Moving least-square reproducing kernel methods
(I) Methodology and convergence

Wing-Kam Liu[1,*], Shaofan Li[2], Ted Belytschko[3]

Department of Mechanical Engineering, Robert R. McCormick School of Engineering and Applied Science, The Technological Institute, Northwestern University, Evanston, IL 60208-3019, USA

Abstract

This paper formulates the moving least-square interpolation scheme in a framework of the so-called moving least-square reproducing kernel (MLSRK) representation. In this study, the procedure of constructing moving least square interpolation function is facilitated by using the notion of reproducing kernel formulation, which, as a generalization of the early discrete approach, establishes a continuous basis for a partition of unity. This new formulation possesses the quality of simplicity, and it is easy to implement. Moreover, the reproducing kernel formula proposed is not only able to reproduce any mth order polynomial exactly on an irregular particle distribution, but also serves as a projection operator that can approximate any smooth function globally with an optimal accuracy.

In this contribution, a generic m-consistency relation has been found, which is the essential property of the MLSRK approximation. An interpolation error estimate is given to assess the convergence rate of the approximation. It is shown that for sufficiently smooth function the interpolant expansion in terms of sampled values will converge to the original function in the Sobolev norms. As a meshless method, the convergence rate is measured by a new control variable—dilation parameter ϱ of the window function, instead of the mesh size h as usually done in the finite element analysis. To illustrate the procedure, convergence has been shown for the numerical solution of the second-order elliptic differential equations in a Galerkin procedure invoked with this interpolant. In the numerical example, a two point boundary problem is solved by using the method, and an optimal convergence rate is observed with respect to various norms.

1. Introduction

Recently, there has been keen interest in developing meshless approximations for Galerkin procedures to solve partial differential equations. Several versions of the method have been proposed; among them are Smooth Particle Hydrodynamics (SPH) by Gingold and Monaghan [1], Monaghan [2]; Diffuse Element Method by Nayroles et al. [3]; Element Free Galerkin Method by Belytschko et al. [4,5] and Reproducing Kernel Particle Method by Liu et al. [6], Liu and Chen [7], Liu et al. [8] and Liu [9]. It seems to us that the moving least square interpolant based meshless method has special appeals, due to its unique inviting features which are incomparable for the other numerical methods in many aspects. The major attractions include:

(1) It is a mesh-free particle method; consequently, the usual, formidable burden of mesh generation of traditional finite element method is ameliorated, if it is not disposed of completely.

(2) It is a smooth interpolation procedure; by using this method, one can easily achieve a global conforming $C^m(\Omega)$ interpolation field of desired order, which is very difficult to obtain by regular finite element method when $m \geq 2$.

---

* Corresponding author.
[1] Professor.
[2] Graduate Assistant.
[3] Walter P. Murphy Professor.

(3) It has an excellent localization capability in both frequency spectrum and spatial domain, which is generically suitable to the task of multiple scale analysis and multi-grid computation.

Because of its diverse characteristics, it has attracted many researchers to use the method to solve the practical problems for particular benefit. Consequently, different authors tend to label it with different names according to a specific feature which might be their personal preferences.

An early, but lucid formulation and documentation of the moving least square approximation is due to Lancaster and Salkauskas [10]. Nevertheless, Lancaster et al. never apply the method in large scale computation nor in a Galerkin procedure; therefore, it was impossible for them to foresee the problems which occur in practical implementation. In the actual computations, the main concerns are: (1) Convergence: how fast will the numerical solution converge, if it converges at all. (2) Computational efficiency: how long does it take to compute a practical problem? These questions remain to be answered.

Thus, it is our intention to reexamine the formulation from the computational perspective, and to systematically lay the foundation for this special partition of unity—a finite support kernel method.

We start with formulating a different version of moving least-square approximation. Based on the new formulation proposed here, it is revealed that the moving least-square interpolant has an unique inner structure—a m-consistency structure. Moreover, the formulation proposed here is constructed on a continuous base with the emphasis on the notion of reproducing kernel representation. The meaning of '*Moving Least-Square Reproducing Kernel Method*' (MLSRK) is two-fold: first, the shape function are generated by a moving least-square process; secondly, the interpolant of this sort contains a reproducing kernel, which can '*reproduce*' any smooth function accurately in a global least square sense; in particular, it reproduces the m order polynomial exactly, if the generating polynomial basis is order m. Furthermore, by choosing the window function appropriately, the MLSRK may be able to achieve an optimal accuracy in the frequency spectrum approximation as a psuedo-spectral method. As observed by many researchers, one of the impediments in using spectral methods, such as wavelet methods, to solve partial differential equations is the difficulty in handling irregular boundary and the associated boundary conditions. Since the MLSRK method is quite flexible in dealing with irregular boundary and general boundary conditions, this difficulty is much alleviated. More precisely speaking, by carefully selecting window function and dilation parameter, not only do the shape functions constructed by this method form a complete basis of a finite dimensional Hilbert space as conventional finite element shape function, but also the Fourier transform of the shape functions can stay within a given range of frequency spectrum, which naturally makes it as a very special candidate to solve evolution type of PDE in Galerkin procedures.

One of the inadequacies of early moving least-square approximation is the lack of rigorous mathematical foundation.[4] Therefore, the main portion of this paper is devoted to the convergence study; it is shown that the interpolation error decreases as the density of particle distribution increases. In fact, a better interpolation estimation could be achieved than that of the finite element method (see [11]). There is a major distinction between the convergence in this paper and convergence for the conventional finite element method as well as the finite difference method. In this paper, the convergence rate is controlled and measured by the dilation parameter ϱ of the window function, whereas in the finite element method convergence rate is measured by the mesh size h. Although the mesh size h and the dilation parameter ϱ of the window function are intimately related, the dilation parameter ϱ explicitly carries the information about the resolution of numerical solutions, and is certainly a better control parameter than the maximum distance between two particles. This notion of convergence has been widely adopted in the wavelet analysis [12]. Besides, in a random meshless particle distribution, to define or to measure the maximum distance between the two adjacent particles is not a trivial task at all.

In last part of the paper, we have shown that, for the second-order elliptic partial differential equation, the numerical solution converges to the exact solution in the Sobolev norm. A numerical example has been carried out to demonstrated the convergence properties.

---

[4] While finishing this study, we received a preprint from Professor J. Tinsley Oden [13]. A mathematical analysis is presented for a similar method—*hp*-clouds method.

2. Moving least-square reproducing kernel interpolant

To expedite the presentation, multi-index notation is adopted throughout the paper. If $\alpha := (\alpha_1, \ldots, \alpha_n)$ is an n-tuple of nonnegative integers α_j, we call α a multi-index, and its length is defined as

$$|\alpha| := \sum_{i=1}^{n} \alpha_i \tag{2.1}$$

We then denote

$$\varrho^{\alpha} := \varrho_1^{\alpha_1} \varrho_2^{\alpha_2} \cdots \varrho_n^{\alpha_n} \tag{2.2}$$

and the αth (Fréchet) derivative of function u as

$$D^{\alpha} u := \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_n}^{\alpha_n} u \tag{2.3}$$

2.1. Moving least-square reproducing formula

Let $u(x)$ be a sufficiently smooth function[5] that is defined on a simply connected open set $\Omega \in \mathbb{R}^n$. Then, for a fixed point $\bar{x} \in \Omega$, one should always be able to approximate $u(x)$ by a polynomial series locally according to the Stone–Weierstrass theorem [14]. Thus, we can define a local function,

$$u^l(x, \bar{x}) := \begin{cases} u(x), & \forall x \in B(\bar{x}) \\ 0, & \text{otherwise} \end{cases} \tag{2.4}$$

where the open sphere $B(\bar{x})$ is defined as

$$B(\bar{x}) := \{ x \mid |x - \bar{x}| < R, x \in \bar{\Omega} \} \tag{2.5}$$

If the function $u(x)$ is smooth enough as assumed, there exists a local operator, such that

$$u^l(x, \bar{x}) \cong L_{\bar{x}} u(x) := \sum_{i=1}^{\ell} P_i \left(\frac{x - \bar{x}}{\varrho} \right) d_i(\bar{x})$$

$$= P \left(\frac{x - \bar{x}}{\varrho} \right) d(\bar{x}) \tag{2.6}$$

where the operator $L_{\bar{x}}$ is a mapping

$$L_{\bar{x}} : C^0(B(\bar{x})) \mapsto C^m(B(\bar{x})) \tag{2.7}$$

and

$$d^t(x) := \{ d_1, d_2, \ldots, d_{\ell} \}(x) \tag{2.8}$$

$$P(x) := \{ P_1, P_2, \ldots, P_l \}(x) ,$$

$$P_i(x) := (x/\varrho)^{\alpha^i} = \prod_{j=1}^{n} \left(\frac{x_j}{\varrho_j} \right)^{\alpha^i_j} \tag{2.9}$$

with $P_1(x) = 1$.

REMARK 2.1. The formulation proposed here differs slightly with that of Lancaster and Salkauskas [10]. In their original paper, Lancaster and Salkauskas constructed the local approximation function $u^l(x)$ as

---

[5] By this, we mean that $u(x) \in C^0(\bar{\Omega})$ at least.

$$u^I(x, \bar{x}) \cong L_{\bar{x}} u(x) := \sum_{i=1}^{\ell} P_i(x) d_i(\bar{x})$$

$$= P(x) d(\bar{x}) \tag{2.10}$$

One can tell the difference between the two by comparing Eq. (2.6) with (2.10).

Since the polynomial series is finite, there exists a residual error distribution r_ϱ over the ball $B(\bar{x})$,

$$r_\varrho(x, \bar{x}) := u^I(x) - P\left(\frac{x - \bar{x}}{\varrho}\right) d(\bar{x}), \quad x \in B(\bar{x}) \tag{2.11}$$

A functional associated with this residual is defined as

$$J(d(\bar{x})) := \int_{B(\bar{x})} r_\varrho^2(x, \bar{x}) \frac{C_n}{\varrho^n} \Phi\left(\frac{x - \bar{x}}{\varrho}\right) dB \tag{2.12}$$

A similar functional has been given by Belytschko et al. [5]. The window function Φ is chosen in such a way that $\text{supp}\{\Phi_\varrho(x - \bar{x})\} \subset B(\bar{x})$, i.e. for fixed $\bar{x}$

$$\Phi_\varrho(x - \bar{x}) := \frac{C_n}{\varrho^n} \Phi\left(\frac{x - \bar{x}}{\varrho}\right) = \begin{cases} >0 & x \in \text{supp}\{\Phi_\varrho(x - \bar{x})\} \\ 0 & x \notin \{\Phi_\varrho(x - \bar{x})\} \end{cases} \tag{2.13}$$

For rectangular compact support in $\mathbb{R}^n$, $C_n = 1$; and for 'hyper-spherical' compact support, $C_n \neq 1$ generally. For instance, $C_2 = 15/7\pi$, if one uses the cubic spline window function. In the sequel, for the sake of simplicity, we only consider the case $C_n = 1$.

The expression (2.13) automatically guarantees that $J(d(\bar{x}))$ is positive definite, though, for positive definiteness, (2.13) may not be necessary. As a matter of fact, condition (2.13) can further be relaxed, but for the moment, we assume that the condition (2.13) always hold. Thus, Eq. (2.12) can also be viewed as a special L_2 norm of the residual, i.e.

$$J(d(\bar{x})) = |||r_\varrho|||^2 \tag{2.14}$$

which is endowed by the inner product

$$\langle f | \Phi_\varrho | g \rangle_y := \int_{\Omega_y} f(y - x) g(y - x) \Phi_\varrho(y - x) \, d\Omega_y \tag{2.15}$$

and

$$|||f||| := \sqrt{\langle f | \Phi_\varrho | f \rangle_y} \tag{2.16}$$

A similar inner product was proposed in the *discrete* form by both Lancaster and Salkauskas [10] and Duarte and Oden [13]. However, in the discrete case, the positive definiteness of the quadratic summation is not automatically guaranteed unless certain prerequisites on the particle distribution are met.

By minimizing the quadratic functional $J(d(\bar{x}))$, one can obtain the following normal equations

$$\int_{B(\bar{x})} P^t\left(\frac{x - \bar{x}}{\varrho}\right)\left(u^I(x) - P\left(\frac{x - \bar{x}}{\varrho}\right) d(\bar{x})\right) \Phi_\varrho(x - \bar{x}) \, dB = 0 \tag{2.17}$$

Since $\text{supp}\{\Phi_\varrho(x - \bar{x})\} \subseteq \bar{B}$, the integral (2.17) can be extended over the whole domain

$$\int_{\Omega_x} P^t\left(\frac{x - \bar{x}}{\varrho}\right)\left(u(x) - P\left(\frac{x - \bar{x}}{\varrho}\right) d(\bar{x})\right) \Phi_\varrho(x - \bar{x}) \, d\Omega_x = 0, \tag{2.18}$$

which leads to

$$\left(\int_{\Omega_x} P^t\left(\frac{x - \bar{x}}{\varrho}\right) \Phi_\varrho(x - \bar{x}) P\left(\frac{x - \bar{x}}{\varrho}\right) d\Omega_x\right) d(\bar{x}) = \int_{\Omega_x} P^t\left(\frac{x - \bar{x}}{\varrho}\right) u(x) \Phi_\varrho(x - \bar{x}) \, d\Omega_x. \tag{2.19}$$

To this end, one can define the so-called moment matrix $M(\bar{x})$ as follows [6],

$$M(\bar{x}) := \int_{\Omega_x} P^t\left(\frac{x-\bar{x}}{\varrho}\right) \Phi_\varrho(x-\bar{x}) P\left(\frac{x-\bar{x}}{\varrho}\right) d\Omega_x$$

(2.20)

or simply

$$M(\bar{x}) := \langle P^t | \Phi_\varrho | P \rangle_x = \begin{pmatrix} \langle P_1|\Phi_\varrho|P_1\rangle & \langle P_1|\Phi_\varrho|P_2\rangle & \cdots & \langle P_1|\Phi_\varrho|P_m\rangle \\ \langle P_2|\Phi_\varrho|P_1\rangle & \langle P_2|\Phi_\varrho|P_2\rangle & \cdots & \langle P_2|\Phi_\varrho|P_m\rangle \\ \vdots & & \ddots & \vdots \\ \langle P_m|\Phi_\varrho|P_1\rangle & \langle P_m|\Phi_\varrho|P_2\rangle & \cdots & \langle P_m|\Phi_\varrho|P_m\rangle \end{pmatrix}.$$

(2.21)

Clearly, the moment matrix M is a *Gram* matrix. Since $P_i(x)$, $(i = 0, 1, \ldots, \ell)$ are linearly independent, the determinant of M is always positive (see [15]),

$$\Gamma(P_1, P_2, \ldots, P_\ell) := \det(M) > 0.$$

(2.22)

Thus, $M(\bar{x})$ is always invertible; therefore, the unknown vector $d(\bar{x})$ is uniquely determined,

$$d(\bar{x}) = M^{-1}(\bar{x}) \int_{\Omega_x} P^t\left(\frac{x-\bar{x}}{\varrho}\right) u(x) \Phi_\varrho(x-\bar{x}) d\Omega_x$$

(2.23)

Substituting (2.23) back into (2.6) and noticing that the argument x in (2.23) is a dummy variable, one can rewrite the local approximation formula as

$$u^l(x, \bar{x}) \cong L_{\bar{x}} u(x)$$

$$:= P\left(\frac{x-\bar{x}}{\varrho}\right) d(\bar{x})$$

$$= P\left(\frac{x-\bar{x}}{\varrho}\right) M^{-1}(\bar{x}) \int_{\Omega_y} P^t\left(\frac{y-\bar{x}}{\varrho}\right) u(y) \Phi_\varrho(y-\bar{x}) d\Omega_y, \quad \forall x \in B(\bar{x})$$

(2.24)

So far, the manipulation is the standard weighted least-square procedure. In order to extend Eq. (2.24) to the whole domain, the so-called 'moving' procedure is introduced. The central idea of the moving least-square method is that one can achieve a global approximation by going through a '*mysterious moving*' process. The formal procedure consists of two steps: first, one takes an arbitrary fixed point $\bar{x} \in \Omega$ and forms a local approximation formula, i.e. (2.24), which is only valid in a local region $B(\bar{x}) \subseteq \Omega$, as shown above; second, since the fixed point $\bar{x}$ is arbitrary, it can be any point $x \in \Omega$, therefore, one can let it 'move' over the whole domain, $\bar{x} \to x$, which will lead to a global approximation of $u(x)$.

Before sweeping Eq. (2.24) through the whole domain, a global approximation operator G is introduced

$$u(x) \cong Gu(x) \quad \forall x \in \bar{\Omega}$$

(2.25)

where

$$G : C^0(\bar{\Omega}) \mapsto C^m(\bar{\Omega})$$

(2.26)

More precisely, the global approximation operator G is defined in such a way that it is the globalization of the local approximate operator $L_{\bar{x}}$, and the globalization is realized through the moving process, viz.

$$Gu(x) := \lim_{\bar{x} \to x} L_{\bar{x}} u(x) \quad \forall x \in \bar{\Omega}$$

(2.27)

At the final phase of the moving process, one obtains $Gu(x) := L_x u(x)$, i.e.

$$Gu(x) = P(0)M^{-1}(x) \int_{\Omega_y} P^t\left(\frac{y-x}{\varrho}\right) u(y) \Phi_\varrho(y-x) \, d\Omega_y$$

(2.28)

Eq. (2.28) is a global approximation formula, which is named by Liu et al. [6] as the 'reproducing kernel' formulation.

To this end, one can define the components of the moments as

$$M_{ij}(x) := \int_{\Omega(x)} P_i(z) P_j(z) \Phi(z) \, d\Omega_z, \quad i, j = 1, \ldots, \ell$$

(2.29)

where $d\Omega_z = d\Omega/\varrho^n$.

It should be noted that the following normalization convention is adopted in this paper

$$\int_\Omega \Phi(x) \, d\Omega = 1$$

(2.30)

which is equivalent to

$$\int_{\Omega(x)} \frac{1}{\varrho^n} \Phi\left(\frac{y-x}{\varrho}\right) d\Omega_y = \int_{\Omega(x)} \Phi_\varrho(y-x) \, d\Omega_y = 1$$

(2.31)

Nevertheless, there is a non-trivial yet subtle difference between the L^2 norms of the two window functions

$$\frac{1}{\varrho^n} \|\Phi\|^2_{L^2(\Omega)} = \|\Phi_\varrho\|^2_{L^2(\Omega)}$$

(2.32)

The moment matrix can be recast in a compact form

$$M(x) = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1\ell} \\ M_{21} & M_{22} & & M_{2\ell} \\ \vdots & & \ddots & \vdots \\ M_{\ell 1} & M_{\ell 2} & \cdots & M_{\ell\ell} \end{pmatrix}$$

(2.33)

and its inverse yields

$$M^{-1}(x) = \frac{1}{D_\ell} \begin{bmatrix} A_{11} & -A_{12} & \cdots & (-1)^{1+i}A_{1i} & \cdots & (-1)^{1+\ell}A_{1\ell} \\ -A_{21} & A_{22} & \cdots & (-1)^{2+i}A_{2i} & \cdots & (-1)^{2+\ell}A_{2\ell} \\ \vdots & & \ddots & & & \vdots \\ (-1)^{j+1}A_{j1} & & & (-1)^{i+j}A_{ij} & & (-1)^{j+\ell}A_{j\ell} \\ \vdots & & & & \ddots & \vdots \\ (-1)^{\ell+1}A_{\ell 1} & \cdots & \cdots & (-1)^{\ell+i}A_{\ell i} & \cdots & A_{\ell\ell} \end{bmatrix}$$

(2.34)

where

$$D_\ell := \det|M|$$

(2.35)

and A_{ij} are the minors of M_{ij}.

Thus, Eq. (2.28) can be rewritten in a succinct manner,

$$Gu(x) := P(0)M^{-1}(x) \int_{\Omega_y} P'\left(\frac{y-x}{\varrho}\right) u(y)\Phi_\varrho(y-x) \, d\Omega_y$$

$$= \int_{\Omega_y} \mathscr{C}(\varrho, y-x, x) u(y)\Phi_\varrho(y-x) \, d\Omega_y \qquad (2.36)$$

where the function $\mathscr{C}(\varrho, y-x, x)$ is the so-called *correction function* [6], which is defined as follows:

$$\mathscr{C}(\varrho, y-x, x) := P(0)M^{-1}(x)P'\left(\frac{y-x}{\varrho}\right) \qquad (2.37)$$

The expression for the correction function can be further simplified as follows:

$$\mathscr{C}(\varrho, y-x, x) = P(0)M^{-1}(x)P'\left(\frac{y-x}{\varrho}\right)$$

$$= (1,0,0,\ldots,0)\frac{1}{D_\ell}$$

$$\cdot \begin{pmatrix} A_{11} & -A_{12} & \cdots & (-1)^{1+i}A_{1i} & \cdots & (-1)^{1+\ell}A_{1\ell} \\ -A_{21} & A_{22} & \cdots & (-1)^{2+i}A_{2i} & \cdots & (-1)^{2+\ell}A_{2\ell} \\ \vdots & & \ddots & & & \vdots \\ (-1)^{j+1}A_{j1} & & & (-1)^{i+j}A_{ij} & & (-1)^{j+\ell}A_{j\ell} \\ \vdots & & & & \ddots & \vdots \\ (-1)^{\ell+1}A_{\ell1} & \cdots & \cdots & (-1)^{\ell+i}A_{\ell i} & \cdots & A_{\ell\ell} \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_i \\ \vdots \\ P_\ell \end{pmatrix}$$

$$= \frac{1}{D_\ell}[A_{11}P_1 - A_{12}P_2, \ldots, +(-1)^{1+i}A_{1i}P_i, \ldots, +(-1)^{1+\ell}A_{1\ell}P_\ell]$$

$$= P\left(\frac{y-x}{\varrho}\right)b(x) \qquad (2.38)$$

where the vector b is defined as

$$b' := \frac{1}{D_\ell}[A_{11}, -A_{12}, \ldots, (-1)^{1+i}A_{1i}, \ldots, (-1)^{1+\ell}A_{1\ell}] \qquad (2.39)$$

Let

$$\mathscr{H}_\varrho(y-x, x) := \mathscr{C}(\varrho, y-x, x)\Phi_\varrho(y-x) \qquad (2.40)$$

The function $\mathscr{H}_\varrho$ is referred to as the moving least-square reproducing kernel function, and consequently, the moving least-square reproducing kernel integral representation has the form

$$\mathscr{R}_\varrho^m u(x) := \int_\Omega u(y)\mathscr{H}_\varrho(y-x, x) \, d\Omega \qquad (2.41)$$

here the superscript m indicates the highest order of the generating polynomial[6] and the subscript ϱ represents the corresponding dilation parameter.

---

[6] Generally speaking, $m \neq \ell$ unless it is in 1-D case.

REMARK 2.2. The definition of reproducing kernel for a proper Hilbert space $\mathcal{V}(\Omega) := \{f(x)|x \in \Omega;\ \|f\|_{L^2(\Omega)} := \sqrt{\langle f, f \rangle} < +\infty\}$ is described as follows [16]: $\forall f \in \mathcal{V}(\Omega) \ \exists\ K(y, x) \in \mathcal{V}(\Omega)$ such that $\forall x \in \Omega$

$$f(x) = \langle f(y), K(y, x) \rangle_y \tag{2.42}$$

Whereas for the moving least-square kernel representation,

$$\mathcal{R}_\varrho^m f(x) \neq f(x) \tag{2.43}$$

unless $f(x) \in \pi_m(\Omega)$, where $\pi_m(\Omega)$ denotes the collection of polynomial in $\Omega \subset \mathbb{R}^n$ of total degree $\leq m$. Otherwise, $\mathcal{R}_\varrho^m f(x)$ is only a projection of the function f in the following reproducing kernel Hilbert space $\mathcal{V}_\varrho^M := \{f|f \in C^0(\Omega);\ f = \langle f, \mathcal{H}_\varrho^m(y - x, x) \rangle\}$.

It would be helpful to examine Eq. (2.28) through a pedagogic example. In the following, we shall illustrate how to construct a reproducing kernel function with linear polynomial basis in 3-D. Let

$$P(x) = (1, x_1, x_2, x_3) \tag{2.44}$$

It is obvious that

$$P(0) = (1, 0, 0, 0) \tag{2.45}$$

The moment matrix is then in the form

$$M(x) = \int_\Omega \begin{pmatrix} 1 \\ \dfrac{y_1 - x_1}{\varrho_1} \\ \dfrac{y_2 - x_2}{\varrho_2} \\ \dfrac{y_3 - x_3}{\varrho_3} \end{pmatrix} \left(1, \dfrac{y_1 - x_1}{\varrho_1}, \dfrac{y_2 - x_2}{\varrho_2}, \dfrac{y_3 - x_3}{\varrho_3} \right) \Phi_\varrho(y - x)\, d\Omega_y$$

$$= \int_\Omega \begin{pmatrix} 1 & \dfrac{y_1 - x_1}{\varrho_1} & \dfrac{y_2 - x_2}{\varrho_2} & \dfrac{y_3 - x_3}{\varrho_3} \\[2mm] \dfrac{y_1 - x_1}{\varrho_1} & \left(\dfrac{y_1 - x_1}{\varrho_1}\right)^2 & \left(\dfrac{y_1 - x_1}{\varrho_1}\right)\left(\dfrac{y_2 - x_2}{\varrho_2}\right) & \left(\dfrac{y_1 - x_1}{\varrho_1}\right)\left(\dfrac{y_3 - x_3}{\varrho_3}\right) \\[2mm] \dfrac{y_2 - x_2}{\varrho_2} & \left(\dfrac{y_2 - x_2}{\varrho_2}\right)\left(\dfrac{y_1 - x_1}{\varrho_1}\right) & \left(\dfrac{y_2 - x_2}{\varrho_2}\right)^2 & \left(\dfrac{y_2 - x_2}{\varrho_2}\right)\left(\dfrac{y_3 - x_3}{\varrho_3}\right) \\[2mm] \dfrac{y_3 - x_3}{\varrho_3} & \left(\dfrac{y_3 - x_3}{\varrho_3}\right)\left(\dfrac{y_1 - x_1}{\varrho_1}\right) & \left(\dfrac{y_3 - x_3}{\varrho_3}\right)\left(\dfrac{y_2 - x_2}{\varrho_2}\right) & \left(\dfrac{y_3 - x_3}{\varrho_3}\right)^2 \end{pmatrix}$$
$$\cdot \Phi_\varrho(y - x)\, d\Omega_y \tag{2.46}$$

Making change of variable, one may see that the moments defined in Eq. (2.29) are closely related with the conventional definition of moments in classic mechanics. Let $z_0 := 1$. One may have

$$M_{(i+1)(j+1)}(x) = m_{ij}(x) := \int_{\Omega(x)} z_i z_j \Phi(z)\, dz \quad 0 \leq i, j \leq 3 \tag{2.47}$$

where m_{ij} are the moments under the conventional definition.

Thus, in this particular case, one can rewrite Eq. (2.46) as follows:

$$M(x) = \begin{pmatrix} m_0 & m_1 & m_2 & m_3 \\ m_1 & m_{11} & m_{12} & m_{13} \\ m_2 & m_{21} & m_{22} & m_{23} \\ m_3 & m_{31} & m_{32} & m_{33} \end{pmatrix} \tag{2.48}$$

and the associated correction function is expressed as

$$\mathscr{C}(\varrho, y - x, x) = P(0)M^{-1}(x)P'\left(\frac{y - x}{\varrho}\right)$$

$$= (1, 0, 0, 0)\frac{1}{D_3}\begin{pmatrix} A_{11} & -A_{12} & A_{13} & -A_{14} \\ -A_{21} & A_{22} & -A_{23} & A_{24} \\ A_{31} & -A_{23} & A_{33} & -A_{34} \\ -A_{41} & A_{24} & -A_{34} & A_{44} \end{pmatrix}\begin{pmatrix} 1 \\ \dfrac{y_1 - x_1}{\varrho_1} \\ \dfrac{y_2 - x_2}{\varrho_2} \\ \dfrac{y_3 - x_3}{\varrho_3} \end{pmatrix}$$

$$= \frac{1}{D_3}\left[A_{11}(x) - A_{12}(x)\left(\frac{y_1 - x_1}{\varrho_1}\right) + A_{13}(x)\left(\frac{y_2 - x_2}{\varrho_2}\right) - A_{14}(x)\left(\frac{y_3 - x_3}{\varrho_3}\right)\right]$$

$$= P\left(\frac{y - x}{\varrho}\right)b(x) \tag{2.49}$$

where

$$b'(x) := \frac{1}{D_3}[A_{11}(x), -A_{12}(x), A_{13}(x), -A_{14}(x)] \tag{2.50}$$

The reproducing kernel function $\mathscr{K}_\varrho(y - x, x)$ can be obtained subsequently.

In general, the correction function has the following properties.

LEMMA 2.1. *Suppose* $\Phi(\cdot) \in C^k(\Omega)$ *and* $k \leq m$, *and* $p \geq 1$, $(1/p) + (1/q) = 1$, *then the correction function* $\mathscr{C}(\varrho, y - x, x) \in C^k(\Omega)$ *and* $\mathscr{C}(\varrho, y - x, x) \sim \mathcal{O}(1)$.

PROOF. By construction, $P_i(\cdot) \in C^\infty(\Omega)$. From (2.29), it follows that $\forall\, |\alpha| \leq k$,

$$D_x^\alpha M_{ij}(x) = \int_{\Omega(x)} \varrho^{-|\alpha|} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D_z^\beta(P_i(z)P_j(z))(D_z^{\alpha-\beta}\Phi(z))\, d\Omega_z \tag{2.51}$$

where $z = (y - x)/\varrho$.

Then, one can deduce from Eq. (2.51) that $D_x^\alpha M_{ij}$ is continuous and consequently bounded in Ω. As a matter of fact,

$$|D_x^\alpha M_{ij}| \leq C_\alpha \varrho^{-|\alpha|} \int_{\Omega(x)} \sum_{\beta \leq \alpha} |D_z^\beta(P_i(z)P_j(z))|\, |D_z^{\alpha-\beta}\Phi(z)|\, d\Omega_z$$

$$\leq C_\alpha \varrho^{-|\alpha|} \sum_{\beta \leq \alpha} \|D_z^\beta(P_i P_j)\|_{L^p(\Omega)}\|D_z^{\alpha-\beta}\Phi\|_{L^q(\Omega)} \tag{2.52}$$

It is easy to verify that

$$\|D_z^\beta(P_i P_j)\|_{L^p(\Omega)} = \left\{\int_{\Omega(x)} [D_z^\beta(P_i(z)P_j(z))]^p\, d\Omega_z\right\}^{1/p} \sim \mathcal{O}(1) \tag{2.53}$$

$$\|D_z^{\alpha-\beta}\Phi\|_{L^q(\Omega)} = \left\{\int_{\Omega(x)} (D_z^{\alpha-\beta}\Phi(z))^q\, d\Omega_z\right\}^{1/q} \sim \mathcal{O}(1) \tag{2.54}$$

Hence

$$|D_x^\alpha M_{ij}(x)| \sim \mathcal{O}(\varrho^{-|\alpha|}) \tag{2.55}$$

Again, by definition (2.39),

$$\mathscr{C}(\varrho, y - x, x) = P\left(\frac{y - x}{\varrho}\right)b(x) \tag{2.56}$$

thus

$$D_x^\alpha \mathscr{C}(\varrho, y - x, x) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D_x^\beta P\left(\frac{y - x}{\varrho}\right) D_x^{\alpha - \beta} b(x) \tag{2.57}$$

Recall,

$$Mb(x) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{2.58}$$

Then

$$D_x^\alpha (M(x)b(x)) = 0 \quad \Rightarrow \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D_x^\beta M(x) D_x^{\alpha - \beta} b(x) = 0 \tag{2.59}$$

or

$$MD_x^\alpha b(x) + \sum_{\substack{\beta \leq \alpha \\ \beta \neq 0}} (D_x^\beta M(x))(D_x^{\alpha - \beta} b(x)) = 0 \tag{2.60}$$

where $1 \leq |\alpha| \leq k$. Thereby, one can solve $D_x^\alpha b(x)$ recursively; for instance, in the 1-D case,

$$\begin{bmatrix} M & 0 & 0 & \cdots & \cdots & 0 \\ \binom{2}{1}D_x^1 M & M & 0 & \cdots & \cdots & 0 \\ \binom{3}{2}D_x^2 M & \binom{3}{1}D_x^1 M & M & \cdots & \cdots & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ \binom{k}{k-1}D_x^{k-1} M & \binom{k}{k-2}D_x^{k-2} M & \cdots & \cdots & & M \end{bmatrix} \begin{pmatrix} D_x^1 b \\ D_x^2 b \\ D_x^3 b \\ \vdots \\ \vdots \\ D_x^k b \end{pmatrix} = - \begin{pmatrix} (D_x^1 M)b \\ (D_x^2 M)b \\ (D_x^3 M)b \\ \vdots \\ \vdots \\ (D_x^k M)b \end{pmatrix} \tag{2.61}$$

The coefficient matrix of the above equations is a lower triangular block matrix, which can be solved by forward substitution. Therefore, $D_x^\alpha b(x)$ are uniquely determined, and by comparing the coefficients in both sides of Eq. (2.61) or (2.60), one may find that $D_x^\alpha b(x) \sim \mathscr{O}(\varrho^{-|\alpha|})$. Consequently, it can be deduced that

$$\mathscr{C}(\varrho, x, y) \in C^k(\Omega) \tag{2.62}$$

and

$$b(x) \sim \mathscr{O}(1) \Rightarrow \mathscr{C}(\varrho, y - x, x) \sim \mathscr{O}(1) \quad \square \tag{2.63}$$

2.2. Shape function and particle distribution

As mentioned above, the early moving least-square interpolant was formulated directly in a discrete base [10]. Instead of following the earlier derivation, we would like to view the moving least-square interpolant function as a discretized form of the kernel formula,

$$\mathscr{R}_\varrho^m u(x) = \int_\Omega u(y)\mathscr{C}(\varrho, x, y)\varPhi_\varrho(y - x) \, d\Omega_y \tag{2.64}$$

Before proceeding to deriving the expression of the shape function, introducing several preliminary concepts is in order. Although there is no mesh required for this method, there are still some topological requirements on the particle distribution (Fig. 1).

Let us consider a simply connected region, $\overline{\Omega} \subset \mathbb{R}^3$, in which there is a particle distribution. By particle distribution we mean that

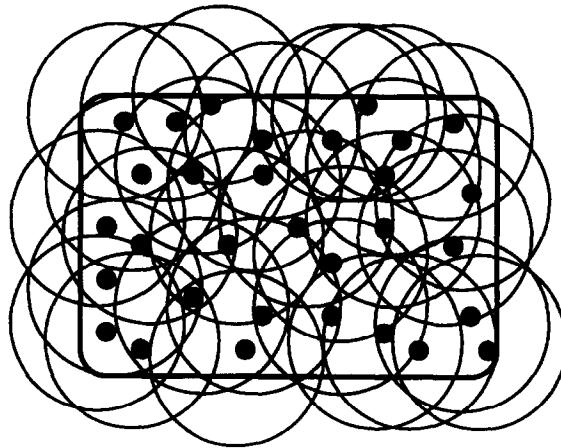DEFINITION 2.1 (Particle Distribution). Each particle inside or on the boundary of Ω is assigned with a

Fig. 1. An admissible particle distribution.

parametric dilation vector ϱ_I, $I = 1, \ldots, NP$. With the parametric dilation vector ϱ_I, one can construct a compact support region around each particle. Suppose the particle I occupies the position x_I in reference coordinate system, the compact support can be constructed as a local 'sphere'

$$S_I := \{x \mid |x - x_I| \leq r|\varrho_I|\}$$ (2.65)

or, a parallelogram box

$$S_i := \{x \mid |x_i - x_{Ii}| \leq r\varrho_{Ii}, \ \varrho_{Ii} > 0, \ 1 \leq i \leq 3\}$$ (2.66)

where constant r is a proportional coefficient.

Not all particle distributions can be used in numerical computation. The valid particle distribution is referred to as the 'admissible particle distribution'. Admissibility of the particle distribution depends on computational feasibility. In what follows, the concept of admissible distribution shall be precisely defined. We start with defining the particle density index.

DEFINITION 2.2 (Particle Density Index). The density of a particle distribution is measured by the indices that are associated with the dilation parameters of all particles,

$$\varrho_{max} := \max\{|\varrho_I|, I = 1, \ldots, NP\}$$ (2.67)

$$\varrho_{min} := \min\{|\varrho_I|, I = 1, \ldots, NP\}$$ (2.68)

In this paper, we restrict our attention on the case in which the dilation parameter is uniform for all particles, i.e.

$$\varrho_1 = \varrho_2 = \cdots = \varrho_I = \cdots = \varrho_{NP} = \varrho$$ (2.69)

which means that

$$\varrho_{min} = \rho_{max} = \varrho$$ (2.70)

Obviously, the smaller ϱ is, the denser the particle distribution will be. Thereby, a density refinement can be defined as a decreasing process of the dilation parameter, provided that in this process the new particle distribution remains as an admissible particle distribution.

DEFINITION 2.3 (Admissible Particle Distribution). An admissible particle distribution is a particle distribution that satisfies the following conditions:

(1) Every particle of a distribution associates with a compact supports

$$S_I := \{x - x_I| \leq r\varrho\}$$ (2.71)

and the union of all the compact support S_I

$$S := \bigcup_{I=1}^{NP} S_I \tag{2.72}$$

generates a covering for the domain $\overline{\Omega}$ in which the particles reside, viz.

$$\overline{\Omega} \subseteq S \tag{2.73}$$

(2) $\forall \overline{x} \in \overline{\Omega}$, there exists a ball

$$\mathscr{B}(x) := \{x \mid |x - \overline{x}| \leq r|\varrho|\} \tag{2.74}$$

so that the number of particles in the $\mathscr{B}$, N_p, satisfies the condition

$$N_{\min} \leq N_p \leq N_{\max} \tag{2.75}$$

where both $N_{\min}$ and $N_{\max}$ are a priori assigned numbers, such that

$$0 < N_{\min} \leq N_{\max} < \infty \tag{2.76}$$

(3) The particle distribution should be non-degenerate, which means that, in 1-D case, there are at least two particles in $\mathscr{B}(x)$, and these two particles cannot overlap, i.e. the difference of their position vector should form a line segment; similarly, in 2-D case, there are at least three particles in $\mathscr{B}(x)$, and the three position vectors form a nonzero triangular element. Generally speaking, for $\mathscr{B}(x) \in \mathbb{R}^n$, there are at least $n + 1$ particles in $\mathscr{B}(x)$, and their position vectors form a nonzero nth rank 'simplex' element.

A non-admissible particle distribution is shown in Fig. 2, which fails to satisfy the Condition 1. In Fig. 2, one can see that in the middle of the area there is a small dark region which is not covered by the compact support of any particles. As a matter of fact, Condition 2 implies Condition 1. Moreover, Condition (2.75) is closely related to computational feasibility. The first part condition $N_p \geq N_{\min}$ guarantees the stability condition of the shape function, or regularity of the moment matrix M. The second part of inequality guarantees the bandedness of the resulting stiffness matrix.

REMARK 2.3.
(1) It can be shown that the shape functions generated from the MLSRK function with the window function $\Phi_\varrho(x)$ form a complete basis for a finite dimensional space $\mathcal{V}_\varrho^h \subset C^m(\Omega)$. As $\varrho \to 0$, there exists a monotonic decreasing sequence
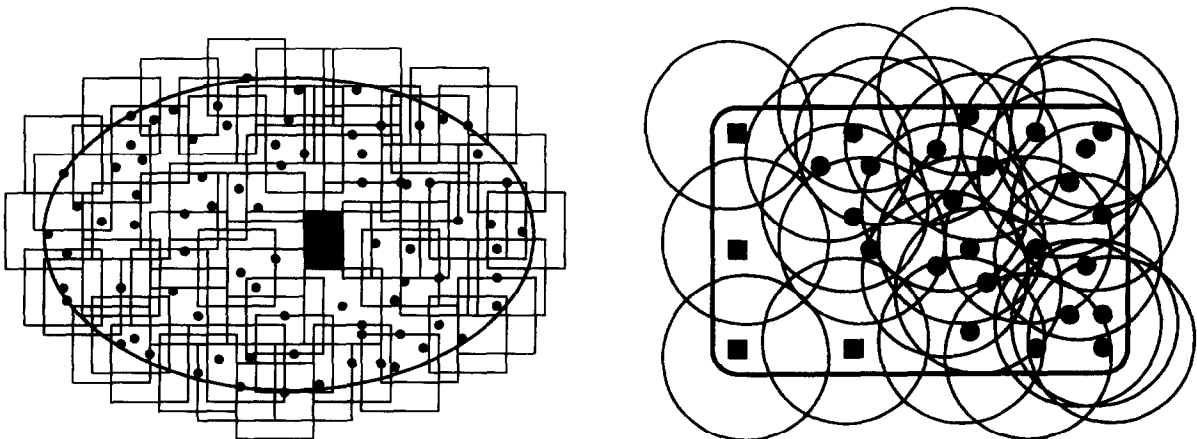

Fig. 2. Inadmissible particle distribution. I: shaded area is not covered by any nodal support.

Fig. 3. Inadmissible particle distribution II.

$$|\varrho^1| \geqslant |\varrho^2| \geqslant |\varrho^3| \geqslant \cdots, \tag{2.77}$$

such that

$$\mathcal{V}_{\varrho^1}^h \subset \mathcal{V}_{\varrho^2}^h \subset \mathcal{V}_{\varrho^3}^h \cdots \subset C^m(\Omega) \tag{2.78}$$

More precisely, for admissible particle distributions as $\varrho \to 0$, or $j \to \infty$

$$\mathcal{V}_{\varrho^j}^h \to C^m(\Omega) \subset L^2(\Omega) \tag{2.79}$$

In fact, this is very similar to the basic concept of multiresolution analysis (see [17]).

(2) It should be noted that the particle distribution with constant dilation parameter does not necessarily mean uniform particle distribution; one can construct non-uniform particle distribution with constant dilation parameter. Fig. 1 illustrates a simple example.

(3) A degenerated particle distribution is shown in Fig. 4. Since all the particles lie along the x-axis, the moment components m_y and m_{xy} will be identically zero, and then the moment matrix will be singular.

Once an admissible particle distribution is set up, one can approximate the reproducing kernel integral by numerical quadrature, i.e.

$$\mathcal{R}_{\varrho,h}^m u(x) := \sum_{I=1}^{NP} u(x_I) \, \mathscr{C}^h(\varrho, x_I - x, x) \Phi_\varrho(x_I - x) \, \Delta V_I \tag{2.80}$$

$$= \sum_{I=1}^{NP} \mathscr{K}_\varrho^h(x_I - x, x) u_I \, \Delta V_I \tag{2.81}$$

In (2.81), ΔV_I is the quadrature weight, or Ith particle's lumped volume. If one chooses

$$\Delta V_I = 1, \tag{2.82}$$

the discrete form of reproducing kernel integral will be equivalent to the formulation proposed by Lancaster and Salkauskas [10]. Although this is truly in the mesh-free spirit, it may not be as accurate as a discrete integral obtained by numerical quadrature. Another advantage of using numerical quadrature to evaluate the reproducing kernel integral is that it approximates the integral with the correct measure, by which we mean

$$\sum_{I=1}^{NP} \Delta V_I = \text{meas}(\Omega) \tag{2.83}$$

whereas Lancaster and Salkauskas' formulation as well as other discrete formulation does not preserve this property, i.e.

$$\sum_{I=1}^{NP} 1 = NP \neq \text{meas}(\Omega) \tag{2.84}$$

However, this choice can be justified by choosing different normalization factors. For simplicity of the presentation, this case is excluded here, though all the results presented here are valid for this case too.

PROPOSITION 2.1 (A Necessary Stability Condition for Correction Function). For a given admissible particle distribution, a necessary condition for stable correction function is
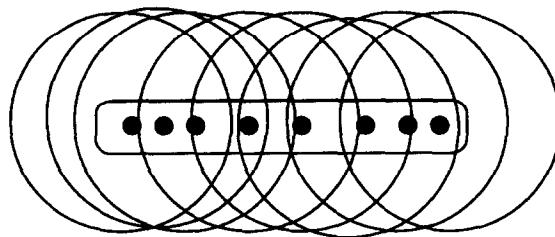

Fig. 4. A degenerate 2-D particle distribution.

$$E(\varrho) \geq \Delta V_I, \qquad 1 \leq I \leq NP \tag{2.85}$$

where $E(\varrho) := r^n \varrho^n$, which represents the volume of the kernel support, and ΔV_I is the quadrature weight. The condition (2.85) implies

$$N_{min} \geq 2 \tag{2.86}$$

PROOF. We first show that when $N_{min} < 2$, the moment matrix $M(x)$ will be singular (not positive definite anymore!).

Since

$$M_{ij}(x) = \int_{\Omega_x} P_i\left(\frac{y - x}{\varrho}\right) P_j\left(\frac{y - x}{\varrho}\right) \Phi_\varrho(y - x)\, d\Omega_x \tag{2.87}$$

then by the quadrature rule—the trapezoidal rule,

$$M_{ij}^h(x) = \sum_{I=1}^{NP} P_i\left(\frac{x_I - x}{\varrho}\right) P_j\left(\frac{x_I - x}{\varrho}\right) \Phi_\varrho(x_I - x)\, \Delta V_I \tag{2.88}$$

If $N_{min} < 2$, $\forall\, i + j > 2$,

$$
\begin{aligned}
M_{ij}^h(x_J) &= \sum_{I=1}^{NP} P_i\left(\frac{x_I - x_J}{\varrho}\right) P_j\left(\frac{x_I - x_J}{\varrho}\right) \Phi_\varrho(x_I - x_J)\, \Delta V_I \\
&= P_i\left(\frac{x_J - x_J}{\varrho}\right) P_j\left(\frac{x_J - x_J}{\varrho}\right) \Phi_\varrho(x_J - x_J)\, \Delta V_I \\
&= 0
\end{aligned}
\tag{2.89}
$$

Consequently, the moment matrix $M^h(x)$ is a singular matrix; thus the correction function $\mathscr{C}^h$ is unbounded and hence the algorithm is unstable.

On the other hand, if $N_{min} \geq 2$, which implies that there are at least two points inside a kernel support; consequently, the quadrature weight ΔV_I has to be smaller than the volume of the kernel support, i.e.

$$E(\varrho) \geq \Delta V_I \tag{2.90}$$

Otherwise, the condition (2.83) will be violated. $\square$

As a matter of fact, one can show that, even in a one-dimensional case, if $m \geq 2$, then $N_{min} \geq 3$. In two dimensions, as mentioned above, one can prove that at least $N_{min} \geq 3$, because if there are less than three particles in a plane region, by rotating the coordinate system, one moment component will always be zero, and hence the moment matrix will be singular. Condition 2 can be also expressed as

$$N_{min} \leq \mathrm{card}\{I \mid x \in \mathrm{supp}\{\mathscr{K}_\varrho(\bar{x}_I - x, x)\}\} \leq N_{max} \tag{2.91}$$

which is partially recognized by Babuška et al. [18] as the pointwise overlap condition.

Fig. 3 shows another type of inadmissible particle distribution, which fails to satisfy the condition $N_{min} > 2$. In Fig. 3, at the left side of problem domain, the compact supports with square shaped particle in the center illustrate this situation, even though in this case the whole interior domain is covered by the union of all compact supports.

The discrete formulation can be viewed as a shape function expansion, which is especially useful in Galerkin procedure; i.e. one can assume the approximation to the solution of PDE as follows:

$$u^\upsilon(x) = \sum_{I=1}^{NP} \xi_I N_I(\varrho, x, x_I) \tag{2.92}$$

where

$$N_I(\varrho, x, x_I) := \mathscr{C}^h(\varrho, x_I - x, x)\Phi_\varrho(x_I - x)\,\Delta V_I$$

$$= P(0)M^{-1}(x)P\left(\frac{x_I - x}{\varrho}\right)\Phi_\varrho(x_I - x)\,\Delta V_I$$

$$= \mathscr{K}_\varrho^h(x_I - x, x)\,\Delta V_I \tag{2.93}$$

From this definition, it is clear that

$$\mathrm{supp}\{N_I\} = \{x \mid \|x - x_I\| \le r\varrho\} \tag{2.94}$$

In the discrete formulation, as Eqs. (2.92) and (2.93) indicate, each particle corresponds to a correction function. In Fig. 5, we display the distribution of a group of correct functions that correspond to a group of eleven particles residing evenly in the interval [0,1]. From Fig. 5, one can see that far away from the end point, i.e. the boundary, the correction functions take the unit value.

REMARK 2.4. It should be noted that once a particular quadrature rule is chosen it should be carried out through all the integrals consistently. It is particularly important that the moments integrals should be carried out by the same quadrature rule, i.e.

$$m_{ij}^h(x) := \sum_{I=1}^{NP} z_i z_j \Phi(z)\,\Delta V_I \tag{2.95}$$

In most mathematical literature, the term 'interpolation' is exclusively preserved for the sampling representation which satisfies the condition that

$$f^h(x) := \sum_{i=1}^{n} f_i N_i(x) \tag{2.96}$$

which implies,

$$f^h(x_i) = f_i, \quad \text{and} \quad N_i(x_j) = \delta_{ij} \tag{2.97}$$

In the context of this paper, the term 'interpolation' is used in a very broad sense, is not restricted in its
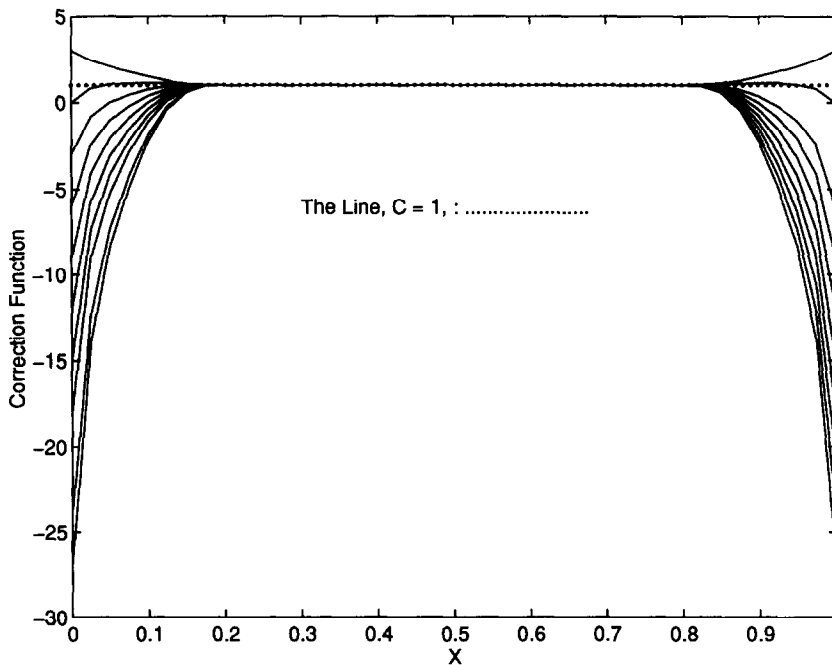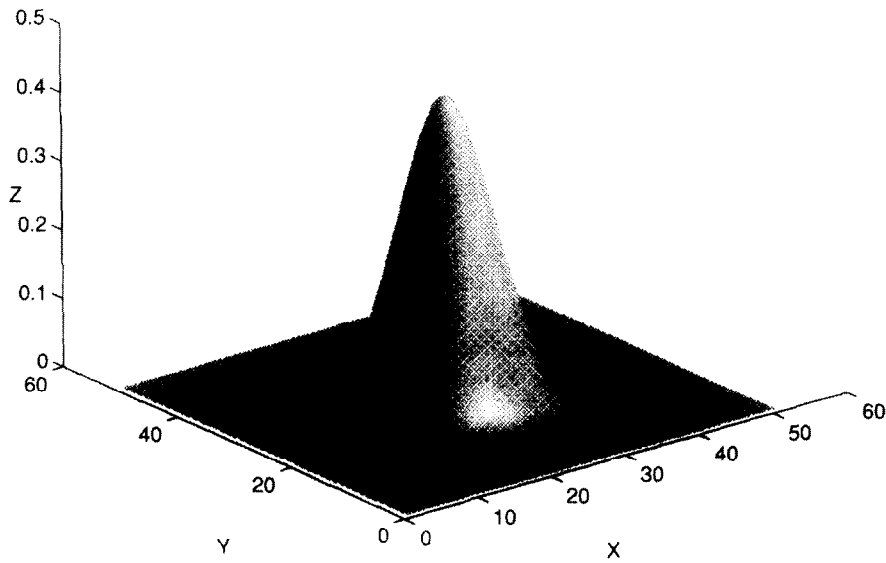

Fig. 5. The profiles of correction functions.

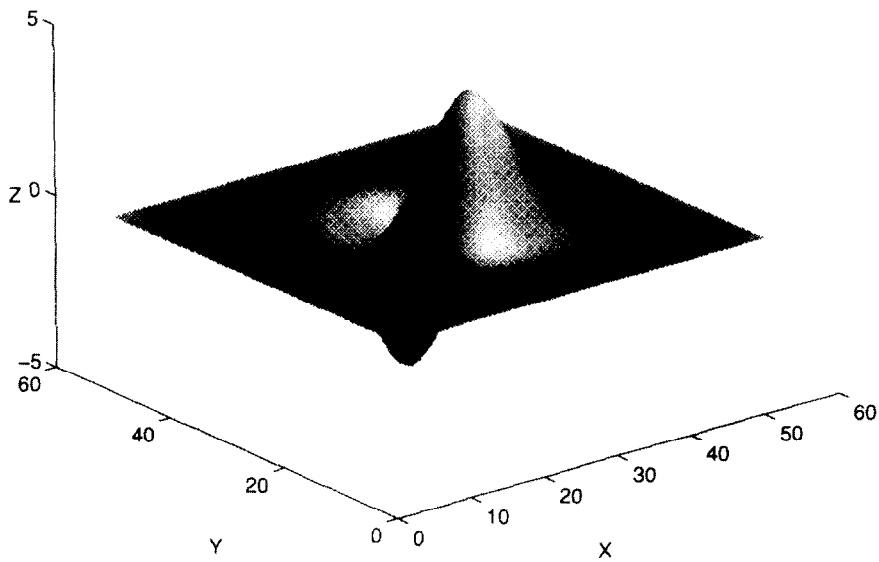Fig. 6. A 2-D shape function based on the cubic spline window function.

Fig. 7. The first derivative of the shape function, dN_I/dy.

conventional connotation, i.e. sampling. The moving least-square kernel Galerkin method discussed here is basically a pseudospectral method; hence, the proposed interpolation function does not satisfy the usual interpolation condition,

$$u^e = \sum_{I=1}^{NP} \xi_I N_I(\varrho, x, x_I) \tag{2.98}$$

$$N_I(\varrho, x_J, x_I) \neq \delta_{IJ} \tag{2.99}$$

this means that

$$u^e(x_I) \neq \xi_I \tag{2.100}$$

where ξ_I is the dual base of N_I. This brings us to another crucial point, i.e. there is an essential difference between the MLSRK interpolation and moving least-square reproducing representation. That is
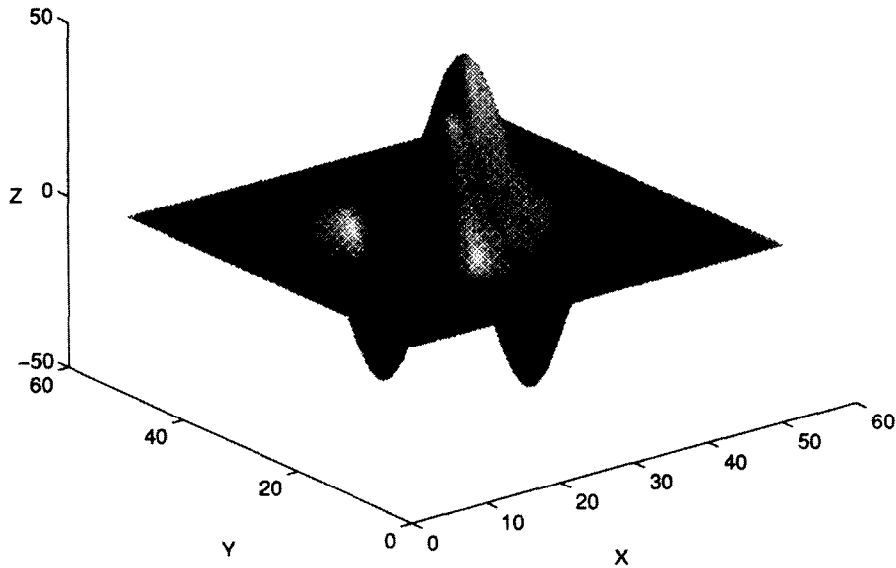
Fig. 8. The mixed derivative of shape function, $d^2 N_I / dx\, dy$.

$$u^e(x) \neq \mathcal{R}^m_{\varrho,h} u(x) \tag{2.101}$$

As a matter of fact, Lancaster and Salkauskas [10] called the former as the non-interpolation interpolant. This type of approximation is also called as quasi-interpolation in spline approximation [19]. On how to construct an exact interpolant with singular weight functions, readers may consult Lancaster and Salkauskas [10].

3. Convergence analysis

In this section, we shall discuss the consistency condition, or the completeness of interpolation and then discuss the error estimation of the proposing method. Before going into the details, we need some machineries to set the stage.
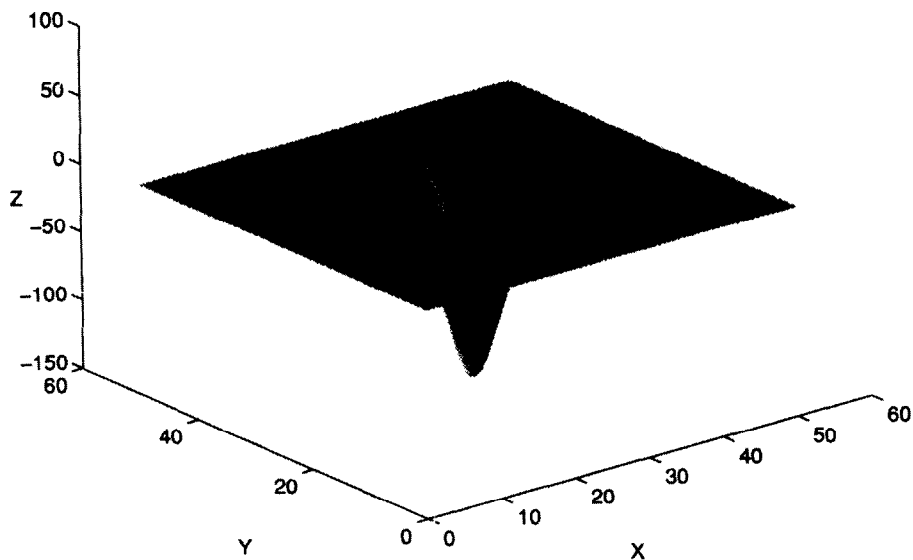
Fig. 9. The second derivative of the shape function, $d^2 N_I / dx^2$.

3.1. Preliminary

If α and β are two multi-indices, we say $\beta \leq \alpha$ provided $\beta_i \leq \alpha_i$, $\forall\ 1 \leq i \leq n$. Let

$$\alpha! := \alpha_1! \alpha_2! \cdots \alpha_n! \tag{3.1}$$

By the same token, one can denote

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha - \beta)!} = \binom{\alpha_1}{\beta_1}\binom{\alpha_2}{\beta_2} \cdots \binom{\alpha_n}{\beta_n} \tag{3.2}$$

It is then easy to verify the Leibniz formula

$$D_x^\alpha(uv)(x) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D_x^\beta u(x) D_x^{\alpha - \beta} v(x) \tag{3.3}$$

Assume $u \in C^{m+1}(\Omega) \cap W_p^{m+1}(\Omega)$, and denote

$$h^\alpha := h_1^{\alpha_1} h_2^{\alpha_2} \cdots h_n^{\alpha_n} \tag{3.4}$$

The Taylor's formula of order α, $|\alpha| = m$ can then be written as

$$u(a + h) = \sum_{|\alpha|=0}^{m} \frac{1}{\alpha!} D_x^\alpha u(a) h^\alpha + \frac{1}{(m+1)!} \sum_{|\alpha|=m+1} D_x^\alpha u(a + \theta h) h^\alpha \tag{3.5}$$

where $\theta \in {]}0, 1{[}$.

For convenience, the definitions of Lebesgue space and Sobolev space are listed as follows.

DEFINITION 3.1. The Lebesgue space is defined as

$$L^p(\Omega) := \{ f \mid \|f\|_{L^p(\Omega)} < \infty \}, \quad 1 \leq p \leq \infty \tag{3.6}$$

where the L_p norm is defined as

$$\|f\|_{L^p(\Omega)} := \left(\int_\Omega |f(x)|^p \, dx \right)^{1/p} \quad \forall\ 1 \leq p < \infty \tag{3.7}$$

$$\|f\|_{L^\infty(\Omega)} := \text{ess. sup}\{ |f(x)| \mid x \in \Omega \} \tag{3.8}$$

where dx is Lebesgue measure.

DEFINITION 3.2. Let k be a non-negative integer, and let $f \in L_{loc}^1(\Omega)$. For $1 \leq p < \infty$, the Sobolev norm is defined as

$$\|f\|_{W_p^k(\Omega)} := \left(\sum_{|\alpha| \leq k} \|D_w^\alpha f\|_{L^p(\Omega)}^p \right)^{1/p} \tag{3.9}$$

and for $p = \infty$

$$\|f\|_{W_\infty^k(\Omega)} := \max_{|\alpha| \leq k} \|D_w^\alpha f\|_{L^\infty(\Omega)} \tag{3.10}$$

In either case, we define the Sobolev space via

$$W_p^k(\Omega) := \{ f \in L_{loc}^1(\Omega) \mid \|f\|_{W_p^k(\Omega)} < \infty \} \tag{3.11}$$

where $L_{loc}^1(\Omega)$ is the set of locally integrable functions defined as

$$L_{loc}^1(\Omega) := \{ f \mid f \in L^1(K)\ \forall\ K \subset \text{interior } \Omega \} \tag{3.12}$$

Particularly, we denote $H^m(\Omega) = W_2^m(\Omega)$. Later, we also need the definition of Sobolev semi-norm,

$$|u|_{W_p^m(\Omega)} := \left(\sum_{|\alpha|=m} \|D^\alpha f\|_{L^p(\Omega)}^p \right)^{1/p}$$ (3.13)

and if $p = \infty$,

$$|u|_{W_\infty^m(\Omega)} := \max_{|\alpha| \le m} \{\text{ess. sup}_{x \in \Omega} |D^\alpha u|\}$$ (3.14)

In later sections, several standard inequalities will be used for error estimation. For the sake of convenience, they are listed below.

Hölder's inequality

For $1 \le p$, $q \le \infty$, p and q are conjugate to each other such that $(1/p) + (1/q) = 1$,

(I) For a_i, $b_i \ge 0$, $i = 1, \ldots . n$. Suppose both sums $\sum_{i=1}^n a_i^p$ and $\sum_{i=1}^n b_i^q$ are finite, then

$$\sum_{i=1}^n a_i b_i \le \left(\sum_{i=1}^n a_i^p \right)^{1/p} \left(\sum_{j=1}^n b_j^q \right)^{1/q}$$ (3.15)

(II) If functional $f \in L^p(\Omega)$, $g \in L^q(\Omega)$, then $fg \in L^1(\Omega)$ and

$$\|fg\| \le \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}$$ (3.16)

Cauchy–Schwarz inequality

This is a special case of Hölder's inequality. Let $p = q = 2$,

(I) If both summations $\sum_{i=1}^n a_i^2$ and $\sum_{i=1}^n b_i^2$ are finite, then

$$\sum_{i=1}^n a_i b_i \le \sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{j=1}^n b_j^2}$$ (3.17)

(II) If $f \in L^2(\Omega)$, $g \in L^2(\Omega)$; then $fg \in L^1(\Omega)$ and

$$\int_\Omega f(x)g(x) \, d\Omega \le \|f\|_{L^2(\Omega)} \|g\|_{L^2(\Omega)}$$ (3.18)

By using the multiple index notation, we construct a complete base vector for $\mathcal{V}_\varrho^h(\Omega)$, where $\Omega \subset \mathbb{R}^n$.

DEFINITION 3.3. Assume $\Omega \subset \mathbb{R}^n$. A complete m-order, ℓ-component polynomial vector $P(x_I - x)/\varrho = [P_1, P_2, \ldots, P_l](x_I - x)/\varrho$ consists of the following components

$|\alpha| = 0$; $P_1 = 1$;

$|\alpha| = 1$; $P_2 = \left(\dfrac{x_{1I} - x_1}{\varrho_1} \right), \ldots, \ldots, P_{n+1} = \left(\dfrac{x_{nI} - x_n}{\varrho_n} \right)$;

$\ldots, \ldots, \ldots,$

$|\alpha| = m$; $\ldots, P_k = \left(\dfrac{x_{1I} - x_1}{\varrho_1} \right)^{\alpha_1} \left(\dfrac{x_{2I} - x_2}{\varrho_2} \right)^{\alpha_2} \cdots \left(\dfrac{x_{nI} - x_n}{\varrho_n} \right)^{\alpha_n}, \ldots, P_\ell \left(\dfrac{x_{nI} - x_n}{\varrho_n} \right)^m$ (3.19)

Particularly, for $n = 2$, $\ell = \frac{1}{2}m(m + 1)$; and $n = 3$, $\ell = 1 + \sum_{k=1}^m \frac{1}{2}k(k + 1)$.   □

For example, in the two-dimensional case $(n = 2)$, the components of a m-order, ℓ-component polynomial base vector, $P(x_I - x)/\varrho = [P_1, P_2, \ldots, P_\ell](x_I - x)/\varrho$, forms the following triangular pyramid,

$$P_1 = 1$$

$$P_2 = \left(\frac{x_I - x}{\varrho} \right); \quad P_3 = \left(\frac{y_I - y}{\varrho} \right)$$

$$P_4 = \left(\frac{x_I - x}{\varrho} \right)^2; \quad P_5 = \left(\frac{x_I - x}{\varrho} \right)\left(\frac{y_I - y}{\varrho} \right); \quad P_6 = \left(\frac{y_I - y}{\varrho} \right)^2$$

$$\ldots, \ldots, \ldots, \ldots, \ldots,$$

$$\cdots, \ \cdots, \ \cdots, \ \cdots, \ \cdots, \ \cdots, \ \cdots,$$

$$P_{\ell-m} = \left(\frac{x_I - x}{\varrho}\right)^m; \quad \cdots, \cdots, \cdots; \quad \cdots, \cdots, \cdots; \quad \cdots, \cdots, \cdots; \quad P_\ell = \left(\frac{y_I - y}{\varrho}\right)^m \quad (3.20)$$

where $\ell = [m(m + 1)]/2$.

3.2. Completeness of interpolants

It is well known that in order to meet convergence requirements, finite element shape functions have to satisfy the completeness condition, i.e. they have to be able to represent rigid body motion and a constant strain field exactly, if the terminology of solid mechanics is used. For non-conforming elements, due to lack of analytic estimates, this completeness requirement is tested through the so-called patch test. Following the same idea, Belytschko et al. [4] have systematically conducted various patch tests for moving least square interpolant functions. All the MLS interpolants have passed the patch test without exception.

Liu et al. [6] proved analytically that for Lancaster–Salkauskas' type moving least-square interpolant, the completeness of the interpolant shape function is automatic. As a matter of fact, one can show that the reproducing kernel formula can reproduce any type of polynomial at a designed manner, which might be called 'super-completeness' if one wishes to exaggerate it. In other words, the moving least-square interpolants enjoy an extra luxury of completeness.

In what follows, we shall discuss this completeness—a m-consistency structure. First, we show the completeness in the one-dimensional case, since it offers the physical insight and the essence of the mathematical structure. The multi-dimensional case will follow as a natural extension. Furthermore, departing from this point, we shall show that the structure of this type of moving least-square interpolant yields a nice property, which is essential to the convergence property of the proposing method.

For a one-dimensional case, the polynomial base vector P is

$$P\left(\frac{x_I - x}{\varrho}\right) := \left(1, \frac{x_I - x}{\varrho}, \left(\frac{x_I - x}{\varrho}\right)^2, \left(\frac{x_I - x}{\varrho}\right)^3, \ldots, \left(\frac{x_I - x}{\varrho}\right)^m\right) \quad (3.21)$$

the corresponding moment matrix is

$$M(x) = \begin{pmatrix} m_0 & m_1 & \cdots & & m_m \\ m_1 & m_2 & \cdots & \cdots & m_{m+1} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ m_m & m_{m+1} & \cdots & \cdots & m_{2m} \end{pmatrix} \quad (3.22)$$

where $m_i := \int_\Omega z^i \Phi(z)\, \mathrm{d}z$.

Its inverse matrix has the form

$$M^{-1}(x) = \frac{1}{D_m} \begin{pmatrix} A_{11} & -A_{12} & \cdots & \cdots & (-1)^{1+m} A_{1m} \\ -A_{21} & A_{22} & \cdots & \cdots & (-1)^{2+m} A_{2m} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ (-1)^{1+m} A_{m1} & (-1)^{2+m} A_{m2} & \cdots & \cdots & A_{mm} \end{pmatrix} \quad (3.23)$$

REMARK 3.1. The matrix (3.22) is a Hankel type matrix, which can be always viewed as a product of the 'backward identity' permutation matrix and a Toeplitz matrix [20].

Thus, the number of operations needed in inverting matrix (3.22) should be of the same order as needed in inverting the corresponding Toeplitz matrix. Therefore, numerically inverting matrix (3.22) only requires $\mathcal{O}(m^2)$ operations [21].

LEMMA 3.1. The moving least-square kernel interpolant function with basis (3.21) can reproduce any polynomials $f(x) \in \pi_m$ exactly by using the sample values, viz.

$$\sum_{I=1}^{NP} N_I(\varrho, x_I, x) f(x_I) = f(x)$$

(3.24)

PROOF. It is sufficient to show that for $0 \le k \le m$,

$$\sum_{I=1}^{NP} N_I(\varrho, x, x_I) x_I^k = x^k \quad \forall\, 0 \le k \le m$$

(3.25)

$$\sum_{I=1}^{NP} N_I(\varrho, x, x_I) x_I^k = P(0) M^{-1}(x) \sum_{I=1}^{NP} \left(P' \left(\frac{x_I - x}{\varrho} \right) \Phi_\varrho(x_I - x) \, \Delta V_I \right) x_I^k$$

(3.26)

Let

$$z_I = \frac{x_I - x}{\varrho} \Rightarrow x_I^k = (\varrho z_I + x)^k$$

(3.27)

Thus,

$$\sum_{I=1}^{NP} N_I(\varrho, x, x_I) x_I^k = \sum_{I=1}^{NP} N_I(\varrho, x, x_I) \left(\sum_{j \le k} \binom{k}{j} \varrho^{k-j} z_I^{k-j} x^j \right)$$

$$= \sum_{j \le k} \binom{k}{j} \varrho^{k-j} x^j P(0) M^{-1}(x) \sum_{I=1}^{NP} (P'(z_I) \Phi_\varrho(z_I) \, \Delta V_I) z_I^{k-j}$$

$$= \sum_{j \le k} \binom{k}{j} \varrho^{k-j} x^j P(0) M^{-1}(x) \sum_{I=1}^{NP} \begin{pmatrix} z_I^{k-j} \\ z_I^{k-j+1} \\ \vdots \\ z_I^{k-j+m} \end{pmatrix} \Phi_\varrho(z_I) \, \Delta V_I$$

(3.28)

By the *Laplace* Theorem,

$$P(0) M^{-1}(x) \sum_{I=1}^{NP} \begin{pmatrix} z_I^{k-j} \\ z_I^{k-j+1} \\ \vdots \\ z_I^{k-j+m} \end{pmatrix} \Phi_\varrho(z_I) \, \Delta V_I = \delta_{kj}$$

(3.29)

it then follows

$$\sum_{I=1}^{N} N_I(\varrho, x, x_I) x_I^k = \sum_{j \le k} \binom{k}{j} \varrho^{k-j} x^j \delta_{kj} = x^k \quad \square$$

(3.30)

More generally, in $\mathbb{R}^n$, the following completeness conditions hold.

LEMMA 3.2. The moving least-square shape function, which is generated by a complete m-order, ℓ-term polynomial $P(x_I - x)/\varrho$, satisfies the following consistency condition:

$$\sum_{I=1}^{NP} P_k \left(\frac{x_I - x}{\varrho} \right) N_I(\varrho, x, x_I) = \delta_{k1}, \quad 1 \le k \le \ell$$

(3.31)

$\square$

PROOF. The proof is straightforward.

$$\sum_{l=1}^{NP} N_l(\varrho, x, x_l) P_k\left(\frac{x_l - x}{\varrho}\right) = P(0) M^{-1}(x) \sum_{l=1}^{NP} P^t\left(\frac{x_l - x}{\varrho}\right) P_k\left(\frac{x_l - x}{\varrho}\right) \Phi_\varrho(x_l - x) \Delta V_l$$

$$= P(0) M^{-1}(x) \sum_{l=1}^{NP} \begin{pmatrix} P_1 P_k \\ P_2 P_k \\ \vdots \\ \vdots \\ P_\ell P_k \end{pmatrix} \Phi_\varrho(x_l - x) \Delta V_l$$

$$= P(0) M^{-1}(x) \begin{pmatrix} M_{1k} \\ M_{2k} \\ \vdots \\ M_{\ell k} \end{pmatrix}$$

$$= \frac{1}{D_l} (A_{11}, -A_{12}, \dots, (-1)^{1+\ell} A_{1\ell}) \begin{pmatrix} M_{1k} \\ M_{2k} \\ \vdots \\ M_{\ell k} \end{pmatrix}$$

$$= \delta_{1k} . \tag{3.32}$$

Once again the *Laplace* theorem is used in the last step. □

Direct consequences of Lemma 3.2 are the following.

COROLLARY 3.1 (*m-consistency condition 1*). *For multi-index* α, $1 \le |\alpha| \le m$, *the* α*-order moments of the MLSRK shape function are identical to zero, i.e.*

$$\sum_{l=1}^{NP} (x_l - x)^\alpha N_l(\varrho, x, x_l) = \sum_{l=1}^{NP} (x_l - x)^\alpha \mathcal{H}_\varrho^h(x_l - x, x) \Delta V_l = \delta_{\alpha 0} ; \tag{3.33}$$

$$\sum_{l=1}^{NP} x_l^\alpha N_l(\varrho, x, x_l) = \sum_{l=1}^{NP} x_l^\alpha \mathcal{H}_\varrho^h(x_l - x, x) \Delta V_l = x^\alpha \tag{3.34}$$

PROOF. We first show (3.33). By construction, for a complete mth order MLSRK approximation, there is a one-to-one correspondence between the polynomial basis $P_l(z)$ and the function z^α. Therefore, for $0 \le |\alpha| \le m$, $\exists \, 1 \le k \le \ell$ such that

$$P_k\left(\frac{x_l - x}{\varrho}\right) = \left(\frac{x_l - x}{\varrho}\right)^\alpha$$

$$= \left(\frac{x_{1l} - x_1}{\varrho_1}\right)^{\alpha_1} \left(\frac{x_{2l} - x_2}{\varrho_2}\right)^{\alpha_2} \cdots \left(\frac{x_{nl} - x_n}{\varrho_n}\right)^{\alpha_n} \tag{3.35}$$

Then, for $0 \le |\alpha| \le m$,

$$\sum_{l=1}^{NP} \left(\frac{x_l - x}{\varrho}\right)^\alpha N_l(\varrho, x, x_l) = \sum_{l=1}^{NP} P_k\left(\frac{x_l - x}{\varrho}\right) N_l(\varrho, x, x_l)$$

$$= \delta_{k1} = \delta_{\alpha 0} \tag{3.36}$$

which implies

$$\sum_{l=1}^{NP} (x_l - x)^\alpha N_l(\varrho, x, x_l) = \sum_{l=1}^{NP} (x_l - x)^\alpha \mathcal{H}_\varrho^h(x_l - x, x) \Delta V_l = \delta_{\alpha 0} \tag{3.37}$$

Now, we show (3.34).

$$\sum_{I=1}^{NP} x_I^{\alpha} N_I(\varrho, x, x_I) = \sum_{I=1}^{NP} (x_I - x + x)^{\alpha} N_I(\varrho, x, x_I)$$

$$= \sum_{I=1}^{NP} \left(\sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (x_I - x)^{\alpha - \beta} x^{\beta} \right) N_I(\varrho, x, x_I)$$

$$= \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} x^{\beta} \sum_{I=1}^{NP} (x_I - x)^{\alpha - \beta} N_I(\pi, x, x_I)$$

$$= \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} x^{\beta} \delta_{(\alpha - \beta)0} = x^{\alpha} \quad \square \tag{3.38}$$

Lemma 3.2 leads a profound consequence for the moments of the derivatives of the shape function, which plays a key role in the convergence behavior of the MLSRK approximation.

LEMMA 3.3 (m-consistency condition II). Let $\Phi \in C^m(\Omega)$; for $\Omega \subset \mathbb{R}^n$, the shape functions generated by the complete m-order, ℓ-component polynominal basis (3.19), satisfies the following conditions, $\forall\, 0 \leq |\alpha|, |\beta| \leq m$,

$$\boxed{\sum_{I=1}^{NP} (x_I - x)^{\alpha} D_x^{\beta} N_I(\varrho, x, x_I) = \sum_{I=1}^{NP} (x_I - x)^{\alpha} D_x^{\beta} \mathcal{H}_{\varrho}^h(\varrho, x_I - x, x)\, \Delta V_I = \alpha!\, \delta_{\alpha\beta}} \tag{3.39}$$

or equivalently,

$$\boxed{\sum_{I=1}^{NP} x_I^{\alpha} D_x^{\beta} N_I(\varrho, x, x_I) = \sum_{I=1}^{NP} x_I^{\alpha} D_x^{\beta} \mathcal{H}_{\varrho}^h(x_I - x, x)\, \Delta V_I = \frac{\alpha!}{(\alpha - \beta)!} x^{\alpha - \beta}} \tag{3.40}$$

PROOF. We first prove (3.39). The proof proceeds by induction.

(1) $|\beta| = 0$, $|\alpha| \leq m$; by Lemma 3.1,

$$\sum_{I=1}^{NP} (x_I - x)^{\alpha} N_I(\varrho, x, x_I) = \delta_{\alpha 0} = \alpha!\, \delta_{\alpha 0} \tag{3.41}$$

(2) $|\beta| = 1$, $|\alpha| \leq m$; without loss of generality, one can assume that

$$\beta := (0, 0, \ldots, \beta_j, 0, \ldots, 0) \quad \text{and} \quad \beta_j = 1$$

$$\Rightarrow D_x^{\beta} = \partial_{x_j} \tag{3.42}$$

Hence,

$$D_x^{\beta} \left\{ \sum_{I=1}^{NP} (x_I - x)^{\alpha} N_I(\varrho, x, x_I) \right\} = \partial_{x_j} \left(\sum_{I=1}^{NP} (x_I - x)^{\alpha} N_I(\varrho, x, x_I) \right)$$

$$= -\alpha_j \sum_{I=1}^{NP} (x_{1I} - x_1)^{\alpha_1} \cdots (x_{jI} - x_j)^{\alpha_j - 1} \cdots (x_{nI} - x_n)^{\alpha_n} N_I(\varrho, x, x_I)$$

$$+ \sum_{I=1}^{NP} (x_I - x)^{\alpha} \partial_{x_j} N_I(\varrho, x, x_I)$$

$$= 0 \tag{3.43}$$

This leads to

$$\sum_{l=1}^{NP} (x_l - x)^\alpha D_x^\beta N_l(\varrho, x, x_l) = \alpha_1! \cdots \alpha_j(\alpha_j - 1)! \cdots \alpha_n! \delta_{\alpha_1 0} \cdots \delta_{(\alpha_j - 1)0} \cdots \delta_{\alpha_n 0}$$

$$= \alpha_1! \cdots \alpha_j! \cdots \alpha_n! \delta_{\alpha_1 0} \cdots \delta_{\alpha_j 1} \cdots \delta_{\alpha_n 0}$$

$$= \alpha! \delta_{\alpha\beta} \tag{3.44}$$

Hence, Eq. (3.39) is true for $|\beta| = 0, 1$. Assume that (3.39) holds for $|\beta| \leq m - 1$ and $|\alpha| \leq m$, we need to that show it is true for the case $|\beta'| = m$, and $|\alpha| \leq m$. By assumption

$$\sum_{l=1}^{NP} (x_l - x)^\alpha D_x^\beta N_l(\varrho, x, x_l) = \alpha! \delta_{\alpha\beta} , \tag{3.45}$$

we let

$$\gamma := (0, 0, \ldots, \gamma_j, \ldots, 0) , \qquad \gamma_j = 1 \tag{3.46}$$

Then

$$D_x^\gamma = \partial_{x_j} \Rightarrow D_x^\gamma D_x^\beta = D_x^{\beta'} \tag{3.47}$$

where $|\beta'| = m$.

Differentiation Eq. (3.45) yields

$$- \alpha_j \sum_{l=1}^{NP} (x_{1l} - x_1)^{\alpha_1}(x_{2l} - x_2)^{\alpha_2} \cdots (x_{jl} - x_j)^{\alpha_j - 1} \cdots (x_{jl} - x_j)^{\alpha_n} D_x^\beta N_l(\varrho, x, x_l)$$

$$+ \sum_{l=1}^{NP} (x_l - x)^\alpha D_x^\gamma D_x^\beta N_l(\varrho, x, x_l) = 0 \tag{3.48}$$

It follows that

$$\sum_{l=1}^{NP} (x_l - x)^\alpha D_x^{\beta'} N_l(\varrho, x, x_l)$$

$$= \alpha_j \sum_{l=1}^{NP} (x_{1l} - x_1)^{\alpha_1}(x_{2l} - x_2)^{\alpha_2} \cdots (x_{jl} - x_j)^{\alpha_j - 1} \cdots (x_{jl} - x_j)^{\alpha_n} D_x^\beta N_l(\varrho, x, x_l)$$

$$= \alpha_j(\alpha_1! \alpha_2! \cdots (\alpha_j - 1)! \cdots \alpha_n!) \delta_{\alpha_1 \beta_1} \delta_{\alpha_2 \beta_2} \cdots \delta_{\alpha_j - 1 \beta_j} \cdots \delta_{\alpha_n \beta_n}$$

$$= (\alpha_1! \alpha_2! \cdots \alpha_j! \cdots \alpha_n!) \delta_{\alpha_1 \beta_1} \delta_{\alpha_2 \beta_2} \cdots \delta_{\alpha_j \beta_{j+1}} \cdots \delta_{\alpha_n \beta_n} = \alpha! \delta_{\alpha\beta'} \tag{3.49}$$

We now show (3.3).

$$\sum_{l=1}^{NP} x_l^\alpha D_x^\beta N_l(\varrho, x, x_l) = \sum_{l=1}^{NP} (x_l - x + x)^\alpha D_x^\beta N_l(\varrho, x, x_l)$$

$$= \sum_{l=1}^{NP} \sum_{\gamma \leq \alpha} \binom{\alpha}{\gamma}(x_l - x)^{\alpha - \gamma} x^\gamma D_x^\beta N_l(\varrho, x, x_l)$$

$$= \sum_{\gamma \leq \alpha} \binom{\alpha}{\gamma} x^\gamma \sum_{l=1}^{NP} (x_l - x)^{\alpha - \gamma} D_x^\beta N_l(\varrho, x, x_l)$$

$$= \sum_{\gamma \leq \alpha} \binom{\alpha}{\gamma} x^\gamma (\alpha - \gamma)! \delta_{(\alpha - \gamma)\beta} \tag{3.50}$$

There can be only one term left; i.e. the term $\beta = \alpha - \gamma$. It then follows immediately that

$$\sum_{l=1}^{NP} x_l^\alpha D_x^\beta N_l(\varrho, x, x_l) = \binom{\alpha}{\alpha - \beta} \beta! x^{\alpha - \beta} = \frac{\alpha!}{(\alpha - \beta)!} x^{\alpha - \beta} \qquad \square \tag{3.51}$$

3.3. Interpolation estimate in Hilbert space

In this section, an interpolation estimate for the MLSRK interpolant is given, which is central to the convergence proof of the corresponding Galerkin method and error estimation. Since the reproducing kernel is defined only for the Hilbert space [27], it is pertinent to discuss the interpolation estimation in Hilbert space separately, instead as a special case of Sobolev spaces. The interpolation estimation in Sobolev space will follow shortly.

THEOREM 3.1. Assume $\Phi(x) \in C^m(\bar{\Omega}) \cap W_\infty^m(\Omega)$, and $u(x) \in C^{m+1}(\bar{\Omega}) \cap H^{m+1}(\Omega)$, where Ω is a bounded open set in $\mathbb{R}^n$. Let

$$\mathcal{R}_{\varrho,h}^m u(x) = \sum_{I \in \Lambda} u(x_I) N_I(\varrho, x, x_I) \tag{3.52}$$

where interpolation shape function $N_I(\varrho, x, x_I)$ is generated by a complete m-order, ℓ-component polynominal bases vector (see 3.19); the index set $\Lambda := \{I | 1 \leqslant I \leqslant NP,\ supp\{N_I\} \cap \Omega \neq \emptyset\}$ represents an admissible particle distribution over Ω. Suppose the boundary $\partial\Omega$ is smooth enough, for $0 \leqslant k \leqslant m$, then the following interpolation estimates hold

$$\left\| u - \mathcal{R}_{\varrho,h}^m u \right\|_{H^k(\Omega)} \leqslant C_k \varrho^{m+1-k} \|u\|_{H^{m+1}(\Omega)} \tag{3.53}$$

particularly, for $k = 0$

$$\left\| u - \mathcal{R}_{\varrho,h}^m u \right\|_{L^2(\Omega)} \leqslant C_0 \varrho^{m+1} \|u\|_{H^{m+1}(\Omega)} \tag{3.54}$$

where C_k, C_0 are constants, which are independent with the dilation parameter ϱ.

PROOF. It suffices to show

$$\left| u - \mathcal{R}_{\varrho,h}^m u \right|_{H^k(\Omega)} \leqslant C \varrho^{m+1-k} |u|_{H^{m+1}(\Omega)} \tag{3.55}$$

By definition

$$\mathcal{R}_{\varrho,h}^m u(x) = \sum_{I \in \Lambda} u(x_I) N_I(\varrho, x, x_I) \tag{3.56}$$

Taking the deivative of (3.56) yields

$$D_x^\beta \mathcal{R}_{\varrho,h}^m u(x) = \sum_{I \in \Lambda} u(x_I) D_x^\beta N_I(\varrho, x, x_I) \tag{3.57}$$

By Taylor expansion

$$u(x_I) = \sum_{|\alpha| \leqslant m} \frac{1}{\alpha!} (x_I - x)^\alpha D_x^\alpha u(x) + \sum_{|\alpha| = m+1} \frac{1}{\alpha!} (x_I - x)^\alpha D_x^\alpha u(x + \theta(x_I - x)) \tag{3.58}$$

where $0 < \theta < 1$. Substituting (3.58) back to (3.57) yields

$$D_x^\beta(\mathcal{R}_{\varrho,h}^m u(x)) = \sum_{I \in \Lambda} \left\{ \sum_{|\alpha| \leqslant m} \frac{(x_I - x)^\alpha}{\alpha!} D_x^\alpha u(x) + \sum_{|\alpha| = m+1} \frac{(x^I - x)^\alpha}{\alpha!} D_x^\alpha u(x + \theta(x_I - x)) \right\} \cdot D_x^\beta N_I(\varrho, x, x_I)$$

$$= \sum_{|\alpha| \leqslant m} \frac{1}{\alpha!} D_x^\alpha(x) \left(\sum_{I \in \Lambda} (x_I - x)^\alpha D_x^\beta N_I(\varrho, x, x_I) \right)$$

$$+ \sum_{I \in \Lambda} \sum_{|\alpha| = m+1} \frac{1}{\alpha!} (x_I - x)^\alpha D_x^\alpha u(x + \theta(x_I - x)) D_x^\beta N_I(\varrho, x, x_I)$$

$$= \sum_{|\alpha| \leqslant m} \frac{1}{\alpha!} D_x^\alpha u(x) \alpha! \delta_{\alpha\beta} \quad \leftarrow \text{by Lemma 3.3}$$

$$+ \sum_{I \in A} \sum_{|\alpha| = m+1} \frac{1}{\alpha!} (x_I - x)^\alpha D_x^\alpha u(x + \theta(x_I - x)) D_x^\beta N_I(\varrho, x, x_I) \tag{3.59}$$

If $\partial \Omega$ is smooth enough, one can carefully choose a particle distribution and dilation parameter ϱ such that each sub-domain $\text{supp}\{N_I\} \cap \Omega$ is *star-shaped* with respect to x_I, i.e. $\forall x \in \text{supp}\{N_I\} \cap \Omega$,

$$x + \theta(x_I - x) \in \text{supp}\{N_I\} \cap \Omega \tag{3.60}$$

thus, the above expansion always make sense. It should be noted that this is a rather loose condition, which does not require domain Ω to be convex.

It follows then

$$|D_x^\beta - D_x^\beta \mathcal{R}_{\varrho,h}^m u| \leqslant \sum_{I \in A} \sum_{|\alpha| = m+1} |x_I - x|^\alpha |D_x^\alpha u(x + \theta(x_I - x))| \cdot |D_x^\beta N_I(\varrho, x, x_I)| \tag{3.61}$$

If $x \in \text{supp}\{N_I(\varrho, x, x_I)\}$ then $\exists \, r > 0$, such that $|x_I - x| \leqslant r\varrho$; therefore

$$|D_x^\beta u - D_x^\beta \mathcal{R}_{\varrho,h}^m u| \leqslant \sum_{I \in A(x)} \sum_{|\alpha| = m+1} |\alpha_I - x|^\alpha |D_x^\alpha u(x + \theta(x_I - x))| \, |D_x^\beta N_I(\varrho, x, x_I)|$$

$$\leqslant C_m r^{m+1} \varrho^{m+1} \sum_{I \in A(x)} \sum_{|\alpha| = m+1} |D_x^\alpha u(x + \theta(x_I - x))| \, |D_x^\beta N_I(\varrho, x, x_I)| \tag{3.62}$$

where $A(x) := \{I \in A | x \in \text{supp}\{N\} \cap \overline{\Omega}\}$.

By the Cauchy's inequality, one will have the estimate,

$$|D_x^\beta u - D_{\varrho,h}^\beta u)|^2 C \left\{ \sum_{I \in A(x)} \left(\sum_{|\alpha| = m+1} D_x^\alpha u(x + \theta(x_I - x)) \right)^2 (D_x^\beta N_I(\varrho, x, x_I)) \right\} \cdot \left\{ \sum_{J \in A(x)} D_x^\beta N_J(\varrho, x, x_J) \right\} \tag{3.63}$$

Let

$$z := \frac{x}{\varrho}, \quad \text{or} \quad z_i = \frac{x_i}{\varrho_i} \tag{3.64}$$

One has

$$D_x^\beta N_I(\varrho, x, x_I) = D_z^\beta \tilde{N}_I(z, x_I) D_x^\beta z \tag{3.65}$$

where

$$D_x^\beta z := \left(\frac{\partial z_1}{\partial x_1} \right)^{\beta_1} \left(\frac{\partial z_2}{\partial x_2} \right)^{\beta_2} \cdots \left(\frac{\partial x_n}{\partial x_n} \right)^{\beta_n}$$

$$= \varrho_1^{-\beta_1} \varrho_2^{-\beta_2} \cdots \varrho_n^{-\beta_n} = \varrho^{-\beta} \tag{3.66}$$

$$\Rightarrow \quad |D_x^\beta z| = \varrho^{-k} \tag{3.67}$$

Thereby

$$|D_x^\beta N_I(\varrho, x, x_I)| \leqslant |D_x^\beta z| |D_z^\beta \tilde{N}_I(z, x_I)|$$

$$\leqslant \varrho^{-k} |D_z^\beta \tilde{N}_I(z, x_I)| \tag{3.68}$$

On the other hand, by the Leibniz rule

$$D_z^\beta \tilde{N}_I(z, x_I) = \left(\sum_{\gamma \leqslant \beta} \binom{\beta}{\gamma} D_z^\gamma \mathcal{C}(z, z_I) D_z^{\beta - \gamma}(z_I - z) \right) \left(\frac{\Delta V_I}{\varrho^n} \right); \tag{3.69}$$

Considering the stability condition (2.85), $\Delta_I \leqslant r^n \varrho^n$, then

$$\left(\frac{\Delta V_I}{\varrho^n} \right) \sim \mathcal{O}(1) \tag{3.70}$$

and Cauchy's inequality, we have

$$|D_z^\beta \tilde{N}_I(z, x_I)| \leq C \sqrt{\sum_{\gamma \leq \beta} \left[\binom{\beta}{\gamma} D_z^\gamma \mathscr{C}(z, z_I)\right]^2} \sqrt{\sum_{\gamma \leq \beta} [D_z^{\beta-\gamma} \Phi_I(z_I - z)]^2} . \tag{3.71}$$

Since $\Phi(\cdot) \in C^m(\Omega)$, by Lemma 2.1, $\mathscr{C} \in C^m(\Omega)$, and $|D_z^\beta(z, z_I)|$ is bounded over Ω. Furthermore, from Lemma 2.1, one can observe that the correction function $\mathscr{C}$ is a function of ϱ at zero order, as $\varrho \to 0$. Thus, $D_z^\gamma \mathscr{C}(z, z_I)$ will be also a function of ϱ at zero order. Therefore, there exists a constant $\mathscr{C}_{z_1}$ such that

$$\mathscr{C}_{z_1} := \sup_{x \in \Omega, I \in \Lambda} \sqrt{\sum_{\gamma \leq \beta} \left[\binom{\beta}{\gamma} D_z^\gamma \mathscr{C}\right]^2} < \infty \tag{3.72}$$

where $\mathscr{C}_{z_1}$ is independent of ϱ.

Hence, the derivatives of the shape function can be always bounded as follows:

$$|D_z^\beta \tilde{N}_I(z, x_I)| \leq C_{\beta, \Omega} |\Phi|_{W_\infty^k(\Omega)} \tag{3.73}$$

From the pointwise overlap condition (2.91), for fixed x,

$$\text{card}\{\Lambda(x)\} \leq N_{\max} , \tag{3.74}$$

then $\exists I \in \Lambda(x)$, the following pointwise estimate holds

$$|D_x^\beta(u - \mathscr{R}_{\varrho,h}^m u)|^2 \leq C_{\alpha,\beta,\Omega} N_{\max}^2 |\Phi|_{W_\infty^k(\Omega)}^2 r^{2(m+1)} \varrho^{2(m+1-k)} \left(\sum_{|\alpha|=m+1} (D^\alpha u(x + \theta(x_I - x)))^2\right) \tag{3.75}$$

One can readily show that

$$\|D_x^\beta u - D_x^\beta \mathscr{R}_{\varrho,h}^m u\|_{L^2(\Omega)}^2 \leq C_{\alpha,\beta,\Omega,\phi,\theta} N_{\max}^2 r^{2(m+1)} \varrho^{2(m+1-k)} \int_\Omega \sum_{|\alpha|=m+1} |D_x^\alpha u(x)|^2 \, \mathrm{d}\Omega_x \tag{3.76}$$

Then, we conclude that $\exists 0 < C < \infty$,

$$|u - \mathscr{R}_{\varrho,h}^m u|_{H^k(\Omega)} \leq C \varrho^{m+1-k} |u|_{H^{m+1}(\Omega)} \quad k = 0, 1, \ldots, m \tag{3.77}$$

or

$$\|u - \mathscr{R}_{\varrho,h}^m u\|_{H^k(\Omega)} \leq C_k \varrho^{m+1-k} \|u\|_{H^m(\Omega)} \quad k = 0, 1, \ldots, m \quad \square \tag{3.78}$$

3.4. Interpolation estimate in Sobolev space

From theoretical perspective, the continuous version of the MLSRK representation is more fundamental than its discrete counterpart. In this section, a Bramble–Hilbert type theorem is asserted for the MLSRK integral approximation.

THEOREM 3.2. Assume that Ω is a convex, compact region in $\mathbb{R}^n$. $\forall 1 \leq p < \infty$, $(1/p) + (1/q) = 1$, assume $u \in C^{m+1}(\Omega) \cap W_p^{m+1}(\Omega)$ and $\Phi \in C^m(\Omega) \cap W_q^m(\Omega)$. Let

$$\mathscr{R}_\varrho^m u(x) := \int_\Omega u(y) \mathscr{K}_\varrho(y - x, x) \, \mathrm{d}\Omega_y \tag{3.79}$$

for $0 \leq k \leq m$, $m > n/p$, the following interpolation estimate holds

$$\boxed{|u - \mathscr{R}_\varrho^m u|_{W_p^k(\Omega)} \leq C_1 \varrho^{m+1-k} \|\Phi\|_{W_q^k(\Omega)} |u|_{W_p^{m+1}(\Omega)}} \tag{3.80}$$

or

$$\|u - \mathcal{R}_\varrho^m u\|_{W_p^k(\Omega)} \le C_2 \varrho^{m+1-k} \|\Phi\|_{W_q^k(\Omega)} \|u\|_{W_p^{m+1}(\Omega)}$$ (3.81)

Before we prove Theorem 3.2, few lemmas are needed. However these lemmas are the continuous counterparts of Lemmas 3.2 and 3.3. The proofs of following Lemmas are omitted, because they are similar to the proofs of Lemmas 3.2 and 3.3.

LEMMA 3.4. *The MLSRK approximation, generated by the complete m-order, ℓ-component polynomial basis vector* $P(y - x)/\varrho := [P_1, P_2, \ldots, P_\ell](y - x)/\varrho$, *has the properties*

$$\int_\Omega (y - x)^\alpha \mathcal{K}_\varrho(y - x, x) \, d\Omega_y = \delta_{\alpha 0} \quad \forall \, 0 \le |\alpha| \le m$$ (3.82)

or equivalently,

$$\int_\Omega y^\alpha \mathcal{K}_\varrho(y - x, x) \, d\Omega_y = x^\alpha .$$ (3.83)

LEMMA 3.5. *The derivatives of the MLSRK function generated by the complete m-order, ℓ-component polynomial satisfy the following conditions,* $\forall \, 0 \le |\alpha|, |\beta| \le m$,

$$\int_\Omega (y - x)^\alpha D_x^\beta \mathcal{K}_\varrho(y - x, x) \, d\Omega_y = \alpha! \delta_{\alpha\beta}$$ (3.84)

or equivalently

$$\int_\Omega y^\alpha D_x^\beta \mathcal{K}_\varrho(y - x, x) \, d\Omega_y = \frac{\alpha!}{(\alpha - \beta)!} x^{\alpha-\beta}$$ (3.85)

PROOF [Theorem 3.2]. It suffices to show (3.80). By Taylor's theorem, $\forall f(x) \in C^{m+1}([0, 1])$, one has

$$f(1) = \sum_{k=0}^m \frac{1}{k!} f^{(k)}(0) + \frac{(m+1)}{(m+1)!} \int_0^1 s^m f^{(m+1)}(1 - s) \, ds$$ (3.86)

Assume Ω is convex. $\text{supp}\{\mathcal{K}_\varrho(y - x, x)\} \cap \Omega$ must be also convex; therefore $\forall s \in [0, 1]$,

$$z = x + s(y - x) \in \text{supp}\{\mathcal{K}_\varrho(y - x, x)\} \cap \Omega$$ (3.87)

Suppose $u(x) \in C^{m+1}(\Omega) \cap W_p^{m+1}(\Omega)$. One can define the real function $f(s)$ as

$$f(s) := u(x + s(y - x))$$ (3.88)

Then by the chain rule, it can be verified that

$$\frac{1}{k!} f^{(k)}(s) = \sum_{|\alpha|=k} \frac{1}{\alpha!} D^\alpha u(x + s(y - x))(y - x)^\alpha$$ (3.89)

Let $s = 1$ in (3.88) and considering (3.86) and (3.89), it yields

$$u(y) = \sum_{|\alpha| \le m} \frac{1}{\alpha!} D^\alpha u(x)(y - x)^\alpha + \sum_{|\alpha|=m+1} \frac{(y - x)^\alpha}{\alpha!} \int_0^1 (m + 1) s^m D^\alpha u(y + s(x - y)) \, ds$$ (3.90)

By definition,

$$\mathcal{R}_\varrho^m u(x) = \int_\Omega u(y) \mathcal{K}_\varrho(y - x, x) \, d\Omega_y .$$ (3.91)

then

$$D_x^\beta(\mathcal{R}_\varrho^m u(x)) = \int_\Omega u(y) D_x^\beta \mathcal{H}_\varrho(y - x, x) \, d\Omega_y \tag{3.92}$$

Substituting (3.90) into (3.92) yields

$$
\begin{aligned}
D_x^\beta(\mathcal{R}_\varrho^m u(x)) &= \int_\Omega \left\{ \sum_{|\alpha| \leqslant m} \frac{1}{\alpha!} D_x^\alpha u(x)(y - x)^\alpha \right. \\
&\quad \left. + \sum_{|\alpha| = m+1} (m+1) \frac{(y-x)^\alpha}{\alpha!} \int_0^1 s^m D^\alpha u(y + s(x - y)) \, ds \right\} D_x^\beta \mathcal{H}_\varrho(y - x, x) \, d\Omega_y \\
&= \sum_{|\alpha| \leqslant m} D_x^\alpha u(x) \left(\int_\Omega \frac{1}{\alpha!} (y - x)^\alpha D_x^\beta \mathcal{H}(y - x, x) \, d\Omega_y \right) \\
&\quad + \int_\Omega \sum_{|\alpha| = m+1} \frac{(y-x)^\alpha}{\alpha!} \left\{ \int_0^1 (m+1) s^m D^\alpha u(y + s(x - y)) \, ds \right\} D_x^\beta \mathcal{H}_\varrho(y - x, x) \, d\Omega_y \\
&= \sum_{|\alpha| \leqslant m} D_x^\alpha u(x) \delta_{\alpha\beta} \quad \Leftarrow \text{By Lemma 3.5} \\
&\quad + \int_\Omega \sum_{|\alpha| = m+1} \frac{(y-x)^\alpha}{\alpha!} \left\{ \int_0^1 (m+1) s^m D^\alpha u(y + s(x - y)) \, ds \right\} D_x^\beta \mathcal{H}_\varrho(y - x, x) \, d\Omega_y
\end{aligned} \tag{3.93}
$$

Let $\zeta := x/\varrho$, $\eta := y/\varrho$ and consider the fact that

$$|y - x| < r\varrho \quad y, x \in \text{supp}\{\mathcal{H}_\varrho(y - x, x)\} \tag{3.94}$$

$\forall |\beta| = k$, $0 \leqslant k \leqslant m$, one has

$$
\begin{aligned}
|D_x^\beta u(x) - D_x^\beta \mathcal{R}_\varrho^m u(x)| &\leqslant \left| \int_\Omega \left\{ \sum_{|\alpha| = m+1} \frac{(y-x)^\alpha}{\alpha!} \left(\int_0^1 (m+1) s^m D^\alpha u(y + s(x - y)) \, ds \right) \right. \right. \\
&\quad \left. \left. \cdot (D_x^\beta \mathcal{H}(y - x, x)) \right\} \, d\Omega_y \right| \\
&\leqslant EC_m \int_\Omega \left\{ \sum_{|\alpha| \leqslant m} |y - x|^\alpha \left(\int_0^1 s^m |D^\alpha u(y + s(x - y))| \, ds \right) \right. \\
&\quad \left. \cdot (|D_\zeta^\beta \mathcal{H}(\eta - \zeta, \zeta)| |D_x^\beta \zeta|) \right\} \, d\Omega_y \\
&\leqslant C_{m,n,r} \varrho^{m+1-k} \int_\Omega \left\{ \sum_{|\alpha| = m+1} \left(\int_0^1 s^m |D^\alpha u(y + s(x - y))| \, ds \right) \right. \\
&\quad \left. \cdot (|D_\zeta^\beta \mathcal{H}(\eta - \zeta, \zeta)|) \frac{1}{\varrho^n} \right\} \, d\Omega_y
\end{aligned} \tag{3.95}
$$

where $\mathcal{H}(\eta - \zeta, \zeta) = \mathcal{C}(1, \eta - \zeta, \zeta) \Phi(\eta - \zeta)$.
Then, by Hölder's inequality

$$
\begin{aligned}
|D_z^\beta \mathcal{H}(\eta - \zeta, \zeta)| &= \left| \sum_{\gamma \leqslant \beta} \binom{\beta}{\gamma} (D_\zeta^{\beta-\gamma} \mathcal{C})(D_\zeta^\gamma \Phi) \right| \\
&\leqslant \left\{ \sum_{\gamma \leqslant \beta} \left[\binom{\beta}{\gamma} D_\zeta^{\beta-\gamma} \mathcal{C} \right]^p \right\}^{1/p} \left\{ \sum_{\gamma \leqslant \beta} [D_\zeta^\gamma \Phi]^q \right\}^{1/q}
\end{aligned} \tag{3.96}
$$

Since

$$\sup_{x \in \Omega} \left(\sum_{\gamma \leqslant \beta} \left[\binom{\beta}{\gamma} D_\zeta^\gamma \mathcal{C} \right]^p \right)^{1/p} < +\infty \tag{3.97}$$

we have

$$|D_x^\beta(u(x) - \mathcal{R}_\varrho^m u(x))| \leq C_{k,m,n,r} \varrho^{m+1-k} \int_\Omega \left\{ \sum_{|\alpha|=m+1} \left(\int_0^1 (m+1) s^m D^\alpha u(y + s(x - y)) \, ds \right) \right.$$

$$\left. \cdot \left[\sum_{\gamma \leq \beta} |D_\zeta^\beta \Phi|^q \right]^{1/q} \frac{1}{\varrho^n} \right\} \, d\Omega_y \tag{3.98}$$

Consequently,

$$\|D_x^\beta(u(x) - \mathcal{R}_\varrho^m u(x))\|_{L^p(\Omega)}^p = \int_\Omega |D_x^\beta(u - \mathcal{R}_\varrho^m u)|^p \, d\Omega_x$$

$$\leq C_{k,m,n,r} \varrho^{p(m+1-k)} \int_{\Omega_x} \left\{ \int_{\Omega_y} \left(\sum_{|\alpha|=m+1} \left(\int_0^1 s^m |D^\beta u(y + s(x - y))| \, ds \right) \right) \right.$$

$$(\chi_{\mathcal{K}_\varrho(y-x,x)}) \left(\sum_{\gamma \leq \beta} |D_\zeta^\beta \Phi|^q \right)^{1/q} d\Omega_y \right\}^p d\Omega_x$$

$$\leq C_{k,m,n,r} \varrho^{p(m+1-k)} \int_{\Omega_x} \left\{ \left(\int_{\Omega_y} \chi_{\mathcal{K}_\varrho(y-x,x)} \left[\sum_{|\alpha|=m+1} \left(\int_0^1 s^m |D^\beta u(y + s(x - y))| \, ds \right) \right]^p d\Omega_y \right) \right.$$

$$\left(\int_{\Omega_y} \left(\sum_{\gamma \leq \beta} |D_\zeta^\gamma \Phi|^q \right) \varrho^{-nq} \, d\Omega_y \right)^{p/q} \right\} d\Omega_x$$

$$\leq C_{k,m,n,r} \varrho^{p(m+1-k)-n} \|\Phi\|_{W_q^k(\Omega)}^p \int_{\Omega_x} \int_{\Omega_y} (\chi_{\mathcal{K}_\varrho(y-x,x)})$$

$$\left(\sum_{|\alpha|=m+1} \left(\int_0^1 s^m |D^\alpha u(y + s(x - y))| \, ds \right) \right)^p d\Omega_y \, d\Omega_x \tag{3.99}$$

where $\chi_{\mathcal{K}_\varrho(y-x,x)}$ is the characteristic function of compact support of $\mathcal{K}_\varrho(y - x, x)$, i.e. for fixed y,

$$\chi_{\mathcal{K}_\varrho(y-x,x)} := \begin{cases} 1 & \forall x \in \operatorname{supp}\{\mathcal{K}_\varrho(y - x, x)\} \\ 0 & \text{otherwise} \end{cases} \tag{3.100}$$

Making change of variable, for fixed y, we let

$$z := y + s(x - y) \tag{3.101}$$

then $d\Omega_z = s^n \, d\Omega_x$, thus

$$\|D_x^\beta(u - \mathcal{R}_\varrho^m u)\|_{L^p(\Omega)}^p \leq C_{(k,m,n,r,\theta)} \varrho^{p(m+1-k)-n} \|\Phi\|_{W_q^k(\Omega)}^p$$

$$\times \left(\left[\int_0^1 s^{m-(n/p)} \, ds \right]^p \int_{\Omega_z} \left[\sum_{|\alpha|=m+1} [D_z^\alpha u(z)]^p \int_{\Omega_y} \chi_{\mathcal{K}_\varrho(y-x,x)} \, d\Omega_y \right] d\Omega_z \right) \tag{3.102}$$

If $m > n/p$,

$$\|D_x^\beta(u - \mathcal{R}_\varrho^m u)\|_{L^p(\Omega)}^p \leq C_{(k,m,n,p,r,\theta)} \varrho^{p(m+1-k)} \|\Phi\|_{W_q^k(\Omega)}^p |u|_{W_p^{m+1}(\Omega)}^p \tag{3.103}$$

Hence, $\exists \, 0 < C < \infty$ such that

$$|u - \mathcal{R}u|_{W_p^k(\Omega)} \leq C\varrho^{m+1-k} \|\Phi\|_{W_q^k(\Omega)} \|u\|_{W_p^{m+1}(\Omega)} \tag{3.104}$$

and (3.81) follows readily. □

4. Galerkin procedures and numerical example

4.1. Error estimation for elliptic partial differential equation

The interpolation estimate obtained in last section can be used to derive the abstract error bound for approximate solutions obtained by using the moving least-square kernel Galerkin method. In fact, a better convergence rate has been observed for the MLSRK method than that of the conventional finite element method. This issue will be further explored in the Part II of this series [11]. To illustrate the general procedure, an error estimate for elliptical partial differential equations is considered.

4.1.1. The Neumann problem

Since natural boundary condition problems require fewer restrictions on both trial functions and weighting functions, it is convenient to consider the following model problem—a Neumann problem for the second-order elliptic partial differential equation

$$L(u) = -\nabla^2 u + u = f(x) \quad x \in \Omega \tag{4.1}$$

$$\frac{\partial u}{\partial n} = g(x) \quad x \in \partial \Omega \tag{4.2}$$

where f, g are assumed to satisfy sufficient regularity requirements.

Let us start with the bilinear form $a(u, v)$

$$a(u, v) := \int_{\Omega} (\nabla u \cdot \nabla v + u \cdot v) \, d\Omega \tag{4.3}$$

It is obvious that $a(u, v)$ is coercive on $H_1(\Omega)$,

$$a(u, u) = \|u\|^2_{H^1(\Omega)} \geq \gamma \|u\|^2_{H^1(\Omega)} \quad \forall u \in H^1(\Omega) \tag{4.4}$$

where $0 < \gamma \leq 1$. Or we should say that $a(u, v)$ is a $H^1(\Omega)$-elliptic form. It is also straightforward to show that $a(u, v)$ is continuous, i.e. $\exists C > 0$ such that

$$a(u, v) \leq C \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \tag{4.5}$$

As a matter of fact, by Cauchy's inequality

$$
\begin{aligned}
a(u, v) &\leq \sqrt{\int_{\Omega} (\nabla u)^2 \, d\Omega} \sqrt{\int_{\Omega} (\nabla v)^2 \, d\Omega} + \sqrt{\int_{\Omega} (u)^2 \, d\Omega} \sqrt{\int_{\Omega} (v)^2 \, d\Omega} \\
&\leq \left(\sqrt{\int_{\Omega} (\nabla u)^2 \, d\Omega} + \sqrt{\int_{\Omega} (u)^2 \, d\Omega} \right) \left(\sqrt{\int_{\Omega} (\nabla v)^2 \, d\Omega} + \sqrt{\int_{\Omega} (v)^2 \, d\Omega} \right) \\
&\leq 2(\|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)})
\end{aligned}
\tag{4.6}
$$

In the last step, the arithmetic–geometric inequality is used.

Then, by the Lax–Milgram theorem, the original problem (4.1)–(4.2) is equivalent to the following variational formulation (weak form),

$$\begin{cases} \text{Find } u \in H^1(\Omega) \text{ such that } \forall v \in H^1(\Omega) \\ \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, d\Omega = \int_{\Omega} fv \, d\Omega + \int_{\partial \Omega} gv \, d\partial \Omega \end{cases} \tag{4.7}$$

Let

$$\mathcal{V}^m_\varrho(\Omega) := \text{span}\{N^m_I | I \in \Lambda; \quad \text{supp}\{N^m_I\} \cap \Omega \neq \emptyset\} \tag{4.8}$$

Here, the superscript m indicates that the shape function is constructed by the complete mth order polynomial. Clearly, $\mathcal{V}^m_\varrho(\Omega) \subset C^m(\Omega) \subset H^1(\Omega)$ provided that $m \geq 1$.

Then, we formulate the moving least-square kernel Galerkin problem I as

$$\text{MLSRKG(I)}\begin{cases}\text{Find } u^\varrho \in \mathscr{V}_\varrho^m(\Omega) \text{ such that } \forall\, v^\varrho \in \mathscr{V}_\varrho^m(\Omega)\\[6pt]\displaystyle\int_\Omega (\nabla u^\varrho \cdot \nabla v^\varrho + u^\varrho v^\varrho)\,\mathrm{d}\Omega = \int_\Omega f v^\varrho\,\mathrm{d}\Omega + \int_{\partial\Omega} g v^\varrho\,\mathrm{d}\partial\Omega\end{cases} \tag{4.9}$$

For the moving least-square Galerkin solution u^ϱ of (4.9), we have the following error estimate, which is based on the celebrated Céa Lemma [22].

THEOREM 4.1. Let $u \in C^{m+1}(\Omega)$; if u is the solution of Neumann problem (4.1)–(4.2), and $u^\varrho \in \mathscr{V}_0^m(\Omega)$ is the solution of weak formulation (4.9), then $\exists\, C_0,\ C_1 > 0$ such that

$$\|u - u^\varrho\|_{H^1(\Omega)} \le C_1 \varrho^m \|u\|_{H^{m+1}(\Omega)} \tag{4.10}$$

and

$$\|u - u^\varrho\|_{L^2(\Omega)} \le C_0 \varrho^{m+1} \|u\|_{H^{m+1}(\Omega)} \tag{4.11}$$

where the constants, C_0, C_1, do not depend on dilation parameter ϱ.

PROOF. We first show (4.10). Since $v^\varrho \in \mathscr{V}_0^m(\Omega) \subset H^1(\Omega)$

$$a(u, v^\varrho) = \int_\Omega f v^\varrho\,\mathrm{d}\Omega + \int_{\partial\Omega} g v^\varrho\,\mathrm{d}\partial\Omega \tag{4.12}$$

$$a(u^\varrho, v^\varrho) = \int_\Omega f v^\varrho\,\mathrm{d}\Omega + \int_{\partial\Omega} g v^\varrho\,\mathrm{d}\partial\Omega \tag{4.13}$$

Subtraction (4.13) from (4.12) yields

$$a(u - u^\varrho, v^\varrho) = 0, \quad \forall\, v^\varrho \in \mathscr{V}_0^m(\Omega) \tag{4.14}$$

thus

$$\begin{aligned}
\|u - u^\varrho\|_{H^1(\Omega)}^2 &= a(u - u^\varrho, u - u^\varrho)\\
&= a((u - u^\varrho), (u - v^\varrho) + (v^\varrho - u^\varrho))\\
&= a(u - u^\varrho, u - v^\varrho) + a(u - u^\varrho, v^\varrho - u^\varrho)\\
&= a(u - u^\varrho, u - v^\varrho) \qquad \Leftarrow v^\varrho - u^\varrho \in \mathscr{V}_0^m(\Omega)\\
&\le C\|u - u^\varrho\|_{H^1(\Omega)}\|u - v^\varrho\|_{H^1(\Omega)} \qquad \Leftarrow \text{ by continuity of } a(u, v)
\end{aligned} \tag{4.15}$$

Thus, $\forall\, v^\varrho \in \mathscr{V}_0^m(\Omega)$,

$$\begin{aligned}
\|u - u^\varrho\|_{H^1(\Omega)} &\le C \inf_{v^\varrho \in \mathscr{V}_0^m(\Omega)} \|u - v^\varrho\|_{H^1(\Omega)}\\
&= C \lim_{v^\varrho \in \mathscr{V}_0^m(\Omega)} \|u - v^\varrho\|_{H^1(\Omega)}
\end{aligned} \tag{4.16}$$

Since v^ϱ is an arbitrary element in $\mathscr{V}_0^m(\Omega)$, let

$$v^\varrho = \mathscr{R}_{\varrho,h}^m u \tag{4.17}$$

Note that $u^\varrho \ne \mathscr{R}_{\varrho,h}^m u$! Consequently, by the interpolation estimate (3.53), we obtain

$$\begin{aligned}
\|u - u^\varrho\|_{H^1(\Omega)} &\le C\|u - \mathscr{R}_{\varrho,h}^m u\|_{H^1(\Omega)}\\
&\le C_1 \varrho^m \|u\|_{H^{m+1}(\Omega)}
\end{aligned} \tag{4.18}$$

Next, we show (4.11). The procedure is the standard duality technique known as Nitsche's trick. Let $e^{\varrho} = u - u^{\varrho}$ and consider the following auxiliary problem,

$$L(w) = e^{\varrho} \tag{4.19}$$

The corresponding weak solution satisfies

$$a(w, v) = (e^{\varrho}, v), \quad \forall v \in \mathcal{V}_{\varrho}^{m}(\Omega) \tag{4.20}$$

Choosing $v = e^{\varrho}$ and considering

$$a(v^{\varrho}, u - u^{\varrho}) = a(v^{\varrho}, e^{\varrho}) \quad \forall v^{\varrho} \in \mathcal{V}_{\varrho}^{m}(\Omega) \tag{4.21}$$

one has

$$\|e^{\varrho}\|_{L^{2}(\Omega)} = a(w, e^{\varrho}) = a(w - v^{\varrho}, e^{\varrho}) \tag{4.22}$$

Let $v^{\varrho} = w^{\varrho}$. By Cauchy's inequality, the above expression can be bounded as follows:

$$
\begin{aligned}
\|e^{\varrho}\|_{L^{2}(\Omega)}^{2} &= a(w - w^{\varrho}, u - u^{\varrho}) \\
&\leq C\|w - w^{\varrho}\|_{H^{1}(\Omega)}\|u - u^{\varrho}\|_{H^{1}(\Omega)} \quad \Leftarrow \text{ by continuity of } a(u, v) \\
&\leq CC'\varrho\|w\|_{H^{2}(\Omega)}\|u - u^{\varrho}\|_{H^{1}(\Omega)} \quad \Leftarrow \text{ by (3.53)}
\end{aligned}
$$

On the other hand, the continuous dependence of the solution on the data requires

$$\|w\|_{H^{2}(\Omega)} \leq C''\|e^{\varrho}\|_{L^{2}(\Omega)} \tag{4.23}$$

Thus,

$$\|e^{\varrho}\|_{L^{2}(\Omega)}^{2} \leq CC'C''\varrho\|e^{\varrho}\|_{L^{2}(\Omega)}\|u - u^{\varrho}\|_{H^{1}(\Omega)} \tag{4.24}$$

which finally leads to

$$\|u - u^{\varrho}\|_{L^{2}(\Omega)} \leq CC'C''\varrho\|u - u^{\varrho}\|_{H^{1}(\Omega)} \leq C_{0}\varrho^{m+1}\|u\|_{H^{m+1}(\Omega)} \quad \square \tag{4.25}$$

In most numerical computations, the interpolated function $f^{\varrho}(x)$, $g^{\varrho}(x)$ are used instead of exact input data function $f(x)$ and $g(x)$. Generally speaking, the moving least-square Galerkin methods adopted are usually 'non-interpolation' schemes; this makes the numerical computation extremely intriguing. In the actual computations, the following scheme is implemented:

$$f^{\varrho}(x) := \mathcal{R}_{\varrho,h}^{m} f(x) \quad x \in \Omega \tag{4.26}$$

$$g^{\varrho}(x) := \mathcal{R}_{\varrho,h}^{m} \tilde{g}(x) \quad x \in \overline{\Omega} \tag{4.27}$$

where function $\tilde{g}(x) \in H^{1}(\overline{\Omega}) \cap C^{1}(\Omega)$ or even more smooth, such that

$$\tilde{g}(x) := \begin{cases} g, & x \in \partial\Omega \\ \text{continuous function}; & x \in \Omega \end{cases} \tag{4.28}$$

In this manner, the so-called moving least-square kernel Galerkin method is based on the following variational formulation

$$\text{MLSRKG(II):} \begin{cases} \text{Find } u^{\varrho} \in \mathcal{V}_{\varrho}^{m}(\Omega) \text{ such that } \forall v^{\varrho} \in \mathcal{V}_{\varrho}^{m}(\Omega) \\ \int_{\Omega} (\nabla u^{\varrho} \nabla v^{\varrho} + u^{\varrho} v^{\varrho}) \, d\Omega = \int_{\Omega} f^{\varrho} v^{\varrho} \, d\Omega + \int_{\partial\Omega} g^{\varrho} v^{\varrho} \, d\partial\Omega \end{cases} \tag{4.29}$$

REMARK 4.1. In (4.29), the expression $\int_{\partial\Omega} g^{\varrho} v^{\varrho} \, d\partial\Omega$ should be interpreted as

$$\int_{\partial\Omega} g^{\varrho}(\varrho, y, x_{I}) v^{\varrho}(y, x_{I}) \, dy \tag{4.30}$$

where $y \in \partial\Omega, x_{I} \in \overline{\Omega}$. $\quad\square$

Based on the formulation MLSRKG(II) (4.29), we have the following error estimation.

THEOREM 4.2. Assume $\partial\Omega$ is Lipschitz; the solution of (4.7) $u(x) \in C^{m+1}(\Omega)$, and also the input data are sufficiently smooth and regular $f \in C^{m+1}(\Omega)$, $g \in C^{m+1}(\partial\Omega) \Rightarrow \tilde{g} \in C^{m+1}(\overline{\Omega})$. The approximation error of the solution of Neumann problem (4.1)–(4.2) based on formulation MLSRKG(II) (4.29) is bounded by

$$\boxed{\|u - u^\rho\|_{H^1(\Omega)} \leqslant C_1 \rho^m}$$

(4.31)

and

$$\boxed{\|u - u^\rho\|_{L^2(\Omega)} \leqslant C_0 \rho^{m+1}}$$

(4.32)

where the constants C_0, C_1 do not depend on the dilation parameter ρ.

Since the proof is standard and it is very similar to the proof of Theorem 4.1, we omit the detail here.

4.1.2. The Dirichlet problem

Here, let us consider the following model problem

$$-\nabla^2 u + u = f(x) \quad x \in \Omega$$

(4.33)

$$u = h(x) \quad x \in \partial\Omega$$

(4.34)

To solve this problem, the conventional finite element method is to minimize the functional

$$J(u) = \int_\Omega [(\nabla u) \cdot (\nabla u) + u \cdot u] \, d\Omega - 2 \int_\Omega fu \, d\Omega$$

(4.35)

Unlike the finite element interpolating shape function, most moving least-square interpolant based trial functions do not satisfy the essential boundary condition generally unless special care has been taken. Thus, the classical technique cannot be legally used in the computation. An early alternative suggested by Belytschko et al. [4] is to seek the stationary point of the following functional,

$$J(u, \lambda) = \int_\Omega [(\nabla u) \cdot (\nabla u) + u \cdot u] \, d\Omega - 2\lambda \int_{\partial\Omega} (u - h) \, d\partial\Omega - 2 \int_\Omega fu \, d\Omega ,$$

(4.36)

by using the Lagrangian multiplier method.

The physical meaning of the Lagrangian multiplier in this case is the boundary flux [23]; if we solve the elastostatic problem, the Lagrangian multiplier would be the boundary traction. In contrast to the finite element shape function, the interior particles, in a MLSRK interpolating field, have non-trivial contribution to the boundary flux, specifically, $\forall x \in \partial\Omega$,

$$\lambda^\rho(x) := \sum_{l=1}^{NP} \eta_l N_l(\rho, x, x_l) \quad x_l \in \overline{\Omega}$$

(4.37)

Thereby, we are basically dealing with a typical mixed problem here. It is well known that the mixed algorithm is not always stable unless additional restrictions are imposed on its weighting function space, or trial function space [23].

It has been discovered recently that the correction function can actually enforce the shape function to fulfill the essential boundary condition by carefully choosing the dilation parameters.

Technically speaking, at least in one dimensional problem, it is possible to construct a finite-dimensional trial function space, $\mathcal{U}_\rho^m$, such that

$$u^{\varrho} \in \mathscr{U}_{\varrho}^{m}(\Omega) \tag{4.38}$$

$$\mathscr{U}_{\varrho}^{m}(\Omega) := \mathrm{span}\{N_I^m | \mathrm{supp}\{N_I^m\} \cap \Omega \neq \emptyset; \; N_I(x_J) = \delta_{IJ}, \; \forall \, x_I \in \partial\Omega\} \tag{4.39}$$

In terms of weighting function, one can show that it is feasible to construct a finite-dimensional weighting function space,

$$\mathscr{W}_{\varrho,0}^m(\Omega) := \mathrm{span}\{N_I^m | \mathrm{supp}\{N_I^m\} \cap \Omega \neq \emptyset; \; N_I(x) = 0, \; \forall \, x \in \partial\Omega\} \tag{4.40}$$

where $\mathscr{W}_{\varrho,0}^m \subset \mathscr{U}_{\varrho}^m$. For one-dimensional problems, two pictorial examples of such finite-dimensional trial function basis are shown in Figs. 10 and 11. Some other examples of such finite-dimensional spaces may be constructed by proper choice of dilation parameter.

The family of shape functions in Fig. 10[7] are constructed based on cubic spline window function with linear generating polynomial, and the family of shape functions in Fig. 11 are constructed based on the fifth-order spline window function with quadratic generating polynomial. In both cases, the particle distributions are uniform. The group of shape functions in Fig. 10 have the support radius $r_{\varrho} = 2\,\Delta x$; for those shape functions in Fig. 11, the support radius is $r_{\varrho} = 3\,\Delta x$. For the multiple dimensional problem, proper trial function basis or weighting function basis can be devised with care. The significance of this finding is not just a happy ending, it shows that the moving least-square kernel Galerkin method still preserves some essential advantages that the traditional finite element method possesses.

By employing the classic variational technique, the moving least-square reproducing kernel Galerkin formulation for Dirichlet problem (4.33) and (4.34) can be set as follows:

$$\mathrm{MLSRKG(III):} \begin{cases} \text{Find } u^{\varrho} \in \mathscr{U}_{\varrho}^m(\Omega) \text{ such that } \forall \, w^{\varrho} \in \mathscr{W}_{\varrho,0}^m(\Omega) \\ \displaystyle\int_{\Omega} (\nabla u^{\varrho} \cdot \nabla w^{\varrho} + u^{\varrho} \cdot w^{\varrho})\, \mathrm{d}\Omega = \int_{\Omega} f w^{\varrho}\, \mathrm{d}\Omega \end{cases} \tag{4.41}$$

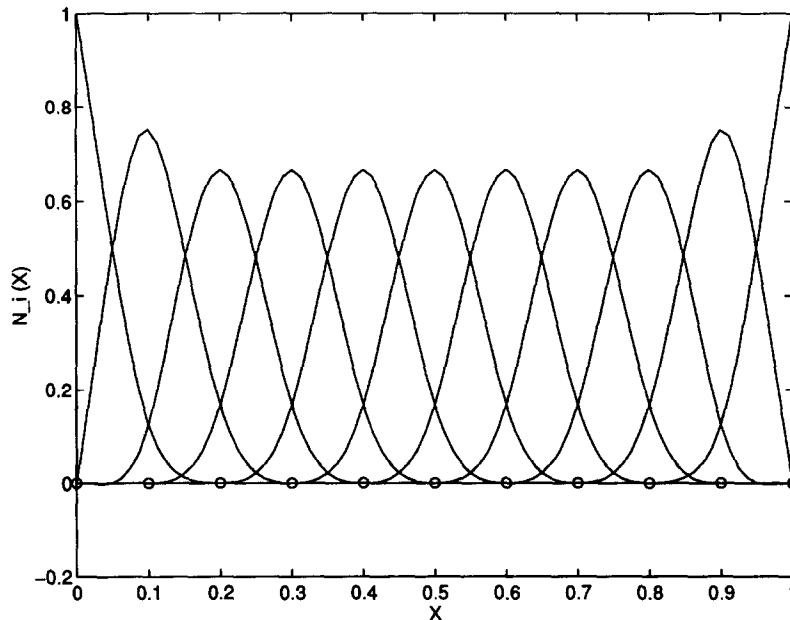Following the same procedure in the last section, one can show that the following statement holds.

Fig. 10. A finite-dimensional trial function basis that belongs to the space $\mathscr{U}_{\varrho}^1[0, 1]$.

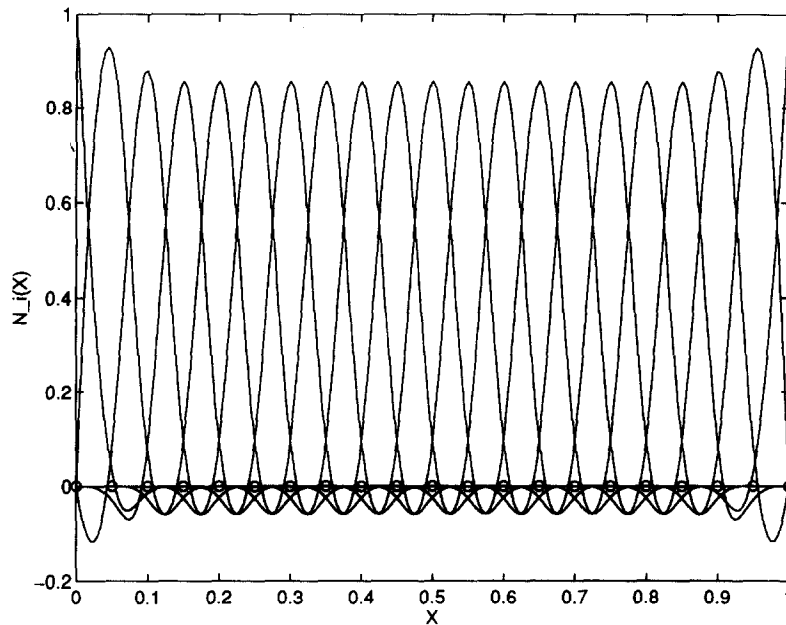[7] This example is discovered by Jeffrey Gosz, a graduate student at Northwestern University.

Fig. 11. A finite-dimensional trial function basis that belongs to space $\mathscr{U}_\varrho^2[0, 1]$.

THEOREM 4.3. *Let $u \in C^{m+1}(\Omega)$. If u is the solution of Dirichlet problem (4.33)–(4.34), and u^ϱ is the solution of weak formulation (4.41), then $\exists\, C$ such that*

$$\left\| u - u^\varrho \right\|_{H^1(\Omega)} C_{d1} \varrho^m \left\| u \right\|_{H^{m+1}(\Omega)}$$

(4.42)

and

$$\left\| u - u^\varrho \right\|_{L^2(\Omega)} \leq C_{d0} \varrho^{m+1} \left\| u \right\|_{H^{m+1}(\Omega)}$$

(4.43)

where the constants C_{d0}, C_{d1} do not depend on dilation parameter ϱ.

The above nice finite-dimensional trial function subspace can be constructed for nonuniform particle distribution. On the other hand, it has been found recently that one can always construct the following finite-dimensional shape function basis for highly irregular particle distribution with special care,

$$\mathscr{V}_\varrho^{o,m} := \mathrm{span}\{N_I^m | \mathrm{supp}\{N_I^m \cap \Omega \neq \emptyset, \, N_I(x) = 0 \, \forall\, x_I \in \Omega \text{ and } x \in \partial\Omega\}$$

(4.44)

This is good enough to handle the general Dirichlet problem. Following Hughes [24], we can define a new function,

$$h^\varrho(x) := \begin{cases} h(x), & x \in \partial\Omega \\ \text{continuous function} & x \in \Omega \end{cases}$$

(4.45)

such that the trial solution can be chosen as

$$u^\varrho = v^\varrho + h^\varrho$$

(4.46)

where $v^\varrho \in \mathscr{V}_\varrho^{o,m}(\Omega)$.

Therefore, for general multi-dimensional problem, we can have the following Bubnov–Galerkin formulation

$$\text{MLSRKG(IV):} \begin{cases} \text{Given } f(x), \text{ find } u^{\varrho} = v^{\varrho} + h^{\varrho}, \\ \text{where } v^{\varrho} \in \mathscr{V}_{\varrho}^{\hat{m}}(\Omega), \text{ such that } \forall w^{\varrho} \in \mathscr{V}_{\varrho}^{\hat{m}}(\Omega) \\ \int_{\Omega} (\nabla v^{\varrho} \cdot \nabla w^{\varrho} + v^{\varrho} \cdot w^{\varrho}) \, d\Omega = \int_{\Omega} (f w^{\varrho} - \nabla h^{\varrho} - h^{\varrho} \cdot \nabla w^{\varrho} \cdot w^{\varrho}) \, d\Omega \end{cases} \tag{4.47}$$

The proof of convergence of above algorithm is almost the same as the proofs in the last section.

4.2. Numerical example

There are many problems which have been tested and solved by using the moving least-square kernel Galerkin method. In order to compare with other numerical algorithms, here, a special benchmark problem is tested. This problem was originally proposed by Rachford and Wheeler [25] to test the convergence property of the H^{-1}-Galerkin method, and was used again by Babuška et al. [26] to test the mixed-hybrid finite element method. It is a two point boundary value problem,

$$\begin{cases} -u_{,xx} + u = f(x) & x \in (0, 1) \\ u(0) = u(1) = 0 \end{cases} \tag{4.48}$$

where

$$f(x) = \frac{2\alpha(1 + \alpha^2(1 - \bar{x})(x - \bar{x}))}{(1 + \alpha^2(x - \bar{x})^2)^2}$$

$$+ (1 - x)(\arctan(\alpha(x - \bar{x})) + \arctan(\alpha\bar{x})) \tag{4.49}$$

The exact solution of Eq. (4.48) is

$$u(x) = (1 - x)(\arctan(\alpha(x - \bar{x})) + \arctan(\alpha\bar{x})) \tag{4.50}$$

According to the designed feature, the solution (4.50) changes its roughness as the parameter α varies. When α is relatively small, the solution (4.50) is smooth; as α becomes large, there will be a sharp knee arising close to the location $x = \bar{x}$. Thus, it provides quite a challenging test for numerical computations.

Following the choice of Babuška et al. [26], the two representative parameter groups chosen are as follows:

$$\text{the smooth solution: } \begin{cases} \alpha = 5.0 \\ \bar{x} = 0.2 \end{cases} \tag{4.51}$$

$$\text{the rough solution: } \begin{cases} \alpha = 50.0 \\ \bar{x} = 0.40 \end{cases} \tag{4.52}$$

In Fig. 12, the exact solutions—both smooth and rough are plotted in comparison with numerical results. One can see that the numerical solutions agree with exact solutions fairly well in both cases—the smooth solution as well as the rough solution. In Fig. 13, the comparison between exact solution and numerical solution is made for the first-order derivatives. As mentioned above, two types of shape functions have been used in numerical computation: the shape functions generated by cubic spline window function and those generated by fifth-order spline window function, i.e. the shape function families shown in both Figs. 10 and 11. The computation is carried out for four different particle distributions: 11 particles, 21 particles, 41 particles and 81 particles. The results shown in Figs. 12 and 13 is obtained by using the first group of shape function with 41 particles uniformly distributed in the domain.

Based on numerical results, the convergence rate of the algorithm is also shown with respect to different norms: L_2 norm, H_1 norm, and $|\cdot|_{\max}$ norm. For the shape functions generated by cubic spline window function, the corresponding convergence results are plotted in both Figs. 14 and 15, and the convergence results for the shape function based on fifth order spline window function are displayed in Figs. 16 and 17.

As mentioned above, the shape function family in Fig. 10 is generated by linear polynomials, i.e. $m = 1$, and the shape function family in Fig. 11 is generated by quadratic polynomials, i.e. $m = 2$. Based on Theorem 4.42, the convergence rates with respect to L_2 norm are 2 and 3, respectively. The numerical results in Figs. 14–17
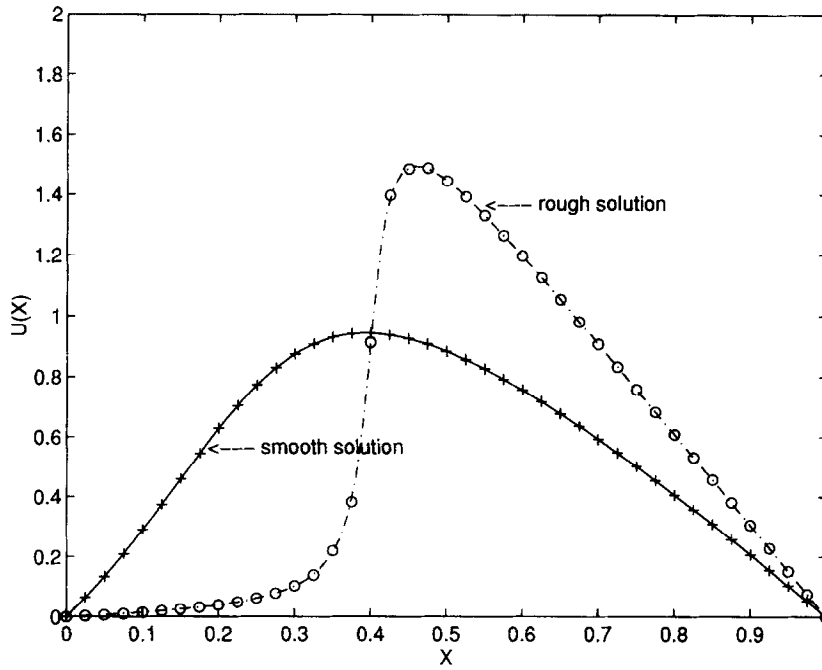
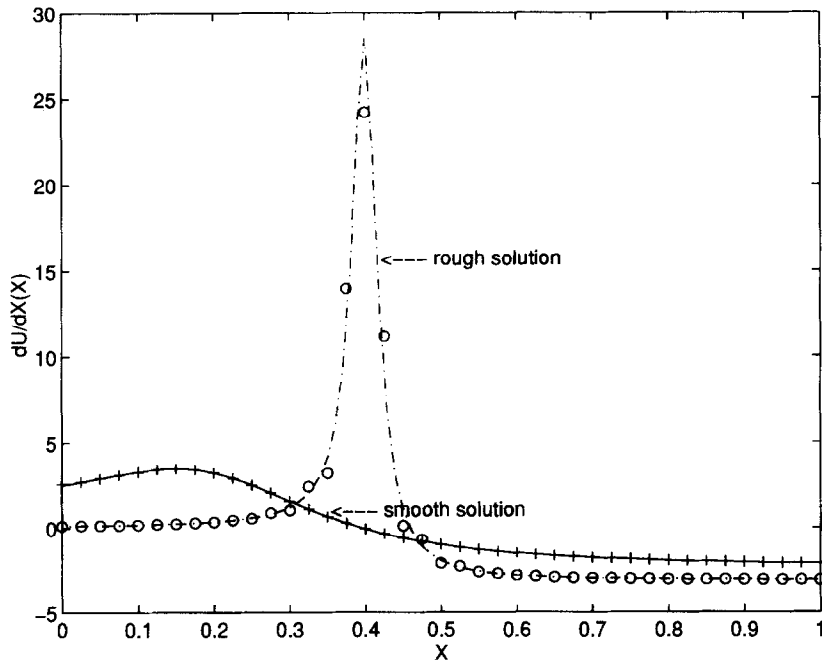Fig. 12. The exact solutions and numerical solutions of the benchmark problem.

Fig. 13. The derivatives of exact solutions and numerical solutions.

show that the numerical computation is doing far better than the estimate. Nevertheless, when the particle density increases, the theoretical bound will become evident. On the other hand, one may observe that there seems to be a tendency that H_1 error norm converges almost as fast as the L_2 error norm.

One may also notice an interesting fact that both L_2 norm and H_1 norm have faster convergence rate than that
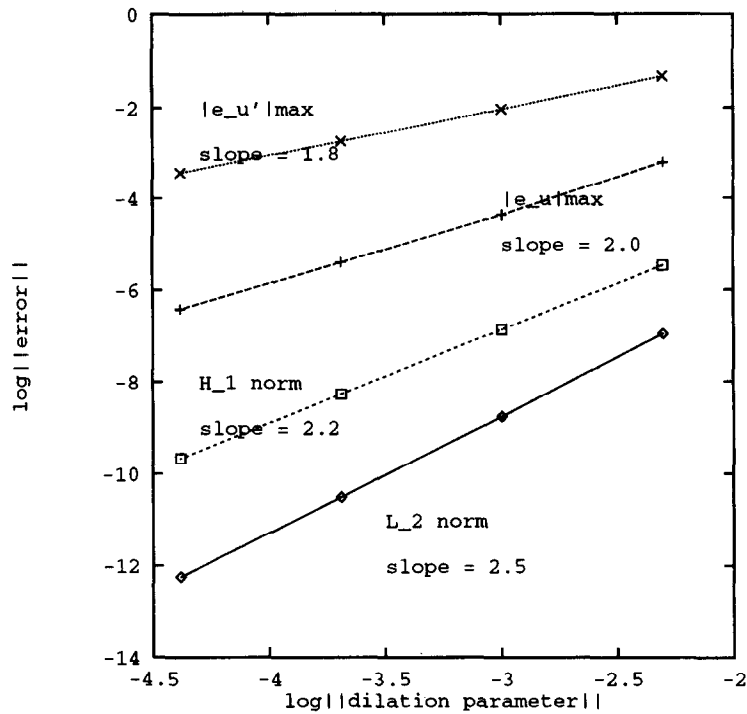
Fig. 14. The convergence rate for the smooth problem, $\alpha = 5.0$, $\bar{x} = 0.2$, for shape functions with basis of linear polynomial.
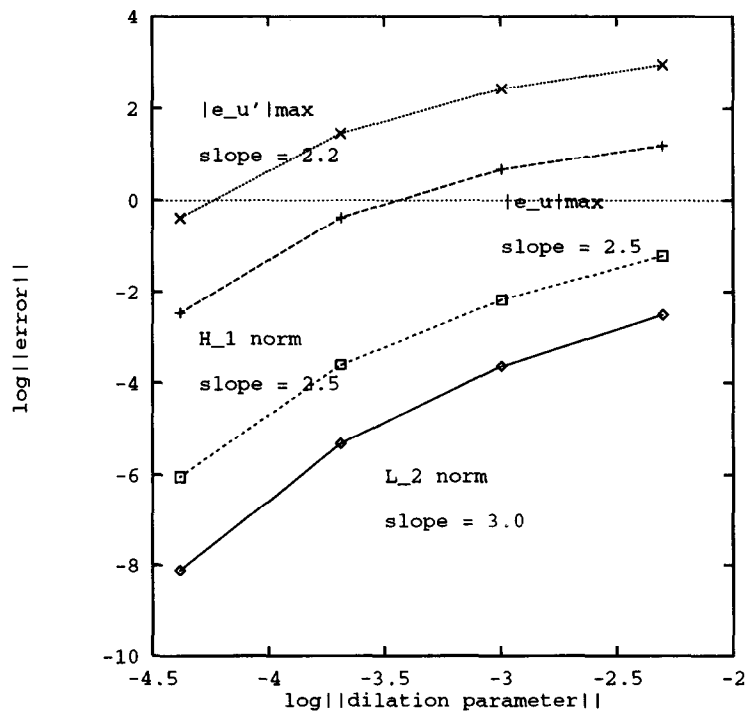

Fig. 15. The convergence rate for the rough problem, $\alpha = 5.0$, $\bar{x} = 0.2$, for shape functions with basis of linear polynomial.

of the maximum norm, which is totally in contrast with the conventional finite element method; the regular finite element method has an opposite tendency that maximum norm always have a faster convergence rate than H norms.
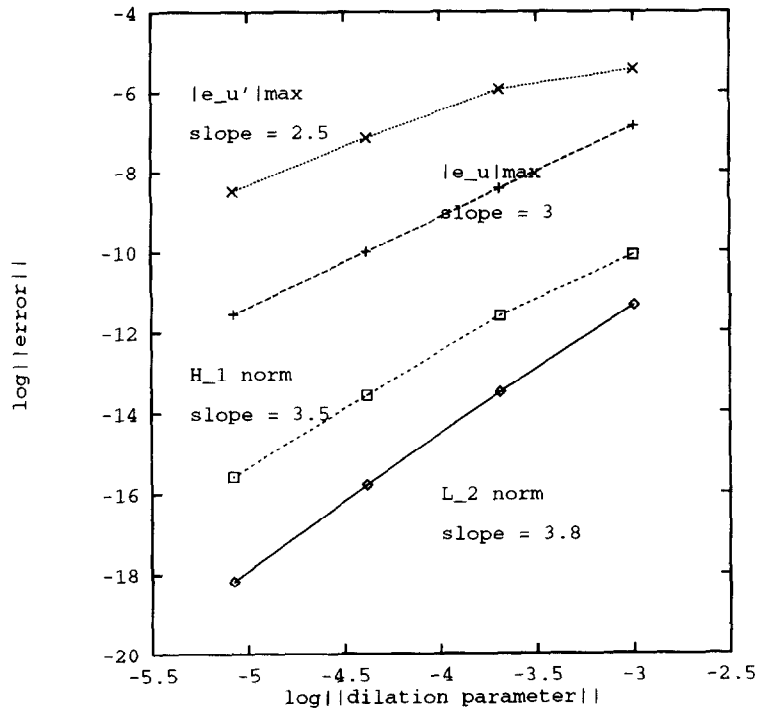
Fig. 16. The convergence rate for the smooth problem. $\alpha = 50.0$, $\bar{x} = 0.4$, while using shape functions with the base of quadratic polynomial.
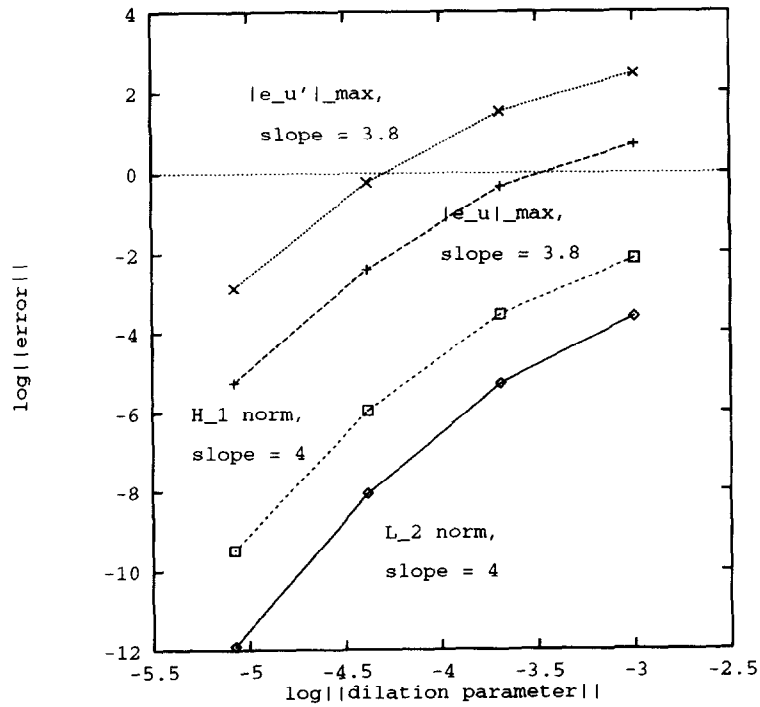

Fig. 17. The convergence rate for the rough problem, $\alpha = 50.0$, $\bar{x} = 0.4$, while using shape functions with base of quadratic polynomial.

5. Closure

After Nayroles et al. [3], Belytschko et al. [4] and Liu et al. [6] applied and modified the moving least-square approximation, a class of meshless methods have emerged with a completely different outlook. This is particularly evident after the connection between the reproducing kernel particle method and spectral analysis was made [7]. The use of moments, dilation parameter, and spline window function, . . . , etc. show its affinity to spectral method and, in particular, the wavelet method.

In this paper, a formal documentation of the MLSRK method has been presented. The analysis has shown that the numerical solutions obtained by using this method will converge as the dilation parameter ϱ approaches to zero. This is also confirmed by various numerical experiments conducted in Northwestern University in the last three years.

There are several points that we would like to reiterate. First, the MLSRK method builds a bridge between the traditional interpolation method and the spectral method; that is very desirable for the next generation of finite element methods. Second, by using moving least-square interpolant, higher-order conforming shape functions, such as $u^e \in C^2(\Omega)$ or higher, can be easily constructed, which is difficult to realize by using the regular finite element method. This could have a direct impact on computational structural mechanics, such as numerical simulations of plates and shells. In addition, as we hope, the complete avoidance of 'variational crimes' may be possible. Third, as a meshless or semi-meshless method, a tremendous work reduction has been achieved in mesh data preparation. This is a significant advance for numerical computations that involve complex geometrical objects and mesh refinement and adaptivity procedures. However, the tradeoff is that there is also an increase of computer time to generate and evaluate these shape functions. How to balance this tradeoff still remains an issue of future research, which is crucial for this method to gain its popularity.

From our perspective, up to this point, the MLSRK method is not yet a mature numerical tool. There is still much room for improvement so that the method can fit various computational tasks. Nevertheless, whatever the modification might be, the key issue is to increase the computation efficiency. If this problem can be properly handled without losing its original technical merits, there is no doubt that the meshless methods could become powerful numerical tools.

Acknowledgments

References

[1] R.A. Gingold and J.J. Monaghan, Kernel estimates as a basis for general particle methods in hydrodynamics, J. Comput. Phys. 46 (1982) 429–453.

[2] J.J. Monaghan, Why particle methods work, SIAM J. Scient. Stat. Comput. 3 (1982) 422–433.

[3] B. Nayroles, G. Touzot and P. Villon, Generalizing the finite element method: Diffuse approximation and diffuse elements, Comput. Mech. 10 (1992) 307–318.

[4] T. Belytschko, Y.Y. Lu and L. Gu, Element free Galerkin methods, Int. J. Numer. Methods Engrg. 37 (1994) 229–256.

[5] T. Belytschko, L. Gu and Y.Y. Lu, Fracture and crack growth by element free Galerkin methods, Modelling Simul. Mater. Sci. Engrg. 2 (1994b) 519–534.

[6] W.K. Liu, S. Jun and Y. Zhang, Reproducing kernel particle methods, Int. J. Numer. Methods Fluids 20 (1995) 1081–1106.

[7] W.K. Liu and Y. Chen, Wavelet and multiple scale reproducing kernel methods, Int. J. Numer. Methods Fluids 21 (1995) 901–933.

[8] W.K. Liu, S. Jun, S. Li, J. Adee and T. Belytschko, Reproducing kernel particle methods for structural dynamics, Int. J. Numer. Methods Engrg. 38 (1995) 1655–1679.

[9] W.K. Liu, An introduction to wavelet reproducing kernel particle methods, USACM Bulletin 8(1) (1995) 3–16.

[10] P. Lancaster and K. Salkauskas, Surfaces generated by moving least square methods, Math. Comput. 37(155) (1980) 141–158.
[11] S. Li and W.K. Liu, Moving least square reproducing kernel method. Part II: Fourier analysis, Comput. Methods Appl. Mech. Engrg. 139 (1996) 159–194.
[12] R.O. Wells Jr. and X. Zhou, Wavelet interpolation and approximate solutions of elliptic partial differential equations, (1993) preprint.
[13] C.A. Duarte and J.T. Oden, Hp Clouds—a meshless method to solve boundary-value problems, (1995) preprint.
[14] W. Rudin, Principles of Mathematical Analysis, 3rd edition (McGraw-Hill, Inc., 1976).
[15] D.S. Mitrinović, Analytic Inequalities (Springer-Verlag, Berlin, Heidelberg, New York, 1970).
[16] P.J. Davis, Interpolation and Approximation (Dover Publications, Inc., New York, 1975).
[17] C.K. Chui, An Introduction to Wavelet (Academic Press, New York, 1992).
[18] I. Babuška and J.M. Melenk, The partition of unity finite element method, Technical Note BN-1185, Institute for Physical Science and Technology, University of Maryland, 1995.
[19] C. de Boor and G.J. Fix, Spline approximation by quasi-interpolants, J. Approx. Theory 8 (1975) 19–45.
[20] R.A. Horn and C.R. Johnson, Matrix Analysis (Cambridge University Press, 1985).
[21] G.H. Golub and C.F. Van Loan, Matrix Computations, 2nd edition (The Johns Hopkins University Press, Baltimore, MD, 1989).
[22] P.G. Ciarlet, The Finite Element Method For Elliptic Problems (North-Holland Publishing Company, Amsterdam, New York, 1978).
[23] I. Babuška, The finite element method with Lagrangian multipliers, Numer. Math. (20)(1973) 179–192.
[24] T.J.R. Hughes, The Finite Element Method: Linear Static and Dynamic Finite Element Analysis (Prentice Hall, Englewood Cliffs, NJ, 1987).
[25] H.H. Rachford Jr. and M.F. Wheeler, An H^{-1}-Galerkin procedure for the two-point boundary value, in: C. de Boor, ed., Mathematical Aspects of Finite Elements in Partial Differential Equations (Academic Press, New York, 1974).
[26] I. Babuška, J.T. Oden and J.K. Lee, Mixed-hybrid finite element approximations of second-order elliptic boundary-value problems, Comput. Methods Appl. Mech. Engrg. 11 (1977) 175–206.
[27] I. Daubechies, Wavelets, CBMS-NSF Series in Applied Mathematics (SIAM Publication, Philadelphia, 1992).