# Cell-based maximum-entropy approximants

Daniel Millán[a], N. Sukumar[b,*], Marino Arroyo[a,*]

[a] *LaCàN, Universitat Politècnica de Catalunya-BarcelonaTech, Barcelona 08034, Spain*
[b] *Department of Civil and Environmental Engineering, University of California, Davis, CA 95616, USA*

---

**Highlights**

- Smooth, nonnegative maximum-entropy approximants are constructed on unstructured Delaunay meshes.
- Relative entropy measure is minimized, subject to linear reproducing conditions.
- Nodal prior weight functions are formed by taking powers of smooth approximations to the distance function to the boundary of the support.
- Method delivers optimal convergence rates in Sobolev norms for two-dimensional elliptic boundary-value problems.

---

**Abstract**

In this paper, we devise cell-based maximum-entropy (max-ent) basis functions that are used in a Galerkin method for the solution of partial differential equations. The motivation behind this work is the construction of smooth approximants with controllable support on unstructured meshes. In the variational scheme to obtain max-ent basis functions, the nodal prior weight function is constructed from an approximate distance function to a polygonal curve in $\mathbb{R}^2$. More precisely, we take powers of the composition of R-functions via Boolean operations. The basis functions so constructed are nonnegative, smooth, linearly complete, and compactly-supported in a neighbor-ring of segments that enclose each node. The smoothness is controlled by two positive integer parameters: the normalization order of the approximation of the distance function and the power to which it is raised. The properties and mathematical foundations of the new compactly-supported approximants are described, and its use to solve two-dimensional elliptic boundary-value problems (Poisson equation and linear elasticity) is demonstrated. The sound accuracy and the optimal rates of convergence of the method in Sobolev norms are established.
© 2014 Elsevier B.V. All rights reserved.

*Keywords:* Delaunay mesh; Relative entropy; Smooth and nonnegative basis functions; Compact-support; R-functions; Approximate distance function

---

## 1. Introduction

Classical finite element (FE) approximations are built on basis functions that possess $C^0$-continuity. The advent of meshfree methods [1] provided $C^k$ ($k \geq 0$) basis functions that have been adopted in Galerkin methods for the solution

---

\* Corresponding authors.
*E-mail addresses:* nsukumar@ucdavis.edu (N. Sukumar), marino.arroyo@upc.edu (M. Arroyo).

of second- and higher-order partial differential equations (PDEs). Over the past decade, the importance of smooth basis functions that are also nonnegative has come to the forefront in applications that utilize B-splines, nonuniform rational B-splines (NURBS), isogeometric analysis (IGA) [2,3], subdivision surfaces [4,5] and maximum-entropy (max-ent) approximants [6–9]. The emergence of IGA has highlighted that such smooth basis functions are attractive for the numerical solution of PDEs. The strength of IGA is in high-fidelity boundary-representation, whereas its weakness is in realizing bulk discretizations; hence, coupling it with more flexible methods in the bulk has been explored [10,11]. Such coupling has been done with local max-ent (LME) [6] and with reproducing kernel particle method [12] that use moving least squares (MLS) approximants [13].

A common drawback of meshfree methods, which is particularly acute for large-scale three-dimensional problems, is the very dense nodal-connectivity structure. This leads to many nonzero entries in the stiffness matrix, which in turn places greater demands on memory usage and increases the CPU-times for the numerical simulations. Hence, to realize an efficient and viable Galerkin method on very large nodal sets, smooth basis functions that possess small (minimal) support are desirable. To this end, a method that leverages the complementary strengths of finite elements and spline-based techniques can provide significant advantages vis-à-vis the current state-of-the-art. In this paper, on adopting R-functions [14–16] within the relative entropy variational formulation, we build smooth local approximants that are linearly complete on unstructured meshes. As the first contribution of the new approach presented here, we describe the construction based on two-dimensional Delaunay meshes, and present numerical results for two-dimensional boundary-value problems on such meshes.

In the finite element literature, $C^1$ shape functions with different number of degrees of freedom (DOFs) for the triangle have been conceived: Argyris triangle (21 DOFs), Bell triangle (18 DOFs) and the composite Hsieh–Clough–Tocher (HCT) triangle (12 DOFs) are well-known [17]. The shape functions in the Argyris triangle span the space of complete fifth-order bivariate polynomials, whereas the Bell triangle is a reduced Argyris triangle. In both constructions, the function value, and its first- and second-order partial derivatives are the nodal unknowns. The HCT triangle is partitioned into three subtriangles and the approximation is a cubic polynomial in each subtriangle. Papanicolopulos and Zervos [18] have presented a systematic framework to derive $C^1$ shape functions on a triangle. It is noted that compared to linear finite elements on triangular meshes, use of $C^1$ formulations on such meshes require many more degrees of freedom. Meshfree methods have been tailored for unstructured meshes: Liu et al. [19] proposed the reproducing kernel element method, whereas Duarte et al. [20] constructed arbitrarily smooth generalized finite element approximations. In Ref. [20], a Shepard partition-of-unity function for polygonal domains is constructed using R-functions, which is then multiplied by a linear combination of monomials to form the discrete approximation. In all these previous developments that use finite elements and meshfree approximants, the basis functions that accrue are wiggly and change sign in general.

A simple illustration to demonstrate the effects of varying the support of the nodal basis functions follows. Let us denote the number of rings that surrounds a node by $N_R$: the one-ring ($N_R = 1$) for a node is the support of standard Delaunay interpolants (piecewise linear finite elements basis functions); the two-ring ($N_R = 2$) is that of subdivision basis functions [4]. In Fig. 1, the nodal basis functions that contribute at a sample point (marked by a ×) are shown for LME and for the cell-based max-ent (CME) approach that we propose. LME approximants have a Gaussian decay that is modulated by a nondimensional parameter $\gamma$, which controls the aspect ratio and effective support of the resulting basis functions as $h\sqrt{-\log(\epsilon_0)/\gamma}$, where $h$ is the nodal spacing and $\epsilon_0$ is a cutoff-tolerance below which the basis functions are considered to be zero. From Fig. 1, we observe that the nodal neighbors of a sample point for LME approximants with $\gamma = 0.8$ and $\epsilon_0 = 10^{-6}$ corresponds to about the four-ring ($N_R = 4$). The neighbors for $N_R = 2$ and $N_R = 3$ are also depicted, which we use in Section 3 for the CME approximants.

A dense connectivity structure leads to many entries in the system matrix. The sparsity of the system matrix is quantified by the total number of nonzeros (NNZ) of the matrix. We consider the NNZ generated by a set of $N$ points in 2D (3D) that are uniformly distributed in a square (cube), with one degree of freedom per node. In Fig. 2, we show the ratio between the NNZ for max-ent approximants and that of piecewise linear finite element basis functions ($N_R = 1$) in 2D (triangles) and 3D (tetrahedra). A dramatic rise in NNZ with increasing $N_R$ is observed, which is especially pronounced in three dimensions. In 3D, in comparison to the LME approximant ($\gamma = 0.8$, $\epsilon_0 = 10^{-6}$), the CME ($N_R = 2$) approximants leads to an order of magnitude fewer nonzeros in the system matrix. If cell-based approximants with tighter support can deliver comparable accuracy to standard meshfree basis functions at a significantly less computational cost, then these new basis functions are an attractive choice in Galerkin methods. Furthermore, due to the element-based support, the CME approach can be more easily integrated within existing finite element codes.
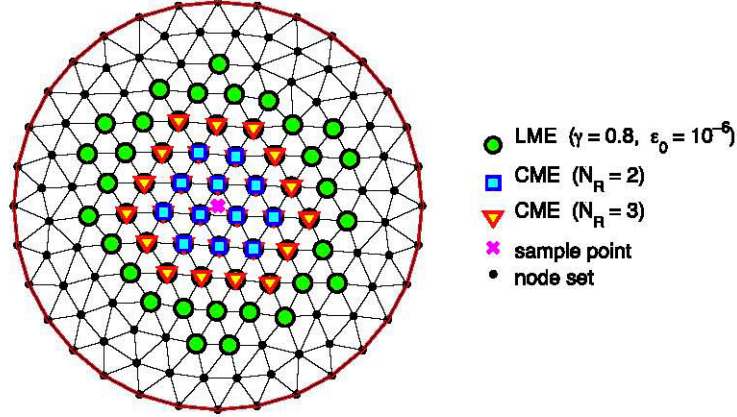
Fig. 1. Nearest nodes affecting a given sample point ($\times$) for local and cell-based maximum-entropy approximation schemes.
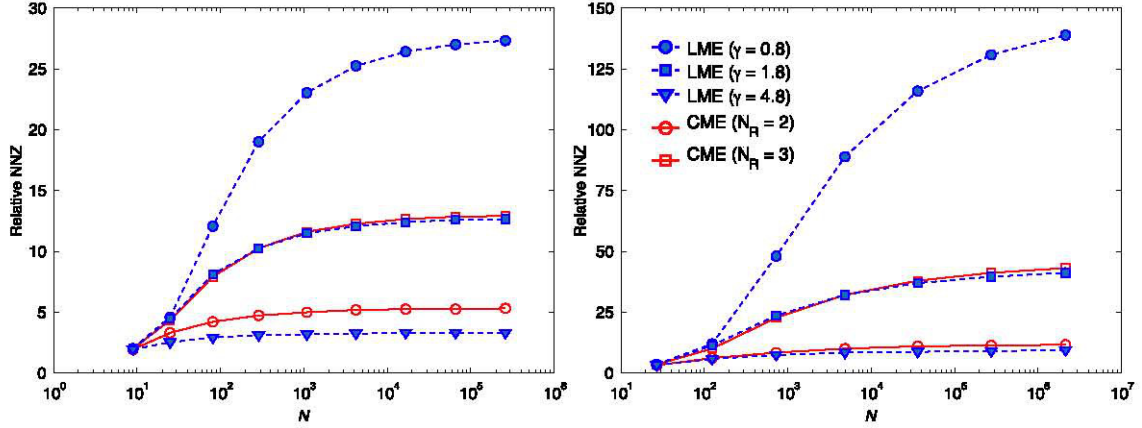


Fig. 2. Relative number of nonzeros (NNZ) of sparse matrices as a function of the number of nodes $N$. The NNZ are scaled by the NNZ of piecewise linear elements on triangles/tetrahedra, which is equivalent to considering $N_R = 1$. We show curves of the relative NNZ for CME with $N_R = 2, 3$, and for LME with $\gamma = 0.8, 1.8, 4.8$ and $\epsilon_0 = 10^{-6}$. Results for uniform grids in 2D (left) and 3D (right) are depicted.

This paper is organized as follows. Section 2 provides the constrained variational formulation for max-ent using the relative entropy functional. We also describe the construction of the nodal *prior* weight function, which is formed by joining R-functions that approximate the distance to a polygonal curve. The nodal weight function depends on two integer parameters $m$ and $s$: increasing values of $m$ provides a better approximation to the exact distance function, and the exponent $s$ controls the smoothness of the weight function. Applications to the Galerkin solution of elliptic PDEs are showcased in Section 3, which demonstrate the accuracy and the optimal rate of convergence of the method in Sobolev norms. Finally, some concluding remarks with promising directions for future work are mentioned in Section 4.

## 2. Method

We describe the main ingredients towards the conception of cell-based max-ent approximants. We begin with the notion of minimizing the relative entropy (Kullback–Leibler divergence) [21,22], where a discrete probability distribution $\{p_a\}_{a=1}^n$ is sought, given a known initial guess (*prior probabilities*) $\{q_a\}_{a=1}^n$ and a set of linear constraints involving $\{p_a\}_{a=1}^n$. For a nodal set $\{x_a\}_{a=1}^n \subset \mathbb{R}^d$ and a point $x \in \mathbb{R}^d$ that lies in the convex hull of this set, the connection is readily established to obtain nonnegative max-ent meshfree basis functions [6,23], where $\phi_a(x) \geq 0$ is $a$th basis function evaluated at $x$, and the prior probabilities are referred to as nodal prior weight functions $w_a(x) \geq 0$, $a = 1, \ldots, n$. The constant and linear reproducing conditions on $\{\phi_a(x)\}_{a=1}^n$ constitute the $d + 1$ equality constraints. The solution of the variational problem provides $\{\phi_a(x)\}_{a=1}^n$ that are closest to the chosen weight functions

and satisfy all the constraints, which can also be interpreted as a *correction* [12] of the nodal weight functions. The resulting convex approximant possesses a weak Kronecker-delta property on the boundary of the domain [6]. Furthermore, subject to some restrictions on the support-width of basis functions, the basis function $\phi_a(\boldsymbol{x})$ inherits the smoothness of the *a priori* known weight function $w_a(\boldsymbol{x})$ [24,25].

The CME approximant that we desire should be smooth and compactly-supported in polygons subordinate to a mesh. The polygons are built using $N_R$ number of nearest neighbors defined through the connectivity of the mesh. Let $\Omega_a^{N_R}$ be the polygonal region of the mesh that is within $N_R$ rings of $\boldsymbol{x}_a$, and if $\boldsymbol{x} \notin \Omega_a^{N_R}$, then we set $w_a(\boldsymbol{x}) := 0$. Formally, to achieve our stated objectives, we require prior weight functions $w_a(\boldsymbol{x})$ that must fulfill the following three conditions:

(i) supp $w_a(\boldsymbol{x}) = \Omega_a^{N_R}$

(ii) $w_a(\boldsymbol{x}) \geq 0$ and smooth in $\Omega_a^{N_R}$

(iii) $\partial^p w_a(\boldsymbol{x}) = 0$ on $\partial \Omega_a^{N_R} \setminus \partial \Omega$ for $p = 0, 1, \ldots, k$.

The last condition enforces that $w_a(\boldsymbol{x})$ and all its derivatives up to order $k$ vanish along the portion of the boundary of $\Omega_a^{N_R}$ that is interior to the domain (interior edges of the mesh). This ensures that $w_a(\boldsymbol{x}) \in C^k(\Omega)$, where $w_a(\boldsymbol{x})$ is extended to $\Omega \setminus \Omega_a^{N_R}$ by zero. If we extended condition (iii) to $\partial \Omega_a^{N_R}$ instead of $\partial \Omega_a^{N_R} \setminus \partial \Omega$, then our approximants would have zero gradient on the boundary of the domain, thereby excessively restricting the approximation space. We consider second- and third-rings of neighbors; the first-ring of neighbors, as in standard piecewise linear finite elements, is too narrow to achieve our objectives. First, the relative entropy formulation for max-ent basis functions is presented in Section 2.1. The main idea behind our construction of prior weight functions is as follows. Assume that there exists a function $d_a(\boldsymbol{x})$ that is smooth within $\Omega_a^{N_R}$, provides a sufficiently good approximation of the distance function to the boundary of the polygonal domain $\Omega_a^{N_R}$, and extends to zero outside the polygon. This function is only $C^0(\Omega)$, but setting $w_a(\boldsymbol{x}) = d_a^s(\boldsymbol{x})$ ($s \in \mathbb{N} \geq 2$) yields a compactly-supported function that is $C^{s-1}(\Omega)$. This observation, taken together with the construction of approximate distance functions based on R-functions [16] is used to form $w_a(\boldsymbol{x})$ in Section 2.2. Finally, we graphically illustrate the resulting CME basis functions and their gradients in Section 2.3.

### 2.1. Minimum relative entropy approximants

Maximum-entropy approximation schemes [6] fall within the class of convex approximants—other examples include natural neighbor approximants [7], subdivision schemes [4], NURBS and B-splines basis functions [2]. Approximants formed from convex basis functions $\{\phi_a(\boldsymbol{x})\}_{a=1}^n$ that reproduce affine functions are intimately linked to convex geometry. The max-ent approximants that we construct are only linearly complete. Extensions to second-order approximation schemes based on adding extra constraints is discussed in Refs. [9,26], whereas in Ref. [27], a method based on de Boor's algorithm is proposed. The max-ent approximation of a scalar-valued function $u(\boldsymbol{x})$ is written as

$$u^h(\boldsymbol{x}) := \sum_{a=1}^n \phi_a(\boldsymbol{x}) u_a, \qquad (1)$$

where $u_a$ are nodal coefficients, and the nonnegative basis functions $\{\phi_a(\boldsymbol{x})\}_{a=1}^n$ fulfill the zeroth- and first-order reproducing conditions:

$$\phi_a(\boldsymbol{x}) \geq 0, \qquad \sum_{a=1}^n \phi_a(\boldsymbol{x}) = 1, \qquad \sum_{a=1}^n \phi_a(\boldsymbol{x}) \boldsymbol{x}_a = \boldsymbol{x}. \qquad (2)$$

The definition of the max-ent approximant is not explicit, but rather follows from an optimization problem that is set up at each evaluation point $\boldsymbol{x}$, where $\{\phi_a(\boldsymbol{x})\}_{a=1}^n$ are the unknowns, and (2) are the linear constraints. As shown in Ref. [6], the constraints are only feasible within the convex hull of the nodal set.

Since the basis functions are nonnegative and form a partition of unity (zeroth-order reproducing condition) at each point $\boldsymbol{x}$, they can be interpreted as a discrete probability distribution. This information-theoretic viewpoint allows us to pose a statistical inference problem where the $\{\phi_a(\boldsymbol{x})\}_{a=1}^n$ are the unknowns. A canonical measure of the uncertainty associated with a discrete probability distribution is the informational entropy, and the principle of maximum entropy provides the least-biased approximation scheme that is consistent with the constraints.

When nodal prior weight functions are used, the formulation of max-ent approximants is based on maximizing the Shannon–Jaynes entropy measure [28]:

$$H(\boldsymbol{\phi}|\boldsymbol{w}) = -\sum_{a=1}^{n} \phi_a(\boldsymbol{x}) \ln\left(\frac{\phi_a(\boldsymbol{x})}{w_a(\boldsymbol{x})}\right), \tag{3}$$

where $D(\boldsymbol{\phi}|\boldsymbol{w}) := -H(\boldsymbol{\phi}|\boldsymbol{w}) \geq 0$ is the relative entropy measure and the principle of minimum relative (cross) entropy is the corresponding variational principle [22]. The variational formulation for max-ent basis functions can now be stated as:

$$
\begin{aligned}
(ME)_w \quad &\text{For fixed } \boldsymbol{x}, \text{ maximize} \quad H(\boldsymbol{\phi}|\boldsymbol{w}) \\
&\text{subject to} \quad \phi_a(\boldsymbol{x}) \geq 0 \quad (a = 1, \ldots, n) \\
&\qquad\qquad\quad \sum_{a=1}^{n} \phi_a(\boldsymbol{x}) = 1 \\
&\qquad\qquad\quad \sum_{a=1}^{n} \phi_a(\boldsymbol{x})\boldsymbol{x}_a = \boldsymbol{x}.
\end{aligned}
\tag{4}
$$

This concave optimization problem is efficiently and robustly solved using duality methods [6,23]:

$$\phi_a(\boldsymbol{x}) = \frac{w_a(\boldsymbol{x}) \exp\left[\boldsymbol{\lambda}(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{x}_a)\right]}{Z\left(\boldsymbol{x}, \boldsymbol{\lambda}(\boldsymbol{x})\right)},$$

where

$$Z(\boldsymbol{x}, \boldsymbol{\lambda}) = \sum_{b=1}^{n} w_b(\boldsymbol{x}) \exp\left[\boldsymbol{\lambda}(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{x}_b)\right]$$

is the partition function and $\boldsymbol{\lambda}$ is the Lagrange multiplier vector. The unconstrained convex optimization problem for the dual vector $\boldsymbol{\lambda}$ is:

$$\boldsymbol{\lambda}^*(\boldsymbol{x}) = \arg \min_{\boldsymbol{\lambda} \in \mathbb{R}^d} \ln Z(\boldsymbol{x}, \boldsymbol{\lambda}), \tag{5}$$

where $\boldsymbol{\lambda}^*$ is the converged solution for the Lagrange multiplier vector at $\boldsymbol{x}$.

The basis functions that result from the $(ME)_w$ convex scheme – and hence the approximant in (1) – inherit the smoothness of the nodal prior weight functions [24,25]. These approximants are noninterpolating in general, except on the boundary of the convex hull of the nodal set, where a weak Kronecker-delta property holds. This property renders it straightforward to impose essential boundary conditions unlike other meshfree methods such as those based on MLS approximants [29]. The expression for the gradient of the basis functions is provided in Ref. [30]. In the interest of being self-contained and complete, we present the derivation for $\nabla \phi_a(\boldsymbol{x})$ in Appendix A.

For the convex problem stated in (5) to have a unique solution for the Lagrange multiplier vector, every point $\boldsymbol{x} \in \Omega$ must be overlapped by at least $d+1$ nodal weight functions. The choice of the nodal prior weight functions provides flexibility to construct tailored approximants with desired properties—Gaussian weight function [6,31,32] and quartic polynomial weight function [30,33,34] in meshfree analysis, level set based nodal weight function (representing the union of edges) to construct linearly and quadratically precise shape functions on arbitrary planar polygons [35,36], and exponential nodal weight functions for convection–diffusion problems [37,38].

## 2.2. Nodal prior weight functions built from approximate distance functions

The desirable conditions that nodal prior weight functions must meet – smoothness, compact-support in a few rings of neighbors and unimodality (unique maxima) – motivates us to consider approximations of the distance function to planar curves [16]. Such distance functions find use in many applications: for example, heterogeneous material modeling [39], to satisfy Dirichlet boundary conditions in a Rayleigh–Ritz approximation [40], and to construct geometric enrichment for partition-of-unity approximations [41]. Since the nodal support of a basis function terminates on a polygon, we consider polygonal (closed and open) curves in $\mathbb{R}^2$ that are described by the union of piecewise linear

segments. We point out two well-known choices to approximate the distance function to polygons: (1) inverse of the sum of the nodal weight functions that appear in the formula for mean value coordinates [42] (see also Ref. [43]), and (2) the theory of R-functions [14]. Our focus is on smooth distance approximations developed for polygonal curves using R-functions [15]. While we adopt this particular approach, the field of constructive geometric modeling with implicit functions is rapidly evolving and may offer better solutions [44–47].

### 2.2.1. R-functions

Consider a real-valued function $F(\omega_1, \omega_2, \ldots, \omega_q)$, where $\omega_i(\boldsymbol{x}) : \mathbb{R}^2 \to \mathbb{R}$ $(i = 1, \ldots, q)$ are also real-valued functions. If the sign of $F(\cdot)$ is completely determined by only the signs and not the magnitude of its arguments $\omega_i(\boldsymbol{x})$, then $F(\cdot)$ is known as an R-function. For example, $F_1(x, y) = xy$ and $F_2(x, y) = x^2 + y^2$ are R-functions, whereas $F_3(x, y) = x^2 + y^2 - 1$ is not. On combining Boolean operations with such functions, the inverse problems of analytic geometry (solid modeling) are readily tackled using algebraic functions. The construction of implicit function representation for complex geometries permits fast evaluation of predicates (*on, in* or *out*) for geometric objects. Just as logical functions are written using the symbols: $\neg$ (negation), $\vee$ (union), $\wedge$ (intersection), and $\sim$ (equivalence), every R-function can be written as the composition of the corresponding elementary R-functions: R-negation ($\bar{x} \equiv -x$) that assumes the negative sign, R-disjunction ($\omega_1 \vee \omega_2$), R-conjunction ($\omega_1 \wedge \omega_2$), and R-equivalence ($\omega_1 \sim \omega_2$). We defer the discussion of R-equivalence to Section 2.2.4. A representative example of R-disjunction and R-conjunction is [15]:

$$R_\alpha(\omega_1, \omega_2) := \frac{1}{1 + \alpha} \left( \omega_1 + \omega_2 \pm \sqrt{\omega_1^2 + \omega_2^2 - 2\alpha\omega_1\omega_2} \right), \tag{6}$$

with (+) and (−) signs defining R-disjunction and R-conjunction, respectively. For $\alpha = 1$, $\omega_1 \vee \omega_2 = \max(\omega_1, \omega_2)$ and $\omega_1 \wedge \omega_2 = \min(\omega_1, \omega_2)$, which are the simplest examples of R-disjunction and R-conjunction, respectively. Note that the triangle inequality is used in (6) to write the term within the braces with $-1 < \alpha < 1$ being the cosine of the angle between the two sides. Hence, if $\omega_1$ and $\omega_2$ are positive, then so are $\omega_1 \vee \omega_2$ and $\omega_1 \wedge \omega_2$. The R-functions defined in (6) are not analytic at points where $\omega_1 = \omega_2 = 0$. Smoothness can be obtained by defining the function ($\alpha = 0$ is selected) [15]

$$R_s(\omega_1, \omega_2) := \omega_1 + \omega_2 \pm \sqrt{\omega_1^2 + \omega_2^2} \left( \omega_1^2 + \omega_2^2 \right)^{\frac{s}{2}}, \tag{7}$$

which renders these functions $C^s$-continuous. Eqs. (6) and (7) serve as prototypical examples of *R-functions*.

### 2.2.2. Distance functions and normalization

The distance function provides an implicit representation for curves and surfaces. Since the exact distance function is not smooth, we seek a smooth approximation to the exact distance field. The theory of R-functions is an appealing choice to construct smooth approximations to exact distance fields. Let $S \subset \mathbb{R}^d$ denote an object with boundary $\partial S$ and unit normal $\boldsymbol{n}$. We denote by $d \equiv d(\boldsymbol{x})$ an approximation to the exact distance function to $\partial S$. If $\partial S$ is a polygonal curve that is composed of piecewise linear segments, then we use $\rho \equiv \rho(\boldsymbol{x})$ to denote the approximate distance function to each segment. For a point $\boldsymbol{x} \in \mathbb{R}^d$ on $\partial S$, any meaningful approximation to the distance function should satisfy $d = 0$. Furthermore, to mimic the distance function, the normal derivative at the boundary should be unity ($\partial d/\partial n = 1$), and all higher-order normal derivatives should be zero. An $m$th order approximate distance function satisfies that the second- to $m$th-order normal derivatives vanish on all regular points (unit normal is well-defined) on $\partial S$ [16]:

$$\frac{\partial d}{\partial n} = 1; \qquad \frac{\partial^k d}{\partial n^k} = 0 \quad (k = 2, 3, \ldots, m), \tag{8}$$

and such a function is said to be *normalized* to the $m$th order. For finite $m$, the normalized function matches the exact distance function only in the vicinity of the boundary; for points that are away from the boundary, it deviates from the exact distance. As discussed in Ref. [16], normalized distance functions mitigate the undesirable *bulging* (undulations near the surface) phenomenon [44]. The R-conjunction and R-disjunction functions given in (7) (used in Ref. [20]) are not normalized.
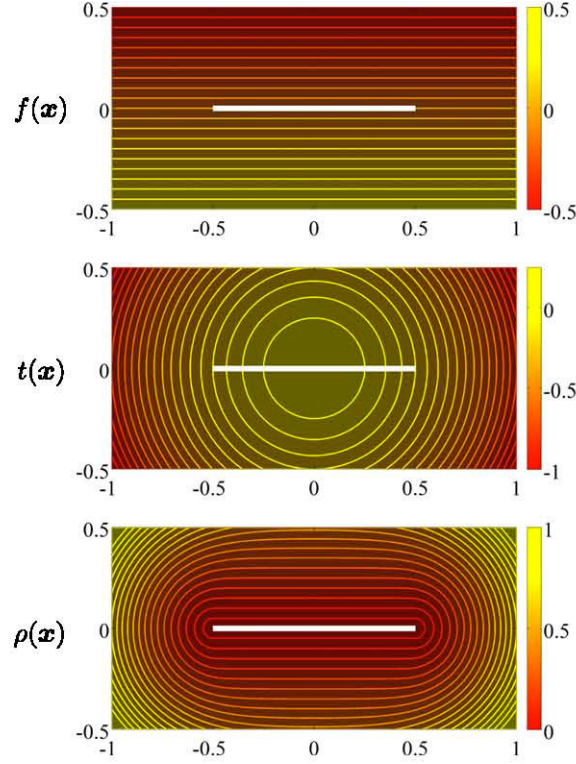
Fig. 3. The top plot shows the signed distance function (9) to a line enclosing a segment, the middle plot shows the trimming function (10), and the bottom plot shows the resulting approximation of the distance function (11) to a line segment.

### 2.2.3. Normalized functions for line segments

Given a node $a$ with coordinate $\boldsymbol{x}_a$, we compute the polygonal region $\Omega_a^{N_R}$ that lies within the $N_R$-ring of $\boldsymbol{x}_a$. The boundary, $\partial \Omega_a^{N_R}$, defines a polygonal curve that is composed of piecewise linear segments. Let us consider one line segment of this polygon with end-points $\boldsymbol{x}_1 \equiv (x_1, y_1)$ and $\boldsymbol{x}_2 \equiv (x_2, y_2)$. The center of this segment is denoted by $\boldsymbol{x}_c := (\boldsymbol{x}_1 + \boldsymbol{x}_2)/2$, and the length of the segment is: $L = \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|$. Now, we define [16]

$$f \equiv f(\boldsymbol{x}) := \frac{(x - x_1)(y_2 - y_1) - (y - y_1)(x_2 - x_1)}{L}, \tag{9}$$

which is the signed distance function from point $\boldsymbol{x}$ to the line that passes through $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.

Since the representation of the segment can be viewed as the intersection of an infinite line with a disk of radius $L/2$, we consider the following *trimming* function that is normalized to first-order [16]:

$$t \equiv t(\boldsymbol{x}) = \frac{1}{L}\left[\left(\frac{L}{2}\right)^2 - \|\boldsymbol{x} - \boldsymbol{x}_c\|^2\right], \tag{10}$$

where $t \geq 0$ defines a disk with center at $\boldsymbol{x}_c$. Now, with $f(\boldsymbol{x})$ and $t(\boldsymbol{x})$ on-hand, we define a normalized function (up to first-order) $\rho(\boldsymbol{x})$ that is differentiable for points away from the line segment [16]:

$$\rho \equiv \rho(\boldsymbol{x}) := \sqrt{f^2 + \frac{1}{4}\left(\sqrt{t^2 + f^4} - t\right)^2}. \tag{11}$$

This function is an approximation of the distance function to the segment with end points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. Fig. 3 provides a graphical illustration of $f$, $t$ and $\rho$.
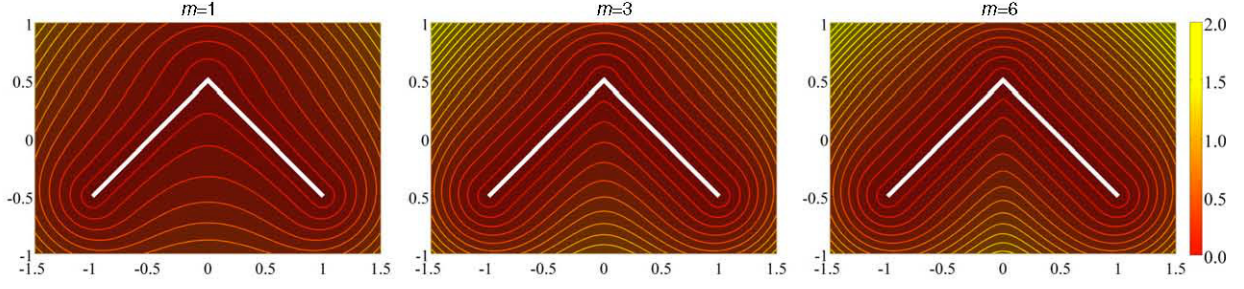
Fig. 4. Approximation of the distance function to two line segments by R-equivalence composition (12) for three different values of the normalization parameter $m$ ($m = 1, 3, 6$).

### 2.2.4. R-equivalence for line segments

Given the normalized functions $\rho_1$ and $\rho_2$ for two line segments $\ell_1$ and $\ell_2$ (see Section 2.2.3), then a distance field $F(\rho_1, \rho_2)$ for the union $\ell_1 \cup \ell_2$ must be zero when either $\rho_1 = 0$ or $\rho_2 = 0$ and positive otherwise. The naive R-equivalence formula $F(\rho_1, \rho_2) = \rho_1 \rho_2$ is no longer normalized at the regular points of the original lines; in Ref. [35] this is used for the union of edges of a polygon. An R-equivalence solution that preserves normalization up to order $m$ of the distance function to two line segments on all regular points (nonvertices) is provided in Biswas and Shapiro [16]:

$$\rho_1 \sim \rho_2 = \frac{\rho_1 \rho_2}{\sqrt[m]{\rho_1^m + \rho_2^m}} = \frac{1}{\sqrt[m]{\frac{1}{\rho_1^m} + \frac{1}{\rho_2^m}}}, \tag{12}$$

which is illustrated in Fig. 4.

The formula for the case when there are $n$ line segments that define a polygonal curve follows.

**Claim 1.** *If $\rho_1, \rho_2, \ldots, \rho_n$ are normalized functions for $n$ line segments that define a polygonal curve, then the approximation of the distance function $d$ that is normalized up to order $m$ is given by*

$$d(\ell_1, \ldots, \ell_n) := \rho_1 \sim \rho_2 \sim \cdots \sim \rho_n = \frac{1}{\sqrt[m]{\frac{1}{\rho_1^m} + \frac{1}{\rho_2^m} + \cdots + \frac{1}{\rho_n^m}}}. \tag{13}$$

**Proof.** We establish the above using mathematical induction. For $n = 2$, (13) holds since it reduces to the formula in (12). Let (13) hold for $n = k$ and we proceed to show that (13) is true for $n = k + 1$. For $n = k$, we have

$$d(\ell_1, \ldots, \ell_k) := \rho_1 \sim \rho_2 \sim \cdots \sim \rho_k = \frac{1}{\sqrt[m]{\frac{1}{\rho_1^m} + \frac{1}{\rho_2^m} + \cdots + \frac{1}{\rho_k^m}}}$$

and therefore

$$d(\ell_1, \ldots, \ell_k, \ell_{k+1}) := \rho_1 \sim \rho_2 \sim \cdots \sim \rho_k \sim \rho_{k+1} = (\rho_1 \sim \rho_2 \sim \cdots \sim \rho_k) \sim \rho_{k+1}$$

$$= \left( \frac{1}{\sqrt[m]{\frac{1}{\rho_1^m} + \frac{1}{\rho_2^m} + \cdots + \frac{1}{\rho_k^m}}} \right) \sim \rho_{k+1}$$

$$= \left( \left[ \sqrt[m]{\frac{1}{\rho_1^m} + \frac{1}{\rho_2^m} + \cdots + \frac{1}{\rho_k^m}} \right]^m + \frac{1}{\rho_{k+1}^m} \right)^{-\frac{1}{m}}$$

$$= \frac{1}{\sqrt[m]{\frac{1}{\rho_1^m} + \frac{1}{\rho_2^m} + \cdots + \frac{1}{\rho_k^m} + \frac{1}{\rho_{k+1}^m}}},$$
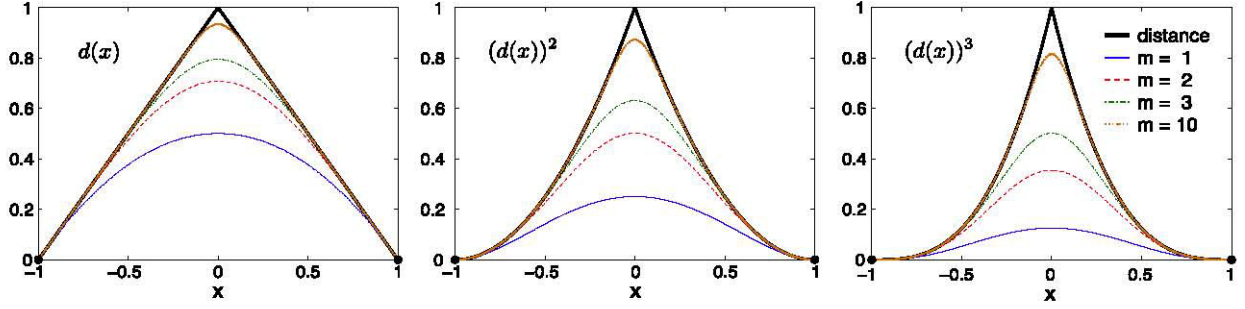
which completes the proof. □

Fig. 5. Approximate distance functions $d(x)$ in one-dimension to $x = \pm 1$ (shown as filled circles) are obtained by composition via R-equivalence up to normalization order $m = 1, 2, 3, 10$ (left). We also show the powers of the approximate distance functions $\left(d(x)\right)^s$ for $s = 2, 3$ (middle, right). In all plots, the exact distance function to a power $s$ is shown as a thick black line.

The R-conjunction, $\rho_1 \wedge \rho_2 = \rho_1 + \rho_2 - \sqrt[m]{\rho_1^m + \rho_2^m}$, is a function that is normalized to the $(m-1)$th order [16]. However, the joining operation is *not associative*, which makes this choice undesirable. The R-equivalence joining relation, as (13) reveals, is associative.

### 2.2.5. Normalized distance function for a polygonal curve and prior weight functions

As mentioned earlier, for a node $a$, we seek a smooth nodal prior weight function $w_a \equiv w_a(\boldsymbol{x})$ that vanishes on the boundary of the polygon, $\partial \Omega_a^{N_R} \setminus \partial \Omega$. Let $\partial \Omega_a^{N_R} \setminus \partial \Omega$ be composed of $n_a$ line segments. As the first step, we consider the joining of normalized distance functions for these $n_a$ line segments via (13):

$$d_a \equiv d_a(\boldsymbol{x}) := \rho_1 \sim \rho_2 \sim \cdots \sim \rho_{n_a}. \tag{14}$$

Now to modulate the smoothness of the desired weight function, we raise $d_a$ to a power $s \geq 2$, and to ensure that the weights are bounded, they are normalized so that $0 \leq w_a \leq 1$ form a partition-of-unity:

$$w_a(\boldsymbol{x}) = \frac{d_a^s(\boldsymbol{x})}{\displaystyle\sum_{b \in \mathscr{N}_{\boldsymbol{x}}} d_b^s(\boldsymbol{x})}, \tag{15}$$

where $\mathscr{N}_{\boldsymbol{x}}$ are the indices of the $N_R$-ring nodal neighbors of point $\boldsymbol{x}$. The gradients of $w_a(\boldsymbol{x})$ and $d_a(\boldsymbol{x})$ are presented in Appendix B.

We now illustrate these functions via a one-dimensional example. Consider the prior weight function $w(x)$ for a node located at $x = 0$; $w(x)$ must vanish at $x = \pm 1$. Plots of powers of the actual distance and its R-function approximations, $d^s(x)$, are shown in Fig. 5. The plot on the left depicts curves for $s = 1$ and varying $m$ (order of normalization); as we approach $x = \pm 1$, the nodal weight functions agree with the exact piecewise linear distance function. The effect of the higher-order normalization is clear in this figure. The smoothness near the boundary is accomplished by raising $d(x)$ to a power $s$, which is shown in Fig. 5 (right). Indeed, since

$$\left(d^s(x)\right)' = s d^{s-1}(x) d'(x), \tag{16a}$$

$$\left(d^s(x)\right)'' = s(s-1) d^{s-2}(x) \left[d'(x)\right]^2 + s d^{s-1}(x) d''(x) \tag{16b}$$

and so on for any $m \geq 1$, we obtain $C^{s-1}$ extensions of $d^s(x)$ on the real line for $s \geq 2$.

Now we present an example in two dimensions. Consider a polygonal curve that is formed by the union of line segments from the boundary of the two-ring neighbors ($\partial \Omega^{N_R}$ with $N_R = 2$) of a point located at the origin $\boldsymbol{x}_a \equiv (0,0)$. Fig. 6 shows powers of the approximated distance function $d_a^s(\boldsymbol{x})$ for $s = 1, 2$ and their associated prior weight functions $w_a(\boldsymbol{x})$ computed using (15). As expected, for $s = 1$, the approximate distance function is only $C^0(\Omega)$. The partition of unity propagates the derivative discontinuities in $w_a(\boldsymbol{x})$ to all the edges in the mesh. Instead, for $s = 2$, the prior weight function is $C^1(\Omega)$. As discussed in Ref. [16], joining operations such as the R-equivalence in (13) introduce undesirable bulging and creasing effects at vertices. Here, this effect is particularly noticeable when
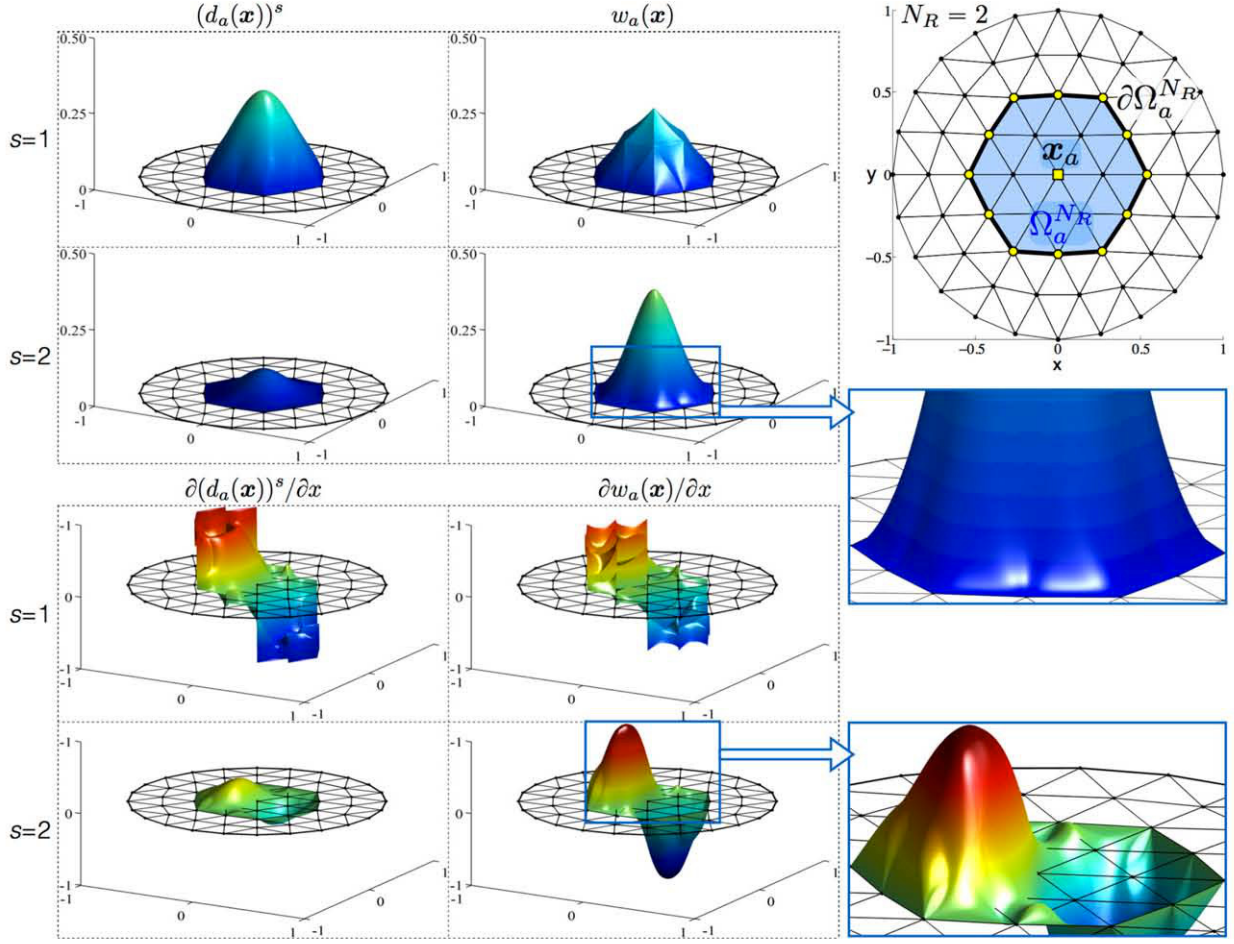
Fig. 6. The figure on the upper-right corner shows the mesh considered in this example. The boundary of the polygonal curve $\partial\Omega_a^{N_R}$ for a node $x_a$ located at the origin ($N_R = 2$) has been highlighted and is composed of $n_a = 12$ line segments. The four panels on the top-left show the approximate distance function raised to different powers $d_a^s(x)$ and the associated prior weight function $w_a(x)$ (Shepard-form) for the node located at the origin. The approximate distance functions have been obtained by using composition via R-equivalence up to normalization order $m = 2$, and powers $s = 1, 2$. The bottom four panels display the derivative along the $x$-direction, where the bulging/creasing effects are apparent close to the vertices of the polygon; see zoom views.

two nearly collinear segments are joined, which is observable in Fig. 6. When segments are exactly collinear, they can be merged to avoid bulging. As shown later, increasing $m$ or $s$ makes bulging less prominent. While the resulting prior weight functions are smooth, alternative approaches to geometric modeling with implicit functions may lead to even smoother functions [44–47].

## 2.3. Basis functions

Having shown that the technology of R-functions allows us to produce prior weight functions $w_a(x)$ satisfying the requirements listed at the beginning of this section, we now proceed to examine linearly complete basis functions that result from entropy maximization in (4). As mentioned earlier, by construction these functions inherit the smoothness and support of the prior weight functions, and therefore the properties of the basis functions can be controlled by the number of rings $N_R$, the normalization order $m$, and the exponent $s$.

We consider a convex domain, and second- and third-rings. The basis function of an interior node is shown in Fig. 7 for $N_R = 2$ and different choices for $m$ and $s$. One of the partial derivatives is also depicted. We observe that the support of the basis functions coincide with the two-ring set of elements. Mathematically, the smoothness of the basis functions increases with $s$; however, increasing this parameter also flattens out the basis functions at the
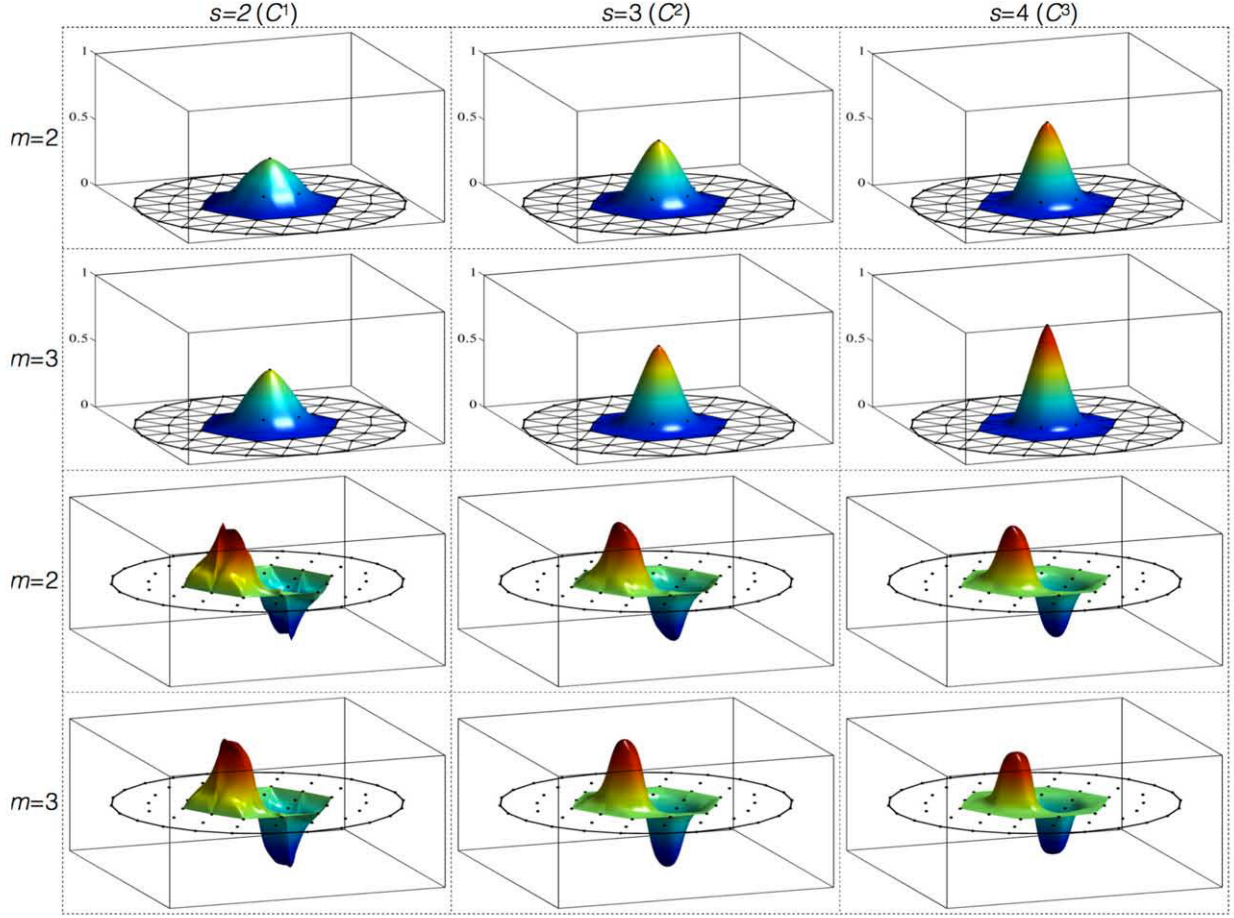
Fig. 7. Cell-based max-ent basis functions (arbitrary scale) for several values of the normalization order $m$ and the exponent $s$. The number of neighbor rings is two ($N_R = 2$). The top two rows display plots of the basis functions, and the bottom two rows show the derivative along the $x$-direction.

boundary of their support. Computational experiments reveal that when $s$ increases, while remaining smooth, the basis functions tend to the piecewise linear finite element basis functions. Increasing $m$ improves the approximation to the exact distance function, and hence leads to smeared creases in the basis functions. However, there is a broad range of parameters that lead to smooth and visually pleasant basis functions. Fig. 8 shows the basis functions obtained for $N_R = 3$. From these plots, we recommend parameters in the range $s = 3, 4$, and $m = 2, 3$.

One of the original motivations for developing CME was to create smooth basis functions, similar to LME, but with a smaller support conforming to cells. In Fig. 9, we compare a LME basis function ($\gamma = 1.7$, $\epsilon_0 = 10^{-6}$) with an analogous CME basis function ($N_R = 2$, $s = 3$, $m = 2$). Both basis functions are visually very similar, and have nearly the same maximum value. The figure clearly shows that the support of the CME function is significantly smaller than that of the LME function, which approximately extends to the three-ring neighborhood.

## 3. Numerical examples

The performance of CME approximants for the Galerkin solution of elliptic boundary-value problems is assessed. We consider two Poisson problems, and a standard benchmark problem in linear elasticity. In all the examples, we compare the proposed CME approximants with two and three rings to LME approximants with aspect ratio parameters $\gamma = 0.8, 1.8, 4.8$, and also to linear finite elements. For the LME computations, $\epsilon_0 = 10^{-6}$ is used. Gauss–Legendre integration rules for triangular domains are considered, and the number of quadrature points per element is selected in accordance with the discretization technique. For linear finite elements, 3 Gauss points are used to accurately integrate
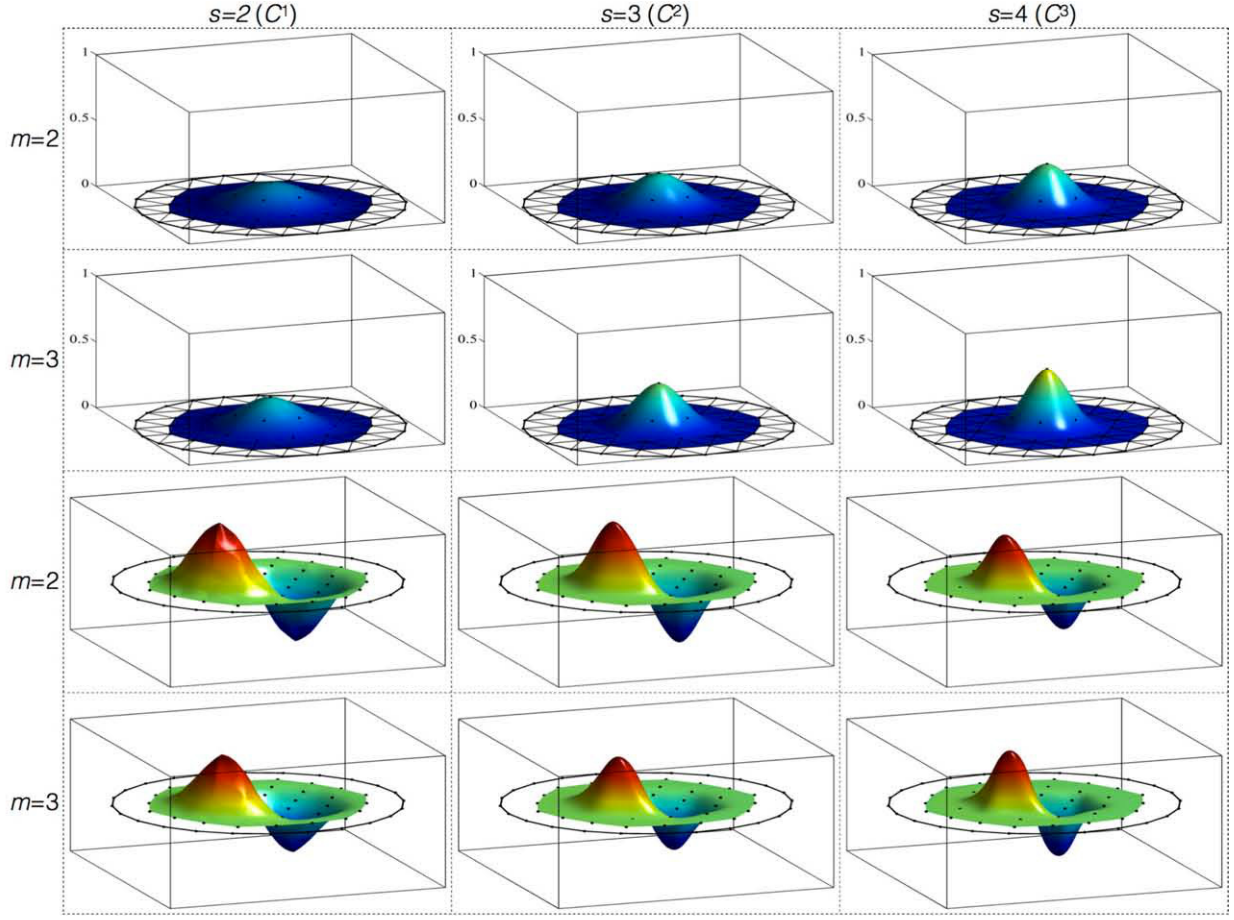
Fig. 8. Cell-based max-ent basis functions (arbitrary scale) for several values of the normalization order $m$ and the exponent $s$. The number of neighbor rings is three ($N_R = 3$). The top two rows display plots of the basis functions, and the bottom two rows show the derivative along the $x$-direction.
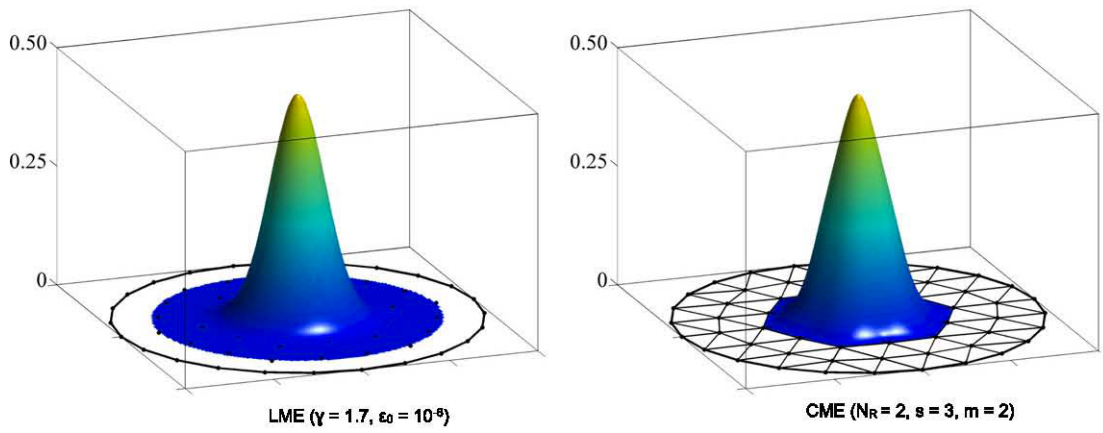


Fig. 9. An example of visually similar LME and CME basis functions. The support size of the CME basis function is significantly smaller.

the source terms, whereas for the LME approach the number of integration points are 12, 7, 6 for $\gamma = 0.8, 1.8, 4.8$, respectively. The appropriate quadrature for CME computations is investigated in the next section.

### 3.1. Homogeneous Poisson problem

Consider the following Poisson boundary-value problem with nonhomogeneous Dirichlet boundary conditions:

$$\Delta u = 0 \quad \text{in } \Omega = (0, 1)^2 \subset \mathbb{R}^2,$$
$$u(0, y) = u(1, y) = u(x, 0) = 0, \tag{17}$$
$$u(x, 1) = x(1 - x),$$

whose exact solution is:

$$u(x, y) = \frac{4}{\pi^3} \sum_{i=1}^{\infty} \frac{1 - (-1)^i}{i^3 \sinh(i\pi)} \sin(i\pi x) \sinh(i\pi y). \tag{18}$$

Maximum-entropy approximants satisfy a weak Kronecker-delta property on the boundary of the convex hull of the nodes, which facilitates the imposition of essential boundary conditions. For this example, this weak Kronecker-delta property implies that only nodes on a given face contribute to the approximation on that particular face. Since a quadratic field is imposed on the top edge, we compute the nodal coefficients involved in imposing the nonhomogeneous Dirichlet boundary condition via a least-squares fit involving the basis functions of nodes on that edge only.

We first examine the sensitivity of the CME solutions to the accuracy of the quadrature. A Delaunay mesh of the domain $\Omega$ is used, with Gauss–Legendre quadrature rules of order 2, 3, 4, 5, 6, 7, 8, 10, and 15 on each triangle; these integration schemes have 3, 4, 6, 7, 12, 13, 16, 25 and 54 quadrature points per triangle, respectively. Fig. 10 (left) shows the relative error in the $L_2$ norm for two different levels of refinement ($h = 1/64, 1/128$), and different number of quadrature points. These results suggest that rules with at least 12 quadrature points are adequate to ensure that the integration error is significantly less than the approximation error. From hereon, we use 12 Gauss points for all calculations involving CME approximants.

Results (not shown) for the CME approximants are relatively insensitive to combinations of parameters $s = 3, 4$ and $m = 2, 3$. We have selected as representative values the pair ($s = 4, m = 2$) for $N_R = 2$, and the pair ($s = 3, m = 2$) for $N_R = 3$. The convergence plots depicted in Fig. 10 (right) compare FE, LME ($\gamma = 0.8, 1.8, 4.8$) and CME schemes. The relative error in the $L_2$ norm converges at the optimal rate of 2 for all methods. The accuracy of CME is on par with LME ($\gamma = 1.8$). In particular, CME with $N_R = 2$ is almost as accurate as LME with $\gamma = 1.8$, but at a substantially less computational cost (see Fig. 2).

### 3.2. Poisson problem with localized sources

We present the convergence in the $L_2$ norm of the FE, LME, and CME trial approximants in the Galerkin solution of the following Poisson boundary-value problem:

$$-\Delta u = s(x, y) \quad \text{in } \Omega = (0, 1)^2$$
$$u = \bar{u} \quad \text{on } \partial\Omega,$$

where the source $s(x, y)$ and the boundary data $\bar{u}$ are chosen such that the exact solution is:

$$u(x, y) = \sum_{i=1}^{4} A_i \, e^{-\beta_i \left[(x - x_i)^2 + (y - y_i)^2\right]}.$$

The coefficients that appear in the exact solution are provided in Table 1. This example was considered by Rosolen et al. [9], and is chosen here to showcase how CME approximants on unstructured meshes can readily resolve the localized sharp features in the exact solution.

The numerical solutions are computed on meshes with an adapted distribution of nodes as illustrated in Fig. 11 (left). The adapted grids are obtained with centroidal Voronoi tessellations (CVT) [48,49], which allow the distribution of nodes to follow a prescribed density function. A density function that is proportional to the norm of the gradients of the exact solution is used. Note that the CVT method does not create nested grids upon refinement, and the notion of nested approximants does not apply to CME or meshfree methods in general.

Table 1
Coefficients to calculate the exact solution for the Poisson problem.

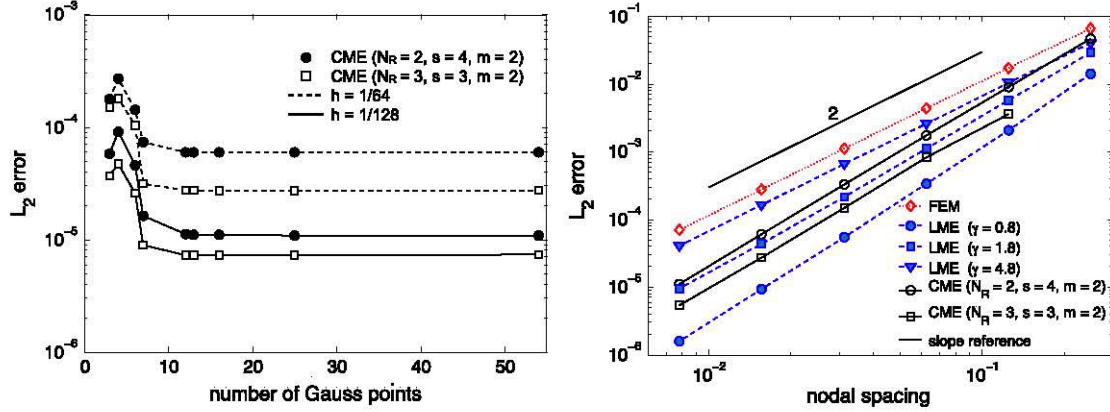| $i$ | $A_i$ | $\beta_i$ | $x_i$ | $y_i$ |
|---|---|---|---|---|
| 1 | 10 | 180 | 0.51 | 0.52 |
| 2 | 50 | 450 | 0.31 | 0.34 |
| 3 | 100 | 800 | 0.73 | 0.71 |
| 4 | 50 | 1000 | 0.28 | 0.72 |



Fig. 10. Comparison of FE, LME and CME approximants for a Poisson problem with nonhomogeneous Dirichlet boundary conditions. Left plot shows the error sensitivity with respect to the number of quadrature points for CME approximants. The results on the left panel correspond to two uniform meshes of $65 \times 65$ nodes ($h = 1/64$, dashed line) and $129 \times 129$ nodes ($h = 1/128$, solid line). The plot on the right shows convergence curves for the relative error in the $L_2$ norm. Results for CME on the right panel have been obtained with 12 quadrature points.
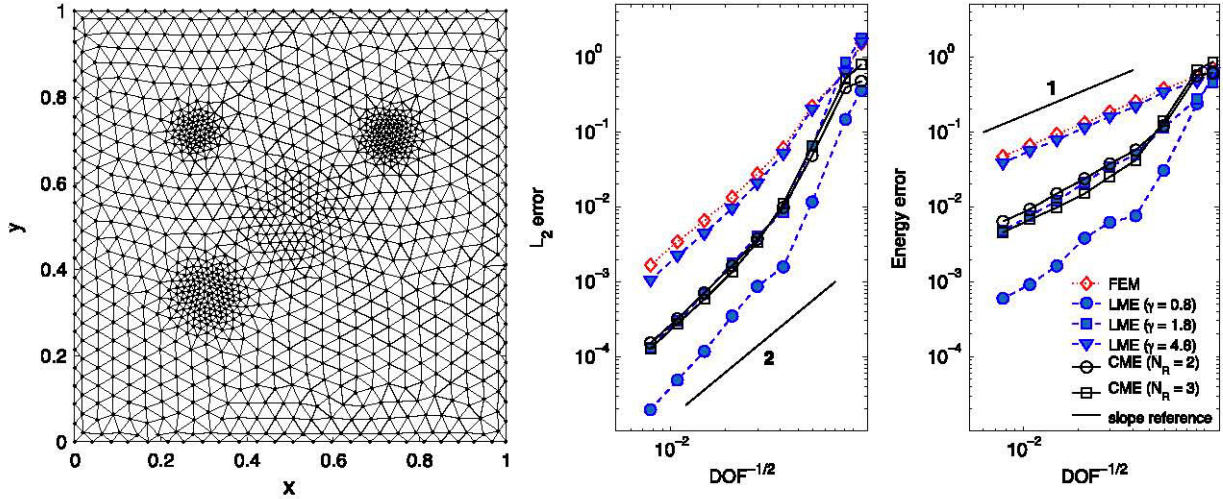


Fig. 11. Illustration of an adapted mesh (set of nodes) for the Poisson problem with localized sources (left). Relative error in the $L_2$ norm and the energy seminorm (right) on the adapted meshes: FE, LME ($\gamma = 0.8, 1.8, 4.8$; $\epsilon_0 = 10^{-6}$) and CME ($N_R = 2, s = 4, m = 2$; $N_R = 3, s = 3, m = 2$) approximants are used..

The numerical solution is computed for each grid with FE, LME ($\gamma = 0.8, 1.8, 4.8$; $\epsilon_0 = 10^{-6}$), and CME ($N_R = 2, s = 4, m = 2$; $N_R = 3, s = 3, m = 2$) approximation schemes. The convergence curves of the relative error in the $L_2$ norm and the energy seminorm are shown in Fig. 11 (right). The guiding straight lines with slopes of two and one are also depicted. It can be seen that all methods exhibit the optimal rate of convergence. In this example, we again observe that CME and LME ($\gamma = 1.8$) approximants perform similarly.
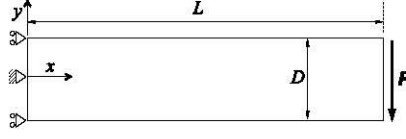
14

Fig. 12. Schematic representation of the cantilever beam problem.

### 3.3. Cantilever beam problem

We consider the standard benchmark problem of an isotropic, linear elastic cantilever beam, which is built-in on one end and subjected to a parabolic shear traction at the other end (Fig. 12). The analytical solution for the displacement field is [50]:

$$u(x, y) = -\frac{Py}{6\bar{E}I}\left[(6L - 3x)x + (2 + \bar{\nu})\left(y^2 - \frac{D^2}{4}\right)\right],$$ (19a)

$$v(x, y) = \frac{P}{6\bar{E}I}\left[3\bar{\nu}y^2(L - x) + (4 + 5\bar{\nu})\frac{D^2}{4}x + (3L - x)x^2\right],$$ (19b)

where $\bar{E} = E/(1 - \nu^2)$ and $\bar{\nu} = \nu/(1 - \nu)$ are the effective Young's modulus and Poisson's ratio under plane strain conditions. The exact solution for the stresses are [50]:

$$\sigma_1 = -\frac{P}{I}(L - x)y,$$ (20a)

$$\sigma_2 = 0,$$ (20b)

$$\tau_{12} = \frac{P}{2I}\left(\frac{D^2}{4} - y^2\right),$$ (20c)

where $L$ and $D$ are the length and height of the beam, $I = D^3/12$ is the moment of inertia of the beam with cross-section of unit width, and $P$ is the total shear load at the right end.

In the numerical computations, we choose $L = 4$, $P = 1000$, $E = 10^7$, and Poisson's ratio $\nu = 0.3$. We consider a beam whose length-to-height ratio is 4:1. The problem is symmetric with respect to the neutral axis of the beam, and therefore only the upper-half of the beam is analyzed. In Fig. 13, the rates of convergence in the $L_2$ norm and the energy seminorm are plotted for FE, LME and CME approximants. The $L_2$ norm of the error is defined as:

$$\|e\|_{L_2} = \left(\int_\Omega \|u - u_h\|^2 \, dx\right)^{1/2},$$ (21)

whereas the energy seminorm is:

$$\|e\|_E = \left(\frac{1}{2}\int_\Omega (\varepsilon - \varepsilon_h) : (\sigma - \sigma_h) \, dx\right)^{1/2}.$$ (22)

In (21) and (22), $u$, $\varepsilon$, and $\sigma$ denote the exact displacement, strain and stress fields, whereas $u_h$, $\varepsilon_h$, and $\sigma_h$ denote the corresponding numerical approximations.

Convergence plots are shown in Fig. 13. The CME approximant converges optimally: method delivers rates of 2 and 1 in the $L_2$ norm and the energy seminorm, respectively. Numerical results reveal that the accuracy of CME is comparable to that obtained with LME approximants, the best known scheme among convex meshfree schemes. In particular, CME with $N_R = 2$ is comparable in accuracy to the LME approximants with $\gamma = 1.8$, which exhibits a significantly denser connectivity structure. Strikingly, CME with $N_R = 3$ provides similar accuracy in the $L_2$ norm than the very expensive LME approximants with $\gamma = 0.8$ (see Fig. 2).

## 4. Concluding remarks

In this paper, we have proposed an approximation method with smooth and nonnegative basis functions, which are supported on groups of elements of a mesh. Our method to obtain the basis functions relies on minimizing the relative
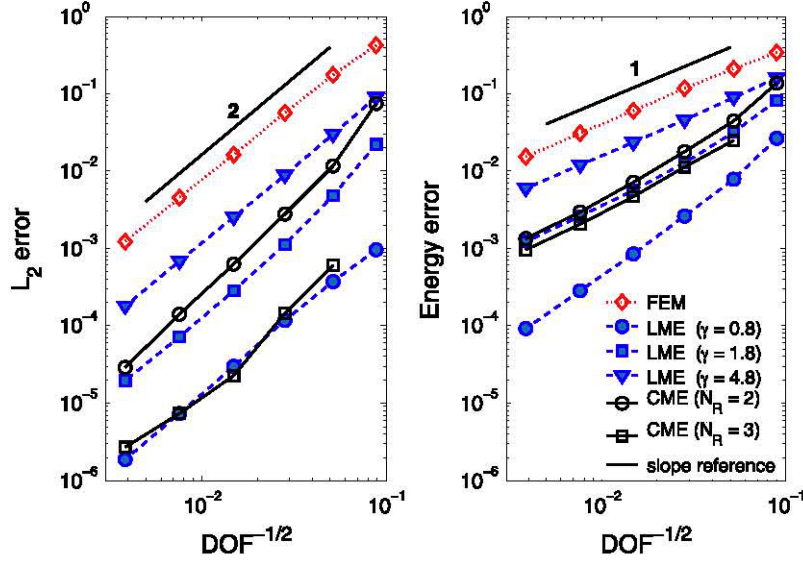
Fig. 13. Rates of convergence in the $L_2$ norm and energy seminorm for linear FE, LME ($\gamma = 0.8, 1.8, 4.8$; $\epsilon_0 = 10^{-6}$) and CME approximants for $N_R = 2$ ($s = 4, m = 2$) and $N_R = 3$ ($s = 3, m = 2$).

entropy functional – starting from properly designed prior weight functions that have the desired support and smoothness – and subject to the linear reproducing conditions as the constraints. To construct these prior weight functions, we have considered powers of smooth approximations to the distance function to the boundary of the support. We have shown that these cell-based max-ent (CME) functions perform very well (accurate and deliver optimal rates of convergence in Sobolev norms) in the Galerkin approximation of two-dimensional elliptic partial differential equations. The method is closely related to maximum-entropy and other meshfree methods. While the proposed method relies on a mesh, it retains the high-accuracy of meshfree approximations for smooth problems, with the advantage of having tight and well-controlled supports of the basis functions. This greatly reduces the computational costs in terms of memory and operation counts for large-scale problems. In the context of convex approximations, our method provides a general route to smooth and nonnegative approximants supported on general unstructured grids. This issue is not settled at this point. Subdivision schemes are limited to two dimensions, and spline-based techniques are quite rigid with regards to domain topology or local refinement. We have implemented the method in two dimensions, but it easily extends to three dimensions, once the approximate distance function to element faces is defined. We believe that the proposed method can be refined by incorporating advances from constructive geometric modeling with implicit functions in the construction of the prior weight functions.

The present work provides many new opportunities for future research. For example, some of the problems that we are currently pursuing are: (1) extending CME schemes to nonconvex domains; (2) implementing these approximants on meshes with quadrangular and hexahedral elements; (3) blending CME approximants with a high-fidelity description of the boundaries; and (4) developing approximants that satisfy the second-order reproducing conditions.

## Acknowledgments

### Appendix A. Derivatives of the basis functions

We present the derivation for the derivatives of the max-ent basis functions. The spatial gradient of a scalar-valued function $f(\boldsymbol{x})$ is denoted by $\nabla f(\boldsymbol{x})$, whereas for vector-valued functions, $\boldsymbol{Du}(\boldsymbol{x})$ represents the matrix of partial derivatives of $\boldsymbol{u}(\boldsymbol{x})$. The symbol $\partial$ denotes partial differentiation, and subindices $a$, $b$ and $c$ refer to nodes. Within the scope of the appendix, we define the following functions:

$$f_a(\boldsymbol{x}, \boldsymbol{\lambda}) := \exp[\boldsymbol{\lambda} \cdot (\boldsymbol{x} - \boldsymbol{x}_a)], \qquad Z(\boldsymbol{x}, \boldsymbol{\lambda}) := \sum_{b=1}^{n} w_b(\boldsymbol{x}) f_b(\boldsymbol{x}, \boldsymbol{\lambda}) \tag{A.1a}$$

$$g_a(\boldsymbol{x}, \boldsymbol{\lambda}) := \frac{f_a(\boldsymbol{x}, \boldsymbol{\lambda})}{Z(\boldsymbol{x}, \boldsymbol{\lambda})} \tag{A.1b}$$

$$\phi_a(\boldsymbol{x}, \boldsymbol{\lambda}) := \frac{w_a(\boldsymbol{x}) f_a(\boldsymbol{x}, \boldsymbol{\lambda})}{Z(\boldsymbol{x}, \boldsymbol{\lambda})} = w_a(\boldsymbol{x}) g_a(\boldsymbol{x}, \boldsymbol{\lambda}) \tag{A.1c}$$

$$\boldsymbol{r}(\boldsymbol{x}, \boldsymbol{\lambda}) := \sum_{a=1}^{n} \phi_a(\boldsymbol{x}, \boldsymbol{\lambda}) \, (\boldsymbol{x} - \boldsymbol{x}_a) \tag{A.1d}$$

$$\boldsymbol{J}(\boldsymbol{x}, \boldsymbol{\lambda}) := \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\lambda}} = \sum_{a=1}^{n} \phi_a(\boldsymbol{x}, \boldsymbol{\lambda}) \, (\boldsymbol{x} - \boldsymbol{x}_a) \otimes (\boldsymbol{x} - \boldsymbol{x}_a) - \boldsymbol{r}(\boldsymbol{x}, \boldsymbol{\lambda}) \otimes \boldsymbol{r}(\boldsymbol{x}, \boldsymbol{\lambda}). \tag{A.1e}$$

From hereon, to ease the notation, the dependence of the above functions on the evaluation point $\boldsymbol{x}$ and the Lagrange multiplier vector $\boldsymbol{\lambda}$ is suppressed. A superscript $*$ on any function indicates that the function is evaluated at $\boldsymbol{\lambda}^*(\boldsymbol{x})$, which is given in (5). This introduces explicit and implicit dependences on $\boldsymbol{x}$ in all functions with a superscript $*$. Note that what has been denoted by $\phi_a$ in the main body of the paper is denoted by $\phi_a^*$ in this appendix. No implied sum is assumed for repeated nodal indices.

Let $\nabla \phi_a^*$ denote the gradient of the basis function $\phi_a^*$. It follows that

$$\nabla \phi_a^* = \nabla w_a \, g_a^* + w_a \nabla g_a^*. \tag{A.2}$$

On applying the chain rule, we have

$$\nabla g_a^* = \left( \frac{\partial g_a}{\partial \boldsymbol{x}} \right)^* + \left( \frac{\partial g_a}{\partial \boldsymbol{\lambda}} \right)^* \boldsymbol{D}\boldsymbol{\lambda}^*, \tag{A.3a}$$

where

$$\left( \frac{\partial g_a}{\partial \boldsymbol{x}} \right)^* = -g_a^* \sum_{b=1}^{n} g_b^* \nabla w_b, \qquad \left( \frac{\partial g_a}{\partial \boldsymbol{\lambda}} \right)^* = g_a^* \, (\boldsymbol{x} - \boldsymbol{x}_a) \tag{A.3b}$$

are obtained using (A.1b). The only term that is not explicitly available in Eq. (A.3) is $\boldsymbol{D}\boldsymbol{\lambda}^*$. In order to compute it, we note that since $\boldsymbol{r}^*$ is identically the zero-vector:

$$\boldsymbol{0} = \boldsymbol{D}\boldsymbol{r}^* = \left( \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}} \right)^* + \left( \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\lambda}} \right)^* \boldsymbol{D}\boldsymbol{\lambda}^*,$$

where

$$\left( \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\lambda}} \right)^* = \boldsymbol{J}^*, \qquad \left( \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}} \right)^* = \sum_{a=1}^{n} g_a^* \nabla w_a \otimes (\boldsymbol{x} - \boldsymbol{x}_a) + \boldsymbol{I}.$$

It follows that

$$\boldsymbol{D}\boldsymbol{\lambda}^* = -(\boldsymbol{J}^*)^{-1} \left( \sum_{a=1}^{n} g_a^* \nabla w_a \otimes (\boldsymbol{x} - \boldsymbol{x}_a) + \boldsymbol{I} \right). \tag{A.4}$$

From (A.2) and (A.3), we finally obtain the expression for the spatial gradient of $\phi_a^*$:

$$\nabla \phi_a^* = g_a^* \, \nabla w_a + \phi_a^* \left[ (\boldsymbol{x} - \boldsymbol{x}_a) \cdot D\boldsymbol{\lambda}^* - \sum_{b=1}^{n} g_b^* \, \nabla w_b \right], \tag{A.5}$$

where $D\boldsymbol{\lambda}^*$ is given in (A.4). Expressions for $w_a$ and $\nabla w_a$ are given in (15) and (B.1), respectively.

## Appendix B. Gradient of the prior weight functions

We present the expression for the gradient of the prior weight function $w_a(\boldsymbol{x})$, which requires the gradient of the approximation of the distance function $d_a(\boldsymbol{x})$ that is computed via the R-equivalence joining relation, and the gradient of the normalized function for a line segment, $\rho(\boldsymbol{x})$, which is given in (11).

In the interest of clarity, we once again suppress the dependence on $\boldsymbol{x}$ for the functions that follow. The gradient of the prior weight function $w_a$ is:

$$\nabla w_a = \frac{s}{Z_w} \left( d_a^{s-1} \nabla d_a - w_a \sum_{b \in \mathscr{N}_{\boldsymbol{x}}} d_b^{s-1} \nabla d_b \right), \tag{B.1}$$

where $\mathscr{N}_{\boldsymbol{x}}$ contains the indices of the $N_R$-ring nodal neighbors of $\boldsymbol{x}$, and the partition function $Z_w \equiv Z_w(\boldsymbol{x})$ is:

$$Z_w = \sum_{b \in \mathscr{N}_{\boldsymbol{x}}} d_b^s. \tag{B.2}$$

The gradient of the normalized distance function $d_a$ (see (13) and (14)) is:

$$\nabla d_a = \frac{\sum_{i=1}^{n_a} (\rho_i)^{-m-1} \nabla \rho_i}{\left( \sum_{i=1}^{n_a} (\rho_i)^{-m} \right)^{\frac{m+1}{m}}}, \tag{B.3}$$

where $\rho_i$ are normalized functions (see Section 2). The gradient of this function for a given segment is:

$$\nabla \rho = \frac{1}{\rho} \left( f \nabla f + \left[ \frac{\sqrt{t^2 + f^4} - t}{4} \right] \left[ \frac{t \nabla t + 2 f^3 \nabla f}{\sqrt{t^2 + f^4}} - \nabla t \right] \right), \tag{B.4}$$

where for ease of notation we have omitted the subindex $i$. Finally, the gradient of the distance function to a line $f$ in (9) and that of the trimming function $t$ in (10) are:

$$\nabla f = \frac{1}{L} \left[ (y_2 - y_1) \boldsymbol{e}_1 - (x_2 - x_1) \boldsymbol{e}_2 \right], \qquad \nabla t = -\frac{2}{L} (\boldsymbol{x} - \boldsymbol{x}_c). \tag{B.5}$$

## References

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

[2] T. Hughes, J. Cottrell, Y. Bazilevs, Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement, Comput. Methods Appl. Mech. Engrg. 194 (2005) 4135–4195.

[3] Y. Bazilevs, V. Calo, J. Cottrell, J. Evans, T. Hughes, S. Lipton, M. Scott, T. Sederberg, Isogeometric analysis using T-Splines, Comput. Methods Appl. Mech. Engrg. 199 (5–8) (2010) 229–263.

[4] F. Cirak, M. Ortiz, P. Schröder, Subdivision surfaces: a new paradigm for thin-shell finite-element analysis, Int. J. Numer. Methods Eng. 47 (2000) 2039–2072.

[5] F. Cirak, Q. Long, Subdivision shells with exact boundary control and non-manifold geometry, Int. J. Numer. Methods Eng. 88 (9) (2011) 897–923.

[6] M. Arroyo, M. Ortiz, Local maximum-entropy approximation schemes: a seamless bridge between finite elements and meshfree methods, Int. J. Numer. Methods Eng. 65 (2006) 2167–2202.

[7] N. Sukumar, Construction of polygonal interpolants: a maximum entropy approach, Int. J. Numer. Methods Eng. 61 (12) (2004) 2159–2181.

[8] A. Rosolen, D. Millán, M. Arroyo, On the optimum support size in meshfree methods: a variational adaptivity approach with maximum entropy approximants, Int. J. Numer. Methods Eng. 82 (7) (2010) 868–895.

[9] A. Rosolen, D. Millán, M. Arroyo, Second order convex *maximum entropy* approximants with applications to high order PDE, Int. J. Numer. Methods Eng. 94 (2) (2013) 150–182.

[10] A. Rosolen, M. Arroyo, Blending isogeometric analysis and maximum entropy meshfree approximants, Comput. Methods Appl. Mech. Engrg. 264 (2013) 95–107.

[11] D. Wang, H. Zhang, A consistently coupled isogeometric–meshfree method, Comput. Methods Appl. Mech. Engrg. 268 (0) (2014) 843–870.

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

[13] P. Lancaster, K. Salkauskas, Surfaces generated by moving least squares methods, Math. Comp. 37 (155) (1981) 141–158.

[14] V. Shapiro, Theory of R-functions and applications: A primer, Tech. Rep. CPA88-3, in: Cornell Programmable Automation, Sibley School of Mechanical Engineering, Ithaca, NY 14853, 1991.

[15] V. Shapiro, Semi-analytic geometry with R-functions, Acta Numer. 16 (2007) 239–303.

[16] A. Biswas, V. Shapiro, Approximate distance fields with non-vanishing gradients, Graph. Models 66 (3) (2004) 133–159.

[17] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

[18] S.-A. Papanicolopulos, A. Zervos, A method for creating a class of triangular $C^1$ finite elements, Int. J. Numer. Methods Eng. 89 (2012) 1437–1450.

[19] W.K. Liu, W. Han, H. Lu, S. Li, J. Cao, Reproducing kernel element method. Part I: Theoretical formulation, Comput. Methods Appl. Mech. Engrg. 193 (2004) 933–951.

[20] C.A. Duarte, D.-J. Kim, D.M. Quaresma, Arbitrarily smooth generalized finite element approximations, Comput. Methods Appl. Mech. Engrg. 196 (1) (2006) 33–56.

[21] S. Kullback, R.A. Leibler, On information and sufficiency, Ann. Math. Stat. 22 (1) (1951) 79–86.

[22] J.E. Shore, R.W. Johnson, Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, IEEE Trans. Inform. Theory 26 (1) (1980) 26–36.

[23] N. Sukumar, R.W. Wright, Overview and construction of meshfree basis functions: From moving least squares to entropy approximants, Int. J. Numer. Methods Eng. 70 (2) (2007) 181–205.

[24] N. Sukumar, R.J.-B. Wets, Deriving the continuity of maximum-entropy basis functions via variational analysis, SIAM J. Optim. 18 (3) (2007) 914–925.

[25] M. Arroyo, M. Ortiz, Local maximum-entropy approximation schemes, in: Meshfree Methods for Partial Differential Equations III, in: Lecture Notes in Computational Science and Engineering, vol. 57, Springer, 2007.

[26] C. Cyron, M. Arroyo, M. Ortiz, Smooth, second order, non-negative meshfree approximants selected by maximum entropy, Int. J. Numer. Methods Eng. 79 (13) (2009) 1605–1632.

[27] D. González, E. Cueto, M. Doblaré, A higher-order method based on local maximum entropy approximation, Int. J. Numer. Methods Eng. 83 (6) (2010) 741–764.

[28] E.T. Jaynes, Information theory and statistical mechanics, in: K. Ford (Ed.), Statistical Physics: The 1962 Brandeis Lectures, W. A. Benjamin, New York, 1963, pp. 181–218.

[29] S. Fernández-Méndez, A. Huerta, Imposing essential boundary conditions in mesh-free methods, Comput. Methods Appl. Mech. Engrg. 193 (2004) 1257–1275.

[30] L.L. Yaw, N. Sukumar, S.K. Kunnath, Meshfree co-rotational formulation for two-dimensional continua, Int. J. Numer. Methods Eng. 79 (8) (2009) 979–1003.

[31] B. Li, F. Habbal, M. Ortiz, Optimal transportation meshfree approximation schemes for fluid and plastic flows, Int. J. Numer. Methods Eng. 83 (2010) 1541–1579.

[32] C. Peco, A. Rosolen, M. Arroyo, An adaptive meshfree method for phase-field models of biomembranes. Part I: Approximation with maximum-entropy basis functions, J. Comput. Phys. 249 (2013) 303–319.

[33] A. Ortiz, M.A. Puso, N. Sukumar, Maximum-entropy meshfree method for compressible and near-incompressible elasticity, Comput. Methods Appl. Mech. Engrg. 199 (25–28) (2010) 1859–1871.

[34] J.S. Hale, P.M. Baiz, A locking-free meshfree method for the simulation of shear-deformable plates based on a mixed variational formulation, Comput. Methods Appl. Mech. Engrg. 241–244 (2012) 311–322.

[35] K. Hormann, N. Sukumar, Maximum entropy coordinates for arbitrary polytopes, Comput. Graph. Forum 27 (5) (2008) 1513–1520.

[36] N. Sukumar, Quadratic maximum-entropy serendipity shape functions for arbitrary planar polygons, Comput. Methods Appl. Mech. Engrg. 263 (2013) 27–41.

[37] K. Nissen, C.J. Cyron, V. Gravemeier, W.A. Wall, Information-flux method: a meshfree maximum-entropy Petrov–Galerkin method including stabilised finite element methods, Comput. Methods Appl. Mech. Engrg. 241–244 (2012) 225–237.

[38] C.T. Wu, D.L. Young, H.K. Hong, Adaptive meshless local maximum-entropy finite element method for convection–diffusion problems, Comput. Mech. 53 (2014) 189–200.

[39] A. Biswas, V. Shapiro, I. Tsukanov, Heterogeneous material modeling with distance fields, Comput. Aided Geom. Des. 21 (3) (2004) 215–242.

[40] M. Freytag, V. Shapiro, I. Tsukanov, Field modeling with sampled distances, Comput.-Aided Des. 38 (2006) 87–100.

[41] A. Tambat, G. Subbarayan, Isogeometric enriched field approximations, Comput. Methods Appl. Mech. Engrg. 245–246 (2012) 1–21.

[42] K. Hormann, M.S. Floater, Mean value coordinates for arbitrary planar polygons, ACM Trans. Graph. 25 (4) (2006) 1424–1441.

[43] A. Belyaev, P.-A. Fayolle, A. Pasko, Signed $L_p$-distance fields, Comput.-Aided Des. 45 (2013) 523–528.

[44] J. Bloomenthal, Bulge elimination in convolution surfaces, Comput. Graph. Forum 16 (1) (1997) 31–41.

[45] A. Sherstyuk, Kernel functions in convolution surfaces: a comparative analysis, Vis. Comput. 15 (4) (1999) 171–182.

[46] L. Barthe, B. Wyvill, E. De Groot, Controllable binary csg operators for "soft objects", Int. J. Shape Model. 10 (02) (2004) 135–154.

[47] O. Gourmel, L. Barthe, M.-P. Cani, B. Wyvill, A. Bernhardt, M. Paulin, H. Grasberger, A gradient-based implicit blend, ACM Trans. Graph. 32 (2) (2013) 12:1–12:12.

[48] Q. Du, V. Faber, M. Gunzburger, Centroidal Voronoi tessellations: Applications and algorithms, SIAM Rev. 41 (4) (1999) 637–676.

[49] Y. Liu, W. Wang, B. Lévy, F. Sun, D.-M. Yan, L. Lu, C. Yang, On centroidal Voronoi tessellation—energy smoothness and fast computation, ACM Trans. Graph. 28 (4) (2009) 1–17.

[50] S. Timoshenko, S. Woinowsky-Kreiger, Theory of Plates and Shells, McGraw-Hill, 1959.