# Comparison of 2D Regular Lattices for the CPWL Approximation of Functions


1st Mehrsa Pourya
*EPFL*
Lausanne, Switzerland
mehrsa.pourya@epfl.ch

2nd Maïka Nogarotto
*EPFL*
Lausanne, Switzerland
maika.nogarotto@epfl.ch

3rd Michael Unser
*EPFL*
Lausanne, Switzerland
michael.unser@epfl.ch



***Abstract*—We investigate the approximation error of functions with continuous and piecewise-linear (CPWL) representations. We focus on the CPWL search spaces generated by translates of box splines on two-dimensional regular lattices. We compute the approximation error in terms of the stepsize, length ratio, and angle that define the lattice. Our results show that hexagonal lattices are optimal, in the sense that they minimize the asymptotic approximation error.**



***Index Terms*—Approximation error bounds, Cartesian grids, continuous and piecewise linear, Fourier-domain analysis, hexagonal grids.**


## I. Introduction

CONTINUOUS and piecewise-linear (CPWL) representations play a fundamental role in signal processing, computer graphics, and computational mathematics [1]–[3]. They are widely appreciated for their simplicity, computational efficiency, and ability to approximate complex structures. They are also of interest in machine learning because they encompass the same class of functions generated by neural networks with the rectified linear unit (ReLU) activation functions [4]. Box splines extend univariate B-splines to multiple dimensions and provide a structured framework for the construction of CPWL search spaces [5]–[9].

A significant body of research investigates the approximation error associated with CPWL representations [10], [11]. However, much of this work focuses on the upper bounds of the error, and less attention is paid to the exact form of the asymptotic error. Fourier-domain methods such as [12] provide powerful tools to analyze such asymptotic behaviors. However, their final results are limited to the one-dimensional scenario.

In this paper, we focus on the two-dimensional case. We rely on box splines to construct each CPWL function over a domain that is partitioned by triangulations where the vertices are on regular lattices. These lattices and the edges of the triangulation form an underlying grid for each CPWL function. Cartesian and hexagonal grids are the well-known examples. Cartesian grids are simple to use, whilst hexagonal grids are known to have better sampling properties [13]–[15]. Our goal is to investigate the effect of the grid on the approximation error. Our main contributions are as follows.


This work was supported by the European Research Council (ERC Project FunLearn) under Grant 101020573, in part by the Swiss National Science Foundation, Grant 200020_219356.


1) Computation of the approximation error for a general grid: We compute the approximation error for a box-spline-based CPWL search space with a general grid in terms of its length ratio, angle, and stepsize. We provide this result in Theorem 1. In Theorem 2, we present an upper bound for the dominant term of this error in the asymptotic regime. Notably, we compute an asymptotic error constant that depends on the length ratio and the angle that define the grid.
2) Optimality of the hexagonal grid: We show that the asymptotic error constant is minimized for the hexagonal grid. We present examples that validate this result.
3) Relation to ReLU neural networks: We provide in Theorem 3 a concise representation of box splines as two-layer neural networks with ReLU activation functions. To our knowledge, this is the simplest construction of two-dimensional box splines with ReLU functions.

In Section II, we define our basis function and its associated CPWL search spaces, and present our proposed parameterization for the grid. We then formally define the approximation error. In Section III, we present our results for the computation and analysis of this approximation error.

## II. Preliminaries and Problem Formulation

### A. Definition of the CPWL Basis Function

Let $\{\boldsymbol{\xi}_n\}_{n=1}^3$ be three vectors, where $\{\boldsymbol{\xi}_n\}_{n=1}^2$ are linearly independent and $\boldsymbol{\xi}_3 = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2$. We refer to $\boldsymbol{\Xi}$ as the grid matrix and define it as $\boldsymbol{\Xi} = [\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2] \in \mathbb{R}^{2\times 2}$. In this paper, we define our basis function $B_{\boldsymbol{\Xi}}$ through its Fourier transform as

$$\hat{B}_{\boldsymbol{\Xi}}(\boldsymbol{\omega}) = |\det \boldsymbol{\Xi}| \prod_{r=1}^{3} \operatorname{sinc}\left(\frac{\boldsymbol{\xi}_r^\top \boldsymbol{\omega}}{2}\right), \tag{1}$$

where $|\det \boldsymbol{\Xi}|$ is the determinant of $\boldsymbol{\Xi}$ and $\operatorname{sinc}(\omega) = \frac{\sin \omega}{\omega}$. Note that $B_{\boldsymbol{\Xi}}$ is a shifted CPWL box spline that is centered around the origin.

If we choose $\boldsymbol{\Xi} = \mathbf{I}$ (identity matrix), we obtain the Cartesian basis with Fourier transform

$$\hat{\varphi}(\boldsymbol{\omega}) = \operatorname{sinc}\left(\frac{\omega_1 + \omega_2}{2}\right) \prod_{r=1}^{2} \operatorname{sinc}\left(\frac{\omega_r}{2}\right). \tag{2}$$

The formula for the spatial evaluation of $\varphi$ ensues, with

$$\varphi(\boldsymbol{x}) = \max\big(1 + \min(x_1, x_2, 0) - \max(x_1, x_2, 0), 0\big), \tag{3}$$

taken from [9]. The basis $B_{\boldsymbol{\Xi}}$ can be expressed in terms of the (centered) Cartesian box spline $\varphi$ as $B_{\boldsymbol{\Xi}}(\boldsymbol{x}) = \varphi(\boldsymbol{\Xi}^{-1}\boldsymbol{x})$. This follows from a simple change of variable in the Fourier domain. In Figure 1, we present two examples of (centered) box splines: Cartesian box spline with $\boldsymbol{\Xi}_{\text{Cart}} = T\mathbf{I}$, and hexagonal box spline with $\boldsymbol{\Xi}_{\text{Hex}} = T\mathbf{D}_{\text{Hex}}$ with $\mathbf{D}_{\text{Hex}} = (\frac{\sqrt{3}}{2})^{-0.5}\begin{bmatrix}1 & -0.5\\ 0 & \frac{\sqrt{3}}{2}\end{bmatrix}$ for some $T > 0$.

### B. CPWL Search Space

We define the CPWL search space

$$\mathcal{V}_{\boldsymbol{\Xi}} = \left\{ \sum_{\boldsymbol{k}\in\mathbb{Z}^2} c[\boldsymbol{k}]\, B_{\boldsymbol{\Xi}}(\cdot - \boldsymbol{\Xi}\boldsymbol{k}) : c[\cdot] \in \ell_2(\mathbb{Z}^2) \right\}, \quad (4)$$

or equivalently ,

$$\mathcal{V}_{\boldsymbol{\Xi}} = \left\{ \sum_{\boldsymbol{k}\in\mathbb{Z}^2} c[\boldsymbol{k}]\, \varphi\left(\boldsymbol{\Xi}^{-1}\cdot - \boldsymbol{k}\right) : c[\cdot] \in \ell_2(\mathbb{Z}^2) \right\}. \quad (5)$$

There, we have used the Cartesian box spline $\varphi$ and the relation $B_{\boldsymbol{\Xi}}(\boldsymbol{x}) = \varphi(\boldsymbol{\Xi}^{-1}\boldsymbol{x})$. The translated basis functions $\{B_{\boldsymbol{\Xi}}(\cdot - \boldsymbol{\Xi}\boldsymbol{k})\}_{\boldsymbol{k}\in\mathbb{Z}^2}$ form a Riesz basis of $\mathcal{V}_{\boldsymbol{\Xi}}$, which guarantees a unique link between each mapping $s\colon \mathbb{R}^2 \to \mathbb{R}$ and its expansion coefficients $c$. It can further reproduce any affine mapping [7]. The domain of each function $s \in \mathcal{V}_{\boldsymbol{\Xi}}$ thus consists of triangles with vertices located on the regular lattice $\{\boldsymbol{\Xi}\boldsymbol{k}\}_{\boldsymbol{k}\in\mathbb{Z}^2}$. This lattice and the edges of the triangulation form an underlying grid for each function. The grid lines are parallel to the directions $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$, and $\boldsymbol{\xi}_3 = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2$ where we had that $\boldsymbol{\Xi} = [\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2]$.

### C. Parameterization of the Grid Matrix

We parameterize $\boldsymbol{\Xi}$ by the the scalar $\lambda > 0$, the angle $0 < \delta < 2\pi$ and, the stepsize $T$ as

$$\boldsymbol{\Xi} = \frac{T}{\sqrt{\lambda \sin\delta}} \begin{bmatrix}\lambda & \cos\delta\\ 0 & \sin\delta\end{bmatrix}. \quad (6)$$

Here, $\delta$ controls the angle between the two vectors of the grid and $\lambda$ is the ratio of their lengths. It follows that $|\det \boldsymbol{\Xi}| = T^2$, which ensures that the number of grid points per area remains invariant for grids constructed with different parameters $\lambda$ and $\delta$. This invariance allows for a fair comparison of grids. Our parameterization encompasses the Cartesian grid matrix $\boldsymbol{\Xi}_{\text{Cart}}$ with $\lambda = 1, \delta = \frac{\pi}{2}$, and the hexagonal grid matrix $\boldsymbol{\Xi}_{\text{Hex}}$ with $\lambda = 1, \delta = \frac{2\pi}{3}$. We represent the central parts of the grids constructed with $\boldsymbol{\Xi}_{\text{Cart}}$ and $\boldsymbol{\Xi}_{\text{Hex}}$ in Figure 2.

### D. Formulation of the Problem

For $f \in L_2(\mathbb{R}^2)$, we are interested in the minimum $L_2$-error solution

$$f_{\text{CPWL}} := \underset{s\in\mathcal{V}_{\boldsymbol{\Xi}}}{\arg\min} \|f - s\|_{L_2}. \quad (7)$$

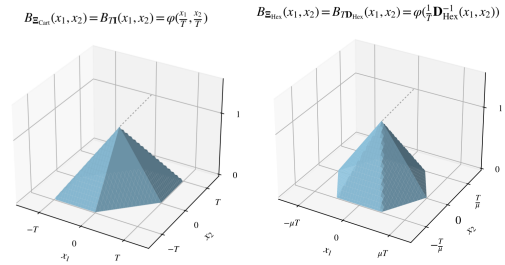

Fig. 1. (Centered) Cartesian (left) and hexagonal ($\mu = (\frac{\sqrt{3}}{2})^{-0.5}$) (right) box splines.

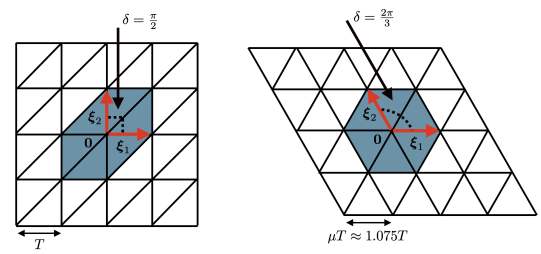


Fig. 2. Cartesian grid with $\boldsymbol{\Xi}_{\text{Cart}} = [\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2] = \begin{bmatrix}T & 0\\ 0 & T\end{bmatrix}$ (left) and hexagonal grid with $\boldsymbol{\Xi}_{\text{Hex}} = [\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2] = \begin{bmatrix}\mu T & -0.5\mu T\\ 0 & 0.5\sqrt{3}\mu T\end{bmatrix}$ and $\mu = (\frac{\sqrt{3}}{2})^{-0.5}$ (right). The highlighted region depicts the support of the corresponding box splines $\varphi_{\Xi_{\text{Cart}}}$ and $\varphi_{\Xi_{\text{Hex}}}$.

This solution could be computed through the projector $P_{\mathcal{V}_{\boldsymbol{\Xi}}}$ as

$$f_{\text{CPWL}}(\boldsymbol{x}) = P_{\mathcal{V}_{\boldsymbol{\Xi}}}\{f\}(\boldsymbol{x}) \quad (8)$$

$$= \sum_{\boldsymbol{k}\in\mathbb{Z}^2} \left\langle f, \frac{1}{|\det \boldsymbol{\Xi}|}\phi(\boldsymbol{\Xi}^{-1}\cdot - \boldsymbol{k}) \right\rangle_{L_2} \varphi(\boldsymbol{\Xi}^{-1}\boldsymbol{x} - \boldsymbol{k}), \quad (9)$$

which is taken from [16] and is a classic result in approximation theory [17]–[19]. The dual basis $\phi$ is defined through its Fourier transform

$$\hat{\phi}(\boldsymbol{\omega}) = \frac{\hat{\varphi}(\boldsymbol{\omega})}{A_{\varphi}(\boldsymbol{\omega})}, \quad (10)$$

with

$$A_{\varphi}(\boldsymbol{\omega}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^2} |\hat{\varphi}(\boldsymbol{\omega} + 2\pi\boldsymbol{k})|^2. \quad (11)$$

We define

$$\epsilon_{\boldsymbol{\Xi}}(f) = \|f - P_{\mathcal{V}_{\boldsymbol{\Xi}}}\{f\}\|_{L_2}. \quad (12)$$

In this paper, we want to quantify the error $\epsilon_{\boldsymbol{\Xi}}(f)$ in terms of the stepsize $T$, the length ratio $\lambda$ and the angle $\delta$ found in (6). In simple words, we want to quantify the effect of the grid on the approximation error.

## III. Methods

We first present our results for the computation of the error $\epsilon_{\boldsymbol{\Xi}}(f)$ and define an asymptotic error constant that depends on the length ratio and the angle of the grid. Then, we investigate the effect of the grid parameters on the approximation error. Mainly, we show that the asymptotic

error constant is minimized for hexagonal grids. Finally, we relate our findings to the approximation that results from ReLU neural networks.

### A. Computation of the Approximation Error

We first present an explicit formula for $A_\varphi$ in Proposition 1, which is crucial for the calculation of the error. Then, in Theorem 1, we provide our result for the computation of $\epsilon_{\boldsymbol{\Xi}}(f)$. Finally, in Theorem 2, we provide an upper bound for the dominant term which involves the asymptotic error constant $C(\lambda,\delta)$.

**Proposition 1.** *Let $\varphi$ be the Cartesian box spline. Then, one has that*

$$A_\varphi(\boldsymbol{\omega}) = \frac{1}{2} + \frac{1}{6}\big(\cos(\omega_1)+\cos(\omega_2)+\cos(\omega_1+\omega_2)\big). \quad (13)$$

*Proof.* The cumbersome computation of the infinite sum in (11) is simplified by the equality

$$\sum_{\boldsymbol{k}\in\mathbb{Z}^2} \left|\hat{\varphi}(\boldsymbol{\omega}+2\pi\boldsymbol{k})^2\right| = \mathrm{DTFT}\Big\{\big(\varphi * \varphi(-\cdot)\big)(\boldsymbol{k})\Big\}. \quad (14)$$

To compute the autocorrelation sequence $a[\boldsymbol{k}] = \big(\varphi * \varphi(-\cdot)\big)(\boldsymbol{k})$, $\boldsymbol{k}\in\mathbb{Z}^2$, we use the spatial evaluation of $\varphi$ in (3) and compute the result integrals to obtain that

$$a[\boldsymbol{k}] = \begin{cases} \frac{1}{2}, & \boldsymbol{k}=\boldsymbol{0} \\ \frac{1}{12}, & \boldsymbol{k}\in\{(0,\pm1),(\pm1,0),(-1,-1),(1,1)\} \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

Now, to complete the proof and obtain (13), we use the definition $\mathrm{DTFT}\big\{a[\boldsymbol{k}]\big\}(\omega) = \sum_{\boldsymbol{k}\in Z^2} a[\boldsymbol{k}]\mathrm{e}^{-\mathrm{j}\boldsymbol{k}^\top\boldsymbol{\omega}}$ and the equality $\cos(\cdot) = \frac{1}{2}(\mathrm{e}^{\mathrm{j}\cdot}+\mathrm{e}^{-\mathrm{j}\cdot})$. □

**Theorem 1.** *For a band-limited $f\in W_2^\rho$ (Sobolov space of order $\rho$) with $\rho>2$ and for a general grid $\boldsymbol{\Xi}$ defined with $0<T<1$ and $\lambda$ and $\delta$ from* (6)*, it holds that*

$$\epsilon_{\boldsymbol{\Xi}}(f) = \epsilon_{\boldsymbol{\Xi},asym}(f) + \mathcal{O}(T^{\min(\rho,3)}), \quad (16)$$

*where*

$$\epsilon_{\boldsymbol{\Xi},asym}(f) = \frac{T^2}{12\sqrt{5}}\Big(\int_{\mathbb{R}^2}\langle \mathbf{H}_f(\boldsymbol{x}), \begin{bmatrix}\alpha & \gamma\\ 0 & \beta\end{bmatrix}\rangle^2 \mathrm{d}\boldsymbol{x}\Big)^{\frac{1}{2}}. \quad (17)$$

*There, $\mathbf{H}_f$ denotes the Hessian of $f$ and*

$$\alpha = \frac{\cos^2\delta+\lambda\cos\delta+\lambda^2}{|\lambda\sin\delta|}, \quad \beta = \frac{\sin^2\delta}{|\lambda\sin\delta|},$$
$$\gamma = \frac{\sin 2\delta+\lambda\sin\delta}{|\lambda\sin\delta|}. \quad (18)$$

*Proof.* For a general grid $\boldsymbol{\Xi}$, a change of variables lead to

$$\epsilon_{\boldsymbol{\Xi}}(f) = \epsilon_{T\mathbf{I}}\left(f\left(\frac{1}{T}\boldsymbol{\Xi}\cdot\right)\right). \quad (19)$$

From Section 2.4 of [16], with $f\in W_2^\rho$ and $\rho>2$, it holds that

$$\epsilon_{T\mathbf{I}}(f) = \epsilon_{\mathrm{dom},T\mathbf{I}}(f) + O(T^\rho), \quad (20)$$

where

$$\epsilon_{\mathrm{dom},T\mathbf{I}}(f) = \Big[\frac{1}{4\pi^2}\int_{\mathbb{R}^2}\epsilon_{\phi,\varphi}(\boldsymbol{\omega}T)\left|\hat{f}(\boldsymbol{\omega})\right|^2\mathrm{d}\boldsymbol{\omega}\Big]^{\frac{1}{2}}. \quad (21)$$

The error kernel $\epsilon_{\phi,\varphi}$ is defined as

$$\epsilon_{\phi,\varphi}(\boldsymbol{\omega}) = 1 - \frac{|\hat{\varphi}(\boldsymbol{\omega})|^2}{A_\varphi(\boldsymbol{\omega})}. \quad (22)$$

In our case, from (2) and Proposition 1, (22) simplifies to

$$\epsilon_{\phi,\varphi}(\boldsymbol{\omega}) = 1 \quad (23)$$
$$- \frac{\prod_{r=1}^2 \mathrm{sinc}^2(\frac{\omega_r}{2})\mathrm{sinc}^2(\omega_1+\omega_2)}{\frac{1}{2}+\frac{1}{6}\big(\cos(\omega_1)+\cos(\omega_2)+\cos(\omega_1+\omega_2)\big)}. \quad (24)$$

By a Taylor series around $\mathbf{0}$, we get that

$$\epsilon_{\phi,\varphi}(\boldsymbol{\omega}T) = \frac{T^4}{720}(\omega_1^2+\omega_1\omega_2+\omega_2^2)^2 + \mathcal{O}(T^6). \quad (25)$$

Then, it follows that

$$\epsilon_{\mathrm{dom},T\mathbf{I}}(f)^2 = \epsilon_{\mathrm{Taylor},T\mathbf{I}}(f)^2 + \mathcal{O}(T^6), \quad (26)$$

where we used that $\int_{\mathbb{R}^2}\mathcal{O}(T^6)\left|\hat{f}(\boldsymbol{\omega})\right|^2\mathrm{d}\boldsymbol{\omega} = \mathcal{O}(T^6)$ due to $f$ being band-limited, and

$$\epsilon_{\mathrm{Taylor},T\mathbf{I}}(f)^2 = \frac{T^4}{2880\pi^2}\int_{\mathbb{R}^2}\big((\omega_1^2+\omega_1\omega_2+\omega_2^2)\hat{f}(\boldsymbol{\omega})\big)^2\mathrm{d}\boldsymbol{\omega}. \quad (27)$$

From (19) and (20), we have that

$$\epsilon_{\boldsymbol{\Xi}}(f) = \epsilon_{\mathrm{Taylor},T\mathbf{I}}(f(\frac{1}{T}\boldsymbol{\Xi}\cdot)) + \mathcal{O}(T^\rho) + \mathcal{O}(T^3)$$
$$= \epsilon_{\mathrm{Taylor},T\mathbf{I}}(f(\frac{1}{T}\boldsymbol{\Xi}\cdot)) + \mathcal{O}(T^{\min(\rho,3)}). \quad (28)$$

Next, we define $\mathbf{D} = \frac{1}{T}\boldsymbol{\Xi}$ and compute

$$\epsilon_{\mathrm{Taylor},T\mathbf{I}}(f(\frac{1}{T}\boldsymbol{\Xi}\cdot))^2 = \frac{1}{4\pi^2}\int_{\mathbb{R}^2}\epsilon_{\phi,\varphi}(\boldsymbol{\omega}T)\big(\hat{f}(\mathbf{D}^{-\top}\boldsymbol{\omega})\big)^2\mathrm{d}\boldsymbol{\omega}$$
$$= \frac{1}{4\pi^2}\int_{\mathbb{R}^2}\epsilon_{\phi,\varphi}(\mathbf{D}^\top\boldsymbol{z}T)\big(\hat{f}(\boldsymbol{z})\big)^2\mathrm{d}\boldsymbol{z}$$
$$= \frac{T^4}{2880\pi^2}\int_{\mathbb{R}^2}\big((\alpha z_1^2+\gamma z_1z_2+\beta z_2^2)\hat{f}(\boldsymbol{z})\big)^2\mathrm{d}\boldsymbol{z}$$
$$= \frac{T^4}{720}\int_{\mathbb{R}^2}\big(\alpha\frac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2}+\gamma\frac{\partial^2 f(\boldsymbol{x})}{\partial x_1\partial x_2}+\beta\frac{\partial^2 f(\boldsymbol{x})}{\partial x_2^2}\big)^2\mathrm{d}\boldsymbol{x}$$
$$= \frac{T^4}{720}\int_{\mathbb{R}^2}\langle\mathbf{H}_f(\boldsymbol{x}), \begin{bmatrix}\alpha & \gamma\\ 0 & \beta\end{bmatrix}\rangle^2\mathrm{d}\boldsymbol{x}. \quad (29)$$

□

For $T<1$, $\epsilon_{\boldsymbol{\Xi},\mathrm{asym}}(f)$ is the dominant term of the error. Moreover, in the asymptotic case $T\to 0$, we have that $\epsilon_{\boldsymbol{\Xi}}(f) = \epsilon_{\boldsymbol{\Xi},\mathrm{asym}}(f)$. Now, we present an upper bound for $\epsilon_{\boldsymbol{\Xi},\mathrm{asym}}(f)$.

**Theorem 2.** *For $\epsilon_{\boldsymbol{\Xi},asym}(f)$ under the same conditions as there in Theorem* (1)*, it holds that*

$$\epsilon_{\boldsymbol{\Xi},asym}(f) \le \frac{1}{12\sqrt{5}}C(\lambda,\delta)T^2\left\|\mathbf{H}_f\right\|_{F,L_2}, \quad (30)$$

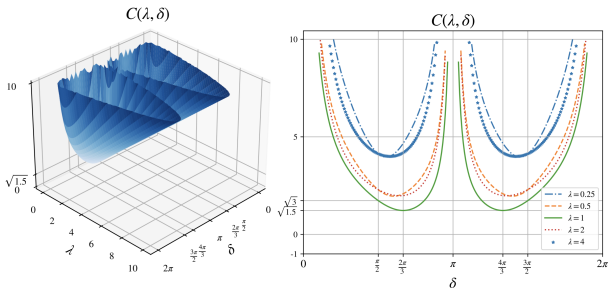


Fig. 3. Error constant $C(\lambda,\delta)$ (left) and one-dimensional profiles of $C(\lambda,\delta)$ for $\lambda \in \{0.25, 0.5, 1, 2, 4\}$ (right). In both plots, we only show $C(\lambda,\delta)$ where $C(\lambda,\delta) < 10$ for better visual representation.

*where we define the asymptotic error constant* $C(\lambda,\delta) = (\alpha^2+\beta^2+\gamma^2)^{\frac{1}{2}}$ *through* $\alpha, \beta,$ *and* $\gamma$ *given in* (1)*, along with the mixed norm* $\|\mathbf{A}(\cdot)\|_{F,L_2} := \big(\int_{\mathbb{R}^2} \|\mathbf{A}(\boldsymbol{x})\|_F^2 \,\mathrm{d}\boldsymbol{x}\big)^{\frac{1}{2}}$.

*Proof.* Hölder's inequality for matrices yields that

$$\langle \mathbf{H}_f(\boldsymbol{x}), \begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix} \rangle^2 \leq \|\mathbf{H}_f(\boldsymbol{x})\|_F^2 \,(\alpha^2+\beta^2+\gamma^2). \quad (31)$$

Then, through (17), we have that

$$\epsilon_{\boldsymbol{\Xi},\text{asym}}(f)^2 \leq \frac{T^4(\alpha^2+\beta^2+\gamma^2)}{720} \int_{\mathbb{R}^2} \|\mathbf{H}_f(\boldsymbol{x})\|_F^2 \,\mathrm{d}\boldsymbol{x}. \quad (32)$$

□

### B. Analysis of the Effect of the Grid

Now, we investigate the effect of the length ratio $\lambda$ and the angle $\delta$ on the error constant $C(\lambda,\delta)$. In Figure 3, we plot $C(\lambda,\delta)$. We observe that for different values of $\delta$, the error constant is minimized at $\lambda = 1$. This suggests that the equilateral grids perform better in terms of approximation error. For such grids, the minimum of $C(\lambda,\delta)$ is $\sqrt{1.5}$ and is achieved when $\lambda = 1$ and $\delta = \frac{2\pi}{3}$ or $\delta = \frac{4\pi}{3}$, which corresponds to a hexagonal grid. Note that $\sqrt{1.5}$ is also the minimum value of $C(\lambda,\delta)$.

We now present explicit formulas for $\epsilon_{\boldsymbol{\Xi},\text{asym}}(f)$ in the case of the Cartesian ($\boldsymbol{\Xi} = \boldsymbol{\Xi}_{\text{Cart}}$) and hexagonal ($\boldsymbol{\Xi} = \boldsymbol{\Xi}_{\text{Hex}}$) grids that are defined in Section II.1. In these two cases, we illustrate the computation of $\epsilon_{\boldsymbol{\Xi},\text{asym}}(f)$ for a function $f$ whose Fourier response is a disk function.

*1) Error for the Cartesian and Hexagonal Grids:* For the Cartesian grid ($\lambda = 1, \delta = \frac{\pi}{2}$), we have that $\alpha = \beta = \gamma = 1$ and $C(1, \frac{\pi}{2}) = \sqrt{3}$; therefore, (17) simplifies to

$$\epsilon_{\boldsymbol{\Xi}_{\text{Cart}},\text{asym}}(f) = \frac{T^2}{12\sqrt{5}} \Big( \int_{\mathbb{R}^2} \big( \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2} + \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_2} + \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2^2} \big)^2 \mathrm{d}\boldsymbol{x} \Big)^{\frac{1}{2}}. \quad (33)$$

For the hexagonal grid ($\lambda = 1, \delta = \frac{2\pi}{3}$), we have that $\alpha = \beta = \frac{\sqrt{3}}{2}$ and $\gamma = 0$, which then yields the elegant formula

$$\epsilon_{\boldsymbol{\Xi}_{\text{Hex}},\text{asym}}(f) = \frac{T^2}{8\sqrt{15}} \Big( \int_{\mathbb{R}^2} \big( \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2} + \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2^2} \big)^2 \mathrm{d}\boldsymbol{x} \Big)^{\frac{1}{2}} = \frac{T^2}{8\sqrt{15}} \|\Delta f\|_{L_2}. \quad (34)$$

where $\Delta$ represents the Laplacian operator. We recall that in this case $C(1, \frac{2\pi}{3}) = \sqrt{1.5}$, which is the minimum achieved by $C(\lambda,\delta)$.

*2) Error Computation for a Fourier Disk:* In this example, we define the function $f$ through its Fourier transform as

$$\hat{f}(\boldsymbol{\omega}) = \begin{cases} c, & \sqrt{w_1^2 + w_2^2} \leq \omega_{\text{max}} \\ 0, & \text{otherwise.} \end{cases} \quad (35)$$

for some $\omega_{\text{max}} > 0$ and $c \in \mathbb{R}$. We then compute $\epsilon_{\boldsymbol{\Xi},\text{asym}}(f)$ for Cartesian and hexagonal grids with the help of (29) as

$$\epsilon_{\boldsymbol{\Xi}_{\text{Cart}},\text{asym}}(f) = \frac{T^2 \,|c|\, \omega_{\text{max}}^3}{\sqrt{7680\pi}}$$
$$\epsilon_{\boldsymbol{\Xi}_{\text{Hex}},\text{asym}}(f) = \frac{T^2 \,|c|\, \omega_{\text{max}}^3}{\sqrt{11520\pi}}. \quad (36)$$

Therefore, we have that $\epsilon_{\boldsymbol{\Xi}_{\text{Hex}},\text{asym}}(f) < \epsilon_{\boldsymbol{\Xi}_{\text{Cart}},\text{asym}}(f)$ for a Fourier disk. This observation confirms our claim that hexagonal grids are better in terms of the approximation error.

### C. Box Splines as ReLU Networks

We now focus on interpolation with box splines on the compact domain $\Omega = (0,1)^2$. Consequently, each function $s \in \mathcal{V}_{T\mathbf{I}}$ can be constructed using $N = (\frac{1}{T} - 1)^2$ nonzero basis functions for $T < 1$.

**Theorem 3.** *We can represent the Cartesian box spline* $\varphi$ *using the ReLU*$(\cdot) := \max(\cdot, 0)$ *function as*

$$\varphi(\boldsymbol{x}) = \mathit{ReLU}(1 - \mathit{ReLU}(x_1 - x_2) - \mathit{ReLU}(x_2) - \mathit{ReLU}(-x_1))). \quad (37)$$

*Proof.* Let us enumerate all configurations of the inequality between $x_1$ and $x_2$ and 0. Then, we observe that (3) and (37) are equal in all cases. □

While the relation between box splines and the ReLU function has been investigated in many works [4], [20]–[22], Theorem 3 provides the simplest representation for a two-dimensional box spline (a.k.a finite-element basis) through ReLU networks, up to our knowledge. Combining this theorem with 5, we conclude that any $s \in \mathcal{V}_{T\mathbf{I}}$ can be constructed using a ReLU network with two hidden layers and $M = 4N$ neurons in total. Moreover, the interpolation error decays at the same rate as (33) where $T = (0.5\sqrt{M}+1)^2$ now depends on the total number of neurons $M$. One can generalize this result to any spline $B_{\boldsymbol{\Xi}}$ by using $B_{\boldsymbol{\Xi}}(\boldsymbol{x}) = \varphi(\boldsymbol{\Xi}^{-1}\boldsymbol{x})$, and proper handling of the domain $\Omega$.

## IV. Conclusion

We have presented an analysis of the approximation error with continuous and piecewise-linear (CPWL) representations using box splines on two-dimensional grids. By deriving explicit error bounds in terms of the grid parameters, we have shown that hexagonal grids minimize the upper bound of the asymptotic error, which emphasizes their optimality in CPWL-based applications.

## REFERENCES


[1] K. Eriksson, D. Estep, and C. Johnson, *Piecewise Linear Approximation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 741–753.

[2] G. Ciarlet Philippe, "The finite element method for elliptic problems," *Studies in Mathematics and Its Aplications. North-Holland Publishing Company Amsterdam. New York-Oxford*, 1978.

[3] M. Pourya, A. Goujon, and M. Unser, "Delaunay-triangulation-based learning with Hessian total-variation regularization," *IEEE Open Journal of Signal Processing*, vol. 4, pp. 167–178, 2023.

[4] R. Arora, A. Basu, P. Mianjy, and A. Mukherjee, "Understanding deep neural networks with rectified linear units," *International Conference on Learning Representations (ICLR)*, 2018.

[5] C. De Boor, K. Höllig, and S. Riemenschneider, *Box Splines*. Springer Science & Business Media, 2013, vol. 98.

[6] M. Kim and J. Peters, "A practical box spline compendium," *Applied Mathematics and Computation*, vol. 464, p. 128376, 2024.

[7] A. Goujon, J. Campos, and M. Unser, "Stable parameterization of continuous and piecewise-linear functions," *Applied and Computational Harmonic Analysis*, vol. 67, p. 101581, 2023.

[8] M. Pourya, Y. Haouchat, and M. Unser, "A continuous-domain solution for computed tomography with Hessian total-variation regularization," in *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, 2024, pp. 1–5.

[9] M. Pourya, A. Boquet-Pujadas, and M. Unser, "A box-spline framework for inverse problems with continuous-domain sparsity constraints," *IEEE Transactions on Computational Imaging*, 2024.

[10] S. Bertoluzza, R. H. Nochetto, A. Quarteroni, K. G. Siebert, A. Veeser, R. H. Nochetto, and A. Veeser, "Primer of adaptive finite element methods," *Multiscale and Adaptivity: Modeling, Numerics and Applications: CIME Summer School, Cetraro, Italy 2009, Editors: Giovanni Naldi, Giovanni Russo*, pp. 125–225, 2012.

[11] R. A. DeVore, "Nonlinear approximation," *Acta numerica*, vol. 7, pp. 51–150, 1998.

[12] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques. i. Interpolators and projectors," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2783–2795, 1999.

[13] R. M. Mersereau, "The processing of hexagonally sampled two-dimensional signals," *Proceedings of the IEEE*, vol. 67, no. 6, pp. 930–949, 1979.

[14] D. Van De Ville, T. Blu, M. Unser, W. Philips, I. Lemahieu, and R. Van de Walle, "Hex-splines: A novel spline family for hexagonal lattices," *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 758–772, 2004.

[15] J. Campos, S. Aziznejad, and M. Unser, "Learning of continuous and piecewise-linear functions with Hessian total-variation regularization," *IEEE Open Journal of Signal Processing*, vol. 3, pp. 36–48, 2022.

[16] S. Ramani, *Nonideal Sampling and Regularized Interpolation of Noisy Data*. EPFL, 2009.

[17] C. Deboor, R. Devore, and A. Ron, "The structure of finitely generated shift-invariant spaces in l2(Rd)," *Journal of Functional Analysis*, vol. 119, no. 1, pp. 37–78, 1994.

[18] M. Unser, "Approximation power of biorthogonal wavelet expansions," *IEEE Transactions on Signal Processing*, vol. 44, no. 3, pp. 519–527, 1996.

[19] T. Blu and M. Unser, "Approximation error for quasi-interpolators and (multi-)wavelet expansions," *Applied and Computational Harmonic Analysis*, vol. 6, no. 2, pp. 219–251, 1999.

[20] P. R. Massopust and P. J. Van Fleet, "Fractional cone splines and hex splines," *arXiv preprint arXiv:1504.00546*, 2015.

[21] L. Condat and D. Van De Ville, "Three-directional box-splines: characterization and efficient evaluation," *IEEE Signal Processing Letters*, vol. 13, no. 7, pp. 417–420, 2006.

[22] J. He, L. Li, and J. Xu, "ReLu deep neural networks from the hierarchical basis perspective," *Computers & Mathematics with Applications*, vol. 120, pp. 105–114, 2022.