Influence of Pareto optimality on the Maximum Entropy Methods

Sreehari Peddavarapu[1, a)], Gujjalapudi Venkata Sai Sunil[1, b)] and Raghuraman S[1]

[1]*School of Mechanical Engineering, SASTRA University, Thanjavur, Tamil Nadu-613401*
[a)]Corresponding author: sreehari@mech.sastra.edu
[b)]gvssunil1997@sastra.ac.in

Abstract. Galerkin meshfree schemes are emerging as a viable substitute to finite element method to solve partial differential equations for the large deformations as well as crack propagation problems. However, the introduction of Shanon-Jayne's entropy principle in to the scattered data approximation has deviated from the trend of defining the approximation functions, resulting in maximum entropy approximants. Further in addition to this, an objective functional which controls the degree of locality resulted in Local maximum entropy approximants. These are based on information-theoretical Pareto optimality between entropy and degree of locality that are defining the basis functions to the scattered nodes. The degree of locality in turn relies on the choice of locality parameter and *prior* (weight) function. The proper choices of both plays vital role in attain the desired accuracy. Present work is focused on the choice of locality parameter which defines the degree of locality and priors: Gaussian, Cubic spline and quartic spline functions on the behavior of local maximum entropy approximants.

Key words: Local maximum entropy approximation, meshfree scheme, prior, basis function.

1. INTRODUCTION

Recently Maximum Entropy meshfree methods have come to forefront in the area of meshfree methods. They does not fall short of the properties of well-established Finite Element Method, Where other deeply researched Moving least square approximants are having laborious process of applying Dirichlet boundary conditions due to their non-interpolating nature on the boundary [1,7]. Local maximum approximations are having weak kronecker delta property which is sufficient to apply their boundary conditions [1-6].

It was Sukumar.N [1], who applied the Shannon-Jaynes entropy functional as subjective measure of conditional probability to deduct the shape functions for polygons. The result is maximum entropy approximants (Max-Ent) to the polygonal elements. Successively local maximum approximations (LME) are coined by Arroyo and Ortiz [2], by introducing local support width functions to convex approximations. These are similar to moving least square approximants (MLS) where weight functions (priors) are used to localise the shape functions [7] but they (LME) are interpolating in nature on the boundary. However, MLS does not have this weak kronecker delta property implying that boundary conditions could not be applied straight forward. Later Sukumar and Wright [3] unified the construction of global (polygonal Max-Ent shape functions) and locally supported (LME shape functions) convex approximation schemes by using Relative maximum entropy of Shannon-Jayne's principle. The proper choice of locality parameter is depending on problem and not easy in general. So to optimise the locality parameter Rosolen et al [8], adapted the variational approach to find the optimum support size with the LME approximants. Max-Ent approximants were successfully blended with other convex approximants such as B-splines or NURBS basis functions by Rosolen A, Arroyo [9] in their recent work to attain the high geometric fidelity on the convex boundary of the domain. The performance of the method can be improved in those problems with this approach and verified that the Kronecker-Delta property can be obtained also on non-convex domains.

Max-Ent methods have been successfully applied to a range of problems, from thin shell analysis to fracture mechanics, including non-linear structural analysis, phase-field models applied to bio membranes, biasing of molecular simulations, and convection-diffusion problems [10, 12, and 13].

However, selection of the appropriate prior or weight function for the construction of LME approximation functions was not addressed properly in the literature. In the current work Gaussian, Cubic spline and quartic spline functions are considered as Priors to show their effect on the continuity and smoothness for the LME approximants. Further, the effect of locality parameter on each prior is also addressed on one dimensional as well as two dimensional space.

2. APPROXIMATION SCHEMES

2.1. Convex Approximation Scheme

Consider a convex hull set of distinct nodes $X = \{x_a, a = 1, \dots, N\} \subset \mathbb{R}^d$, to be referred to as the node set. Let $u : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ be a function whose values $\{u_a, a = 1, \dots, N\}$ are known on a node set X. Assume that the affine hull of the node set is $\mathbb{R}^d$ to construct the approximations of the form

$$u^h(x) = \sum_{a=1}^{N} p_a(x) u_a \quad (1)$$

Where the functions $p_a: \Omega \to \mathbb{R}$ will be referred to as shape functions and they are strictly nonnegative. A particular choice of shape functions defines an approximation scheme. The shape functions are required to satisfy the zeroth and first-order consistency conditions:

$$\sum_{a=1}^{N} p_a(x) = 1, \forall x \in \Omega, \quad \sum_{a=1}^{N} p_a(x) x_a = x, \forall x \in \Omega. \quad (2)$$

These properties of the approximation scheme guarantee that the affine functions are exactly reproduced. In the finite element method, the shape functions defined on the reference elements, comprise grid of nodes, satisfy these key properties, lead to the shape functions which are interpolating in nature. This kronecker delta property of shape functions facilitates the exact imposition of linear Dirichlet boundary conditions. For instance, to solve the second-order partial differential equations (PDEs) such as the equations of linear elasticity of the Poisson equation, approximants that possess constant and linear completeness are sufficient for convergence [1]. However, when $N > d + 1$, the shape functions are not uniquely determined by the consistency conditions [2].

2.2. Maximum Entropy Approximation Scheme

Consider a convex hull set of distinct nodes $X = \{x_a, a = 1, \dots, N\} \subset \mathbb{R}^d$, to be referred to as the node set. Let $\phi : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ be a function whose values $\{\phi_a, a = 1, \dots, N\}$ are the probabilities associated with node set X. Canonical Entropy measure of the uncertainty associated with the probabilities ϕ_a :

$$H(\phi) = -\sum_{a=1}^{N} \phi_a \log \phi_a \quad (3)$$

The probabilities ϕ_a associated with above entropy are realized as approximation functions by invoking Jayne's principle of maximum entropy [1, 2]. The entropy measures the lack of information about the data set i.e., the set of shape functions which results in the least biased possible choice of convex scheme, devoid of artifacts or hidden assumptions. So, the approximation of a function from scattered data is viewed as the probability of influence of other nodes at node x and thus becomes a problem of statistical inference, and is mathematically formulated through the convex program:

$$(ME) \qquad Maximize \quad H(\phi) = -\sum_{a=1}^{N} \phi_a \log \phi_a \quad (4)$$

Subject to $\phi_a \geq 0, a = 1, \dots, N,$

$$\sum_{a=1}^{N}\phi_a = 1, \quad \sum_{a=1}^{N}\phi_a x_a = x, \quad \sum_{a=1}^{N}\phi_a y_a = y, \tag{5}$$

The solutions of (ME*)*, denoted by $\phi_a(\boldsymbol{x})$ referred to as Max-Ent basis functions.

Sukumar [1] implemented the Lagrange multipliers method to solve above constrained optimization problem and the resulting shape functions:

$$\phi_a = \frac{Z_a}{\sum_{i=1}^{n} Z_i},\ a=1,\ldots,N, \qquad Z(\lambda_1,\lambda_2) = \sum_{a=1}^{n} e^{-\lambda_1 x_a - \lambda_2 y_a} \qquad \phi_a = \frac{e^{-\lambda_1 x_a - \lambda_2 y_a}}{\sum_{i=1}^{n} e^{-\lambda_1 x_i - \lambda_2 y_i}},\ a=1,\ldots.N, \tag{6}$$

Where λ_1 and λ_2 are Lagrange multipliers, maximizes the entropy. In this dual problem, the point x is fixed, and the Lagrange multipliers are varied to attain numerical stability while solving for [3]. Newton's method is implemented along with steepest descent algorithm to find the λ_1 and λ_2. So the nodal coordinates are shifted to the fixed point x results in $\tilde{x}_a = x_a - x$. Therefore the consistency conditions are

$$\sum_{a=1}^{N}\phi_a(x) = 1, \forall x \in \Omega, \quad \sum_{a=1}^{N}\phi_a(x)\tilde{x}_a = x, \forall x \in \Omega. \qquad \sum_{a=1}^{N}\phi_a(y)\tilde{y}_a = y, \forall y \in \Omega. \tag{7}$$

Results in shape functions with shifted coordinates:

$$Z(\lambda_1,\lambda_2) = \sum_{a=1}^{n} e^{-\lambda_1 \tilde{x}_a - \lambda_2 \tilde{y}_a} \qquad \phi_a = \frac{e^{-\lambda_1 \tilde{x}_a - \lambda_2 \tilde{y}_a}}{\sum_{i=1}^{n} e^{-\lambda_1 \tilde{x}_i - \lambda_2 \tilde{y}_i}},\ a=1,\ldots,N, \tag{8}$$

Here Z_a can be viewed as a weight function [3] and discussed further in section 2.3.

2.3.Local Maximum Entropy Approximation Scheme

Arroyo and Ortiz [2], proposed LME basis functions, adding special correlation (degree of locality) to Max-Ent convex approximations to define information-theoretical optimal approximants, least biased by locality, i.e. the property that the approximation at a given point should depend on nodal values of nearby node set. In general, it is not possible to find the approximation scheme which is maximizing both locality and entropy simultaneously, i.e. locality and unbiased estimation are competing objective functions.

To harmonize this Pareto set, is to seek the Pareto optimality between two competing objectives [2, 3], subject to the consistency constraints. Optimization program with these requirements, to select the approximants:

(LME*)* For a fixed x minimize $f_\beta(x,\phi) = \beta U(x,\phi) - H(\phi)$ subject to $\phi_a \geq 0$, $a=1,\ldots,N$, (9)

$$\sum_{a=1}^{N}\phi_a = 1, \quad \sum_{a=1}^{N}\phi_a\tilde{x}_a = x, \quad \sum_{a=1}^{N}\phi_a\tilde{y}_a = y, \tag{10}$$

Where

$$U(x,\phi) = \sum_{a=1}^{N}\phi_a |x - x_a|^2, \quad H(\phi) = -\sum_{a=1}^{N}\phi_a \log\phi_a, \quad \beta = (0,+\infty) \tag{11}$$

The non-negative parameter $\beta = \gamma/h^2$ defines the decay or locality of the approximants. γ is a dimensionless parameter, characterizes the degree of locality along with nodal spacing h. These can be made to vary with each nodal point to accommodate unstructured nodal set to discretize the problem domain. Larger value of β gives the global support and it was found that when the first order consistency condition is withdrawn resembling Shepard approximations with Gaussian weight function [2] which implies that the Lagrange multipliers correcting the local maximum entropy approximants to satisfy the first order consistency condition.

As the dimensionless parameter γ increases, the basis function becomes sharper and the Delaunay approximants recover for $\gamma \geq 4$ in practice [4], which will be detailed further in the section. 3. $U(x,\phi)$, a weight function (prior) recovers the degree of locality. Priors also varied (Non-uniform prior) for each node i to give suitable weight to x_i than to other nodal locations. Approximation scheme optimizes the Pareto set:

$$Z(\lambda_1,\lambda_2)=\sum_{a=1}^{n} e^{-\beta|x-x_a|^2-\lambda_1\tilde{x}_a-\lambda_2\tilde{y}_a} \tag{12}$$

$$\phi_a=\frac{e^{-\beta|x-x_a|^2-\lambda_1\tilde{x}_a-\lambda_2\tilde{y}_a}}{\sum_{i=1}^{n} e^{-\beta|x-x_a|^2-\lambda_1\tilde{x}_i-\lambda_2\tilde{y}_i}}\text{ , } a=1,\ldots., N, \tag{13}$$

On generalization of the entropy functional, Sukumar and Wright [3] unified both convex maximum entropy approximation schemes; Max-Ent [1] and LME [2], by Relative maximum entropy of Shannon-Jayne's principle. The modified convex optimization problem:

$$Maximize \quad H(\phi)=-\sum_{a=1}^{N}\phi_a \log(\frac{\phi_a}{w_a}) \tag{14}$$

Subject to linear reproducing conditions $\phi_a \geq 0$, *a=1,…., N,*

$$\sum_{a=1}^{N}\phi_a=1,\ \sum_{a=1}^{N}\phi_a\tilde{x}_a=x,\ \sum_{a=1}^{N}\phi_a\tilde{y}_a=y, \tag{15}$$

Where w_a weight function associated with node *a*. It is evident that any weight function that is at least C^0 continuous and compactly supported may be used for the second-order partial differential equations (refer section 2.1). Typical weight functions are Gaussian radial basis functions [2]

$$w_a(x)=e^{-\beta_a\|x_a-x\|^2} \tag{16}$$

Quartic polynomials [11]

$$w_a(q)=\begin{cases}1-6q^2+8q^3-3q^4, & 0\leq q\leq 1\\ 0, & q>1\end{cases} \tag{17}$$

Cubic splines [3]

$$w_a(q)=\begin{cases}\frac{2}{3}-4q^2+4q^3, & q\leq\frac{1}{2}\\ \frac{4}{3}-4q+4q^2-\frac{4}{3}q^3, & \frac{1}{2}<q\leq 1\\ 0, & q>1\end{cases} \tag{18}$$

Where $q=\frac{\|x-x_a\|}{r_a}$,

$\|x-x_a\|$ is the L^2 norm of its argument, $r_a = \gamma\, h$ is the radius of support for each node. Alternative compactly supported weight functions, R-functions [3] are also suitable.

We refer [2, 12] for the expression to compute $\nabla\phi_a(x)$ of the LME approximants:

$$\nabla\phi_a=\phi_a\left\{\tilde{x}_a.\left[(H)^{-1}-(H)^{-1}.A\right]+\frac{\nabla w_a}{w_a}-\sum_{b=1}^{n}\phi_b\frac{\nabla w_b}{w_b}\right\} \tag{19}$$

Where

$$A=\sum_{b=1}^{n}\phi_b\tilde{x}_b\otimes\frac{\nabla w_b}{w_b} \tag{20}$$

And *H* is the hessian matrix defined by

$$H=\nabla_\lambda\nabla_\lambda \ln Z=\sum_{b=1}^{n}\phi_b\tilde{x}_b\otimes\tilde{x}_b \tag{21}$$

3. LOCAL MAXIMIZATION

This section is devoted to the discussion on the parameters, influencing the Pareto optimality between entropy and degree of locality. From the equations (9, 14), it is evident that optimality is controlled by locality parameter β and prior $U(x,\phi)$, defines the approximation for node x_a. Java applet developed by sukumar and wright [3] is used to generate the Max-Ent approximants and MATLAB code is developed to draw the LME approximants.

For a one dimension node set (Fig.1), smooth and seamless transition is observed from non-Unity to Unity (interpolant nature of finite elements). The transition is controlled by the non-dimensional nodal parameter $\gamma = \beta h^2$ which here take linearly varying values from 0 (left) to 6.2 (right).For the maximum value of $\gamma = 6$ shown in Fig.1, the shape function become sharp and taking unity value .

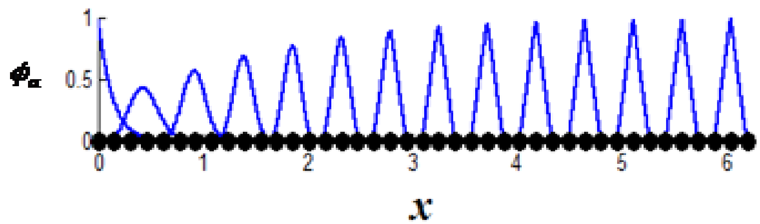

FIGURE 1. One dimensional example of seamless transition from meshless to finite elements achieved by tuning the value γ.

One dimensional Max-Ent basis functions are shown with uniform prior for regular grid and random grid in the Fig.2. It is observed that the interior Max-Ent basis functions are zero contribution on the boundary, contrary to MLS basis functions [7], whereas boundary Max-Ent basis functions satisfy the Kronecker-delta property. This weak Kronecker-delta property made them different from other meshfree methods [10, 12, and 13].

Next we consider the two dimensional Max-Ent basis functions shown in Fig.3. .For the value of $\gamma = 1.8$, the shape function is smooth, whereas for $\gamma = 3.8$, become polygonal and sharp. The approximation at the convex boundary face is independent of interior nodes (Weak Kronecker-delta property) is also evident on two dimensional node set. This made the imposition of dirichlet boundary conditions straight forward as standard finite element method [3].

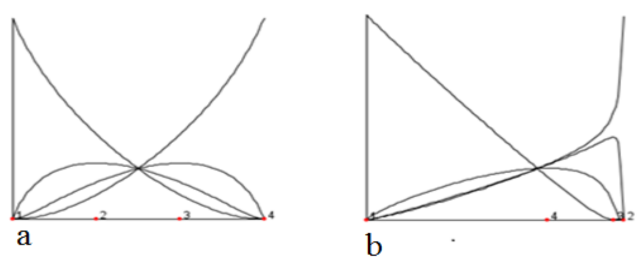

FIGURE 2. One dimensional Max-Ent basis functions with a uniform prior (a) with regular and (b) random grid.

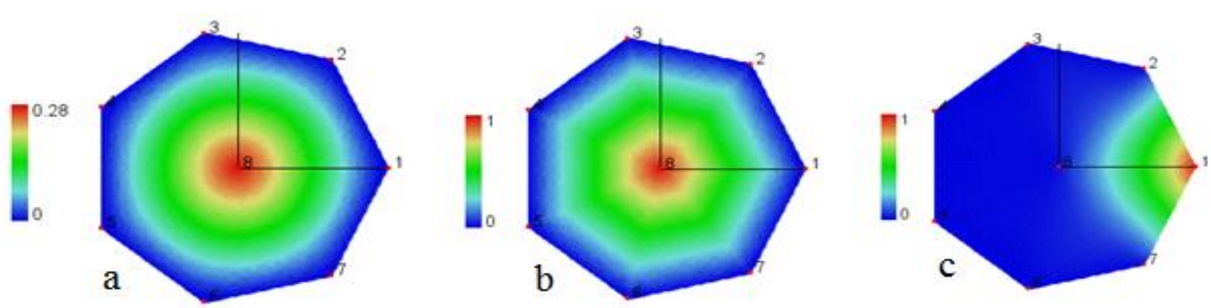

FIGURE 3. Two dimensional Max-Ent basis functions: (a) For interior node, *γ=1.8 (b)* For interior node, *γ=3.8 (c)* For boundary node

Behavior of the priors; Cubic spline (Fig.4) and Gaussian (Fig.5) on the one dimensional Max-Ent approximants is shown for locality parameter γ values of 4, 2 & 0.8. Weak Kronecker-delta property and the influence of γ is evident for regular grid. Similarly, Behavior of priors in the case of two dimensional Max-Ent basis functions are shown in Fig.6. Contribution of the prior to achieve smoothness is evident and Interpolating in nature is observed in all the approximants with different priors. Cubic spline C^1 continuity in Fig.6 (b) and Gaussian decay (bell shape) in Fig.6 (c) are observed. Figure 7 shows the effect of locality parameter on the bell shape localization for the γ values of 0.8, 2 & 2.8 and Weak Kronecker-delta property is also observed.

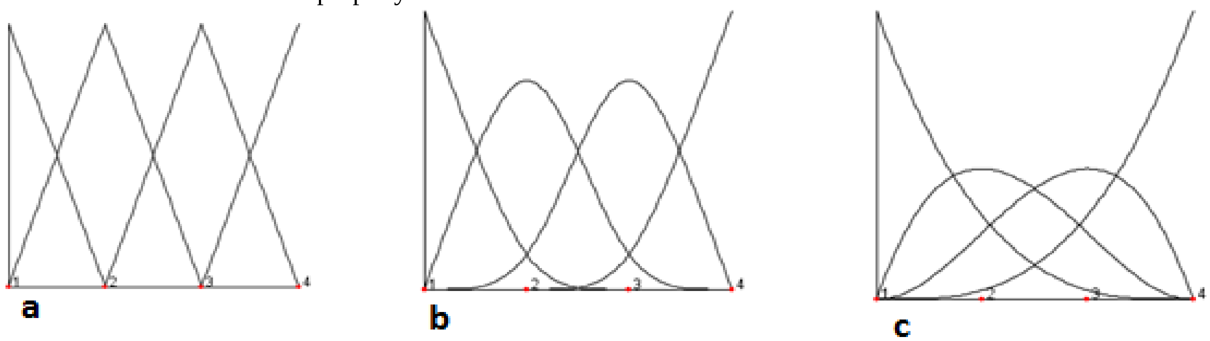

FIGURE 4. One dimensional Max-Ent basis with Compact cubic spline priors with (a) γ=4 (b) γ=2 and (c) γ=0.8

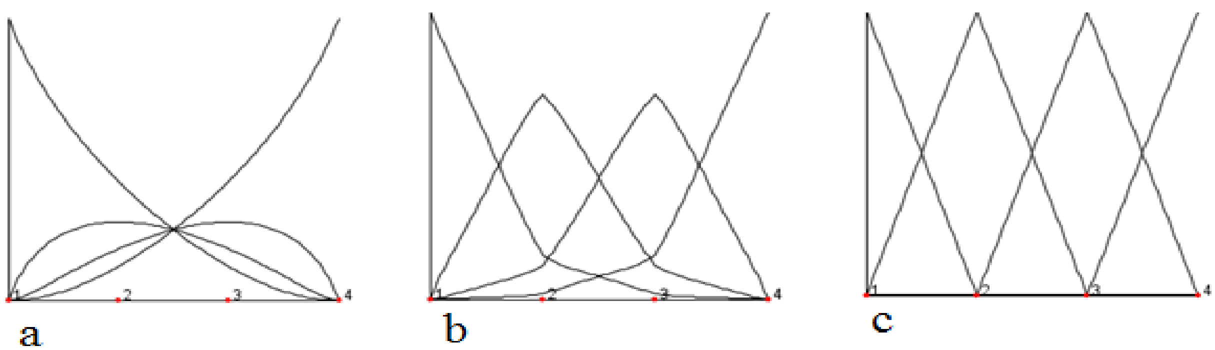

FIGURE 5. One dimensional Max-Ent basis functions with Gaussian prior for (a) γ=0.8 (b) γ=2 and (c) γ=4

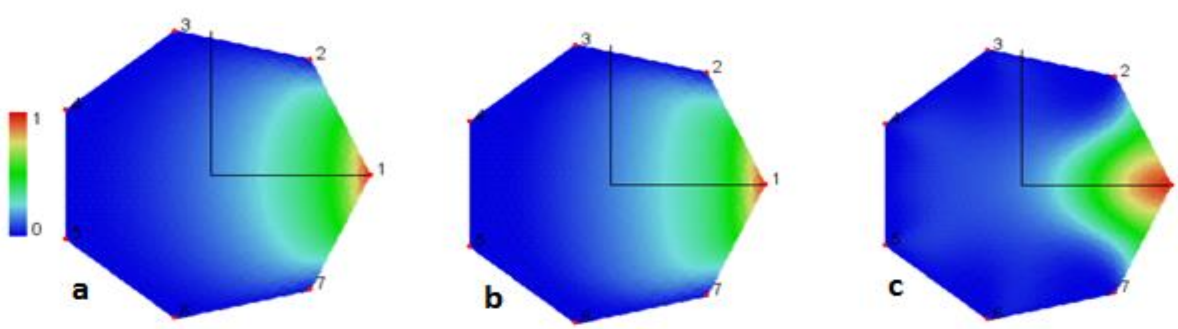

FIGURE 6. Two dimensional Max-Ent Basis functions for priors: (a) Constant (b) Cubic spline (c) Gaussian

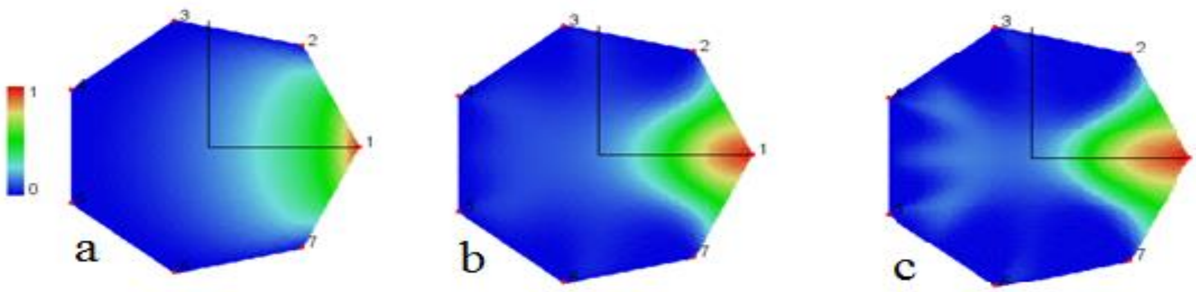

FIGURE 7. Two dimensional Max-Ent Basis functions for Gaussian priors: (a) γ =0.8(b) γ =2 and (c) γ = 2.8.

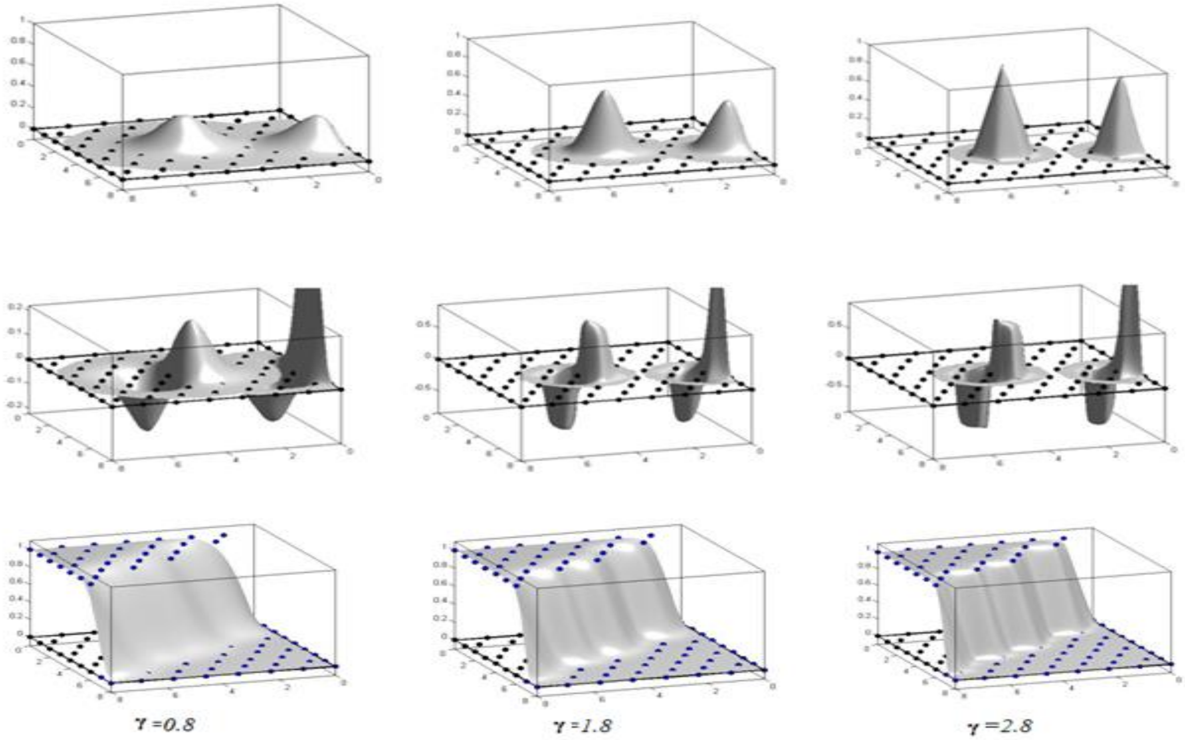

FIGURE 8. Two dimensional LME Basis functions for Gaussian prior ϕ (first row) and derivatives $\phi,_x$ (second row) and $\phi,_y$ (third row) for different values of γ

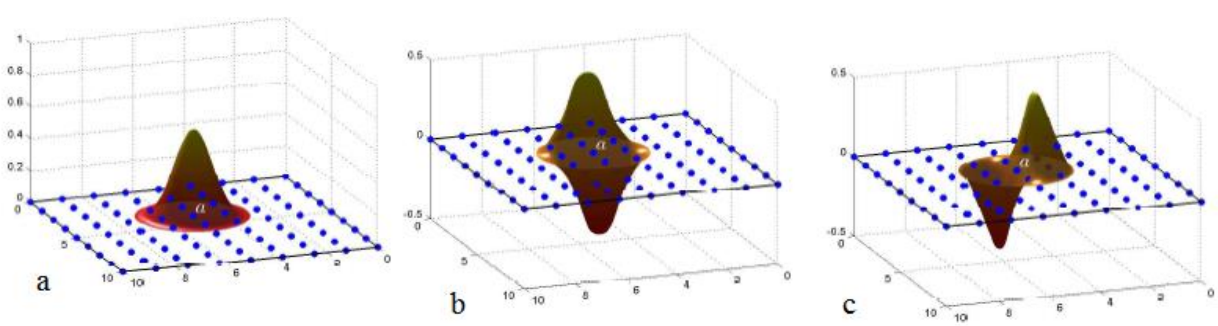

FIGURE 9. Two dimensional LME Basis functions for Quartic spline prior and derivatives for γ =1.6 [6]

For a Two dimensional nodal set, LME Basis functions and their derivatives for interior node and boundary node are shown in Fig.8 for Gaussian prior. As the γ value increases, the number of nodes in the neighborhood (bell shape of Gaussian decay) which influence the approximation is reduced and become polygonal. Derivatives of the basic functions are discontinuous for γ=2.8 [5]. Figure 9 shows LME basis functions for Quartic spline prior and derivatives for γ =1.6 and observed the C^2continuity and smoothness [6].

Typically, low values of γ lead to more meticulous results for problems with smooth solutions, but also result in significantly more expensive calculations due to the wider band [13]. The optimum value for γ is in between 0.8 to 1.8, provides tradeoff between accuracy and computational cost [2, 3, and 10].

It is observed that, as the β value increases, the Newton–Raphson method needs more iterations for convergence implies the rise in computational time for maximum entropy approximation functions. This slowdown in convergence, which is expected since the objective function tends to a faceted polyhedral convex function [2, 13].

4. SUMMARY

Similarity between global (Max-Ent) and local (LME) approximation schemes is presented. Pareto optimality between the Locality and Maximum Entropy is discussed for one dimensional as well as two dimensional nodal set with various values of γ and priors. Weak Knonecker delta property of the approximants is discussed for all the parameters. Derivatives of LME approximants are shown to ensure the continuity and smoothness due to the imposition of priors.

REFERENCES

1. N. Sukumar, *Int. J. Numer. Meth. Engng.* **61**, 2159–2181 (2004).
2. M. Arroyo and M. Ortiz, *Int. J. Numer. Meth. Engng*, **65**, 2167–2202 (2006).
3. N. Sukumar and R.W. Wright, *Int. J. Numer. Meth. Engng.* **70**, 181–205 (2007).
4. F. Greco, L. Filice, C. Peco and M. Arroyo, *Int. J. Mater. Form.* 8, 341–353 (2015).
5. A. Ortiz, M.A. Puso and N. Sukumar, *Comput. Methods. Appl. Mech. Eng*. **199**, 1859–1871 (2010).
6. A. Ortiz, M.A. Puso and N. Sukumar, *Finite elem. Anal. Des.* **47**, 572-585 (2011).
7. T. Belytschko, Y. Y. Lu and L. Gu, *Int. J. Numer. Meth. Engng.* **37**, 229–256 (1994).
8. A. Rosolen, D. Mill´an and M. Arroyo, *Int. J. Numer. Meth. Engng.* **82**, 868–895 (2010).
9. A. Rosolen, and M. Arroyo, *Comput. Methods. Appl. Mech. Eng*. **264**, 95–107 (2013).
10. Elias Cueto and Francisco Chinesta, *Int. J. Mater. Form.* **8**, 25–43 (2015).
11. L. L. Yaw, N. Sukumar, S. K. Kunnath, *Int. J. Numer. Meth. Engng.* **79**, 979–1003 (2009).
12. F. Greco and N. Sukumar, *Int. J. Numer. Meth. Engng.* **94**, 1123–1149 (2013).
13. A. Rosolen, C. Peco and M. Arroyo, *J. comput. Phys.* **249**, 303-319 (2013).