



**HAL**  
open science

## Macsum: A new interval-valued linear operator

Olivier Strauss, Agnès Rico, Yassine Hmidy

► **To cite this version:**

Olivier Strauss, Agnès Rico, Yassine Hmidy. Macsum: A new interval-valued linear operator. International Journal of Approximate Reasoning, 2022, 145, pp.121-138. 10.1016/j.ijar.2022.03.003 . hal-03837036

**HAL Id: hal-03837036**

**<https://hal.science/hal-03837036v1>**

Submitted on 2 Nov 2022

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

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

# Macsum: a new interval-valued linear operator

Olivier Strauss<sup>a</sup>, Agnès Rico<sup>b</sup>, Yassine Hmidy<sup>a</sup>

<sup>a</sup>*LIRMM, Université de Montpellier, CNRS, France*

<sup>b</sup>*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 measure, predict or 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 measure. What we propose in this article is an imprecise linear model, so the simplicity of representation and use is quite comparable to that of a linear model. This model is imprecise in the sense that the output is imprecise, although the inputs are precise, thus potentially reflecting the inadequacy of a linear model to represent the behavior of the system: the more imprecise the output, the less likely a single linear model is to correctly describe the system. 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

---

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

a function of the current in an electronic circuit [2] or the budget constraint of a homo-economicus in a model of microeconomics [8], 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 [18], in mechanics to predict the performance of an assembly and thus optimize the manufacturing of objects [20], in chemistry to develop new products without real tests thanks to simulation techniques [6], in econometrics to improve predictive models [10], 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 prominence 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. However, there are so called sensitivity analysis methods [21, 43] which allow estimation of the influence of a variation of one of the system inputs on its output. Nevertheless, this type of approach is beyond the scope of this article. 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 cost of a lesser 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 [41] proposes to consider solving ill-conditioned systems of linear algebraic equations by *intervalizing* the parameters of the sought 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 [9]. On a more applied level, Rinner and Weiss [39] propose using this interval-valued representation to deal with incomplete knowledge about a system to be supervised. In [3], 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 sophisticated representation [4]. In a completely different manner, in [24], 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 [25] with applications to robotics [23]. In a recent article [5], 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 [48] first and then many other authors in multi-criteria decision making [7, 27, 49, 40] or random sets approximation [17] 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 [26]. 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 [28]. 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 [38]. 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 predicted and partially controlled.

After this introductory section, the article is organized as follows. Section 2 recalls settings we consider as well as few preliminary considerations and notations. Section 3 presents the **macsum** operator as a signed generalization of the maxitive operator. Section 4 shows how this operator can be used to represent a convex set of linear operators, thus leading to an interval-valued aggregation. In section 5, we propose an index of cardinality of this convex set. A very simple algorithm is proposed in Section 6 for computing a macsum-

based aggregation. Section 7 proposes an experiment that illustrates, through a practical application, the properties described in the previous parts of the article. We finally conclude in Section 8.

## 2. Theoretical background

### 2.1. Notations

- $\Omega = \{1, \dots, N\} \subset \mathbb{N}$ .
- $x : \Omega \rightarrow \mathbb{R}$ , is a function defined by a discrete subset of  $\mathbb{R}^N$ :  
 $x = (x_1, \dots, x_N) \in \mathbb{R}^N$ .
- $[\cdot]$  is the permutation that sorts the  $x_i$ 's in increasing order:  
 $x_{[1]} \leq x_{[2]} \leq \dots \leq x_{[N]}$ .
- $[\cdot]$  is the permutation that sorts the  $x_i$ 's in decreasing order:  
 $x_{[1]} \geq x_{[2]} \geq \dots \geq x_{[N]}$ .
- $A_{[i]}$  ( $i \in \Omega$ ) is the coalition of  $\Omega$  defined by  $A_{[i]} = \{[i], \dots, [N]\}$ .
- $A_{[i]}$  ( $i \in \Omega$ ) is the coalition of  $\Omega$  defined by  $A_{[i]} = \{[i], \dots, [N]\}$

### 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  $\Omega$  is a discrete function  $\varphi : \Omega \rightarrow \mathbb{R}$  defined by  $\varphi = (\varphi_1, \dots, \varphi_N)$ .
- The set of kernels of  $\Omega$  is denoted  $\mathcal{K}(\Omega) \equiv \mathbb{R}^N$ .
- A set function is a function  $\vartheta : 2^\Omega \rightarrow \mathbb{R}$  that associates a real value to any subset of  $\Omega$ .
- A set function  $\vartheta$  of  $\Omega$  is said to be **concave** or **supermodular** if  $\forall A, B \subseteq \Omega$ ,  $\vartheta(A \cup B) + \vartheta(A \cap B) \geq \vartheta(A) + \vartheta(B)$ .
- A set function  $\vartheta$  of  $\Omega$  is said to be **convex** or **submodular** if  $\forall A, B \subseteq \Omega$ ,  $\vartheta(A \cup B) + \vartheta(A \cap B) \leq \vartheta(A) + \vartheta(B)$ .
- A set function  $\vartheta$  of  $\Omega$  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  $\vartheta$  of  $\Omega$  can be associated a complementary set function  $\vartheta^c$  defined by  $\forall A \subseteq \Omega$ ,  $\vartheta^c(A) = \vartheta(\Omega) - \vartheta(A^c)$ ,  $A^c$  being the complementary set of  $A$  in  $\Omega$ .
- If a set function  $\vartheta$  is concave (rsp. convex) then  $\vartheta^c$  is convex (rsp. concave).

- A capacity is a normalized increasing set function  $v : 2^\Omega \rightarrow \mathbb{R}^+$  with  $v(\emptyset) = 0$ . Normalized means  $v(\Omega) = 1$  and increasing means that  $\forall A \subseteq B \subseteq \Omega$ ,  $v(A) \leq v(B)$ . To a capacity  $v$  is associated its complementary capacity  $v^c$ :  $\forall A \in \Omega$ ,  $v^c(A) = 1 - v(A^c)$ .
- A **maxitive kernel** of  $\Omega$  is a discrete function  $\pi : \Omega \rightarrow [0, 1]$  defined by  $\pi = (\pi_1, \dots, \pi_N) \in [0, 1]^N$  such that  $\max_{i \in \Omega} \pi_i = 1$ .
- The set of maxitive kernels of  $\Omega$  is denoted  $\mathcal{K}_m(\Omega) \subseteq \mathcal{K}(\Omega)$ .
- To a maxitive kernel can be associated a capacity  $\Pi_\pi$  defined by  $\forall A \subseteq \Omega$   $\Pi_\pi(A) = \max_{i \in A} \pi_i$ . This function is called a **possibility** measure when  $\pi$  is interpreted as a possibility distribution [13]. 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  $\Omega$  is a discrete function  $\rho : \Omega \rightarrow (\mathbb{R}^+)^+$  defined by  $\rho = (\rho_1, \dots, \rho_N) \in (\mathbb{R}^N)^+$  such that  $\sum_{i \in \Omega} \rho_i = 1$ .
- The set of summative kernels of  $\Omega$  is denoted  $\mathcal{K}_s(\Omega) \subseteq \mathcal{K}(\Omega)$ .
- To a summative kernel can be associated an additive set function  $P_\rho$  defined by  $\forall A \subseteq \Omega$ ,  $P_\rho(A) = \sum_{i \in A} \rho_i$ .  $P_\rho$  is called a **probability** measure in the context of confidence measure. The complementary set function  $P_\rho^c$  of  $P_\rho$  is  $P_\rho$  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  $\varphi = (\varphi_1, \dots, \varphi_N)$  be an unnormalized ( $\sum_{i \in \Omega} \varphi_i \neq 1$ ) or signed (i.e.  $\exists i, j \in \Omega$  such that  $\varphi_i \cdot \varphi_j < 0$ ) kernel. The additive set function associated to  $\varphi$  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  $\forall A \subseteq \Omega$ ,  $\Pi_\pi(A) \geq P_\rho(A)$ . [12]
- The set of summative kernels dominated by a maxitive kernel  $\pi$ , 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  $v$  denoted as  $\mathcal{M}(v)$  is the set of probability measures that it dominates. When  $v$  is concave, this can be written  $\mathcal{M}(v) = \{P \in \mathcal{P}(\Omega) : \forall A \subseteq \Omega, v^c(A) \leq P(A) \leq v(A)\}$ , where  $\mathcal{P}(\Omega)$  is the set of probability measures defined on  $\Omega$ .

### 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) [11].

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  $v$ :  $y = \mathbb{C}_v(x)$  [33].

$$y = \mathbb{C}_v(x) = \sum_{k=1}^N x_{\lceil k \rceil} \cdot (v(A_{\lceil k \rceil}) - v(A_{\lceil k+1 \rceil})), \text{ with } A_{\lceil N+1 \rceil} = \emptyset. \quad (1)$$

$$y = \mathbb{C}_v(x) = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) \cdot v(A_{\lceil k \rceil}), \text{ with } x_{\lceil 0 \rceil} = 0. \quad (2)$$

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

**Proposition 2.1.** (see [34] corollary C4, p. 381) Equations (1) and (2) can be used to compute the asymmetric Choquet integral.

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. [46], 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  $\mu : 2^\Omega \rightarrow \mathbb{R}$  such that  $\mu(\emptyset) = 0$ .

Several work, including [16, 33, 47, 36] proposed to extend Choquet integral to non-monotonic set functions. In particular, in [47], 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 [46] 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 comonotonically additive, which means that if  $f$  and  $g$  are two comonotonic 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:  $y = \mathbb{E}_\rho(x) = \sum_{k \in \Omega} \rho_k \cdot x_k$ , where  $\rho \in \mathcal{K}_s(\Omega)$  is a summative kernel of  $\Omega$ . When  $\rho$  is interpreted as a probability distribution, then the value  $y$  is called the expectation of  $x$ . Let  $P_\rho$  be the additive set function associated to  $\rho$ , then  $\mathbb{E}_\rho(x) = \check{\mathbb{C}}_{P_\rho}(x)$ .

In [29], Loquin et al. propose a new aggregation, under the name of **max-itive expectation**, denoted  $\overline{\mathbb{E}}_\pi$ , where  $\pi \in \mathcal{K}_m(\Omega)$  is a maxitive kernel of  $\Omega$ . 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,  $\overline{\mathbb{E}}_\pi$  leads to an imprecise expectation  $\overline{\mathbb{E}}_\pi(x) = [\underline{y}, \overline{y}] = [\check{\mathbb{C}}_{\Pi_\pi^c}(x), \check{\mathbb{C}}_{\Pi_\pi}(x)]$ . It has been shown in [32] Theorem 1 that:

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

Therefore, the maxitive aggregation  $\overline{\mathbb{E}}_\pi(x)$  of  $x$  w.r.t. the maxitive kernel  $\pi$  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 [31], image processing [30, 14, 15], statistics [28], 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 [44].

To circumvent this problem, in [45] 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  $(\pi^+, \pi^-)$ , 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  $\mu_{\pi^+, \pi^-}$  that is still normalized (i.e.  $\mu_{\pi^+, \pi^-}(\Omega) = 1$ ) but that is not increasing w.r.t. union in  $\Omega$ . Aggregating  $x$  w.r.t.  $\mu_{\pi^+, \pi^-}$  requires the use of an extension of the Choquet integral as proposed in [47]. 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  $\mu_\varphi : 2^\Omega \rightarrow \mathbb{R}$ , with  $\varphi \in \mathcal{K}(\Omega)$  ( $\mu_\varphi(\emptyset) = 0$ ) where,  $\forall A \subseteq \Omega$ , the value of  $\mu_\varphi(A)$  only depends on the  $N$  values of the kernel  $\varphi$ . As for capacities, we can associate to  $\mu_\varphi$  a complementary operator  $\mu_\varphi^c$ :  $\mu_\varphi^c(A) = \mu_\varphi(\Omega) - \mu_\varphi(A^c)$ ,  $A^c$  being the complementary set of  $A$  in  $\Omega$  [33].

The additive set function presented in Section 2.2 is a good example of an operator: a set function  $P_\varphi$  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  $P_\varphi^c = P_\varphi$ .

The maxitive set function proposed in [29] 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 [45] 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  $\nu$  that associated to any kernel  $\varphi \in \mathcal{K}(\Omega)$  a set function  $\nu_\varphi$  as follows:

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

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

As the max operator is an increasing set function while the min operator is a decreasing set function, we propose the following definition:

**Definition 3.1.** The max over the empty set of a positive function is equal to 0:  $\max_{i \in \emptyset} \varphi_i^+ = 0$  and the min over the empty set of a negative function is equal to 0:  $\min_{i \in \emptyset} \varphi_i^- = 0$ .

This definition is in compliance with :  $\forall A \subseteq B \subseteq \Omega$ ,  $\max_{i \in \emptyset} \varphi_i^+ \leq \max_{i \in A} \varphi_i^+ \leq \max_{i \in B} \varphi_i^+$  and  $\min_{i \in B} \varphi_i^- \leq \min_{i \in A} \varphi_i^- \leq \min_{i \in \emptyset} \varphi_i^-$  since – in line with the standard axiomatic set theory – the empty set is a subset of every subset of  $\Omega$ . It is also in line with the usual convention, whereby the sup (resp. inf) over an empty family is the smallest (resp. largest) element, when  $\varphi$  is a signed kernel (i.e.  $\exists i, j \in \Omega / \varphi_i \leq 0$  and  $\varphi_j \geq 0$ ).

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

The complementary set function of  $\nu_\varphi$ , denoted  $\nu_\varphi^c$ , is given by:

$$\forall A \subseteq \Omega, \nu_\varphi^c(A) = \min_{i \in A} \varphi_i^- + \bar{\alpha} - \max_{i \in A^c} \varphi_i^+. \quad (4)$$

This simply comes from the definition of a complementary set function:

$$\begin{aligned} \nu_\varphi^c(A) &= \nu_\varphi(\Omega) - \nu_\varphi(A^c) = \bar{\alpha} + \underline{\alpha} - \left( \max_{i \in A^c} \varphi_i^+ - \min_{i \in A} \varphi_i^- + \underline{\alpha} \right) \\ &= \min_{i \in A} \varphi_i^- + \bar{\alpha} - \max_{i \in A^c} \varphi_i^+ = \min_{i \in A} \varphi_i^- + \min_{i \in A^c} (\bar{\alpha} - \varphi_i^+). \end{aligned}$$

Thus  $\nu_\varphi(\Omega) = \nu_\varphi^c(\Omega) = \underline{\alpha} + \bar{\alpha}$ .

Moreover,  $\nu_\varphi(\emptyset) = \max_{i \in \emptyset} \varphi_i^+ + \min_{i \in \Omega} \varphi_i^- - \min_{i \in \Omega} \varphi_i^- = 0$  and  $\nu_\varphi^c(\emptyset) = \min_{i \in \emptyset} \varphi_i^- + \max_{i \in \Omega} \varphi_i^+ - \min_{i \in \Omega} \varphi_i^+ = 0$ .

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

**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  $\tau = \max_{i \in A \cup B} \varphi_i + \max_{i \in A \cap B} \varphi_i - \max_{i \in A} \varphi_i - \max_{i \in B} \varphi_i$ .  
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  $\tau = b + c - a - b = c - a \leq 0$ .  $\square$

**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  $\tau = \min_{i \in A \cup B} \varphi_i + \min_{i \in A \cap B} \varphi_i - \min_{i \in A} \varphi_i - \min_{i \in B} \varphi_i$ .  
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  $\tau = a + c - a - b = c - b \geq 0$ .  $\square$

**Proposition 3.3.**  $\nu_\varphi$  is concave and thus  $\nu_\varphi^c$  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  $\nu_\varphi$  is concave. Proving  $\nu_\varphi^c$  being convex can be done in the same way.  $\square$

**Remark 1.** The macsum operator defined by Expression (3) is an extension of the one proposed by Loquin et al. [29] since if  $\pi \in \mathcal{K}_m(\Omega)$  is a maxitive kernel, then  $\nu_\pi = \Pi_\pi$  is a maxitive aggregation fonction defined by:  $\forall A \subseteq \Omega, \nu_\pi(A) = \Pi_\pi(A) = \max_{i \in A} \pi_i$ .

### 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. [29]:

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

**Remark 2.** This definition coincides with the one of [29] when  $\varphi$  is a maxitive kernel.

**Remark 3.** To be inline with the work of [29], let  $\psi, \varphi \in \mathcal{K}(\Omega)$ , if  $\psi \in \mathcal{M}(\varphi)$  we say that  $\varphi$  **dominates**  $\psi$  because the macsum operator based on  $\varphi$  dominates the additive operator base on  $\psi$ .  $\psi$  belongs to the convex set represented by  $\varphi$ .

**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  $\Omega$ , 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  $\Omega$ , is there a simple way to define a kernel  $\varphi \in \mathcal{K}(\Omega)$  such that  $\psi \in \mathcal{M}(\varphi)$ ?

In [29], Loquin et al. have used some known properties defined in the context of possibility theory [13] 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 3.4.** (*weak domination*) Let  $\rho \in \mathcal{K}_s(\Omega)$  be a summative kernel of  $\Omega$ , then the maxitive kernel  $\check{\pi} \in \mathcal{K}_m(\Omega)$ , defined by  $\forall i \in \Omega, \check{\pi}_i = \sum_{j \in \Omega} \min(\rho_i, \rho_j)$ , dominates  $\rho$ , i.e.  $\rho \in \mathcal{M}(\check{\pi})$ .

**Lemma 3.5.** (*strong domination*) Let  $\rho \in \mathcal{K}_s(\Omega)$  be a summative kernel of  $\Omega$ , then the maxitive kernel  $\hat{\pi} \in \mathcal{K}_m(\Omega)$ , defined by  $\forall i \in \Omega, \hat{\pi}_i = \sum_{j \in A_i} \rho_j$ , where  $A_i = \{j \in \Omega / \rho_j \leq \rho_i\}$ , dominates  $\rho$  i.e.  $\rho \in \mathcal{M}(\hat{\pi})$ .

As shown in [29],  $\check{\pi}$  is said to weaker dominate  $\rho$  than  $\hat{\pi}$  since  $\mathcal{M}(\hat{\pi}) \subseteq \mathcal{M}(\check{\pi})$ .

What we propose here is to use Lemma 3.4 and 3.5 to define also weak and strong domination relations between two kernels of  $\mathcal{K}(\Omega)$ .

**Proposition 3.6.** (*general weak domination*) Let  $\psi \in \mathcal{K}(\Omega)$  be a kernel of  $\Omega$ , then the kernel  $\check{\varphi} \in \mathcal{K}(\Omega)$  defined by  $\forall i \in \Omega, \check{\varphi}_i = \sum_{j \in \Omega} \min(\psi_i^+, \psi_j^+) + \sum_{j \in \Omega} \max(\psi_i^-, \psi_j^-)$ , where  $\psi^+ = \max(0, \psi)$  and  $\psi^- = \min(0, \psi)$  dominates  $\psi$  i.e.  $\psi \in \mathcal{M}(\check{\varphi})$ .

*Proof.* Let us define  $\alpha^+ = \sum_{i \in \Omega} \psi_i^+$  and  $\alpha^- = \sum_{i \in \Omega} \psi_i^-$ . Let  $\rho^+$  and  $\rho^-$  be two kernels defined by:  $\forall i \in \Omega, \rho_i^+ = \psi_i^+ / \alpha^+$  and  $\rho_i^- = -\psi_i^- / \alpha^-$ . By construction  $\rho^+$  and  $-\rho^-$  are summative in the sense of Loquin [29] (i.e. positive and normalized). Thus, according to Lemma 3.4, two maxitive kernels  $\pi$  and  $\tau$  can be defined by:

$$\forall i \in \Omega, \pi_i = \sum_{j \in \Omega} \min(\rho_i^+, \rho_j^+) \text{ and } \tau_i = \sum_{j \in \Omega} \min(-\rho_i^-, -\rho_j^-),$$

such that  $\forall A \subseteq \Omega$ ,

$$\begin{aligned} \Pi_\pi(A) &= \max_{i \in A} \pi_i \geq P_{\rho^+}(A) = \sum_{i \in A} \rho_i^+ \text{ and} \\ \Pi_\tau(A) &= \max_{i \in A} \tau_i \geq P_{(-\rho^-)}(A) = -P_{\rho^-}(A) = -\sum_{i \in A} \rho_i^-. \end{aligned}$$

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

$$\begin{aligned}\forall i \in \Omega, \check{\varphi}_i^+ &= \alpha^+ \cdot \pi_i = \sum_{j \in \Omega} \min(\psi_i^+, \psi_j^+) \text{ and} \\ \check{\varphi}_i^- &= \alpha^- \cdot \tau_i = \sum_{j \in \Omega} \max(\psi_i^-, \psi_j^-).\end{aligned}$$

By construction  $\forall i \in \Omega, \check{\varphi}_i^+ \geq 0$  and  $\check{\varphi}_i^- \leq 0$ .

Let us define  $\check{\varphi} = \check{\varphi}^+ + \check{\varphi}^-$ .

Let us consider the two operators  $\vee$  and  $\wedge$  that associate two set functions  $\vee_\phi$  and  $\wedge_\phi$  to each  $\phi \in \mathcal{K}(\Omega)$  as follows:

$$\forall A \subseteq \Omega, \vee_\phi(A) = \max_{i \in A} \phi_i \text{ and } \wedge_\phi(A) = \min_{i \in A} \phi_i.$$

We have,  $\forall A \subseteq \Omega$ ,

$$\begin{aligned}\vee_{\check{\varphi}^+}(A) &= \max_{i \in A} \check{\varphi}_i^+ = \alpha^+ \max_{i \in A} \pi_i \geq \alpha^+ \cdot P_{\rho^+}(A) = \sum_{i \in A} \alpha^+ \cdot \rho_i^+ = P_{\psi^+}(A) \text{ and} \\ \wedge_{\check{\varphi}^-}(A) &= \min_{i \in A} \check{\varphi}_i^- = \alpha^- \max_{i \in A} \tau_i \leq -\alpha^- \cdot P_{\rho^-}(A) = \sum_{i \in A} -\alpha^- \cdot \rho_i^- = P_{\psi^-}(A).\end{aligned}$$

Considering that  $\wedge_{\check{\varphi}^-}(\Omega) = \min_{i \in \Omega} \check{\varphi}_i^- = \sum_{i \in \Omega} \psi_i^- = \alpha^-$ , we also have:

$$\wedge_{\check{\varphi}^-}^c(A) = \wedge_{\check{\varphi}^-}(\Omega) - \wedge_{\check{\varphi}^-}(A^c) \geq P_{\psi^-}(\Omega) - P_{\psi^-}(A^c) = P_{\psi^-}(A).$$

Now,  $\forall A \subseteq \Omega$ ,

$$\begin{aligned}\nu_{\check{\varphi}}(A) &= \max_{i \in A} \check{\varphi}_i^+ + \min_{i \in \Omega} \check{\varphi}_i^- - \min_{i \in A^c} \check{\varphi}_i^- \\ &= \vee_{\check{\varphi}^+}(A) + \wedge_{\check{\varphi}^-}^c(A) \geq P_{\psi^+}(A) + P_{\psi^-}(A),\end{aligned}$$

and therefore  $\nu_{\check{\varphi}}(A) \geq P_{\psi}(A)$ .  $\square$

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

The proof of Proposition 3.7 follows the same pattern as the proof of Proposition 3.6.

*Proof.* As previously, we define two kernels  $\rho^+$  and  $\rho^-$  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 previously,  $\rho^+$  and  $-\rho^-$  are summative kernels.

As proposed in Lemma 3.5, we can define two maxitive kernels  $\pi$  and  $\tau$  by:

$$\begin{aligned}\forall i \in \Omega, \pi_i &= \sum_{j \in A_i^+} \rho_j^+, \text{ with } A_i^+ = \{j \in \Omega / \rho_j^+ \leq \rho_i^+\} = \{j \in \Omega / \psi_j^+ \leq \psi_i^+\}, \\ \tau_i &= \sum_{j \in A_i^-} -\rho_j^-, \text{ with } A_i^- = \{j \in \Omega / -\rho_j^- \leq -\rho_i^-\} = \{j \in \Omega / \psi_j^- \geq \psi_i^-\},\end{aligned}$$

such that,  $\forall A \subseteq \Omega$ ,

$$\begin{aligned}\Pi_\pi(A) &= \max_{i \in A} \pi_i \geq P_{\rho^+}(A) = \sum_{i \in A} \rho_i^+, \text{ and} \\ \Pi_\tau(A) &= \max_{i \in A} \tau_i \geq P_{-\rho^-}(A) = - \sum_{i \in A} \rho_i^-. \end{aligned}$$

Let us define the two kernels  $\hat{\varphi}^+$  and  $\hat{\varphi}^-$  by:  
 $\hat{\varphi}^+ = \alpha^+ \cdot \pi$  (i.e.  $\forall i \in \Omega$ ,  $\hat{\varphi}_i^+ = \sum_{j \in A_i^+} \psi_j^+$ ) and  
 $\hat{\varphi}^- = -\alpha^- \cdot \tau$  (i.e.  $\forall i \in \Omega$ ,  $\hat{\varphi}_i^- = \sum_{j \in A_i^-} \psi_j^-$ ).  
Let  $\hat{\varphi} = \hat{\varphi}^+ + \hat{\varphi}^-$ .

Let us consider the two previously defined operators  $\vee$  and  $\wedge$  that associate two set functions  $\vee_\phi$  and  $\wedge_\phi$  to each  $\phi \in \mathcal{K}(\Omega)$  as follows:  
 $\forall A \subseteq \Omega$ ,  $\vee_\phi(A) = \max_{i \in A} \phi_i$  and  $\wedge_\phi(A) = \min_{i \in A} \phi_i$ .

By construction we have,  $\forall A \subseteq \Omega$ :  
 $\nu_{\hat{\varphi}}(A) = \vee_{\hat{\varphi}^+}(A) + \wedge_{\hat{\varphi}^-}(A) = \vee_{\hat{\varphi}^+}(A) + \wedge_{\hat{\varphi}^-}^c(A)$ .

Following the same scheme as in the previous proof, we can deduce:  
 $\vee_{\hat{\varphi}^+}(A) \geq P_{\psi^+}(A)$  and  $\wedge_{\hat{\varphi}^-}^c(A) \geq P_{\psi^-}(A)$  and thus  $\nu_{\hat{\varphi}}(A) \geq P_\psi(A)$   $\square$

Finally, we have the following property:

**Proposition 3.8.**  $\forall \varphi \in \mathcal{K}(\Omega)$ ,  $\mathcal{M}(\varphi)$  is not empty.

*Proof.* To prove Proposition 3.8, 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  $\psi$  such that  $\varphi$  strongly dominates  $\psi$  as suggested in Proposition 3.7.

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

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

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

$$\begin{aligned}\forall i \in \{2, \dots, N\}, \quad \psi_i^+ &= \varphi_i^+ - \sum_{j=1}^{i-1} \psi_j^+ = \varphi_i^+ - \varphi_{i-1}^+, \text{ with } \psi_1^+ = \varphi_1^+, \text{ and} \\ \forall i \in \{N-1, \dots, 1\}, \quad \psi_i^- &= \varphi_i^- - \sum_{j=i+1}^N \psi_j^- = \varphi_i^- - \varphi_{i+1}^-, \text{ with } \psi_N^- = \varphi_N^-. \end{aligned}$$

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

#### 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  $\overline{F}_\varphi(x)$ .

The linear aggregation is simply a weighted sum defined by:

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

The macsum aggregation is defined by:

$$\overline{F}_\varphi(x) = [E_\varphi(x), \overline{F}_\varphi(x)] = [\check{C}_{\nu_\varphi^c}(x), \check{C}_{\nu_\varphi}(x)]. \quad (7)$$

This extends the work of Loquin et al. since:

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

*Proof.* The fact that  $\forall \psi \in \mathcal{M}(\varphi)$ ,  $\forall x \in \mathbb{R}^N$ ,  $F_\psi(x) \in \overline{F}_\varphi(x)$  is simply a reformulation of the Schmeidler-Denneberg theorem [11] 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 let's prove that  $\forall y \in \overline{F}_\varphi(x)$ ,  $\exists \psi$  such that  $y = F_\psi(x)$ . We have  $\check{C}_{\nu_\varphi}(x) = \sum_{k=1}^N x_{[k]} \cdot (\nu_\varphi(A_{[k]}) - \nu_\varphi(A_{[k+1]}))$  and  $\check{C}_{\nu_\varphi^c}(x) = \sum_{k=1}^N x_{[k]} \cdot (\nu_\varphi^c(A_{[k]}) - \nu_\varphi^c(A_{[k+1]}))$  as  $\overline{F}_\varphi(x) = [E_\varphi(x), \overline{F}_\varphi(x)] = [\check{C}_{\nu_\varphi^c}(x), \check{C}_{\nu_\varphi}(x)]$  is a convex set, and  $y \in \overline{F}_\varphi(x)$  we then have  $y = \lambda \check{C}_{\nu_\varphi^c}(x) + (1 - \lambda) \check{C}_{\nu_\varphi}(x)$  with  $\lambda \in [0, 1]$  therefore  $y = \lambda \sum_{k=1}^N x_{[k]} \cdot (\nu_\varphi^c(A_{[k]}) - \nu_\varphi^c(A_{[k+1]})) + (1 - \lambda) \sum_{k=1}^N x_{[k]} \cdot (\nu_\varphi(A_{[k]}) - \nu_\varphi(A_{[k+1]}))$  which gives that  $y = \sum_{k=1}^N x_{[k]} \cdot (\lambda(\nu_\varphi^c(A_{[k]}) - \nu_\varphi^c(A_{[k+1]})) + (1 - \lambda)(\nu_\varphi(A_{[k]}) - \nu_\varphi(A_{[k+1]})))$ . Thus we have that  $\psi = \{\psi_{[k]}\}_{k \in \Omega}$  is a kernel such that  $\psi_{[k]} = \lambda(\nu_\varphi^c(A_{[k]}) - \nu_\varphi^c(A_{[k+1]})) + (1 - \lambda)(\nu_\varphi(A_{[k]}) - \nu_\varphi(A_{[k+1]}))$  and finally  $y = \sum_{k=1}^N \psi_{[k]} \cdot x_{[k]} = F_\psi(x)$  as the sum is commutative.  $\square$

**Corollary.** *Let  $\phi, \varphi \in \mathcal{K}(\Omega)$ , such that  $\mathcal{M}(\varphi) \subseteq \mathcal{M}(\phi)$ , then  $\forall x \in \mathbb{R}^N$   $\overline{F}_\varphi(x) \subseteq \overline{F}_\phi(x)$  since  $\overline{F}_\varphi(x) = \{\psi \in \mathcal{M}(\varphi), F_\psi(x)\}$ .*

**Remark 5.** *The operator defined in Expression (3) may seem to come out of nowhere. In fact its construction is based on an attempt to dominate a linear operator as done in [45]. Let us consider a kernel  $\psi$  and the linear operator  $P_\psi$ . Let us consider  $\psi^+ = \max(\psi, 0)$  and  $\psi^- = \min(\psi, 0)$ . A weight  $P_\psi(A) = \sum_{i \in A} \psi_i = \sum_{i \in A} \psi_i^+ + \sum_{i \in A} \psi_i^-$  is associated to each set  $A \subseteq \Omega$ . Thus, when considering a linear aggregation based on a Choquet integral, the term  $\sum_{i \in A} \psi_i^+$  can be considered as a positive weight, and  $\sum_{i \in A} \psi_i^-$  as a negative weight. The macsum operator considers dominating the positive part by a maxitive operator*

defined by positive weights defined by  $\varphi^+$  (i.e.  $\sum_{i \in A} \psi_i^+ \leq \max_{i \in A} \varphi_i^+$ ) and dominating the negative part by a maxitive operator defined by negative weights defined by  $\varphi^-$ . For this last part, this is obtained by rewriting  $\sum_{i \in A} \psi_i^- = \underline{\alpha} - \sum_{i \in A^c} \psi_i^-$ , with  $\underline{\alpha} = \sum_{i \in \Omega} \psi_i^-$  (i.e.  $\sum_{i \in A} \psi_i^- \leq \max_{i \in A^c} (\underline{\alpha} - \varphi_i^-)$ ).

## 5. Macsum kernel granularity

The term granularity was introduced by Pawlak [37] who defined the granularity of a rough set as its power of resolution. This notion was extended by Loquin [31] to express the cardinal of the core of a maxitive kernel, i.e. a measure of its non-resolution power. In [31] the granularity  $\gamma$  of a maxitive kernel  $\pi \in \mathcal{K}_m(\Omega)$  (i.e the cardinal of its core) is defined as:  $\gamma(\pi) = \sum_{i=1}^N \pi_i$ . Therefore, with the less specific maxitive kernel being the vacuous kernel  $\varpi$  defined by  $\forall i \in \Omega, \varpi_i = 1$ , and the most specific maxitive kernel being the Krøenecker impulse  $\delta^k$  ( $k \in \Omega$ ) defined by  $\delta_k = 1$  and  $\forall i \in \Omega/k, \delta_i = 0$ , then  $\forall \pi \in \mathcal{K}_m(\Omega)$ ,  $\gamma(\pi) \in [1, N]$ .

In this section, we propose to extend this notion of granularity to macsum kernels in order to characterize the cardinal of the core of a macsum-based set function.

**Proposition 5.1.** *Let  $\varphi \in \mathcal{K}(\Omega)$ , let  $\bar{\alpha} = \max_{i \in \Omega} \varphi_i$  and  $\underline{\alpha} = \min_{i \in \Omega} \varphi_i$ , then if  $\psi \in \mathcal{M}(\varphi)$ ,  $\sum_{i \in \Omega} \max(0, \psi_i) = \bar{\alpha}$  and  $\sum_{i \in \Omega} \min(0, \psi_i) = \underline{\alpha}$ . Moreover,  $\forall i \in \Omega$ ,  $\varphi_i > 0 \Rightarrow \psi_i \geq 0$  and  $\varphi_i < 0 \Rightarrow \psi_i \leq 0$ .*

*Proof.* Let  $i \in \Omega$ . By construction  $\nu_\varphi^c(\{i\}) \leq P_\psi(\{i\}) = \psi_i \leq \nu_\varphi(\{i\})$ .

– If  $\varphi_i > 0$ , then  $\nu_\varphi^c(\{i\}) = \min_{k \in \{i\}} \varphi_k^- + \bar{\alpha} - \max_{k \in \{i\}^c} \varphi_k^+ = \bar{\alpha} - \max_{k \in \{i\}^c} \varphi_k^+$   
Thus  $0 \leq \nu_\varphi^c(\{i\}) \leq \psi_i$ .

– If  $\varphi_i < 0$ , then  $\nu_\varphi(\{i\}) = \max_{k \in \{i\}} \varphi_k^+ + \underline{\alpha} - \min_{k \in \{i\}^c} \varphi_k^- = \underline{\alpha} - \min_{k \in \{i\}^c} \varphi_k^-$   
Thus  $\psi_i \leq \nu_\varphi(\{i\}) \leq 0$ .

Let  $A = \{i \in \Omega / \varphi_i \geq 0\}$ ,  $\nu_\varphi(A) = \max_{i \in A} \varphi_i^+ + \underline{\alpha} - \min_{i \in A^c} \varphi_i^- = \max_{i \in \Omega} \varphi_i^+ = \bar{\alpha}$ , and  $\nu_\varphi^c(A) = \min_{i \in A} \varphi_i^- + \bar{\alpha} - \max_{i \in A^c} \varphi_i^+ = \bar{\alpha}$ .

Since  $P_\psi(A) = \sum_{i \in \Omega} \max(0, \psi_i)$  and  $\nu_\varphi^c(A) \leq P_\psi(A) \leq \nu_\varphi(A)$

thus  $\sum_{i \in \Omega} \max(0, \psi_i) = \bar{\alpha}$ .

In the same manner we have  $\nu_\varphi(A^c) = \nu_\varphi^c(A^c) = \underline{\alpha}$

and thus  $P_\psi(A^c) = \sum_{i \in \Omega} \min(0, \psi_i) = \underline{\alpha}$ . □

**Proposition 5.2.** *Let  $\varphi \in \mathcal{K}(\Omega)$ ,  $\psi \in \mathcal{M}(\varphi) \Leftrightarrow \forall \lambda \in \mathbb{R}^*$ ,  $\lambda \cdot \psi \in \mathcal{M}(\lambda \cdot \varphi)$ .*

*Proof.*  $\varphi \in \mathcal{K}(\Omega)$  means  $\forall A \in \Omega$ ,  $\nu_\varphi(A) \leq P_\psi(A) \leq \nu_\varphi(A)$ .

Let  $\phi = \lambda \cdot \varphi$ .

– If  $\lambda > 0$ ,

$$\begin{aligned} \lambda \nu_\varphi(A) &= \lambda \max_{i \in A} \varphi_i^+ + \lambda \min_{i \in \Omega} \varphi_i^- - \lambda \min_{i \in A^c} \varphi_i^-, = \max_{i \in A} \lambda \varphi_i^+ + \min_{i \in \Omega} \lambda \varphi_i^- - \min_{i \in A^c} \lambda \varphi_i^-, \\ &= \max_{i \in A} \phi_i^+ + \min_{i \in \Omega} \phi_i^- - \min_{i \in A^c} \phi_i^- = \nu_\phi(A). \end{aligned}$$

The same goes for  $\nu^c$ :  $\lambda\nu_\varphi^c(A) = \nu_\phi^c(A)$ .

– If  $\lambda < 0$ ,

$$\begin{aligned}\lambda\nu_\varphi(A) &= \lambda \max_{i \in A} \varphi_i^+ + \lambda \min_{i \in \Omega} \varphi_i^- - \lambda \min_{i \in A^c} \varphi_i^-, = \min_{i \in A} (\lambda \varphi_i^+) + \max_{i \in \Omega} (\lambda \varphi_i^-) - \max_{i \in A^c} (\lambda \varphi_i^-), \\ &= \min_{i \in A} \phi_i^- + \max_{i \in \Omega} \phi_i^+ - \max_{i \in A^c} \phi_i^+ = \nu_\phi^c(A).\end{aligned}$$

The same goes for  $\nu^c$ :  $\lambda\nu_\varphi^c(A) = \nu_\phi^c(A)$ .

Thus  $\nu_\varphi^c(A) \leq P_\psi(A) \leq \nu_\varphi(A) \Leftrightarrow \nu_\phi^c(A) \leq P_{\lambda.\psi(A)} \leq \nu_\phi(A)$ .  $\square$

**Definition 5.1.** Let  $\varphi \in \mathcal{K}(\Omega)$ , then its granularity is defined as:

- $\gamma(\varphi) = \frac{1}{\bar{\alpha}} \cdot \sum_{i \in \Omega} \max(0, \varphi_i) + \frac{1}{\underline{\alpha}} \sum_{i \in \Omega} \min(0, \varphi_i)$ , if  $\bar{\alpha} > 0$  and  $\underline{\alpha} < 0$ ,
- $\gamma(\varphi) = \frac{1}{\bar{\alpha}} \sum_{i \in \Omega} \max(0, \varphi_i)$ , if  $\bar{\alpha} > 0$  and  $\underline{\alpha} \geq 0$ ,
- $\gamma(\varphi) = \frac{1}{\underline{\alpha}} \sum_{i \in \Omega} \min(0, \varphi_i)$ , if  $\bar{\alpha} \leq 0$  and  $\underline{\alpha} > 0$ ,
- $\gamma(\varphi) = 0$ , if  $\bar{\alpha} = \underline{\alpha} = 0$ .

**Proposition 5.3.** Let  $\varphi, \phi \in \mathcal{K}(\Omega)$ , if  $\max_{i \in \Omega} \varphi_i = \max_{i \in \Omega} \phi_i$  and  $\min_{i \in \Omega} \varphi_i = \min_{i \in \Omega} \phi_i$  and  $\forall i \in \Omega$  we have  $\max(0, \varphi_i) \leq \max(0, \phi_i)$  and  $\min(0, \varphi_i) \geq \min(0, \phi_i)$ , then  $\mathcal{M}(\varphi) \subseteq \mathcal{M}(\phi)$  and  $\gamma(\varphi) \leq \gamma(\phi)$ .

*Proof.* As  $\forall i \in \Omega$ ,  $\max(0, \varphi_i) \leq \max(0, \phi_i)$  and  $\min(0, \varphi_i) \geq \min(0, \phi_i)$ , then we have:  $\forall A \subseteq \Omega$ ,  $\nu_\varphi^c(A) \geq \nu_\phi^c(A)$  and  $\nu_\varphi(A) \leq \nu_\phi(A)$ . Thus, if  $\psi \in \mathcal{M}(\varphi)$  then  $\psi \in \mathcal{M}(\phi)$ , i.e.  $\mathcal{M}(\varphi) \subseteq \mathcal{M}(\phi)$ .

Now, let  $\bar{\alpha} = \max_{i \in \Omega} \varphi_i = \max_{i \in \Omega} \phi_i$  and  $\underline{\alpha} = \min_{i \in \Omega} \varphi_i = \min_{i \in \Omega} \phi_i$ . Let us suppose that  $\bar{\alpha} > 0$  and  $\underline{\alpha} < 0$ . Since  $\sum_{i \in \Omega} \max(0, \varphi_i) \leq \sum_{i \in \Omega} \max(0, \phi_i)$  and  $\sum_{i \in \Omega} \min(0, \varphi_i) \geq \sum_{i \in \Omega} \min(0, \phi_i)$ , then  $\frac{1}{\bar{\alpha}} \sum_{i \in \Omega} \max(0, \varphi_i) + \frac{1}{\underline{\alpha}} \sum_{i \in \Omega} \min(0, \varphi_i) \leq \frac{1}{\bar{\alpha}} \sum_{i \in \Omega} \max(0, \phi_i) + \frac{1}{\underline{\alpha}} \sum_{i \in \Omega} \min(0, \phi_i)$  and thus  $\gamma(\varphi) \leq \gamma(\phi)$ . The cases where either  $\bar{\alpha} = 0$  or  $\underline{\alpha} = 0$  are trivial.  $\square$

**Proposition 5.4.** Let  $\psi \in \mathcal{K}(\Omega)$ , the kernel  $\hat{\varphi} \in \mathcal{K}(\Omega)$ , defined in Proposition 3.7, is the most specific kernel such that  $\psi \in \mathcal{M}(\hat{\varphi})$ , i.e.  $\forall \phi \in \mathcal{K}(\Omega)$ , if  $\psi \in \mathcal{M}(\phi)$  then  $\gamma(\phi) \geq \gamma(\hat{\varphi})$ .

*Proof.* Let  $\hat{\varphi}$  be the kernel defined by  $\forall i \in \Omega$ ,  $\hat{\varphi}_i = \sum_{j \in A_i^+} \max(0, \psi_j) + \sum_{j \in A_i^-} \min(0, \psi_j)$ , where  $A_i^+ = \{j \in \Omega / \psi_j^+ \leq \psi_i^+\}$  and  $A_i^- = \{j \in \Omega / \psi_j^- \geq \psi_i^-\}$ . Let  $B_i^+ = \{j \in \Omega / \psi_j > 0 \text{ and } \psi_j^+ \leq \psi_i^+\}$  and  $B_i^- = \{j \in \Omega / \psi_j < 0 \text{ and } \psi_j^- \geq \psi_i^-\}$ . We have  $\forall i \in \Omega$ ,  $\sum_{j \in B_i^+} \psi_j = \sum_{j \in A_i^+} \max(0, \psi_j)$  and  $\sum_{j \in B_i^-} \psi_j = \sum_{j \in A_i^-} \min(0, \psi_j)$  thus  $\hat{\varphi}_i = \sum_{j \in B_i^+} \psi_j + \sum_{j \in B_i^-} \psi_j$ .

By construction we have

$$\forall i \in \Omega, \nu_{\hat{\varphi}}(B_i^+) = \max_{j \in B_i^+} \hat{\varphi}_j^+ = \hat{\varphi}_i^+ \text{ since } \{j / \psi_j \leq 0\} \subseteq (B_i^+)^c \text{ and}$$

$$\forall i \in \Omega, \nu_{\hat{\varphi}}^c(B_i^-) = \min_{j \in B_i^-} \hat{\varphi}_j^- = \hat{\varphi}_i^- \text{ since } \{j / \psi_j \geq 0\} \subseteq (B_i^-)^c.$$

Now, let us suppose that  $\exists \phi \in \mathcal{K}(\Omega)$  such that  $\psi \in \mathcal{M}(\phi)$  and  $\gamma(\phi) < \gamma(\hat{\varphi})$ .

We know from Proposition 5.1 that since  $\psi \in \mathcal{M}(\phi)$  and  $\psi \in \mathcal{M}(\hat{\varphi})$ , then

$$\bar{\alpha} = \max_{i \in \Omega} \phi_i = \max_{i \in \Omega} \hat{\varphi}_i = \sum_{i \in \Omega} \max(0, \psi_i) \text{ and}$$

$$\underline{\alpha} = \min_{i \in \Omega} \phi_i = \min_{i \in \Omega} \hat{\varphi}_i = \sum_{i \in \Omega} \min(0, \psi_i).$$

The inequality  $\gamma(\phi) < \gamma(\varphi)$  is equivalent to

$$\frac{1}{\underline{\alpha}} \sum_{i \in \Omega} \max(0, \phi_i) + \frac{1}{\underline{\alpha}} \sum_{i \in \Omega} \min(0, \phi_i) < \frac{1}{\bar{\alpha}} \sum_{i \in \Omega} \max(0, \varphi_i) + \frac{1}{\bar{\alpha}} \sum_{i \in \Omega} \min(0, \varphi_i).$$

The above inequality is of the form  $a + b < c + d$  where  $a, b, c, d$  are positive real values. Hence we have either  $a < c$  or  $a \geq c$  which implies  $b < d$ . So within our context ( $\underline{\alpha} \leq 0$ ), we have either  $\sum_{i \in \Omega} \max(0, \varphi_i) > \sum_{i \in \Omega} \max(0, \phi_i)$  or  $\sum_{i \in \Omega} \min(0, \varphi_i) < \sum_{i \in \Omega} \min(0, \phi_i)$ .

– If  $\sum_{i \in \Omega} \max(0, \varphi_i) > \sum_{i \in \Omega} \max(0, \phi_i)$  then  $\exists k \in \Omega$  such that  $\max(0, \varphi_k) > \max(0, \phi_k)$ . This property is due to the structural separation of negative and positive values in the kernels (i.e.  $\forall i \in \Omega, \psi_i \leq 0 \Leftrightarrow \phi_i \leq 0 \Leftrightarrow \varphi_i \leq 0$  and conversely). As,  $\nu_{\varphi}(B_k^+) = \varphi_k^+ = \sum_{j \in A_k^+} \max(0, \psi_j)$  and  $\nu_{\phi}(B_k^+) = \phi_k^+$  we have  $\nu_{\varphi}(B_k^+) > \nu_{\phi}(B_k^+)$ . Consequently, we have  $\nu_{\phi}(B_k^+) < \sum_{i \in B_k^+} \psi_i$  thus  $\psi \notin \mathcal{M}(\phi)$ .

– If  $\sum_{i \in \Omega} \min(0, \varphi_i) \leq \sum_{i \in \Omega} \min(0, \phi_i)$ . The same goes when considering  $k \in \Omega$  such that  $\min(0, \varphi_k) \leq \min(0, \phi_k)$ , leading to the conclusion that  $\nu_{\phi}(B_k^-) \geq \sum_{i \in B_k^-} \psi_i$  thus  $\psi \notin \mathcal{M}(\phi)$ .

So by contraposition we can conclude that  $\psi \in \mathcal{M}(\phi) \Rightarrow \gamma(\phi) \geq \gamma(\varphi)$ .  $\square$

**Proposition 5.5.** *Let  $\psi \in \mathcal{K}(\Omega)$ , the kernel  $\varphi \in \mathcal{K}(\Omega)$ , defined by  $\forall i \in \Omega$  if  $\psi_i \geq 0$ ,  $\varphi_i = \sum_{k \in \Omega} \max(\psi_k, 0)$  else  $\varphi_i = \sum_{k \in \Omega} \min(\psi_k, 0)$ , is the least specific kernel such that  $\psi \in \mathcal{M}(\varphi)$ , i.e.  $\forall \phi \in \mathcal{K}(\Omega)$ , if  $\psi \in \mathcal{M}(\phi)$  then  $\gamma(\phi) \leq \gamma(\varphi)$ .*

*Proof.* Lets first prove that  $\psi \in \mathcal{M}(\varphi)$ .

Let  $\mathcal{P}(\Omega) = 2^\Omega$  be the power set of  $\Omega$  and let  $\mathcal{P}^+(\Omega) = \{A \subseteq \Omega | \forall i \in A \psi_i \geq 0\}$ ,  $\mathcal{P}^-(\Omega) = \{A \subseteq \Omega | \forall i \in A \psi_i < 0\}$ , and  $\mathcal{P}^\pm(\Omega) = \{A \subseteq \Omega | \exists i \in A \psi_i \geq 0 \text{ and } \exists j \in A \psi_j < 0\}$ . By construction, these three subsets form a partition of  $\mathcal{P}(\Omega)$ . Let  $A^+ = \{i \in \Omega / \psi_i > 0\}$ ,  $A^- = \{i \in \Omega / \psi_i < 0\}$  and  $A \subseteq \Omega$ . Let  $\underline{\alpha} = \sum_{k \in \Omega} \min(\psi_k, 0) = \min_{k \in \Omega} \varphi_k$  and  $\bar{\alpha} = \sum_{k \in \Omega} \max(\psi_k, 0) = \max_{k \in \Omega} \varphi_k$ .

– If  $A \in \mathcal{P}^\pm(\Omega)$  and  $A^c \in \mathcal{P}^\pm(\Omega)$  we have:

$$\nu_{\varphi}(A) = \max_{i \in A} \varphi_i^+ + \min_{i \in \Omega} \varphi_i^- - \min_{i \in A^c} \varphi_i^- = \bar{\alpha} + \underline{\alpha} - \underline{\alpha} = \bar{\alpha} \text{ and}$$

$$\nu_{\varphi}^c(A) = \min_{i \in A} \varphi_i^- + \max_{i \in \Omega} \varphi_i^+ - \max_{i \in A^c} \varphi_i^+ = \underline{\alpha} + \bar{\alpha} - \bar{\alpha} = \underline{\alpha}.$$

Thus, since by construction  $\underline{\alpha} \leq \sum_{i \in A} \psi_i \leq \bar{\alpha}$ ,  $\nu_{\varphi}^c(A) \leq P_{\psi}(A) \leq \nu_{\varphi}(A)$ .

– If  $A \in \mathcal{P}^-(\Omega)$  and  $A^c \in \mathcal{P}^\pm(\Omega)$  we have:  $\nu_{\varphi}(A) = \underline{\alpha} - \underline{\alpha} = 0$ ,  $\sum_{i \in A} \psi_i \leq 0$ .

Now, since  $A \in \mathcal{P}^-(\Omega)$ ,  $A^+ \subseteq A^c$ ,  $\nu_{\varphi}^c(A) = \underline{\alpha} + \bar{\alpha} - \bar{\alpha} = \underline{\alpha}$  and  $\sum_{i \in A} \psi_i \geq \underline{\alpha}$ .

Thus,  $\nu_{\varphi}^c(A) \leq P_{\psi}(A) \leq \nu_{\varphi}(A)$ .

– If  $A \in \mathcal{P}^+(\Omega)$  and  $A^c \in \mathcal{P}^\pm(\Omega)$  we have:  $\nu_{\varphi}(A) = \bar{\alpha}$ ,  $\sum_{i \in A} \psi_i \leq \bar{\alpha}$ .

Now, since  $A \in \mathcal{P}^+(\Omega)$ ,  $A^- \subseteq A^c$ ,  $\nu_{\varphi}^c(A) = \bar{\alpha} - \bar{\alpha} = 0$  and  $\sum_{i \in A} \psi_i \geq 0$ .

Thus,  $\nu_{\varphi}^c(A) \leq P_{\psi}(A) \leq \nu_{\varphi}(A)$ .

– If  $A \in \mathcal{P}^-(\Omega)$  and  $A^c \in \mathcal{P}^+(\Omega)$  we have:  $\nu_{\varphi}(A) = \underline{\alpha}$ ,  $\nu_{\varphi}^c(A) = \underline{\alpha}$  and

$\sum_{i \in A} \psi_i = \underline{\alpha}$ , thus  $\nu_{\varphi}^c(A) = P_{\psi}(A) = \nu_{\varphi}(A)$ .

Cases where  $(A \in \mathcal{P}^\pm(\Omega) \text{ and } A^c \in \mathcal{P}^+(\Omega))$  or  $(A \in \mathcal{P}^\pm(\Omega) \text{ and } A^c \in \mathcal{P}^-(\Omega))$  or  $(A \in \mathcal{P}^+(\Omega) \text{ and } A^c \in \mathcal{P}^-(\Omega))$  can be proved in the same way as above.

Case where  $A \in \mathcal{P}^+(\Omega)$  and  $A^c \in \mathcal{P}^+(\Omega)$  (which implies  $\forall i \in \Omega, \psi_i \geq 0$ ) as well as case where  $A \in \mathcal{P}^-(\Omega)$  and  $A^c \in \mathcal{P}^-(\Omega)$  (which implies  $\forall i \in \Omega, \psi_i \leq 0$ ) are trivial.

Thus  $\psi \in \mathcal{M}(\varphi)$ .

Now let us suppose that  $\exists \phi \in \mathcal{K}(\Omega)$  such that  $\psi \in \mathcal{M}(\phi)$  and  $\gamma(\phi) > \gamma(\varphi)$ . We know from Proposition 5.1 that with  $\psi \in \mathcal{M}(\phi)$  and  $\psi \in \mathcal{M}(\varphi)$  then:  $\bar{\alpha} = \max_{i \in \Omega} \phi_i = \max_{i \in \Omega} \varphi_i$  and  $\underline{\alpha} = \min_{i \in \Omega} \phi_i = \min_{i \in \Omega} \varphi_i$ . Therefore  $\gamma(\phi) > \gamma(\varphi)$  implies that  $\exists k \in \Omega$  such that  $\max(0, \phi_k) \geq \max(0, \varphi_k)$  (i.e.  $\phi_k \geq \bar{\alpha}$ ) or that  $\exists k \in \Omega$  such that  $\min(0, \phi_k) \leq \min(0, \varphi_k)$  (i.e.  $\phi_k \leq \underline{\alpha}$ ), which contradicts the hypothesis. Thus,  $\gamma(\phi) \leq \gamma(\varphi)$ .  $\square$

**Remark 6.**  $\forall \varphi \in \mathcal{K}(\Omega), \gamma(\varphi) \in [0, N]$ .

## 6. Computing the macsum aggregation

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

### 6.1. Formulæ

**Proposition 6.1.**

$$\begin{aligned} \bar{y} = \bar{F}(x) &= \check{C}_{\nu_\varphi}(x) = \sum_{k=1}^N x_{\lfloor k \rfloor} \cdot (\alpha_k - \alpha_{k-1}) + \sum_{k=1}^N x_{\lceil k \rceil} \cdot (\beta_k - \beta_{k-1}), \\ \underline{y} = \underline{F}(x) &= \check{C}_{\nu_\varphi}(x) = \sum_{k=1}^N x_{\lceil k \rceil} \cdot (\alpha_k - \alpha_{k-1}) + \sum_{k=1}^N x_{\lfloor k \rfloor} \cdot (\beta_k - \beta_{k-1}), \\ &\text{with } \alpha_k = \max_{i=N-k+1}^N \varphi_{\lfloor i \rfloor}^+, \text{ and } \beta_k = \min_{i=1}^k \varphi_{\lceil i \rceil}^-, \end{aligned}$$

where  $\lfloor \cdot \rfloor$  is the permutation that sorts the instances of  $x$  in increasing order and  $\lceil \cdot \rceil$  is the permutation that sorts the instances of  $x$  in decreasing order.

*Proof.* Let us first consider  $\bar{y}$ .

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

$$\nu_\varphi(A_{\lceil k \rceil}) = \max_{i \in A_{\lceil k \rceil}} \varphi_i^+ - \min_{i \in A_{\lceil k \rceil}^c} \varphi_i^- + \min_{i \in \Omega} \varphi_i^-,$$

$$A_{\lceil k \rceil} = \{\lceil k \rceil, \dots, \lceil N \rceil\} \text{ and } A_{\lceil k \rceil}^c = \{\lceil 1 \rceil, \dots, \lceil k-1 \rceil\}.$$

$$\text{Thus } \nu_\varphi(A_{\lceil k \rceil}) = \max_{i=k}^N \varphi_{\lceil i \rceil}^+ - \min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^- + \min_{i \in \Omega} \varphi_i^- \text{ with, by convention, } \min_{i=1}^0 \varphi_{\lceil i \rceil}^- = 0.$$

Therefore  $\bar{y}$  can be rewritten as:

$$\bar{y} = \check{C}_{\nu_\varphi}(x) = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) \cdot \left( \max_{i=k}^N \varphi_{\lceil i \rceil}^+ - \min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^- + \min_{i \in \Omega} \varphi_i^- \right).$$

Let us decompose this sum:

$$\bar{y} = \check{C}_{\nu_\varphi}(x) = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) \cdot \max_{i=k}^N \varphi_{\lceil i \rceil}^+ + \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) \cdot \min_{i \in \Omega} \varphi_i^- -$$

$$\sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) \cdot \min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^- = \bar{y}' + \bar{y}'', \text{ with } \bar{y}' = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) \cdot \max_{i=k}^N \varphi_{\lceil i \rceil}^+ \\ \text{and } \bar{y}'' = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) \cdot \min_{i \in \Omega} \varphi_i^- - \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) \cdot \min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^-.$$

First note that:  $x_{\lceil k \rceil} = x_{\lfloor N-k+1 \rfloor}$ .  
 So by making a change in index:  $k' = N - k + 1 \rightarrow k = N - k' + 1$  we get:  
 $\bar{y}' = \sum_{k'=1}^N (x_{\lfloor k' \rfloor} - x_{\lfloor k'+1 \rfloor}) \cdot \max_{i=N-k'+1}^N \varphi_{\lfloor i \rfloor}^+$ .  
 Now, let  $\alpha_k = \max_{i=N-k+1}^N \varphi_{\lfloor i \rfloor}^+$ , then:  
 $\bar{y}' = \sum_{k=1}^N (x_{\lfloor k \rfloor} - x_{\lfloor k+1 \rfloor}) \cdot \alpha_k = \sum_{k=1}^N x_{\lfloor k \rfloor} \cdot (\alpha_k - \alpha_{k-1})$ .

Second, note that:  $\sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) = x_{\lceil N \rceil}$ .  
 Thus,  $\bar{y}'' = x_{\lceil N \rceil} \cdot \min_{i \in \Omega} \varphi_i^- - \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k-1 \rceil}) \cdot \min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^-$ .

Let us now rearrange the expression of  $\bar{y}''$ .  
 $\bar{y}'' = x_{\lceil N \rceil} \cdot \min_{i=1}^N \varphi_{\lceil i \rceil}^- - x_{\lceil N \rceil} \cdot \min_{i=1}^{N-1} \varphi_{\lceil i \rceil}^- + x_{\lceil N-1 \rceil} \cdot \min_{i=1}^{N-1} \varphi_{\lceil i \rceil}^- - \dots + x_{\lceil 1 \rceil} \cdot \min_{i=1}^1 \varphi_{\lceil i \rceil}^-$ .  
 $\bar{y}'' = \sum_{k=1}^N x_{\lceil k \rceil} \cdot \left( \min_{i=1}^k \varphi_{\lceil i \rceil}^- - \min_{i=1}^{k-1} \varphi_{\lceil i \rceil}^- \right)$ .  
 Now let  $\beta_k = \min_{i=1}^k \varphi_{\lceil i \rceil}^-$ , then:  
 $\bar{y}'' = \sum_{k=1}^N (x_{\lceil k \rceil} - x_{\lceil k+1 \rceil}) \cdot \beta_k = \sum_{k=1}^N x_{\lceil k \rceil} \cdot (\beta_k - \beta_{k-1})$ .

Therefore,  $\bar{y} = \bar{y}' + \bar{y}'' = \sum_{k=1}^N x_{\lfloor k \rfloor} \cdot (\alpha_k - \alpha_{k-1}) + \sum_{k=1}^N x_{\lceil k \rceil} \cdot (\beta_k - \beta_{k-1})$ .

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

The values of  $\alpha_k$  and  $\beta_k$  can be computed recursively.

**Proposition 6.2.** *Let  $\alpha, \beta \in \mathbb{R}^N$  be defined by  $\forall k \in \Omega$ ,  $\alpha_k = \max_{i=N-k+1}^N \varphi_{\lfloor i \rfloor}^+$  and  $\beta_k = \min_{i=1}^k \varphi_{\lceil i \rceil}^-$ , then, let  $\alpha_0 = 0$  and  $\beta_0 = 0$ ,  $\alpha_k = \max(\alpha_{k-1}, \varphi_{\lfloor k \rfloor}^+)$  and  $\beta_k = \min(\beta_{k-1}, \varphi_{\lceil k \rceil}^-)$ .*

The proof is trivial and follows directly from the definition  $\alpha$  and  $\beta$ .

The algorithm for computing  $\bar{y}$  is provided in Algorithm 1, while the algorithm for computing  $\underline{y}$  is provided in Algorithm 2.

## 6.2. 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  $\bar{y}$ .

---

**Algorithm 1:** Computation of  $\bar{y}$ 

---

**Input:**  $\mathbf{x} = \{x_i\}_{i=1\dots N}$ ,  $\boldsymbol{\varphi} = \{\varphi_i\}_{i=1\dots N}$

**Output:**  $\bar{y}$

sort  $(\mathbf{x}, \boldsymbol{\varphi})$  w.r.t.  $\mathbf{x}$  in decreasing order ;

$\alpha = 0$ ,  $\beta = 0$  ;

$\bar{y} = 0$  ;

**for**  $k = 1 \dots N$  **do**

$\beta = \alpha$  ;  
     $\alpha = \max(\alpha, \varphi_k)$  ;  
     $\bar{y} = \bar{y} + (\alpha - \beta) \cdot x_k$  ;

sort  $(\mathbf{x}, \boldsymbol{\varphi})$  w.r.t.  $\mathbf{x}$  in increasing order (i.e. reverse the sorting) ;

$\alpha = 0$ ,  $\beta = 0$  ;

**for**  $k = 1 \dots N$  **do**

$\beta = \alpha$  ;  
     $\alpha = \min(\alpha, \varphi_k)$  ;  
     $\bar{y} = \bar{y} + (\alpha - \beta) \cdot x_k$  ;

---

---

**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**

$\beta = \alpha$  ;  
     $\alpha = \max(\alpha, \varphi_k)$  ;  
     $\underline{y} = \underline{y} + (\alpha - \beta) \cdot x_k$  ;

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**

$\beta = \alpha$  ;  
     $\alpha = \min(\alpha, \varphi_k)$  ;  
     $\underline{y} = \underline{y} + (\alpha - \beta) \cdot x_k$  ;

---

## 7. Experiment

In this section, we propose to illustrate the properties of the macsum operator via an experiment that consists of deriving a digital signal. A digital signal is generally a continuous signal that has been sampled and quantified (i.e. mapped onto a countable set of values). Being sampled, a digital signal is not continuous and thus cannot be derived. Therefore, deriving a digital signal must be understood as estimating, at each sampling location, the value of the 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. The reconstruction step is essential since it enables us to transform a quantized sampled integer-valued signal into a real-valued continuous signal. It is performed by convoluting the original digital signal with a (summative) kernel that allows us to estimate the value of the real signal at any location outside the sampling locations.

As explained by J. Shen and S. Castan in [42] 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 the chosen 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  $\kappa$  defined by:

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

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

A conventional Dirac sampling of this reconstruction kernel leads to a discrete kernel  $\rho$  defined by:

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

with  $\frac{1 - \beta}{1 + \beta}$  being a normalisation factor that makes the discrete kernel  $\rho$  summative in the sense of Loquin [31], i.e.  $\lim_{K \rightarrow \infty} \sum_{k=-K}^K \rho_k = 1$  (see [42]) since sampling a summative kernel should lead to a discrete summative kernel.

Deriving the kernel  $\kappa$  leads to the continuous kernel  $\eta$  defined by:

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

And therefore, sampling this derivation leads to the derivation kernel  $\psi$  defined by:

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

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

Let us now consider computing  $y$  the discrete signal that is the derivative of  $x$ . This derivation can be obtained by convoluting  $x$  with  $\psi$ :  $y = x \star \psi$ ,  $\star$  being the convolution. This operation can be written [29]:

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

$\psi^n$  being the kernel  $\psi$  translated in  $n$  and truncated on  $\Omega$  defined by:  $\psi_k^n = \psi_{n-k}$ ,  $k \in \{1, \dots, N\}$ . 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. This may sound strange in the decision theory context.

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  $\beta$ . This choice is generally made in a rather arbitrary way or could be based on visual analyses (or by computers) of results (of segmentation, for example) on the basis of a bank of representative signals of the problem to solve.

First note that the granularity (in the sense of Loquin or Shannon) of the Shen-Castan kernel increases with  $\beta$ , i.e. the bigger  $\beta$ , the smoother the reconstructed continuous signal. On the other hand, the smoother the reconstruction, the less sensitive the algorithm will be to fast variations in the signal.

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:

$$s(t) = \sin(21.t) + \sin(12.t + 5) + \sin(-4.t + 3) + 10.(1 + t)^{-1}.$$

The digital signal  $\{x_k\}_{k=1 \dots N}$  resulting from sampling is pictured in Figure (1).

The imprecision of the interval-valued derivatives can be used to define thresholds for detecting local maxima, like in [22] or use a posteriori criteria, other than frequency criteria, to define the optimal candidate to be the derivation of  $x$  (see e.g. [19]).

Now, let us assume that some previous experiments have shown that  $\beta = 0.8$  can be considered as a choice that achieves a good compromise between the quality of the reconstruction obtained by the Shen-Castan kernel and the sensitivity of the derivation algorithm. To account for the arbitrary aspect of this choice, we propose to calculate an imprecise derivative of this signal by using three approaches: the macsum approach, proposed in this paper, the transfer of variance used to compute uncertainty in measurement [35, 1] and an interval-valued method based on a Monte Carlo approach [43].

- For the macsum approach, we use Proposition 3.7 to define, for each  $n \in \Omega$ ,  $\hat{\varphi}^n$ , the kernel with the smallest granularity that dominates  $\psi^n$  (i.e.  $\psi^n \in \mathcal{M}(\hat{\varphi}^n)$ ) while using Proposition 3.6 to define  $\check{\varphi}^n$  that weakly dominates  $\psi^n$ .

- For the variance transfer approach, we suppose that  $\beta$  is a Gaussian random variable of mean 0.8 and variance 0.004 (a bigger value would lead to too imprecise signals) and we compute, for each value of the output signal, the variance of its estimate; we consider the one-sigma interval.
- For the Monte Carlo based approach, we suppose that the value of  $\beta$  comes within the  $[0.7, 0.9]$  interval and we compute both maximum and minimum values by randomly choosing 1000 values of  $\beta$  in this interval.

There is no question here of considering these three approaches in a competitive sense (which would not be logical), but the idea is simply to gain insight into the purpose of each method.

Figures (2), (3) and (4) show the results of this experiment. Details of those Figures are plotted in Figure (5). On each figure is plotted (in black) the result of deriving the original signal  $x(t)$  by using a precise kernel with  $\beta = 0.8$ . For the interval-valued signal, the upper values  $\{\bar{y}_n\}_{n=1\dots N}$  are plotted in blue, while the lower values  $\{\underline{y}_n\}_{n=1\dots N}$  are in red. The *mean imprecision* of the interval valued signal is actually the mean average value of the difference between  $\bar{y}_n$  and  $\underline{y}_n$ .

On Figure (2) the results of the macsum approach using  $\hat{\varphi}$ , (solid lines) and  $\check{\varphi}$  (dashed lines) are superimposed. Since, by construction  $\psi \in \mathcal{M}(\hat{\varphi}) \subseteq \mathcal{M}(\check{\varphi})$ ,  $\psi \in \mathcal{M}(\hat{\varphi}) \subseteq \mathcal{M}(\check{\varphi})$ ,  $F_{\psi_k^n}(x) \in [F_{\hat{\varphi}_k^n}(x), F_{\check{\varphi}_k^n}(x)] \subseteq [F_{\check{\varphi}_k^n}(x), F_{\hat{\varphi}_k^n}(x)]$ , i.e. the output using  $\hat{\varphi}$  is included in that using  $\check{\varphi}$ . We also have  $\gamma(\hat{\varphi}) = 10 \leq \gamma(\check{\varphi}) = 18$ . This illustrates that the higher the granularity of the kernel, the less precise the interval-valued reconstructed signal. The mean imprecision of the reconstructed imprecise signal is 1.06 for the strong domination approach and 1.89 for the weak domination approach. It is interesting to note that the ratio of mean imprecision (weak/strong) is close to the ratio of granularity of the used kernels. This would warrant further investigation.

On Figure (3) the results of the variance transfer approach are plotted. The variance transfer consists of computing, for each  $k, n \in \Omega$ , the value of  $\frac{\partial \psi_k^n}{\partial \beta}$  and using it to transfer the variance of  $\beta$  (here  $\sigma_\beta^2 = 0.004$ ) to the variance of  $\psi_k^n$  then to the variance of  $y_n$ . The variance transfer method leads to very qualitatively different results. Indeed, by hypothesis, we reconstruct an imprecise signal whose upper and lower bounds are symmetrical with respect to the precise signal. This is easier to compare when looking at Figure (5).a and Figure (5).b. In Figure (5).a the higher the value of the precise signal, the closer it is to the upper bound of the imprecise signal (and reverse for the lower bound). The macsum approach better accounts for the dissymmetry of the variation in the reconstructed value when there is a variation in the  $\beta$  parameter. Moreover, the mean imprecision within this approach is 1.94, although the variance of the  $\beta$  parameter is assumed to be very small.

On Figure (4) the results of the Monte Carlo approach are plotted. The mean imprecision within this approach is 0.70, i.e. the convex set of reconstructed values is much more specific than the two previous methods. Note, by comparing Figure (5).a and Figure (5).c, that the asymmetry of the interval-

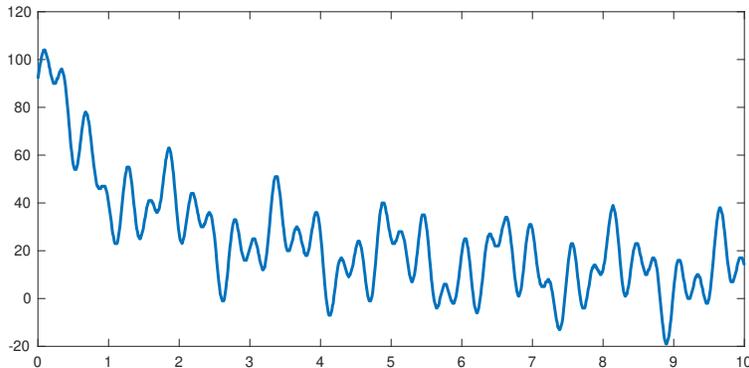


Figure 1: The signal to be derivated.

valued reconstructed signal with respect to the precise signal is very similar to what we obtain with the macsum method. Moreover, all signals obtained using the Monte Carlo approach are within the interval obtained by the macsum approach with the strong domination (and thus also with the weak domination).

The comparison between those three approaches is rather limited. In fact, each method has different endpoints with different advantages and drawbacks. With the macsum approach, the granularity of the kernel cannot be completely controlled, thus making it difficult to control the specificity of the reconstructed imprecise signal. Whereas with the two other approaches, the definition of  $\sigma_{\beta}^2$  for the transfer approach and  $[\underline{\beta}, \overline{\beta}]$  for the Monte Carlo approach allows us to get as close as required to the value obtained by using the precise approach. However, the macsum approach has the advantage of having a reduced computation time compared to the Monte Carlo approach, and a controlled and less arbitrary specificity compared to the variance transfer approach ( $\sigma_{\beta}^2 = 0.01$  would result in a completely meaningless result). In addition, the macsum approach computes the exact requested interval, while the Monte Carlo approach only gives an interior approximation of this interval. Moreover, the macsum approach requires  $N$  parameters (as with the precise approach) while the other approaches requires  $2.N$  values ( $(\psi_k, \sigma_{\psi_k})$  for variance transfer and  $(\underline{\psi}_k, \overline{\psi}_k)$  for Monte Carlo approach).

## 8. 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

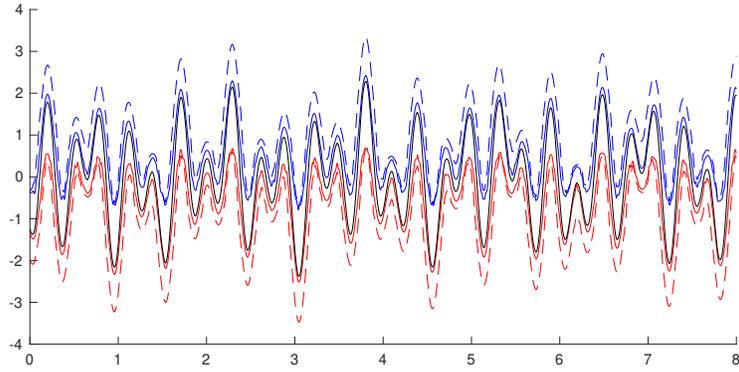


Figure 2: Precise derivative (black) and imprecise derivatives (blue upper, red lower) of the signal depicted in Figure (1) with  $\beta = 0.8$  obtained with  $\tilde{\varphi}$  (solid) and  $\check{\varphi}$  (dahed).

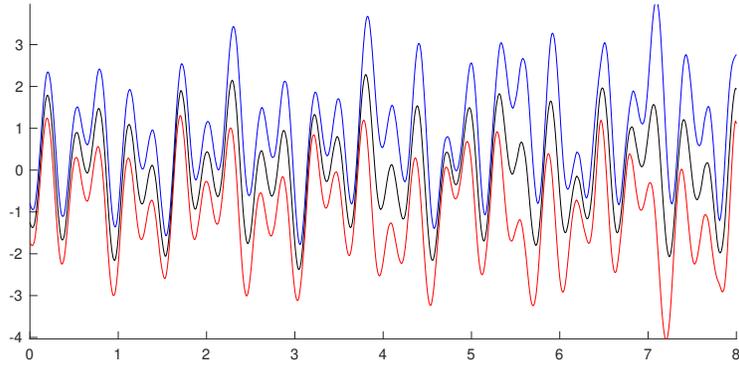


Figure 3: Precise derivative (black) and imprecise derivative (blue upper, red lower) of the signal depicted in Figure (1) with  $\beta = 0.8$  obtained by transferring a  $\sigma_{\beta}^2 = 4.10^{-6}$  variance of  $\beta$  to the variance of the output.

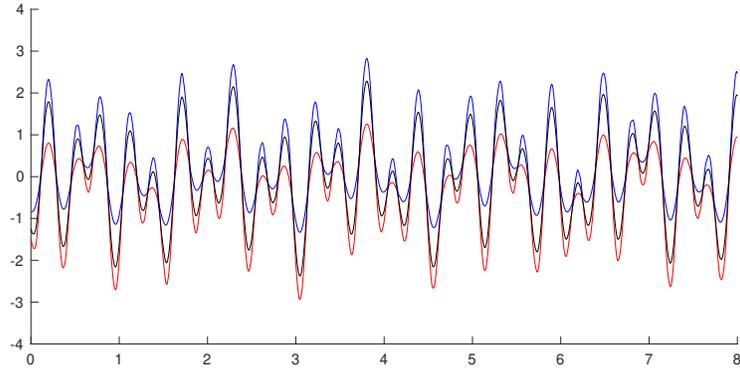


Figure 4: Precise derivative (black) and imprecise derivative (blue upper, red lower) of the signal depicted in Figure (1) with  $\beta = 0.8$  for the precise derivative and  $\beta \in [0.7, 0.9]$  for the imprecise derivative obtained by using a Monte Carlo simulation.

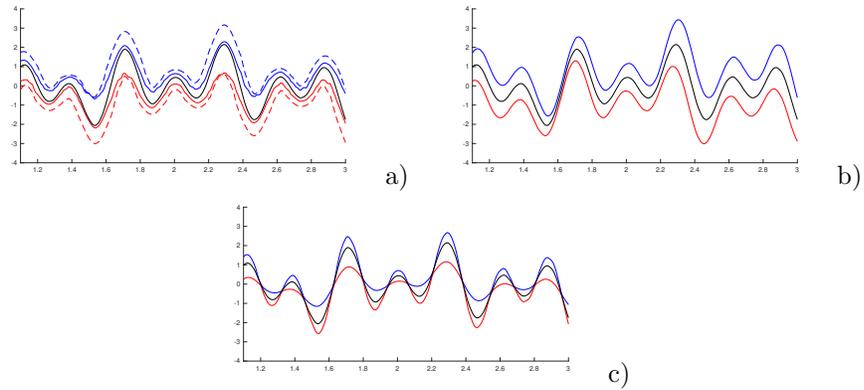


Figure 5: Details of Figure (2) (a), Figure (3) (b) and Figure (4) (c).

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 predict or control this approximation.

The approach we have proposed in this article exploits an interesting new avenue. Instead of trying to describe a system precisely, we propose to represent it by imprecise linear modeling, i.e. a convex set of linear models. Within this model, a system is associated with a set of weights, as in a linear model, but we obtain the description of a convex set of linear relations. One major limitation of this approach is that, due to the simplicity of the modeling, it may be hard to control the extent to which the interval-valued output is close to the output of the system to be represented. However, thanks to the granularity concept, we can know *a priori* the accuracy of the output of a macsum model with two different kernels.

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, if needed, to get a more accurate approximation while keeping the simplicity of the model.

## References

- [1] K. Arras. Experimental uncertainty estimation and statistics for data having interval uncertainty. Unlimited release, Sandia National Laboratories, 1998. Technical Report Nr. EPFL-ASL-TR-98-01 R3, Autonomous Systems Lab, EPFL.
- [2] R. McCullough B. Mitchell1, R. Ekey and W. Reitz. A fantastic quantitative exploration of ohm’s law. *The Physics Teacher*, pages 56–75, 2018.
- [3] R. Boukezzoula, S. Galichet, and A. Bissierier. A midpoint-radius approach to regression with interval data. *International Journal of Approximate Reasoning*, 52:1257–1271, 2011.
- [4] R. Boukezzoula, S. Galichet, and D. Coquin. From fuzzy regression to gradual regression: Interval-based analysis and extensions. *Information Sciences*, 441:18–40, 2018.

- [5] 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.
- [6] J. Canongia Lopes. Simulation and modeling in computational chemistry: A molecular portfolio. In Manuel Seabra Pereira, editor, *A Portrait of State-of-the-Art Research at the Technical University of Lisbon*, pages 121–135, Dordrecht, 2007. Springer Netherlands.
- [7] 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.
- [8] L. Coppock and D. Mateer. *Principles of Microeconomics*. New York : W.W. Norton & Company, 2017.
- [9] 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.
- [10] J. Hunter D. Jorgenson and M. Nadiri. The predictive performance of econometric models of quarterly investment behavior. *Econometrica*, 38:213–224, 1970.
- [11] D Denneberg. *Non-additive Measure and Integral*. Kluwer Academic Publishers, Dordrecht, 1994.
- [12] 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.
- [13] D. Dubois and H. Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Springer, 1988.
- [14] K. Loquin F. Jacquy, F. Comby, and O. Strauss. Non-additive approach for gradient-based detection. In *ICIP: International Conference on Image Processing '07*, pages 49–52, San Antonio, Texas, 2007.
- [15] F. Graba, K. Loquin, F. Comby, and O. Strauss. Guaranteed reconstruction for image super-resolution. In *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–8. IEEE, 2013.
- [16] M. Grabisch. Fuzzy measures and integrals: recent developments. *Fifty years of fuzzy logic and its applications*, pages 125–151, 2015.
- [17] J. Hall and J. Lawry. Generation, combination and extension of random set approximations to coherent lower and upper probabilities. *Reliability Engineering and System Safety*, 85(1-3):89–101, 2004.

- [18] 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.
- [19] V. Itier, F. Kucharczak, O. Strauss, and W. Puech. Interval-valued jpeg decompression for artifact suppression. In *2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 1–6, 2018.
- [20] D. Chen J. Mao and L. Zhang. Mechanical assembly quality prediction method based on state space model. *The International Journal of Advanced Manufacturing Technology*, 86:107–116, 2016.
- [21] A. O’Hagan J. Oakley. Probabilistic sensitivity analysis of complex models: A bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66:751 – 769, 08 2004.
- [22] F. Jacquy, F. Comby, and O. Strauss. Fuzzy edge detection for omnidirectional images. *Fuzzy Sets and Systems*, 159(15):1991–2010, 2008.
- [23] 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.
- [24] 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.
- [25] L. Jaulin and E. Walter. Set inversion via interval analysis for nonlinear bounded-error estimation. *Automatica (29)*, pages 1053–1064, 1996.
- [26] G. Kleiter. Propagating imprecise probabilities in bayesian networks. *Artificial Intelligence (88)*, pages 143–161, 1996.
- [27] 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.
- [28] K. Loquin and O. Strauss. Histogram density estimators based upon a fuzzy partition. *Statistics and Probability Letters*, 78(13):1863–1868, September 2008.
- [29] K. Loquin and O. Strauss. On the granularity of summative kernels. *Fuzzy Sets and Systems*, 159(15):1952–1972, August 2008.
- [30] 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.

- [31] 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.
- [32] 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.
- [33] M. Sugeno M. Grabisch and T. Murofushi. Fuzzy measures and integrals: theory and applications. *Heidelberg: Physica*, 2000.
- [34] M. Troffaes and G. De Cooman. *Lower previsions*. Wiley series in probability and statistics. John Wiley and Sons, 2014.
- [35] Gilles Mauris. Possibility distributions: A unified representation of usual direct-probability-based parameter estimation methods. *International Journal of Approximate Reasoning*, 52(9):1232–1242, 2011.
- [36] Y. Narukawa and V. Torra. Non-monotonic fuzzy measures and intuitionistic fuzzy sets. In V. Torra, Y. Narukawa, A. Valls, and J. Domingo-Ferrer, editors, *Modeling Decisions for Artificial Intelligence*, pages 150–160. Springer Berlin Heidelberg, 2006.
- [37] Z. Pawlak. *Rough sets: Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publisher, 1991.
- [38] A. Rico and O. Strauss. Imprecise expectations for imprecise linear filtering. *International Journal of Approximate Reasoning*, 51(8):933–947, October 2010.
- [39] 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.
- [40] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.
- [41] P. Shary. *Interval Regularization for Inaccurate Linear Algebraic Equations*, pages 375–398. Springer International Publishing, Cham, 2020.
- [42] J. Shen and S. Castan. Towards the Unification of Band-limited Derivative Operators for Edge Detection. *Signal Processing*, 31(2):103–119, 1993.
- [43] I. Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55(1):271–280, 2001.
- [44] O. Strauss, K. Loquin, and F. Kucharczak. *On Maxitive Image Processing*, pages 201–215. Springer International Publishing, Cham, 2021.

- [45] O. Strauss and A. Rico. Where the domination of maxitive kernels is extended to signed values. In *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–6, 2019.
- [46] M. Sugeno T. Murofushi and M. Machida. Non-monotonic fuzzy measures and the choquet integral. *Fuzzy Sets and Systems*, 64(1):73–86, 1994.
- [47] 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.
- [48] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall, 1991.
- [49] R. Yager and V. Kreinovich. Decision making under interval probabilities. *International Journal of Approximate Reasoning*, 22:195–215, 1999.