Mapped Variably Scaled Kernels: Applications to Solar Imaging

Francesco Marchetti[1(✉)], Emma Perracchione[2], Anna Volpara[3], Anna Maria Massone[3], Stefano De Marchi[1], and Michele Piana[3]

[1] University of Padova, Padua, Italy
{francesco.marchetti,stefano.demarchi}@unipd.it
[2] Polytechnic of Torino, Turin, Italy
emma.perracchione@polito.it
[3] University of Genova, Genoa, Italy
{volpara,massone,piana}@dima.unige.it

Abstract. Variably scaled kernels and mapped bases constructed via the so-called fake nodes approach are two different strategies to provide adaptive bases for function interpolation. In this paper, we focus on kernel-based interpolation and we present what we call mapped variably scaled kernels, which take advantage of both strategies. We present some theoretical analysis and then we show their efficacy via numerical experiments. Moreover, we test such a new basis for image reconstruction tasks in the framework of hard X-ray astronomical imaging.

Keywords: Variably scaled kernels · Mapped bases interpolation · Hard X-ray imaging

1 Introduction

Kernel-based interpolation is an effective approach to deal with the scattered data interpolation problem, where data sites do not necessarily belong to some particular structure or grid [17,30]. Therefore, because of its flexibility and the achievable accuracy, it finds application in many different contexts [20], including image reconstruction and the numerical solution of partial differential equations. The effectiveness of radial kernel-based interpolation, which is also known as Radial Basis Function (RBF) interpolation, very often relies on a good choice of the so-called *shape parameter*, which rules the shape of basis functions. However, in many situations, a fine tuning of this shape parameter is not sufficient to construct basis functions that are tailored with respect to both the data distribution and the target function to be recovered.

Therefore, in order to gain more adaptivity in the interpolation process, Variably Scaled Kernels (VSKs) have been introduced in [6], and further analyzed in [8] in the more general framework of kernel-based regression networks, and in [16] in the context of persistent homology. The VSK setting has also been employed in the reconstruction of functions presenting jumps, leading to the

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
O. Gervasi et al. (Eds.): ICCSA 2023 Workshops, LNCS 14108, pp. 577–592, 2023.
https://doi.org/10.1007/978-3-031-37117-2_39

definition of Variably Scaled Discontinuous Kernels (VSDKs) [13], which turned out to be effective in medical image reconstruction tasks in the field of Magnetic Particle Imaging (MPI) [12].

An alternative approach for constructing data-dependent or target-dependent basis functions, the so-called Fake Nodes Approach (FNA), has been introduced in [14] in the framework of univariate polynomial interpolation, and then extended to rational barycentric approximation [4] and general multivariate interpolation, including the RBF framework [15]. In particular, in latter paper the authors showed that the VSDK setting and the FNA can lead to very similar results when facing the Gibbs phenomenon in the reconstruction of discontinuous functions. Moreover, in [11] an effective scheme for dealing with both Gibbs and Runge's phenomena has been proposed.

In this paper, our purpose is to design a unified kernel-based approach for function interpolation by taking advantage of both the VSK setting and the FNA, which are recalled in Sect. 2 and 3, respectively. The resulting kernels that we call Mapped VSKs (MVSKs) are defined in Sect. 4 and provided with an original theoretical contribution, which includes a focus on the discontinuous case tested in Sect. 5. In Sect. 6 we present the application of the proposed method to the inverse problems of solar hard X-ray imaging and, specifically, we consider data from the on the ESA *Spectrometer/Telescope for Imaging X-rays (STIX)* telescope, on board of Solar Orbiter mission. Finally, in Sect. 7 we draw some conclusions.

2 Variably Scaled Kernel-Based Approximation

We refer to [17,30] for the following introduction.

Let $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, and $X = X_N = \{\boldsymbol{x}_i,\ i = 1, \dots, N\} \subset \Omega$ be a set of possible scattered distinct nodes, $N \in \mathbb{N}$. Suppose that we wish to reconstruct an unknown function $f : \Omega \longrightarrow \mathbb{R}$ from its values at X_N, i.e. from the vector $\boldsymbol{f} = (f(\boldsymbol{x}_1), \dots, f(\boldsymbol{x}_N))^\intercal = (f_1, \dots, f_N)^\intercal$. In kernel-based interpolation, this is obtained by considering an approximating function of the form

$$R_{f,X}(\boldsymbol{x}) = \sum_{i=1}^{N} c_i \kappa_\varepsilon(\boldsymbol{x}, \boldsymbol{x}_i), \quad \boldsymbol{x} \in \Omega,$$

where $\boldsymbol{c} = (c_1, \dots, c_N)^\intercal \in \mathbb{R}^N$ and $\kappa_\varepsilon : \Omega \times \Omega \longrightarrow \mathbb{R}$ is a strictly positive definite kernel, which depends on a *shape parameter* $\varepsilon > 0$. In the following, we may use the shortened notation $\kappa = \kappa_\varepsilon$. The interpolation conditions are imposed by employing $\boldsymbol{c}$ so that

$$\mathsf{K}\boldsymbol{c} = \boldsymbol{f}, \tag{1}$$

where $\mathsf{K} = (\mathsf{K}_{i,j}) = \kappa(\boldsymbol{x}_i, \boldsymbol{x}_j)$, $i, j = 1, \dots, N$, is the so-called interpolation (or collocation or simply kernel) matrix. The vector $\boldsymbol{c}$ that satisfies (1) is unique as long as κ is strictly positive definite. Moreover, we assume κ_ε to be *radial*, i.e. there exists a univariate function $\varphi_\varepsilon : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}$ such that $\kappa_\varepsilon(\boldsymbol{x}, \boldsymbol{y}) = \varphi_\varepsilon(r)$, with $r := \|\boldsymbol{x} - \boldsymbol{y}\|_2$.

The interpolant $R_{f,X}$ belongs to the dot-product space

$$\mathcal{H}_\kappa = \operatorname{span}\left\{\kappa(\cdot, \boldsymbol{x}),\ \boldsymbol{x} \in \Omega\right\},$$

whose completion with respect to the norm $\|\cdot\|_{\mathcal{H}_\kappa} = \sqrt{(\cdot,\cdot)_{\mathcal{H}_\kappa}}$ induced by the bilinear form $(\cdot,\cdot)_{\mathcal{H}_\kappa}$ is the *native space* $\mathcal{N}_\kappa$ associated to κ. The well-known pointwise error bound [17, Theorem 14.2, p. 117]

$$|f(\boldsymbol{x}) - R_{f,X}(\boldsymbol{x})| \leq P_{\kappa,X}(\boldsymbol{x})\|f\|_{\mathcal{N}_\kappa}, \quad f \in \mathcal{N}_\kappa, \quad \boldsymbol{x} \in \Omega, \tag{2}$$

involves the *power function* $P_{\kappa,X} = \|\kappa(\cdot, x) - \boldsymbol{\kappa}(\cdot)^\intercal \mathsf{K}^{-1}\boldsymbol{\kappa}(\boldsymbol{x})\|_{\mathcal{H}_\kappa}$, where $\boldsymbol{\kappa}(\boldsymbol{x}) := (\kappa(\boldsymbol{x}, \boldsymbol{x}_1), \ldots, \kappa(\boldsymbol{x}, \boldsymbol{x}_N))^\intercal$. In the estimate (2), a noteworthy property is that the error is split into a first term that only depends on the nodes and on the kernel, and a second term that relies on the underlying function f.

A different perspective is provided by the following error bound, which takes into account the so-called *fill distance*

$$h_{X,\Omega} := \sup_{\boldsymbol{x}\in\Omega} \min_{\boldsymbol{x}_k \in X} \|\boldsymbol{x} - \boldsymbol{x}_k\|_2.$$

Assuming Ω to be bounded and satisfying an interior cone condition, and $\kappa \in C^{2k}(\Omega \times \Omega)$, there exists $C_\kappa(\boldsymbol{x})$ such that

$$|f(\boldsymbol{x}) - R_{f,X}(\boldsymbol{x})| \leq C_\kappa(\boldsymbol{x}) h_{X,\Omega}^k \|f\|_{\mathcal{N}_\kappa(\Omega)}. \tag{3}$$

The factor $C_\kappa(\boldsymbol{x})$ depends on the maximum of kernel derivatives of degree $2k$ in a neighborhood of $\boldsymbol{x} \in \Omega$, and $h_{X,\Omega}^k$ needs to be *small enough*; we refer to [17, Section 14.5] for a detailed presentation of this result.

While the theoretically achievable convergence rate is influenced by the fill distance and the smoothness of the kernel, two terms play an important role in affecting the conditioning of the interpolation process:

- The *separation distance* $q_X := \dfrac{1}{2}\min_{i\neq j}\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2$.
- The value of the shape parameter ε of the kernel κ_ε.

Precisely, the interpolation process gets more ill-conditioned as the separation distance becomes smaller, which is what usually happens in practice when increasing the number of interpolation nodes, thus reducing the value of the fill-distance (this is often denoted as a *trade-off* principle in RBF literature). As far as the shape parameter is concerned, a large value produces very localized basis functions that lead to a well-conditioned but likely inaccurate approximation scheme, while by lowering such value we may obtain a more accurate reconstruction at the price of an ill-conditioned setting. As a consequence, the tuning of the shape parameter is a non trivial problem in kernel-based interpolation.

In order to partially overcome such instability issues, in [6] the authors introduced Variably Scaled Kernels (VSKs), which are defined as follows. Letting $\psi : \Omega \longrightarrow \mathbb{R}$ be a scaling or shape function and $\kappa : (\Omega \times \mathbb{R}) \times (\Omega \times \mathbb{R}) \longrightarrow \mathbb{R}$ a kernel, a VSK $\kappa_\psi : \Omega \times \Omega \longrightarrow \mathbb{R}$ is defined as

$$\kappa_\psi(\boldsymbol{x}, \boldsymbol{y}) := \kappa((\boldsymbol{x}, \psi(\boldsymbol{x})), (\boldsymbol{y}, \psi(\boldsymbol{y}))),$$

for $\boldsymbol{x}, \boldsymbol{y} \in \Omega$. Note that we can consider the related native space $\mathcal{N}_{\kappa_\psi}$ spanned by the functions $\kappa_\psi(\cdot, \boldsymbol{x})$, $\boldsymbol{x} \in \Omega$. Among the properties of VSKs, we recall that:

A1. The theoretical analysis of the VSK setting reduces to the analysis of the classical framework in the augmented space $\Omega \times \mathbb{R}$.
A2. The spaces $\mathcal{N}_\kappa$ and $\mathcal{N}_{\kappa_\psi}$ are isometrically isomorphic.

Although we may consider different values of ε, as done e.g. in the context of moving least squares [21], in the VSK framework the shape parameter is often set to $\varepsilon = 1$. Therefore, the role played by the shape function is duplex. On the one hand, the tuning of the shape parameter is substituted with the choice of the shape function. On the other hand, ψ can lead to an improvement in the conditioning of the approximation process by increasing the value of the separation distance in the augmented domain $\Omega \times \mathbb{R}$. Note that κ_ψ is (strictly) positive definite if so is κ.

3 Interpolation via Mapped Bases

In this section, we present the so called Fake Nodes Approach (FNA) for the kernel-based approximation framework; for a more general and comprehensive treatment we refer to [15].

Let $S : \Omega \longrightarrow \mathbb{R}^d$ be an injective map. Our purpose is to construct an interpolant $R^S_{f,X}$ of the function f in the space

$$\mathcal{H}_{\kappa^S} = \operatorname{span}\left\{\kappa^S(\cdot, \boldsymbol{x}),\ \boldsymbol{x} \in \Omega\right\},$$

where $\kappa^S(\boldsymbol{x}, \boldsymbol{y}) \coloneqq \kappa(S(\boldsymbol{x}), S(\boldsymbol{y}))$. We have

$$R^S_{f,X}(\boldsymbol{x}) = \sum_{i=1}^{N} c_i^S \kappa^S(\boldsymbol{x}, \boldsymbol{x}_i) = \sum_{i=1}^{N} c_i^S \kappa(S(\boldsymbol{x}), S(\boldsymbol{x}_i)) = R_{g,X}(S(\boldsymbol{x})), \quad \boldsymbol{x} \in \Omega,$$

where $g_{|S(X)} = f_{|X}$. In other words, the construction of the interpolant $R^S_{f,X} \in \mathcal{H}_{\kappa^S}$ is equivalent to the construction of a classical interpolant $R_{g,S(X)} \in \mathcal{H}_\kappa$ at the *fake* nodes $S(X)$. Similarly to the VSK setting, the FNA is provided with the following properties (cf. A1 and A2):

B1. The theoretical analysis of the FNA interpolant reduces to the analysis of the interpolant in the classical framework. This was proved in [15, Proposition 3.4] and it is linked to the inheritance property of the *Lebesgue constant* [7].
B2. The spaces $\mathcal{N}_\kappa$ and $\mathcal{N}_{\kappa^S}$ are isometrically isomorphic (a direct consequence of Theorem 1 and Proposition 1 stated below).

In previous works, the FNA has been mainly employed for two main purposes. The first is using S to obtain an interpolation design $S(X)$ that leads to a more stable interpolation process with respect to the original set of nodes. For example, this can be achieved in a polynomial-based framework by mapping onto the

set of (tensor-product) Chebyshev-Lobatto nodes in the (multi) one-dimensional case. In addition, as we will discuss in Subsect. 4.2, S can be constructed in a target-dependent fashion to emulate the possible discontinuities of the underlying function and thus recover accuracy near jump points.

4 Mapped VSKs

4.1 The General Framework

In Sects. 2 and 3, we outlined two different approaches, which however present some similarities and are employed for analogous purposes. In the following, we discuss how the VSK setting and the FNA can be merged in a unified framework. We start by giving the following definition.

Definition 1. *Let $S : \Omega \longrightarrow \mathbb{R}^d$ be an injective map, $\psi : \Omega \longrightarrow \mathbb{R}$ be a shape function and let $\kappa : (S(\Omega) \times \mathbb{R}) \times (S(\Omega) \times \mathbb{R}) \longrightarrow \mathbb{R}$ be a kernel. Then, a Mapped VSK (MVSK) $\kappa_\psi^S : \Omega \times \Omega \longrightarrow \mathbb{R}$ is defined as*

$$\kappa_\psi^S(\boldsymbol{x}, \boldsymbol{y}) \coloneqq \kappa((S(\boldsymbol{x}), \psi(\boldsymbol{x})), (S(\boldsymbol{y}), \psi(\boldsymbol{y}))),$$

for $\boldsymbol{x}, \boldsymbol{y} \in \Omega$.

We remark that κ_ψ^S might be defined in different possible equivalent manners. However, the advantage of Definition 1 lies in the separation between the actions of S and ψ: The function S works in the original dimension $\mathbb{R}^d$, while ψ rules the coordinate of the input in the augmented dimension. Under certain assumptions, a MVSK reduces to a mapped or VSK.

Proposition 1. *Let κ_ψ^S be a MVSK on $\Omega \times \Omega$ built upon a radial kernel κ. Then:*

1. *If $\psi(\boldsymbol{x}) \equiv \alpha \in \mathbb{R}$, then $\kappa_\psi^S = \kappa^S$.*
2. *If $S(\boldsymbol{x}) - S(\boldsymbol{y}) = \boldsymbol{x} - \boldsymbol{y}$, (e.g., S is the identity map), then $\kappa_\psi^S = \kappa_\psi$.*

Proof. By hypothesis, there exists $\varphi : [0, +\infty) \longrightarrow \mathbb{R}$ such that $\kappa(\boldsymbol{x}, \boldsymbol{y}) = \varphi(\|\boldsymbol{x} - \boldsymbol{y}\|_2)$. Therefore, we can write

$$\begin{aligned}\kappa_\psi^S(\boldsymbol{x}, \boldsymbol{y}) &= \varphi(\|(S(\boldsymbol{x}), \psi(\boldsymbol{x})) - (S(\boldsymbol{y}), \psi(\boldsymbol{y}))\|_2) \\ &= \varphi(\sqrt{(S(\boldsymbol{x}) - S(\boldsymbol{y}))^2 + (\psi(\boldsymbol{x}) - \psi(\boldsymbol{y}))^2}),\end{aligned}$$

from which the two theses follow.

We also prove the following.

Theorem 1. *The spaces $\mathcal{H}_\kappa$ and $\mathcal{H}_{\kappa_\psi^S}$ are isometric.*

Proof. By defining the map $\Lambda_\psi^S(\boldsymbol{x}) \coloneqq (S(\boldsymbol{x}), \psi(\boldsymbol{x}))$ we can see κ_ψ^S as the pull-back of κ in the sense provided in [29, Equation 2.52]. Then, the proof follows from [29, Theorem 2.9]. Indeed, since S is injective then so is Λ_ψ^S, which implies

$$\bigcap_{\boldsymbol{x}\in\Omega} \ker\left(\delta_{\Lambda_\psi^S(\boldsymbol{x})}\right) = \{\mathbf{0}\},$$

where we consider the evaluation functional $\delta_{\Lambda_\psi^S(\boldsymbol{x})}(f) = f(\Lambda_\psi^S(\boldsymbol{x}))$, $f \in \mathcal{H}_\kappa$.

As a consequence of Theorem 1, the native spaces $\mathcal{N}_\kappa$ and $\mathcal{N}_{\kappa_\psi^S}$ are isometrically isomorphic (cf. [6, Section 3]).

Independently of the target function to be recovered, we recall that the conditioning of the interpolation problem is related to the ℓ_2-conditioning of the kernel matrix $\mathrm{cond}(\mathsf{K}) = \lambda_{\max}/\lambda_{\min}$, being $\lambda_{\max}$, $\lambda_{\min}$ the maximum and minimum eigenvalue, respectively. It is known that $\lambda_{\min}$ decays according to the separation distance q_X, in a way that is influenced by the regularity of the kernel (see [17, Chapter 16]). In this direction, MVSKs can be employed in order to increase the separation distance and thus improve the conditioning of the interpolation scheme, e.g., by separating clustered nodes. On the other hand, diminishing the fill distance may improve the accuracy of the method (see (3)). We will experiment on this in Sect. 5.

4.2 Working with Mapped Discontinuous Kernels

In the following, we focus on the case of discontinuous functions. First, we briefly review in which manners VSKs and the FNA were used in the discontinuous setting.

Variably Scaled Discontinuous Kernels (VSDKs). The idea proposed in [13] and further investigated in [12] was to define the scaling function ψ to be discontinuous at the jumps of the target function $f : \Omega \longrightarrow \mathbb{R}$. In order to do so, we assume $\Omega \subset \mathbb{R}^d$ to be a bounded set such that:

- Ω is the union of m pairwise disjoint sets Ω_k, $k \in \{1, \ldots, m\}$.
- Each subset Ω_k, $k = 1, \ldots, m$, has a Lipschitz boundary.
- The discontinuity points of f are contained in the union of the boundaries of the subsets Ω_k, $k = 1, \ldots, m$.

Then, letting $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$, $\alpha_i \in \mathbb{R}$, the scaling function ψ is such that:

- ψ is piecewise constant, such that $\psi(\boldsymbol{x})\big|_{\Omega_k} = \alpha_k$.
- $\alpha_i \neq \alpha_j$ if Ω_i and Ω_j are neighboring sets.

The theoretical analysis carried out in the referring papers then focused on radial kernels whose related univariate function φ has the following Fourier decay

$$(\mathrm{F}\varphi)(\boldsymbol{\omega}) \sim (1 + \|\boldsymbol{\omega}\|^2)^{-s-\frac{1}{2}}, \quad s > \frac{d-1}{2}. \tag{4}$$

The native space of the kernels that satisfy (4), e.g. Matérn and Wendland kernels, is a Sobolev space [30, Chapter 10]. In order to present an error bound in terms of Sobolev spaces norm for the VSDK setting, we introduce two necessary ingredients:

1. The *regional* fill-distance

$$h_k := h_{X,\Omega_k} = \sup_{\boldsymbol{x}\in\Omega_k} \min_{\boldsymbol{x}_k\in X\cap\Omega_k} \|\boldsymbol{x}-\boldsymbol{x}_k\|_2$$

and the *global* fill distance

$$h := \max_{k\in\{1,\dots,m\}} h_k.$$

2. Letting $s \geq 0$ and $1 \leq p \leq \infty$, we define the space

$$\mathcal{WP}_p^s(\Omega) := \left\{ f : \Omega \longrightarrow \mathbb{R} \mid f|_{\Omega_k} \in \mathcal{W}_p^s(\Omega_k),\ k \in \{1,\dots,m\} \right\},$$

which contains the piecewise smooth functions f on Ω whose restriction to any subregion Ω_k, $k = 1,\dots,m$, is contained in the standard Sobolev space $\mathcal{W}_p^s(\Omega_k)$. The space $\mathcal{WP}_p^s(\Omega)$ is endowed with the norm

$$\|f\|^p_{\mathcal{WP}_p^s(\Omega)} = \sum_{k=1}^m \|f|_{\Omega_k}\|^p_{\mathcal{W}_p^s(\Omega_k)}.$$

In the outlined assumptions and letting $R^{\psi}_{f,X}$ be the kernel-based interpolant of f at X built upon the VSDK κ_ψ, in [12, Theorem 3.4] the authors proved the following. Let $s > 0$, $1 \leq q \leq \infty$ and $t \in \mathbb{N}_0$ such that $\lfloor s \rfloor > t + \frac{d}{2}$. Then, for $f \in \mathcal{WP}_2^s(\Omega)$ and *sufficiently small* h we have that

$$\|f - R^{\psi}_{f,X}\|_{\mathcal{WP}_q^t(\Omega)} \leq Ch^{s-t-d(1/2-1/q)_+}\|f\|_{\mathcal{WP}_2^s(\Omega)}, \tag{5}$$

where the constant $C > 0$ is independent of h.

The S-Gibbs Map in the FNA. In the mapped bases approach, similarly to the VSDK framework, the intuition is to map the nodes in order to create some *gaps* in presence of discontinuities. To do so, considering the collection of subsets $\Omega_1,\dots,\Omega_m$ employed to construct VSDKs, the so-called *S-Gibbs map* is designed as

$$S(\boldsymbol{x}) = \boldsymbol{x} + \sum_{k=1}^m \boldsymbol{\beta}_k \chi_{\Omega_k}(\boldsymbol{x}),$$

where $\boldsymbol{\beta}_k = (k\beta,\dots,k\beta) \in \mathbb{R}^d$, $\beta \in \mathbb{R}$, and χ_{Ω_k} is the characteristic function corresponding to Ω_k. In [15, Section 4.2], the authors discussed the analogies between a VSDK κ_ψ and a mapped kernels κ^S constructed via the S-Gibbs map. It turned out that these two approaches can lead to *similar* results for certain values of the vectors of parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_k$. This is due to the fact that the interpolation matrices are *close* being the kernel radial.

Mapped VSDKs (MVSDKs). In the mixed approach with the mapped VSK kernel κ_ψ^S, we deal with the jumps of the underlying function as follows.

- We define ψ as in the VSDKs framework. Therefore, the role of the shape function is to mimic the jumps of the target function and thus to prevent the appearance of the Gibbs phenomenon.
- Since the discontinuities are already addressed by ψ, we employ S to map the set of nodes X to obtain improved distributions locally on each Ω_k, meaning that we aim at diminishing the global fill distance and increasing the global separation distance defined as

$$q \coloneqq \min_{k\in\{1,\dots,m\}} q_k,$$

being $q_k \coloneqq q_{X_k} = \frac{1}{2}\min_{i\neq j}\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2,\ \boldsymbol{x}_i, \boldsymbol{x}_j \in \Omega_k$.

Recalling the error estimate in (5), this proposed construction for MVSDKs may lead to better results, being h smaller. Moreover, by increasing the separation distance, an improvement in the conditioning of the scheme with respect to classical VSDKs is likely to be obtained. We test these aspects in some numerical examples in the next section.

5 Numerical Tests with MVSDKs

The purpose of this section is to provide a numerical example to show the benefits of the MVSDK framework in comparison to VSDKs and classical RBF interpolation. To do so, we set $\Omega = [-1, 1]^2$ and letting $\boldsymbol{x} = (x_1, x_2)$ we consider the target function

$$f : \Omega \longrightarrow \mathbb{R}, \quad f(\boldsymbol{x}) = \begin{cases} x_1 + x_2, & x_1 < -0.3, \\ \sin(x_1 - 2x_2), & 0 \leq x_1 < 0.5, \\ 0, & \text{otherwise.} \end{cases}$$

To deal with the jumps of f, we define the shape function

$$\psi : \Omega \longrightarrow \mathbb{R}, \quad \psi(\boldsymbol{x}) = \begin{cases} 0, & x_1 < -0.3, \\ 1, & -0.3 \leq x_1 < 0, \\ 2, & 0 \leq x_1 < 0.5, \\ 3, & x_1 \geq 0.5. \end{cases}$$

As far as the nodes are concerned, we let G_N be a set of N nodes that are sampled from a bivariate normal distribution with mean $\boldsymbol{\mu} = (0, 0)$ and covariance matrix $\boldsymbol{\Sigma} = 0.1 \cdot \mathsf{I}$, being I the 2×2 identity matrix. We can then consider the following mapping function $S : \Omega \longrightarrow \Omega$ defined as

$$S(x_1, x_2) = \left(1 + \operatorname{erf}\left(\frac{x_1}{\sqrt{0.2}}\right), 1 + \operatorname{erf}\left(\frac{x_2}{\sqrt{0.2}}\right)\right) - 1,$$

where erf is the well-known *error function*. To clarify the idea behind the construction of S, it is known from classical probability theory that if $z_1, \ldots, z_n$ are sampled according to a normal distribution of mean μ and standard deviation σ, then $0.5(1 + \text{erf}((z_i - \mu)/(\sqrt{2}\sigma)))$, $i = 1, \ldots, n$, are distributed uniformly in $[0, 1]$. Therefore, our set of nodes G_N is mapped to a uniform distribution in the square $[-1, 1]^2$, as displayed in Fig. 1. We point out that possible nodes in G_N that are not in Ω are removed from the set.

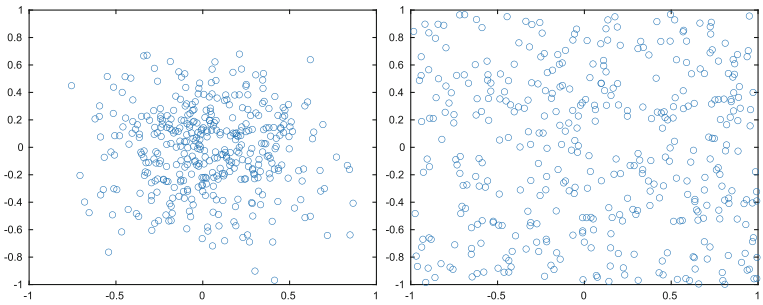


Fig. 1. $N = 400$. Left: G_N. Right: $S(G_N)$.

Consequently, the mapped set $S(G_N)$ is very likely to present smaller fill distance and larger separation distance than G_N. The interpolation results are evaluated on a finer $M \times M$ equispaced grid Ξ_{M^2} in Ω. Precisely, we compute the Root Mean Square Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{M^2} \sum_{k=1}^{M^2} \big(f(\boldsymbol{\xi}_k) - \iota(\boldsymbol{\xi}_k)\big)^2}, \tag{6}$$

where $\boldsymbol{\xi}_k \in \Xi_{M^2}$ and $\iota = R_{f,G_N}$, R^{ψ}_{f,G_N}, $R^{\psi,S}_{f,G_N}$ are the interpolants constructed via the classical, the VSDK and the MVSDK. The shape parameter ε is chosen between 200 equispaced values in the interval $[0.01, 50]$ via Leave-One-Out Cross Validation (LOOCV) [23]. Finally, we let N vary between 10 and 500 and we test two radial kernels:

$$\varphi_W(r) = (1 - \varepsilon r)_+^2 \quad \text{Wendland } C^0,$$
$$\varphi_M(r) = e^{-\varepsilon r}(15 + 15\varepsilon r + 6(\varepsilon r)^2 + (\varepsilon r)^3) \quad \text{Matérn } C^6 \ .$$

In Fig. 2, we show the behavior of the separation and fill distances, while in Fig. 3 we show the RMSEs achieved with both φ_W and φ_M.

The plots in Fig. 2 display the benefits in employing S in the MVSDK in terms of diminished fill distance and increased separation distance. In Fig. 3, we observe that MVSDKs are more effective in the case of the chosen Matérn kernel. This is due to the fact that this kernel is more regular than the chosen Wendland

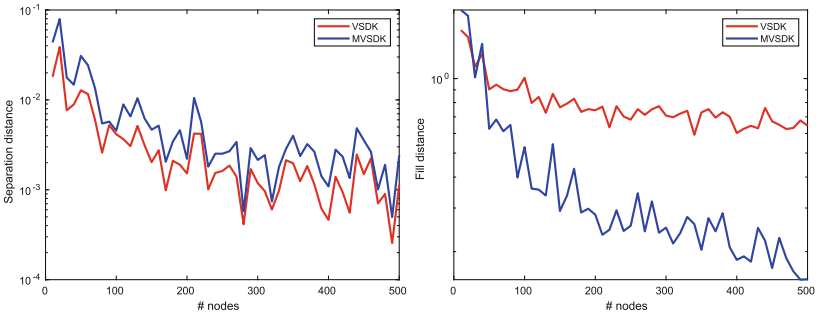


Fig. 2. The separation distance (left) and the fill distance (right) varying N.

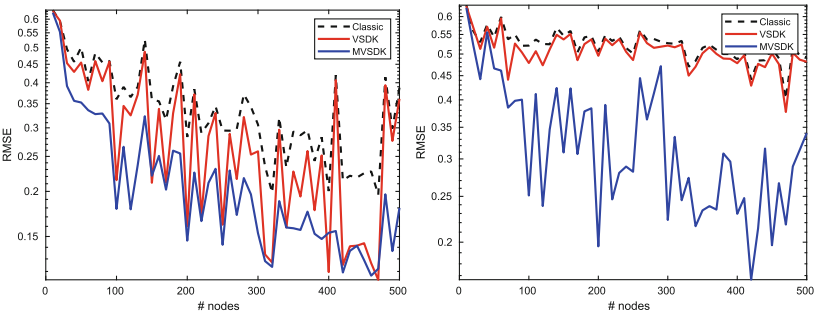


Fig. 3. The RMSE obtained using φ_W (left) and φ_M (right) varying N.

kernel, therefore it is more prone to provide an ill-conditioned interpolation process. Furthermore, we remark that the sets of nodes G_N, $N = 10, \ldots, 500$ are not nested, and they are clustered around the origin. Therefore, we can not expect an accurate recovering of the function f, nor convergence increasing N.

6 Applications to the STIX Imaging Framework

In order to test the proposed MVSKs in real applied sciences, we focus on solar hard X-ray imaging and, specifically, on the ESA STIX telescope [28], on board of Solar Orbiter mission (see Fig. 4). Hard X-ray telescopes provide experimental measurements, named visibilities, of the Fourier transform of the incoming photon flux at specific points of the spatial frequency plane. In the case of STIX, 30 subcollimators relying on the Moiré pattern technology provide $N = 60$ visibilities on 10 circles of the frequency plane with increasing radii from about 2.79×10^{-3} arcsec^{-1} to 7.02×10^{-2} arcsec^{-1} (see Fig. 4). We observe that the visibilities lying in the lower half plane are obtained by reflecting the visibilities in the upper half with respect to the origin.

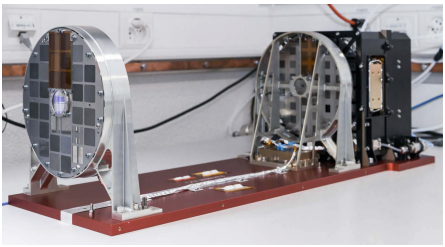


Fig. 4. The Spectrometer/Telescope for Imaging X-rays (STIX).

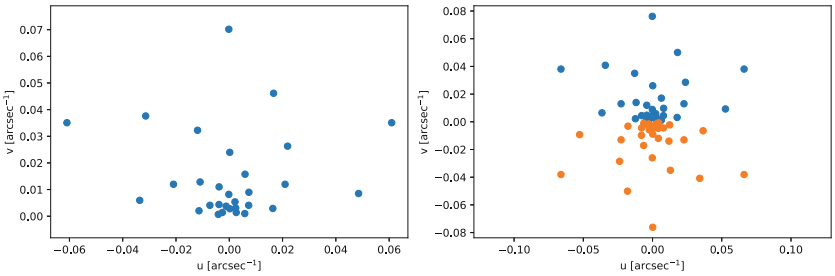


Fig. 5. STIX visibilities: reflecting with respect to the origin leads to a total number of 60 visibilities.

We denote by $\mathbf{f}$ the vector whose components are the discretized values of the incoming flux, by $\mathbf{F}$ the discretized Fourier transform sampled at the set of points $\{\mathbf{u}_i = (u_i, v_i)\}_{i=1}^N$ in the (u, v)-plane and with $\mathbf{V}$ the vector whose n components are the observed visibilities. Then, the image formation model in this framework can be approximated by

$$\mathbf{V} = \mathbf{F}\mathbf{f} \ . \tag{7}$$

6.1 The Imaging Process

Many inversion methods have been formulated to express the STIX observations as images, see e.g. [2,3,5,9,18,24,25]. The approach that we propose consists of two steps: interpolation of the visibilities so that we obtain the visibility surfaces and the inversion of the so generated surfaces with rather standard techniques. This idea was already used for the dismissed NASA telescope Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) [22], and its IDL implementation called uv_smooth [25], can be find in the NASA Solar SoftWare (SSW) tree.

The uv_smooth code addresses Eq. (7) by means of an interpolation and extrapolation procedure in which the interpolation step is carried out via an algorithm based on spline functions and the extrapolation step is realized by means of a soft-thresholding scheme [1,10]. As far as the second step is concerned, we will rely on the well-established projected Landweber iterative method [27]. The first step instead will be carried out with our MVSKs.

6.2 MVSKs for STIX

To employ MVSKs, we need to define a scaling function ψ and a mapping S.

The choice of ψ. In order to define the scaling function, we take advantage of a first approximation of the inverse problem obtained via a standard back-projection algorithm [26] that computes the discretized inverse Fourier transform of the visibilities by means of the IDL source code `vis_bpmap` available in the NASA SSW tree. The so-constructed image is then forward Fourier transformed to obtain ψ. Once the interpolated visibility surface $\overline{\mathbf{V}}$ has been computed, the image reconstruction problem reads as follows

$$\overline{\mathbf{V}} = \overline{\mathbf{F}}\overline{\mathbf{f}}\ , \tag{8}$$

where $\overline{\mathbf{F}}$ is the $N^2 \times N^2$ discretized Fourier transform and $\overline{\mathbf{f}}$ is the $N^2 \times 1$ vector to reconstruct. In the following we will point out the advantages of interpolating the visibilities with our technique.

The Choice of S. To present in details the mapping S chosen for the STIX imaging framework, let us first deepen the definition of the visibilities represented in Fig. 5 (left). We have

$$u_i = (L_1 + L_2)\frac{\cos(\alpha_i^f)}{\rho_i^f} - L_2\frac{\cos(\alpha_i^r)}{\rho_i^r},\ v_i = (L_1 + L_2)\frac{\sin(\alpha_i^f)}{\rho_i^f} - L_2\frac{\sin(\alpha_i^r)}{\rho_i^r},$$

where $L_1 = 550$, $L_2 = 47$ and $\alpha_i^f, \rho_i^f, \alpha_i^r, \rho_i^r$ are discussed in [19]. We define the map

$$S(u, v) = C\log(\|(u, v)\|_2)(\cos(\arctan(v/u)), \sin(\arctan(v/u))),$$

where C is a normalizing factor used to retain the order of magnitude of the original visibilities after the mapping. In Fig. 6, we can observe that the mapped visibilities are distributed in a circular crown with no clustering around the origin.

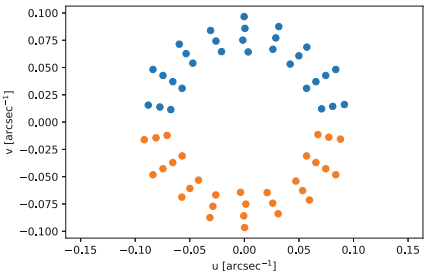


Fig. 6. Mapped STIX visibilities.

6.3 Numerical Results

On Jul 2022 STIX recorded a flare during the during the time interval 23:41:15–23:41:41 UT. The energy range of the event is 15–25 keV. In Fig. 7 we reported the reconstruction carried out with the interpolation/extrapolation algorithm where we respectively interpolate with: classical radial kernels, VSKs and MVSKs. Moreover, as a further comparison we also consider mem_ge [24] which is a well established implementation of the maximum entropy approach and it is used by the solar physics community. We note that all methods present artifacts but the shape of the source reconstructed by interpolating with MVSKs is more similar to the one computed with mem_ge. To have a quantitative feedback on the accuracy, we show in Fig. 8 the visibility fits obtained with the four different approaches. We further observe that the chi squares values of mem_ge and uv_smooth + MVSKs are similar and are the lowest.

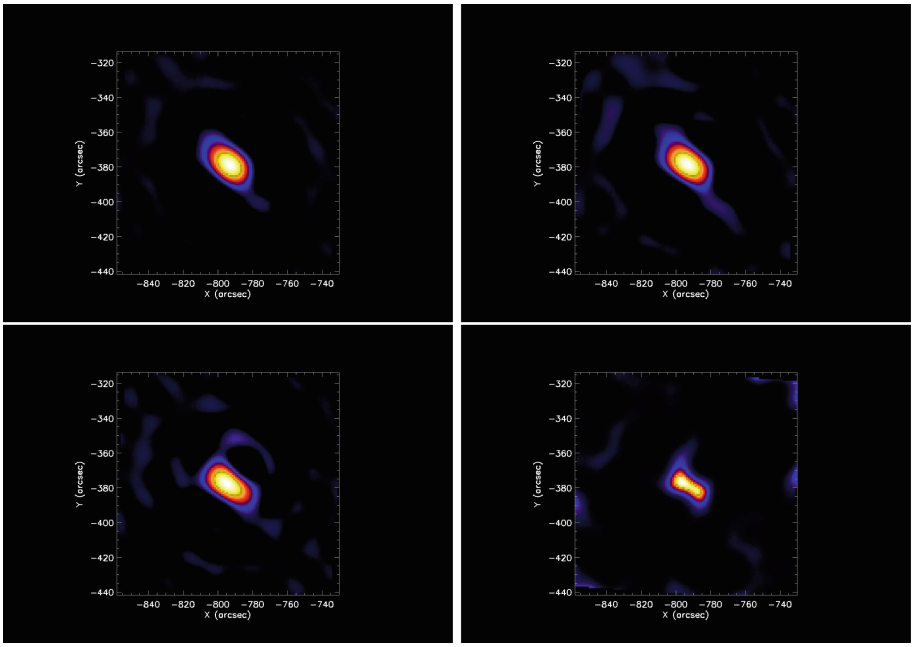


Fig. 7. Left to right, top to bottom: reconstruction of the flaring source with: uv_smooth + classical radial kernels, uv_smooth + VSKs, uv_smooth + MVSKs and mem_ge.

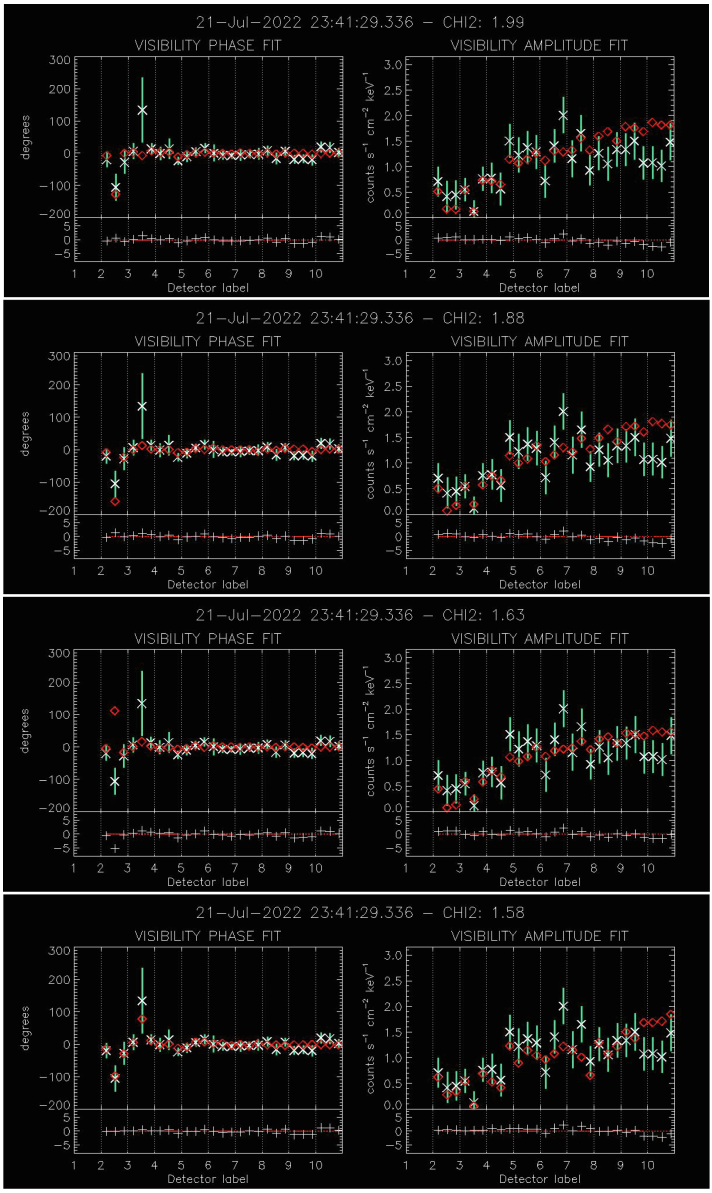


Fig. 8. Top to bottom: visibility fits of the flaring source with: uv_smooth + classical radial kernels, uv_smooth + VSKs, uv_smooth + MVSKs and mem_ge.

7 Conclusions

In this work, we defined MVSKs by mixing VSKs with the FNA. After providing some theoretical details, we performed some numerical simulations that showed

the advantages of using MVSDKs with respect to classical RBF kernels and VSDKs. Then, we applied our method in the framework of hard X-ray astronomical imaging. The obtained results encourage further investigations on this research line.

Acknowledgments. EP and AM are supported by the project: "Physics-based AI for predicting extreme weather and spaceweather events (AIxtreme)", Founded by Fondazione Compagnia di San Paolo. FM acknowledges the financial support of the Programma Operativo Nazionale (PON) "Ricerca e Innovazione" 2014–2020. The authors kindly acknowledge the financial contribution from the AI-FLARES ASI-INAF n.2018-16-HH.0 agreement. This research has been accomplished within GNCS-INδAM.

References

1. Allavena, S., Piana, M., Benvenuto, F., Massone, A.M.: An interpolation/extrapolation approach to X-ray imaging of solar flares. Inverse Probl. Imag. **6**, 147 (2012)
2. Aschwanden, M.J., Schmal, E.: The RHESSI team: reconstruction of RHESSI solar flare images with a forward-fitting method. Sol. Phys. **210**, 193–211 (2002)
3. Benvenuto, F., Schwartz, R., Piana, M., Massone, A.M.: Expectation maximization for hard X-ray count modulation profiles. Astron. Astrophys. **555**, A61 (2013)
4. Berrut, J.P., De Marchi, S., Elefante, G., Marchetti, F.: Treating the Gibbs phenomenon in barycentric rational interpolation and approximation via the S-Gibbs algorithm. Appl. Math. Lett. **103**, 106196, 7 (2020)
5. Bonettini, S., Anastasia, C., Prato, M.: A new semiblind deconvolution approach for Fourier-based image restoration: an application in astronomy. SIAM J. Imaging Sci. **6**, 1736–1757 (2013)
6. Bozzini, M., Lenarduzzi, L., Rossini, M., Schaback, R.: Interpolation with variably scaled kernels. IMA J. Numer. Anal. **35**(1), 199–219 (2015)
7. Brutman, L.: Lebesgue functions for polynomial interpolation - a survey. Ann. Numer. Math. **4**(1/4), 111–128 (1996)
8. Campi, C., Marchetti, F., Perracchione, E.: Learning via variably scaled kernels. Adv. Comput. Math. **47**(4), 23 (2021). Paper No. 51. https://doi.org/10.1007/s10444-021-09875-6
9. Cornwell, T., Evans, K.F.: A simple maximum entropy deconvolution algorithm. Astron. Astrophys. **143**(1), 77–83 (1985)
10. Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math. **57**(11), 1413–1457 (2004)
11. De Marchi, S., Elefante, G., Marchetti, F.: Stable discontinuous mapped bases: the Gibbs-Runge-avoiding stable polynomial approximation (GRASPA) method. Comput. Appl. Math. **40**(8), 17 (2021). Paper No. 299
12. De Marchi, S., Erb, W., Marchetti, F., Perracchione, E., Rossini, M.: Shape-driven interpolation with discontinuous kernels: error analysis, edge extraction, and applications in magnetic particle imaging. SIAM J. Sci. Comput. **42**(2), B472–B491 (2020)
13. De Marchi, S., Marchetti, F., Perracchione, E.: Jumping with variably scaled discontinuous kernels (VSDKs). BIT Numer. Math. **60**, 441–463 (2020)

14. De Marchi, S., Marchetti, F., Perracchione, E., Poggiali, D.: Polynomial interpolation via mapped bases without resampling. J. Comput. Appl. Math. **364**, 112347, 12 (2020)
15. De Marchi, S., Marchetti, F., Perracchione, E., Poggiali, D.: Multivariate approximation at fake nodes. Appl. Math. Comput. **391**, 17 (2021). Paper No. 125628
16. De Marchi, S., Lot, F., Marchetti, F., Poggiali, D.: Variably scaled persistence kernels (VSPKs) for persistent homology applications. J. Comput. Math. Data Sci. **4**, 100050 (2022)
17. Fasshauer, G.E.: Meshfree Approximation Methods with MATLAB. World Scientific (2007)
18. Felix, S., Bolzern, R., Battaglia, M.: A compressed sensing-based image reconstruction algorithm for solar flare x-ray observations. Astrophys. J. **849**(1), 10 (2017)
19. Giordano, S., Pinamonti, N., Piana, M., Massone, A.M.: The process of data formation for the Spectrometer/Telescope for Imaging X-rays (STIX) in Solar Orbiter. SIAM J. Imaging Sci. **8**(2), 1315–1331 (2015)
20. Guastavino, S., Benvenuto, F.: Convergence rates of spectral regularization methods: a comparison between ILL-posed inverse problems and statistical kernel learning. SIAM J. Numer. Anal. **58**(6), 3504–3529 (2020)
21. Karimnejad Esfahani, M., De Marchı, S., Marchetti, F.: Moving least squares approximation using variably scaled discontinuous weight function. Constr. Math. Anal. **6**(1), 38–54 (2023)
22. Lin, R.P., et al.: The Reuven Ramaty High-Energy Solar Spectroscopic Imager (RHESSI). Sol. Phy. **210**(1–2), 3–32 (2002)
23. Ling, L., Marchetti, F.: A stochastic extended Rippa's algorithm for LpOCV. Appl. Math. Lett. **129**, 107955 (2022)
24. Massa, P., et al.: MEM_GE: a new maximum entropy method for image reconstruction from solar X-ray visibilities. Astrophys. J. **894**(1), 46 (2020)
25. Massone, A.M., Emslie, A.G., Hurford, G.J., Prato, M., Kontar, E.P., Piana, M.: Hard X-ray imaging of solar flares using interpolated visibilities. Astrophys. J. **703**, 2004–2016 (2009)
26. Mersereau, R., Oppenheim, A.: Digital reconstruction of multidimensional signals from their projections. Proc. IEEE **62**(10) (1974)
27. Piana, M., Bertero, M.: Projected Landweber method and preconditioning. Inverse Probl. **13**(2), 441–463 (1997)
28. Krucker, S., et al.: The Spectrometer/Telescope for Imaging X-rays (STIX). Astron. Astrophys. **642**, A15 (2020)
29. Saitoh, S., Sawano, Y.: Theory of Reproducing Kernels and Applications, Developments in Mathematics, vol. 44. Springer, Singapore (2016). https://doi.org/10.1007/978-981-10-0530-5
30. Wendland, H.: Scattered Data Approximation, Cambridge Monographs on Applied and Computational Mathematics, vol. 17. Cambridge University Press, Cambridge (2005)