Approximating general kernels by extended fuzzy measures: application to filtering

Sébastien Destercke, Agnès Rico, Olivier Strauss

Approximating general kernels by extended fuzzy measures: application to filtering

Sébastien Destercke[1], Agnès Rico[2], and Olivier Strauss[3]

[1] Sorbonne universités, UTC, CNRS, Heudiasyc, 57 Avenue de Landshut, Compiègne
sebastien.destercke@hds.utc.fr
[2] ERIC – Université Claude Bernard Lyon 1, Lyon, France
agnes.rico@univ-lyon1.fr
[3] LIRMM – Université de Montpellier, Montpellier, France
olivier.strauss@lirmm.fr

Abstract. Convolution kernels are essential tools in signal processing: they are used to filter noisy signal, interpolate discrete signals, ... However, in a given application, it is often hard to select an optimal shape of the kernel. This is why, in practice, it may be useful to possess efficient tools to perform a robustness analysis, talking the form in our case of an imprecise convolution. When convolution kernels are positive, their formal equivalence with probability distributions allows one to use imprecise probability theory to achieve such an imprecise convolution. However, many kernels can have negative values, in which case the previous equivalence does not hold anymore. Yet, we show mathematically in this paper that, while the formal equivalence is lost, the computational tools used to describe sets of probabilities by intervals on the singletons still retain their key properties when used to approximate sets of (possibly) non-positive kernels. We then illustrate their use on a single application that consists of filtering a human electrocardiogram signal by using a low-pass filter whose order is imprecisely known. We show, in this experiment, that the proposed approach leads to tighter bounds than previously proposed approaches.

Keywords: Signal filtering · Probability intervals · Signed fuzzy measures · Interval-valued filtering

1 Introduction

Filtering a signal aims at removing some unwanted components or features, i.e. other signals or measurement noise. It is a common problem in both digital analysis and signal processing [5]. In this context, kernels are used for impulse response modelling, interpolation, linear and non-linear transformations, stochastic or band-pass filtering, etc. However, how to choose a particular kernel and its parameters to filter a given signal is often a tricky question.

A way to circumvent this difficulty is to filter the signal with a convex set of kernels, thus ending up with a set-valued signal, often summarized by lower/upper bounds. The set of kernels has to be convex due to the fact that, if two kernels are suitable to achieve the filtering, a combination of those two kernel should be suitable too. A key problem is

then to propose a model to approximate this convex set of kernels in a reasonable way (i.e., without losing too much information) that will perform the set-valued filtering in an algorithmic efficient way guaranteeing the provided bounds (in the sense that the sets of signals obtained by filtering with each kernel of the set is contained within the bounds).

In the case of positive summative kernels, i.e., positive functions summing up to one, previous works used the formal equivalence between such kernels and probabilities at their advantage, and have proposed to use well-known probability set models as approximation tools. For example, maxitive [6] and cloudy [4] kernels respectively use possibility distributions and generalized p-boxes to model sets of kernels, and have used the properties of the induced lower measure on events to propose efficient filtering solutions from a computational standpoint.

However, when the kernel set to approximate contains functions that are not positive everywhere (but still sum up to one), this formal analogy is lost, and imprecise probabilistic tools cannot be used straightforwardly to model the set of kernels. Yet, recent works [9] have shown that in some cases such imprecise models can be meaningfully extended to accommodate negative values, while preserving the properties that makes them interesting for signal filtering (i.e., the guarantees of obtained bounds and the algorithmic efficiency). More formally, this means that we have to study the extension of Choquet-integral based digital filtering to the situation where kernels $\kappa : \mathscr{X} \to [-A,B] \subseteq \mathbb{R}$ can be any (bounded) function.

In this work, we show that this is also true for another popular model, namely probability intervals [1], that consists in providing lower/upper bound on singleton probabilities. In particular, while principle applied to sets of additive but possibly negative measures lead to a model inducing a non-monotone set function, called *signed measure*, we show that the Choquet integral of such a measure still leads to interesting bounds for the filtered signal, in the sense that these bounds are guaranteed and are obtained for specific additive measure dominated by the signed measure. Let us call this new kind of interval-valued kernels *imprecise kernels*.

The paper is structured as follows. Section 2 recalls the setting we consider, as well as a few preliminaries. We demonstrate in Section 3 that probability intervals can be meaningfully extended to accommodate sets of additive but non-positive measures. Section 4 shows how these results can be applied to numerical signal filtering.

2 Preliminaries: filtering, signed kernels and fuzzy measures

We assume a finite space $\mathscr{X} = \{x_1, \ldots, x_n\}$ of n points, that is a subset of an infinite discrete space Ω that may be a discretization of a continuous space (e.g., the real line), and an observed signal whose values at these points are $f(x_1), \ldots, f(x_n)$, that can represent a time- or space-dependent record (EEG signal, sound, etc.). A kernel is here a bounded discrete function $\eta : \Omega \to [-A,B]$, often computed from a continuous kernel (corresponding, for example, to assumed filter impulse response). This kernel is such that $\sum_{x\in\Omega} \eta(x) = 1$, and for a given kernel we will denote b_η and a_η the sum of the

positive and negative parts of the kernel, respectively. That is:

$$b_\eta = \sum_{x\in\Omega} \max(0,\eta(x)) \quad a_\eta = \sum_{x\in\Omega} \max(0,-\eta(x))$$

Filtering the signal f by using the kernel η consists of estimating the filtered signal $\hat{f}$ at each point x of $\mathscr{X}$ by:

$$\hat{f}(x) = \sum_{i=1}^{n} f(x_i)\eta(x_i - x) = \sum_{i=1}^{n} f(x_i)\eta^x(x_i), \tag{1}$$

where $\forall y \in \mathscr{X}$, $\eta^x(y) = \eta(y-x)$.

Since filtering f amounts to compute the value of $\hat{f}$ at each point of $\mathscr{X}$, let us simplify the previous formal statements by assuming that, at each point x of $\mathscr{X}$ exists a kernel $\kappa = \eta^x$. Note that the domain of κ can be restricted to $\mathscr{X}$ without any loss of generality. Moreover, $b_\kappa = \sum_{x\in\Omega} \max(0,\kappa(x)) = b_\eta$ and $a_\kappa = \sum_{x\in\Omega} \max(0,-\kappa(x)) = a_\eta$.

From any kernel κ, we can build a set function $\mu_\kappa : 2^{\mathscr{X}} \to [-a_\kappa, b_\kappa]$ such that $\mu_\kappa(A) = \sum_{x\in A} \kappa(x)$ for any $A \subseteq \mathscr{X}$. It is clear that we have $\mu_\kappa(\mathscr{X}) = 1$ and $\mu_\kappa(\emptyset) = 0$, but that we can have $\mu_\kappa(A) > \mu_\kappa(B)$ with $A \subset B$ (simply consider $A = \emptyset$ and $B = x$ where $\kappa(x) < 0$). Note also that if $A \cap B = \emptyset$, we keep $\mu_\kappa(A \cup B) = \mu_\kappa(A) + \mu_\kappa(B)$, hence the additivity of the measure.

Estimating the value of the signal $\hat{f}$ in a given point $x \in \mathscr{X}$ requires to compute a value $\mathscr{C}_\kappa$ that can be written as a weighted sum $\mathscr{C}_\kappa(f) = \sum_{i=1}^{n} f(x_i)\kappa(x_i)$. If we order and rank the values of f such that $f(x_{(1)}) \leq f(x_{(2)}) \leq \ldots \leq f(x_{(n)})$, this weighted sum can be rewritten

$$\hat{f}(x) = \mathscr{C}_\kappa(f) = \sum_{i=1}^{n} \left(f(x_{(i)}) - f(x_{(i-1)})\right) \mu_\kappa(A_{(i)}) \tag{2}$$

where $A_{(i)} = \{x_{(i)}, \ldots, x_{(n)}\}$, $f(x_{(0)}) = 0$ and $A_{(1)} = \mathscr{X}$. One can already notice the similarity with the usual Choquet integral.

Example 1. Consider a kernel η that is a Hermite polynomial of degree 3, and its sampled version pictured in Figure 1, with $a = -2$. From the picture, it is obvious that some of the values are negative.

$$\eta(x) = \begin{cases} (a+2)|x|^3 - (a+3)|x|^2 + 1 & \text{if } |x| \leq 1 \\ a|x|^3 - 5a|x|^2 + 8a|x| - 4a & \text{if } 1 \leq |x| \leq 2 \\ 0 & \text{else} \end{cases}$$

In this paper, we consider the case where the ideal η is ill-known, that is we only know that, for each $x \in \mathscr{X}$, η belongs to a convex set $\mathscr{N}$, which entails that κ belongs to a convex set $\mathscr{K}$. This set $\mathscr{N}$ (and $\mathscr{K}$) can reflect, for example, our uncertainty about which kernel should be ideally used (they can vary in shape, bandwidth, etc.). In the following section, we propose to approximate this set $\mathscr{K}$ of kernels that we deem suitable to filter f at point x by a measure $\underline{\mu}_{\mathscr{K}}$ that is close to a fuzzy measure: it

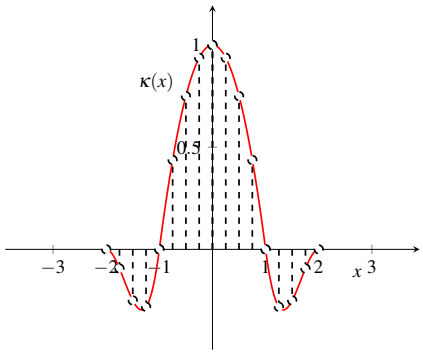

Fig. 1. Sampled Hermite polynomial of degree 3

shares with such measures the fact that $\underline{\mu}_{\mathscr{K}}(\emptyset) = 0$ and $\underline{\mu}_{\mathscr{K}}(\mathscr{X}) = 1$, and will be a standard fuzzy measure in the case where all $\kappa \in \mathscr{K}$ are positive. However, in the case of negative $\kappa \in \mathscr{K}$, we may have $\underline{\mu}_{\mathscr{K}}(A) < 0$ for some A, as well as have for some couple $A \subset B$ non-monotonicity in the form $\underline{\mu}_{\mathscr{K}}(A) > \underline{\mu}_{\mathscr{K}}(B)$.

3 Approximating set of kernels with extended probability intervals

We now consider a set $\mathscr{K}$ of kernels κ defined on $\mathscr{X}$, that can be discretised versions of a set of continuous kernels. We make no assumptions about this set, except for the fact that each kernel $\kappa \in \mathscr{K}$ is bounded and such that $\sum_{x \in \mathscr{X}} \kappa(x) = 1$. If $\mathscr{K}$ contains a large amount of kernels, filtering the signal with each of them and getting the set answer $\{\mathscr{C}_\kappa | \kappa \in \mathscr{K}\}$ for each possible point of the signal can be untractable. Rather than doing so, we may search some efficient way to find some lower and upper bounds of $\{\mathscr{C}_\kappa | \kappa \in \mathscr{K}\}$ that are not too wide.

To achieve such a task, we propose here to use tools inspired by the imprecise probabilistic literature, namely non-additive set functions and the Choquet integral. To use such tools, we must first build a set-function approximating the set $\mathscr{K}$. To do so, we will extend probabilistic intervals [1] that consist in associating lower and upper bounds $[l(x), u(x)]$ to each atom, given a set of probabilities. In our case, for each elements $x \in \mathscr{X}$, we consider the interval-valued kernel

$$\overline{\underline{\rho}}(x) = [\underline{\rho}(x), \overline{\rho}(x)] \tag{3}$$

such that the bounds are given, for each $x \in \mathscr{X}$, as

$$\underline{\rho}(x) = \inf_{\kappa \in \mathscr{K}} \kappa(x) \text{ and } \overline{\rho}(x) = \sup_{\kappa \in \mathscr{K}} \kappa(x). \tag{4}$$

Clearly, both can be negative and are not classical probability intervals. Nevertheless, we will show that set-functions induced by these bounds enjoy properties similar to those of standard probability intervals, and hence can be used to efficiently approximate $\{\mathscr{C}_\kappa | \kappa \in \mathscr{K}\}$.

From the imprecise kernel $\overline{\rho}$, we propose to build the set function $\underline{\mu} : \mathscr{X} \to \mathbb{R}$ such that, for $A \subseteq \mathscr{X}$

$$\underline{\mu}_{\mathscr{K}}(A) = \max\{\sum_{x\in A} \underline{\rho}(x), 1 - \sum_{x\notin A} \overline{\rho}(x)\}, \tag{5}$$

which is formally the same equation as the one used for probability intervals. We still have that $\underline{\mu}(\emptyset) = 0$ and $\underline{\mu}(\mathscr{X}) = 1$ (as $\underline{\rho}(x) \leq \kappa(x)$ for any $\mu \in \mathscr{K}$, we necessarily have $\sum_{x\in\mathscr{X}} \underline{\rho}(x) \leq 1$). However, we can have $\underline{\mu}(A) > \underline{\mu}(B)$ with $A \subset B$, meaning that $\underline{\mu}$ is not a classical fuzzy measure, and not a so-called coherent lower probability[4]. It is also non-additive, as we have $\underline{\mu}(A \cup B) \neq \underline{\mu}(A) + \underline{\mu}(B)$ for $A \cap B = \emptyset$ in general.

Simply replacing μ_κ by $\underline{\mu}_{\mathscr{K}}$ in Equation 2 gives us

$$\underline{\mathscr{C}}_{\mathscr{K}}(f) = \sum_{i=1}^{n} \left(f(x_{(i)}) - f(x_{(i-1)})\right) \underline{\mu}_{\mathscr{K}}(A_{(i)}) \tag{6}$$

$$= \sum_{i=1}^{n} f(x_{(i)}) \left(\underline{\mu}(A_{(i)}) - \underline{\mu}(A_{(i+1)})\right), \tag{7}$$

where $A_{(n+1)} = \emptyset$ and $f(x_0) = 0$.

In the positive case, $\underline{\mathscr{C}}_{\mathscr{K}}(f)$ is well-known to outer-approximate $\{\mathscr{C}_\kappa | \kappa \in \mathscr{K}\}$, as $\underline{\mathscr{C}}_{\mathscr{K}}(f) = \inf_{\mu_\kappa \geq \underline{\mu}_{\mathscr{K}}} \mathscr{C}_\kappa(f)$. In the rest of this section, we show that this is also true when the set $\mathscr{K}$ contains non-positive functions. In particular, we will show:

1. that Equation (6) still provides a lower bound of $\inf_{\mu_\kappa \geq \underline{\mu}_{\mathscr{K}}} \mathscr{C}_\kappa(f)$ and,
2. that this lower bound corresponds to an infimum, meaning that it is obtained for a peculiar additive measure of $\mathscr{K}$.

To show the first point, we will first prove the following proposition concerning $\underline{\mu}$:

Proposition 1. *Given a set $\mathscr{K}$, we have for any $A \subseteq \mathscr{X}$*

$$\underline{\mu}_{\mathscr{K}}(A) \leq \inf_{\kappa\in\mathscr{K}} \mu_\kappa(A)$$

Proof. To prove it, consider a given $\kappa \in \mathscr{K}$, we have $\mu_\kappa(A) = \sum_{x\in A} \kappa(x) = 1 - \sum_{x\notin A} \kappa(x)$ since $\sum_{x\in\mathscr{X}} \kappa(x) = 1$. Now, from Eq. (5), we have either

- $\underline{\mu}_{\mathscr{K}}(A) = \sum_{x\in A} \underline{\rho}(x) \leq \sum_{x\in A} \kappa(x)$
- $\underline{\mu}_{\mathscr{K}}(A) = 1 - \sum_{x\notin A} \overline{\rho}(x) \leq 1 - \sum_{x\notin A} \kappa(x)$

and since this is true for any $\kappa \in \mathscr{K}$, we have the inequality.

Note that $\underline{\mu}$ is a tight measure on singletons, since $\underline{\mu}(x) = \underline{\rho}(x) = \inf_{\kappa\in\mathscr{K}} \kappa(x)$, hence any set-function higher than $\underline{\mu}$ on singletons would not be a lower envelope of $\mathscr{K}$. The fact that $\underline{\mathscr{C}}_{\mathscr{K}}(f) \leq \mathscr{C}_\kappa(f)$ for any $\kappa \in \mathscr{K}$ then simply follows from Proposition 1. If $\mathscr{K}$ reduces to a single kernel κ, then we find back the classical filtering result. Note that $\underline{\mathscr{C}}_{\mathscr{K}}(f)$ is equivalent to filter f, that is to compute $\mathscr{C}_\kappa$, with the specific kernel $\kappa(x_{(i)}) = \underline{\mu}(A_{(i)}) - \underline{\mu}(A_{(i+1)})$. To prove that $\underline{\mathscr{C}}_{\mathscr{K}}$ is a tight lower bound, it remains to show that such a kernel is within $\mathscr{K}$.

To show that the bound obtained by Equation (6) is actually obtained by an additive measure dominated by $\underline{\mu}_{\mathscr{K}}$, we will first show that it still satisfies a convexity property

[4] The lower envelope of a set of probability measures.

Proposition 2. *Given a set $\mathscr{K}$ of kernels, the measure $\underline{\mu}_{\mathscr{K}}$ is 2-monotone and convex, as for every pair $A,B \subseteq \mathscr{X}$ we have*

$$\underline{\mu}_{\mathscr{K}}(A\cup B)+\underline{\mu}_{\mathscr{K}}(A\cap B) \geq \underline{\mu}_{\mathscr{K}}(A)+\underline{\mu}_{\mathscr{K}}(B)$$

Proof. We will mainly adapt the proof from [1, Proposition 5] to the case of non-positive kernels and signed measures, as its mechanism still works in this case.

A key element will be to show that for any two subsets C,D with $C\cap D=\emptyset$, there exists a **single** additive measure μ_κ with $\kappa \in \mathscr{K}$ such that

$$\underline{\mu}_{\mathscr{K}}(C) = \mu_\kappa(C) \text{ and } \underline{\mu}_{\mathscr{K}}(C\cup D) = \mu_\kappa(C\cup D), \tag{8}$$

as if we then take $C = A\cap B$ and $D = (A\cup B)\setminus(A\cap B)$ and choose κ so that it coincides on $\underline{\mu}_{\mathscr{K}}$ for events C,D, we do have

$$\begin{aligned}\underline{\mu}_{\mathscr{K}}(A\cap B)+\underline{\mu}_{\mathscr{K}}(A\cup B) &= \mu_\kappa(A\cap B)+\mu_\kappa(A\cup B)\\ &= \mu_\kappa(A)+\mu_\kappa(B)\\ &\geq \underline{\mu}_{\mathscr{K}}(A)+\underline{\mu}_{\mathscr{K}}(B).\end{aligned}$$

By Equation (5), we know that

$$\underline{\mu}_{\mathscr{K}}(A) = \max\{\sum_{x\in A}\underline{\rho}(x), 1-\sum_{x\notin A}\overline{\rho}(x)\}$$

which means that for any event A, we have two possibilities (the two terms of the max). This means four possibilities when considering C and $C\cup D$ together. Here, we will only show that Equation (8) is true for one of those case, as the proofs for the other cases follow similar reasoning.

So let us consider the case where

$$\begin{aligned}\underline{\mu}_{\mathscr{K}}(C) &= \sum_{x\in C}\underline{\rho}(x) \geq 1-\sum_{x\notin C}\overline{\rho}(x),\\ \underline{\mu}_{\mathscr{K}}(C\cup D) &= 1-\sum_{x\notin C\cup D}\overline{\rho}(x) \geq \sum_{x\in C\cup D}\underline{\rho}(x).\end{aligned}$$

Let us now consider the κ distribution such that $\kappa(x)=\underline{\rho}(x)$ if $x\in C$, $\kappa(x)=\overline{\rho}(x)$ if $x\in(C\cup D)^c$, that fits the requirements of Equation (8) and so far satisfy the constraints on $\mathscr{K}$. To get an additive kernel whose weights sum up to one, we must still assign $\lambda = 1-\sum_{x\in C}\underline{\rho}(x)-\sum_{x\in(C\cup D)^c}\overline{\rho}(x)$ mass over the singletons composing D. One can see that $\sum_{x\in D}\underline{\rho}(x)\leq\lambda\leq\sum_{x\in D}\overline{\rho}(x)$: for instance, that $\sum_{x\in D}\underline{\rho}(x)\leq\lambda$ immediately follows from the fact that in this sub-case $1-\sum_{x\notin C\cup D}\overline{\rho}(x)\geq\sum_{x\in C\cup D}\underline{\rho}(x)$. This means that one can choose values $\kappa(x)\in[\underline{\rho}(x),\overline{\rho}(x)]$ for each $x\in D$ such that $\sum_{x\in D}\kappa(x)=\lambda$. So in this case we can build an additive $\kappa\in\mathscr{K}$ with $\sum\kappa(x)=1$.

That a single additive $\kappa\in\mathscr{K}$ reaching the bounds $\underline{\mu}_{\mathscr{K}}(C)$ and $\underline{\mu}_{\mathscr{K}}(C\cup D)$ can be built in other sub-cases can be done similarly (we refer to [1, Proposition 5], as the proofs are analogous).

Hence $\underline{\mathscr{C}}_{\mathscr{K}}$ is a signed Choquet integral with respect to the convex capacity $\underline{\mu}_{\mathscr{K}}$. We have $\underline{\mu}_{\mathscr{K}}(\emptyset) = 0$, so according to [10, Theorem 3], $\underline{\mathscr{C}}_{\mathscr{K}}$ is the minimum of the integrals or expectations taken with respect to the additive measures dominating $\underline{\mu}_{\mathscr{K}}$, i.e., $\underline{\mathscr{C}}_{\mathscr{K}}(f) = \min\{\mathscr{C}_\mu | \mu \in core(\underline{\mu}_{\mathscr{K}}\}$, where the core of a capacity is the set of additive set function that lie above the capacity everywhere. This allows us to state the following property

Proposition 3. $\underline{\mathscr{C}}_{\underline{\mu}_{\mathscr{K}}}(f) = \min\{\mathscr{C}_\mu | \mu \in core(\underline{\mu}_{\mathscr{K}})\}$.

We have therefore shown that, to approximate the result of filtering with any set of kernels (bounded and with no gain), it is still possible to use tools issued from imprecise probabilistic literature. However, it is even clearer in this case that such tools should not be interpreted straightforwardly as uncertainty models (as set-functions are not monotone, a property satisfied by standard fuzzy measures and coherent lower probabilities), but as convenient and efficient tools to perform robust filtering.

Remark 1. An upper bound $\overline{\mathscr{C}}_{\mathscr{K}}(f)$ can be obtained by using the conjugate capacity $\overline{\mu}_{\mathscr{K}}(A) = 1 - \underline{\mu}_{\mathscr{K}}(A^c)$ in Equation (6). As $\overline{\mu}_{\mathscr{K}}$ is a concave capacity, we also have $\overline{\mathscr{C}}_{\mathscr{K}}(f) = \max\{\mathscr{C}_\mu | \mu \in anticore(\overline{\mu}_{\mathscr{K}})\}$, where the anticore of a capacity is the set of additive set function that lie below the capacity everywhere. Note that $\mu \in core(\underline{\mu}_{\mathscr{K}})$ is equivalent to $\mu \in anticore(\overline{\mu}_{\mathscr{K}})$.

Example 2. Consider the imprecise kernel $\overline{\rho}$ whose values on $\mathscr{X} = \{x_1, \ldots, x_5\}$ are given in Table 1. For such an imprecise kernel, we have for instance $\underline{\mu}_{\mathscr{K}}(\{x_1\}) =$

	x_1	x_2	x_3	x_4	x_5
$\overline{\rho}$	−0.1	0.3	0.9	0.3	−0.1
$\underline{\rho}$	−0.2	0.2	0.8	0.2	−0.2

Table 1. Imprecise kernel example

$\max(-0.2, 1-(1.4)) = -0.2$ and $\underline{\mu}_{\mathscr{K}}(\{x_1, x_5\}) = \max(-0.4, 1-(1.5)) = -0.4$, showing that the defined measure is not monotonic with inclusion.

Finally, it should be noted that applying Equation (6) does not require to evaluate our lower measure on every possible events, but only in a linear number of them (once function values have been ordered). Moreover, evaluating the value of this lower measure on any interval is quite straightforward given Equation (5). So, even though the measure is non-additive (and not necessarily monotonic), evaluating the filtered values can be done quite efficiently.

4 Illustration of signed filtering on a real case

We now illustrate the use of our method on a real case scenario involving the filtering of human electrocardiogram (ECG) signals, using data initially collected to detect heart

conditions under different settings [7]. ECG signals contain many types of noises - e.g. baseline wander, power-line interference, electromyographic (EMG) noise, electrode motion artifact noise, etc. Baseline wander is a low-frequency noise of around 0.5 to 0.6 Hz that is usually removed during the recording by a high-pass filtering of cut-off frequency 0.5 to 0.6 Hz. EMG noise, which is a high frequency noise of above 100 Hz, may be removed by a digital low-pass filter with an appropriate cut-off frequency. In [7] they propose to use a cut-off frequency of 45 Hz to preprocess the ECG signals. The noisy ECG signal to be filtered is presented in Figure 2.

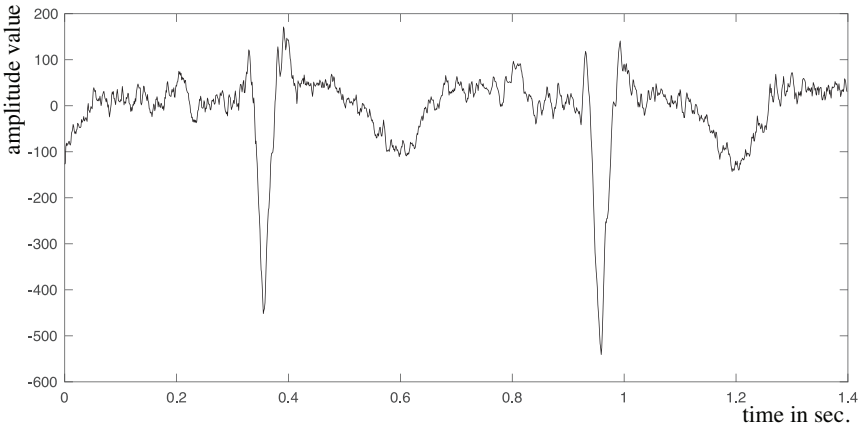

Fig. 2. The ECG signal to be low-pass filtered.

To prevent phase distortion in the bandpass, a Butterworth kernel should preferably be chosen. Moreover, since the signal has not to be processed on line, a symmetric Butterworth filter can be used, that is the combination of a causal and an anti-causal Butterworth kernel. Using such an even kernel prevents from phase delay.

In this experiment, we suppose that the 45 Hz cutoff frequency proposed in [7] is appropriate while the suitable order of the kernel is imprecisely known. Figure 3 presents the superposition of 13 kernels that are the impulse responses of the 13 symmetric low-pass Butterworth kernels of orders 1 to 13. Figure 3.a shows the superimposed kernels, that constitute the set $\mathscr{N}$ of kernels we have to approximate.

Applying our approximation to $\mathscr{N}$ provides the upper ($\overline{\rho}$) and lower ($\underline{\rho}$) bounds of the imprecise kernel that are pictured in Figure 3.c, with the lower in red, the upper in blue. These bounds are simply obtained by computing $\underline{\rho} = \min_{n=1\ldots13} \eta_n$ and $\overline{\rho} = \max_{n=1\ldots13} \eta_n$ where η_n is the impulse response of the lowpass symmetric Butterworth kernel of order n with cutoff frequency equal to 45 Hz.

To have a comparison point, we will also apply to the same signal the maxitive approach proposed in [9], where a signed kernel is approximated by a couple of extended possibility distributions (π^-, π^+). This couple of functions is computed in this way: $\pi^+ = \max_{n=1\ldots13} \pi_n^+$, where π_n^+ is the most specific maxitive kernel that dominates $\eta_n^+ = \max(0, \eta_n)$ and $\pi^- = \max_{n=1\ldots13} \pi_n^-$, where π_n^- is the most specific maxitive ker-

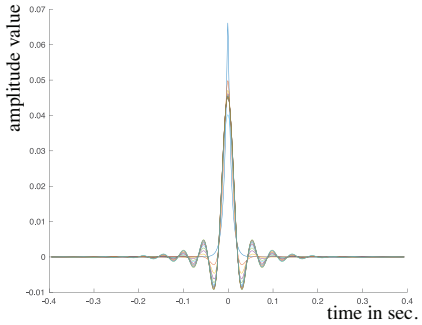

Fig. 3.a: 13 superimposed symmetric Butterworth kernel impulse responses.

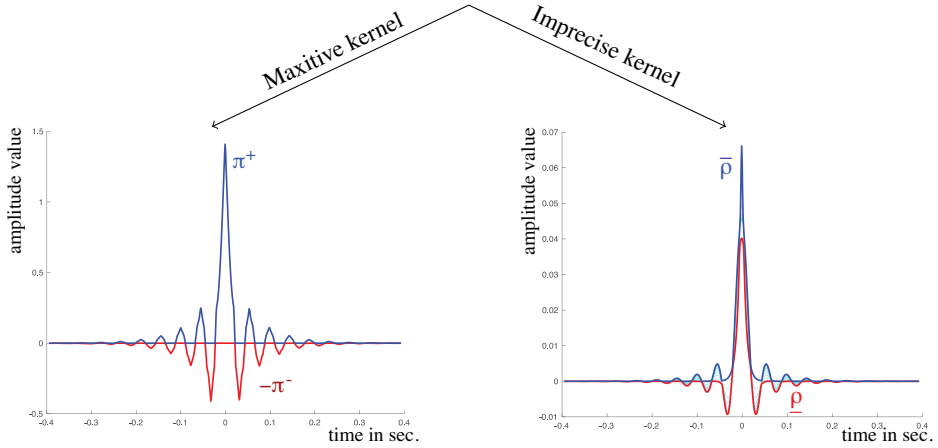

Fig. 3.b: maxitive kernel.

Fig. 3.c: lower (red) and upper (blue) bounds of our imprecise kernel

Fig. 3. Original kernel family and corresponding imprecise models.

nel that dominates $\eta_n^- = \max(0, -\eta_n)$ (see [9] Equation (4)). Figure 3.b plots π^+ (in blue) and $-\pi^-$ (in red). One can readily notice that, if their shape are similar, their boundary values are quite different (the imprecise maxitive kernel varying between -0.5 and 1.5, and our imprecise kernel between -0.01 and 0.07).

In Figures 5 and 4, we have plotted the ECG signal of Figure 2 filtered by the 13 kernels of Figure 3.a, as well as the imprecise signal obtained by using the most specific signed maxitive kernel defined by the couple of functions (π^-, π^+), and the imprecise signal obtained by using the imprecise kernel plotted in Figure 3.c, respectively. The upper bounds of the imprecise filtered signals are plotted in blue while their lower bounds is plotted in red.

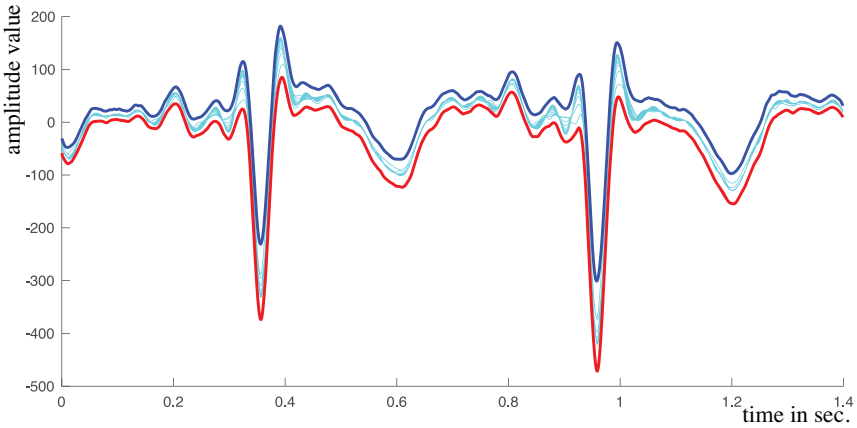

Fig. 4. The ECG signal being low-pass filtered by the 13 classical symmetric Butterworth (in cyan) and by the imprecise kernel (in red - lower and blue - upper).

It seems obvious, by looking at Figure 4, that the imprecise signal obtained by this new approach reaches our pursued goal, i.e. the obtained imprecise signal contains all the signals that would have been obtained by using the conventional approach. Moreover, the bounds are reasonably tight, which means that the core of the imprecise kernel is specific enough as an approximation. Indeed, non-parametric imprecise representation of kernels always leads to include unwanted kernels, and may lead to over-conservative bounds.

This is even more patent if we compare it to the signal bounds obtained with the maxitive approach, as this latter one leads to a less specific interval-valued signal. For instance, the values spanned by the interval-valued signal in our approach span from -500 to 200, and -800 to 500 for the maxitive approach. Another possible advantage of our approach is that the Krœnecker impulse is not necessarily included in the described set of kernels, while it is systematically included in a maxitive kernel, meaning that in this latter case the interval-valued signal always include the noisy original signal itself.

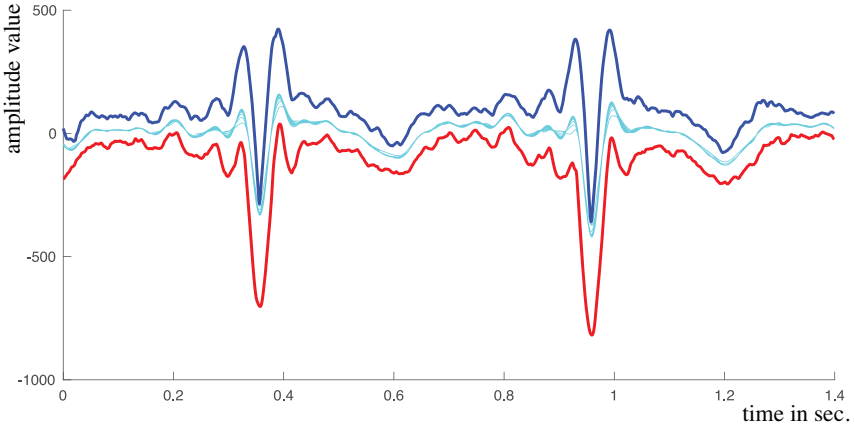

Fig. 5. The ECG signal being low-pass filtered by the 13 classical symmetric Butterworth (in cyan) and by the signed maxitive kernel (in red - lower and blue - upper).

5 Discussion and conclusion

In this paper, we have explored to which extent some of the tools usually used to model and reason with sets of probabilities can still be used when considering sets of additive measures that can be negative and fail the monotonicity condition. Such a situation happens, for instance, when filtering a signal.

We have proved that approximating such sets with interval-valued bounds on singletons by extending probability intervals still provides tools that allow on the one hand to use efficient algorithms, and on the other hand to get tight bounds (in the sense that obtained bounds are reached by specific additive measures). We have provided some preliminary experiments showing how our results could be used in filtering problems.

Future works could include the investigation of other imprecise probabilistic models that also offer computational advantages in the case of positive measures, such as using lower and upper bounds over sequences of nested events [3, 8]. Complementarily, we could investigate whether computations with some parametric sets of signed kernels can be achieved exactly and efficiently without resorting to an approximation, as can be sometimes done for positive kernels [2].

Bibliography

[1] L. M. De Campos, J. F. Huete, and S. Moral. Probability intervals: a tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2(02):167–196, 1994.

[2] G. Dendievel, S. Destercke, and P. Wachalski. Density estimation with imprecise kernels: Application to classification. In *Uncertainty Modelling in Data Science, SMPS 2018, Compiègne, France, 17-21 September 2018*, pages 59–67, 2018.

[3] S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations: I. Generalized p-boxes. *International Journal of Approximate Reasoning*, 49(3):649–663, 2008.

[4] S. Destercke and O. Strauss. Filtering with clouds. *Soft Computing*, 16(5):821–831, 2012.

[5] J. Jan. *Digital Signal Filtering, Analyses and Restoration*. IET, 2000.

[6] K. Loquin and O. Strauss. On the granularity of summative kernels. *Fuzzy Sets and Systems*, 159(1952-1972), 2008.

[7] F. Poree, J.-Y. Bansard, G. Kervio, and G. Carrault. Stability analysis of the 12-lead ecg morphology in different physiological conditions of interest for biometric applications. In *Proceedings of the 36th Annual Computers in Cardiology Conference (CinC)*, pages 285–288, 2009.

[8] O. Strauss and S. Destercke. F-boxes for filtering. In *Proceedings of the 7th conference of the European Society for Fuzzy Logic and Technology, EUSFLAT 2011, Aix-Les-Bains, France, July 18-22, 2011*, pages 935–942, 2011.

[9] 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. IEEE, 2019.

[10] A. D. Waegenaere and P. Wakker. Nonmonotonic choquet integrals. *Journal of Mathematical Economics*, 36:45–60, 2001.