{i\in\Omega} \breve{\varphi}_i^- = \sum_{i\in\Omega} \psi_i^-$, we also have $\mu^c_{\breve{\varphi}^-}(A) = \min_{i\in\Omega} \breve{\varphi}_i^- - \min_{i\in A^c} \breve{\varphi}_i^- \geq \sum_{i\in\Omega} \psi_i^- - \sum_{i\in A^c} \psi_i^- = \sum_{i\in A} \psi_i^- = P_{\psi_i^-}(A)$. Thus, $\forall A \subseteq \Omega$, $\nu_{\breve{\varphi}}(A) = \max_{i\in A} \breve{\varphi}_i^+ + \min_{i\in\Omega} \breve{\varphi}_i^- - \min_{i\in A^c} \breve{\varphi}_i^- = \mu_{\breve{\varphi}^+}(A) + \mu^c_{\breve{\varphi}^-}(A)$, and therefore $\nu_{\breve{\varphi}}(A) \geq P_{\psi^+}(A) + P_{\psi^-}(A) = P_\psi(A)$. □

Proposition 3.8. *(general strong domination) Let $\psi \in \mathcal{K}(\Omega)$ be a kernel of Ω, then the kernel $\mathring{\varphi} \in \mathcal{K}(\Omega)$ defined by $\forall i \in \Omega$, $\mathring{\varphi}_i = \sum_{j\in A_i^+} \psi_j^+ + \sum_{j\in A_i^-} \psi_j^-$, where $\psi^+ = \max(0, \psi)$, $\psi^- = \min(0, \psi)$, $A_i^+ = \left\{j \in \Omega \,/\, \psi_j^+ \leq \psi_i^+\right\}$ and $A_i^- = \left\{j \in \Omega \,/\, \psi_j^- \geq \psi_i^-\right\}$ dominates ψ i.e. $\psi \in \mathcal{M}(\mathring{\varphi})$.*

The proof of Proposition 3.8 follows the same pattern as the proof of Proposition 3.7.

Proof. As previously, we define the summative kernels ρ^+ and ρ^- by: $\forall i \in \Omega$, $\rho_i^+ = \psi_i^+/\alpha^+$ and $\rho_i^- = \psi_i^-/\alpha^-$, with $\alpha^+ = \sum_{i\in\Omega} \psi_i^+$ and $\alpha^- = \sum_{i\in\Omega} \psi_i^-$.

As proposed in Lemma 3.6 we can define two maxitive kernels $\mathring{\pi}^+$ and $\mathring{\pi}^-$ by: $\forall i \in \Omega$, $\mathring{\pi}_i^+ = \sum_{j\in A_i^+} \rho_j^+$, $\mathring{\pi}_i^- = \sum_{j\in A_i^-} \rho_j^-$, with $A_i^+ = \left\{j \in \Omega / \rho_j^+ \leq \rho_i^+\right\} = \left\{j \in \Omega \,/\, \psi_j^+ \leq \psi_i^+\right\}$ and $A_i^- = \left\{j \in \Omega \,/\, \rho_j^- \leq \rho_i^-\right\} = \left\{j \in \Omega \,/\, \psi_j^- \geq \psi_i^-\right\}$. According to Lemma 3.5 we have, $\forall A \subseteq \Omega$, $\Pi_{\mathring{\pi}^+}(A) = \max_{i\in A} \mathring{\pi}_i^+ \geq P_{\rho^+}(A) = \sum_{i\in A} \rho_i^+$ and $\Pi_{\mathring{\pi}^-}(A) = \max_{i\in A} \mathring{\pi}_i^- \geq P_{\rho^-}(A) = \sum_{i\in A} \rho_i^-$. We also have $\Pi^c_{\mathring{\pi}^-}(A) = 1 - \max_{i\in A^c} \mathring{\pi}_i^- \leq P_{\rho^-}(A)$.

Let $\mathring{\varphi}^+ = \alpha^+.\pi^+ = \sum_{j\in A_i^+} \psi_j^+$ and $\mathring{\varphi}^- = \alpha^-.\pi^- = \sum_{j\in A_i^-} \psi_j^-$.

We define $\mathring{\varphi} = \mathring{\varphi}^+ + \mathring{\varphi}^-$. We have $\forall A \subseteq \Omega$, $\nu_{\mathring{\varphi}}(A) = \mu_{\mathring{\varphi}^+}(A) + \mu^c_{\mathring{\varphi}^+}(A)$, with $\mu_{\mathring{\varphi}^+}(A) = \max_{i\in A} \mathring{\varphi}_i^+ = \alpha^+.\Pi_{\mathring{\pi}^+}(A)$ and $\mu_{\mathring{\varphi}^-}(A) = \min_{i\in A} \mathring{\varphi}_i^- = \alpha^-.\Pi_{\mathring{\pi}^-}(A)$. Thus $\mathring{\varphi}(A) = \alpha^+.\Pi_{\mathring{\pi}^+}(A) + \alpha^-.\Pi^c_{\mathring{\pi}^-}(A) \geq \alpha^+.P_{\rho^+}(A) + \alpha^-.P_{\rho^-}(A) = P_{\psi^+}(A) + P_{\psi^-}(A) = P_\psi(A)$. □

Finally, we have the following property:

Proposition 3.9. *$\forall \varphi \in \mathcal{K}(\Omega)$, $\mathcal{M}(\varphi)$ is not empty.*

Proof. To prove Proposition 3.9, it is enough to be able to associate to each kernel $\varphi \in \mathcal{K}(\Omega)$ a kernel $\psi \in \mathcal{K}(\Omega)$ such that $\psi \in \mathcal{M}(\varphi)$. This proof is based

on constructing ψ such that φ strongly dominates ψ as suggested in Proposition 3.8.

If φ is the kernel that strongly dominates ψ then: $\varphi_i = \sum_{j\in A_i^+} \psi_j^+ + \sum_{j\in A_i^-} \psi_j^-$, where $\psi^+ = \max(0,\psi)$, $\psi^- = \min(0,\psi)$, $A_i^+ = \left\{j \in \Omega \,/\, \psi_j^+ \leq \psi_i^+\right\}$ and $A_i^- = \left\{j \in \Omega \,/\, \psi_j^- \geq \psi_i^-\right\}$.

We can suppose, without any loss of generality, that φ is sorted in increasing order: $\varphi_1 \leq \varphi_2 \leq \cdots \leq \varphi_N$. Let $\varphi^+ = \max(0,\varphi)$ and $\varphi^- = \min(0,\varphi)$ – i.e. $\varphi = \varphi^+ + \varphi^-$. Then, if φ strongly dominates ψ, $\varphi_i^+ = \sum_{j=1}^{i} \psi_i^+$ and $\varphi_i^- = \sum_{j=i}^{N} \psi_i^-$.

Based on this, we can built iteratively the values of ψ^+ and $\psi-$:

$$\forall i \in \{2,\dots N\},\ \ \psi_i^+ = \varphi_i^+ - \sum_{j=1}^{i-1} \psi_j^+ = \varphi_i^+ - \varphi_{i-1}^+,\ \text{with } \psi_1^+ = \varphi_1^+,\ \text{and}$$

$$\forall i \in \{N-1,\dots 1\},\ \ \psi_i^- = \varphi_i^- - \sum_{j=i+1}^{N} \psi_j^- = \varphi_i^- - \varphi_{i+1}^-,\ \text{with } \psi_N^- = \varphi_N^-.$$

This guarantees that the so-constructed $\psi = \psi^+ + \psi^-$ belongs to $\mathcal{M}(\varphi)$ and thus $\mathcal{M}(\varphi) \neq \varnothing$. □

4. Representing a convex set of linear aggregations

In this section, we define two operator-based aggregations of the function $x \in \mathbb{R}^N$ w.r.t. a kernel $\varphi \in \mathcal{K}(\Omega)$: the linear aggregation denoted $F_\varphi(x)$ and the macsum aggregation denoted $\underline{F}_\varphi(x)$.

The linear aggregation is simply a weighted sum defined by:

$$F_\varphi(x) = \sum_{i\in\Omega} \varphi_i.x_i = \check{\mathbb{C}}_{P_\varphi}(x). \tag{6}$$

The macsum aggregation is defined by:

$$\overline{\underline{F}}_\varphi(x) = \left[\underline{F}_\varphi(x), \overline{F}_\varphi(x)\right] = \left[\check{\mathbb{C}}_{\nu_\varphi^c}(x), \check{\mathbb{C}}_{\nu_\varphi}(x)\right]. \tag{7}$$

This extend the work of Loquin et al. since:

Proposition 4.1. *Let $\varphi \in \mathcal{K}(\Omega)$, $\forall \psi \in \mathcal{M}(\varphi)$, $\forall x \in \mathbb{R}^N$, $F_\psi(x) \in \overline{\underline{F}}_\varphi(x)$. Moreover, $\forall y \in \overline{\underline{F}}_\varphi(x)$, $\exists \psi$ such that $y = F_\psi(x)$.*

Proof. The fact that $\forall \psi \in \mathcal{M}(\varphi)$, $\forall x \in \mathbb{R}^N$, $F_\psi(x) \in \overline{\underline{F}}_\varphi(x)$ is simply a reformulation of the Schmeidler-Denneberg theorem [9] in a more particular case. Indeed, whereas this theorem has been proven for any concave set function, we here use it only for kernel-based set functions. Now lets prove that $\forall y \in \overline{\underline{F}}_\varphi(x)$, $\exists \psi$ such that $y = F_\psi(x)$. We have $\check{\mathbb{C}}_{\nu_\varphi}(x) = \sum_{k=1}^{N} x_{\lceil k\rceil}.(\nu_\varphi(A_{\lceil k\rceil}) - \nu_\varphi(A_{\lceil k+1\rceil}))$ and $\check{\mathbb{C}}_{\nu_\varphi^c}(x) = \sum_{k=1}^{N} x_{\lceil k\rceil}.(\nu_\varphi^c(A_{\lceil k\rceil}) - \nu_\varphi^c(A_{\lceil k+1\rceil}))$ as $\overline{\underline{F}}_\varphi(x) = \left[\underline{F}_\varphi(x), \overline{F}_\varphi(x)\right] =$

$\left[\check{\mathbb{C}}_{\nu_\varphi^c}(x), \check{\mathbb{C}}_{\nu_\varphi}(x)\right]$ is a convex set, and $y \in \overline{F}_\varphi(x)$ we then have $y = \lambda \underline{F}_\varphi(x) + (1-\lambda)\overline{F}_\varphi(x)$ with $\lambda \in [0,1]$ therefore $y = \lambda \check{\mathbb{C}}_{\nu_\varphi^c}(x) + (1-\lambda)\check{\mathbb{C}}_{\nu_\varphi}(x)$
and $y = \lambda \sum_{k=1}^N x_{\lceil k \rceil}.(\nu_\varphi^c(A_{\lceil k \rceil}) - \nu_\varphi^c(A_{\lceil k+1 \rceil})) + (1-\lambda)\sum_{k=1}^N x_{\lceil k \rceil}.(\nu_\varphi(A_{\lceil k \rceil}) - \nu_\varphi(A_{\lceil k+1 \rceil}))$ which gives that $y = \sum_{k=1}^N x_{\lceil k \rceil}.(\lambda((\nu_\varphi^c(A_{\lceil k \rceil}) - \nu_\varphi^c(A_{\lceil k+1 \rceil})) + (1-\lambda)((\nu_\varphi(A_{\lceil k \rceil}) - \nu_\varphi(A_{\lceil k+1 \rceil}))))$.
Thus we have that $\psi = \{\psi_{\lceil k \rceil}\}_{k \in \Omega}$ is a kernel such that
$\psi_{\lceil k \rceil} = \lambda((\nu_\varphi^c(A_{\lceil k \rceil}) - \nu_\varphi^c(A_{\lceil k+1 \rceil})) + (1-\lambda)((\nu_\varphi(A_{\lceil k \rceil}) - \nu_\varphi(A_{\lceil k+1 \rceil}))$
and finally $y = \sum_{k=1}^N \psi_{\lceil k \rceil}.x_{\lceil k \rceil} = \boldsymbol{F}_\psi(x)$ as the sum is commutative. □

5. Computing the macsum aggregation

This section aims at proposing a simple algorithm for computing $\overline{y} = [\underline{y}, \overline{y}] = \overline{\underline{F}}[x]$.

5.1. Formulæ

Let us first consider $\overline{y}$.

$\overline{y} = \check{\mathbb{C}}_{\nu_\varphi}(x) = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\nu_\varphi(A_{\lceil k \rceil})$, with $x_{\lceil 0 \rceil} = 0$.

Remember that $\nu_\varphi(A_{\lceil k \rceil}) = \max_{i \in A_{\lceil k \rceil}} \varphi_i^+ - \min_{i \in A^c_{\lceil k \rceil}} \varphi_i^- + \min_{i \in \Omega} \varphi_i^-$, with $A_{\lceil k \rceil} = \{\lceil k \rceil, \dots \lceil N \rceil\}$ and $A^c_{\lceil k \rceil} = \{\lceil 1 \rceil, \dots \lceil k-1 \rceil\}$ thus $\nu_\varphi(A_{\lceil k \rceil}) = \max_{i=k}^N \varphi_{\lceil i \rceil}^+ - \min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^- + \min_{i \in \Omega} \varphi_i^-$
(by convention $\min_{i=1}^0 \varphi_{\lceil i \rceil}^- = 0$).
Therefore $\check{\mathbb{C}}_{\nu_\varphi}(x) = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\left(\max_{i=k}^N \varphi_{\lceil i \rceil}^+ - \min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^- + \min_{i \in \Omega} \varphi_i^-\right)$.

Let us decompose this sum:
$\check{\mathbb{C}}_{\nu_\varphi}(x) = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\max_{i=k}^N \varphi_{\lceil i \rceil}^+ - \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^-$
$+ \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\min_{i \in \Omega} \varphi_i^-$,
Note that $\sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\min_{i \in \Omega} \varphi_i^- = \min_{i \in \Omega} \varphi_i^-.\sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) = x_{\lceil N \rceil}.\min_{i \in \Omega} \varphi_i^-$.

Thus, $\overline{y} = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\max_{i=k}^N \varphi_{\lceil i \rceil}^+ - \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^- + x_{\lceil N \rceil}.\min_{i \in \Omega} \varphi_i^-$.

Let $\overline{y}^+ = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\max_{i=k}^N \varphi_{\lceil i \rceil}^+$.

and $\overline{y}^- = x_{\lceil N \rceil}.\min_{i \in \Omega} \varphi_i^- - \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}).\min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^-$

First, let us consider $\overline{y}^+$, and $\lfloor . \rfloor$ the permutation that sorts the x in decreasing order:
$x_{\lceil k \rceil} = x_{\lfloor N-k+1 \rfloor}$. Thus let $k' = N - k + 1 \to k = N - k' + 1$.
$\overline{y}^+ = \sum_{k'=1}^N \left(x_{\lfloor k' \rfloor} - x_{\lfloor k'+1 \rfloor}\right).\max_{i=N-k'+1}^N \varphi_{\lfloor i \rfloor}^+$

Now, let $\alpha_k = \max_{i=N-k+1}^N \varphi_{\lfloor i \rfloor}^+$, then, by construction, $\alpha_{k+1} = \max(\alpha_k, \varphi_{\lfloor k+1 \rfloor}^+)$ and $\alpha_0 = 0$. Thus $\overline{y}^+ = \sum_{k=1}^N \left(x_{\lfloor k \rfloor} - x_{\lfloor k+1 \rfloor}\right).\alpha_k = \sum_{k=1}^N x_{\lfloor k \rfloor}.(\alpha_k - \alpha_{k-1})$.

Le us now rearrange the expression of $\overline{y}^-$.

$\overline{y}^- = x_{\lceil N \rceil}.\min_{i=1}^{N} \varphi^-_{\lceil i \rceil} - x_{\lceil N \rceil}.\min_{i=1}^{N-1} \varphi^-_{\lceil i \rceil} + x_{\lceil N-1 \rceil}.\min_{i=1}^{N-1} \varphi^-_{\lceil i \rceil} - \cdots + x_{\lceil 1 \rceil}.\min_{i=1}^{1} \varphi^-_{\lceil i \rceil}$.

$\overline{y}^- = \sum_{k=1}^{N} x_{\lceil k \rceil}.\left(\min_{i=1}^{k} \varphi^-_{\lceil i \rceil} - \min_{i=1}^{k-1} \varphi^-_{\lceil i \rceil}\right)$.

Let $\beta_k = \min_{i=1}^{k} \varphi^-_{\lceil i \rceil}$, by construction $\beta_{k+1} = \min(\beta_k, \varphi^-_{\lceil k+1 \rceil})$

Therefore $\overline{y}^- = \sum_{k=1}^{N} x_{\lceil k \rceil}.(\beta_k - \beta_{k-1}) = \sum_{k=1}^{N} \left(x_{\lceil k \rceil} - x_{\lceil k+1 \rceil}\right).\beta_k$.

The complete algorithm is reported in Algorithm 1.

Computation of $\underline{y}$ can be obtained in the same way by simply remembering that $\underline{y} = \check{\mathbb{C}}_{\nu^c_\varphi}(x) = -\check{\mathbb{C}}_{\nu_\varphi}(-x)$. Thus the algorithms are the same, we have simply to exchange the permutation that sorts the x in decreasing order by the permutation that sorts the x in increasing order and vice versa.

The complete algorithm is reported in Algorithm 2.

5.2. Algorithms

Computing the upper and lower values of the interval-valued output of an maxitive operator based aggregation is provided by the Algorithm 2 to compute $\underline{y}$ and Algorithm 1 to compute $\overline{y}$.

Algorithm 1: Computation of $\overline{y}$

Input: $\mathbf{x} = \{x_i\}_{i=1\ldots N}$, $\boldsymbol{\varphi} = \{\varphi_i\}_{i=1\ldots N}$
Output: $\overline{y}$
sort ($\mathbf{x}$, $\boldsymbol{\varphi}$) w.r.t. $\mathbf{x}$ in decreasing order ;
$\alpha = 0$, $\beta = 0$;
$\overline{y} = 0$;
for $k = 1 \ldots N$ **do**
  $\beta = \alpha$;
  $\alpha = \max(\alpha, \varphi_k)$;
  $\overline{y} = \overline{y} + (\alpha - \beta).x_k$;
sort ($\mathbf{x}$, $\boldsymbol{\varphi}$) w.r.t. $\mathbf{x}$ in increasing order (i.e. reverse the sorting) ;
$\alpha = 0$, $\beta = 0$;
for $k = 1 \ldots N$ **do**
  $\beta = \alpha$;
  $\alpha = \min(\alpha, \varphi_k)$;
  $\overline{y} = \overline{y} + (\alpha - \beta).x_k$;

6. Experiment

In this section, we propose an experiment that consists in deriving a digital signal. A digital signal is generally a continuous signal that has been sampled and quantified. Being sampled, a digital signal is not continuous and thus

Algorithm 2: Computation of $\underline{y}$

Input: $\mathbf{x} = \{x_i\}_{i=1\dots N}$, $\boldsymbol{\varphi} = \{\varphi_i\}_{i=1\dots N}$
Output: $\underline{y}$
sort $(\mathbf{x}, \boldsymbol{\varphi})$ w.r.t. $\mathbf{x}$ in increasing order ;
$\alpha = 0$, $\beta = 0$;
$\underline{y} = 0$;
for $k = 1 \dots N$ **do**
  $\beta = \alpha$;
  $\alpha = \max(\alpha, \varphi_k)$;
  $\underline{y} = \underline{y} + (\alpha - \beta).x_k$;
sort $(\mathbf{x}, \boldsymbol{\varphi})$ w.r.t. $\mathbf{x}$ in decreasing order (i.e. reverse the sorting) ;
$\alpha = 0$, $\beta = 0$;
for $k = 1 \dots N$ **do**
  $\beta = \alpha$;
  $\alpha = \min(\alpha, \varphi_k)$;
  $\underline{y} = \underline{y} + (\alpha - \beta).x_k$;

cannot be derived. Therefore, deriving a digital signal refers to as estimating the sampled derivative of the original continuous signal.

The estimation of this derivative can be seen as the succession of three steps: first reconstructing a continuous signal based on the samples of the digital signal, then deriving this continuous signal and finally sampling the result of this derivation.

As explained in [37] in the context of image processing, this three steps procedure can be achieved in one step by convoluting the input digital signal by the sampled derivative of a reconstruction signal. In their paper, the authors propose the exponential filter for being a good candidate since it can be derived and, as a reconstruction operator, it lows the effect of noise and quantization on the obtained continuous signal.

The Shen-Castan derivating kernel is obtained by deriving then sampling the reconstruction kernel κ defined by:

$$\forall x \in \mathbb{R}, \quad \kappa(x) = -\frac{2}{\ln(\beta)}.\beta^{|x|}, \tag{8}$$

with $\beta \in]0, 1[$ being a smoothing factor: the bigger β the smoother the output reconstructed signal.

Sampling this reconstruction kernel leads to a discrete (summative) kernel ρ defined by:

$$\forall k \in \mathbb{Z}, \quad \rho_k = \frac{1-\beta}{1+\beta}.\beta^{|k|}, \tag{9}$$

this kernel is summative in the sense of Loquin [28] since $\sum_{k\in\mathbb{Z}} \rho_k = 1$.

Deriving the kernel κ leads to the continuous kernel η defined by:

$$\forall x \in \mathbb{R}, \quad \eta(x) = 2.\mathrm{sign}(x).\beta^{|x|} = -\mathrm{sign}(x).\ln(\beta).\kappa(x). \tag{10}$$

And therefore, sampling this derivation leads to the derivation kernel ψ defined by:

$$\forall k \in \mathbb{Z},\ \ \psi_k = -\text{sign}(k).\ln(\beta).\rho_k = -\text{sign}(k).\ln(\beta).\frac{1-\beta}{1+\beta}.\beta^{|k|} = -\text{sign}(k).\gamma.\beta^{|k|}, \tag{11}$$

with $\gamma = \ln(\beta).\frac{1-\beta}{1+\beta}$.

Let us now consider computing y the discrete signal that is the derivative of x. This derivation can be obtained by convoluting x with ψ: $y = x \star \psi$, $\star$ being the convolution. This operation can be written [26]:

$$y_n = (x \star \psi)_n = \sum_{k=1}^{N} x_k.\psi_k^n = \check{\mathbb{C}}_{P_{\psi_k^n}}(x), \tag{12}$$

ψ^n being the kernel ψ translated in n and truncated on Ω defined by: $\psi_k^n = \psi_{n-k}$, $k \in \{1, \ldots, N\}$. Remark that this assumes that un-sampled values of the underlying unknown signal whose sampled values are x are null. It is interesting to note that $P_{\psi_k^n}(\Omega) = \beta^{n-N} - \beta^n \in [\beta^N - 1, 1 - \beta^N]$ can be positive, negative or even null.

Now, we can use the Proposition 3.8 to define, for each $n \in \Omega$, $\mathring{\varphi}^n$, a kernel that is the smallest that dominates ψ^n (i.e. $\psi^n \in \mathcal{M}(\mathring{\varphi}^n)$).

In most applications using signal derivation, one of the main difficulties is to define both shape and parameter of the kernel that is optimal for this application. Assuming, with the authors, that the Shen-Castan kernel is optimal for this application, the remaining problem is the smoothing parameter β.

First note that the granularity (in the sense of Loquin or Shannon) of the Shen-Castan kernel increases with β, i.e. the bigger β, the smoother the reconstructed continuous signal.

Therefore, let ψ and ψ' be two derivating kernels defined by Expression (11) using β and β', with $\beta' < \beta$. Let $\mathring{\varphi}$ and $\mathring{\varphi}'$ the two kernels that strongly dominate ψ and ψ'. Then, by construction, the set of derivatives computed by using $\mathring{\varphi}$ are the derivatives of smoother signals than the set of derivatives computed by using $\mathring{\varphi}'$.

Moreover, as shown in [29], the width of the interval-valued signal is representative of the conflict between the different kernels dominated in the derivative computation, and thus sensitive to noise. The imprecision of the interval-valued derivative can be used to define thresholds for detecting local maxima, like in [19] or use a posteriori criteria, other than frequency criteria, to define the optimal candidate to be the derivation of x (see e.g. [17]).

In this experiment, we propose to compute the derivative of a digital signal obtained by sampling (sampling frequency 100Hz) and quantizing the composite signal whose equation is: $x(t) = \sin(21.t) + \sin(12.t + 5) + \sin(-4.t + 3) + 10.(1 + t)^{-1}$, that is pictured in Figure (1).

In this experiment, we compute both precise and imprecise derivatives by using either the kernel ψ with Equation (12) or the kernel $\mathring{\varphi}$ with Equation (7).

This experiment illustrates two properties. First, the more appropriate the reconstruction kernel, the closer the precise derivative from the bounds of the

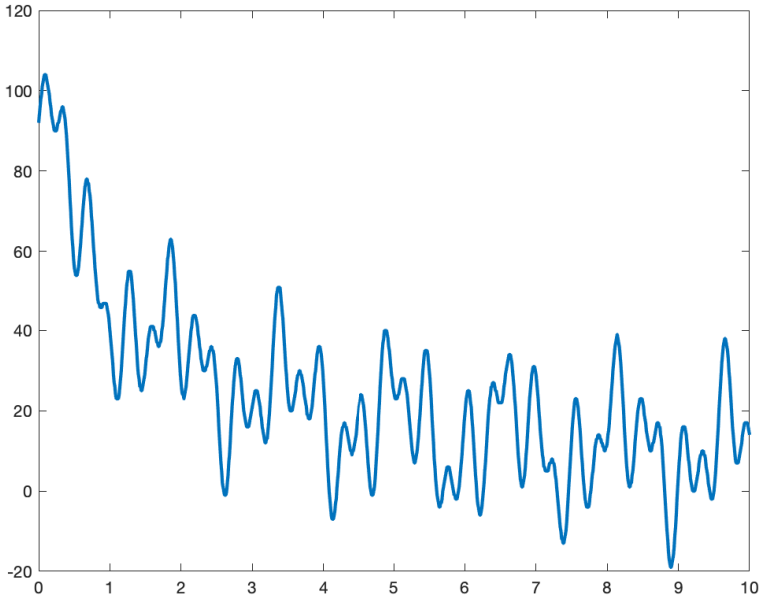


Figure 1: The signal to be derivated.

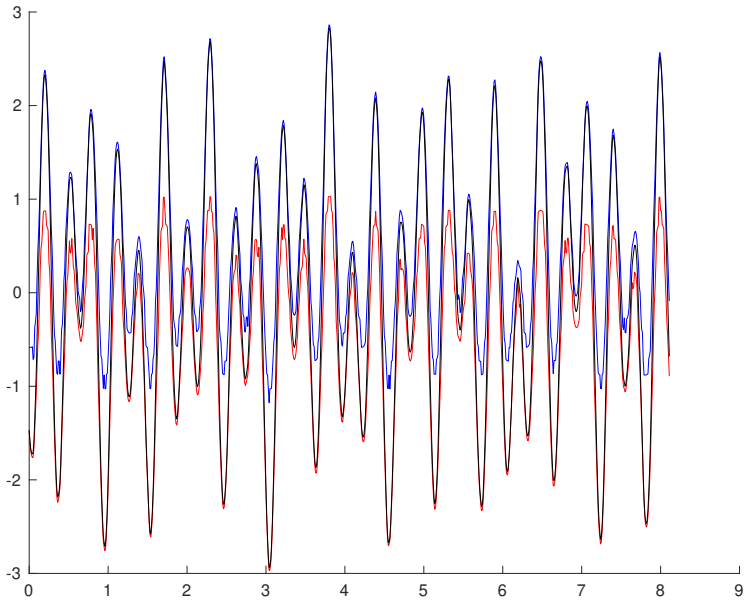


Figure 2: Precise derivative (black) and imprecise derivative (blue upper, red lower) of the signal depicted in Figure (1) with $\beta = 0.7$

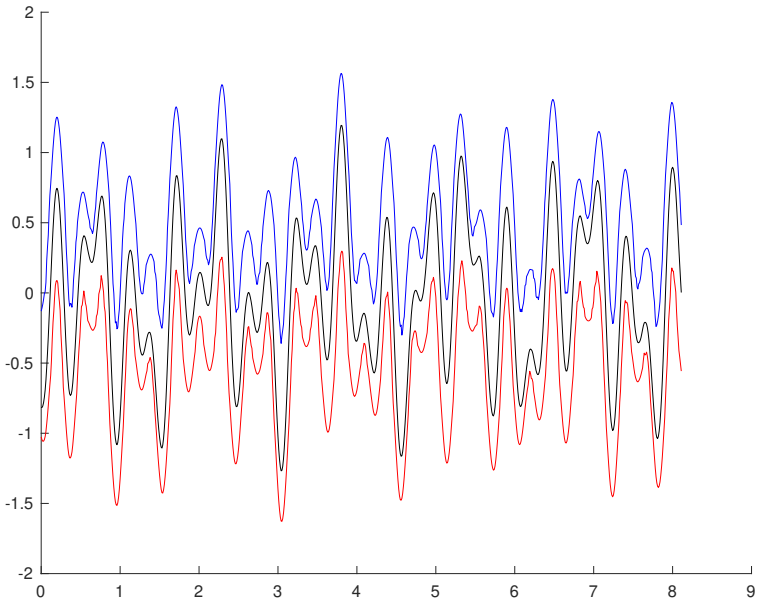
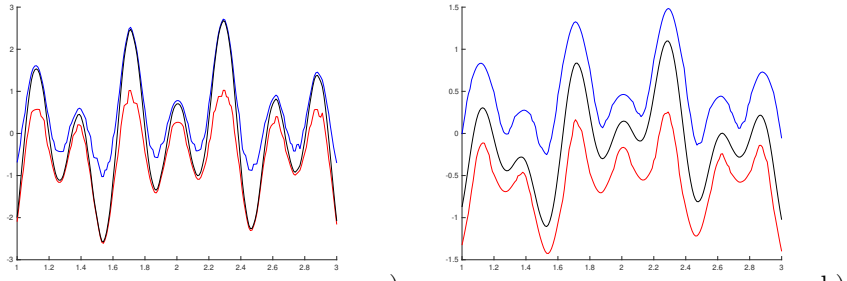


Figure 3: Precise derivative (black) and imprecise derivative (blue upper, red lower) of the signal depicted in Figure (1) with $\beta = 0.9$



a) b)

Figure 4: Details of Figure (2) (a) and Figure (3) (b)

imprecise derivative (see for example, in Figure (4.a) that the precise (black) curve is pretty close to the bounds near the extrema). Second, the less appropriate the kernel, the higher the imprecision of the imprecise derivative. This is obvious when comparing Figure (4.a) with Figure (4.b).

We also have computed precise derivatives by varying β in $[0.88, 0.96]$. Those derivatives are superimposed on the imprecise derivative with $\beta = 0.9$ on Figure (5.a) and zoomed in Figure (5.b). What we can observed is that the variation of β corresponds to a variation of the estimation of the derivative in the interval-valued derivative computed by using the macsum approach. Note, however, that when β varies outside of $[0.88, 0.96]$ some precise derivatives fall outside the interval-valued signal computed with $\beta = 0.9$.

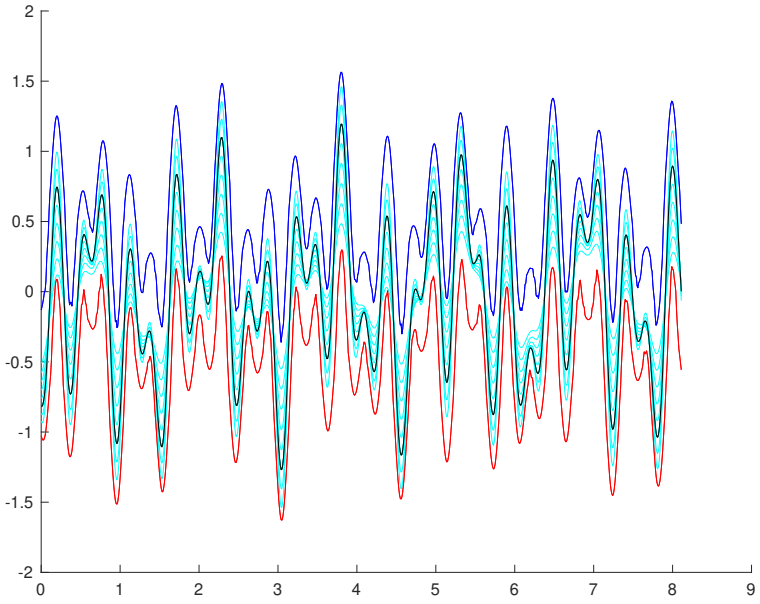
7. Conclusion

In this paper, we are interested in a new way of representing a relation between the inputs and outputs of a system, or, to be more precise, in a new representation of a function linking N inputs to one output. The classical approach is based on the notion of accurate representation: defining a parametric model and trying, by an optimization process, to find the values of the parameters of the model that best fits its behaviour. One of the most used models is the linear model. It has the advantage of being simple to set up and use, but the disadvantage of leading to a description of the functioning of the system that is too approximate, without it being possible to control this approximation. The use of non-linear models makes the representation more accurate but less simple and robust, and with more parameters to tune (or learn). Controlling how the modeling approximates the real system is even less easy.

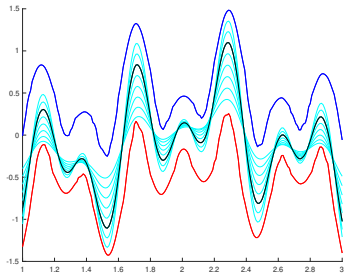
The approach we have proposed in this article exploits a completely different paradigm. Instead of trying to describe a system precisely, we propose to represent it by an imprecise linear modeling. Within this model, a system is associated to a set of weights, as in a linear model, but what we obtain is the description of a convex set of linear relations. Within this approach, how this modeling is close to the behaviour of the system to be represented can be controlled.

We believe that this approach can allow extending the notion of domination, widely used in decision theory, to many other domains. We have mainly given examples in signal processing. However, the macsum approach can be used to make approximate representations in any domain where the linear representation is relevant and where it would be interesting to model how well it describes a real aggregation process.

One important remaining question, that will be our main track for future work, is *how to identify a system?* i.e. how to learn the weights of a macsum representation of a function with benchmark values as in the classical approach? Finally, this modeling can suffer from lack of specificity (i.e. representing a too wide set of linear relations). Thus another path of investigation would be to tighten these boundaries to get a more accurate approximation while keeping the simplicity of the model.



a)



b)

Figure 5: Precise derivatives with $\beta \in [0.88, 0.96]$ (in cyan) superimposed on Figure (3) (a) with detailled view (b)

References

[1] R. McCullough B. Mitchell1, R. Ekey and W. Reitz. A fantastic quantitative exploration of ohm's law. *The Physics Teacher*, pages 56 – 75, 2018.

[2] R. Boukezzoula, S. Galichet, and A. Bisserier. A midpoint-radius approach to regression with interval data. *International Journal of Approximate Reasoning*, 52:1257–1271, 2011.

[3] R. Boukezzoula, S. Galichet, and D. Coquin. From fuzzy regression to gradual regression: Interval-based analysis and extensions. *Information Sciences*, 441:18–40, 2018.

[4] R. Boukezzoula, L. Jaulin, and L. Foulloy. Thick gradual intervals: An alternative interpretation of type-2 fuzzy intervals and its potential use in type-2 fuzzy computations. *Eng. Appl. Artif. Intell.*, 85:691–712, 2019.

[5] T. Chen. An electre-based outranking method for multiple criteria group decision making using interval type-2 fuzzy sets. *Information Sciences Elsevier (263)*, pages 1–21, 2014.

[6] L. Coppock and D. Mateer. *Principles of Microeconomics*. 2017.

[7] I. Couso and D Dubois. Statistical reasoning with set-valued information: Ontic vs. epistemic views. *International Journal of Approximate Reasoning*, 55(7):1502–1518, 2014. Special issue: Harnessing the information contained in low-quality data sources.

[8] J. Hunter D. Jorgenson and M.Nadiri. The predictive performance of econometric models of quarterly investment behavior. *Econometrica*, 38:213–224, 1970.

[9] D Denneberg. *Non-additive Measure and Integral*. Kluwer Academic Publishers, Dordrecht, 1994.

[10] S. Destercke, A. Rico, and O. Strauss. Approximating general kernels by extended fuzzy measures: application to filtering. In *18th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, volume 2, pages 112–123, 2020.

[11] D. Dubois and H. Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Springer, 1988.

[12] K. Loquin F. Jacquey, F. Comby, and O. Strauss. Non-additive approach for gradient-based detection. In *ICIP'07*, pages 49–52, San Antonio, Texas, 2007.

[13] F. Graba, K. Loquin, F. Comby, and O. Strauss. Guaranteed reconstruction for image super-resolution. In *Fuzzy Systems (FUZZ), 2013 IEEE International Conference on*, pages 1–8. IEEE, 2013.

[14] M. Grabisch. Fuzzy measures and integrals: Recent developments. fifty years of fuzzy logic and its applications. *Fifty years of fuzzy logic and its applicationss*, pages 125 – 151, 2015.

[15] J. Hall and J. Lawry. Generation, combination and extension of random set approximations to coherent lower and upper probabilities. *Reliability Engeneering and System Safety*, 85(1-3):89–101, 2004.

[16] K. Hampshire, H. Hamill, S. Mariwah, J. Mwanga, and D. Amoako-Sakyie. The application of signalling theory to health-related trust problems: The example of herbal clinics in ghana and tanzania. *Social Science & Medicine*, 188:109–118, 2017.

[17] V. Itier, F. Kucharczak, O. Strauss, and W. Puech. Interval-valued jpeg decompression for artifact suppression. pages 1–6, 2018.

[18] D. Chen J. Mao and L. Zhang. Mechanical assembly quality prediction method based on state space model. *The International Journal of Advanced Manufacturing Technology volume*, 86:107–116, 2016.

[19] F. Jacquey, F. Comby, and O. Strauss. Fuzzy edge detection for omnidirectional images. *Fuzzy sets and Systems*, 159(15):1991–2010, 2008.

[20] L. Jaulin, M. Kieffer, E. Walter, and D. Meizel. Guaranteed robust nonlinear estimation with application to robot localization. *IEEE Transactions on systems, man and cybernetics; Part C - Applications and Reviews*, 32:374–382, 2003.

[21] L. Jaulin and E. Walter. Guaranteed nonlinear parameter estimation from bounded-error data via interval analysis. *Mathematics and Computers in Simulation*, 35:123–127, 1993.

[22] L. Jaulin and E. Walter. Set inversion via interval analysis for nonlinear bounded-error estimation. *Automatica (29)*, pages 1053–1064, 1996.

[23] G. Kleiter. Propagating imprecise probabilities in bayesian networks. *Artificial Intelligence (88)*, pages 143–161, 1996.

[24] P. Liu. A weighted aggregation operators multi-attribute group decision-making method based on interval-valued trapezoidal fuzzy numbers. *Expert Systems with Applications (38)*, pages 1053–1060, 2011.

[25] K. Loquin and O. Strauss. Histogram density estimators based upon a fuzzy partition. *Statistics and Probability Letters*, 78(13):1863–1868, September 2008.

[26] K. Loquin and O. Strauss. On the granularity of summative kernels. *Fuzzy Sets and Systems*, 159(15):1952–1972, August 2008.

[27] K. Loquin and O. Strauss. Linear filtering and mathematical morphology on an image: a bridge. In *Proc. IEEE International Conference on Image Processing*, pages 3965–3968, Le Caire, Egypt, 2009.

[28] K. Loquin and O. Strauss. Noise quantization via possibilistic filtering. In *Proc. 6th Int. Symp. on Imprecise Probability:Theories and Application*, pages 297–306, Durham, United Kingdom, July 2009.

[29] K. Loquin, O. Strauss, and J.F. Crouzet. Possibilistic signal processing: How to handle noise? *International Journal of Approximate Reasoning*, 51(9):1129–1144, 2010.

[30] M. Sugeno M. Grabisch and T. Murofushi. Fuzzy measures and integrals: theory and applications. *Heidelberg: Physica*, 2000.

[31] Y. Narukawa and V. Torra. Non-monotonic fuzzy measures and intuitionistic fuzzy sets. In Vicenç Torra, Yasuo Narukawa, Aïda Valls, and Josep Domingo-Ferrer, editors, *Modeling Decisions for Artificial Intelligence*, pages 150–160, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[32] J. Nuno and C. Lopes. *Simulation and Modeling in Computational Chemistry: A Molecular Portfolio*. 2007.

[33] A. Rico and O. Strauss. Imprecise expectations for imprecise linear filtering. *International Journal of Approximate Reasoning*, 51(8):933–947, October 2010.

[34] B. Rinner and U. Weiss. Online monitoring by dynamically refining imprecise models. *IEEE Transactions on Systems, Man and cybernetics – part B: Cybernetics*, 34:213–224, 2004.

[35] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

[36] P. Shary. *Interval Regularization for Inaccurate Linear Algebraic Equations*, pages 375–398. Springer International Publishing, Cham, 2020.

[37] J. Shen and S. Castan. Towards the unification of band-limited derivative operators for edge detection. *Signal Processing*, 31(2):103–119, 1993.

[38] O. Strauss, K. Loquin, and F. Kucharczak. *On Maxitive Image Processing*, pages 201–215. Springer International Publishing, Cham, 2021.

[39] O. Strauss and A. Rico. Where the domination of maxitive kernels is extended to signed values. pages 1–6, 2019.

[40] M. Sugeno T. Murofushi and M. Machida. Non-monotonic fuzzy measures and the choquet integral. *Fuzzy Sets and Systems*, 64(1):73–86, 1994.

[41] A. De Waegenaere and P. Wakker. Choquet integrals with respect to non-monotonic set functions. Discussion Paper 1997-44, Tilburg University, Center for Economic Research, 1997.

[42] A. De Waegenaere and P. Wakker. Choquet integrals with respect to non-monotonic set functions. *Tilburg University*, 1997.

[43] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. 1990.

[44] R. Yager and V. Kreinovich. Decision making under interval probabilities. *International Journal of Approximate Reasoning*, 22:195–215, 1999.