
Two-step Generalized RBF-Generated Finite Difference Method on Manifolds

Rongji Li · Haichuan Di · Shixiao W.
Jiang

Received: date / Accepted: date

Abstract Solving partial differential equations (PDEs) on manifolds has broad applications in various fields. In this paper, we develop a two-step generalized radial basis function-generated finite difference (gRBF-FD) method for solving PDEs on manifolds without boundaries, identified by randomly sampled point cloud data. The gRBF-FD is based on polyharmonic spline kernels and multivariate polynomials (PHS+Poly) defined over the tangent space in a local Monge coordinate system. The first step is to regress the local target function using a generalized moving least-squares (GMLS) while the second step is to compensate for the residual using a PHS interpolation. Our gRBF-FD method has the same interpolant form with the standard RBF-FD but differs in interpolation coefficients. Our approach utilizes a specific weight function in both the GMLS and PHS steps and implements an automatic tuning strategy for the stencil size K (i.e., the number of nearest neighbors) at each point. These strategies are designed to produce a Laplacian matrix with a specific coefficient structure, thereby enhancing stability and reducing the solution error. We establish an error bound for the operator approximation in terms of the so-called local stencil diameter as well as in terms of the number of data. We further demonstrate the high accuracy of gRBF-FD through numerical tests on various smooth manifolds from 1D to 4D.

Rongji Li
School of Information Science and Technology, ShanghaiTech University, Shanghai 201210,
China
E-mail: lirj2022@shanghaitech.edu.cn

Haichuan Di
School of Information Science and Technology, ShanghaiTech University, Shanghai 201210,
China
E-mail: dihch2024@shanghaitech.edu.cn

Shixiao W. Jiang
Institute of Mathematical Sciences, ShanghaiTech University, Shanghai 201210, China
E-mail: jiangshx@shanghaitech.edu.cn

Keywords Generalized RBF-generated finite difference, Randomly sampled point cloud data, Screened Poisson equations on manifolds, Numerical stability

Mathematics Subject Classification (2020) 41A10 · 41A15 · 58J05 · 65N12

1 Introduction

Partial differential equations (PDEs) on manifolds have a wide range of applications in physics, biology, and engineering problems. For example, the applications include phase separation on biomembranes [17], liquid crystal alignment on deformable interfaces [55], morphogen transport on cell surfaces [5], image processing on curved domains [9] and surface reconstruction [73], just to list a few. These broad applications motivate the development of accurate and efficient numerical solvers. In this paper, we focus on a type of mesh-free approach based on local approximations using radial basis functions (RBFs) and multivariate polynomials for solving PDEs on manifolds.

The early application of global RBF collocation methods for solving PDEs can be traced back to Kansa's method in 1990 [37]. Following this success, various RBF collocation methods were subsequently developed for solving PDEs in Euclidean spaces and on manifolds [25, 18, 57, 27, 22, 32]. With the dawn of the 21st century, the local radial basis function-generated finite difference (RBF-FD) method was introduced by several groups as a replacement of global RBFs [63, 66, 70, 11, 69, 61, 41, 23, 62, 36]. RBF-FD, as a local method, utilizes both radial basis functions and multivariate polynomials to approximate functions locally.

The principal advantages of both global RBF and local RBF-FD methods include meshlessness, geometric flexibility, dimension independence, high-order accuracy, and implementation simplicity [16, 24, 21]. However, RBF-FD is superior to global RBF for the following reasons. (1) While powerful for small problems (e.g., the number of points $N < 10^4$), global RBF methods do not scale well to larger cases due to their expensive computational costs induced by dense matrix systems [29]. However, RBF-FD employs sparse matrices enabling efficient scaling to relatively large problems (e.g., $10^4 < N < 10^6$) while maintaining high-order accuracy of solutions. (2) While global RBFs face fundamental challenges such as ill-conditioned matrices and shape parameter tuning, RBF-FD, in contrast, achieves high accuracy by leveraging local processing to handle these issues efficiently during local interpolation and differentiation (see [27] and ref therein). (3) The global nature of RBFs makes local refinement inefficient, while RBF-FD naturally supports adaptive node placement via compact stencils without global reconfiguration [40]. (4) Solving dense and large-scale systems in global RBFs challenges parallelization, whereas RBF-FD's sparse structure readily accommodates modern parallel computing frameworks [10, 65]. (5) Global RBFs need specialized techniques

like ghost points to handle boundaries [20, 39, 12, 49], whereas RBF-FD maintains high accuracy without special techniques for boundaries as demonstrated in references [7, 6].

Other mesh-free collocation methods for solving PDEs include generalized moving least-squares (GMLS) [46, 30, 42], graph-based approaches [43, 44, 28, 34, 72], and generalized finite difference method (GFDM) [13, 64, 35, 31], just to name a few. Some characteristics of these methods are listed in [35] for practical applications. In particular, GMLS method was early found as a powerful tool for surface fitting [38] and was later extended to the solution of partial differential equations [54, 8]. The theoretical framework of GMLS was established in [69, 50] for linear functional approximations. We will review the GMLS method on manifolds, which is based on local polynomial regression in the tangent space [46, 30, 36, 42]. In our numerical experiments, we will compare our RBF-FD results with those from GMLS for approximating differential operators and solving PDEs on manifolds.

In this work, we consider a two-step generalized RBF-generated finite difference (gRBF-FD) approach for solving screened Poisson equations on manifolds without boundaries, identified by randomly sampled point cloud data. The basic idea of gRBF-FD is to regress the local function using GMLS followed by fitting the residual using an RBF interpolation. The first GMLS regression step is to capture the smooth, leading Taylor's expansion component of the target function while the second RBF interpolation step is to compensate for the expansion's remainder contained within the residual. Compared to standard RBF-FD, our gRBF-FD method employs the same interpolant form but differs in the interpolation coefficients. In addition, for problems on manifolds, we follow the tangent-plane method in [62, 71, 36], that is, local polyharmonic spline kernels and polynomials (PHS+Poly) are both defined over the projected tangent-space coordinates. It is natural for the tangent-plane PHS+Poly method to approximate the local representation of a function which is defined over the tangent space (see Sections 3.1 and 3.2). Moreover, the tangent-plane method reduces the computational cost as seen from Remark 3.

Other innovations of this work can be summarized as follows.

- To stabilize the approximation, we employ a specific weight function [45, 46, 42] in both GMLS regression and PHS interpolation steps to achieve a Laplacian matrix that is nearly diagonally dominant. Here, we say a matrix is nearly diagonally dominant if the absolute value of each diagonal entry is relatively large compared to those of the off-diagonal entries (see Definition 2). Numerical results demonstrate that the usage of the weight function ensures the stability of the Laplacian matrix.
- The choice of the stencil size K for each point is important for randomly sampled data. We propose an automatic strategy to tune the stencil size K at each base point, delivering a numerically stable approximation of the discrete Laplacian matrix with a small operator approximation error. Indeed, we aim to select an appropriate K for each point that navigates a trade-off: an overly small K leads to instability, while an unnecessarily large K increases the op-

erator approximation error.

- Theoretically, we show the approximation error bound of our gRBF-FD method in terms of the so-called stencil diameter for randomly sampled data. We further show that the approximation error rate of RBF-FD is the same as that of GMLS, which is of order $(\frac{\log N}{N})^{(l-1)/d}$. Here, N is the number of data points, l is the polynomial degree, and d is the intrinsic dimension. Numerically, we observe that by enhancing the approximation space, PHS functions in gRBF-FD help reduce both approximation and solution errors compared to using GMLS.

The paper is organized as follows. In Section 2, we review the GMLS and the tangent-plane RBF-FD approaches on manifolds. In Section 3, we present the two-step gRBF-FD method for approximating the Laplace–Beltrami operator and also point out its relation to the standard RBF-FD. In Section 4, we detail our numerical implementation, including normalization of the tangent-plane coordinates, the specific weight function used in two steps, and the auto-tuning strategy for the stencil size K . In Section 5, we report supporting numerical results. We conclude the paper with a summary and several open problems in Section 6. We include some proofs in Appendix A and Appendix B.

2 Preliminaries

We consider a d -dimensional manifold M embedded in the ambient Euclidean space \mathbb{R}^n . For a point $\mathbf{x} \in M$, the tangent space is denoted by $T_{\mathbf{x}}M$ with orthonormal basis $\{\mathbf{t}_1(\mathbf{x}), \dots, \mathbf{t}_d(\mathbf{x})\}$. Let $f : M \rightarrow \mathbb{R}$ be a smooth function. We are given a set of N distinct points $\mathbf{X}_M = \{\mathbf{x}_i\}_{i=1}^N \subset M$ and the corresponding function values $\mathbf{f} = (f(\mathbf{x}_1), \dots, f(\mathbf{x}_N))^{\top}$. We briefly review the GMLS and RBF-FD methods in Sections 2.1 and 2.2, respectively, for approximating functions on manifolds using the set of scattered data points \mathbf{X}_M and the function values \mathbf{f} .

2.1 GMLS regression of functions on manifolds

We first briefly review the basic idea of the GMLS approach, which utilizes a set of local polynomials to approximate functions on manifolds [46, 30, 35]. For an arbitrary base point $\mathbf{x}_0 \in \mathbf{X}_M \subset M$, its K -nearest neighbors in \mathbf{X}_M are denoted by $S_{\mathbf{x}_0} = \{\mathbf{x}_{0,k}\}_{k=1}^K \subset \mathbf{X}_M$. The set $S_{\mathbf{x}_0}$ is referred to as a "stencil" (see e.g., [23, 41]). By definition, $\mathbf{x}_{0,1} = \mathbf{x}_0$ is the base point. Let $\mathbb{P}_{\mathbf{x}_0}^{l,d}$ be the space of local "intrinsic" polynomials with degree up to l in d variables at the base point \mathbf{x}_0 , i.e.,

$$\mathbb{P}_{\mathbf{x}_0}^{l,d} = \text{span} \left\{ p_{\boldsymbol{\alpha}} (\boldsymbol{\theta}(\mathbf{x})) := \prod_{i=1}^d \theta_i^{\alpha_i} (\mathbf{x}) \mid 0 \leq |\boldsymbol{\alpha}| \leq l \right\}, \quad (1)$$

where $\boldsymbol{\theta}(\mathbf{x})$ is a projected coordinate system of the local tangent space $T_{\mathbf{x}_0} M$ with

$$\boldsymbol{\theta}(\mathbf{x}) = (\theta_1(\mathbf{x}), \dots, \theta_d(\mathbf{x})), \quad \theta_i(\mathbf{x}) = \mathbf{t}_i(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0), \quad i = 1, \dots, d. \quad (2)$$

Here, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{N}^d$ is the multi-index notation with $|\boldsymbol{\alpha}| = \alpha_1 + \dots + \alpha_d$. By definition, the dimension of the space $\mathbb{P}_{\mathbf{x}_0}^{l,d}$ is $m = \binom{l+d}{d}$.

Given $K > m$ and $\mathbf{f}_{\mathbf{x}_0} := (f(\mathbf{x}_{0,1}), \dots, f(\mathbf{x}_{0,K}))^\top$, we can define an operator $\mathcal{I}_p : \mathbf{f}_{\mathbf{x}_0} \in \mathbb{R}^K \rightarrow \mathcal{I}_p \mathbf{f}_{\mathbf{x}_0} \in \mathbb{P}_{\mathbf{x}_0}^{l,d}$ such that $\mathcal{I}_p \mathbf{f}_{\mathbf{x}_0}$ is the optimal solution of the following moving least-squares problem:

$$\min_{\hat{f} \in \mathbb{P}_{\mathbf{x}_0}^{l,d}} \frac{1}{2} \sum_{i,j=1}^K \lambda_{ij} (f(\mathbf{x}_{0,i}) - \hat{f}(\mathbf{x}_{0,i})) (f(\mathbf{x}_{0,j}) - \hat{f}(\mathbf{x}_{0,j})), \quad (3)$$

where $[\lambda_{ij}]_{i,j=1}^K$ is a symmetric positive-definite matrix. In most literature, $[\lambda_{ij}]_{i,j=1}^K$ is taken to be a diagonal matrix in which the diagonal element λ_{kk} is a positive weight function of the distance $\|\mathbf{x}_{0,k} - \mathbf{x}_0\|$ (see e.g., [69, 47, 46, 30, 36, 42]). GMLS is a regression technique for approximating functions by solving the above least-squares problem with a weighted ℓ^2 -norm. The solution to the least-squares problem (3) can be represented as

$$\hat{f}(\mathbf{x}) := (\mathcal{I}_p \mathbf{f}_{\mathbf{x}_0})(\mathbf{x}) = \sum_{j=1}^m b_j p_{\boldsymbol{\alpha}(j)}(\boldsymbol{\theta}(\mathbf{x})), \quad \mathbf{x} \in M, \quad (4)$$

where each b_j is the coefficient associated with the corresponding j th polynomial basis function $p_{\boldsymbol{\alpha}(j)}(\boldsymbol{\theta}(\mathbf{x}))$. Denote the concatenated coefficients $\mathbf{b} = (b_1, \dots, b_m)^\top$. Then the moving least-squares problem (3) can be formulated in matrix-vector form as

$$\min_{\mathbf{b} \in \mathbb{R}^m} \frac{1}{2} (\mathbf{f}_{\mathbf{x}_0} - \mathbf{P}\mathbf{b})^\top \boldsymbol{\Lambda} (\mathbf{f}_{\mathbf{x}_0} - \mathbf{P}\mathbf{b}), \quad (5)$$

where $\boldsymbol{\Lambda} = [\lambda_{ij}]_{i,j=1}^K$ is the $K \times K$ positive-definite matrix and \mathbf{P} is the $K \times m$ Vandermonde-type matrix with components $\mathbf{P}_{kj} = p_{\boldsymbol{\alpha}(j)}(\boldsymbol{\theta}(\mathbf{x}_{0,k}))$ for $1 \leq k \leq K, 1 \leq j \leq m$. Thus, these expansion coefficients $\mathbf{b} = (b_1, \dots, b_m)^\top$ satisfy the normal equation,

$$(\mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{P})\mathbf{b} = \mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{f}_{\mathbf{x}_0}. \quad (6)$$

The above normal equation admits a unique solution if \mathbf{P} has full column rank ($\text{rank}(\mathbf{P}) = m$).

2.2 Local RBF-FD approximation of functions on manifolds using the tangent plane method

In most previous studies on RBF-FD methods, the RBF interpolant was defined over the Euclidean radial distance in \mathbb{R}^n to approximate differential operators on manifolds (see e.g., [61, 41, 56, 3]). Recently, in order to approximate the Laplace-Beltrami operator on manifolds, authors of [62, 71, 36] first consider a PHS+Poly interpolant of the function f to the scattered data in the stencil $S_{\mathbf{x}_0}$ using the so-called tangent-plane RBF-FD method as follows:

$$(\mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0})(\mathbf{x}) := \sum_{k=1}^K a_k \phi(\|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|) + \sum_{j=1}^m b_j p_{\alpha(j)}(\boldsymbol{\theta}(\mathbf{x})), \quad \mathbf{x} \in M, \quad (7)$$

where ϕ is the radial function, $\|\cdot\|$ is the standard Euclidean norm and $p_{\alpha(j)}$ are multivariate polynomials. Here, both the PHS radial function ϕ and the polynomial $p_{\alpha(j)}$ (equation (1)) are defined over the projected tangent-plane coordinate $\boldsymbol{\theta}(\mathbf{x})$ in (2). Notice that this is similar to GMLS on manifolds, in which the polynomials are also defined over the projected coordinate $\boldsymbol{\theta}(\mathbf{x})$ as seen in (4). By definition, we have $\boldsymbol{\theta}(\mathbf{x}_{0,1}) = \boldsymbol{\theta}(\mathbf{x}_0) = \mathbf{0}$ to be center of the projected coordinate system. One significant advantage of the tangent-plane method, as noted in [36], is that it defines interpolation functions locally on the tangent space of the manifold, which can simplify the calculation of their derivatives. We will discuss this advantage later in Remark 3.

Common choices of the radial function ϕ include Polyharmonic spline (PHS) function [23, 36], Gaussian function [19], inverse quadratic function [32], inverse multiquadric function [27] and Matérn class function [27]. For most bell-shaped positive-definite RBFs, including the Gaussian and Matérn classes, they contain a shape parameter that controls the flatness of the RBF. In general, the shape parameter needs appropriate tuning in a certain range; specifically, a small value achieves better accuracy but also results in an ill-conditioned interpolation matrix. When the shape parameter is properly chosen, the local RBF-FD can achieve the convergence results (see e.g., [61, 41, 56, 3]).

On the other hand, the *U*-shaped PHS function is attractive due to its absence of the shape parameter. For RBF-FD methods, the PHS function has been used together with polynomials in many studies (see e.g., [23, 7, 6, 36]). In this article, we focus on the following shape-parameter-free PHS function,

$$\text{Polyharmonic spline (PHS)} : \quad \phi(r) = r^{2\kappa+1}, \quad (8)$$

where the variable r is the radial distance and the PHS parameter $\kappa \in \mathbb{N}$ controls the smoothness of ϕ . While some theories suggest using $\phi(r) = r^{2\kappa+1}$ in odd dimension, and $\phi(r) = r^{2\kappa} \log r$ in even dimension, that distinction may not be necessary [23]. In the RBF-FD interpolant (7), the radial distance r is taken to be the norm in \mathbb{R}^d , i.e., $r = \|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|$.

Given data $f(\mathbf{x}_{0,k})$ at node $\mathbf{x}_{0,k}$ ($k = 1, \dots, K$), the expansion coefficients are obtained by enforcing K interpolation conditions together with m additional moment conditions ensuring uniqueness:

$$\begin{aligned} (\mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0})(\mathbf{x}_{0,k}) &= f(\mathbf{x}_{0,k}), \quad \text{for } k = 1, \dots, K, \\ \sum_{k=1}^K a_k p_{\alpha(j)}(\boldsymbol{\theta}(\mathbf{x}_{0,k})) &= 0, \quad \text{for } j = 1, \dots, m. \end{aligned} \quad (9)$$

Denote the coefficients as $\mathbf{a} = [a_1, a_2, \dots, a_K]^\top$ and $\mathbf{b} = [b_1, b_2, \dots, b_m]^\top$. In matrix form, the $K + m$ conditions above can be written as

$$\begin{bmatrix} \boldsymbol{\Phi} & \mathbf{P} \\ \mathbf{P}^\top & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{\mathbf{x}_0} \\ \mathbf{0} \end{bmatrix}, \quad (10)$$

where $\mathbf{P} \in \mathbb{R}^{K \times m}$ is defined in (6) and the symmetric $\boldsymbol{\Phi} \in \mathbb{R}^{K \times K}$ is defined as

$$\boldsymbol{\Phi} = [\phi(\|\boldsymbol{\theta}(\mathbf{x}_{0,i}) - \boldsymbol{\theta}(\mathbf{x}_{0,j})\|)]_{i,j=1}^K. \quad (11)$$

For the well-posed linear system (10), its solution can be obtained using the Schur complement of $\boldsymbol{\Phi}$ by

$$\begin{aligned} \mathbf{a} &= \boldsymbol{\Phi}^{-1} \left(\mathbf{I} - \mathbf{P} (\mathbf{P}^\top \boldsymbol{\Phi}^{-1} \mathbf{P})^{-1} \mathbf{P}^\top \boldsymbol{\Phi}^{-1} \right) \mathbf{f}_{\mathbf{x}_0} \\ \mathbf{b} &= (\mathbf{P}^\top \boldsymbol{\Phi}^{-1} \mathbf{P})^{-1} \mathbf{P}^\top \boldsymbol{\Phi}^{-1} \mathbf{f}_{\mathbf{x}_0}. \end{aligned} \quad (12)$$

For the PHS $\phi(r) = r^{2\kappa+1}$, the chosen parameter κ should be less than or equal to the degree of the polynomial l as defined in (1). Under this constraint on κ , it can be shown that $\boldsymbol{\Phi}$ is conditionally positive definite on the subspace satisfying the m moment conditions in (9) [33, 69, 36]. If the stencil points satisfy $\text{rank}(\mathbf{P}) = m$, then the system (10) is non-singular and the PHS+Poly interpolant is well-posed.

3 Two-step Generalized RBF-FD method

Motivated by the tangent-plane RBF-FD method in [36], we develop the generalized RBF-FD (gRBF-FD) method for solving PDEs on manifolds, identified by randomly sampled data. For gRBF-FD method, the local approximation space matches that of the tangent-plane RBF-FD but the distinction lies in how the expansion coefficients are calculated. That is, we still consider using the PHS+Poly interpolant (7), in which both the PHS function and the polynomials are defined over the projected coordinate system. It is natural, as the geometry on the manifold M can be pulled back to define these functions on the projected tangent space, which is spanned by the local d -dimensional coordinates $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$. We next introduce how we approximate the expansion coefficients $\{a_k\}_{k=1}^K$ and $\{b_j\}_{j=1}^m$ in (7).

3.1 Approximation of the expansion coefficients in gRBF-FD method

The gRBF-FD method employs a two-step approximation for the target function f . For the first step, we apply the GMLS regression technique to approximate the target function f . Then the coefficients \mathbf{b} can be solved from the normal equation in (6),

$$\mathbf{b} = (\mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{P})^{-1} \mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{f}_{\mathbf{x}_0}. \quad (13)$$

This provides a first-step approximation to the function f and then the residual of the GMLS approximation becomes

$$\mathbf{s}_{\mathbf{x}_0} := \mathbf{f}_{\mathbf{x}_0} - (\mathcal{I}_p \mathbf{f}_{\mathbf{x}_0})|_{S_{\mathbf{x}_0}} = \mathbf{f}_{\mathbf{x}_0} - \mathbf{P}\mathbf{b},$$

where the operator $(\mathcal{I}_p \mathbf{f}_{\mathbf{x}_0})(\mathbf{x}) = \sum_{j=1}^m b_j p_{\boldsymbol{\alpha}(j)}(\boldsymbol{\theta}(\mathbf{x}))$ is defined in (4).

For the second step, we fit the above residual $\mathbf{s}_{\mathbf{x}_0}$ using the PHS interpolant, that is,

$$(\mathcal{I}_\phi \mathbf{s}_{\mathbf{x}_0})(\mathbf{x}) = \sum_{k=1}^K a_k \phi(\|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|). \quad (14)$$

The coefficients $\mathbf{a} = [a_1, a_2, \dots, a_K]^\top$ are determined by the interpolation condition,

$$\boldsymbol{\Phi} \mathbf{a} = \mathbf{s}_{\mathbf{x}_0} = \mathbf{f}_{\mathbf{x}_0} - \mathbf{P}\mathbf{b} = \mathbf{f}_{\mathbf{x}_0} - \mathbf{P}(\mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{P})^{-1} \mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{f}_{\mathbf{x}_0}, \quad (15)$$

where $\boldsymbol{\Phi}$ is given by (11). Then we can compute the coefficients \mathbf{a} as

$$\mathbf{a} = \boldsymbol{\Phi}^{-1}(\mathbf{s}_{\mathbf{x}_0} - \mathbf{P}\mathbf{b}) = \boldsymbol{\Phi}^{-1} \left(\mathbf{I} - \mathbf{P}(\mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{P})^{-1} \mathbf{P}^\top \boldsymbol{\Lambda} \right) \mathbf{f}_{\mathbf{x}_0}. \quad (16)$$

The basic idea of gRBF-FD is to regress the function using GMLS followed by fitting the residual $\mathbf{s}_{\mathbf{x}_0}$ using the PHS interpolant. Intuitively, GMLS captures the smooth, leading component of the Taylor expansion of f , while PHS compensates for the expansion's remainder contained within the residual. As a result, we arrive at a PHS+Poly interpolant for the gRBF-FD method,

$$\begin{aligned} (\mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0})(\mathbf{x}) &= (\mathcal{I}_\phi \mathbf{s}_{\mathbf{x}_0})(\mathbf{x}) + (\mathcal{I}_p \mathbf{f}_{\mathbf{x}_0})(\mathbf{x}) \\ &:= \sum_{k=1}^K a_k \phi(\|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|) + \sum_{j=1}^m b_j p_{\boldsymbol{\alpha}(j)}(\boldsymbol{\theta}(\mathbf{x})), \quad \mathbf{x} \in M, \end{aligned} \quad (17)$$

where the coefficients \mathbf{a} and \mathbf{b} are determined by (16) and (13), respectively. We note that this interpolant formulation is exactly the same as the one in (7) for RBF-FD while the only difference comes from the evaluation of the coefficients \mathbf{a} and \mathbf{b} .

Remark 1 The relationship between RBF-FD and gRBF-FD can be revealed by a direct comparison of the coefficients in (12) with those in (16) and (13). Specifically, the coefficients \mathbf{a} and \mathbf{b} are identical for both methods if the weight matrix in gRBF-FD is chosen as $\Lambda = \Phi^{-1}$. We therefore refer to this more general formulation as the generalized RBF-FD method. Moreover, for RBF-FD, those moment conditions in (9) admit a natural interpretation. By applying the RBF interpolant to the residual, we arrive at the expression, $\Phi\mathbf{a} = \mathbf{f}_{\mathbf{x}_0} - \mathbf{P}(\mathbf{P}^\top \Phi^{-1} \mathbf{P})^{-1} \mathbf{P}^\top \Phi^{-1} \mathbf{f}_{\mathbf{x}_0}$. This follows immediately from equation (15) by selecting the weight $\Lambda = \Phi^{-1}$. Left multiplying both sides of this equation by $\mathbf{P}^\top \Phi^{-1}$ causes the right-hand-side term to vanish, yielding the moment conditions: $\mathbf{P}^\top \Phi^{-1} \Phi \mathbf{a} = \mathbf{P}^\top \mathbf{a} = \mathbf{0}$.

Remark 2 There are many choices of the weight $\Lambda = \{\lambda_{ij}\}_{i,j=1}^K$ in the moving least-squares (5) as shown in [69, 46, 30]. As reported in previous works [45, 46, 35, 42], the diagonal weight function,

$$\Lambda_K = \{\lambda_{ij}\}_{i,j=1}^K, \quad \text{with diagonal entries } \lambda_{kk} = \begin{cases} 1, & \text{if } k = 1 \\ 1/K, & \text{if } k = 2, \dots, K \end{cases}, \quad (18)$$

can numerically provide a stable approximation to the Laplace-Beltrami operator across various data sets. In the following, this weight is referred to as the diagonal $1/K$ weight function. For the gRBF-FD method, we will focus on this special weight function and compare its performance with other weight functions (see Figs. 1 and 2 in the following Section 4.3 for the numerical performance among different choices of Λ).

3.2 Monge parametrization

To facilitate the computation of local geometry, we employ a Monge parametrization of the local manifold near the base point \mathbf{x}_0 [51, 58, 46, 36, 30]. A local coordinate chart for the manifold near the base point is defined using the embedding map \mathbf{q} ,

$$\mathbf{x} = \mathbf{q}(\theta_1, \dots, \theta_d; \mathbf{x}_0) = \mathbf{x}_0 + \sum_{i=1}^d \theta_i \mathbf{t}_i(\mathbf{x}_0) + \sum_{s=1}^{n-d} q_s(\theta_1, \dots, \theta_d) \mathbf{n}_s(\mathbf{x}_0), \quad (19)$$

where $(\theta_1, \dots, \theta_d)$ are the local coordinates as defined in (2), $q_s(\theta_1, \dots, \theta_d) = \mathbf{n}_s(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$ ($s = 1, \dots, n-d$), and $\{\mathbf{n}_s\}_{s=1}^{n-d}$ are the $n-d$ orthonormal vectors that are orthogonal to the tangent space $T_{\mathbf{x}_0} M$. This is known as the Monge parametrization of manifold. This representation allows us to express differential operators in terms of the local coordinates $(\theta_1, \dots, \theta_d)$, which is essential for computing the Laplace-Beltrami operator in our mesh-free framework.

Indeed, all above multivariate polynomials and polyharmonic spline functions in RBF-FD (7) and gRBF-FD (17) are defined over the tangent plane

and have been formulated as their local representations in local Monge coordinates $(\theta_1, \dots, \theta_d)$. That is, the local representation of $(\mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0})$ is $(\mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0}) \circ \boldsymbol{\theta}^{-1} : \boldsymbol{\theta}(U_0) \rightarrow \mathbb{R}$ for some open set $U_0 \subset M$, and the interpolant $\sum_{k=1}^K a_k \phi(\cdot) + \sum_{j=1}^m b_j p_{\alpha(j)}(\cdot)$ is approximating the local representation of the target function f , which is defined over $\boldsymbol{\theta}(U_0)$. This is natural in differential geometry because differentiating a function on a manifold requires pulling it back to a local coordinate chart and then computing the derivative with respect to those local coordinates.

Then the tangent vector bases for $\mathbf{x} \in S_{\mathbf{x}_0}$ in the Monge coordinate system can be calculated as

$$\frac{\partial \mathbf{q}}{\partial \theta_k} = \mathbf{t}_k(\mathbf{x}_0) + \sum_{s=1}^{n-d} \frac{\partial q_s}{\partial \theta_k} \mathbf{n}_s(\mathbf{x}_0), \quad \text{for } k = 1, \dots, d. \quad (20)$$

Each component g_{ij} of the local Riemannian metric $g = (g_{ij})_{i,j=1}^d$ for $\mathbf{x} \in S_{\mathbf{x}_0}$ becomes

$$g_{ij} = \left\langle \frac{\partial \mathbf{q}}{\partial \theta_i}, \frac{\partial \mathbf{q}}{\partial \theta_j} \right\rangle = \delta_{ij} + \sum_{s=1}^{n-d} \frac{\partial q_s}{\partial \theta_i} \frac{\partial q_s}{\partial \theta_j}, \quad (21)$$

where δ_{ij} is the Kronecker delta function. Let $(g^{ij})_{i,j=1}^d$ be the inverse of the metric tensor. Then we have the following properties in a Monge coordinate system centered at \mathbf{x}_0 :

Proposition 1 *Let $\{U_0; \theta_1, \dots, \theta_d\}$ be a Monge coordinate system centered at \mathbf{x}_0 where $U_0 \subset M$ is an open neighborhood of \mathbf{x}_0 and M is a d -dimensional manifold with Riemannian metric g . Then*

- (1) *For all $1 \leq i, j \leq d$, $g_{ij}(\mathbf{x}_0) = g^{ij}(\mathbf{x}_0) = \delta_{ij}$.*
- (2) *For all $1 \leq i, j, k \leq d$, $\Gamma_{ij}^k(\mathbf{x}_0) = 0$.*
- (3) *For all $1 \leq i, j, k \leq d$, $\partial_k g_{ij}(\mathbf{x}_0) = 0$.*

Proof We notice that the embedding map (19) restricted on the base point should be $\mathbf{q}|_{\mathbf{x}_0} = \mathbf{x}_0$, which implies that all coordinates vanish at the base point, that is, $\theta_1|_{\mathbf{x}_0} = \dots = \theta_d|_{\mathbf{x}_0} = 0$ and $q_1|_{\mathbf{x}_0} = \dots = q_{n-d}|_{\mathbf{x}_0} = 0$. Also notice that the tangent vector bases (20) should live in the tangent space at the base point, which implies that the first-order derivative $\frac{\partial q_s}{\partial \theta_k}|_{\mathbf{x}_0} = 0$ for all $s = 1, \dots, n-d$ and $k = 1, \dots, d$. Then we obtain the property for the metric g_{ij} :

$$g_{ij}(\mathbf{x}_0) = \left(\delta_{ij} + \sum_{s=1}^{n-d} \frac{\partial q_s}{\partial \theta_i} \frac{\partial q_s}{\partial \theta_j} \right) \Big|_{\mathbf{x}_0} = \delta_{ij},$$

as well as for the inverse $g^{ij}(\mathbf{x}_0) = \delta_{ij}$. The derivative can be calculated as

$$\partial_k g_{ij}(\mathbf{x}_0) = \left(\sum_{s=1}^{n-d} \frac{\partial q_s}{\partial \theta_i} \frac{\partial q_s}{\partial \theta_k} + \frac{\partial q_s}{\partial \theta_i} \frac{\partial q_s}{\partial \theta_j \partial \theta_k} \right) \Big|_{\mathbf{x}_0} = 0.$$

Thus, by definition, $\Gamma_{ij}^k = \frac{1}{2} g^{km} \left(\frac{\partial g_{jm}}{\partial \theta_i} + \frac{\partial g_{im}}{\partial \theta_j} - \frac{\partial g_{ij}}{\partial \theta_m} \right)$, all the Christoffel symbols vanish at the base point.

In the next section, we will use these properties to simplify the calculation of the Laplace-Beltrami operator.

3.3 Approximation of the Laplace-Beltrami operator

Our goal is to approximate the Laplace-Beltrami operator Δ_M acting on a function f at the base point \mathbf{x}_0 by a linear combination of the function values in the stencil $S_{\mathbf{x}_0}$, i.e., find the coefficients $\{w_k\}_{k=1}^K$ such that

$$\Delta_M f(\mathbf{x}_0) \approx \sum_{k=1}^K w_k f(\mathbf{x}_{0,k}). \quad (22)$$

Arranging the coefficients at each point $\mathbf{x}_0 \in \mathbf{X}_M \subset M$ into each row of a sparse $N \times N$ matrix $\mathbf{L}_{\mathbf{X}_M}$, we can approximate $\Delta_M f$ at all points by $\mathbf{L}_{\mathbf{X}_M} \mathbf{f}$. We now provide the calculation details for the coefficients $\{w_k\}_{k=1}^K$.

The Laplace-Beltrami operator can be formulated in the local Monge coordinate system as

$$\Delta_M f = \sum_{i,j=1}^d \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial \theta_i} \left(\sqrt{|g|} g^{ij} \frac{\partial f}{\partial \theta_j} \right) = \sum_{i,j=1}^d g^{ij} \left(\frac{\partial^2 f}{\partial \theta_i \partial \theta_j} - \Gamma_{ij}^k \frac{\partial f}{\partial \theta_k} \right),$$

where $|g| = \det(g)$ is the determinant of the Riemannian metric tensor and d is the intrinsic dimension of manifold. The two formulas are equivalent for the Laplace-Beltrami operator. Using Proposition 1, the Laplace-Beltrami operator at the base point can be simplified as

$$\Delta_M f(\mathbf{x}_0) = \sum_{i=1}^d \frac{\partial^2}{\partial \theta_i^2} f(\mathbf{x}_0) := \Delta_{\boldsymbol{\theta}} f(\mathbf{x}_0). \quad (23)$$

Here, we use $\Delta_{\boldsymbol{\theta}}$ to denote the Laplace operator in the local Monge coordinate chart, $\{U_0; \boldsymbol{\theta} = (\theta_1, \dots, \theta_d)\}$, where $\boldsymbol{\theta} : U_0 \subset M \rightarrow V_0 \subset \mathbb{R}^d$ is a diffeomorphism onto an open subset of \mathbb{R}^d . The local representation of f becomes $f \circ \boldsymbol{\theta}^{-1} : \boldsymbol{\theta}(U_0) \rightarrow \mathbb{R}$. Indeed, for gRBF-FD, the linear combination of multivariate polynomials and PHS functions in the interpolant (17) approximates the local representation $f \circ \boldsymbol{\theta}^{-1}$ in the local Monge coordinate chart (on the tangent plane).

Then the Laplace-Beltrami operator at the base point can be approximated using the gRBF-FD method,

$$\begin{aligned} \Delta_M f(\mathbf{x}_0) &\approx \Delta_{\boldsymbol{\theta}}(\mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0})(\mathbf{x}_0) \\ &= \left. \left(\sum_{k=1}^K a_k \Delta_{\boldsymbol{\theta}} \phi(\|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|) + \sum_{s=1}^m b_s \Delta_{\boldsymbol{\theta}} p_{\alpha(s)}(\boldsymbol{\theta}(\mathbf{x})) \right) \right|_{\mathbf{x}=\mathbf{x}_0}, \end{aligned} \quad (24)$$

where (17) has been used. Here, the first term in (24) can be calculated as

$$\begin{aligned} \Delta_{\theta}\phi(\|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|) \Big|_{\mathbf{x}=\mathbf{x}_0} &= \sum_{i=1}^d \frac{\partial}{\partial \theta_i} \left[(2\kappa+1) \|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|^{2\kappa-1} (\theta_i(\mathbf{x}) - \theta_i(\mathbf{x}_{0,k})) \right] \\ &= (4\kappa^2 + 2d\kappa + d - 1) \|\boldsymbol{\theta}(\mathbf{x}_0) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|^{2\kappa-1}, \end{aligned} \quad (25)$$

where the PHS $\phi(\|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|) = \|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|^{2\kappa+1}$ has been used. The second term in (24) can be calculated as

$$\Delta_{\theta} p_{\alpha(s)}(\boldsymbol{\theta}(\mathbf{x})) \Big|_{\mathbf{x}=\mathbf{x}_0} = \begin{cases} 2, & \alpha(s) \in \mathbf{E}, \\ 0, & \alpha(s) \notin \mathbf{E}, \end{cases} \quad (26)$$

where the index set \mathbf{E} is defined as

$$\mathbf{E} = \left\{ (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d \mid \text{each } \alpha_i \in \{0, 2\} \text{ } (i = 1, \dots, d) \text{ and } |\alpha| = \sum_{j=1}^d \alpha_j = 2 \right\}.$$

In fact, for the index in \mathbf{E} , only one entry $\alpha_i = 2$ for some $i \in \{1, \dots, d\}$ and all the others are zero. Note that \mathbf{E} contains d number of indices. Substituting (25) and (26) into (24), we arrive at

$$\begin{aligned} \Delta_M f(\mathbf{x}_0) &\approx \sum_{k=1}^K a_k (4\kappa^2 + 2d\kappa + d - 1) \|\boldsymbol{\theta}(\mathbf{x}_0) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|^{2\kappa-1} + \sum_{\{s \mid \alpha(s) \in \mathbf{E}\}} 2b_s \\ &:= (\Delta_{\theta}\phi_{\mathbf{x}_0})\mathbf{a} + (\Delta_{\theta}p_{\mathbf{x}_0})\mathbf{b}, \end{aligned} \quad (27)$$

where $\Delta_{\theta}\phi_{\mathbf{x}_0} \in \mathbb{R}^{1 \times K}$ with the k th entry given by (25) and $\Delta_{\theta}p_{\mathbf{x}_0} \in \mathbb{R}^{1 \times m}$ with d number of 2 and $m-d$ number of 0. Substituting \mathbf{a} (equation (16)) and \mathbf{b} (equation (13)) into above (27), we obtain the coefficients in (22):

$$\begin{aligned} \Delta_M f(\mathbf{x}_0) &\approx \sum_{k=1}^K w_k f(\mathbf{x}_{0,k}) \\ &= \left[(\Delta_{\theta}\phi_{\mathbf{x}_0}) \boldsymbol{\Phi}^{-1} (\mathbf{I} - \mathbf{P}(\mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{P})^{-1} \mathbf{P}^\top \boldsymbol{\Lambda}) + (\Delta_{\theta}p_{\mathbf{x}_0})(\mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{P})^{-1} \mathbf{P}^\top \boldsymbol{\Lambda} \right] \mathbf{f}_{\mathbf{x}_0}. \end{aligned} \quad (28)$$

Remark 3 Here we do not employ PHS functions defined in Euclidean coordinates due to additional computational cost required to evaluate their derivatives and the Laplace-Beltrami operator. In particular, let $r_{\mathbf{x},k}(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_{0,k}\|$ be the radial distance function defined for $\mathbf{x} \in \mathbb{R}^n$ and then we compute the first-order derivative of the PHS function as

$$\frac{\partial}{\partial \theta_j} \phi(\|\mathbf{x} - \mathbf{x}_{0,k}\|) = \frac{\partial}{\partial \theta_j} (r_{\mathbf{x},k}^2(\mathbf{x}))^{\frac{2\kappa+1}{2}} = \frac{2\kappa+1}{2} (r_{\mathbf{x},k}(\mathbf{x}))^{2\kappa-1} \frac{\partial r_{\mathbf{x},k}^2(\mathbf{x})}{\partial \theta_j}.$$

Using the embedding map (19), we can calculate the derivative of the square of the radial distance by

$$\begin{aligned} \frac{\partial r_{\mathbf{x},k}^2(\mathbf{x})}{\partial \theta_j} &= \frac{\partial}{\partial \theta_j} \left[\sum_{i=1}^d (\theta_i - \theta_i(\mathbf{x}_{0,k}))^2 + \sum_{s=1}^{n-d} (q_s - q_s(\mathbf{x}_{0,k}))^2 \right] \\ &= 2(\theta_j - \theta_j(\mathbf{x}_{0,k})) + 2 \sum_{s=1}^{n-d} (q_s - q_s(\mathbf{x}_{0,k})) \frac{\partial q_s}{\partial \theta_j}, \end{aligned}$$

where we see that the calculation involves many derivatives of normal components q_s with respect to tangent components θ_j especially when the ambient dimension n is large. In contrast, as seen from the first line of (25) for the PHS function defined over the tangent plane, the first-order derivative does not contain any term of $\frac{\partial q_s}{\partial \theta_j}$ for $s = 1, \dots, n-d$ and $j = 1, \dots, d$, which simplifies our computation.

4 Application to solving screened Poisson problems

4.1 Screened Poisson problems on closed manifolds

In this section, we will show an application of the proposed two-step gRBF-FD discretization method to solve screened Poisson problems on closed manifolds. For a closed manifold M , we consider the following screened Poisson problem,

$$(a - \Delta_M) f = h, \quad \mathbf{x} \in M, \quad (29)$$

where $a > 0$ and h are defined such that the problem is well-posed. Here, we set the parameter $a = 1$ for simplicity. Numerically, we will approximate the solution to the PDE problem (29) pointwisely on the scattered dataset \mathbf{X}_M , solving an $N \times N$ linear algebra system

$$\mathbf{L}_{\mathbf{X}_M, \mathbf{I}} \mathbf{F} := (\mathbf{I} - \mathbf{L}_{\mathbf{X}_M}) \mathbf{F} = \mathbf{h}, \quad (30)$$

for $\mathbf{F} \in \mathbb{R}^{N \times 1}$ to approximate the true solution $\mathbf{f} := (f(\mathbf{x}_1), \dots, f(\mathbf{x}_N))^\top \in \mathbb{R}^{N \times 1}$. Here, $\mathbf{h} := (h(\mathbf{x}_1), \dots, h(\mathbf{x}_N))^\top \in \mathbb{R}^{N \times 1}$ and $\mathbf{L}_{\mathbf{X}_M} \in \mathbb{R}^{N \times N}$ is a sparse Laplacian matrix constructed from a discretization scheme, such as GMLS and gRBF-FD. To verify the convergence, we define the forward error (**FE**) for the operator approximation and the inverse error (**IE**) for the solution approximation as

$$\mathbf{FE} = \max_{1 \leq i \leq N} |\Delta_M f(\mathbf{x}_i) - (\mathbf{L}_{\mathbf{X}_M} \mathbf{f})_i|, \quad \mathbf{IE} = \max_{1 \leq i \leq N} |\mathbf{F}_i - f(\mathbf{x}_i)|,$$

where $\mathbf{L}_{\mathbf{X}_M} \mathbf{f} \in \mathbb{R}^{N \times 1}$, $(\mathbf{L}_{\mathbf{X}_M} \mathbf{f})_i$ is the i th component to approximate the value of $\Delta_M f$ at \mathbf{x}_i , and \mathbf{F}_i is the i th component of the numerical solution \mathbf{F} .

In order to numerically verify the validity of operator approximation, stability and convergence, we test on a one-dimensional full ellipse example in this section.

Example 1 Consider a 1D full ellipse in \mathbb{R}^2 parametrized by

$$\mathbf{x} := (x^1, x^2) = (\cos \theta, 2 \sin \theta) \in M \subset \mathbb{R}^2, \quad (31)$$

defined with the Riemannian metric $g = \sin^2 \theta + 4 \cos^2 \theta$ for $0 \leq \theta < 2\pi$. The true solution is set to be $f = \sin \theta \cos \theta$ and the right-hand-side $h := (1 - \Delta_M) f$ can be calculated. Then, we approximate the numerical solution \mathbf{F} in (30) for the PDE problem in (29), subjected to the manufactured h .

Numerically, the convergence is examined on the ellipse using the number $N = [400, 800, 1600, 3200, 6400]$ of dataset. Here, two types of dataset are considered for convergence examination:

- *Well-sampled data.* The grid points are well-ordered and all consecutive points have equal intrinsic distance in θ space (see Fig. 1(a)).
- *Randomly sampled data.* The data $\{\theta_i\}_{i=1}^N$ are randomly generated from an independent and identical distribution with density $p(\theta) = \theta/(4\pi^2) + 1/(4\pi)$ on $[0, 2\pi)$ in the intrinsic coordinate and then mapped to the point cloud $\{\mathbf{x}_i\}_{i=1}^N$ via the parametrization (31) (see Fig. 1(d)).

4.2 Normalization of Monge coordinates

In our numerical implementation, several technical details need to be addressed to ensure the stability and convergence using the gRBF-FD approach. First, we need to normalize the Monge coordinates since the PHS function is sensitive to changes in the magnitude of the input radial distance.

Definition 1 For a stencil $S_{\mathbf{x}_0}$ centered at the base point \mathbf{x}_0 , its diameter and radius are defined as

$$\begin{aligned} D_{K,\max}(\mathbf{x}_0) &:= \max_{1 \leq i, j \leq K} \|\boldsymbol{\theta}(\mathbf{x}_{0,i}) - \boldsymbol{\theta}(\mathbf{x}_{0,j})\|, \\ R_{K,\max}(\mathbf{x}_0) &:= \max_{1 \leq k \leq K} \|\boldsymbol{\theta}(\mathbf{x}_{0,k}) - \boldsymbol{\theta}(\mathbf{x}_0)\|, \end{aligned}$$

respectively, over the projected tangent space of \mathbf{x}_0 .

To compute multivariate polynomials and PHS functions, we normalize the Monge coordinates $\boldsymbol{\theta}(\mathbf{x})$ and the radial distances $r_{\boldsymbol{\theta},k}(\mathbf{x}) := \|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|$ in the tangent plane by the stencil diameter as follows,

$$\begin{aligned} \tilde{\boldsymbol{\theta}}(\mathbf{x}) &= (\tilde{\theta}_1(\mathbf{x}), \dots, \tilde{\theta}_d(\mathbf{x})), \\ \tilde{r}_{\boldsymbol{\theta},k}(\mathbf{x}) &= \left\| \tilde{\boldsymbol{\theta}}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}(\mathbf{x}_{0,k}) \right\|, \\ \tilde{\theta}_i(\mathbf{x}) &= \frac{\theta_i(\mathbf{x})}{D_{K,\max}} = \frac{\mathbf{t}_i(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)}{D_{K,\max}}, \end{aligned} \tag{32}$$

where the normalized quantities are scaled to lie in the interval $[0, 1]$, that is, $0 \leq \|\tilde{\boldsymbol{\theta}}(\mathbf{x})\| = \|\boldsymbol{\theta}(\mathbf{x})/D_{K,\max}\| \leq 1$ and $0 \leq \tilde{r}_{\boldsymbol{\theta},k}(\mathbf{x}) = r_{\boldsymbol{\theta},k}(\mathbf{x})/D_{K,\max} \leq 1$. Here we use the shorthand $D_{K,\max}$ for $D_{K,\max}(\mathbf{x}_0)$. Then the coefficients in (28) for approximating the Laplace-Beltrami operator become

$$\begin{aligned} \Delta_M f(\mathbf{x}_0) &\approx \frac{1}{D_{K,\max}^2} \left[(\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\phi}_{\mathbf{x}_0}) \tilde{\boldsymbol{\Phi}}^{-1} (\mathbf{I} - \tilde{\mathbf{P}} (\tilde{\mathbf{P}}^\top \boldsymbol{\Lambda} \tilde{\mathbf{P}})^{-1} \tilde{\mathbf{P}}^\top \boldsymbol{\Lambda}) \right. \\ &\quad \left. + (\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\mathbf{p}}_{\mathbf{x}_0}) (\tilde{\mathbf{P}}^\top \boldsymbol{\Lambda} \tilde{\mathbf{P}})^{-1} \tilde{\mathbf{P}}^\top \boldsymbol{\Lambda} \right] \mathbf{f}_{\mathbf{x}_0}, \end{aligned} \tag{33}$$

where $\tilde{\mathbf{P}} \in \mathbb{R}^{K \times m}$, $\tilde{\boldsymbol{\Phi}} \in \mathbb{R}^{K \times K}$, $\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\mathbf{p}}_{\mathbf{x}_0} \in \mathbb{R}^{1 \times m}$ and $\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\phi}_{\mathbf{x}_0} \in \mathbb{R}^{1 \times K}$ are all defined over the normalized Monge coordinates with components

$$\begin{aligned}\tilde{\mathbf{P}}_{kj} &= p_{\alpha(j)}(\tilde{\boldsymbol{\theta}}(\mathbf{x}_{0,k})), \quad [\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\mathbf{p}}_{\mathbf{x}_0}]_j = \begin{cases} 2, & \alpha(j) \in \mathbf{E} \\ 0, & \alpha(j) \notin \mathbf{E} \end{cases}, \\ \tilde{\boldsymbol{\Phi}}_{ks} &= \phi\left(\|\tilde{\boldsymbol{\theta}}(\mathbf{x}_{0,k}) - \tilde{\boldsymbol{\theta}}(\mathbf{x}_{0,s})\|\right), \\ [\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\phi}_{\mathbf{x}_0}]_k &= (4\kappa^2 + 2d\kappa + d - 1) \left\| \tilde{\boldsymbol{\theta}}(\mathbf{x}_0) - \tilde{\boldsymbol{\theta}}(\mathbf{x}_{0,k}) \right\|^{2\kappa-1},\end{aligned}\tag{34}$$

for $j = 1, \dots, m$ and $k, s = 1, \dots, K$.

4.3 Weighted ridge regression for the inverse of the PHS matrix

Since the PHS matrix $\tilde{\boldsymbol{\Phi}}$ might be singular, we use a weighted ridge regression or Tikhonov regularization for gRBF-FD to compute the inverse $\tilde{\boldsymbol{\Phi}}^{-1}$ in (33),

$$\min_{\tilde{\mathbf{a}} \in \mathbb{R}^K} \frac{1}{2} \left(\mathbf{s}_{\mathbf{x}_0} - \tilde{\boldsymbol{\Phi}} \tilde{\mathbf{a}} \right)^T \boldsymbol{\Lambda}_K \left(\mathbf{s}_{\mathbf{x}_0} - \tilde{\boldsymbol{\Phi}} \tilde{\mathbf{a}} \right) + \frac{1}{2} \delta^2 \|\tilde{\mathbf{a}}\|^2,\tag{35}$$

where $\boldsymbol{\Lambda}_K$ is the diagonal $1/K$ weight function defined in (18) and δ is the regularization parameter. For the gRBF-FD method, the inverse $\tilde{\boldsymbol{\Phi}}^{-1}$ in (33) can be found by solving the normal equation to be,

$$\tilde{\boldsymbol{\Phi}}^{-1} := \tilde{\boldsymbol{\Phi}}_{\boldsymbol{\Lambda}_K}^\dagger = \left(\tilde{\boldsymbol{\Phi}}^T \boldsymbol{\Lambda}_K \tilde{\boldsymbol{\Phi}} + \delta^2 \mathbf{I} \right)^{-1} \tilde{\boldsymbol{\Phi}}^T \boldsymbol{\Lambda}_K,\tag{36}$$

Here, we set $\delta = 10^{-6}$ for a stable and unique inverse of $\tilde{\boldsymbol{\Phi}}$. Notably, we emphasize that the choice of $\boldsymbol{\Lambda}_K$ is very important for the stability of the discretized Laplacian matrix as can be seen from Figs. 1 and 2. The weight $\boldsymbol{\Lambda}_K$ in this optimization problem (35) stabilizes the Laplacian analogous to its role in the GMLS problem in (5) (see e.g., [46, 35] for the numerical performance).

For comparison, we choose three GMLS methods and two types of RBF-FD methods. The Laplace-Beltrami operator using GMLS methods can be calculated by

$$\Delta_M f(\mathbf{x}_0) \approx \frac{1}{D_{K,\max}^2} \left((\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\mathbf{p}}_{\mathbf{x}_0}) \left(\tilde{\mathbf{P}}^T \boldsymbol{\Lambda} \tilde{\mathbf{P}} \right)^{-1} \tilde{\mathbf{P}}^T \boldsymbol{\Lambda} \right) \mathbf{f}_{\mathbf{x}_0},\tag{37}$$

which corresponds to the GMLS part of equation (33). One may refer to [46, 30, 36] for more calculation details. GMLS methods can be applied with three different weight functions:

- GMLS-1/ K . The first weight function is the diagonal $1/K$ as discussed in equation (18) in Remark 2.
- GMLS-SW. The second is a diagonal smooth weight function with diagonal entries $\lambda_{kk} = (1 - \frac{r_k}{R})^2$, where r_k is defined as $r_k = \|\boldsymbol{\theta}(\mathbf{x}_{0,k}) - \boldsymbol{\theta}(\mathbf{x}_0)\|$ and $R = 1.5R_{K,\max}$, where $R_{K,\max}$ is the stencil radius.

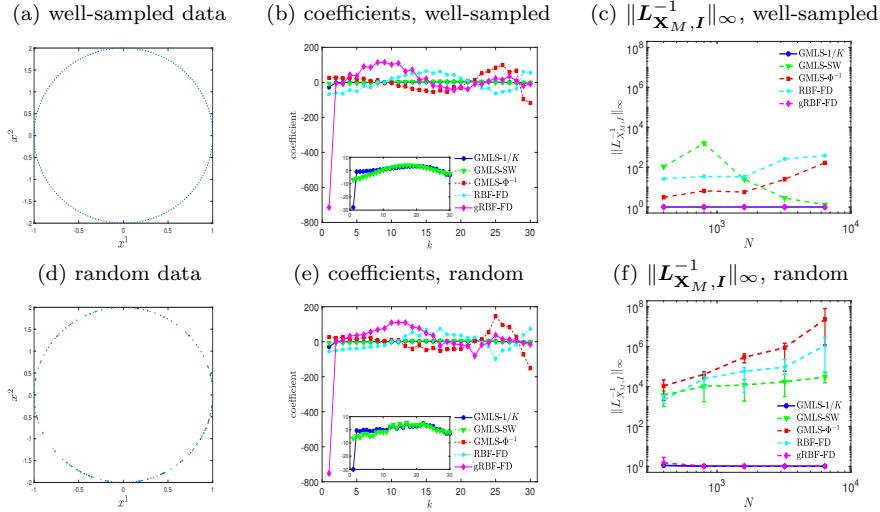


Fig. 1 1D ellipse in \mathbb{R}^2 . Comparison among GMLS (with three different weights), RBF-FD and gRBF-FD. The upper panels are well-sampled data and the bottom panels are random data. The left(a)(d), middle(b)(e), right(c)(f) columns correspond to the scattered dataset, coefficient w_k vs. k th nearest neighbor, $\|L_{X_M, I}^{-1}\|_\infty$ vs. N , respectively. We fix $K = 30$, polynomial degree 4 and PHS parameter $\kappa = 3$.

- GMLS- Φ^{-1} . For the third one, we take the weight Λ to be the inverse of the symmetric PHS matrix, that is, $\tilde{\Phi}_I^\dagger := (\tilde{\Phi}^\top \tilde{\Phi} + \delta^2 I)^{-1} \tilde{\Phi}^\top$ for singular $\tilde{\Phi}$ and $\tilde{\Phi}^{-1}$ for invertible $\tilde{\Phi}$. Here $\tilde{\Phi}$ is defined in (34) and the $\delta^2 I$ term is used for the uniqueness of the inverse of the singular $\tilde{\Phi}$. Indeed, this set of Laplacian coefficients corresponds to the GMLS part of the RBF-FD coefficients as can be derived from equation (12).

While there are many gRBF-FD methods depending on the choice of weight functions and radial basis functions, we here only consider the following two typical gRBF-FD methods.

- RBF-FD. The RBF-FD approach follows from the review in Section 2.2 and references in [62, 36]. The Laplacian coefficients can be constructed from equation (33) by substituting Λ with the inverse of the PHS matrix, $\tilde{\Phi}^{-1}$. When $\tilde{\Phi}$ is invertible, its inverse can be calculated directly by $\tilde{\Phi}^{-1}$. When $\tilde{\Phi}$ is singular, its inverse can be approximated either from a pseudo-inverse based on its singular values or from a ridge regression $\tilde{\Phi}_I^\dagger := (\tilde{\Phi}^\top \tilde{\Phi} + \delta^2 I)^{-1} \tilde{\Phi}^\top$. Here we simply employ a common technique to approximate the inverse for singular $\tilde{\Phi}$.

- gRBF-FD using the $1/K$ weight. The Laplacian coefficients are calculated exactly following equation (33) with the $1/K$ weight $\Lambda = \Lambda_K$ in (18) and the specific inverse $\tilde{\Phi}^{-1} = \tilde{\Phi}_{\Lambda_K}^\dagger$ in (36). The difference between RBF-FD and gRBF-FD lies in the usage of the weight function. In particular, for gRBF-FD using the $1/K$ weight, we allocate more weights to the base point than all

other neighboring points in the first GMLS regression step as well as in the second PHS interpolation step.

Figures 1(b)(e) display the Laplacian coefficient w_k as a function of the k th nearest neighbor for a fixed stencil. It can be seen that for GMLS-1/ K (blue) and gRBF-FD (magenta), the base point coefficient $|w_1|$ is significantly larger than all other neighboring coefficients $|w_k|$ ($k = 2, \dots, K$). We refer to this characteristic as the *central spike pattern* of the coefficients w_k . However, for GMLS-SW (green), GMLS- Φ^{-1} (red), RBF-FD (cyan), their coefficients are nearly uniform within a certain range across the stencil, with no central spike pattern.

Figures 1(c) and 1(f) display the infinity norm of the inverse of the Laplacian matrix, $\|\mathbf{L}_{\mathbf{X}_M, I}^{-1}\|_\infty$, for well-sampled data and randomly-sampled data, respectively. Indeed, the quantity $\|\mathbf{L}_{\mathbf{X}_M, I}^{-1}\|_\infty$ reveals the stability of the Laplacian matrix (see e.g., [2, 67, 28, 35]). As shown in Fig. 1(c), for well-sampled data, all values of $\|\mathbf{L}_{\mathbf{X}_M, I}^{-1}\|_\infty$ are bounded above by a constant, which implies stability; in contrast, the values of GMLS-SW (green), GMLS- Φ^{-1} (red), and RBF-FD (cyan) are significantly larger than those of GMLS-1/ K (blue) and gRBF-FD (magenta). However, for randomly-sampled data (Fig. 1(f)), the values of $\|\mathbf{L}_{\mathbf{X}_M, I}^{-1}\|_\infty$ for GMLS-1/ K (blue) and gRBF-FD (magenta) remain a small constant; whereas those of the other three methods increase with N , potentially leading to their instability.

Moreover, a close connection can be observed between the pattern of coefficients w_k and the value of $\|\mathbf{L}_{\mathbf{X}_M, I}^{-1}\|_\infty$. In particular, the coefficients for GMLS-1/ K and gRBF-FD exhibit the central spike pattern, resulting in a small, constant $\|\mathbf{L}_{\mathbf{X}_M, I}^{-1}\|_\infty$ value and consequently demonstrating good stability. In fact, according to the Lax Equivalence Theorem, these numerical observations will be connected to the convergence of the solution, as demonstrated below in Fig. 2.

In Fig. 2(a) and Fig. 2(c), we illustrate the consistency of the Laplace-Beltrami approximation and the convergence of the solution approximation for well-sampled data. All the five methods show the consistency rate of $\mathcal{O}(N^{-3})$ as well as the super-convergence rate of $\mathcal{O}(N^{-4})$. For well-sampled data or quasi-uniform data, the convergence of GMLS has been numerically examined [46, 30, 36, 35], as has that of RBF-FD [61, 23, 41, 36]. The super-convergence phenomenon has also been discussed in [46].

For randomly-sampled data, the results become different among GMLS, RBF-FD and gRBF-FD. All the five methods still show the consistency rate of $\mathcal{O}(N^{-3})$ as expected (Fig. 2(b)). However, only GMLS-1/ K and gRBF-FD achieve convergent solutions on random data, exhibiting both small errors and small error fluctuations. This good performance stems from the stability of the approximate Laplacians as well as from the central spike pattern of the coefficients w_k . Consequently, employing the 1/ K weight function is very important for ensuring the stability and convergence for both the GMLS and gRBF-FD approaches.

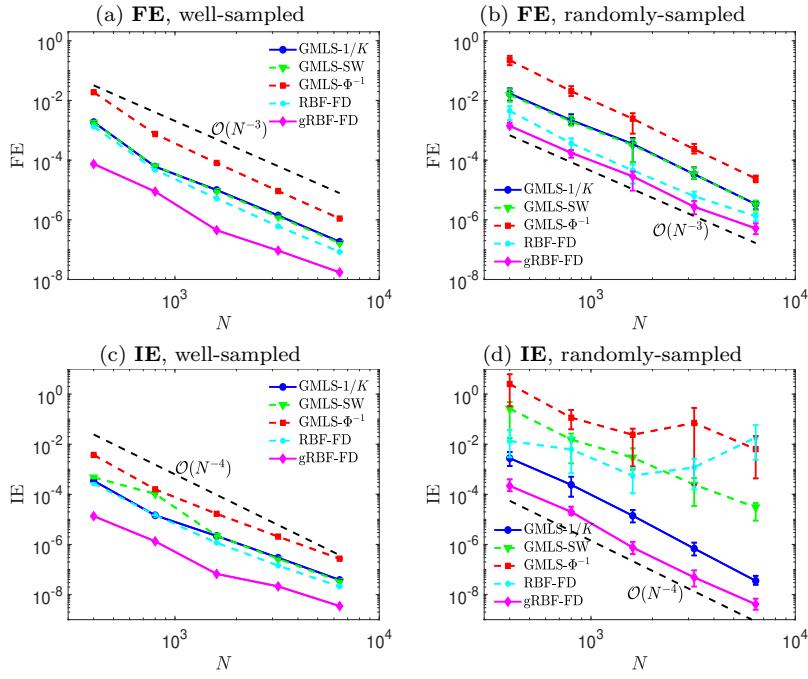


Fig. 2 1D ellipse in \mathbb{R}^2 . Comparison among GMLS (with three weight functions), RBF-FD and gRBF-FD. The left panels are well-sampled data and the right panels are random data. The upper panels are forward error (**FE**) vs. N while the bottom panels are inverse error (**IE**) vs. N . We fix $K = 30$, polynomial degree 4 and PHS parameter $\kappa = 3$.

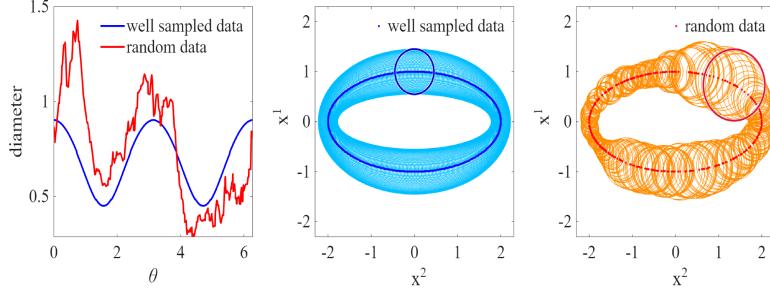


Fig. 3 1D ellipse in \mathbb{R}^2 . (Left) Stencil diameter as a function θ for $N = 400$ points. Plotted are the sizes of all stencils for well-sampled data (middle, light blue circles) and random data (right, orange circles). Each circle is plotted with a center \mathbf{x}_i and a radius $D_{K,\max}(\mathbf{x}_i)/2$. The circle plotted using dark blue (middle) and dark red (right) corresponds to the largest stencil diameter. We fix $K = 30$, polynomial degree 4 and PHS parameter $\kappa = 3$.

In addition, Fig. 2 shows that the **FEs** for random data are nearly 10 times those for well-sampled data. As shown in Lemma 2 and Theorem 1 below, the **FEs** are of order $O(D_{K,\max}^{l-1})$, where $D_{K,\max}$ is the stencil diameter and l is the polynomial degree. From Fig. 3, the largest stencil diameter for random data is nearly twice that for well-sampled data. Therefore, theoretically, the **FEs** for random data should be approximately $2^3 = 8$ times larger. It is a prediction that agrees well with our numerical observations.

4.4 Automatic specification of K -nearest neighbors

We now discuss an automated method of tuning the K -nearest neighbors for our gRBF-FD method. This automatic tuning method can also be applied to the GMLS method.

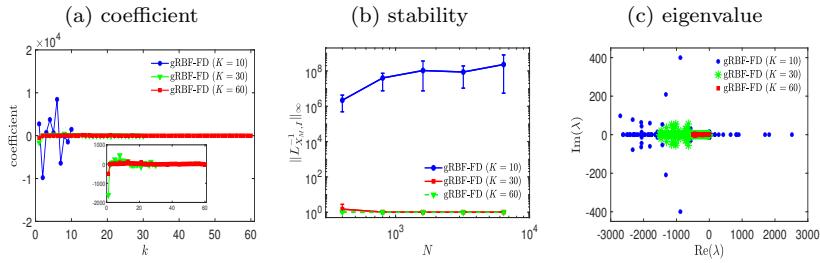


Fig. 4 1D ellipse in \mathbb{R}^2 using random data. Shown are (a) the Laplacian coefficient w_k vs. k -th nearest neighbor, (b) $\|L_{X_M,I}^{-1}\|_\infty$ vs. N , and (c) the leading 200 eigenvalues of L_{X_M} for different K neighbors. We fix polynomial degree 4 and PHS parameter $\kappa = 3$ for all panels. In panels (a)(c), we use $N = 1600$ random data points.

Why do we need to tune K automatically?

We first note that relatively small K may provide unstable approximation to the Laplacian. For the 1D ellipse example, we know that K should be chosen such that $K \geq m = l + 1 = 5$ if the polynomial degree is $l = 4$. We now investigate the numerical performance for three values of K (≥ 5) in terms of the Laplacian coefficients w_k (Fig. 4(a)), the stability $\|L_{X_M,I}^{-1}\|_\infty$ (Fig. 4(b)), and the leading eigenvalues (Fig. 4(c)). For a relatively small value of K (e.g., $K = 10$), the numerical results show the following closely connected characteristics: a positive coefficient at the base point ($w_1 > 0$) (Fig. 4(a)), an unstable approximation to the Laplacian (Fig. 4(b)), and the presence of spurious eigenvalues in the right half of the complex plane (Fig. 4(c)). Here, the Laplace-Beltrami operator is negative semi-definite. For relatively large values of K (e.g., 30, 60), the Laplacian approximation becomes stable and all leading eigenvalues lie in the left half complex plane. Hence, avoiding relatively small values of K is crucial for the successful implementation of the gRBF-FD method (as well as the GMLS method).

There are other minor reasons for auto-tuning K . For instance, a global choice of K for all points might not be a good strategy, especially for randomly sampled data. Also for random data, it can be challenging to get K right on the first try, especially in three-dimensional or higher dimensional problems. Here, we only conjecture some potential drawbacks of manually-tuned K , without showing evidence or numerical validation.

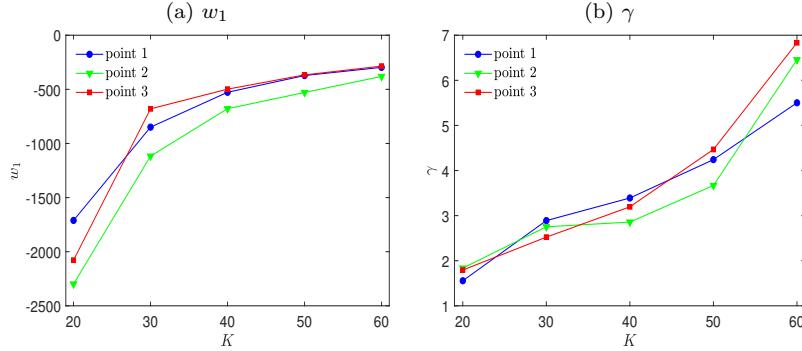


Fig. 5 1D ellipse in \mathbb{R}^2 using random data. Shown are (a) the coefficient of the center point w_1 vs. the stencil size K and (b) the ratio γ vs. K . In both panels (a) and (b), we show the results on three different points. We fix the number of points $N = 1600$, polynomial degree 4 and PHS parameter $\kappa = 3$.

How to automatically tune K ?

As can be seen from Fig. 4, there is a close connection among the Laplacian coefficients w_k , the stability $\|\mathbf{L}_{\mathbf{X}_M, \mathbf{I}}^{-1}\|_\infty$, and the leading eigenvalues. We hope to reduce the upper bound of $\|\mathbf{L}_{\mathbf{X}_M, \mathbf{I}}^{-1}\|_\infty$ as well as to ensure that all eigenvalues lie in the left half plane. However, it might not be easy to predict these two properties of the Laplacian matrix before it is constructed. Fortunately, we are able to control the central spike pattern of the Laplacian coefficients by tuning the K for each point so that the Laplacian approximation is expected to be stable.

Definition 2 To quantitatively characterize the central spike pattern, we define a ratio $\gamma(\mathbf{x}_i)$ as

$$\gamma := \frac{|w_1|}{\max_{2 \leq k \leq K} |w_k|},$$

for each $\mathbf{x}_i \in \mathbf{X}_M$. The Laplacian matrix generated by the GMLS or gRBF-FD methods is said to be nearly diagonally dominant if the ratio is such that $\gamma \geq \gamma_{\text{th}} = 3$ for each point in \mathbf{X}_M .

Our goal is to find the parameter K such that the Laplacian coefficients $\{w_k\}_{k=1}^K$ satisfy two conditions: a negative coefficient at the base point ($w_1 < 0$) and a ratio γ exceeding a given threshold γ_{th} ($\gamma \geq \gamma_{\text{th}}$). It can be seen from Fig. 5(a) that the first condition ($w_1 < 0$) can be easily satisfied for the

coefficient to be negative. This can be expected since the Laplace-Beltrami is negative semi-definite. It can be further seen from Fig. 5(b) that the second condition, $\gamma \geq \gamma_{\text{th}}$, can also be satisfied when the threshold $\gamma_{\text{th}} = 3$. This can be expected because assigning a higher weight to the base point in both optimization problems (5) and (35) is designed to increase γ . The threshold $\gamma_{\text{th}} = 3$ was chosen empirically because it consistently leads to a stable approximation in our experiments. Moreover, in all numerical examples in Section 5, we successfully found a parameter K for each base point that satisfies both conditions: $w_1 < 0$ and $\gamma \geq \gamma_{\text{th}} = 3$.

Now we summarize the automatic tuning strategy of K as follows. For each point, the parameter K is kept growing from an initial value $K_0 (> m)$ until an appropriate K is identified such that a set of Laplacian coefficients $\{w_k\}_{k=1}^K$ satisfies both $w_1 < 0$ and $\gamma \geq 3$. Hence the Laplacian matrix becomes nearly diagonally dominant for the GMLS-1/ K and gRBF-FD methods.

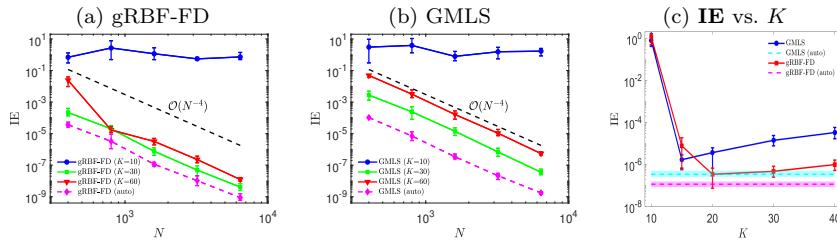


Fig. 6 1D ellipse in \mathbb{R}^2 using random data. Comparison of **IEs** for different stencil size K . Panels (a) and (b) display **IE** vs. N for gRBF-FD and GMLS, respectively. Panel (c) displays **IE** vs. K using $N = 1600$ points. The errors are plotted with error bars obtained from the standard deviations. We fix polynomial degree 4 and PHS parameter $\kappa = 3$. For gRBF-FD (auto) and GMLS (auto), we set the initial $K_0 = 10$.

What is the numerical performance of auto-tuned K ?

In Fig. 6(a) and Fig. 6(b), we compare the performance of **IEs** between manually-tuned K and auto-tuned K for the GMLS-1/ K and the gRBF-FD, respectively. It can be seen that a manually chosen large $K (= 60)$ provides large **IEs**, whereas a manually chosen small $K (= 10)$ fails to maintain the stability. In Fig. 6(c), we show **IEs** as a function of K for GMLS-1/ K and gRBF-FD. It can be observed that gRBF-FD (red solid) not only exhibits smaller **IEs** than GMLS (blue solid) but is also less sensitive to variations in K for $K \geq 20$. This advantage of gRBF-FD primarily results from the improved operator approximation and the enhanced stability using the $1/K$ weight Λ_K . It can be further observed from Fig. 6 that approaches using auto-tuned K always provide smaller **IEs** than those using manually-tuned K for both GMLS-1/ K and gRBF-FD. Indeed, the auto-tuning strategy provides an appropriate value of K that stabilizes the approximation and avoids unnecessary stencil growth for each point. Therefore, using auto-tuned K is important for achieving good numerical performance of GMLS-1/ K and gRBF-FD.

4.5 GRBF-FD algorithm using auto-tuned K -nearest neighbors

Combining the computational techniques of the normalization of Monge coordinates (Sec. 4.2), weighted ridge regression (Sec. 4.3), and auto-tuned K -nearest neighbors (Sec. 4.4), we now provide our gRBF-FD algorithm with auto-tuned K in Algorithm 1. The GMLS algorithm using auto-tuned K can be applied similarly following the similar procedure.

Algorithm 1 gRBF-FD using auto-tuned K for the Laplace-Beltrami operator

- 1: **Input:** A point cloud $\{\mathbf{x}_i\}_{i=1}^N \subset M$, bases of analytic tangent vectors at each node $\{t_1(\mathbf{x}_i), \dots, t_d(\mathbf{x}_i)\}_{i=1}^N$, the degree l of polynomials, the PHS smoothness parameter κ , and a parameter $K (= K_0 > m)$ nearest neighbors where $m = \binom{l+d}{d}$ is the number of monomial basis functions $\{\theta^\alpha | 0 \leq |\alpha| \leq l\}$.
- 2: Set $\mathbf{L}_{\mathbf{X}_M}$ to be a sparse $N \times N$ matrix with initial NK nonzeros.
- 3: **for** $i \in \{1, 2, \dots, N\}$ **do**
- 4: Set $\gamma = 0$ and $w_1 = 0$.
- 5: **while** $\gamma < 3$ or $w_1 > 0$ **do**
- 6: Find the K nearest neighbors of the point \mathbf{x}_i , denoted by the stencil $S_{\mathbf{x}_i} = \{\mathbf{x}_{i,k}\}_{k=1}^K$.
- 7: Compute the stencil diameter $D_{K,\max}(\mathbf{x}_i)$ and the normalized Monge coordinates $\tilde{\theta}(\mathbf{x}_{i,k})$ using (32).
- 8: Construct the K by m Vandermonde-type matrix $\tilde{\mathbf{P}}$ with entries $\tilde{\mathbf{P}}_{kj} = p_{\alpha(j)}(\tilde{\theta}(\mathbf{x}_{i,k})) = \prod_{s=1}^d \tilde{\theta}_s^{\alpha_s(j)}(\mathbf{x}_{i,k})$.
- 9: Construct the K by K PHS matrix $\tilde{\Phi}$ with entries $\tilde{\Phi}_{ks} = \|\tilde{\theta}(\mathbf{x}_{i,k}) - \tilde{\theta}(\mathbf{x}_{i,s})\|^{2\kappa+1}$.
- 10: Calculate the Laplacian coefficients $\{w_k\}_{k=1}^K$ by the formula in (33),

$$(w_1, \dots, w_K) = \frac{1}{D_{K,\max}^2} \left((\Delta_{\tilde{\theta}} \tilde{\phi}_{\mathbf{x}_0}) \tilde{\Phi}_{\Lambda_K}^\dagger (\mathbf{I} - \tilde{\mathbf{P}} (\tilde{\mathbf{P}}^\top \Lambda_K \tilde{\mathbf{P}})^{-1} \tilde{\mathbf{P}}^\top \Lambda_K) \right. \\ \left. + (\Delta_{\tilde{\theta}} \tilde{p}_{\mathbf{x}_0}) (\tilde{\mathbf{P}}^\top \Lambda_K \tilde{\mathbf{P}})^{-1} \tilde{\mathbf{P}}^\top \Lambda_K \right),$$

where $\tilde{\Phi}_{\Lambda_K}^\dagger$ is given (36), Λ_K is given in (18), and $\Delta_{\tilde{\theta}} \tilde{\phi}_{\mathbf{x}_0}$ and $\Delta_{\tilde{\theta}} \tilde{p}_{\mathbf{x}_0}$ are given in (34).

- 11: Calculate the ratio $\gamma(\mathbf{x}_i) = \frac{|w_1|}{\max_{2 \leq k \leq K} |w_k|}$ to characterize the central spike pattern at \mathbf{x}_i .
 - 12: Increase the K by a small positive integer, e.g., 2.
 - 13: **end while**
 - 14: Arrange the Laplacian coefficients $\{w_k\}_{k=1}^K$ into the corresponding rows and columns of $\mathbf{L}_{\mathbf{X}_M}$.
 - 15: **end for**
 - 16: **Output:** The approximate operator matrix $\mathbf{L}_{\mathbf{X}_M}$.
-

4.6 Error analysis for operator approximation

In this section, we provide the error estimate for the Laplace-Beltrami operator approximation. Note that for quasi-uniform data, the error estimate is written

in terms of the fill distance and separation distance (see e.g., [69, 50, 26, 27, 30]). Here, for random data, we first write the error in terms of the stencil diameter $D_{K,\max}$, and then relate it to the number of data points, N .

4.6.1 Error estimate in terms of the stencil diameter

The following local polynomial reproduction property is presented in terms of the stencil diameter, in contrast to the definition using the fill distance in [69, 50]. For simplicity, we only consider the derivative to be the Laplace-Beltrami operator.

Lemma 1 Consider a process that defines for each stencil $S_{\mathbf{x}_0} = \{\mathbf{x}_{0,k}\}_{k=1}^K$ a family of function $u_{k,\Delta} : M \rightarrow \mathbb{R}, 1 \leq k \leq K$ to approximate

$$\Delta_M f(\mathbf{x}) \approx \sum_{k=1}^K u_{k,\Delta}(\mathbf{x}) f(\mathbf{x}_{0,k}) := \Delta_M \mathcal{I}_p \mathbf{f}_{\mathbf{x}_0}(\mathbf{x}), \quad (38)$$

for functions $f \in C^l(M)$. Then we say that the process provides a local polynomial reproduction of degree l on M if there exists a constant $C_1 > 0$ such that

- (1) $\sum_{k=1}^K u_{k,\Delta}(\mathbf{x}) p(\mathbf{x}_{0,k}) = \Delta_M p(\mathbf{x}), \text{ for all } p \in \mathbb{P}_{\mathbf{x}_0}^{l,d}, \mathbf{x} \in M,$
- (2) $\sum_{k=1}^K |u_{k,\Delta}(\mathbf{x})| \leq C_1 D_{K,\max}^{-2}, \forall \mathbf{x} \in M,$
- (3) $u_{k,\Delta}(\mathbf{x}) = 0 \text{ if } \|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\| > D_{K,\max}.$

For the first step of our gRBF-FD, the process satisfies the local polynomial reproduction if we use the GMLS interpolant (4) with the regression coefficients in (13) (see Appendix A). Then we have the interpolation result following from Theorem 4.2 and Theorem 4.3 of [50].

Lemma 2 Define $\mathcal{I}_p \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})$ (equation (4)) and $\Delta_M \mathcal{I}_p \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})$ (equation (38)) to be the GMLS approximation to $f(\mathbf{x})$ and $\Delta_M f(\mathbf{x})$, respectively, using local polynomials up to degree l . Let $\{u_{k,\Delta}(\mathbf{x})\}$ in (38) be a local polynomial reproduction of order l . Then for any $f \in C^{l+1}(M)$ with $l \geq 2$, there is an error bound

$$\begin{aligned} |f(\mathbf{x}) - \mathcal{I}_p \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})| &\leq C_2 D_{K,\max}^{l+1}(\mathbf{x}_0) |f|_{C^{l+1}(M)}, \\ |\Delta_M f(\mathbf{x}) - \Delta_M \mathcal{I}_p \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})| &\leq \tilde{C}_2 D_{K,\max}^{l-1}(\mathbf{x}_0) |f|_{C^{l+1}(M)}, \end{aligned}$$

for all $\mathbf{x} \in M$ and some $C_2, \tilde{C}_2 > 0$. Here the semi-norm, $|f|_{C^{l+1}(M)} := \max_{|\boldsymbol{\alpha}|=l+1} \|D^{\boldsymbol{\alpha}} f\|_{L^\infty(M)}$, is defined over M , where $D^{\boldsymbol{\alpha}}$ denotes a general multi-dimensional derivative with multi-index $\boldsymbol{\alpha}$.

To study the second step of gRBF-FD in (14), we next provide a local reproduction property for a PHS function.

Lemma 3 Let ϕ be a PHS function and its interpolant (14) satisfies $\mathcal{I}_\phi : \mathbb{R}^K \rightarrow C^{l_2}$, where l_2 denotes the smoothness of ϕ . Consider a process that defines for each stencil $S_{\mathbf{x}_0} = \{\mathbf{x}_{0,k}\}_{k=1}^K$ a family of functions $v_{k,\Delta} : M \rightarrow \mathbb{R}, 1 \leq k \leq K$ to approximate

$$\Delta_M s(\mathbf{x}) \approx \sum_{k=1}^K v_{k,\Delta}(\mathbf{x}) s(\mathbf{x}_{0,k}) := \Delta_M \mathcal{I}_\phi \mathbf{s}_{\mathbf{x}_0}(\mathbf{x}), \quad (39)$$

for functions $s \in C^{l_2}(M)$. Then we say that the process provides a local PHS weak reproduction on M if there exist $C_3, C_4 > 0$ such that

(1) for $\mathbf{x} \in M$ and $\varphi = \sum_{k=1}^K \tilde{c}_k \tilde{\phi}_k \in \text{Span}\{\tilde{\phi}_1, \dots, \tilde{\phi}_K\}$,

$$\begin{aligned} |\Delta_M \mathcal{I}_\phi \varphi_{\mathbf{x}_0}(\mathbf{x}) - \Delta_M \varphi(\mathbf{x})| &= \left| \sum_{k=1}^K v_{k,\Delta}(\mathbf{x}) \varphi(\mathbf{x}_{0,k}) - \Delta_M \varphi(\mathbf{x}) \right| \\ &\leq C_3 \delta^2 D_{K,\max}^{-2} \max_{1 \leq k \leq K} |\tilde{c}_k|, \end{aligned}$$

where δ is the regularization parameter in (35) and $\tilde{\phi}_k(\mathbf{x}) = \phi(\|\tilde{\boldsymbol{\theta}}(\mathbf{x}) - \tilde{\boldsymbol{\theta}}(\mathbf{x}_{0,k})\|) = \phi(\|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\|/D_{K,\max})$ for $k = 1, \dots, K$,

(2) $\sum_{k=1}^K |v_{k,\Delta}(\mathbf{x})| \leq C_4 D_{K,\max}^{-2}$, $\forall \mathbf{x} \in M$,

(3) $v_{k,\Delta}(\mathbf{x}) = 0$ if $\|\boldsymbol{\theta}(\mathbf{x}) - \boldsymbol{\theta}(\mathbf{x}_{0,k})\| > D_{K,\max}$.

Here, C_3, C_4 are independent of the stencil diameter $D_{K,\max}$, but depend on the regularization parameter δ .

We partially prove the above Lemma 3 and examine some of the results by numerical verification. See Appendix A in detail.

Theorem 1 Let $\mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})$ in equation (17) be the gRBF-FD approximation to $f(\mathbf{x})$ using local polynomials up to degree l and radial basis functions with smoothness $l_2 \geq l$. Let $\{u_{k,\Delta}(\mathbf{x})\}$ in (38) be a local polynomial reproduction of order l and let $\{v_{k,\Delta}(\mathbf{x})\}$ in (39) be a local PHS weak reproduction. Then for any $f \in C^{l+1}(M)$ with $l \geq 2$, there is an error bound

$$|\Delta_M f(\mathbf{x}) - \Delta_M \mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})| \leq C_5 D_{K,\max}^{l-1}(\mathbf{x}_0) |f|_{C^{l+1}(M)}, \quad (40)$$

for all $\mathbf{x} \in M$ and some $C_5 > 0$.

Proof Define the residual function $s(\mathbf{x}) = f(\mathbf{x}) - \mathcal{I}_p \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})$ for the first step using the GMLS regression. Using the gRBF-FD interpolant (17), we have the error bound

$$\begin{aligned} |\Delta_M f(\mathbf{x}) - \Delta_M \mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})| &= |\Delta_M f(\mathbf{x}) - \Delta_M \mathcal{I}_p \mathbf{f}_{\mathbf{x}_0}(\mathbf{x}) - \Delta_M \mathcal{I}_\phi \mathbf{s}_{\mathbf{x}_0}(\mathbf{x})| \\ &= |\Delta_M s(\mathbf{x}) - \Delta_M (\mathcal{I}_\phi \mathbf{s}_{\mathbf{x}_0})(\mathbf{x})|, \end{aligned}$$

where $\mathbf{s}_{\mathbf{x}_0} = s(\mathbf{x})|_{\mathbf{x} \in S_{\mathbf{x}_0}} \in \mathbb{R}^K$. Let $\varphi \in \text{Span}\{\tilde{\phi}_1, \dots, \tilde{\phi}_K\}$ be an arbitrary function spanned by the K -dimensional subspace. Then the error bound becomes

$$\begin{aligned} & |\Delta_M f(\mathbf{x}) - \Delta_M \mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})| \\ & \leq |\Delta_M s(\mathbf{x}) - \Delta_M \varphi(\mathbf{x})| + |\Delta_M \varphi(\mathbf{x}) - \Delta_M (\mathcal{I}_\phi \varphi_{\mathbf{x}_0})(\mathbf{x})| + |\Delta_M (\mathcal{I}_\phi \varphi_{\mathbf{x}_0})(\mathbf{x}) - \Delta_M (\mathcal{I}_\phi \mathbf{s}_{\mathbf{x}_0})(\mathbf{x})| \\ & \leq |\Delta_M s(\mathbf{x}) - \Delta_M \varphi(\mathbf{x})| + |\Delta_M \varphi(\mathbf{x}) - \Delta_M (\mathcal{I}_\phi \varphi_{\mathbf{x}_0})(\mathbf{x})| + \sum_{k=1}^K |v_{k,\Delta}(\mathbf{x})| |\varphi(\mathbf{x}_{0,k}) - s(\mathbf{x}_{0,k})| \\ & \leq \|\Delta_M s - \Delta_M \varphi\|_{L^\infty(\mathcal{D})} + C_3 \delta^2 D_{K,\max}^{-2} \max_{1 \leq k \leq K} |\tilde{c}_k| + C_4 D_{K,\max}^{-2} \|\varphi - s\|_{L^\infty(\mathcal{D})}. \end{aligned} \quad (41)$$

where $\mathcal{D} = B(\mathbf{x}_0, R_{K,\max}(\mathbf{x}_0))$. In the last inequality, we have used the properties in Lemma 3. Now we choose φ to be

$$\varphi = \arg \min_{\hat{\varphi} \in \text{Span}\{\tilde{\phi}_1, \dots, \tilde{\phi}_K\}} \|\Delta_M s - \Delta_M \hat{\varphi}\|_{L^\infty(\mathcal{D})} + C_4 D_{K,\max}^{-2} \|\hat{\varphi} - s\|_{L^\infty(\mathcal{D})}.$$

Notice that the optimal value must be less than the objective function with $\hat{\varphi} = 0$,

$$\begin{aligned} & \|\Delta_M s - \Delta_M \varphi\|_{L^\infty(\mathcal{D})} + C_4 D_{K,\max}^{-2} \|\varphi - s\|_{L^\infty(\mathcal{D})} \\ & \leq \|\Delta_M s\|_{L^\infty(\mathcal{D})} + C_4 D_{K,\max}^{-2} \|s\|_{L^\infty(\mathcal{D})} \\ & \lesssim D_{K,\max}^{l-1} |f|_{C^{l+1}(M)}, \end{aligned} \quad (42)$$

where the results in Lemma 2 have been used.

Next we estimate the second error term of (41). From (42), we see that

$$\|\varphi - s\|_{L^\infty(\mathcal{D})} \lesssim D_{K,\max}^{l+1} |f|_{C^{l+1}(M)}.$$

Since $\|s\|_{L^\infty(\mathcal{D})} \lesssim D_{K,\max}^{l+1} |f|_{C^{l+1}(M)}$ using Lemma 2, we have that $\|\varphi\|_{L^\infty(\mathcal{D})} \lesssim D_{K,\max}^{l+1} |f|_{C^{l+1}(M)}$ and each $|\tilde{c}_k| \lesssim D_{K,\max}^{l+1} |f|_{C^{l+1}(M)}$. Then the second term in (41) can be bounded by

$$C_3 \delta^2 D_{K,\max}^{-2} \max_{1 \leq k \leq K} |\tilde{c}_k| \lesssim \delta^2 D_{K,\max}^{l-1} |f|_{C^{l+1}(M)}. \quad (43)$$

Substituting (42) and (43) into (41), we arrive at the desired error bound in (40). In particular, if the second error term (43) is negligible, then the error bound becomes

$$\begin{aligned} & |\Delta_M f(\mathbf{x}) - \Delta_M \mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})| \\ & \leq \inf_{\hat{\varphi} \in \text{Span}\{\tilde{\phi}_1, \dots, \tilde{\phi}_K\}} \left\{ \|\Delta_M s - \Delta_M \hat{\varphi}\|_{L^\infty(\mathcal{D})} + C_4 D_{K,\max}^{-2} \|\hat{\varphi} - s\|_{L^\infty(\mathcal{D})} \right\} \\ & \leq C_5 D_{K,\max}^{l-1} |f|_{C^{l+1}(M)}. \end{aligned}$$

4.6.2 Error estimate in terms of the number of data

We now estimate the error bound in terms of the number of data.

Lemma 4 *For some integer K and define the stencil diameter $D_{K,\max}(\mathbf{x}_0)$ and stencil radius $R_{K,\max}(\mathbf{x}_0)$ as in Definition 1 for each base point \mathbf{x}_0 . Assume that there exists a constant $C_K > 0$ such that $R_{K,\max}(\mathbf{x}_i) \leq D_{K,\max}(\mathbf{x}_i) \leq C_K R_{K,\max}(\mathbf{x}_i)$ for all stencils $\{S_{\mathbf{x}_i}\}_{i=1}^N$. Let $\mathbf{x}_1, \dots, \mathbf{x}_N$ be i.i.d. random samples from the distribution Q with the density $q \in L^1(M)$ such that $\int_M q(\mathbf{x})dV(\mathbf{x}) = 1$ and $q \geq q_{\min} > 0$ for some positive q_{\min} . Then, we have for each \mathbf{x}_i and all $\delta > 0$,*

$$\mathbb{P}_{\mathbf{X}_M \sim Q}(R_{K,\max}(\mathbf{x}_i) > \delta) \leq \exp(-q_{\min} C_d(N - K)\delta^d),$$

where C_d is a constant depending on the dimension d . Moreover, with probability higher than $1 - \frac{1}{N}$, we have

$$R_{K,\max}(\mathbf{x}_i) = O\left(\left(\frac{\log N}{N}\right)^{\frac{1}{d}}\right), \quad D_{K,\max}(\mathbf{x}_i) = O\left(\left(\frac{\log N}{N}\right)^{\frac{1}{d}}\right),$$

where the constant in the big-oh is independent of N . Subsequently, we also have

$$ED_{K,\max}(\mathbf{x}_i) = O\left(\left(\frac{\log N}{N}\right)^{\frac{1}{d}}\right), \quad \sigma(D_{K,\max}(\mathbf{x}_i)) = O\left(\left(\frac{\log N}{N}\right)^{\frac{1}{d}}\right),$$

where the expectation E and the standard deviation σ are taken with respect to the distribution of $D_{K,\max}(\mathbf{x}_i)$ at a fixed \mathbf{x}_i across different samplings of the point cloud data $\{\mathbf{x}_i\}_{i=1}^N \subset M$.

Proof See Appendix B.

Using the results in Theorem 1 and Lemma 4, we arrive at the following the error bound in terms of the number of data.

Theorem 2 *Suppose that the conditions and assumptions in Theorem 1 and Lemma 4 hold. Then for any $f \in C^{l+1}(M)$ with $l \geq 2$, there is an error bound*

$$|\Delta_M f(\mathbf{x}) - \Delta_M \mathcal{I}_{\phi p} \mathbf{f}_{\mathbf{x}_0}(\mathbf{x})| = O\left(\left(\frac{\log N}{N}\right)^{\frac{l-1}{d}}\right), \quad (44)$$

for all $\mathbf{x} \in M$, where the constant in the big-oh error bound is independent of N .

5 Numerical experiments

To support the performance of gRBF-FD, we present numerical results for solving the screened Poisson equations in (29) across various manifolds, identified by randomly sampled point clouds. We also compare with the results using the GMLS method. For all examples below, the manifolds are assumed to be known with analytic tangent planes. We fix the PHS parameter to be $\kappa = 3$. For both GMLS and gRBF-FD, we apply the auto-tuned K -nearest neighbors starting from an initial value K_0 , as detailed in Section 4.4 and Algorithm 1.

This section is organized as follows. In Sections 5.1 and 5.2, we examine the numerical performance on two 2D smooth surfaces: a red blood cell (RBC) and a bumpy sphere (BSP), respectively. In Section 5.3, we report the numerical results on two higher-dimensional manifolds, a 3D flat torus embedded in \mathbb{R}^{12} and a 4D flat torus embedded in \mathbb{R}^{16} . In Section 5.4, we verify the performance on Stanford bunny and armadillo models. The following numerical results demonstrate that our gRBF-FD approach can persistently provide stable and convergent solutions with smaller errors than GMLS across multiple trials of randomly sampled data points.

5.1 Red blood cell

Consider a red blood cell (RBC) with the parametrization:

$$\mathbf{x} = \left(r \cos \theta \cos \phi, r \cos \theta \sin \phi, \frac{1}{2} \sin \theta (c_0 + c_2 \cos^2 \theta + c_4 \cos^4 \theta) \right), \quad (45)$$

where $-\pi/2 \leq \theta \leq \pi/2$, $-\pi \leq \phi < \pi$, $r = 3.91/3.39$, $c_0 = 0.81/3.39$, $c_2 = 7.83/3.39$ and $c_4 = -4.39/3.39$, which is the same as the RBC in [27]. In our numerical experiment, we set the true solution to be $f = \cos^2(\theta)$ and then we can calculate the RHS $h := (1 - \Delta_M) f$. Next, we approximate the numerical solution for the PDE problem subjected to the manufactured h . Numerically, the points $\{\mathbf{x}_i\}_{i=1}^N$ are generated from the parametrization (45) using randomly sampled $\{\theta_i, \phi_i\}_{i=1}^N$.

In Fig. 7, we plot the **FEs** and **IEs** as a function of N over 4 independent trials. It can be seen that the **FEs** decrease on the order of $N^{-(l-1)/2}$ for both GMLS and gRBF-FD. This error rate is in good agreement with the theoretical one in (44). It can be further observed that the **IEs** for $l = 2$ and $l = 4$ decay with respective rates N^{-1} and N^{-2} , which are half an order faster than the corresponding **FEs**. This reveals that the **IEs** possess the super-convergence phenomenon, as also observed in [46, 35, 42]. Moreover, it can be seen that the **IEs** of gRBF-FD are smaller than those of GMLS by factors of approximately 10% and 20% for $l = 2$ and $l = 4$, respectively.

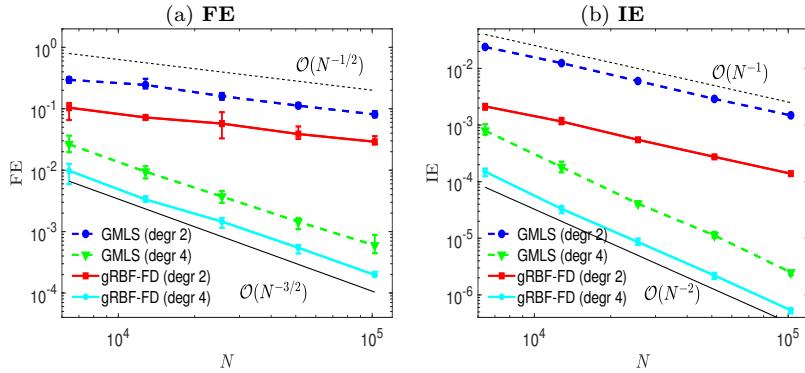


Fig. 7 2D RBC in \mathbb{R}^3 using random data. Shown are (a) **FEs** vs. the number of points N and (b) **IEs** vs. N . We show the results using different polynomial degrees $l = 2, 4$. We fix the PHS parameter $\kappa = 3$ and the initial $K_0 = 40$.

5.2 Bumpy sphere

We consider a smooth but complex surface, bumpy sphere (BSP), parameterized by

$$\mathbf{x} = (x^1, x^2, x^3) = (r(\theta, \phi) \sin \theta \cos \phi, r(\theta, \phi) \sin \theta \sin \phi, r(\theta, \phi) \cos \theta), \quad (46)$$

where $0 \leq \theta \leq \pi$, $0 \leq \phi < 2\pi$ and $r(\theta, \phi) = 1 + 0.1(\sin^7 4\theta)(\sin 4\phi)$. This surface exhibits C^6 smoothness at the two poles and C^∞ smoothness elsewhere. Its geometric complexity arises from the rapidly varying curvature in some regions (see Fig. 8(a)). In our numerical experiment, we set the true solution to be the third coordinate of the surface $f = x^3$ and then construct the RHS $h := (1 - \Delta_M) f$. The points are generated randomly from the parameterization (46) in intrinsic coordinates, where (θ, ϕ) are i.i.d. sampled with the density $p(\theta, \phi) = \frac{\sin \theta}{4\pi}$ on $[0, \pi] \times [0, 2\pi]$.

In Fig. 8 (a), we show the surface of bumpy sphere as well as the true solution with its value color coded. In panels (b)(c), we plot **FEs** and **IEs** as functions of N over 4 independent trials. It can be seen that **FEs** for both GMLS and gRBF-FD decay on the order of $N^{-(l-1)/2}$, which agrees with the theoretical prediction. Again, the **IEs** still decay half an order faster than the **FEs**. Moreover, gRBF-FD achieves smaller errors than GMLS in both operator and solution approximations.

5.3 Flat tori

We consider a 3-dimensional flat torus embedded in \mathbb{R}^{12} with the parameterization,

$$\mathbf{x} = \frac{1}{\sqrt{1^2 + 2^2}} \begin{pmatrix} \cos(\phi_1), \sin(\phi_1), \cos(2\phi_1), \sin(2\phi_1), \\ \cos(\phi_2), \sin(\phi_2), \cos(2\phi_2), \sin(2\phi_2), \\ \cos(\phi_3), \sin(\phi_3), \cos(2\phi_3), \sin(2\phi_3) \end{pmatrix}, \quad (47)$$

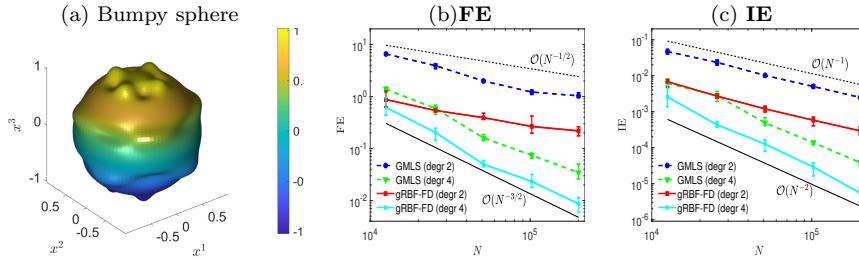


Fig. 8 2D BSP in \mathbb{R}^3 using random data. (a) The true solution $f = x^3$ with its value color coded. Shown are (b) FEes vs. N and (c) IEes vs. N . In both panels (b) and (c), we show the results with polynomial degree $l = 2, 4$. We fix the PHS parameter $\kappa = 3$ and the initial $K_0 = 40$.

with $0 \leq \phi_1, \phi_2, \phi_3 < 2\pi$. The Riemannian metric is given by a 3×3 identity matrix \mathbf{I}_3 . The true solution f is set to be $f = \sin(\phi_1) \sin(\phi_2) \sin(\phi_3)$. Numerically, the points $\{\mathbf{x}_i\}_{i=1}^N$ are generated from the parametrization (47) using randomly sampled $\{\phi_1, \phi_2, \phi_3\}_{i=1}^N$ with the uniform distribution on $[0, 2\pi]^3$. We use the degree $l = 2, 4$ and the initial $K_0 = 60$ nearest neighbors for illustrating the convergence of solutions.

As shown in Fig. 9(a), the FEes decay at the rate $N^{-(l-1)/3}$ for both GMLS and gRBF-FD, which is consistent with the theory in (44) for dimension $d = 3$. As shown in Fig. 9(b), the solutions again show super-convergence, that is, IEes decay with respective rates $N^{-2/3}$ and $N^{-4/3}$ for $l = 2$ and $l = 4$. Moreover, the IEes of gRBF-FD are approximately 30% and 40% of the GMLS errors for $l = 2$ and $l = 4$, respectively.

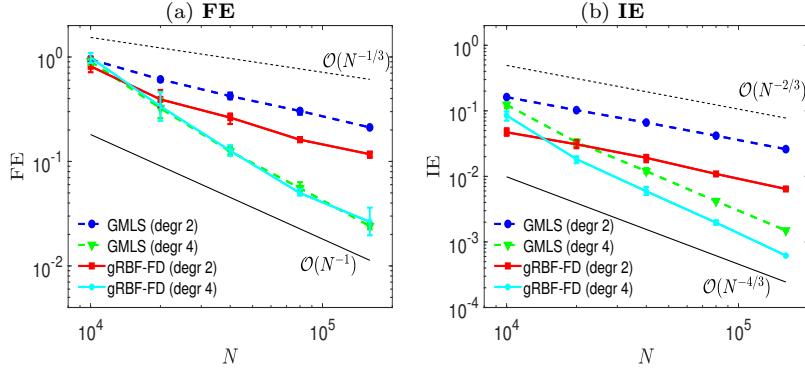


Fig. 9 3D flat torus in \mathbb{R}^{12} using random data. Shown are (a) FEes vs. N and (b) IEes vs. N . In both panels (a) and (b), we show the results with polynomial degree $l = 2, 4$. We fix the PHS parameter $\kappa = 3$ and the initial $K_0 = 60$.

We further consider a 4-dimensional flat torus embedded in \mathbb{R}^{16} with the parameterization,

$$\mathbf{x} = \frac{1}{\sqrt{1^2 + 2^2}} \begin{pmatrix} \cos(\phi_1), \sin(\phi_1), \cos(2\phi_1), \sin(2\phi_1), \\ \cos(\phi_2), \sin(\phi_2), \cos(2\phi_2), \sin(2\phi_2), \\ \cos(\phi_3), \sin(\phi_3), \cos(2\phi_3), \sin(2\phi_3), \\ \cos(\phi_4), \sin(\phi_4), \cos(2\phi_4), \sin(2\phi_4), \end{pmatrix}, \quad (48)$$

where $0 \leq \phi_1, \phi_2, \phi_3, \phi_4 < 2\pi$. The Riemannian metric is given by a 4×4 identity matrix \mathbf{I}_4 . The true solution f is set to be $f = \sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \sin(\phi_4)$. We did not test with polynomial degree $l = 4$ due to the large number of monomial basis functions required compared to the 2D examples above. We use the polynomial degree $l = 3$ and the initial stencil size $K_0 = 75$ nearest neighbors for illustrating the convergence of solutions. It can be seen from Fig. 10 that **FEs** and **IEs** both decay with the rate $N^{-1/2}$, which agrees with the theory in (44). The **IE** of gRBF-FD is still smaller than that of GMLS, despite its larger **FE**.

Moreover, we numerically examine the stability of both approaches when $N = 40000$. In our implementation, we use the command `eigs(L_{X_M}, 6, 10)` in MATLAB to compute the leading six eigenvalues close to 10. Then the leading eigenvalues of GMLS are 0, -1.30, -1.30, -1.30, -1.30, -1.30 while the leading ones of gRBF-FD are 0, -1.21, -1.21, -1.21, -1.21, -1.21. No spurious eigenvalues are observed in the right half of the complex plane. These numerical eigenvalues are approximations of the analytic true eigenvalues of the 4D flat torus, 0, -1, -1, -1, -1, -1 (see e.g., [32]).

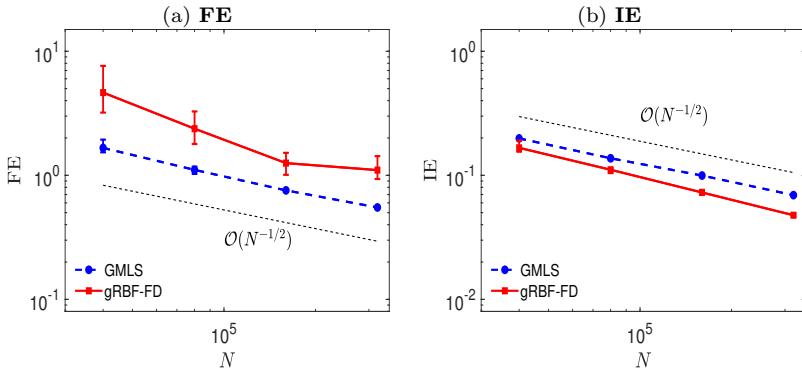


Fig. 10 4D flat torus in \mathbb{R}^{16} using random data. Shown are (a) **FEs** vs. N and (b) **IEs** vs. N . In both panels (a) and (b), we fix the polynomial degree $l = 3$, the PHS parameter $\kappa = 3$ and the initial stencil size $K_0 = 75$.

5.4 Bunny and Armadillo models

We now consider solving the elliptic problems on the Bunny and Armadillo models which are both two-dimensional surfaces embedded in \mathbb{R}^3 . The data of these two models are downloaded from the Stanford 3D Scanning Repository [1]. The original dataset of the Bunny comprises a triangle mesh with 34,817 vertices and the Armadillo comprises a triangle mesh with 172,974 vertices. To resolve the singular regions of the original dataset, we first generate new meshes of the surfaces using the Marching Cubes algorithm [48] that is implemented in MeshLab [14]. Notice that the Marching Cubes algorithm does not smooth the surface. We will compare the solutions among surface finite element method (FEM), GMLS, and gRBF-FD on smoothed surfaces for both Bunny and Armadillo models. Then, we use the Screened Poisson surface reconstruction algorithm to smooth the surfaces that fit the point clouds of models. Next, we apply the Poisson-disk sampling algorithm via the Meshlab to generate point clouds over the surfaces. Subsequently, new meshes could be obtained by using the Marching Cubes algorithm again. As a result, we generate a point cloud of $N = 34,596$ points over the Stanford Bunny model and generate $N = 172,974$ points over the Stanford Armadillo model. For the Armadillo model, we normalize its dataset into a unit box $[0, 1]^3$ for computational convenience. Finally, after all above processes, we use MeshLab to clean the vertices and meshes of the two models for further application in FEM.

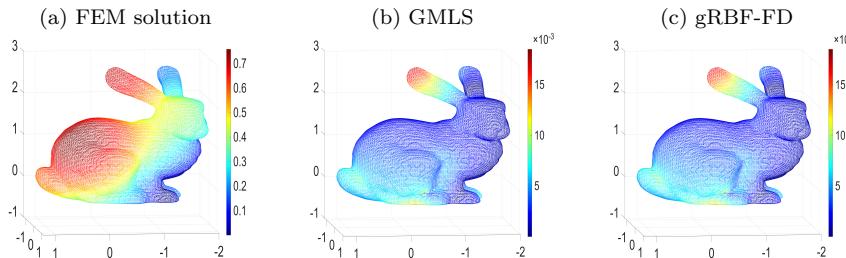


Fig. 11 Bunny example. $N = 34,596$. (a) FEM solution as a reference. (b) Pointwise absolute difference between FEM and GMLS with degree $l = 4$ and $K_0 = 41$ -nearest neighbors. (c) Pointwise absolute difference between FEM and gRBF-FD with $l = 4$ and $K_0 = 41$. GMLS and gRBF-FD both have maximum norm differences of 0.018 and relative differences of 2.4%.

We solve the PDE problem in (29) with $a = 1$ and $f = 0.6(x_1 + x_2 + x_3)$ for both models. Here we have no access to the analytic true solutions due to the unknown embedding functions. For comparison, we take the FEM solution obtained from the FELICITY FEM Matlab toolbox [68] as the reference (see Fig. 11(a) for Bunny and Fig. 12(a) for Armadillo). For GMLS and gRBF-FD, the used parameters can be found in the captions of Figs. 11 and 12. For Bunny, it can be seen from Figs. 11(b) and (c) that GMLS and gRBF-FD show comparable results, both with maximum norm differences of 0.018 and

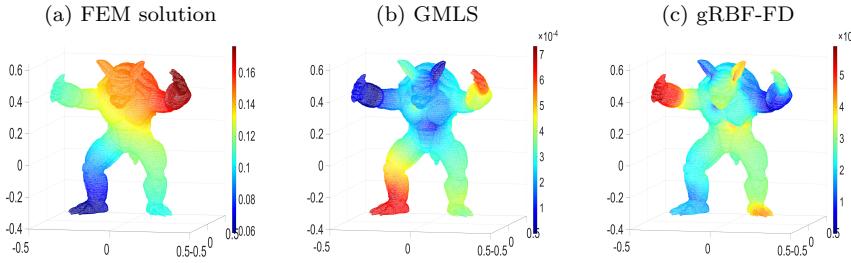


Fig. 12 Armadillo example. $N = 172,974$. (a) FEM solution as a reference. (b) Pointwise absolute difference between FEM and GMLS with degree $l = 4$ and $K_0 = 41$ -nearest neighbors. GMLS has the maximum norm difference of 7.4×10^{-4} and the relative difference of 0.41%. (c) Pointwise absolute difference between FEM and gRBF-FD with $l = 4$ and $K_0 = 41$. GRBF-FD has the maximum norm difference of 5.9×10^{-4} and the relative difference of 0.33%.

relative differences of 2.4%. For Armadillo, it can be seen from Figs. 12(b) and (c) that the absolute difference between the FEM and the gRBF-FD is slightly smaller than that of FEM and GMLS. For GMLS, the maximum norm difference is 7.4×10^{-4} and the relative difference is 0.41%. For gRBF-FD, the maximum norm difference is 5.9×10^{-4} and the relative difference is 0.33%.

6 Conclusions

In this paper, we considered a two-step generalized RBF-FD approach for solving screened Poisson problems on smooth manifolds, identified by randomly sampled point clouds. In the first step, we applied a GMLS regression to capture the smooth, leading component in the Taylor's expansion of the target function f . In the second step, we employed a PHS interpolation to handle the residual from the GMLS regression, which contains the high-order remainder of the Taylor's expansion. We established the error bound of gRBF-FD in Theorem 2 for the consistency analysis of the Laplace-Beltrami operator, demonstrating that the method can achieve arbitrary-order algebraic accuracy on smooth manifolds. For the stability of the Laplacian matrix, we employed both GMLS regression and PHS interpolation within a weighted norm which assigns a larger weight to the base point. Moreover, we proposed an automated method of tuning the K -nearest neighbors, thereby preventing the instability caused by a small K . Once K is auto-tuned, the resulting Laplacian matrix becomes nearly diagonally dominant, leading to a numerically stable approximation. Combining the weighted norm and auto-tuned K technique, we provided supporting numerical examples to verify the convergence of solutions. Numerically, our gRBF-FD approach performs well for i.i.d. randomly sampled data on manifolds—outperforming the GMLS approach.

Several open questions remain for future investigation. First, our study used the PHS function in the second step of the gRBF-FD approach to enable a fair comparison with the standard RBF-FD method, which also typically

uses this parameter-free function (see e.g., [23, 7, 6, 36, 71]). A natural extension is to investigate the use of bell-shaped positive-definite RBFs, such as the Gaussian and Matérn class functions, in the second step of the gRBF-FD framework. Second, a theoretical justification is lacking for the stability of gRBF-FD, although we observed numerical stability for the nearly diagonally dominant Laplacian matrices. Third, we realized that the specification of the $1/K$ weight function is important for the stability of Laplacian-type operators for randomly sampled point cloud data. A promising direction is to investigate weight function selection for gRBF-FD to solve other types of PDEs on manifolds with boundaries, extending the Euclidean-space results of [7]. This includes tackling equations such as the advection equation, along with handling derivative boundary conditions like Neumann and Robin types.

Acknowledgements This work was partially supported by the ShanghaiTech University Grant No. 2024X0303-902-01. S. J. was supported by the NSFC Grant No. 12471412 and the HPC Platform of ShanghaiTech University.

CRediT authorship contribution statement

Rongji Li: Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Haichuan Di:** Investigation, Validation, Visualization. **Shixiao Willing Jiang:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Data availability

Data will be made available on request.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

This work was partially supported by the ShanghaiTech University Grant No. 2024X0303-902-01. S. J. was supported by the NSFC Grant No. 12471412 and the HPC Platform of ShanghaiTech University.

A Proof of reproduction properties

Proof of Lemma 1. For the first property (1) of Lemma 1, we use the normalized polynomial basis functions for the GMLS regression of an arbitrary $p = \sum_{j=1}^m \bar{b}_j p_{\alpha(j)}(\hat{\theta}(\mathbf{x})) \in$

$$\mathbb{P}_{\mathbf{x}_0}^{l,d},$$

$$\sum_{k=1}^K u_{k,\Delta}(\mathbf{x}) p(\mathbf{x}_{0,k}) = \Delta_M \mathcal{I}_p \mathbf{p}_{\mathbf{x}_0}(\mathbf{x}) = \Delta_M \tilde{\mathbf{p}}(\mathbf{x}) (\tilde{\mathbf{P}}^\top \Lambda \tilde{\mathbf{P}})^{-1} \tilde{\mathbf{P}}^\top \Lambda \mathbf{p}_{\mathbf{x}_0},$$

where $\tilde{\mathbf{p}}(\mathbf{x}) = (p_{\alpha(1)}(\tilde{\boldsymbol{\theta}}(\mathbf{x})), \dots, p_{\alpha(m)}(\tilde{\boldsymbol{\theta}}(\mathbf{x}))) \in \mathbb{R}^{1 \times m}$, $\mathbf{p}_{\mathbf{x}_0} = (p(\mathbf{x}_{0,1}), \dots, p(\mathbf{x}_{0,K}))^\top \in \mathbb{R}^{K \times 1}$ and others are defined in (34). Here we notice that $\mathbf{p}_{\mathbf{x}_0} = \tilde{\mathbf{P}} \tilde{\mathbf{b}}$ with $\tilde{\mathbf{b}} = (\tilde{b}_1, \dots, \tilde{b}_m)^\top \in \mathbb{R}^{m \times 1}$. Then we arrive at

$$\sum_{k=1}^K u_{k,\Delta}(\mathbf{x}) p(\mathbf{x}_{0,k}) = \Delta_M \tilde{\mathbf{p}}(\mathbf{x}) \tilde{\mathbf{b}} = \Delta_M p(\mathbf{x}).$$

For the second property (2) of Lemma 1, we observe that

$$\begin{aligned} (u_{1,\Delta}(\mathbf{x}), \dots, u_{K,\Delta}(\mathbf{x})) &= \Delta_M \tilde{\mathbf{p}}(\mathbf{x}) (\tilde{\mathbf{P}}^\top \Lambda \tilde{\mathbf{P}})^{-1} \tilde{\mathbf{P}}^\top \Lambda \\ &= \frac{1}{D_{K,\max}^2} \Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\mathbf{p}}(\mathbf{x}) (\tilde{\mathbf{P}}^\top \Lambda \tilde{\mathbf{P}})^{-1} \tilde{\mathbf{P}}^\top \Lambda, \end{aligned}$$

where $\Delta_{\tilde{\boldsymbol{\theta}}}$ is the Laplace-Beltrami operator with respect to the normalized Monge coordinates. We notice that all above quantities $\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\mathbf{p}}(\mathbf{x})$ and $\tilde{\mathbf{P}}$ are of $O(1)$, where the constant in the big-oh is independent of the stencil diameter $D_{K,\max}$. Hence, we arrive at $\sum_{k=1}^K |u_{k,\Delta}(\mathbf{x})| \leq C_1 D_{K,\max}^{-2}$ for some constant C_1 as desired. The proof is complete.

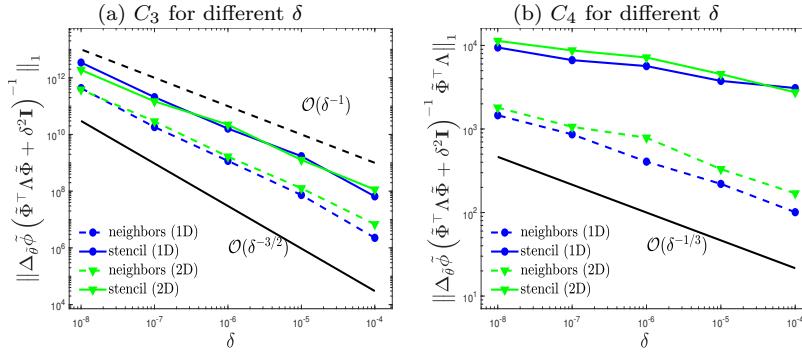


Fig. 13 1D ellipse in \mathbb{R}^2 and 2D red blood cell (RBC) in \mathbb{R}^3 using random data. Shown are (a) the maximum of $\|\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\phi}(\mathbf{x}) (\tilde{\mathbf{P}}^\top \Lambda \tilde{\mathbf{P}} + \delta^2 \mathbf{I})^{-1}\|_1$ vs. δ and (b) the maximum of $\|\Delta_{\tilde{\boldsymbol{\theta}}} \tilde{\phi}(\mathbf{x}) (\tilde{\mathbf{P}}^\top \Lambda \tilde{\mathbf{P}} + \delta^2 \mathbf{I})^{-1} \tilde{\mathbf{P}}^\top \Lambda\|_1$ vs. δ for ellipse (blue lines) and RBC (green lines). Here, given a fixed base point \mathbf{x}_0 , the maximum is taken on its K -nearest neighbors (dashed lines) and the maximum is computed from the resamplings of 200 new points in the local stencil (solid lines). We fix $K_0 = 30, N = 1600$ for ellipse and $K_0 = 40, N = 6400$ for RBC. The polynomial degree $l = 4$ and the PHS parameter $\kappa = 3$.

Partial proof and numerical verification of Lemma 3. The proof for the lemma is closely related to the norm estimates of inverses of interpolation matrices for RBFs (see e.g., [52, 4, 53, 59, 60]). Here, we use the weighted ridge regression (36) to compute the inverse for random data. We only prove the partial results and leave the remaining by numerical verification. For the first property (1), we use the normalized PHS functions for the interpolation

of an arbitrary $\varphi = \sum_{k=1}^K \tilde{c}_k \tilde{\phi}_k \in \text{Span}\{\tilde{\phi}_1, \dots, \tilde{\phi}_K\}$,

$$\begin{aligned} |\Delta_M \varphi(\mathbf{x}) - \Delta_M \mathcal{I}_\phi \varphi_{\mathbf{x}_0}(\mathbf{x})| &= |\Delta_M \tilde{\phi}(\mathbf{x}) \tilde{\mathbf{c}} - \Delta_M \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\Phi}^\top \Lambda \tilde{\Phi} \tilde{\mathbf{c}}| \\ &= |\Delta_M \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I}) \tilde{\mathbf{c}} - \Delta_M \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\Phi}^\top \Lambda \tilde{\Phi} \tilde{\mathbf{c}}| \\ &= \delta^2 |\Delta_M \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\mathbf{c}}| = \delta^2 D_{K,\max}^{-2} |\Delta_{\tilde{\theta}} \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\mathbf{c}}|, \end{aligned}$$

where $\tilde{\phi}(\mathbf{x}) = (\tilde{\phi}_1(\mathbf{x}), \dots, \tilde{\phi}_K(\mathbf{x})) \in \mathbb{R}^{1 \times K}$, $\tilde{\mathbf{c}} = (\tilde{c}_1, \dots, \tilde{c}_K)^\top \in \mathbb{R}^{K \times 1}$, and other quantites can be found in (34). Here, in the last equality, the Laplacian derivative is changed from the unnormalized θ to the normalized $\tilde{\theta}$. Hence we arrive at the result, $|\Delta_M \varphi(\mathbf{x}) - \Delta_M \mathcal{I}_\phi \varphi_{\mathbf{x}_0}(\mathbf{x})| \leq C_3 \delta^2 D_{K,\max}^{-2} \max_{1 \leq k \leq K} |\tilde{c}_k|$. In Fig. 13(a), it can be seen that $C_3 = \max_{\mathbf{x} \in S_{\mathbf{x}_0}} \|\Delta_{\tilde{\theta}} \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1}\|_1$ is between $O(\delta^{-1})$ and $O(\delta^{-3/2})$. Consequently, we have $|\Delta_M \varphi(\mathbf{x}) - \Delta_M \mathcal{I}_\phi \varphi_{\mathbf{x}_0}(\mathbf{x})| = O(\delta^{1/2} D_{K,\max}^{-2})$.

For the second property (2), we have that

$$\begin{aligned} (v_{1,\Delta}(\mathbf{x}), \dots, v_{K,\Delta}(\mathbf{x})) &= \Delta_M \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\Phi}^\top \Lambda \\ &= D_{K,\max}^{-2} \Delta_{\tilde{\theta}} \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\Phi}^\top \Lambda. \end{aligned}$$

Let $\Lambda^{1/2} \tilde{\Phi} \Lambda^{1/2} = \mathbf{U} \Sigma \mathbf{U}^\top$ be diagonalization of the symmetric matrix $\Lambda^{1/2} \tilde{\Phi} \Lambda^{1/2}$, where \mathbf{U} is orthonormal and $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_K)$ is diagonal possibly with negative eigenvalues. Then

$$\begin{aligned} (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\Phi}^\top \Lambda &= \Lambda^{1/2} (\Lambda^{1/2} \tilde{\Phi}^\top \Lambda^{1/2} \Lambda^{1/2} \tilde{\Phi} \Lambda^{1/2} + \delta^2 \Lambda)^{-1} \Lambda^{1/2} \tilde{\Phi}^\top \Lambda^{1/2} \Lambda^{1/2} \\ &= \Lambda^{1/2} (\mathbf{U} \Sigma^2 \mathbf{U}^\top + \delta^2 \Lambda)^{-1} \mathbf{U} \Sigma \mathbf{U}^\top \Lambda^{1/2} \\ &= \Lambda^{1/2} \mathbf{U} (\Sigma^2 + \delta^2 \mathbf{U}^\top \Lambda \mathbf{U})^{-1} \Sigma \mathbf{U}^\top \Lambda^{1/2}. \end{aligned}$$

We now bound the 2-norm of the above matrix,

$$\begin{aligned} \|(\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\Phi}^\top \Lambda\|_2 &\leq \|\Lambda^{1/2}\|_2 \|\mathbf{U}\|_2 \left\| (\Sigma^2 + \delta^2 \mathbf{U}^\top \Lambda \mathbf{U})^{-1} \Sigma \right\|_2 \|\mathbf{U}^\top\|_2 \|\Lambda^{1/2}\|_2 \\ &= \sqrt{\lambda_{\max} \left((\Sigma^2 + \delta^2 \mathbf{U}^\top \Lambda \mathbf{U})^{-1} \Sigma^2 (\Sigma^2 + \delta^2 \mathbf{U}^\top \Lambda \mathbf{U})^{-1} \right)} \\ &= \sqrt{\lambda_{\max} \left(|\Sigma| (\Sigma^2 + \delta^2 \mathbf{U}^\top \Lambda \mathbf{U})^{-2} |\Sigma| \right)} \\ &\leq \sqrt{\lambda_{\max} \left(|\Sigma| (\Sigma^2 + \delta^2 / K \mathbf{I})^{-2} |\Sigma| \right)} \\ &= \max_{1 \leq i \leq K} \frac{|\sigma_i|}{\sigma_i^2 + \delta^2 / K} \leq \frac{\sqrt{K}}{2\delta}, \end{aligned}$$

where we have used the fact that $\Sigma^2 + \delta^2 \mathbf{U}^\top \Lambda \mathbf{U} \succeq \Sigma^2 + \delta^2 / K \mathbf{I}$. Then we obtain that

$$\begin{aligned} \sum_{k=1}^K |v_{k,\Delta}(\mathbf{x})| &= D_{K,\max}^{-2} \|\Delta_{\tilde{\theta}} \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\Phi}^\top \Lambda\|_1 \\ &\leq D_{K,\max}^{-2} \sqrt{K} \|\Delta_{\tilde{\theta}} \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\Phi}^\top \Lambda\|_2 \leq C_4 D_{K,\max}^{-2}, \end{aligned}$$

for some $C_4 = O(1/\delta)$, where Cauchy-Schwarz inequality has been used for the first inequality. However, the above bound for C_4 is not sharp with respect to δ . As can be seen from Fig. 13(b), the maximum, $C_4 = \max_{\mathbf{x} \in S_{\mathbf{x}_0}} \|\Delta_{\tilde{\theta}} \tilde{\phi}(\mathbf{x}) (\tilde{\Phi}^\top \Lambda \tilde{\Phi} + \delta^2 \mathbf{I})^{-1} \tilde{\Phi}^\top \Lambda\|_1$, is only around $O(\delta^{-1/3})$. Numerically, we have $\sum_{k=1}^K |v_{k,\Delta}(\mathbf{x})| = O(\delta^{-1/3} D_{K,\max}^{-2})$.

B Probabilistic stencil size result

To obtain the probabilistic stencil size result, we need the following result from [15]:

Lemma 5 (*Proposition 14 in [15]*) Let $B_\delta(x)$ denote a geodesic ball of radius δ around a point $\mathbf{x} \in M$. For a sufficiently small $\delta < \iota(M)/2$ with $\iota(M)$ the injectivity radius of M , we have

$$\text{Vol}(B_\delta(\mathbf{x})) \geq C_d \delta^d,$$

where C_d is a constant depending only on the dimension d of the manifold.

Proof of Lemma 4. Let $B_\delta(\mathbf{x}_0)$ denote a geodesic ball of radius δ around the base point $\mathbf{x}_0 \in \mathbf{X}_M \subset M$. Let $B_\delta^c(\mathbf{x}_0) = M \setminus B_\delta(\mathbf{x}_0)$ be the complement of $B_\delta(\mathbf{x}_0)$. For a point $\mathbf{x}_0 \in \mathbf{X}_M$ sampled from M with a density $q(\mathbf{x})$, it holds that

$$\mathbb{P}_{\mathbf{X}_M \sim Q}(\mathbf{x}_0 \in \mathbf{X}_M \cap B_\delta^c(\mathbf{x}_0)) = 1 - \int_{B_\delta(\mathbf{x}_0)} q(\mathbf{x}) dV(\mathbf{x}).$$

Let $d_g(\mathbf{x}_{0,k}, \mathbf{x}_0)$ be the geodesic distance from $\mathbf{x}_{0,k}$ to the base \mathbf{x}_0 for $k = 1, \dots, K$. For small δ , we can arrive at

$$\begin{aligned} \mathbb{P}_{\mathbf{X}_M \sim Q} \left(\max_{k \in \{1, \dots, K\}} d_g(\mathbf{x}_{0,k}, \mathbf{x}_0) > \delta \right) &= \mathbb{P}(\text{there are at least } N - K \text{ points of } \mathbf{X}_M \text{ in } B_\delta^c(\mathbf{x}_0)) \\ &\leq \mathbb{P}(\mathbf{x}_{0,K+1} \in B_\delta^c, \dots, \mathbf{x}_{0,N} \in B_\delta^c) \\ &= \left(1 - \int_{B_\delta(\mathbf{x}_0)} q(\mathbf{x}) dV(\mathbf{x}) \right)^{N-K}. \end{aligned}$$

where $\mathbf{x}_{0,K+1}, \dots, \mathbf{x}_{0,N}$ are the $N - K$ points out of the stencil $S_{\mathbf{x}_0}$. Using the assumption $q \geq q_{\min}$ and Lemma 5 from [15], we obtain that

$$\begin{aligned} \mathbb{P}_{\mathbf{X}_M \sim Q} \left(\max_{k \in \{1, \dots, K\}} d_g(\mathbf{x}_{0,k}, \mathbf{x}_0) > \delta \right) &\leq (1 - q_{\min} \text{Vol}(B_\delta(\mathbf{x}_0)))^{N-K} \leq \left(1 - q_{\min} C_d \delta^d \right)^{N-K} \\ &\leq \exp(-q_{\min} C_d (N - K) \delta^d), \end{aligned}$$

where C_d is a constant depending on the dimension d . Using the fact that $\|\theta(\mathbf{x}_{0,k}) - \theta(\mathbf{x}_0)\| \leq d_g(\mathbf{x}_{0,k}, \mathbf{x}_0)$, we arrive at

$$\mathbb{P}_{\mathbf{X}_M \sim Q}(R_{K,\max}(\mathbf{x}_0) > \delta) \leq \exp(-q_{\min} C_d (N - K) \delta^d).$$

Moreover, taking $\exp(-q_{\min} C_d (N - K) \delta^d) = \frac{1}{N}$, we obtain that

$$\delta = \left(\frac{\log N}{q_{\min} C_d (N - K)} \right)^{\frac{1}{d}} = O \left(\left(\frac{\log N}{N} \right)^{\frac{1}{d}} \right).$$

Since $R_{K,\max}(\mathbf{x}_0) \leq D_{K,\max}(\mathbf{x}_0) \leq C_K R_{K,\max}(\mathbf{x}_0)$, we arrive at

$$R_{K,\max}(\mathbf{x}_0) \leq \delta = O \left(\left(\frac{\log N}{N} \right)^{\frac{1}{d}} \right), \quad D_{K,\max} \leq C_K \delta = O \left(\left(\frac{\log N}{N} \right)^{\frac{1}{d}} \right),$$

with probability higher than $1 - \frac{1}{N}$.

Subsequently, we compute the expectation and the standard deviation of $D_{K,\max}(\mathbf{x}_0)$. Let $p_0(z)$ be the density function of the random variable $Z := D_{K,\max}(\mathbf{x}_0)$. Let a_{\max} be

the diameter of the manifold M , that is, $a_{\max} = \max_{\mathbf{x}, \mathbf{y} \in M} \|\mathbf{x} - \mathbf{y}\|$. Then, the expectation can be bounded by

$$\begin{aligned} ED_{K,\max}(\mathbf{x}_0) &= \int_0^{a_{\max}} z p_0(z) dz = \int_0^{C_K \delta} z p_0(z) dz + \int_{C_K \delta}^{a_{\max}} z p_0(z) dz \\ &\leq C_K \delta + a_{\max} P_{\mathbf{X}_M \sim Q}(D_{K,\max}(\mathbf{x}_0) > C_K \delta) \\ &\leq C_K \delta + a_{\max} \exp(-q_{\min} C_d(N - K) \delta^d). \end{aligned}$$

Taking $\delta = \left(\frac{\log N}{dq_{\min} C_d(N - K)} \right)^{1/d}$, then the expectation becomes

$$ED_{K,\max}(\mathbf{x}_0) \leq C_K \left(\frac{\log N}{dq_{\min} C_d(N - K)} \right)^{\frac{1}{d}} + a_{\max} N^{-\frac{1}{d}} = O \left(\left(\frac{\log N}{N} \right)^{\frac{1}{d}} \right).$$

The standard deviation can be calculated similarly,

$$\sigma(D_{K,\max}(\mathbf{x}_i)) \leq (ED_{K,\max}^2(\mathbf{x}_0))^{1/2} \leq [C_K^2 \delta^2 + a_{\max}^2 \exp(-q_{\min} C_d(N - K) \delta^d)]^{1/2}.$$

Taking here $\delta = \left(\frac{2 \log N}{dq_{\min} C_d(N - K)} \right)^{1/d}$, then the standard deviation becomes

$$\sigma(D_{K,\max}(\mathbf{x}_i)) \leq \left[C_K^2 \left(\frac{2 \log N}{dq_{\min} C_d(N - K)} \right)^{\frac{2}{d}} + a_{\max}^2 N^{-\frac{2}{d}} \right]^{1/2} = O \left(\left(\frac{\log N}{N} \right)^{\frac{1}{d}} \right).$$

The proof is complete.

References

1. The stanford 3d scanning repository URL <http://graphics.stanford.edu/data/3Dscanrep/>
2. Ahlberg, J.H., Nilson, E.N.: Convergence properties of the spline fit. *J. Soc. Ind. Appl. Math.* **11**(1), 95–104 (1963)
3. Álvarez, D., González-Rodríguez, P., Kindelan, M.: A local radial basis function method for the laplace–beltrami operator. *J. Sci. Comput.* **86** (2021)
4. Ball, K., Sivakumar, N., Ward, J.D.: On the sensitivity of radial basis interpolation to minimal data separation distance. *J. Approx. Theory* **8**(4), 401–426 (1992)
5. Barreira, R., Elliott, C.M., Madzvamuse, A.: The surface finite element method for pattern formation on evolving biological surfaces. *J. Math. Biol.* **63**(6), 1095–1119 (2011)
6. Bayona, V., Flyer, N., Fornberg, B.: On the role of polynomials in rbf-fd approximations: Iii. behavior near domain boundaries. *J. Comput. Phys.* **380**, 378–399 (2019)
7. Bayona, V., Flyer, N., Fornberg, B., Barnett, G.A.: On the role of polynomials in rbf-fd approximations: II. numerical solution of elliptic pdes. *J. Comput. Phys.* **332**, 257–273 (2017)
8. Belytschko, T., Lu, Y.Y., Gu, L.: Element-free galerkin methods. *Int. J. Numer. Methods Eng.* **37**(2), 229–256 (1994)
9. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: Proc. of the 27th Annu. Conf. on Computer Graphics and Interact. Tech., pp. 417–424 (2000)
10. Bollig, E.F., Flyer, N., Erlebacher, G.: Solution to pdes using radial basis function finite-differences (rbf-fd) on multiple gpus. *J. Comput. Phys.* **231**(21), 7133–7151 (2012)
11. Cecil, T., Qian, J., Osher, S.: Numerical methods for high dimensional hamilton–jacobi equations using radial basis functions. *J. Comput. Phys.* **196**(1), 327–347 (2004)
12. Chen, C., Karageorghis, A., Dou, F.: A novel rbf collocation method using fictitious centres. *Appl. Math. Lett.* **101**, 106069 (2020)
13. Chen, J.S., Hillman, M., Chi, S.W.: Meshfree methods: progress made after 20 years. *J. Eng. Mech.* **143**(4), 04017001 (2017)

14. Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., Ranzuglia, G., et al.: Meshlab: an open-source mesh processing tool. *Eurographics Italian Chapter Conference, Salerno, Italy* **2008**, 129–136 (2008)
15. Croke, C.B.: Some isoperimetric inequalities and eigenvalue estimates. In: *Annales scientifiques de l’École normale supérieure*, vol. Ser. 4, 13, pp. 419–435 (1980)
16. Delengov, V.: Computing eigenmodes of elliptic operators on manifolds using radial basis functions. Ph.D. Thesis (2018)
17. Elliott, C.M., Stinner, B.: Modeling and computation of two phase geometric biomembranes using surface finite elements. *J. Comput. Phys.* **229**(18), 6585–6612 (2010)
18. Fasshauer, G.E.: Solving partial differential equations by collocation with radial basis functions. In: *Proceedings of Chamonix*, vol. 1997, pp. 1–8. Vanderbilt University Press Nashville, TN (1996)
19. Fasshauer, G.E., McCourt, M.J.: Stable evaluation of gaussian radial basis function interpolants. *SIAM J. Sci. Comput.* **34**(2), A737–A762 (2012)
20. Fedoseyev, A., Friedman, M., Kansa, E.: Improved multiquadric method for elliptic partial differential equations via pde collocation on the boundary. *Comput. Mathe. Appl.* **43**(3–5), 439–455 (2002)
21. Flyer, N., Fornberg, B.: Radial basis functions: Developments and applications to planetary scale flows. *Comput. Fluids* **46**(1), 23–32 (2011)
22. Flyer, N., Fornberg, B.: Solving pdes with radial basis functions. *Acta Numer.* **24**, 215–258 (2015)
23. Flyer, N., Fornberg, B., Bayona, V., Barnett, G.A.: On the role of polynomials in rbf-fd approximations: I. interpolation and accuracy. *J. Comput. Phys.* **321**, 21–38 (2016)
24. Flyer, N., Wright, G.B., Fornberg, B.: Radial basis function-generated finite differences: A mesh-free method for computational geosciences. In: *Handbook of geomathematics*, pp. 1–30. Springer (2013)
25. Franke, C., Schaback, R.: Solving partial differential equations by collocation using radial basis functions. *Appl. Math. Comput.* **93**(1), 73–82 (1998)
26. Fuselier, E.J., Wright, G.B.: Scattered data interpolation on embedded submanifolds with restricted positive definite kernels: Sobolev error estimates. *SIAM J. Numer. Anal.* **50**(3), 1753–1776 (2012)
27. Fuselier, E.J., Wright, G.B.: A high-order kernel method for diffusion and reaction-diffusion equations on surfaces. *J. Sci. Comput.* **56**(3), 535–565 (2013)
28. Gilani, F., Harlim, J.: Approximating solutions of linear elliptic pde’s on a smooth manifold using local kernel. *J. Comput. Phys.* **395**, 563 – 582 (2019)
29. Glaubitz, J., Nordström, J., Öffner, P.: Energy-stable global radial basis function methods on summation-by-parts form. *J. Sci. Comput.* **98**(1), 30 (2024)
30. Gross, B.J., Trask, N., Kuberry, P., Atzberger, P.J.: Meshfree methods on manifolds for hydrodynamic flows on curved surfaces: A generalized moving least-squares (gmls) approach. *J. Comput. Phys.* **409**, 109340 (2020)
31. Halada, T., Yaskovets, S., Singh, A., Benes, L., Suchde, P., Sbalzarini, I.F.: An overview of meshfree collocation methods. arXiv preprint arXiv:2509.20056 (2025)
32. Harlim, J., Jiang, S.W., Peoples, J.W.: Radial basis approximation of tensor fields on manifolds: from operator estimation to manifold learning. *J. Mach. Learn. Res.* **24**(345), 1–85 (2023)
33. Iske, A.: On the approximation order and numerical stability of local lagrange interpolation by polyharmonic splines. In: *Modern Developments in Multivariate Approximation: 5th International Conference, Witten-Bommerholz (Germany), September 2002*, pp. 153–165. Springer (2003)
34. Jiang, S.W., Harlim, J.: Ghost point diffusion maps for solving elliptic pdes on manifolds with classical boundary conditions. *Commun. Pure Appl. Math.*, **76**(2), 337–405 (2023)
35. Jiang, S.W., Li, R., Yan, Q., Harlim, J.: Generalized finite difference method on unknown manifolds. *J. Comput. Phys.* **502**, 112812 (2024)
36. Jones, A.M., Bosler, P.A., Kuberry, P.A., Wright, G.B.: Generalized moving least squares vs. radial basis function finite difference methods for approximating surface derivatives. *Comput. Math. Appl.* **147**, 1–13 (2023)
37. Kansa, E.J.: Multiquadratics—a scattered data approximation scheme with applications to computational fluid-dynamics—ii solutions to parabolic, hyperbolic and elliptic partial differential equations. *Comput. Math. Appl.* **19**(8–9), 147–161 (1990)

38. Lancaster, P., Salkauskas, K.: Surfaces generated by moving least squares methods. *Mathematics of computation* **37**(155), 141–158 (1981)
39. Larsson, E., Fornberg, B.: A numerical study of some radial basis function based solution methods for elliptic pdes. *Comput. Mathe. Appl.* **46**(5-6), 891–902 (2003)
40. Le Borne, S., Leinen, W.: Guidelines for rbf-fd discretization: numerical experiments on the interplay of a multitude of parameter choices. *J. Sci. Comput.* **95**(1), 8 (2023)
41. Lehto, E., Shankar, V., Wright, G.B.: A radial basis function (rbf) compact finite difference (fd) scheme for reaction-diffusion equations on surfaces. *SIAM J. Sci. Comput.* **39**(5), A2129–A2151 (2017)
42. Li, R., Yan, Q., Jiang, S.W.: Generalized moving least-squares for solving vector-valued pdes on unknown manifolds. arXiv e-prints pp. arXiv-2406 (2024)
43. Li, Z., Shi, Z.: A convergent point integral method for isotropic elliptic equations on a point cloud. *Multiscale Model. Simul.* **14**(2), 874–905 (2016)
44. Li, Z., Shi, Z., Sun, J.: Point integral method for solving poisson-type equations on manifolds from point clouds with convergence guarantees. *Commun. Comput. Phys.* **22**(1), 228–258 (2017)
45. Liang, J., Lai, R., Wong, T.W., Zhao, H.: Geometric understanding of point clouds using laplace-beltrami operator. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 214–221. IEEE (2012)
46. Liang, J., Zhao, H.K.: Solving partial differential equations on point clouds. *SIAM J. Sci. Comput.* **35**(3), A1461–A1486 (2013)
47. Lipman, Y.: Stable moving least-squares. *J. Approx. Theory* **161**(1), 371–384 (2009)
48. Lorensen, W., Cline, H.: Marching cubes: a high resolution 3d surface construction algorithm. *ACM Siggraph Comput. Graph.* **21**(4), 163–169 (1987)
49. Ma, Z., Li, X., Chen, C.: Ghost point method using rbfs and polynomial basis functions. *Appl. Math. Lett.* **111**, 106618 (2021)
50. Mirzaei, D., Schaback, R., Dehghan, M.: On generalized moving least squares and diffuse derivatives. *IMA J. Numer. Anal.* **32**(3), 983–1000 (2012)
51. Monge, G.: Application de l'analyse à la géométrie à l'usage de l'Ecole impériale polytechnique. Veuve Bernard (1809)
52. Narcowich, F.J., Ward, J.D.: Norms of inverses and condition numbers for matrices associated with scattered data. *J. Approx. Theory* **64**(1), 69–94 (1991)
53. Narcowich, F.J., Ward, J.D.: Norm estimates for the inverses of a general class of scattered-data radial-function interpolation matrices. *J. Approx. Theory* **69**(1), 84–109 (1992)
54. Nayroles, B., Touzot, G., Villon, P.: Generalizing the finite element method: diffuse approximation and diffuse elements. *Comput. Mech.* **10**(5), 307–318 (1992)
55. Nitschke, I., Reuther, S., Voigt, A.: Liquid crystals on deformable surfaces. *Proc. Roy. Soc. Edinburgh Sect. A* **476**(2241), 20200313 (2020)
56. Petras, A., Ling, L., Ruuth, S.J.: An rbf-fd closest point method for solving pdes on surfaces. *J. Comput. Phys.* **370**, 43–57 (2018)
57. Piret, C.: The orthogonal gradients method: A radial basis functions method for solving partial differential equations on arbitrary surfaces. *J. Comput. Phys.* **231**(14), 4662–4675 (2012)
58. Pressley, A.N.: Elementary differential geometry. Springer Science & Business Media (2010)
59. Schaback, R.: Lower bounds for norms of inverses of interpolation matrices for radial basis functions. *J. Approx. Theory* **79**(2), 287–306 (1994)
60. Schaback, R.: Error estimates and condition numbers for radial basis function interpolation. *Adv. Comput. Math.* **3**(3), 251–264 (1995)
61. Shankar, V., Wright, G.B., Kirby, R.M., Fogelson, A.L.: A radial basis function (rbf)-finite difference (fd) method for diffusion and reaction-diffusion equations on surfaces. *J. Sci. Comput.* **63**(3), 745–768 (2015)
62. Shaw, S.B.: Radial basis function finite difference approximations of the Laplace-Beltrami operator, thesis. Boise State University Graduate College (2019)
63. Shu, C., Ding, H., Yeo, K.: Local radial basis function-based differential quadrature method and its application to solve two-dimensional incompressible navier–stokes equations. *Comput. Methods Appl. Mech. Eng.* **192**(7-8), 941–954 (2003)

-
- 64. Suchde, P., Kuhnert, J.: A meshfree generalized finite difference method for surface pdes. *Comput. Math. Appl.* **78**(8), 2789–2805 (2019)
 - 65. Tillenius, M., Larsson, E., Lehto, E., Flyer, N.: A scalable rbf-fd method for atmospheric flow. *J. Comput. Phys.* **298**, 406–422 (2015)
 - 66. Tolstykh, A.I., Shirobokov, D.: On using radial basis functions in a “finite difference mode” with applications to elasticity problems. *Comput. Mech.* **33**(1), 68–79 (2003)
 - 67. Varah, J.M.: A lower bound for the smallest singular value of a matrix. *Linear Algebra Appl.* **11**(1), 3–5 (1975)
 - 68. Walker, S.: Felicity, a matlab/c++ toolbox for developing finite element methods and simulation modeling. *SIAM J. Sci. Comput.* **40**(2), C234–C257 (2018)
 - 69. Wendland, H.: Scattered Data Approximation. Cambridge University Press (2005)
 - 70. Wright, G.B.: Radial basis function interpolation: numerical and analytical developments. University of Colorado at Boulder (2003)
 - 71. Wright, G.B., Jones, A., Shankar, V.: Mgm: a meshfree geometric multilevel method for systems arising from elliptic equations on point cloud surfaces. *SIAM J. Sci. Comput.* **45**(2), A312–A337 (2023)
 - 72. Yan, Q., Jiang, S.W., Harlim, J.: Kernel-based methods for solving time-dependent advection-diffusion equations on manifolds. *J. Sci. Comput.* **94**(1), 5 (2023)
 - 73. Zhao, H.K., Osher, S., Fedkiw, R.: Fast surface reconstruction using the level set method. In: Proc. IEEE Workshop Var. Level Set Methods Comput. Vis., pp. 194–201. IEEE (2001)