# Regularized Jacobi Wavelets Kernel for Support Vector Machines

Nadira Abbassa *, Abdessamad Amir, Sidi Mohamed Bahri

*Laboratory of pure and applied mathematics, Mostaganem University, Algeria*

**Abstract**    A new family of regularized Jacobi wavelets is constructed. Based on this Jacobi wavelets, a new kernel for support vector machines is presented. Using kernel and frame theory, the Reproducing Kernel Hilbert Space of this kernel is identified. We show that without being a universal kernel, the proposed one possesses a good separation property and a big ability to extract more discriminative features. These theoretical results are confirmed and supported by numerical experiments.

**Keywords**    SVM, Jacobi polynomials, Jacobi wavelets, Kernel, Reproducing Kernel Hilbert Space, Frame.

## 1. Introduction

Support vector machines (SVMs) have become a very dominant tool in machine learning used for both data classification and regression ([20]). They have been extensively applied in many fields as pattern recognition, biology, medical diagnosis, chemistry and bioinformatics ([8, 16, 21]). The classification problem can be limited to consideration of the two-class problem namely binary classification. The principle of SVMs is to bring back this binary classification problem in search of a separating hyperplane. SVM's theory is based on using three main concepts :

- The Structural Risk Minimization principle ([20]), which has been shown to be superior, to traditional Empirical Risk Minimization principle ([3]).
- The use of a convex quadratic optimization problem to solve the problem of maximum margin, efficient solvers are available now that can solve such problem even with large size and dense data ([18]).
- When data are nonlinearly separable, SVMs use a kernel function that maps the data into high-dimensional feature space in which the problem becomes linearly separable.

The first kernels investigated for SVM problems were polynomial, Sigmoid and the radial basis kernel ([20]). Since several kernels were introduced and studied in literature and it is well known, that the major task of the SVM approach lies in the selection of its kernel. Recently, special attention has been given to the construction based on the concepts of orthogonal polynomials ( [1, 9, 10, 11, 22, 25]), which have found widespread use in several areas of science and engineering. The fundamental reason for the success of this kernels type is its orthogonality properties, that reduce the redundancy of the original data and make it able to extract more discriminative features, so that SVM can classify the data better and easier ([6]). Inheriting also this property, kernels based on wavelets have given great satisfaction in practice ([12, 24]). The idea of combining orthogonal polynomial and wavelet theories

---

*Correspondence to: Nadira Abbassa (Email: abbassa.nadira@gmail.com). Laboratory of pure and applied mathematics, Faculty of exact science and computer science, University of Abdelhamid Ibn Badis, Mostaganem, Algeria.

gave rise to new wavelets, such as the wavelets of Legendre, Chebyshev, Hermite, etc, which was a powerful technique for signal processing, solving differential equations, optimal control, and calculus variations problems ([2, 7, 14]). Suffering from a continuity problem, kernels based on this type of wavelet have never been introduced before, more in the context of SVMs. In this paper, we built a new kernel based on Jacobi wavelets. Legendre and Chebyshev wavelets represents a particular case of Jacobi wavelets of indices $(\alpha, \beta)$ when $(\alpha = 0, \beta = 0)$ and $(\alpha = \beta = -\frac{1}{2})$ respectively, this shows the general framework of our kernel. Using frame theory ([13]), we identify the Reproducing Kernel Hilbert Space (RKHS) of this kernel and we confirm that this new kernel has the separation properties. These theoretical results are confirmed and supported by numerical experiments. The paper is structured as follows: section 2 is devoted to kernel and frame theories. In Section 3, we present the main results of this work, we describe the method of constructing the regularized Jacobi wavelets kernel and its RKHS using frame theory. Numerical experiments are given in Section 4, the Jacobi wavelets kernel performance has been evaluated on two-dimensional illustrative examples in order to give also a graphical comparison with other kernels. Finally, we conclude the paper in section 5.

## 2. Backgrounds

Given a finite training samples set $\{(x_i, y_i)\}_{i=1}^{l}$ where $X = \{x_i\}_{i=1}^{l}$ are $l$ points in $\mathbb{R}^d$ and the corresponding labels $y_i$ in $\{-1, 1\}$, the goal is to find an hyperplane in the space $\mathbb{R}^d$ that is inserted between the two classes : class 1 and class $-1$. In linear classification, the aim is to construct a classifier $f$ (a decision function) that distinguishes between the two sets. However, in general data are not linearly separable in practice. Then, a non-linear classification is proposed that maps the data into a high- dimensional feature space $H$ via a transformation $\phi$ called feature map ($\phi \colon X \to H$), such that the transformed data are linearly separable. When this hyperplane is back into the original space it describes a surface.

### 2.1. The optimization problem

The classic soft margin SVM formulation ([5]) is defined by

$$\begin{cases} \min_{\alpha} -\frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j) + \sum_{i=1}^{l} \alpha_i \\ \text{s. t.} \quad \sum_{i=1}^{l} \alpha_i y_i = 0 \\ \qquad 0 \leq \alpha_i \leq C, \quad i = 1, \ldots, l \end{cases}, \tag{1}$$

where the parameter $C$ is the penalty which trade-off balancing the merit of model complexity and misclassified points. It is not necessary to know the explicit form of the mapping $\phi$ since we can replace the inner product $\phi(x_i)^T \phi(x_j)$ by the kernel function $K$ such that

$$K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle_H = \mathbf{K}_{ij},$$

where $\langle ., . \rangle_H$ denotes the inner product of $H$ and $(\mathbf{K}_{ij})_{\substack{i=1,\ldots,l \\ j=1,\ldots,l}}$ is called the Gram matrix of the kernel $K$.

The problem (1) becomes

$$\begin{cases} \min_{\alpha} -\frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j \mathbf{K}_{ij} + \sum_{i=1}^{l} \alpha_i \\ \text{s. t.} \quad \sum_{i=1}^{l} \alpha_i y_i = 0 \\ \qquad 0 \leq \alpha_i \leq C, \quad i = 1, \ldots, l \end{cases}. \tag{2}$$

The solution $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_l)$ of the optimization problem (2) and the bias $b$, which can be recovered from the Lagrangian multiplier associated to the equality contraint of (2) ([3]), gives the classifier function

$$f(x) = \sum_{i=1}^{sv} \alpha_i y_i K(x_i, x) + b,$$

which is fundamental for the learning task. The integer $sv$ corresponds to the number of support vectors, which are the points $x_i$ whose associated Lagrangian multipliers are not zeros ($\alpha_i \neq 0$). Thus, the decision function can be represented by only a few numbers of data points ( the so-called support vectors). SVM classifier that yields fewer support vectors for a given kernel is desirable.

### 2.2. Kernels for SVMs

In the sequel, we present the basic definition and some important properties of kernels for SVMs.

*Definition 1*
[17] Let $X$ be a non-empty set. Then a function $K : X \times X \to \mathbb{R}$ is called a kernel on $X$ if there exists an $\mathbb{R}$-Hilbert space $H$ and a map $\phi : X \to H$ such that

$$K(x,y) = \langle \phi(y), \phi(x) \rangle_H, \quad \forall x, y \in X.$$

$\phi$ is called feature map and $H$ a feature space of $K$.

The most used kernels are the radial basis function RBF or Gaussian defined for some positive parameter $\gamma$ by

$$K(x,y) = \exp\left(-\frac{\|x-y\|^2}{\gamma^2}\right),$$

and the polynomial kernel defined for some positive integer $p$ by

$$K(x,y) = (x \times y + 1)^p.$$

One can construct new kernels from scratch using kernel's properties. A good study on kernels theory is given in reference ([17]).

A characterization of kernels for SVMs is given by the following theorem ([17]) :

*Theorem 1*
(Mercer) Let $X$ be a non-empty set and let a function $K : X \times X \to \mathbb{R}$. $K$ is a kernel if and only if it is symmetric and its Gram matrix is positive semidefinite.

It's known that every kernel admits several feature space, the next definition introduces the RKHS, which is in a certain sense, the smallest feature space of this kernel and consequently, it can serve as a canonical one.

*Definition 2*
[17] Let $X \neq \emptyset$ and $H$ be a $\mathbb{R}$-Hilbert space. A function $K : X \times X \to \mathbb{R}$ is called a reproducing kernel of $H$ if

1. $K(.,x) \in H, \ \forall x \in X$.
2. The reproducing property is verified

$$f(x) = \langle f, K(.,x) \rangle_H, \ \forall f \in H, \forall x \in X.$$

3. The space $H$ is called a RKHS over $X$ if the Dirac functional

$$\begin{aligned} \delta_x \quad &: \quad H \to \mathbb{R} \\ f \quad &\to \quad f(x) \end{aligned}$$

is continuous for all $x \in X$.

## 2.3. Frame for SVMs

One can construct kernels and its associated RKHS from frame's theory ([13]). We give the definition of a frame and then a characterization result necessary for the remainder.

### Definition 3

[4] Let $N \in \mathbb{N}$ and $\{\psi_n\}_{n=1,\dots,N}$ be a set of non-zeros functions of an Hilbert space $(H, \langle .,. \rangle_H)$. The familly $\{\psi_n\}_{n=1,\dots,N}$ is a frame for $H$ if there exist constants $B, G > 0$ such that

$$B \|f\|_H^2 \leq \sum_{n=1}^{N} |\langle f, \psi_n \rangle_H|^2 \leq G \|f\|_H^2.$$

The numbers $B$ and $G$ are called frame bounds.

The reconstruction of $f$ from its frame coefficients needs the following definition of a dual frame.

### Definition 4

[13] Let $\{\psi_n\}_{n=1,\dots,N}$ be a frame of $H$ with frame bounds $B$ and $G$. The dual frame $\{\bar{\psi}_n\}_{n=1,\dots,N}$ is defined as

$$\frac{1}{G} \|f\|_H^2 \leq \sum_{n=1}^{N} |\langle f, \bar{\psi}_n \rangle_H|^2 \leq \frac{1}{B} \|f\|_H^2,$$

and $f = \sum_{n=1}^{N} \langle f, \bar{\psi}_n \rangle_H \psi_n$.

### Theorem 2

Let $N \in \mathbb{N}$ and $\{\psi_n\}_{n=1,\dots,N}$ be a set of non-zeros functions of an Hilbert space $(\mathcal{B}, \langle .,. \rangle_{\mathcal{B}})$ such that $\exists \Delta > 0$, $\forall x \in X$, $\forall n \; 1 \leq n \leq N$

$$|\psi_n(x)| \leq \Delta.$$

Let $H$ be the set of functions so that

$$H = \left\{ f = \sum_{n=1}^{N} a_n \psi_n, \; a_n \in \mathbb{R} \text{ for } n = 1,\dots,N \right\}.$$

The space $(H, \langle .,. \rangle_{\mathcal{B}})$ is an RKHS and its reproducing kernel is

$$K(x,y) = \sum_{n=1}^{N} \bar{\psi}_n(x)\psi_n(y),$$

where $\{\bar{\psi}_n\}_{n=1,\dots,N}$ is the dual frame of $\{\psi_n\}_{n=1,\dots,N}$ in $H$.

We will combine SVM and wavelet theories to built a new kernel based on Jacobi wavelets and we will characterize the theoretical properties of this new kernel in the next section.

## 3. Main results

### 3.1. The Jacobi wavelets kernel

The Jacobi polynomials $J_m^{(\alpha,\beta)}$ are defined by the following iterative formula for all $\alpha > -1$, $\beta > -1$:

$$\begin{aligned}
J_m^{(\alpha,\beta)}(x) &= \frac{(\alpha+\beta+2m-1)\left[\alpha^2 - \beta^2 + x(\alpha+\beta+2m)(\alpha+\beta+2m-2)\right]}{2m\,(\alpha+\beta+2m-2)\,(\alpha+\beta+m)} J_{m-1}^{(\alpha,\beta)}(x) \\
&\quad - \frac{(\alpha+m-1)\,(\beta+m-1)\,(\alpha+\beta+2m)}{m\,(\alpha+\beta+2m-2)\,(\alpha+\beta+m)} J_{m-2}^{(\alpha,\beta)}(x),
\end{aligned}$$

where

$$J_0^{(\alpha,\beta)}(x) = 1, \ J_1^{(\alpha,\beta)}(x) = \frac{\alpha+\beta+2}{2}x + \frac{\alpha-\beta}{2}.$$

These polynomials belong to the weight space $L_\omega^2\left([-1,1]\right)$

$$\left\langle J_m^{(\alpha,\beta)}, J_{m'}^{(\alpha,\beta)} \right\rangle_{L_\omega^2} = h_m^{(\alpha,\beta)}\delta_{m,m'}, \ \forall m, m' \in \mathbb{N},$$

where

$$h_m^{(\alpha,\beta)} = \left\| J_m^{(\alpha,\beta)} \right\|^2 = \frac{2^{\alpha+\beta+1}\Gamma\left(\alpha+m+1\right)\Gamma\left(\beta+m+1\right)}{(2m+1+\alpha+\beta)\,m!\Gamma\left(\alpha+\beta+m+1\right)},$$

and $\omega(x) = (1-x)^\alpha (1+x)^\beta$. $\delta_{n,m}$ represents the Kronecker function, $\Gamma$ is the Euler gamma function and $\langle .,.\rangle_{L_\omega^2}$ denotes the inner product of $L_\omega^2\left([-1,1]\right)$.

The family $\left\{J_m^{(\alpha,\beta)}\right\}_{m\in\mathbb{N}}$ forms an orthogonal basis for $L_\omega^2\left([-1,1]\right)$ ([19]).

The Jacobi wavelets are defined by ([23])

$$\psi_{n,m}^{(\alpha,\beta)}(x) = \begin{cases} \dfrac{2^{\frac{k+1}{2}}}{\sqrt{h_m^{(\alpha,\beta)}}} J_m^{(\alpha,\beta)}\left(2^{k+1}x - 2n + 1\right), & \frac{n-1}{2^k} \le x < \frac{n}{2^k}, \\ 0, & \text{othetwise} \end{cases} \tag{3}$$

where $k \in \mathbb{N}$, $n = 1,,...,2^k$ represents the number of decomposition levels, $m = 0, 1, ..., M$ is the degree of the Jacobi polynomials ($M \in \mathbb{N}$). The coefficient $\frac{2^{\frac{k+1}{2}}}{\sqrt{h_m^{(\alpha,\beta)}}}$ is for normality.

*Theorem 3*
The familly $\left\{\psi_{n,m}^{(\alpha,\beta)}\right\}_{\substack{n=1,...,2^k \\ m \ge 0}}$ forms an orthonormal basis of $\left(L_\omega^2\left([0,1[\right), \langle.,.\rangle_{wav}\right)$, where

$$\langle f, g\rangle_{wav} = \sum_{i=1}^{2^k} \int_{\frac{i-1}{2^k}}^{\frac{i}{2^k}} f_i(x)g_i(x)\omega_i(x)dx,$$

$f_i$ is the restriction of $f$ on $\left[\frac{i-1}{2^k}, \frac{i}{2^k}\right[$ and $\omega_i(x) = \omega(2^{k+1}x - 2i + 1)$.

*Proof*
First, we have to prove that $\langle.,.\rangle_{wav}$ is an inner product of $L_\omega^2\left([0,1[\right)$. Let $f$ be a function of $L_\omega^2\left([0,1[\right)$. The symmetry and bilinearity of $\langle.,.\rangle_{wav}$ are obvious, we only show that $\langle.,.\rangle_{wav}$ is positive definite. We have

**i** $\langle f, f\rangle_{wav} = \sum_{i=1}^{2^k} \int_{\frac{i-1}{2^k}}^{\frac{i}{2^k}} (f_i(x))^2 \omega_i(x)dx \ge 0,$

since $\int_{\frac{i-1}{2^k}}^{\frac{i}{2^k}} (f_i(x))^2 \omega_i(x)dx = \langle f_i, f_i\rangle_{L_\omega^2\left[\frac{i-1}{2^k}, \frac{i}{2^k}\right[} \ge 0.$

**ii**

$$\begin{aligned} \langle f, f\rangle_{wav} &= 0 \\ &\Longleftrightarrow \int_{\frac{i-1}{2^k}}^{\frac{i}{2^k}} (f_i(x))^2 \omega_i(x)dx = 0, \forall i = 1, \ldots, 2^k \\ &\Longleftrightarrow f_i = 0, \forall i = 1, \ldots, 2^k \\ &\Longleftrightarrow f = 0. \end{aligned}$$

Now, we show that the family $\left\{\psi_{n,m}^{(\alpha,\beta)}\right\}_{\substack{n=1,...,2^k \\ m \geq 0}}$ forms an orthonormal basis of $L_\omega^2\left([0,1[\right)$

$$
\left\langle \psi_{n,m}^{(\alpha,\beta)}, \psi_{j,l}^{(\alpha,\beta)} \right\rangle_{wav} = \sum_{i=1}^{2^k} \int_{\frac{i-1}{2^k}}^{\frac{i}{2^k}} \left(\psi_{n,m}^{(\alpha,\beta)}\right)_i (x) \left(\psi_{j,l}^{(\alpha,\beta)}\right)_i (x) \omega_i(x) dx
$$

$$
= \delta_{n,j} \int_{\frac{n-1}{2^k}}^{\frac{n}{2^k}} \psi_{n,m}^{(\alpha,\beta)}(x) \psi_{n,l}^{(\alpha,\beta)}(x) \omega_n(x) dx.
$$

By using the change of variable $t = 2^{k+1}x - 2n + 1$ and the fact that the family $\left\{\frac{J_m^{(\alpha,\beta)}}{\left\|J_m^{(\alpha,\beta)}\right\|}\right\}_{m \geq 0}$ is an orthonormal basis for $L_\omega^2\left([-1,1]\right)$, we obtain

$$
\left\langle \psi_{n,m}^{(\alpha,\beta)}, \psi_{j,l}^{(\alpha,\beta)} \right\rangle_{wav} = \delta_{n,j} \frac{2^{k+1}}{\left\|J_m^{(\alpha,\beta)}\right\| \left\|J_l^{(\alpha,\beta)}\right\|} \int_{\frac{n-1}{2^k}}^{\frac{n}{2^k}} J_m^{(\alpha,\beta)}(t) J_l^{(\alpha,\beta)}(t) \omega(t) \frac{dt}{2^{k+1}}
$$

$$
= \delta_{n,j} \delta_{m,l}.
$$

$\square$

We will use the same approach of constructing RKHS by frame theory given in Theorem 2. One can affirm the existence of a positive constant $\Delta$ for a fixed $M, k \in \mathbb{N}$ such that

$$
\left|\psi_{n,m}^{(\alpha,\beta)}(x)\right| \leq \Delta, \quad \text{for } x \in [0,1[, \ n = 1,...,2^k \text{ and } m = 0,...,M. \tag{4}
$$

Let $k, M \in \mathbb{N}$ and $\left\{\psi_{n,m}^{(\alpha,\beta)}\right\}_{\substack{n=1,...,2^k \\ m=0,...,M}}$ defined as (3). The space of functions $(H_{wav}, \langle .,. \rangle_{wav})$ is defined by

$$
H_{wav} = \left\{ f = \sum_{n=1}^{2^k} \sum_{m=0}^{M} a_{n,m} \psi_{n,m}^{(\alpha,\beta)}, \ a_{n,m} \in \mathbb{R} \text{ for } n = 1,...,2^k, \ m = 0,...,M \right\}.
$$

From (4) and the fact that orthogonal basis is frame ([4]), we deduce from Theorem 2. that $H_{wav}$ is a RKHS and its reproducing kernel is given by

$$
\begin{aligned}
k_{wav} \quad &: \quad [0,1] \times [0,1] \to \mathbb{R} : \\
(x,y) \quad &\mapsto \quad k_{wav}(x,y) = \sum_{n=1}^{2^k} \sum_{m=0}^{M} \psi_{n,m}^{(\alpha,\beta)}(x) \psi_{n,m}^{(\alpha,\beta)}(y).
\end{aligned} \tag{5}
$$

In general, functions belonging to $H_{wav}$ are not all continuous.

In fact, for $k = 1$ and $M = 1$, we have

$$
\begin{aligned}
f(x) &= a_{1,0}\psi_{1,0}^{(\alpha,\beta)}(x) + a_{1,1}\psi_{1,1}^{(\alpha,\beta)}(x) + a_{2,0}\psi_{2,0}^{(\alpha,\beta)}(x) + a_{2,1}\psi_{2,1}^{(\alpha,\beta)}(x) \\
&= \begin{cases} \frac{2a_{1,0}}{\sqrt{h_0^{(\alpha,\beta)}}} + \frac{2a_{1,1}}{\sqrt{h_1^{(\alpha,\beta)}}} J_1^{(\alpha,\beta)}(4x-1), & x \in \left[0, \frac{1}{2}\right[ \\ \frac{2a_{2,0}}{\sqrt{h_0^{(\alpha,\beta)}}} + \frac{2a_{2,1}}{\sqrt{h_1^{(\alpha,\beta)}}} J_1^{(\alpha,\beta)}(4x-3), & x \in \left[\frac{1}{2}, 1\right[ \end{cases}.
\end{aligned}
$$

It is clear that $f$ is continuous on $\left[0, \frac{1}{2}\right[$ and on $\left[\frac{1}{2}, 1\right[$. We will study the continuity of $f$ at $\frac{1}{2}$.

$$
\begin{aligned}
\lim_{x \overset{<}{\to} \frac{1}{2}} f(x) &= \lim_{x \overset{<}{\to} \frac{1}{2}} \frac{2a_{1,0}}{\sqrt{h_0^{(\alpha,\beta)}}} + \frac{2a_{1,1}}{\sqrt{h_1^{(\alpha,\beta)}}} J_1^{(\alpha,\beta)}(4x-1) \\
&= \frac{2a_{1,0}}{\sqrt{h_0^{(\alpha,\beta)}}} + \frac{2a_{1,1}}{\sqrt{h_1^{(\alpha,\beta)}}} J_1^{(\alpha,\beta)}(1).
\end{aligned}
$$

$$
f(\frac{1}{2}) = \frac{2a_{2,0}}{\sqrt{h_0^{(\alpha,\beta)}}} + \frac{2a_{2,1}}{\sqrt{h_1^{(\alpha,\beta)}}} J_1^{(\alpha,\beta)}(-1).
$$

We take $a_{n,m} = a_{n',m} = a_m$ for $n, n' = 1, ..., 2^k$, we obtain

$$
\lim_{x \overset{<}{\to} \frac{1}{2}} f(x) = \frac{2a_0}{\sqrt{h_0^{(\alpha,\beta)}}} + \frac{2a_1}{\sqrt{h_1^{(\alpha,\beta)}}} J_1^{(\alpha,\beta)}(1),
$$

$$
f(\frac{1}{2}) = \frac{2a_0}{\sqrt{h_0^{(\alpha,\beta)}}} + \frac{2a_1}{\sqrt{h_1^{(\alpha,\beta)}}} J_1^{(\alpha,\beta)}(-1),
$$

but $J_1^{(\alpha,\beta)}(1) \neq J_1^{(\alpha,\beta)}(-1)$.

Kernel of type (5) has never been used for SVM since continuity is necessary to have a high- performance RKHS ([6]). This continuity problem is addressed in the following subsection.

### 3.2. The regularized Jacobi wavelets

Now, we consider the regularized Jacobi wavelets

$$
\hat{\psi}_{n,m}^{(\alpha,\beta)}(x) = \begin{cases} \frac{2^{\frac{k+1}{2}}}{\sqrt{h_m^{(\alpha,\beta)}}} \left( J_m^{(\alpha,\beta)} \left( 2^{k+1}x - 2n + 1 \right) \right) \\ \quad + \frac{2^{\frac{k+1}{2}}}{\sqrt{h_m^{(\alpha,\beta)}}} (n-1) \left( J_m^{(\alpha,\beta)}(1) - J_m^{(\alpha,\beta)}(-1) \right), & \frac{n-1}{2^k} \leq x < \frac{n}{2^k}, \\ 0. & \text{otherwise} \end{cases} \tag{6}
$$

Let $\hat{H}$ be the space of functions defined by

$$
\hat{H} = \left\{ f = \sum_{n=1}^{2^k} \sum_{m=0}^{M} a_m \hat{\psi}_{n,m}^{(\alpha,\beta)}, \ a_m \in \mathbb{R} \text{ for } m = 0, ..., M \right\}.
$$

One can claim that :

*Theorem 4*
All functions belonging to the space $\hat{H}$ are continuous.

*Proof*
We know that the function $f$ is continuous on each subinterval $\left[\frac{n-1}{2^k}, \frac{n}{2^k}\right[$, then it only remains to study the continuity at the points $\frac{n}{2^k}$, $n = 1, ..., 2^k - 1$. We just need to compare between $\lim_{x \overset{<}{\to} \frac{n}{2^k}} f(x)$ and $f(\frac{n}{2^k})$ :

$$
\begin{aligned}
f(\frac{n}{2^k}) &= \sum_{m=0}^{M} a_m \hat{\psi}_{n+1,m}^{(\alpha,\beta)}(\frac{n}{2^k}) \\
&= \sum_{m=0}^{M} a_m \frac{2^{\frac{k+1}{2}}}{\sqrt{h_m^{(\alpha,\beta)}}} \left( n J_m^{(\alpha,\beta)}(1) - (n-1) \left( J_m^{(\alpha,\beta)}(-1) \right) \right).
\end{aligned}
$$

$$
\begin{aligned}
\lim_{x \searrow \frac{n}{2^k}} f(x) &= \lim_{x \searrow \frac{n}{2^k}} \sum_{m=0}^{M} a_m \hat{\psi}_{n,m}^{(\alpha,\beta)}(x) \\
&= \lim_{x \searrow \frac{n}{2^k}} \sum_{m=0}^{M} a_m \frac{2^{\frac{k+1}{2}}}{\sqrt{h_m^{(\alpha,\beta)}}} \left( J_m^{(\alpha,\beta)} \left( 2^{k+1}x - 2n + 1 \right) + (n-1) \left( J_m^{(\alpha,\beta)}(1) - J_m^{(\alpha,\beta)}(-1) \right) \right) \\
&= f(\frac{n}{2^k}),
\end{aligned}
$$

hence, $f$ is continuous. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

*Remark 1*
The family $\left\{ \hat{\psi}_{n,m}^{(\alpha,\beta)} \right\}_{\substack{n=1,...,2^k \\ m=0,...,M}}$ is not orthogonal, but it is a wavelet frame for $\hat{H} =$ $span \left\{ \hat{\psi}_{n,m}^{(\alpha,\beta)} \right\}_{\substack{n=1,...,2^k \\ m=0,...,M}}$, since it is a finite set of elements of $\left( L_\omega^2([0,1]), \langle .,. \rangle_{wav} \right)$ ([4]).

To apply Theorem 2., we must compute the dual of (6). The frame operator $U$ can be defined as in ([4])

$$
U: \quad
\begin{array}{l}
\hat{H} \to \mathbb{R}^{2^K(M+1)} \\
f \to \left\{ \left\langle f, \hat{\psi}_{n,m}^{(\alpha,\beta)} \right\rangle_{wav} \right\}_{\substack{n=1,...,2^k \\ m=0,...,M}}
\end{array}
\quad .
$$

The adjoint operator is given by

$$
U^*: \quad
\begin{array}{l}
\mathbb{R}^{2^K(M+1)} \to \hat{H} \\
\{c_{nm}\}_{\substack{n=1,...,2^k \\ m=0,...,M}} \to \sum_{n=1}^{2^k} \sum_{m=0}^{M} c_{nm} \hat{\psi}_{n,m}^{(\alpha,\beta)}
\end{array}
\quad .
$$

Let us define the dual frame $\left\{ \bar{\psi}_{n,m}^{(\alpha,\beta)} \right\}_{\substack{n=1,...,2^k \\ m=0,...,M}}$ as

$$
S^{-1} \left( \hat{\psi}_{n,m}^{(\alpha,\beta)} \right) = \bar{\psi}_{n,m}^{(\alpha,\beta)},
$$

where

$$
S = UU^*,
$$

is called the frame operator. This operator is linear and self- adjoint. Note that in terms of the frame operator ([4]) we have

$$
S(f) = \sum_{n=1}^{2^k} \sum_{m=0}^{M} \left\langle f, \hat{\psi}_{n,m}^{(\alpha,\beta)} \right\rangle_{wav} \hat{\psi}_{n,m}^{(\alpha,\beta)}, \quad f \in \hat{H}. \tag{7}
$$

Now, it remains to determine the $\bar{\psi}_{n,m}^{(\alpha,\beta)}$.

*Theorem 5*
The dual frame $\bar{\psi}_{n,m}^{(\alpha,\beta)}$ is defined as

$$
\left\{
\begin{array}{l}
\bar{\psi}_{n,0}^{(\alpha,\beta)} = \psi_{n,0}^{(\alpha,\beta)} - \sum_{j=1}^{M} \frac{(n-1)\sqrt{h_0^{(\alpha,\beta)}}}{\sqrt{h_j^{(\alpha,\beta)}}} \left( J_j^{(\alpha,\beta)}(1) - J_j^{(\alpha,\beta)}(-1) \right) \psi_{n,j}^{(\alpha,\beta)} \\
\bar{\psi}_{n,m}^{(\alpha,\beta)} = \psi_{n,m}^{(\alpha,\beta)}, \text{ for } m = 1,...,M
\end{array}
\right.
\quad , \text{ for } n = 1,...,2^k.
$$

*Proof*
We know that

$$
\begin{aligned}
\hat{\psi}_{n,m}^{(\alpha,\beta)}(x) &= \frac{2^{\frac{k+1}{2}}}{\sqrt{h_m^{(\alpha,\beta)}}}\left(J_m^{(\alpha,\beta)}\left(2^{k+1}x-2n+1\right)+(n-1)\left(J_m^{(\alpha,\beta)}(1)-J_m^{(\alpha,\beta)}(-1)\right)\right) \\
&= \psi_{n,m}^{(\alpha,\beta)}(x)+\frac{(n-1)\sqrt{h_0^{(\alpha,\beta)}}}{\sqrt{h_m^{(\alpha,\beta)}}}\left(J_m^{(\alpha,\beta)}(1)-J_m^{(\alpha,\beta)}(-1)\right)\frac{2^{\frac{k+1}{2}}}{\sqrt{h_0^{(\alpha,\beta)}}}J_0^{(\alpha,\beta)}(x) \\
&= \psi_{n,m}^{(\alpha,\beta)}(x)+A(n,m)\psi_{n,0}^{(\alpha,\beta)}(x),
\end{aligned}
$$

with $A(n,m)=\frac{(n-1)\sqrt{h_0^{(\alpha,\beta)}}}{\sqrt{h_m^{(\alpha,\beta)}}}\left(J_m^{(\alpha,\beta)}(1)-J_m^{(\alpha,\beta)}(-1)\right)$. Using (7), we get for $n=1,...,2^k$, $m=0,...,M$

$$
\begin{aligned}
S(\psi_{n,m}^{(\alpha,\beta)}) &= \sum_{i=1}^{2^k}\sum_{j=0}^{M}\left\langle\psi_{n,m}^{(\alpha,\beta)},\hat{\psi}_{i,j}^{(\alpha,\beta)}\right\rangle_{wav}\hat{\psi}_{i,j}^{(\alpha,\beta)} \\
&= \sum_{i=1}^{2^k}\sum_{j=0}^{M}\left(\left\langle\psi_{n,m}^{(\alpha,\beta)},\psi_{i,j}^{(\alpha,\beta)}+A(i,j)\psi_{i,0}^{(\alpha,\beta)}\right\rangle_{wav}\right)\hat{\psi}_{i,j}^{(\alpha,\beta)} \\
&= \sum_{i=1}^{2^k}\delta_{n,i}\sum_{j=0}^{M}\left(\left\langle\psi_{n,m}^{(\alpha,\beta)},\psi_{i,j}^{(\alpha,\beta)}\right\rangle_{wav}+\left\langle\psi_{n,m}^{(\alpha,\beta)},A(i,j)\psi_{i,0}^{(\alpha,\beta)}\right\rangle_{wav}\right)\hat{\psi}_{i,j}^{(\alpha,\beta)} \\
&= \sum_{j=0}^{M}\left(\delta_{m,j}+A(n,j)\delta_{m,0}\right)\hat{\psi}_{n,j}^{(\alpha,\beta)} \\
&= \hat{\psi}_{n,m}^{(\alpha,\beta)}+\sum_{j=1}^{M}A(n,j)\delta_{m,0}\hat{\psi}_{n,j}^{(\alpha,\beta)}.
\end{aligned}
$$

If $m\neq 0$, we obtain
$$
S(\psi_{n,m}^{(\alpha,\beta)})=\hat{\psi}_{n,m}^{(\alpha,\beta)},
$$

then,
$$
\bar{\psi}_{n,m}^{(\alpha,\beta)}=S^{-1}(\hat{\psi}_{n,m}^{(\alpha,\beta)})=\psi_{n,m}^{(\alpha,\beta)}. \tag{8}
$$

Otherwise,
$$
S(\psi_{n,0}^{(\alpha,\beta)})=\hat{\psi}_{n,0}^{(\alpha,\beta)}+\sum_{j=1}^{M}A(n,j)\hat{\psi}_{n,j}^{(\alpha,\beta)},
$$

so
$$
\hat{\psi}_{n,0}^{(\alpha,\beta)}=S(\psi_{n,0}^{(\alpha,\beta)})-\sum_{j=1}^{M}A(n,j)\hat{\psi}_{n,j}^{(\alpha,\beta)},
$$

then
$$
S^{-1}\left(\hat{\psi}_{n,0}^{(\alpha,\beta)}\right)=\psi_{n,0}^{(\alpha,\beta)}-\sum_{j=1}^{M}A(n,j)S^{-1}\left(\hat{\psi}_{n,j}^{(\alpha,\beta)}\right). \tag{9}
$$

Combining (8) with (9), we find

$$
\bar{\psi}_{n,0}^{(\alpha,\beta)}=S^{-1}(\hat{\psi}_{n,0}^{(\alpha,\beta)})=\psi_{n,0}^{(\alpha,\beta)}-\sum_{j=1}^{M}A(n,j)\psi_{n,j}^{(\alpha,\beta)}.
$$

So, for every function $f$ in $\hat{H}$, it can be written as

$$f(x) = \sum_{n=1}^{2^k} \sum_{m=0}^{M} \left\langle \bar{\psi}_{n,m}^{(\alpha,\beta)}, f \right\rangle_{wav} \hat{\psi}_{n,m}^{(\alpha,\beta)}(x). \tag{10}$$

$\square$

The space $\hat{H}$ is an RKHS for the same arguments given in the subsection 3.1. According to ([13]), the reproducing kernel associated with $\hat{H}$ is

$$
\begin{aligned}
\hat{K}(x,y) &= \sum_{n=1}^{2^k} \sum_{m=0}^{M} \bar{\psi}_{n,m}^{(\alpha,\beta)}(x) \hat{\psi}_{n,m}^{(\alpha,\beta)}(y) \\
&= \sum_{n=1}^{2^k} \left( \bar{\psi}_{n,0}^{(\alpha,\beta)}(x) \hat{\psi}_{n,0}^{(\alpha,\beta)}(y) + \sum_{m=1}^{M} \bar{\psi}_{n,m}^{(\alpha,\beta)}(x) \hat{\psi}_{n,m}^{(\alpha,\beta)}(y) \right) \\
&= \sum_{n=1}^{2^k} \sum_{m=0}^{M} \psi_{n,m}^{(\alpha,\beta)}(x) \psi_{n,m}^{(\alpha,\beta)}(y) = K_{wav}(x,y).
\end{aligned}
$$

Similarly, the associated feature map is given by

$$
\begin{aligned}
\hat{\Phi} : [0,1] &\to \hat{H} \\
x &\mapsto \hat{\Phi}(x) = \sum_{n=1}^{2^k} \sum_{m=0}^{M} \bar{\psi}_{n,m}^{(\alpha,\beta)}(x) \hat{\psi}_{n,m}^{(\alpha,\beta)} \\
&= \sum_{n=1}^{2^k} \sum_{m=0}^{M} \psi_{n,m}^{(\alpha,\beta)}(x) \psi_{n,m}^{(\alpha,\beta)}
\end{aligned}
$$

For the multidimensional case, we will take for every $\mathbf{x}, \mathbf{y} \in [0,1]^d$, $d \in \mathbb{N}^*$ the following set

$$\mathcal{H} = \otimes_{j=1}^{d} \hat{H},$$

where $\otimes$ is the tensor product ([15]). We define the inner product in $\mathcal{H}$ by

$$\langle f, g \rangle_{\mathcal{H}} = \prod_{j=1}^{d} \left\langle f^j, g^j \right\rangle_{wav} \tag{11}$$

where $f = \otimes_{j=1}^{d} f^j$, $g = \otimes_{j=1}^{d} g^j$ and $f^j, g^j \in \hat{H}$, for $j = 1, ..., d$. We also define for $\mathbf{x}, \mathbf{y} \in [0,1]^d$ the following functions (Lemma 4.6 of [17])

$$\mathcal{K}(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^{d} K_{wav}(x^j, y^j), \tag{12}$$

such as $x^j, y^j \in [0,1]$, and

$$
\mathbf{\Phi} : \begin{array}{l} [0,1]^d \to \mathcal{H} \\ \mathbf{x} \longmapsto \mathbf{\Phi}(\mathbf{x}) = \otimes_{j=1}^{d} \hat{\Phi}(x^j). \end{array}
$$

*Theorem 6*
The space $\mathcal{H}$ is the RKHS associated to the Jacobi wavelets for the multidimensional case, and $\mathcal{K}(.,.)$ is its reproducing kernel.

*Proof*

**i** Let $\mathbf{x}, \mathbf{y} \in [0,1]^d$, from relation (12), we have

$$
\begin{aligned}
\mathcal{K}(\mathbf{x}, \mathbf{y}) &= \prod_{j=1}^{d} K_{wav}(x^j, y^j) \\
&= \otimes_{j=1}^{d} K_{wav}(x^j, .)(y^j) \\
&= \mathcal{K}(\mathbf{x}, .)(\mathbf{y}),
\end{aligned}
$$

so

$$
\mathcal{K}(\mathbf{x}, .) = \otimes_{j=1}^{d} K_{wav}(x^j, .). \tag{13}
$$

Since $k_{wav}$ is the reproducing kernel for $\hat{H}$, then $K_{wav}(x^j, .) \in \hat{H}$, so $\mathcal{K}(\mathbf{x}, .) \in \mathcal{H}$.

**ii** Let $f \in \mathcal{H}$ and $\mathbf{x} \in [0,1]^d$, we know that

$$
\begin{aligned}
f(\mathbf{x}) &= \left( \otimes_{j=1}^{d} f^j \right)(\mathbf{x}) \\
&= \prod_{j=1}^{d} f^j(x^j).
\end{aligned}
$$

Since $f^j \in \hat{H}$ and $\hat{H}$ is an RKHS, then

$$
\begin{aligned}
f(\mathbf{x}) &= \prod_{j=1}^{d} \left\langle K_{wav}(x^j, .^j), f^j \right\rangle_{wav} \\
&= \langle \mathcal{K}(\mathbf{x}, .), f \rangle_{\mathcal{H}},
\end{aligned}
$$

this last equality follows from relations (11) and (13). Then, the reproducing property is verified.

**iii** The Dirac function is continuous on $\mathcal{H}$. Indeed, for $\mathbf{x} \in [0,1]^d$ and $f \in \mathcal{H}$, we have

$$
\begin{aligned}
\delta_{\mathbf{x}}(f) &= f(\mathbf{x}) \\
&= \prod_{j=1}^{d} f^j(x^j) \\
&= \prod_{j=1}^{d} \delta_{x^j} \left( f^j \right).
\end{aligned}
$$

We know that the functional $\delta_{x^j}$ is continuous on $\hat{H}$. So, $\delta_{\mathbf{x}}$ is continuous.

$\square$

The following theorem generalizes Remark 1. for the multidimensional case.

*Theorem 7*
The familly $\left\{ \prod_{j=1}^{d} \hat{\psi}_{n,m}^{(\alpha,\beta)}(.^j) \right\}_{\substack{n=1,...,2^k \\ m=0,...,M}}$ forms a frame for $\mathcal{H}$.

*Proof*
We know that

$$
\begin{aligned}
\mathcal{H} &= \otimes_{j=1}^{d} \hat{H} \\
&= \otimes_{j=1}^{d} span \left\{ \hat{\psi}_{n,m}^{(\alpha,\beta)}(.^j) \right\}.
\end{aligned}
$$

Let $f \in \mathcal{H}$, we have

$$
\begin{aligned}
f(\mathbf{x}) &= \left( \otimes_{j=1}^{d} f^{j} \right)(\mathbf{x}) \\
&= \prod_{j=1}^{d} f^{j}(x^{j}).
\end{aligned}
$$

By applying relation (10) on every function $f^{j}$ in $\hat{H}$, we get

$$
\begin{aligned}
f(\mathbf{x}) &= \prod_{j=1}^{d} \sum_{n=1}^{2^{k}} \sum_{m=0}^{M} \left\langle \left( \bar{\psi}_{n,m}^{(\alpha,\beta)} \right)^{j}, f^{j} \right\rangle_{wav} \hat{\psi}_{n,m}^{(\alpha,\beta)}(x^{j}) \\
&= \sum_{n=1}^{2^{k}} \sum_{m=0}^{M} \prod_{j=1}^{d} \left\langle \left( \bar{\psi}_{n,m}^{(\alpha,\beta)} \right)^{j}, f^{j} \right\rangle_{wav} \prod_{j=1}^{d} \hat{\psi}_{n,m}^{(\alpha,\beta)}(x^{j}) \\
&= \sum_{n=1}^{2^{k}} \sum_{m=0}^{M} \left\langle \bar{\Psi}_{n,m}^{(\alpha,\beta)}, f \right\rangle_{\mathcal{H}} \hat{\Psi}_{n,m}^{(\alpha,\beta)}(\mathbf{x}),
\end{aligned}
$$

where

$$
\hat{\Psi}_{n,m}^{(\alpha,\beta)}(\mathbf{x}) = \prod_{j=1}^{d} \hat{\psi}_{n,m}^{(\alpha,\beta)}(x^{j}),
$$

and

$$
\bar{\Psi}_{n,m}^{(\alpha,\beta)}(\mathbf{x}) = \prod_{j=1}^{d} \bar{\psi}_{n,m}^{(\alpha,\beta)}(x^{j}).
$$

So, the familly $\left\{ \hat{\Psi}_{n,m}^{(\alpha,\beta)} \right\}_{\substack{n=1,...,2^{k} \\ m=0,...,M}}$ is a frame for $\mathcal{H}$, and the familly $\left\{ \bar{\Psi}_{n,m}^{(\alpha,\beta)} \right\}_{\substack{n=1,...,2^{k} \\ m=0,...,M}}$ is its dual. $\qquad\square$

*Remark 2*
The space $\hat{H} = span \left\{ \hat{\psi}_{n,m}^{(\alpha,\beta)} \right\}_{\substack{n=1,...,2^{k} \\ m=0,...,M}}$ is the vector space of polynomials of degree less than or equal to $M$. So, we can separate the data by applying theorem 6 of ([6]).

## 4. Numerical tests

In this section, we validate the performance of our regularized Jacobi wavelets on two examples in two-dimensional space and give a geometric comparison between Jacobi kernel, Jacobi wavelets kernel and other kernels : Haar wavelet kernel, Wavelet kernel proposed by Zhang and al ([24]), Hermite and Laguerre polynomial kernel, RBF

and polynomial. In Table 1, we give the kernels which are used for our numerical tests

| Jacobi | $k(x,y) = \prod\limits_{j=1}^{d} \sum\limits_{i=0}^{M} J_i^{(\alpha,\beta)}(x^j) J_i^{(\alpha,\beta)}(y^j)$ |
|---|---|
| Haar | $k(x,y) = \prod\limits_{j=1}^{d} \sum\limits_{i=0}^{k} \sum\limits_{n=1}^{2^i} \psi_{in}^H(x^j) \psi_{in}^H(y^j)$ |
| Hermite | $k(x,y) = \prod\limits_{j=1}^{d} \sum\limits_{i=0}^{M} H_i(x^j) H_i(y^j)$ |
| Laguerre | $k(x,y) = \prod\limits_{j=1}^{d} \sum\limits_{i=0}^{M} L_i^{\alpha}(x^j) L_i^{\alpha}(y^j)$ |
| Wavelet | $k(x,y) = \prod\limits_{j=1}^{d} \cos\left(1.75 \times \left(\frac{x^j - y^j}{a}\right)\right) \exp\left(-\frac{\|x^j - y^j\|^2}{2a^2}\right)$ |
| RBF | $k(x,y) = \prod\limits_{j=1}^{d} \exp\left(-\frac{\|x^j - y^j\|^2}{\gamma^2}\right)$ |
| Polynomial | $k(x,y) = \prod\limits_{j=1}^{d} \left(x^j \times y^j + 1\right)^p$ |

,

Table 1
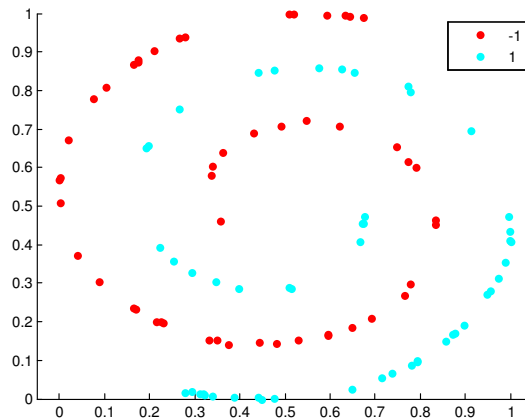


Figure 1. two spiral dataset

with the following functions $\psi_{kn}^H$ : Haar wavelet, $H_i$ : Hermite polynomial, $L_i^\alpha$ : Laguerre polynomial of order $\alpha$. All this kernels have the following tuning parameters, $k$ : dilatation parameter, $n$ : translation parameter, $M$ : the degree of the polynomial, $\gamma$ : RBF parameter and $a$ : Zhang wavelet parameter.

Since Jacobi wavelets are defined on $[0,1]$, the original data should be transformed to $[0,1]$ using this formula

$$newx_i^j = \frac{x_i^j - \min_i\left(x_i^j\right)}{\max_i\left(x_i^j\right) - \min_i\left(x_i^j\right)}, \ \text{ for } i = 1,...,l, \ j = 1,...,d. \tag{14}$$

The kernels parameters are selected by using the ten-folds cross validation. First, the data are normalized by the relation (14). Then, 90% of the normalized data are chosen randomly as training set and the left ones as testing set. The best kernel parameters are the ones that gave the best accuracy over the 10 blocks. The accuracy is given by

the following formula

$$accuracy = \frac{\text{nbr of correctly predicted data}}{\text{nbr of total testing data}} \times 100.$$

We choose the variation of the parameters as follows : the parameter $C$ equal to $2^i$ with $i = -5, ..., 15$. The RBF parameter $\gamma$ equal to $2^i$ with $i = -15, ..., 3$. The degree of the polynomial $N$ varies from 1 to 6. The dilatation parameter of wavelets $k$ varies from 1 to 3 and the parameter $a = 0.5 : 0.5 : 4$. Finally, the parameters $\alpha$ and $\beta$ varie from $-0.75 : 0.25 : 2$.

The programs used are implemented on a pc of 2.2 GHz with 4 Go of RAM for the first example, and on a pc of 3.2 GHz with 4 Go of RAM for the second example. The SVM toolbox has been used is the solver $fitcsvm$ of Matlab version R2014a.

### 4.1. Example 1 : two spiral dataset

We consider the two spiral dataset which consist of points of $\mathbb{R}^2$ on two interwined spirals not lineary separable. The spiral dataset has 100 points, 50 for each class (Figure 1.)

In Table 2, we reported only the best results for each kernel and they are shown in figures 2-9

| Kernel | C | N | K/γ/a | α | β | N SVM | time | accuracy |
|---|---|---|---|---|---|---|---|---|
| Jacobi | 2048 | 6 | / | 0.5 | -0.75 | 30 | 8.3919 | 100% |
| Hermite | 2048 | 3 | / | / | / | 52 | 0.9221 | 90% |
| Laguerre | 128 | 6 | / | -0.5 | / | 58 | 7.2587 | 80% |
| RBF | 2048 | / | γ = 0.5 | / | / | 27 | 0.0807 | 100% |
| polynomial | 8 | 2 | / | / | / | 63 | 0.6528 | 60% |
| Jacobi wavelet | 0.125 | 5 | K=1 | 0 | -0.75 | 47 | 13.1361 | 100% |
| Haar kernel | 0.125 | | K=2 | / | / | 39 | 0.2355 | 90% |
| Wavelet kernel | 512 | / | a =0.5 | / | / | 32 | 6.8874 | 100% |

Table 2

| Kernel | C | N | K/γ/a | α | β | N SVM | time | accuracy |
|---|---|---|---|---|---|---|---|---|
| Jacobi | 2048 | 5 | / | 1.75 | 0.5 | 21 | 4.6894 | 95% |
| Hermite | 32768 | 6 | / | / | / | 125 | 7.4001 | 70% |
| Laguerre | 1024 | 5 | / | -0.25 | / | 56 | 5.1282 | 70% |
| RBF | 8192 | / | γ = 0.1250 | / | / | 56 | 0.039 | 85% |
| polynomial | 512 | 15 | / | / | / | 32 | 2.7863 | 70% |
| Jacobi wavelet | 0.0313 | 1 | K=2 | -0.75 | -0.75 | 175 | 6.817 | 100% |
| Haarkernel | 1 | / | K=0.0313 | / | / | 160 | 0.3193 | 50% |
| Waveletkernel | 32768 | / | a =0.5 | / | / | 26 | 1.5139 | 95% |

Table 3

### 4.2. Example 2 : checkerboard

The second dataset used was the checkerboard consisting of 486 blue points and 514 red points. For calculations reasons, we only consider 200 points of the checkerboard chosen randomly ( see Figure 10. ).
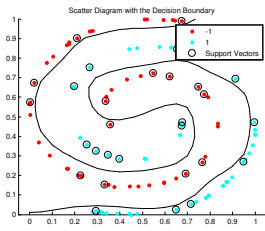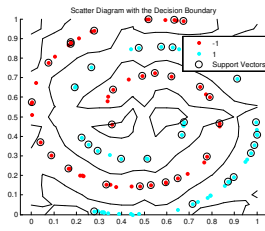
Figure 2. Jacobi kernel
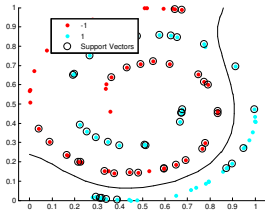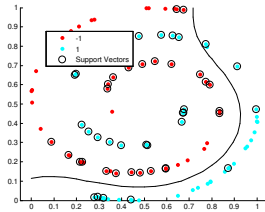
Figure 3. Jacobi wavelet kernel
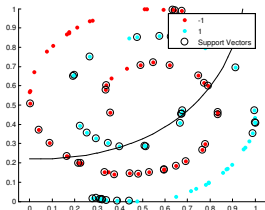
Figure 4. Laguerre kernel
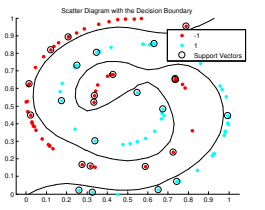
Figure 5. Hermite kernel

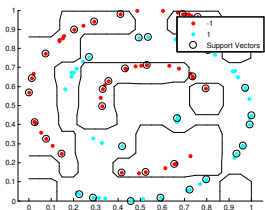Figure 6. polynomial kernel

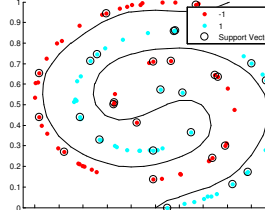Figure 7. RBF kernel

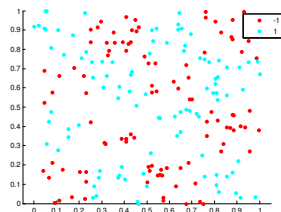Figure 8. Haar wavelet kernel

Figure 9. Wavelet kernel

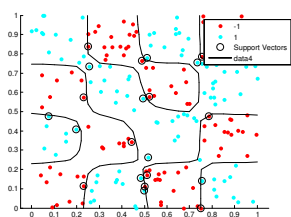Figure 10. checker board with 200 points

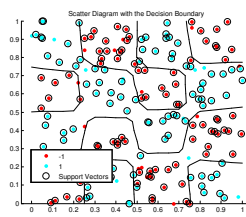Figure 11. Jacobi kernel



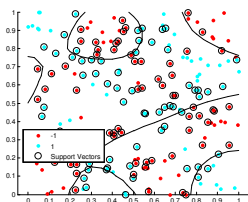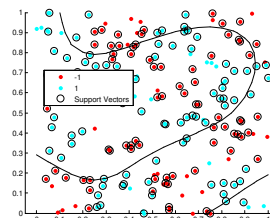Figure 12. Jacobi wavelet kernel



Figure 13. Hermite kernel
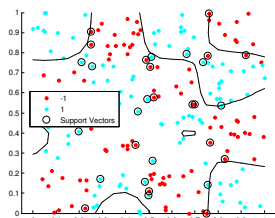


Figure 14. Laguerre kernel
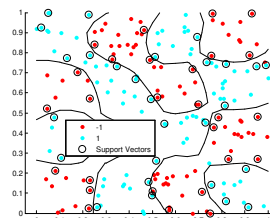


Figure 15. polynomial kernel
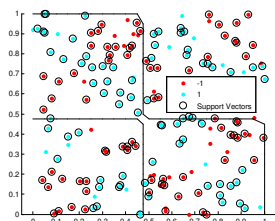


Figure 16. RBF kernel
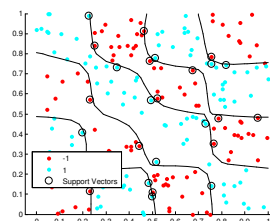


Figure 17. Haar wavelet kernel



Figure 18. Wavelet kernel

The best results are reported in Table 3 and figures 11-18.

## 5. Conclusion

In this paper, a new kernel based on regularized Jacobi wavelets is proposed. We show that the functions belonging to the RKHS defined by the Jacobi wavelets are not all continuous, where the continuity being essential for the separation property. So, we construct another RKHS by introducing regularized Jacobi wavelets that form a frame for the new RKHS. Using frame theory, this new RKHS verified the separation properties. Experimental results show that this kernel can provide competitive results compared to other kernel functions. Kernels based on wavelets and orthogonal polynomials, have the reputation of having an expensive Gram matrix in terms of computation time,

this is known that is a big challenge for this type of kernel. However, we noticed that by reducing the number of the data in the training set, as shown by the numerical tests, our kernel always behaves well with a very competitive execution time.

## REFERENCES

1.  M. Achache  *A weighted full- Newton step primal-dual interior point algorithm for convex quadratic optimization*,  Statistics, Optimization and Information Computing, vol. 2, pp. 21–32, 2014.
2.  A. Bokhari, A. Amir, S. M. Bahri,  *A numerical approach to solve quadratic calculus of variation problems*,  Dynamics of Continuous, Discrete and Impulsive Systems Series B: Applications & Algorithms 25, pp. 427–440, 2018.
3.  C. Burges  *A tutorial on support vector machines for pattern recognition, In Data Mining and Knowledge Discovery*,  Kluwer Academic Publishers, Boston, Vol. 2, 1998.
4.  O. Christensen,  *An introduction to frames and Riesz bases*,  Applied and numerical harmonic analysis, Springer Science Business Media, LLC, 2002.
5.  N. Cristianni, J. Shawe-Taylor,  *An introduction to Support Vector Machines and other kernel-based learning Methods*,  Cambridge University Press, 2000.
6.  N. Djelloul, A. Amir,  *Analysis of Legendre Polynomial Kernel in Support Vector Machines*,  Int. J. Computing Science and Mathematics. To appear (2018).
7.  H. Elaydi, A. A. Abu Haya,  *Solving Optimal Control Problem for Linear Time-invariant Systems via Chebyshev Wavelet*,  International Journal of Electrical Engineering, Vol. 5, no. 5, pp. 541–556, 2012.
8.  Y. Fillali, M. A. Sabri, A. Aarab  *An Improved Segmentation Approach for skin Lesion classification*,  Statistics, Optimization and Information Computing, vol. 7, pp. 456–467, 2019.
9.  V. H. Moghaddam, J. Hamidzadel,  *New Hermite orthogonal polynomial kernel and combined kernels in support vector machine classifier*,  Pattern Recognition 60, Elsevier, pp. 921–935, 2016.
10.  S. Ozer, C. H. Chen and H. A. Cirpan,  *A set of new Chebyshev kernel functions for support vector machine pattern classification*,  Pattern Recognition 44, pp. 1435–1447, Elsevier, 2011.
11.  Z. B. Pan, H. Chan and X. H. You,  *Support vector machine with orthogonal Legendre kernel*,  Proceeding of the 2012 International Conference on wavelet Analysis and Pattern Recognition, Xian, IEEE, 15-17 July 2012.
12.  D. Peijun, T. Kun, X. Xiaoshi,  *Wavelet SVM in Reproducing Kernel Hilbert Space for hyperspectral remote sensing image classification*,  Optics Communications 283, pp. 4978–4984, 2010.
13.  A. Rakotomamonjy, S. Canu,  *Frames, Reproducing Kernels, Regularization and Learning*,  Journal of Machine Learning Research 6, pp. 1485–1515, 2005.
14.  M. Razzaghi, S. Yousefi,  *Legendre wavelets method for constrained optimal control problems*,  Math. Meth. Appl. Sci. 25, pp. 529–539, 2002.
15.  R.A. Ryan,  *Introduction to Tensor Products of Banach Spaces*,  Monographs in Mathematics, Springer, London, 2002.
16.  Spiliopoulou, M. Kruse, R. Borgelt, C. Nürnberger, A. and W. Gaul (Eds.),  *From Data and Information Analysis to Knowledge Engineering*,  Proceedings of the 29th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Magdeburg, Springer- Verlag, 2005.
17.  I. Steinwart, A. Christmann,  *Support vector machines*,  Information Science and Statistics, Springer, 2008.
18.  S. Suvrit, S. Nowozin, and S.J. Wright,  *Optimization for Machine Learning*,  Massachusetts Institute of Technology, London, 2012.
19.  G. Szecö,  *Orthogonal polynomials*,  American mathematical society, colloquim publications, vol. $XXIII$, 1939.
20.  V. Vapnik,  *The Nature of Statistical Learning Theory*,  Springer-Verlag, NewYork, 1995.
21.  L. Wang,  *Support vector machines: theory and applications*,  New York, Springer-Verlag, 2005.
22.  N. YE, R. Sun, Y. Liu and L. Cao,  *Support vector machine with orthogonal Chebyshev kernel*,  IEEE, 2006.
23.  M. A. Zaky, I. G. Ameen and M. A. Abdelkawy,  *A new operational matrix based on Jacobi wavelets for a class of variable- order fractional differential equations*,  proceedings of the Romanian Academy, series A, Vol. 18, no. 4, pp. 315–322, 2017.
24.  L. Zhang, W. Zhou and L. Jiao,  *Wavelet support vector machine*,  IEEE transactions on systems, Man and Cybernetics, Part B, Cybernetics, vol. 34, no. 1, 2004.
25.  F. Zhou, Z. Fang and J. Xu,  *Constructing support vector machine kernels from orthogonal polynomials for face and speaker verification*,  Fourth International Conference on Image and Graphics, IEEE, 2007.