A Petrov-Galerkin RBF method for diffusion equation on the unit sphere*

Mohammadreza Ahmadi Darani†

Department of Applied Mathematics, Faculty of Mathematical Sciences
Shahrekord University, P.O. Box. 115, Shahrekord, Iran.

Davoud Mirzaei[a,b‡]

[a] *Department of Applied Mathematics and Computer Science, Faculty of Mathematics and Statistics, University of Isfahan, 81746-73441 Isfahan, Iran.*

[b] *School of Mathematics, Institute for Research in Fundamental Sciences (IPM), Tehran 19395-5746, Iran.*

July 2020

Abstract

This paper concerns a numerical solution for the diffusion equation on the unit sphere. The given method is based on the spherical basis function (SBF) approximation and the Petrov-Galerkin test discretization. The method is meshless because spherical triangulation is not required neither for approximation nor for numerical integration. This feature is achieved through the SBF approximation and the use of local weak forms instead of a global variational formulation. The local Petrov-Galerkin formulation allows to compute the integrals on small independent spherical caps without any dependence on a connected background mesh. Experimental results show the accuracy and the efficiency of the new method.

Keywords: Spherical basis functions, Radial basis functions, Meshless methods, Petrov-Galerkin method, Diffusion on the sphere.

1 Introduction

A popular and interesting tool in computational sciences, which has received a lot of attention from researchers in recent years, is the meshless method that is based on discretizing a continuum by a set of nodal points without any mesh constraints. Some

*Authorship: Both authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript. Furthermore, each author certifies that this material or similar material has not been and will not be submitted to or published in any other publication.

Conflict of Interest: No conflict of interest exists. We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

†E-mail address: ahmadi.darani@sku.ac.ir

‡E-mail address: d.mirzaei@sci.ui.ac.ir, Corresponding author

examples of meshless approximation methods are the *moving least squares* (MLS), developed by Lancaster and Salkauskas [18] after the initial work of Shepard [40], and the *kernels* –in particular the *radial basis functions* (RBF)– introduced (in some special cases) by Hardy [11] and Harder and Desmarais [10]. Both of these approximation techniques are extensively used in various fields of computational sciences and engineering. In particular, some meshless methods based on the MLS and the RBF approximations have been developed for solving partial differential equations (PDEs) after the leading papers by Kansa [16, 17] on the RBF collocation and Belytschko et. al. [4] on the MLS-based Galerkin methods. As some, we can mention the symmetric collocation method [32, 46], the Galerkin RBF method [43], the meshless local Petrov-Galerkin (MLPG) method [3], the direct meshless local Petrov-Galerkin (DMLPG) method [26], and etc. In a point of view, these methods are alternatives to some mesh-based methods such as finite elements and finite volume methods because they do not rely on a predefined and connected triangulation. Instead, meshless methods write the unknown solution in terms of scattered nodes. Being independent from the dimension, working on more complicated geometries and easily adapting with smoothness are some important advantages of the RBF based meshless methods.

When the underlying domain (for pure approximation or PDE solution) is a non-planar submanifold (a general surface), the kernel based methods become much superior because mesh generation and mesh refinement on such domains are absolutely non trivial. While, working with scattered points is obviously simpler.

The unit sphere is an example of a surface domain for problems from various mathematical modelings such as a physical modeling in a geodesic region. Usually, such models lead to a PDE problem on the whole surface of the sphere. Approximation theory on the unit sphere has a rather long history in the context of meshless methods. Positive definite kernels on the sphere were already characterized by Schoenberg [38] many years ago, but the issue to determine what kind of positive definite kernels are actually strictly positive definite was initiated by Xu and Cheney [47] and continued by [5, 8, 34, 39]. These kernels are called *spherical basis functions (SBF)*. Also, error estimates of scattered data interpolation on spheres are given in [15]. Approximation by *zonal* basis functions and error estimates were studied in [13, 14, 33]. Sampling inequalities were proved in [9, 21]. SBFs were also applied successfully for PDE problems on the sphere [7, 19, 20, 22, 24, 29, 30]. In particular, in [24] a SBF based method in combination with a Petrov-Galerkin test discretization was proposed for solving spherical PDEs. The idea of local Petrov-Galerkin method dates back to [2, 3] for solving problems in solid mechanics. The combination with RBFs in Euclidean spaces can be found in [35, 37, 41]. In this paper we extend the method of [24] for solving diffusion problems on the sphere.

SBF approximation is not the only meshless tool for spherical problems. Wendland in [44] developed the MLS method for approximating the real-valued functions defined on spheres. For solving a spherical PDE by MLS, the PDE operator should act on complicated spherical MLS shape functions. Recently, in [23] a generalized MLS approximation was introduced to approximate differential operators on the sphere and hence to solve spherical PDEs. For a recent application see [6].

This paper is organized as follows. Approximation by kernels on the sphere and some related definitions are stated in section 2. In section 3, some results on differential operators on the sphere are reviewed. The extension of the Petrov-Galerkin method for

diffusion problems is given in section 4, and the stability of the time integration method is discussed in section 5. Finally, in section 6, some test examples are given to show the efficiency of the proposed method.

2 Kernels on Spheres

Consider the Euclidean space $\mathbb{R}^{d+1}$ and let $\mathbb{S}^d$ for $d \geq 1$ be the unit sphere in $\mathbb{R}^{d+1}$ i.e.,

$$\mathbb{S}^d = \{\boldsymbol{x} \in \mathbb{R}^{d+1} : \|\boldsymbol{x}\|_2^2 := \boldsymbol{x}^T\boldsymbol{x} = 1\}.$$

The geodesic distance between two points $\boldsymbol{x}, \boldsymbol{\xi} \in \mathbb{S}^d$ which is denoted by $\mathrm{dist}(\boldsymbol{x}, \boldsymbol{\xi})$, is the shortest curve on $\mathbb{S}^d$ connecting these points and is defined as

$$\mathrm{dist}(\boldsymbol{x}, \boldsymbol{\xi}) = \cos^{-1}(\boldsymbol{x}^T\boldsymbol{\xi}).$$

The spherical cap with radius $\delta \in (0, \pi)$ and center $\boldsymbol{x}$ on $\mathbb{S}^d$ is defined as

$$\mathcal{C}_\delta(\boldsymbol{x}) = \{\boldsymbol{\xi} \in \mathbb{S}^d : \mathrm{dist}(\boldsymbol{x}, \boldsymbol{\xi}) \leq \delta\}. \tag{2.1}$$

Spherical harmonics play the role of polynomials on $\mathbb{S}^d$ which are very important for analysis and approximation theory on $\mathbb{S}^d$. The space of spherical harmonics of degree ℓ is the restriction of the space of all harmonic and homogeneous polynomials of degree ℓ on $\mathbb{R}^{d+1}$ to $\mathbb{S}^d$ and denoted by $\mathbb{H}_\ell(\mathbb{S}^d)$. The dimension of $\mathbb{H}_\ell(\mathbb{S}^d)$ can be obtained as

$$N(d, \ell) = \begin{cases} 1 & \ell = 0, \\ \frac{(2\ell+d-1)\Gamma(\ell+d-1)}{\Gamma(\ell+1)\Gamma(d)} & \ell \geq 0, \end{cases}$$

where Γ is the Gamma function. The space of spherical harmonics of order at most n then is defined as

$$\mathbb{P}_n(\mathbb{S}^d) := \bigoplus_{\ell=0}^{n} \mathbb{H}_\ell(\mathbb{S}^d),$$

which is of dimension $N(d+1, n)$. The corresponding orthonormal basis for $\mathbb{P}_n(\mathbb{S}^d)$ is

$$\{Y_{\ell k}, \ell = 1, 2, \ldots, n,\ k = 1, 2, \ldots, N(d, \ell)\},$$

with the following orthonormal property

$$\int_{\mathbb{S}^d} Y_{ij}(\boldsymbol{x}) Y_{\ell k}(\boldsymbol{x})\, d\boldsymbol{x} = \delta_{i\ell}\delta_{jk}.$$

In additions, the spherical harmonics $Y_{\ell k}$ are eigenfunctions of the Laplace-Beltrami operator Δ_* (see the next section) in the sense that

$$\Delta_* Y_{\ell k} - \lambda_\ell Y_{\ell k} = 0, \quad \lambda_\ell = -\ell(\ell + d - 1),\ k = 1, 2, \ldots, N(d, \ell),\ \ell = 0, 1, 2, \ldots, \tag{2.2}$$

where λ_ℓ is the eigenvalue of Laplace-Beltrami operator corresponding to eigenfunction $Y_{\ell k}$. So, a function $f \in L_2(\mathbb{S}^d)$ can be expanded in terms of spherical harmonics as

$$f(\boldsymbol{x}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widehat{f}_{\ell k} Y_{\ell k}(\boldsymbol{x}), \quad \widehat{f}_{\ell k} = \langle f, Y_{\ell k} \rangle_{L_2(\mathbb{S}^d)}.$$

Kernels are a powerful tool for approximation on spheres. By a kernel we mean a function $\Phi : \mathbb{S}^d \times \mathbb{S}^d \to \mathbb{R}$ which is at least continuous in both its arguments. Assume that a set of scattered points $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\} \subset \mathbb{S}^d$ and values $u_1, u_2, \ldots, u_N \in \mathbb{R}$ are given. The kernel interpolant s_X of this data is coming from the data-dependent approximation space

$$V_{X,\Phi} := \operatorname{span}\{\Phi(\cdot, \boldsymbol{x}_1), \Phi(\cdot, \boldsymbol{x}_2), \ldots, \Phi(\cdot, \boldsymbol{x}_N)\}.$$

If s_X is expanded as

$$s_X(\boldsymbol{x}) = \sum_{j=1}^{N} c_j \Phi(\boldsymbol{x}, \boldsymbol{x}_j)$$

and the interpolation constrains $s_X(\boldsymbol{x}_k) = u(\boldsymbol{x}_k)$, $k = 1, \ldots, N$ are imposed, then the linear system

$$A\boldsymbol{c} = \boldsymbol{u}, \tag{2.3}$$

is resulted where $A_{kj} = \Phi(\boldsymbol{x}_k, \boldsymbol{x}_j)$ for $k, j = 1, \ldots, N$, and $\boldsymbol{u} = (u(\boldsymbol{x}_1), \ldots, u(\boldsymbol{x}_N))^T$. For any choice of set of points X on $\mathbb{S}^d$ the interpolation matrix A is positive definite provided that Φ is a positive definite kernel on $\mathbb{S}^d$.

We are interested in spherical kernels of the form

$$\Phi(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\ell=0}^{\infty} \sum_{k=1}^{N(d,\ell)} \widehat{\phi}_\ell Y_{\ell k}(\boldsymbol{x}) Y_{\ell k}(\boldsymbol{y}) \tag{2.4}$$

for positive Fourier coefficients $\widehat{\phi}_\ell$. Although the intrinsic positive definite kernels on $\mathbb{S}^d$ are well studied and have a rather long history in mathematics [38,45,47], in this paper we use the restriction of positive definite kernels from $\mathbb{R}^{d+1}$ to $\mathbb{S}^d$. The restriction of a positive definite kernel from $\mathbb{R}^{d+1}$ to any embedded submanifold $\mathbb{M} \subset \mathbb{R}^{d+1}$ is a simple way for obtaining a positive definite kernel on $\mathbb{M}$. It is clear that if the original kernel is positive definite, so its restriction to the submanifold is. There exist some other issues, such as corresponding kernel spaces and approximation properties of the restricted kernels on the submanifold that come up with this approach. The case $\mathbb{M} = \mathbb{S}^d$ has been studied in [31], while the general case has been investigated in [9]. Before all, a variation of compactly supported RBFs of Wendland's type on the sphere was introduced in [33].

If the original kernel on $\mathbb{R}^{d+1}$ is a radial function, i.e. it is a univariate function of Euclidian distance $\|\boldsymbol{x} - \boldsymbol{\xi}\|_2$ then the restricted kernel is a univariate function of $\boldsymbol{x}^T\boldsymbol{\xi}$. The reason is clear, because for $\boldsymbol{x}, \boldsymbol{\xi} \in \mathbb{S}^d$ we have $\|\boldsymbol{x} - \boldsymbol{\xi}\|_2 = \sqrt{2 - 2\boldsymbol{x}^T\boldsymbol{\xi}}$. A kernel $\Phi(\boldsymbol{x}, \boldsymbol{\xi})$ on $\mathbb{S}^d$ for which there exists a univariate and continuous function $\psi : [-1, 1] \to \mathbb{R}$ such that $\Phi(\boldsymbol{x}, \boldsymbol{\xi}) = \psi(\boldsymbol{x}^T\boldsymbol{\xi})$ is called a *zonal* kernel. The value of a zonal kernel depends only to the angle between to vectors $\boldsymbol{x}$ and $\boldsymbol{\xi}$.

3 Differential operators on spheres

In this section we briefly review the surface gradient and Laplacian operators which play an important role in diffusion and advection models on the sphere.

The Laplace-Beltrami operator Δ_* is the spherical part of the Euclidean Laplace operator Δ. In fact, in the spherical-polar coordinated we have

$$\Delta = \frac{\partial}{\partial r^2} + \frac{d}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Delta_*,$$

where Δ_* is independent of derivatives with respect to r. For example on $\mathbb{S}^2$ in spherical coordinates (θ, ϕ, r)

$$\Delta_* = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial}{\partial\theta}\Big) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}.$$

Laplace-Beltrami can also be expressed in extrinsic (or Cartesian) coordinates. Here we again restrict ourselves to case $d = 2$, because the extension to other dimensions and even other smooth embedded sub-manifolds will be obvious. For $\boldsymbol{x} = (x, y, z) \in \mathbb{S}^2$ denotes the unit normal vector to $\mathbb{S}^2$ at $\boldsymbol{x}$ by $\boldsymbol{n}(\boldsymbol{x}) = \boldsymbol{n} = (n^x, n^y, n^z)^T$. Obviously on $\mathbb{S}^2$ we have $\boldsymbol{n}(\boldsymbol{x}) = \boldsymbol{x}$. We define $\Pi := I_{3\times3} - \boldsymbol{n}\boldsymbol{n}^T$ which projects vectors in $\mathbb{R}^3$ to space of tangent vectors to $\mathbb{S}^2$ at $\boldsymbol{x}$. Then, the surface gradient on $\mathbb{S}^2$ at $\boldsymbol{x}$ is defined as

$$\nabla_* := \Pi\nabla = (I_{3\times3} - \boldsymbol{n}\boldsymbol{n}^T)\nabla = (I_{3\times3} - \boldsymbol{x}\boldsymbol{x}^T)\nabla,$$

where ∇ is the usual gradient in $\mathbb{R}^3$ and $I_{3\times3}$ is the identity matrix of size 3. In an extensive form ∇_* at $\boldsymbol{x} \in \mathbb{S}^2$ is written as

$$\nabla_* = \begin{bmatrix} (1-x^2)\dfrac{\partial}{\partial x} - xy\dfrac{\partial}{\partial y} - xz\dfrac{\partial}{\partial z} \\ -xy\dfrac{\partial}{\partial x} + (1-y^2)\dfrac{\partial}{\partial y} - yz\dfrac{\partial}{\partial z} \\ -xz\dfrac{\partial}{\partial x} - yz\dfrac{\partial}{\partial y} + (1-z^2)\dfrac{\partial}{\partial z} \end{bmatrix} =: \begin{bmatrix} \mathcal{G}^x \\ \mathcal{G}^y \\ \mathcal{G}^z \end{bmatrix}, \tag{3.5}$$

and the surface divergence of a vector field $\boldsymbol{v} = (v^x, v^y, v^z) : \mathbb{S}^2 \to \mathbb{R}^3$ at $\boldsymbol{x} \in \mathbb{S}^2$ is expressed as

$$\nabla_* \cdot \boldsymbol{v} = \mathcal{G}^x v^x + \mathcal{G}^y v^y + \mathcal{G}^z v^z.$$

Now the Laplace-Beltrami operator Δ_* can be define as

$$\Delta_* := \nabla_* \cdot \nabla_* = \mathcal{G}^x\mathcal{G}^x + \mathcal{G}^y\mathcal{G}^y + \mathcal{G}^z\mathcal{G}^z.$$

If the involved functions are zonal then the expressions become rather simpler. Assume $\Phi : \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}$ is a zonal function that means there exists a continuous function $\psi : [-1, 1] \to \mathbb{R}$ such that

$$\Phi(\boldsymbol{x}, \boldsymbol{\xi}) = \psi(s), \quad s = \boldsymbol{x}^T\boldsymbol{\xi}, \quad \boldsymbol{x}, \boldsymbol{\xi} \in \mathbb{S}^2.$$

If we fix $\boldsymbol{\xi} = (\xi, \eta, \zeta)$ and differentiate with respect to $\boldsymbol{x} = (x, y, z)$ then

$$\frac{\partial\Phi}{\partial x}(\boldsymbol{x}, \boldsymbol{\xi}) = \xi\psi'(s), \quad \frac{\partial\Phi}{\partial y}(\boldsymbol{x}, \boldsymbol{\xi}) = \eta\psi'(s), \quad \frac{\partial\Phi}{\partial z}(\boldsymbol{x}, \boldsymbol{\xi}) = \zeta\psi'(s).$$

Using the definition of $\mathcal{G}^x$, $\mathcal{G}^x$ and $\mathcal{G}^z$ in (3.5) we have

$$\begin{aligned}
\mathcal{G}^x\Phi(\boldsymbol{x}, \boldsymbol{\xi}) &= [1-x^2, -xy, -xz]^T\boldsymbol{\xi}\psi'(s) = \xi - xs\psi'(s) \\
\mathcal{G}^y\Phi(\boldsymbol{x}, \boldsymbol{\xi}) &= [-xy, 1-y^2, -yz]^T\boldsymbol{\xi}\psi'(s) = \eta - ys\psi'(s) \\
\mathcal{G}^z\Phi(\boldsymbol{x}, \boldsymbol{\xi}) &= [-xz, -yz, 1-z^2]^T\boldsymbol{\xi}\psi'(s) = \zeta - zs\psi'(s),
\end{aligned}$$

that simply give

$$\nabla_* \Phi(\boldsymbol{x}, \boldsymbol{\xi}) = \boldsymbol{\xi} - s\psi'(s)\boldsymbol{x}, \quad s = \boldsymbol{x}^T \boldsymbol{\xi}. \tag{3.6}$$

If we proceed with the same argument we would have

$$\Delta_* \Phi(\boldsymbol{x}, \boldsymbol{\xi}) = ((1 - s^2)\psi'(s))', \quad s = \boldsymbol{x}^T \boldsymbol{\xi}. \tag{3.7}$$

Obviously $'$ represents the differentiation with respect to s.

Finally we note that the *Green-Beltrami identity* on $\mathbb{S}^d$ [1] leads to

$$\int_{\mathbb{S}^d} v(\boldsymbol{x}) \Delta_* u(\boldsymbol{x})\, d\boldsymbol{x} = -\int_{\mathbb{S}^d} \nabla_* v(\boldsymbol{x}) \cdot \nabla_* u(\boldsymbol{x})\, d\boldsymbol{x} \tag{3.8}$$

for $u \in H^2(\mathbb{S}^d)$ and $v \in H^1(\mathbb{S}^d)$. This will be important in the sequel for constructing a Petrov-Galerkin test discretization for the spherical PDE of this paper.

4 The Petrov-Galerkin method

Consider the diffusion equation

$$\frac{\partial u}{\partial t}(\boldsymbol{x}, t) = \kappa \Delta_* u(\boldsymbol{x}, t) + f(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \mathbb{S}^d, t > 0, \tag{4.9}$$

with the initial condition

$$u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{S}^d,$$

for the given diffusion constant κ, initial function u_0 and the source function f. Although we only consider the above equation, the method of this paper can be easily extended for other time dependent spherical PDEs such as advection-diffusion problems. This method has been originally proposed in [24] and uses the idea of local Petrov-Galerkin methods [3] in Euclidian spaces.

Consider a set of scattered points $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset \mathbb{S}^d$ and let $\vartheta(\boldsymbol{x}, \boldsymbol{y})$ be a compactly supported positive definite kernel with a support on spherical cap $\mathcal{C}_\delta(\boldsymbol{y})$ where $\delta \in (0, \pi)$. Integrating (4.9) against test functions $\vartheta(\cdot, \boldsymbol{x}_k)$, $k = 1, \ldots, N$ and applying the divergence theorem (3.8) yield

$$\frac{d}{dt} \int_{\mathbb{S}^d} u(\boldsymbol{x}, t) \vartheta(\boldsymbol{x}, \boldsymbol{x}_k) d\boldsymbol{x} + \kappa \int_{\mathbb{S}^d} \nabla_* u(\boldsymbol{x}, t) \cdot \nabla_* \vartheta(\boldsymbol{x}, \boldsymbol{x}_k) d\boldsymbol{x} = \int_{\mathbb{S}^d} f(\boldsymbol{x}, t) \vartheta(\boldsymbol{x}, \boldsymbol{x}_k) d\boldsymbol{x}, \tag{4.10}$$

for $k = 1, \ldots, N$. Since $\vartheta(\cdot, \boldsymbol{x}_k)$ are assumed to be compactly supported on caps $\mathcal{C}_\delta(\boldsymbol{x}_k)$, the integrations on whole $\mathbb{S}^d$ can of course be replaced by local integrations on $\mathcal{C}_\delta(\boldsymbol{x}_k)$. This is a local Petrov-Galerkin test discretization which leads to a semi-discrete system of equations. The discretization in spatial domain will be completed by approximating u in kernel space $V_{X,\Phi}$ via

$$u(\boldsymbol{x}, t) \approx u_X(\boldsymbol{x}, t) = \sum_{j=1}^{N} c_j(t) \Phi(\boldsymbol{x}, \boldsymbol{x}_j), \tag{4.11}$$

for another positive definite kernel Φ on $\mathbb{S}^d$. Substituting u_X instead of u into (4.10) then gives

$$M\dot{\boldsymbol{c}}(t) = S\boldsymbol{c}(t) + \boldsymbol{f}(t), \tag{4.12}$$

for mass and stiffness matrices

$$M_{kj} = \int_{\mathcal{C}_\delta(\boldsymbol{x}_k)} \Phi(\boldsymbol{x}, \boldsymbol{x}_j)\vartheta(\boldsymbol{x}, \boldsymbol{x}_k)d\boldsymbol{x}, \tag{4.13}$$

$$S_{kj} = -\kappa \int_{\mathcal{C}_\delta(\boldsymbol{x}_k)} \nabla_*\Phi(\boldsymbol{x}, \boldsymbol{x}_j) \cdot \nabla_*\vartheta(\boldsymbol{x}, \boldsymbol{x}_k)d\boldsymbol{x}, \tag{4.14}$$

and source vector

$$\boldsymbol{f}_k(t) = \int_{\mathcal{C}_\delta(\boldsymbol{x}_k)} f(\boldsymbol{x}, t)\vartheta(\boldsymbol{x}, \boldsymbol{x}_k)d\boldsymbol{x},$$

for $k, j = 1, \ldots, N$. $\dot{\boldsymbol{c}}(t)$ denotes the first derivative of the vector $\boldsymbol{c}(t)$ with respect to the time variable. We need a numerical integration formula on spherical caps. Recently, a class of quadratures with degree of polynomial (spherical harmonics) exactness for numerical integration over spherical caps on $\mathbb{S}^2$ is presented in [12]. According to the following lemma, these quadratures are only required to be constructed for north cap $\mathcal{C}_\delta([0, 0, 1])$ and then rotate to other caps of the same radius.

Lemma 4.1. *(see [12]) Let $\mathcal{C}_\delta(\boldsymbol{x}_0) \subset \mathbb{S}^d$ be the spherical cap with center $\boldsymbol{x}_0 \in \mathbb{S}^d$ and radius $\delta \in (0, \pi)$. Let $Q_{\mathcal{C}_\delta(\boldsymbol{x}_0),n}$, given by*

$$Q_{\mathcal{C}_\delta(\boldsymbol{x}_0),n}(f) := \sum_{j=1}^{n} w_j f(\boldsymbol{z}_j) \approx \int_{\mathcal{C}_\delta(\boldsymbol{x}_0)} f(\boldsymbol{x})d\boldsymbol{x}, \quad f \in C(\mathcal{C}_\delta(\boldsymbol{x}_0)),$$

where $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n \in \mathcal{C}_\delta(\boldsymbol{x}_0)$ and $w_1, w_2, \ldots, w_n \in \mathbb{R}$, be a rule for numerical integration over the spherical cap $\mathcal{C}_\delta(\boldsymbol{x}_0)$ that is exact on $\mathbb{P}_n(\mathcal{C}_\delta(\boldsymbol{x}_0))$. Let $\mathcal{C}_\delta(\boldsymbol{x}'_0) \subset \mathbb{S}^d$ be another spherical cap with center $\boldsymbol{x}'_0 \in \mathbb{S}^d$ and the same radius δ, and let R denotes any rotation on $\mathbb{R}^{d+1}$ such that $\boldsymbol{x}'_0 = R\boldsymbol{x}_0$. Then the rule $Q_{\mathcal{C}_\delta(\boldsymbol{x}_0),n}$, defined by

$$Q_{\mathcal{C}_\delta(\boldsymbol{x}'_0),n}(f) := \sum_{j=1}^{n} w_j f(R\boldsymbol{z}_j), \quad f \in C(\mathcal{C}_\delta(\boldsymbol{x}'_0)),$$

is a rule for numerical integration over $\mathcal{C}_\delta(\boldsymbol{x}'_0)$, with nodes $R\boldsymbol{z}_1, R\boldsymbol{z}_2, \ldots, R\boldsymbol{z}_n \in \mathcal{C}_\delta(\boldsymbol{x}'_0)$.

Remark 4.2. *Using the change of variables $\boldsymbol{x} = (\sqrt{1-\tau^2}\cos\theta, \sqrt{1-\tau^2}\sin\theta, \tau)$ for $\tau \in [-1, 1]$ and $\theta \in [0, 2\pi]$ we simply have*

$$\int_{\mathcal{C}_\delta([0,0,1])} f(\boldsymbol{x})d\boldsymbol{x} = \int_0^{2\pi} \int_{\cos\delta}^{1} F(\tau, \theta)d\tau\, d\theta, \tag{4.15}$$

allowing to use any available quadrature rules for both τ and θ directions on the right-hand side. In numerical examples of section 6 we use the Gauss-Legendre quadrature in both τ and θ variables.

Coming back to (4.12), this is a linear system of first order initial value problems that can be solved by various time integration methods. We will analyze the Crank-Nicolson and the fully implicit methods. In numerical results, for some comparisons, we may use the explicit 4th order Runge-Kutta (RK4) and the implicit 4th order backward differentiation formula (BDF4) schemes. For time stabilities the properties of mass and stiffness matrices M and S are important. In [24] it is proved that if the test kernel ϑ is zonal then both mass and stiffness integrals take a convolution form on the sphere and if, in addition, both Φ and ϑ are positive definite kernels then M and $-S$ are positive definite matrices. The Petrov-Galerkin method in bounded Euclidean domains is usually known as an unsymmetric method. But thanks to the symmetrical properties of the sphere and the test and the trial kernels, the method of this paper produces symmetric matrices for symmetric PDE operators. For instance, for kernel Ψ corresponding to the matrix M, we may write

$$\begin{aligned}
\Psi(\boldsymbol{z},\boldsymbol{y}) &= \int_{\mathcal{C}_\delta(\boldsymbol{z})} \Phi(\boldsymbol{x},\boldsymbol{y})\vartheta(\boldsymbol{x},\boldsymbol{z})d\boldsymbol{x} = \int_{\mathbb{S}^d} \Phi(\boldsymbol{x},\boldsymbol{y})\vartheta(\boldsymbol{x},\boldsymbol{z})d\boldsymbol{x} \\
&= \int_{\mathbb{S}^d} \Big(\sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} \widehat{\phi}_\ell Y_{\ell\,k}(\boldsymbol{x})Y_{\ell\,k}(\boldsymbol{y})\Big)\Big(\sum_{\ell'=0}^{\infty}\sum_{k'=1}^{N(d,\ell')} \widehat{\theta}_{\ell'} Y_{\ell'\,k'}(\boldsymbol{x})Y_{\ell'\,k'}(\boldsymbol{z})\Big)d\boldsymbol{x} \\
&= \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)}\sum_{\ell'=0}^{\infty}\sum_{k'=1}^{N(d,\ell')} \widehat{\phi}_\ell\widehat{\theta}_{\ell'} Y_{\ell\,k}(\boldsymbol{y})Y_{\ell'\,k'}(\boldsymbol{z}) \int_{\mathbb{S}^d} Y_{\ell\,k}(\boldsymbol{x})Y_{\ell'\,k'}(\boldsymbol{x})d\boldsymbol{x} \\
&= \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)}\sum_{\ell'=0}^{\infty}\sum_{k'=1}^{N(d,\ell')} \widehat{\phi}_\ell\widehat{\theta}_{\ell'} Y_{\ell\,k}(\boldsymbol{y})Y_{\ell'\,k'}(\boldsymbol{z})\delta_{\ell\,\ell'}\delta_{k\,k'} \\
&= \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} \widehat{\phi}_\ell\widehat{\theta}_{\ell} Y_{\ell\,k}(\boldsymbol{y})Y_{\ell\,k}(\boldsymbol{z}).
\end{aligned}$$

This shows that $\Psi(\boldsymbol{y},\boldsymbol{z}) = \Psi(\boldsymbol{z},\boldsymbol{y})$ and thus M is a symmetric matrix. The same argument is true for the matrix S, since by using (2.2) we have

$$-\Delta_*\Phi(\boldsymbol{x},\boldsymbol{y}) = \sum_{\ell=0}^{\infty}\sum_{k=1}^{N(d,\ell)} \widehat{\phi}_\ell \ell(\ell+d-1)Y_{\ell\,k}(\boldsymbol{x})Y_{\ell\,k}(\boldsymbol{y}).$$

There exists another issue concerning the size of local caps $\mathcal{C}_\delta(\boldsymbol{x}_k)$ in the above Petrov-Galerkin method. Two strategies are described in [24] for choosing δ. Let $h_{X,\mathbb{S}^d}$ be the fill distance of points in X defined by

$$h_{X,\mathbb{S}^d} := \sup_{\boldsymbol{x}\in\mathbb{S}^d} \min_{1\le j\le N} \operatorname{dist}(\boldsymbol{x},\boldsymbol{x}_j).$$

If δ is chosen to be proportional to $h_{X,\mathbb{S}^d}$, i.e. $\delta = c_{cap}h_{X,\mathbb{S}^d}$ for a constant c_{cap}, then the method is called the *stationary* Petrov-Galerkin. While the *nonstationary* Petrov-Galerkin method addresses the case $\delta = c_{cap}$ for a constant $c_{cap} \in (0,\pi)$. Theoretical and numerical results of [24] show that the stationary method has no significant advantage over the collocation method. This is the reason why we only use the nonstationary method in numerical results of section 6.

We will compare the method of this paper with the classical collocation method that is briefly discussed here. The numerical solution u_X in (4.11) could be determined by collocating (4.9) in trial points X as

$$\sum_{j=1}^{N} \frac{dc_j(t)}{dt}\Phi(\boldsymbol{x}_k,\boldsymbol{x}_j) = \kappa \sum_{j=1}^{N} c_j(t)\Delta_*\Phi(\boldsymbol{x}_k,\boldsymbol{x}_j) + f(\boldsymbol{x}_k,t), \quad k=1,\ldots,N, \quad t>0, \tag{4.16}$$

which yields a linear system same as (4.12) with the following stiffness and mass matrices

$$\begin{aligned} M_{kj} &= \Phi(\boldsymbol{x}_k,\boldsymbol{x}_j), \\ S_{kj} &= \kappa\Delta_*\Phi(\boldsymbol{x}_k,\boldsymbol{x}_j), \quad k,j=1,2,\ldots,N, \end{aligned} \tag{4.17}$$

and the source vector

$$\boldsymbol{f}_k(t) = f(\boldsymbol{x}_k,t), \quad k=1,2,\ldots,N.$$

Both collocation matrices M and $-S$ are positive definite [25].

5 Stability of time integration

Using the Crank-Nicolson scheme for (4.12) we get

$$M\frac{\boldsymbol{c}^{n+1}-\boldsymbol{c}^n}{\Delta t} = \frac{1}{2}S\big(\boldsymbol{c}^{n+1}+\boldsymbol{c}^n\big) + \frac{1}{2}\big(\boldsymbol{f}^n+\boldsymbol{f}^{n+1}\big),$$

where $\boldsymbol{c}^n$ and $\boldsymbol{f}^n$ are the values of the coefficient vectors $\boldsymbol{c}(t)$ and $\boldsymbol{f}(t)$ at n'th time level. This yields

$$\boldsymbol{c}^{n+1} = \big(I-\frac{\Delta t}{2}M^{-1}S\big)^{-1}\Big(\big(I+\frac{\Delta t}{2}M^{-1}S\big)\boldsymbol{c}^n + \frac{1}{2}\Delta t M^{-1}\big(\boldsymbol{f}^n+\boldsymbol{f}^{n+1}\big)\Big), \tag{5.18}$$

where I is the identity matrix.

Remark 5.1. *Recall that in [24] it is proved that both matrices $-S$ and M which are corresponded to the Laplace-Beltrami and identity operators, respectively, are positive definite. This implies that $I-\frac{\Delta t}{2}M^{-1}S$ is positive definite with eigenvalues all bigger than* 1. *In particular, the inverse matrix in* (5.18) *exists.*

The L_2 stability is resulted if

$$\Big|\lambda\Big(\big(I-\frac{\Delta t}{2}M^{-1}S\big)^{-1}\big(I+\frac{\Delta t}{2}M^{-1}S\big)\Big)\Big| = \left|\frac{1+\frac{\Delta t}{2}\lambda\big(M^{-1}S\big)}{1-\frac{\Delta t}{2}\lambda\big(M^{-1}S\big)}\right| \le 1,$$

or, equivalently

$$-1 \le \frac{1+\frac{\Delta t}{2}\lambda\big(M^{-1}S\big)}{1-\frac{\Delta t}{2}\lambda\big(M^{-1}S\big)} \le 1, \tag{5.19}$$

where $\lambda(A)$ stands for any eigenvalue of matrix A. According to Remark 5.1, (5.19) is satisfied regardless of the size of time step Δt. This proves that the scheme is unconditionally stable.

The fully implicit time integration can also be used as

$$M\frac{\boldsymbol{c}^{n+1}-\boldsymbol{c}^{n}}{\Delta t}=S\boldsymbol{c}^{n+1}+\boldsymbol{f}^{n+1}.$$

A similar eigenvalue analysis gives

$$-1\leq\frac{1+\Delta t\lambda\big(M^{-1}S\big)}{1-\Delta t\lambda\big(M^{-1}S\big)}\leq 1.$$

which means that this method is also unconditionally stable.

Remark 5.2. *For the nonstationary Petrov-Galerkin method on quasi-uniform points, as is proved in [24], for a trial kernel Φ with Fourier coefficients $\widehat{\phi}_\ell$, and a test kernel ϑ with Fourier coefficients $\widehat{\theta}_\ell$ that decay at infinity as $\widehat{\phi}_\ell\sim(1+\ell)^{-2\sigma}$ and $\widehat{\theta}_\ell\sim(1+\ell)^{-2\mu}$, respectively, the spatial error bound*

$$\|u-u_X\|_{H^\beta(\mathbb{S}^d)}\leqslant Ch_{X,\mathbb{S}^d}^{\gamma-\beta}\|u\|_{H^\gamma(\mathbb{S}^d)},$$

holds for all $u\in H^\gamma(\mathbb{S}^d)$, where $\gamma>d/2$, $\gamma\geq 2$, $\gamma\leq 2+2\sigma-2\mu$, and $0\leq\beta\leq\gamma$. Here, $H^\gamma(\mathbb{S}^d)$ stands for the Sobolev space of order γ on $\mathbb{S}^d$. If an unconditionally stable time integration method of order $(\Delta t)^k$ is employed for the diffusion equation, the order of convergence $\mathcal{O}((\Delta t)^k,h_{X,\mathbb{S}^d}^\gamma)$ should be obtained provided that the error is measured in the L_2 norm. Note that, since the solution of the diffusion equation is smoothed out after a time instance, the full order $h_{X,\mathbb{S}^d}^\gamma$ for $\gamma=2+2\sigma-2\mu$ should be attained regardless of the smoothness of the initial heat profile.

6 Numerical Examples

Algorithms are implemented in MATLAB and executed on a machine with an Intel Core i7 processor, 4.00 GHz and 16 GB RAM. In experiments we use the restricted compactly supported Wendland's function

$$\phi_{3,1}(r)=(1-r)_+^4(4r+1),\quad r=\|\boldsymbol{x}\|_2,\quad \boldsymbol{x}\in\mathbb{R}^3$$

on $\mathbb{S}^2$ as a test kernel to define $\vartheta(\boldsymbol{x},\boldsymbol{y})=\phi_{3,1}(\sqrt{2-2\boldsymbol{x}^T\boldsymbol{y}})=:\upsilon(s)$ for $s=\boldsymbol{x}^T\boldsymbol{y}$. The radius of local spherical caps (support size of test functions) is fixed at $\delta=0.1$ in all experiments which addresses the nonstationary Petrov-Galerkin method. Different types (globally or compactly supported) of restricted positive definite functions can be employed as trial kernels. Here, the globally supported Matérn's functions

$$\begin{aligned}\phi(r)&=\phi_{2.5}(r)=\exp(-r)(1+r),\\ \phi(r)&=\phi_{3.5}(r)=\exp(-r)(3+3r+r^2),\end{aligned}$$

and the compactly supported Wendland's function

$$\phi(r)=\phi_{3,2}(r)=(1-r)_+^6(35r^2+18r+3),$$

for $r = \|\boldsymbol{x}\|_2$ and $\boldsymbol{x} \in \mathbb{R}^3$, are used to define the zonal kernels

$$\Phi(\boldsymbol{x}, \boldsymbol{\xi}) = \phi(\sqrt{2 - 2\boldsymbol{x}^T\boldsymbol{\xi}}).$$

In experiments we use $\phi(\varepsilon r)$ where ε is the shape parameter. For the compactly supported function, $1/\varepsilon$ defines the size of the SBF support. Although numerical results are affected by ε, we do not aim to numerically optimize this parameter. In experiments, we use $\varepsilon = 1, 5, 7$ for $\phi_{3,2}$, $\phi_{2.5}$ and $\phi_{3.5}$, respectively. Quasi-uniformly scattered points on $\mathbb{S}^2$ are obtained by the *equal area partitioning* algorithm of [36]. Figure 1 shows sets of $N = 500$ and $N = 1500$ points generated by this algorithm.

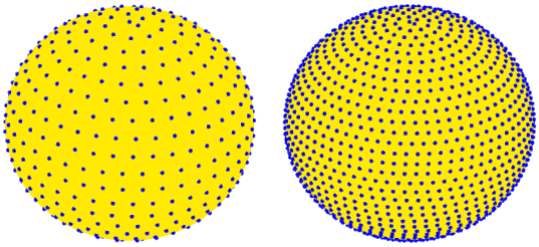

Figure 1: The unit sphere with 500 (left) and 1500 (right) quasi-uniform scattered points.

For numerical integration on spherical caps, as is discussed in Remark 4.2, we use $80 = 8 \times 10$ Gauss-Legendre points in τ and θ variables, respectively.

For comparison, in some examples the results of both collocation and nonstationary Petrov-Galerkin methods are reported. Since the trial points are quasi-uniformly distributed on $\mathbb{S}^2$, the numerical convergence orders are computed via

$$\log\left(\frac{\text{error}_{\text{old}}}{\text{error}_{\text{new}}}\right) \Big/ \log\left(\sqrt{\frac{N_{\text{new}}}{N_{\text{old}}}}\right).$$

The approximate orders are obtained by the linear least squares fitting to error values in terms of $h_{X,\mathbb{S}^2}$ in the logarithmic scale. Since $h = \mathcal{O}(N^{-1/2})$, we plot the errors and the stability numbers vs. $\sqrt{N}$. All convergence plots are on a log-log scale.

In the time domain, the backward Euler (BDF1) and the Crank-Nicolson (CN) schemes are used in almost all experiments. However, for comparison with higher order schemes, we also employ the 4th order Runge-Kutta method (RK4) and the 4th order backward differentiation formula (BDF4). The RK4 scheme is an explicit formula with a small stability region, while the BDF4 is an implicit scheme with a relatively large stability region. Time steps $\Delta t = 2 \times 10^{-5}, 5 \times 10^{-3}, 10^{-2}$ are used for BDF1, CN and BDF4 schemes, respectively. The MATLAB's command `ode45` for relative and absolute errors of 10^{-10} is employed for the RK4 scheme.

Example 6.1. In this example, we consider equation (4.9) with $\kappa = 1$ and the exact solution

$$u(\boldsymbol{x}, t) = \exp(-t) \sum_{i=1}^{23} \cos^4\left(\frac{\pi}{2}\boldsymbol{x}^T\boldsymbol{\xi}_i\right)$$

where $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{23}\}$ is a set of scattered points on $\mathbb{S}^2$ generated by the equal area partitioning algorithm. The forcing term that makes this solution hold is given by

$$f(\boldsymbol{x}, t) = -4\exp(-t)\sum_{k=1}^{23}\cos^2\big(\frac{\pi}{2}\boldsymbol{x}^T\boldsymbol{\xi}_k\big)\Big[\Big(0.25-\pi^2\big(1-(\boldsymbol{x}^T\boldsymbol{\xi}_k)^2\big)\Big)\cos^2\big(\frac{\pi}{2}\boldsymbol{x}^T\boldsymbol{\xi}_k\big) + \pi\boldsymbol{x}^T\boldsymbol{\xi}_k\cos\big(\frac{\pi}{2}\boldsymbol{x}^T\boldsymbol{\xi}_k\big)\sin\big(\frac{\pi}{2}\boldsymbol{x}^T\boldsymbol{\xi}_k\big) + 0.75\pi^2\big(1-(\boldsymbol{x}^T\boldsymbol{\xi}_k)^2\big)\Big].$$

In the results that follow we estimate the order of convergence of the Petrov-Galerkin and the collocation methods versus N, the number of trial points. In Figure 2 the error plots for the Matérn's kernels $\phi_{2.5}$ and $\phi_{3.5}$, and the Wendland's kernel $\phi_{3,2}$ are shown where BDF1, CN, RK4 and BDF4 schemes are used in the time domain. The convergence rate of Petrov-Galerkin method outperforms that of collocation method for all three trial kernels. Since these kernels produce $H^{2.5}(\mathbb{S}^2)$, $H^{3.5}(\mathbb{S}^2)$ and $H^{3.5}(\mathbb{S}^2)$, respectively, as their native spaces on the unit sphere, the observed numerical orders 5, 7 and 7 are expectable for the Petrov-Galerkin method.

In Table 1, the CPU times of both methods to achieve the same accuracy are compared. As we observe, the desired relative errors are obtained using the Petrov-Galerkin method with much fewer (approximately halved) number of trial points than that for the collocation method, leading to smaller CPU times. The results of this table are obtained by trial kernel $\phi_{3,2}$ and the CN scheme with time step $\Delta t = 0.0001$.

Relative Error $\rightarrow$	$\leq 10^{-2}$	$\leq 10^{-3}$	$\leq 10^{-4}$
Petrov-Galerkin	$(500, 1.0)$	$(1000, 5.0)$	$(1850, 24.0)$
Collocation	$(700, 1.0)$	$(1700, 10.0)$	$(4950, 82.0)$

Table 1: Comparing the CPU times (sec.) of the Petrov-Galerkin and the collocation methods to obtain the same accuracy; The kernel $\phi_{3,2}$ and the CN scheme with $\Delta t = 10^{-4}$ are used. Here, $(N, s) =$ (number of trial points, CPU time).

Finally, Figures 3 shows the eigenvalues of $-S$ and M (mass and stiffness matrices) for $N = 1000$. Both M and $-S$ are theoretically positive definite. Small imaginary parts appeared in numerical eigenvalues can be reduced to the machine's precision by using more integration points on spherical caps. See the bottom plots of the figure.

In the numerical algorithm, the mass matrix M, the matrix $I - \frac{1}{2}\Delta t\, M^{-1}S$ in the CN scheme, the matrix $I - \Delta t\, M^{-1}S$ in the the BDF1 scheme and the matrix $I - \frac{12}{25}\Delta t\, M^{-1}S$ in the BDF4 scheme should be inverted. In Table 2 we report the 2-norm condition numbers of M and $I - \frac{1}{2}\Delta t\, M^{-1}S$ for $\Delta t = 0.005$ when the Wendland's kernel $\phi_{3,2}$ is used for trial discretization. A computational numerical order around -4 is observed for the mass matrix. The condition number of the CN matrix is extremely small and enjoys a mildly increasing order. The same holds true for BDF matrices which are not presented here.

Example 6.2. Consider the function

$$G(s) = \sum_{\ell=1}^{K}\frac{P_\ell(s)}{\ell(\ell+1)}, \quad K \in \mathbb{N},\ s \in \mathbb{R}, \tag{6.20}$$

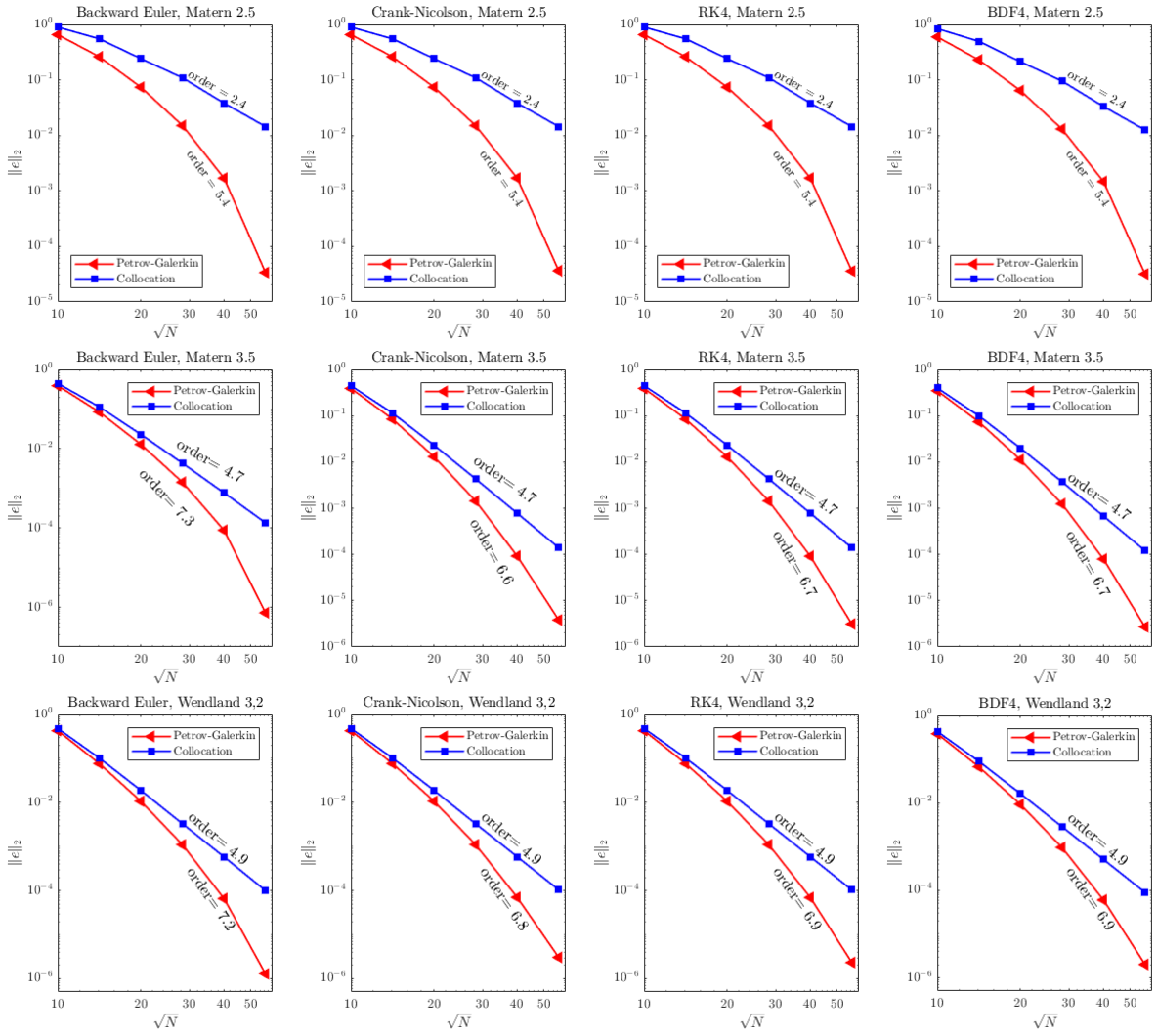

Figure 2: Errors and convergence orders of the Petrov-Galerkin and the collocation methods by using the Matérn's kernel $\phi_{2.5}$ (first row), the Matérn's kernel $\phi_{3.5}$ (second row) and the Wendland's kernel $\phi_{3,2}$ (last row). In the time domain the backward Euler scheme with $\Delta t = 2 \times 10^{-5}$ (first column), the Crank-Nicolson scheme with $\Delta t = 5 \times 10^{-3}$ (second column), the 4th order Runge-Kutta command of Matlab (third column) and the BDF4 scheme with $\Delta t = 10^{-2}$ (last column) are used.

where P_ℓ is the Legendre polynomial of degree ℓ. The exact solution of problem (4.9), for $f(\boldsymbol{x}, t) = 0$, $\kappa = 1$ and $u(\boldsymbol{x}, 0) = G(\boldsymbol{x}^T \boldsymbol{p})$ is

$$u(\boldsymbol{x}, t) = \sum_{\ell=1}^{K} \frac{e^{-\ell(\ell+1)t}}{\ell(\ell+1)} P_\ell(\boldsymbol{x}^T \boldsymbol{p}),$$

where $\boldsymbol{p} = (0, 0, 1)^T$ is the north pole of the unit sphere. When $K \to \infty$, this solution describes the heat diffusion on the unit sphere from its north pole onto its whole surface [20]. Here we set $K = 160$.

Errors and numerical convergence orders with respect to N are shown in Figure 4 for the Mat'ern's kernel $\phi_{2.5}$. Though not illustrated here, the results of the other kernels are

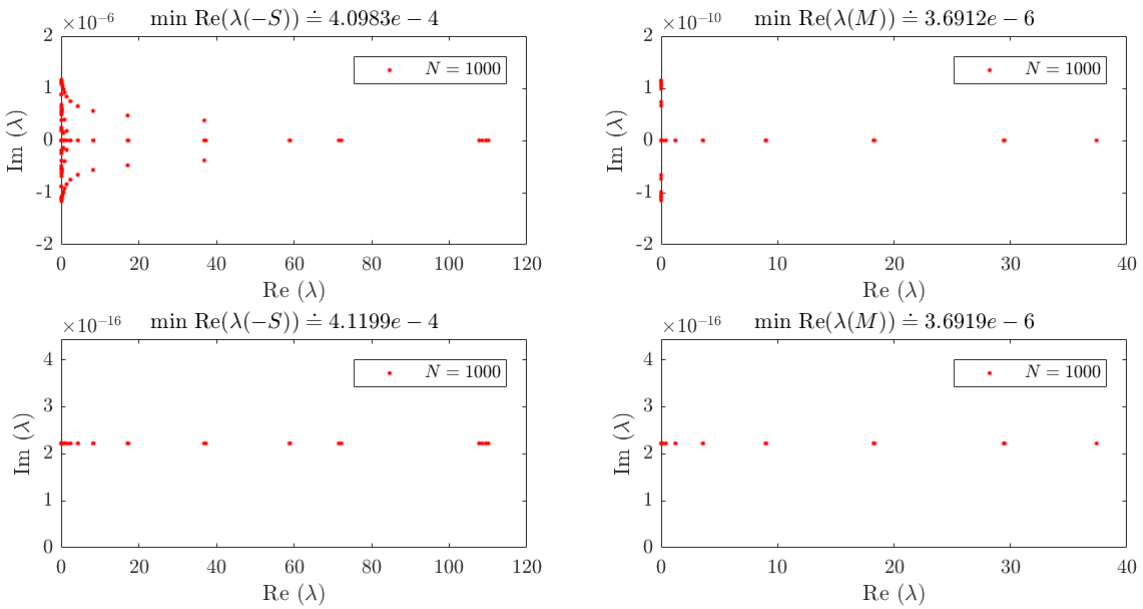

Figure 3: The eigenvalue patterns of positive definite matrices $-S$ (left) and M (right) with $N = 1000$ scattered points. Small imaginary parts in the top plots are due to the error of the numerical integration. In the bottom plots the size of imaginary part is reduced to the machine's precision by using a quadrature with more integration points.

Table 2: Condition numbers of mass matrix M and Crank-Nicolson matrix $A_{CN} = I - \frac{dt}{2}M^{-1}S$ and numerical orders.

N	cond(M)	order	cond(A_{CN})	order
100	$1.96e+3$		$1.34e+0$	
200	$2.45e+4$	-3.6	$1.65e+0$	-0.3
400	$3.12e+5$	-3.7	$2.35e+0$	-0.5
800	$4.25e+6$	-3.8	$3.69e+0$	-0.7
1600	$6.77e+7$	-4.0	$6.27e+0$	-0.8
3200	$1.74e+9$	-4.7	$1.14e+1$	-0.9

more or less the same as those in the first example.

In Figure 5, the diffusion of the initial heat density $u(\boldsymbol{x}, 0)$ to the sphere is drawn at some different time levels.

In Figure 6, the errors and convergence rates of the Petrov-Galerkin method versus Δt are plotted for BDF1 and CN schemes. In this case, to eliminate the error of the space discretization, we use a fine point density with $N = 3000$. Theoretically, the orders should be 1 and 2 for BDF1 and CN schemes, respectively. Experiments show better results in some cases.

Finally, in Table 3 the CPU times for obtaining the same accuracy are compared. The Petrov-Galerkin method uses a fewer number of points and a less CPU time to reach a prescribed accuracy.

Example 6.3. In this example we consider $f = 0$ and $\kappa = 1/42$ with an initial condition for which there is no closed form solution. We set the initial heat profile to a sum of five

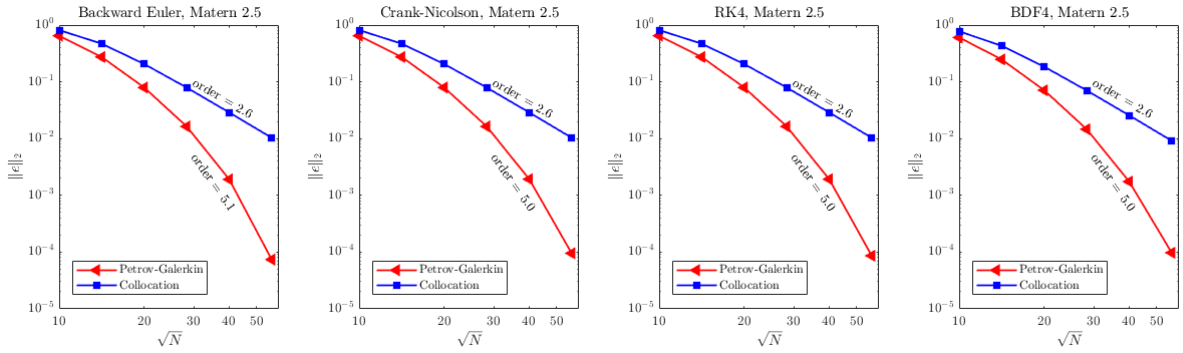

Figure 4: Errors and convergence orders of the Petrov-Galerkin and the collocation methods by using the Matérn's kernel $\phi_{2.5}$ In the time domain the backward Euler scheme with $\Delta t = 2 \times 10^{-5}$ (first column), the Crank-Nicolson scheme with $\Delta t = 5 \times 10^{-3}$ (second column), the 4th order Runge-Kutta command of Matlab (third column) and the BDF4 scheme with $\Delta t = 10^{-2}$ (last column) are used.

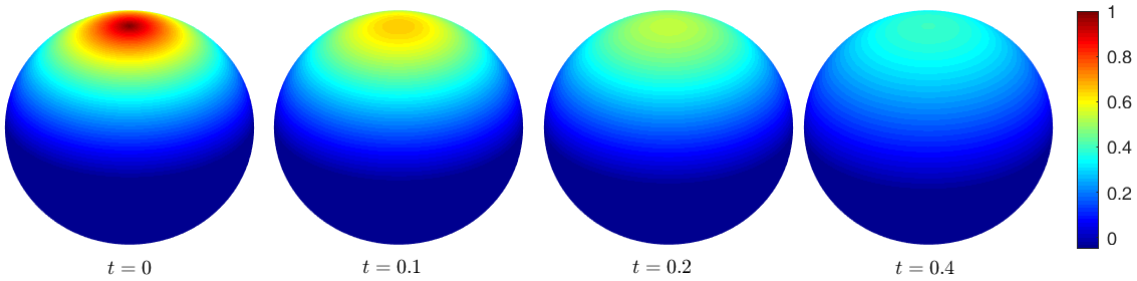

Figure 5: The diffusion of initial heat density $u(\boldsymbol{x}, 0) = G(z)$ at some time values.

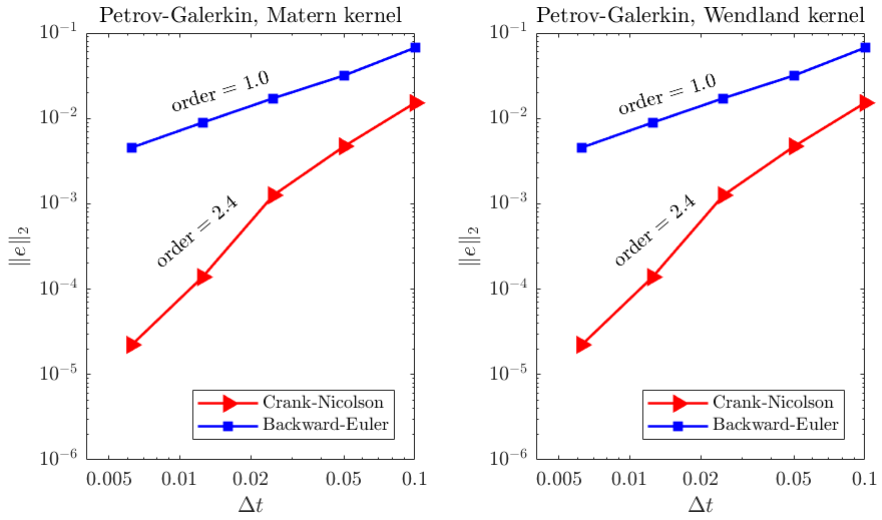

Figure 6: Convergence of the Petrov-Galerkin method with respect to the time step Δt for backward Euler and Crank-Nicolson schemes with Matérn's kernel $\phi_{3.5}$ and Wendland's kernels $\phi_{3,2}$ at $N = 3000$ scattered points.

Relative Error →	$\leq 10^{-2}$	$\leq 10^{-3}$	$\leq 10^{-4}$	$\leq 10^{-5}$
Petrov-Galerkin	$(500, 1.0)$	$(700, 2.0)$	$(1250, 10.0)$	$(5300, 206.0)$
Collocation	$(650, 1.0)$	$(1400, 7.0)$	$(3750, 46.0)$	$(9150, 307.0)$

Table 3: Comparing the CPU times (sec.) of the Petrov-Galerkin and the collocation methods to obtain the same accuracy; The kernel $\phi_{3,2}$ and the CN scheme with $\Delta t = 10^{-4}$ are used. Here, $(N, s) =$ (number of trial points, CPU time).

Gaussian bumps

$$u(\boldsymbol{x}, 0) = \sum_{k=1}^{5} e^{-30((x-\xi_k)^2+(y-\eta_k)^2+(z-\zeta_k)^2)},$$

where $\boldsymbol{x} = (x, y, z) \in \mathbb{S}^2$, $(\xi_k, \eta_k, \zeta_k) = \boldsymbol{\xi}_k \in \mathbb{S}^2$ and $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_5\}$ is a set of 5 random points [42]. For this problem the total amount of heat is conserved because

$$\frac{\partial}{\partial t} \int_{\mathbb{S}^d} u(\boldsymbol{x}, t) d\boldsymbol{x} = \int_{\mathbb{S}^d} \frac{\partial}{\partial t} u(\boldsymbol{x}, t) d\boldsymbol{x} = \kappa \int_{\mathbb{S}^d} \Delta_* u(\boldsymbol{x}, t) d\boldsymbol{x} = 0.$$

The last equality is immediately followed form the Green-Beltrami identity (3.8) by choosing $v = 1$. Thus, if we define the mean value

$$m(t) := \frac{1}{\text{vol}(\mathbb{S}^d)} \int_{\mathbb{S}^d} u(\boldsymbol{x}, t) d\boldsymbol{x}$$

then $m(t) = m(0)$ for $t \geq 0$. This conservation property can be used to verify the accuracy of the numerical method. For computing $m(t)$, we employ a Gauss-Legendre quadrature with 10^4 integration points on the whole $\mathbb{S}^2$.

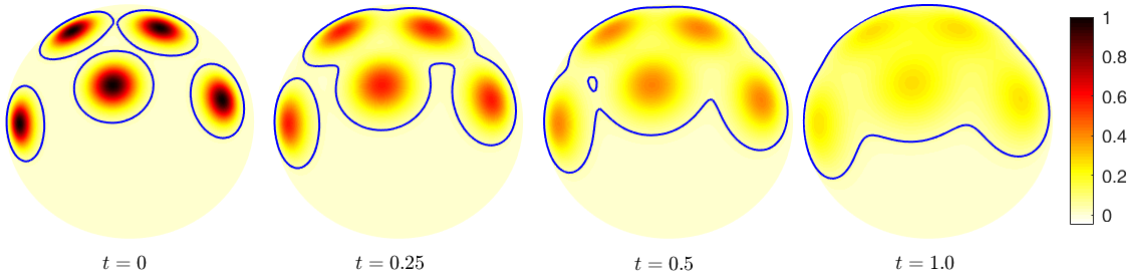

Figure 7: The diffusion of Gaussian bumps at some different time values, and contours to track the mean of the initial solution.

The Wendland's kernel $\phi_{3,2}$ and $N = 2000$ trial points are used for spatial discretization, and the CN scheme with $\Delta t = 0.005$ is applied for time integration. In Figure 7, the diffusion of the heat profile is shown at times $t = 0, 0.25, 0.5$ and 1. We also plot a contour that tracks the mean of the initial condition solution which is obtained as $m(0) \doteq 4.1667 \times 10^{-2}$. As is shown in Figure 8, the numerical scheme preserves the mean value property.

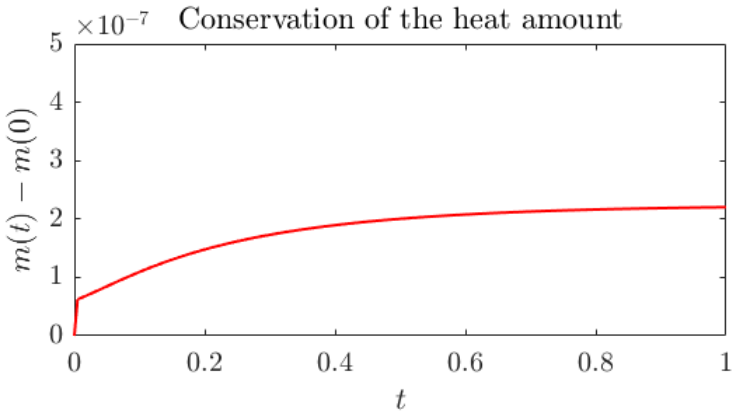

Figure 8: Conservation of the heat amount.

7 Conclusion

In this paper, we present the Petrov-Galerkin SBF approximation for solving the diffusion equation on the unit sphere. The method uses a Petrov-Galerkin test discretization which is different from the well-considered collocation and Galerkin methods on the sphere. For each testing functional equation, numerical integration is simply done on a local spherical cap. This means that spherical caps are used as local integration supports and play the role of spherical triangles in the standard Galerkin finite element. Here, the spherical caps may overlap but they are used independent from each other without any connectivity assumption. In the time domain, the Crank-Nicolson and fully implicit schemes are analyzed. In numerical results the 4th order Runge-Kutta method (RK4) and the 4th order backward differentiation formula (BDF4) are also employed. At the price of numerical integration, numerical results show that the new method is more accurate than the classical collocation method.

The proposed method can also be applied for nonlinear diffusion equations on the unit sphere. For this purpose, one can use the semi-implicit (implicit for linear and explicit for nonlinear terms) time difference scheme that still leads to a final linear system. Another technique is to use a time splitting scheme to separate the PDE into the linear and the nonlinear terms, and then use the method of this paper for the linear part, and an exact ODE solution for the nonlinear part. As some examples of such schemes in the context of meshless methods, we refer the reader to [27, 28].

Acknowledgments

Supports from University of Isfahan and IPM-Isfahan during the first author's sabbatical leave are greatly acknowledged. The second author was in part supported by a Grant from IPM, No. 99650421. Authors are very grateful to reviewers for carefully reading this paper and for their comments and suggestions which have improved the paper.

References

[1] K. Atkinson and W. Han. *Spherical Harminics and Approximations on the Unit Sphere: An Introduction*. Springer, New York, 2012.

[2] S. N. Atluri and S. Shen. *The Meshless Local Petrov-Galerkin (MLPG) Method*. Tech Science Press, Encino, CA, 2002.

[3] S. N. Atluri and T.-L. Zhu. A new meshless local Petrov-Galerkin (MLPG) approach in Computational mechanics. *Comput. Mech.*, 22:117–127, 1998.

[4] T. Belytschko, Y. Y. Lu, and L. Gu. Element-Free Galerkin methods. *Int. J. Numer. Methods Eng.*, 37:229–256, 1994.

[5] D. Chen, V. A. Menegatto, and X. Sun. A necessary and sufficient condition for strictly positive definite functions on spheres. *Proceedings of the American Mathematical Society*, 131:2733–2740, 09 2003.

[6] M. Dehghan and N. Narimani. Approximation of continuous surface differential operators with the generalized moving least-squares (GMLS) method for solving reaction-diffusion equation. *Comp. Appl. Math.*, 37:6955–6971, 2018.

[7] G. E. Fasshauer and L. L. Schumaker. Scattered data fitting on the sphere. In *Proceedings of the International Conference on Mathematical Methods for Curves and Surfaces II Lillehammer, 1997*, pages 117–166, Nashville, TN, USA, 1998. Vanderbilt University.

[8] W. Freeden, T. Gervens, and M. Schreiner. *Constructive Approximation on the Sphere with Applications to Geomathematics*. Clarendon Press, Oxford, 1998.

[9] E. Fuselier and G. B. Wright. Scattered data interpolation on embeded submanifolds with restricted positive definite kernels: Sobolev error estimates. *SIAM Journal on Numerical Analysis*, 50:1753–1776, 2012.

[10] R. L. Harder and R. N. Desmarais. Interpolation using surface splines. *J. Aircraft*, 9:189–191, 1972.

[11] R. L. Hardy. Multiquadric equations of topography and other irregular surfaces. *J. Geophys. Res.*, 76:1905–1915, 1971.

[12] K. Hesse and R. S. Womersley. Numerical integration with polynomial exactness over a spherical cap. *Advances in Computational Mathematics*, 36:451–483, 2012.

[13] S. Hubbert. *Radial basis function interpolation on the sphere*. PhD thesis, Imperial College London, London, 2002.

[14] S. Hubbert and T. M. Morton. A Duchon framework for the sphere. *Journal of Approximation Theory*, 129:28–57, 2004.

[15] K. Jetter, J. Stöckler, and J. D. Ward. Error estimates for scattered data interpolation on spheres. *Mathematics of Computation*, 68:733–747, 1999.

[16] E. J. Kansa. Multiquadrics —A scattered data approximation scheme with applications to computational fluid-dynamics, I. Surface approximations and partial derivative estimates. *Computers & Mathematics with Applications*, 19:127–145, 1990.

[17] E. J. Kansa. Multiquadrics —A scattered data approximation scheme with applications to computational fluid-dynamics, II. Solutions to parabolic, hyperbolic and elliptic partial differential equations. *Computers & Mathematics with Applications*, 19:147–161, 1990.

[18] P. Lancaster and K. Salkauskas. Surfaces generated by moving least squares methods. *Mathematics of Computation*, 37:141–158, 1981.

[19] Q. T. Le Gia. Galerkin approximation for elliptic PDEs on spheres. *Journal of Approximation Theory*, 130:123–147, 2004.

[20] Q. T. Le Gia. Approximation of parabolic PDEs on spheres using spherical basis functions. *Advances in Computational Mathematics*, 22:377–397, 2005.

[21] Q. T. Le Gia, F. J. Narcowich, J. D. Ward, and H. Wendland. Continuous and discrete least-square approximation by radial basis functions on spheres. *J. Approx. Theory*, 143:124–133, 2006.

[22] Q. T. Le Gia, I. H. Sloan, and H. Wendland. Multiscale RBF collocation for solving PDEs on spheres. *Numer. Math.*, 121:99–125, 2012.

[23] D. Mirzaei. Direct approximation on spheres using generalized moving least squares. *BIT Numerical Mathematics*, 57:1041–1063, 2017.

[24] D. Mirzaei. A Petrov–Galerkin kernel approximation on the sphere. *SIAM Journal on Numerical Analysis*, 56:274–295, 2018.

[25] D. Mirzaei. On analysis of kernel collocation methods for spherical PDEs. *Applied Numerical Mathematics*, 150:222–232, 2020.

[26] D. Mirzaei and R. Schaback. Direct Meshless Local Petrov-Galerkin (DMLPG) method: a generalized MLS approximation. *Applied Numerical Mathematics*, 33:73–82, 2013.

[27] V. Mohammadi and M. Dehghan. A meshless technique based on generalized moving least squares combined with the second-order semi-implicit backward differential formula for numerically solving time-dependent phase field models on the spheres. *Applied Numerical Mathematics*, 153:248–275, 2020.

[28] V. Mohammadi, D. Mirzaei, and M. Dehghan. Numerical simulation and error estimation of the time-dependent Allen-Cahn equation on surfaces with radial basis functions. *Journal of Scientific Computing*, 79:493–516, 2019.

[29] T. M. Morton and M. Neamtu. Error bounds for solving pseudodifferential equations on spheres. *Journal of Approximation Theory*, 114:242–268, 2002.

[30] F. J. Narcowich, S. T. Rowe, and J. D. Ward. A novel Galerkin method for solving PDEs on the sphere using highly localized kernel bases. *Mathematics of Computation*, 86:197–231, 2017.

[31] F. J. Narcowich, X. Sun, and J. D. Ward. Approximation power of RBFs and their associated SBFs: A connection. *Advances in Computational Mathematics*, 27:107–124, 2007.

[32] F. J. Narcowich and J. D. Ward. Generalized Hermite interpolation via matrix-valued conditionally positive definite functions. *Mathematics of Computation*, 63:661–687, 1994.

[33] F. J. Narcowich and J. D. Ward. Scattered-data interpolation on spheres: error estimates and locally supported basis functions. *SIAM J. Math. Anal.*, 33:1393–1410, 2002.

[34] A. Ron and X. Sun. Strictly positive definite functions on spheres in Euclidean spaces. *Mathematics of Computation*, 65:1513–1530, 1996.

[35] M. Safarpoor and A. Shirzadi. A localized RBF-MLPG method for numerical study of heat and mass transfer equations in elliptic fins. *Engineering Analysis with Boundary Elements*, 98:35 – 45, 2019.

[36] E. B. Saff and A. B. J. Kuijlaars. Distributing many points on a sphere. *Math. Intell.*, 19:5–11, 1997.

[37] R. Schaback. Unsymmetric meshless methods for operator equations. *Numer. Math.*, 114:629–651, 2010.

[38] I.J. Schoenberg. Positive definite function on spheres. *Duke Math. J.*, 9:96–108, 1942.

[39] M. Schreiner. On a new condition for strictly positive definite functions on spheres. *Proceedings of the American Mathematical Society*, 125(2):531–539, 1997.

[40] D. Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 23th National Conference ACM*, pages 517–523, 1968.

[41] A. Shirzadi and L. Ling. Convergent overdetermined-RBF-MLPG for solving second order elliptic PDEs. *Advances in Applied Mathematics and Mechanics*, 5:78–89, 2013.

[42] A. Townsend and G. B. Wright. Solving the heat equation on the unit sphere. Chebfun Example, 2016.

[43] H. Wendland. Meshless Galerkin methods using radial basis functions. *Mathematics of Computation*, 68:1521–1531, 1999.

[44] H. Wendland. Moving least squares approximation on the sphere. In T. Lyche and L. L. Schumaker, editors, *Mathematical Methods in CAGD*, pages 1–10, Nashville, TN, 2001. Vanderbilt University Press.

[45] H. Wendland. *Scattered Data Approximation*. Cambridge University Press, 2005.

[46] Z. Wu. Hermite Birkhoff interpolation of scattered data by radial basis functions. *Approx. Theory Appl.*, 8:1–10, 1992.

[47] Y. Xu and E. W. Cheney. Strictly positive definite functions on spheres. *Proc. Am. Math. Soc.*, 116:977–981, 1992.