On the optimum support size in meshfree methods: A variational adaptivity approach with maximum-entropy approximants

Adrian Rosolen, Daniel Millán and Marino Arroyo*,†

Department of Applied Mathematics 3, LaCàN, Universitat Politècnica de Catalunya (UPC), Barcelona 08034, Spain

SUMMARY

We present a method for the automatic adaption of the support size of meshfree basis functions in the context of the numerical approximation of boundary value problems stemming from a minimum principle. The method is based on a variational approach, and the central idea is that the variational principle selects both the discretized physical fields and the discretization parameters, here those defining the support size of each basis function. We consider local maximum-entropy approximation schemes, which exhibit smooth basis functions with respect to both space and the discretization parameters (the node location and the locality parameters). We illustrate by the Poisson, linear and non-linear elasticity problems the effectivity of the method, which produces very accurate solutions with very coarse discretizations and finds unexpected patterns of the support size of the shape functions.

KEY WORDS: meshfree methods; local maximum-entropy approximants; variational adaptivity; support size

1. INTRODUCTION

In the numerical approximation of partial differential equations, meshfree methods are numerical strategies combining two elements: (1) basis functions built from a scattered set of nodes and not supported on a mesh or grid and (2) a discretization principle of the governing equations such as the Galerkin method or point collocation. These methods have been applied for over 30

---

years in complex simulations involving impact, crack propagation, large deformations, vibrations and acoustics, and flow problems to name a few [1–3]. Despite a large number of the proposed meshfree methods, these rely on a handful of meshfree basis functions or approximants [4], the most popular being those based on the moving least-squares (MLS) idea [5]. Recently, the information-theoretic concept of maximum-entropy (*max-ent*) has been put forth to develop meshfree first-order and second-order approximants [6, 7], and polygonal approximants [8]. In the meshfree flavor, *max-ent* approximants present some advantages over MLS approximants (e.g. their positivity, the straightforward imposition of boundary data, the simpler quadrature, or the robustness of their evaluation). In both MLS and *max-ent* approximants, one can adjust the *locality* of the shape functions, understood as the extent of the support of the basis functions relative to the nominal nodal spacing. In MLS methods, the locality is controlled through the so-called dilation parameter.

The most visible advantage of meshfree methods over the conventional mesh-based methods is the flexibility in the definition and the adaption of the spacial discretization. Not only it is possible to scatter nodes in the domain at will, without caring about any mesh or connectivity (other than neighbor lists), but one can also select different support sizes for each basis function. An understated advantage of meshfree methods based on smooth basis function, e.g. MLS or *max-ent* approximants, is the very high accuracy that is achieved in problems with smooth solution as compared with C^0 finite elements (FEs) of the same order of polynomial consistency. *Max-ent* C^∞ approximants have even been shown to be considerably more accurate than the B-spline shape basis functions in the Galerkin approximation of vibration and heat conduction problems [7]. This accuracy comes at a cost, in that the meshfree shape functions cannot be, in general, evaluated explicitly but rather require the solution of a local problem, a linear system for MLS, and a convex optimization problem for *max-ent*. More important than the computational cost of the basis function evaluation is the need for expensive quadrature rules, and the extra band-width in the systems of equations as a result of the larger support size. Despite these drawbacks, in many examples the high accuracy for very few degrees of freedom overrules the extra computational cost as compared with standard FEs or B-Splines [6, 7]. Thus, meshfree methods seem to be well suited in problems with smooth solutions and when accurate solutions are needed with stringent limitations on the number of degrees of freedom.

While some authors have exploited the flexibility of meshfree methods in adaptive strategies by arranging the nodes to capture sharp features of the solution or to control the discretization error below a given bound [9, 10], the locality of the shape functions has not been adapted to enhance the accuracy of the numerical solution. Most often, the dilation parameter or locality is kept fixed, partly due to the structure of the *a priori* error estimates for meshfree methods [11], needed in node adaption strategies. However, the locality of the meshfree shape functions has a strong effect on the accuracy of the numerical solutions. By way of an illustration, Figure 1 shows the convergence plots obtained with the local (first order) *max-ent* (LME) approximants in a Galerkin approximation of a 2D linear elasticity problem. In this method, as discussed below in detail, a non-dimensional parameter γ controls the locality. As it tends to infinity, it has been proved [6] that the affine function supported on the Delaunay triangulation of the node set is recovered. In practice, for $\gamma = 4$ the shape functions are visually very close to the Delaunay approximants (see Figure 2). Figure 1 shows that the accuracy of the numerical solution measured in the L_2 norm can change by almost two orders of magnitude depending on the locality. It can be observed that the less local the functions, the more accurate the solution up to a limit. The figure also illustrates an important practical point; it can be noticed that for very low values of γ, the accuracy and convergence rate are degraded. This can be fixed by using a more accurate and very expensive
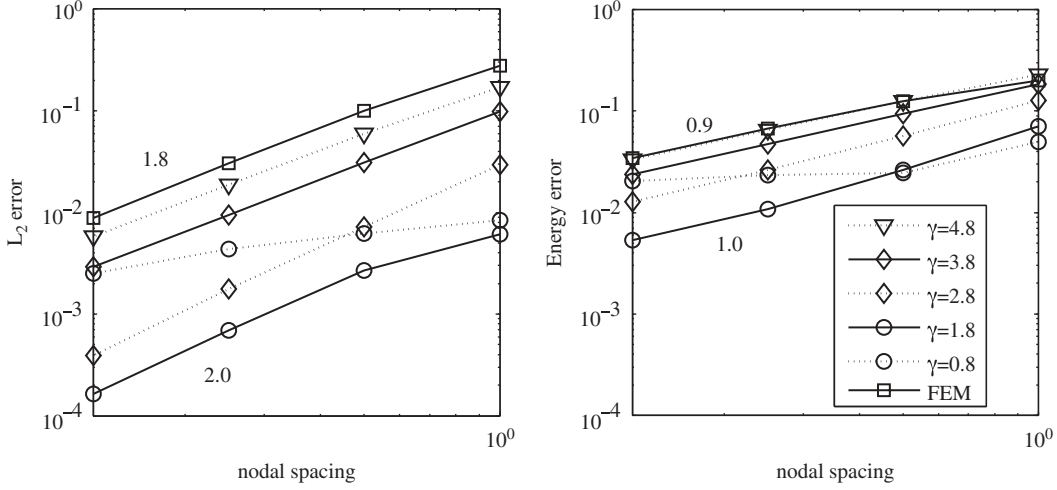
Figure 1. Convergence of first-order *max-ent* approximants in the Galerkin approximation of a 2D linear elasticity problem [6] for several values of the locality parameter γ.
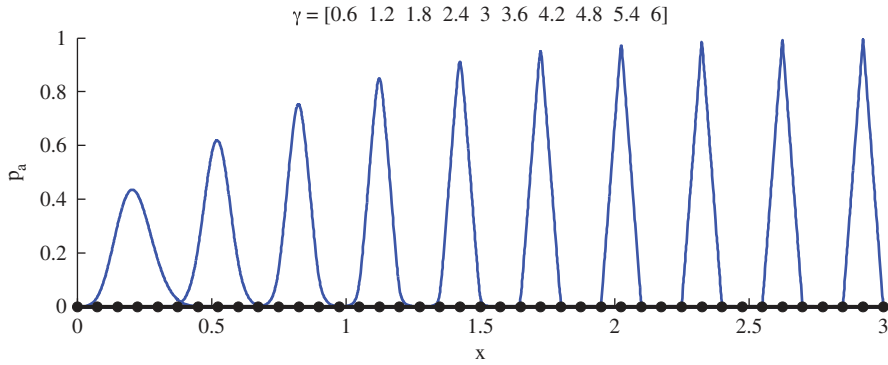

Figure 2. Seamless transition from meshfree to the Delaunay affine basis functions. The transition is controlled by the non-dimensional nodal parameters γ_a, which here take linearly varying values from 0.6 (left) to 6 (right).

quadrature, with 12 Gauss points per triangle in the quadrature mesh used in this example. In summary, in this example widespread shape functions result in very accurate numerical solutions, but the numerical integration rule can become prohibitively expensive.

This example highlights that the accuracy of the numerical solution for a fixed number of nodes is considerably affected by the choice of locality. Unfortunately, the optimal support size of the basis functions is problem-dependent, and even depends on the level of refinement [3, 6, 12]. Furthermore, it is expectable that the optimal locality may be different in different regions of the domain. Yet, *a priori* error estimates do not provide clues on how to adapt the support size [10, 11]. Although experience can be useful in many cases to estimate an appropriate support size for the shape functions, the presence of free and very sensitive parameters in meshfree methods

is an obstacle to a more widespread use. In this paper, we explore the automatic adaption of the support size of the shape functions in the context of the LME approximants and partial differential equations stemming from a minimum principle. Here, we consider Poisson's equation, and linear and non-linear elasticity approximated by the Galerkin method. We follow a variational adaptivity approach, closely related to the variational node relocation methods proposed recently [13–15]. The central idea of the method is that the variational principle governs the selection of both the discretized physical fields and the discretization itself. Although similar methods are conceivable for MLS approximants, LME approximants are particularly well suited for variational adaption because (1) the basis functions depend smoothly (C^∞) on the parameters governing the locality, with explicitly computable sensitivities, and (2) the evaluation of the shape functions is robust with respect to the locality parameters, without the solvability problems of MLS methods for small dilation parameters.

The structure of this paper is as follows. Section 2 introduces a modified formulation of the LME approximants, particularly appropriate for the purpose of the paper. In Section 3, we describe a framework for locality variational adaption. Numerical examples are presented and discussed in Section 4. An algorithm for the optimization problem is also proposed. Some concluding remarks are collected in Section 5.

2. LME APPROXIMANTS

LME meshfree approximants, introduced in [6], fall into the general class of convex approximation schemes, like natural neighbor approximants [16], subdivision approximants [17], or B-spline and NURBS basis functions [18]. Convex basis functions, which we will denote by $p_a(\boldsymbol{x}), a = 1, \ldots, N$ with $\boldsymbol{x} \in \mathbb{R}^d$, are non-negative approximants that fulfil the zeroth-order and first-order consistency conditions, and are intimately related to convex geometry. The consistency conditions are expressed by

$$p_a(\boldsymbol{x}) \geqslant 0, \quad \sum_{a=1}^{N} p_a(\boldsymbol{x}) = 1, \quad \sum_{a=1}^{N} p_a(\boldsymbol{x}) \, \boldsymbol{x}_a = \boldsymbol{x}, \tag{1}$$

where the last equation allows us to identify the vectorial weights $\boldsymbol{x}_a$ with the positions of the *nodes* associated with each basis function. We denote the set of nodes by $X = \{\boldsymbol{x}_a\}_{a=1,\ldots,N}$, and its convex hull by

$$\operatorname{conv} X = \left\{ \boldsymbol{x} \in \mathbb{R}^d \,\middle|\, \boldsymbol{x} = \sum_{a=1}^{N} \eta_a \boldsymbol{x}_a, \text{ with } \eta_a \geqslant 0, \sum_{a=1}^{N} \eta_a = 1 \right\}.$$

As shown in [6], convex approximants can only exist within the convex hull (or subsets of it) and satisfy *ab initio* a weak Kronecker-delta property at the boundary of the convex hull of the nodes. With this property, the imposition of essential boundary conditions in the Galerkin methods is straightforward. Note that convex approximants can be used in non-convex domains, as illustrated later in the paper. However, the weak Kronecker-delta property does not hold in the non-convex parts of the boundary of the domain.

There are infinitely many convex approximants for a node set of $d+2$ or more affinely independent nodes. In *max-ent* approximation schemes, one selects from this set of approximants-distinguished basis functions, whose functional form is not explicitly known in general but can

4

be efficiently computed. A computationally and conceptually appealing procedure is to select the optimal choice from an information-theoretic viewpoint. For this, one only needs to realize that according to Equation (1), we can interpret the shape functions at each space location x as a discrete probability distribution for a scheme with N events associated with the nodes. From this probability distribution, we know that the expectation of the position random variable is precisely x (first-order consistency condition). With this interpretation, the approximation of a function $u(x) \approx \sum_{a=1}^{N} p_a(x) u_a$ from the nodal values $\{u_a\}_{a=1,\dots,N}$ can be understood as computing an expectation or average. The principle of *max-ent* postulates that the least biased distribution consistent with the known information maximizes Shannon's entropy subject to the constraints:

$$\text{For fixed } x \text{ maximize } H(p_1, p_2, \dots, p_N) = -\sum_{a=1}^{N} p_a \ln p_a,$$

$$\text{subject to } p_a \geqslant 0, \quad a = 1, \dots, N,$$

$$\sum_{a=1}^{N} p_a = 1, \quad \sum_{a=1}^{N} p_a x_a = x.$$

The solution of this optimization problem for a given x is the set of basis functions at this point, $p_a(x), a = 1, \dots, N$. This pure entropy maximization problem leads to non-local shape functions [6], and has been proposed and used to develop basis functions for polygonal elements [8].

This kind of optimization problem to select convex approximants is reminiscent of Rajan's variational formulation of the Delaunay triangulation in arbitrary dimensions [19]. In this reference, it was proven that the shape functions of the Delaunay triangulation are the solutions of the following linear program:

$$\text{For fixed } x \text{ minimize } U(x, p_1, p_2, \dots, p_N) = \sum_{a=1}^{N} p_a |x - x_a|^2,$$

$$\text{subject to } p_a \geqslant 0, \quad a = 1, \dots, N,$$

$$\sum_{a=1}^{N} p_a = 1, \quad \sum_{a=1}^{N} p_a x_a = x.$$

The objective function can be interpreted as a measure of the width or locality of the shape functions. Note that, despite that all sums in the above program run over all the nodes or basis functions, for the solution only $d + 1$ shape functions are non-zero at each x.

The idea behind the LME basis functions is to select the convex approximants that exhibit a (Pareto) compromise between competing objectives, entropy maximization, and minimum width. By minimizing the function $\beta U(x, p_1, p_2, \dots, p_N) - H(p_1, p_2, \dots, p_N)$ subject to the usual constraints, meshfree-type approximants of controllable locality can be built. The non-negative parameter β, which weighs the relative importance given to each of the objectives, can in principle be a function of position. As fully detailed in [6], it can be mathematically proved that the optimization problem has a unique solution within $\text{conv} X$, the resulting shape functions are as smooth as $\beta(x)$ with respect to position (e.g. C^∞ if β is constant), are smooth (C^∞) with respect to β, can be efficiently computed using duality methods (see below), and *max-ent* and the Delaunay approximants arise as specialized limits. An extension to second-order schemes has been proposed in [7].

Thus, LME approximants offer the opportunity of seamlessly transitioning from widespread basis functions in parts of the computational domain to tightly supported functions in other parts, by appropriately selecting the function $\beta(\boldsymbol{x})$. In practice, it is not easy in general to define such smooth adapted function on a general domain. Here, we present a modified objective function, conceptually very similar to that presented in [6], but much more convenient from a practical viewpoint. The optimization problems defining the LME approximants considered here take the form

$$\text{For fixed } \boldsymbol{x} \in \text{conv} X, \quad \text{minimize } \sum_{a=1}^{N} \beta_a p_a |\boldsymbol{x}-\boldsymbol{x}_a|^2 + \sum_{a=1}^{N} p_a \ln p_a,$$

$$\text{subject to } p_a \geqslant 0, \quad a=1,\ldots,N,$$

$$\sum_{a=1}^{N} p_a = 1, \quad \sum_{a=1}^{N} p_a \boldsymbol{x}_a = \boldsymbol{x}$$

Here, locality is defined through a set of nodal parameters $\boldsymbol{\beta} = \{\beta_a\}_{a=1,\ldots,N}$ rather than a function. This flavor of the LME approximants shares the main properties previously mentioned. In particular, the methods of [6] allow us to immediately show that the resulting basis functions are C^∞ functions of both $\boldsymbol{x}$ and the discretization parameters, $\boldsymbol{\beta}$ and X. Similarly, the efficient solution of this program follows from standard duality methods. Here, we just summarize the recipe for the final calculation of the basis functions. By analogy with statistical mechanics, we define the partition function

$$Z(\boldsymbol{x}, \lambda) = \sum_{b=1}^{N} \exp[-\beta_b |\boldsymbol{x}-\boldsymbol{x}_b|^2 + \lambda \cdot (\boldsymbol{x}-\boldsymbol{x}_b)].$$

At each evaluation point $\boldsymbol{x}$, the Lagrange multiplier for the linear consistency condition is the unique solution to a solvable, convex, unconstrained optimization problem

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

This optimization problem with d unknowns, where d is the space dimension, is efficiently solved with Newton's method. Then, the basis functions adopt the form

$$p_a(\boldsymbol{x}) = \frac{1}{Z(\boldsymbol{x}, \lambda^*(\boldsymbol{x}))} \exp[-\beta_a |\boldsymbol{x}-\boldsymbol{x}_a|^2 + \lambda^*(\boldsymbol{x}) \cdot (\boldsymbol{x}-\boldsymbol{x}_a)]. \tag{2}$$

The calculation of the spatial gradients of the shape functions can be found in Appendix A.

The smooth transition from widespread meshfree basis functions to linear C^0 basis functions in 1D is illustrated in Figure 2. The transition is controlled by the set of non-negative locality parameters $\boldsymbol{\beta} = \{\beta_a = \gamma_a / h_a^2\}_{a=1,\ldots,N}$, where h_a is the nodal spacing (uniform in this case) and γ_a is a dimensionless parameter that characterizes the degree of locality of the basis function associated with the node $\boldsymbol{x}_a$. It can be noticed that the shape functions become sharper and more local as the value of the dimensionless parameter γ_a increases. For values of γ_a close to 4 and above, the shape functions are nearly indistinguishable from the affine Delaunay basis functions.

The formulation of the LME approximants presented here allows us to control with great flexibility the degree of locality, defined as the extent of the basis functions relative to the typical nodal spacing, through the parameters β_a. Note that as soon as the nodal spacing is non-uniform,
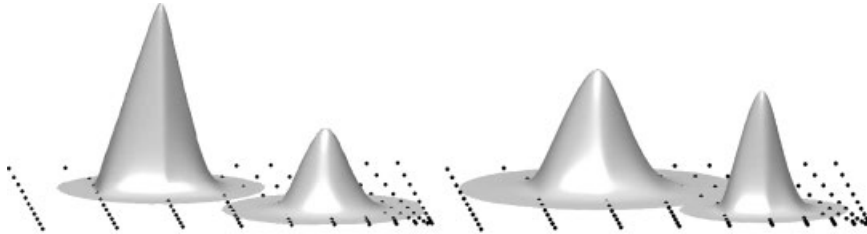
Figure 3. Shape functions for a non-uniform nodal distribution. Non-uniform locality of the basis functions for a constant value of β (left) and basis functions with uniform locality by adapting the nodal parameters β_a to achieve uniform $\gamma_a = \beta_a h_a^2$ (right). Here, h_a denotes the typical nodal spacing close to the ath node.

a spatially varying β is required to define basis functions with uniform value of γ or locality, see Figure 3 for an illustration.

3. VARIATIONAL ADAPTIVITY WITH THE LME APPROXIMANTS

This section outlines the concept of variational adaptivity, and its application to LME approximants. Variational adaptivity is a natural strategy in partial differential equations that can be cast as a minimization problem, such as linear and non-linear elasticity, linear and non-linear diffusion equations, including transient problems appropriately time-discretized, or variational plasticity formulations. In these methods, the energy minimization determines not only the equilibrium discrete solution, but also the optimal node locations or mesh connectivity. When compared with the traditional adaptive methods based on error estimation and remeshing criteria, variational adaptive strategies are particularly well suited for strongly non-linear problems since they do not rely on linearization. In the present setting, one fundamental drawback of adaptivity driven by *a posteriori* error estimation is that current *a priori* error estimates for meshfree methods [11] do not provide any clue on how to adapt the locality of the basis functions once the local error is estimated. Variational adaptivity does not require any remeshing criterion, and the adaption is performed on the sole basis of energy optimality. On the other hand, one limitation of current variational adaptivity techniques is that they optimize the discretization on the basis of a global quantity, whereas most engineering calculations are performed to extract a specific quantity of interest. Error estimation techniques targeted at these quantities have been developed in the recent years [20].

To illustrate the key ideas, we consider non-linear elastostatics. Given the undeformed configuration of an elastic body $\Omega_0 \subset \mathbb{R}^3$ subject to body forces, with prescribed deformation on part of its boundary $\partial\Omega_0^D$ and prescribed tractions in the rest of the boundary of the body $\partial\Omega_0^N$, the goal is to obtain the deformation mapping $\boldsymbol{\varphi} : \Omega_0 \longrightarrow \mathbb{R}^3$ satisfying the following equations (balance of linear momentum and boundary conditions):

$$\text{Div}\,\boldsymbol{P} + \boldsymbol{B} = \boldsymbol{0} \quad \text{in } \Omega_0,$$

$$\boldsymbol{\varphi} = \overline{\boldsymbol{\varphi}} \quad \text{on } \partial\Omega_0^D,$$

$$\boldsymbol{P}\boldsymbol{N} = \overline{\boldsymbol{T}} \quad \text{on } \partial\Omega_0^N.$$

Here Div denotes the divergence in material coordinates, $\boldsymbol{P}$ is the first Piola–Kirchhoff stress tensor, $\boldsymbol{B}$ the body force per unit undeformed volume, $\boldsymbol{N}$ the unit outward normal to the boundary of the undeformed body, $\overline{\boldsymbol{T}}$ is the prescribed traction per unit undeformed area, and $\overline{\boldsymbol{\varphi}}$ is the prescribed deformation mapping in part of the boundary. These equations need to be supplemented by the constitutive relation, which in hyper-elastic materials takes the form

$$\boldsymbol{P}(\boldsymbol{X}) = \frac{\partial W}{\partial \boldsymbol{F}}(D\boldsymbol{\varphi}(\boldsymbol{X})) \quad \forall \boldsymbol{X} \in \Omega_0.$$

Here $W(\boldsymbol{F})$ is the free energy density of the material, whose argument is the deformation gradient, that is the derivative of the deformation mapping $D\boldsymbol{\varphi}$.

The standard FE discretization of this problem relies on the weak form of these equations and the FE interpolation of the deformation mapping and the test functions. The resulting non-linear set of equations can be solved using, for instance, Newton's method. This approach may lead to unstable equilibria, and in many instances, it is useful to recast the non-linear elastostatics problem as a minimization problem. We define the total potential energy of the body and the applied loads as

$$I[\boldsymbol{\varphi}] = \int_{\Omega_0} W(D\boldsymbol{\varphi}) \, \mathrm{d}\Omega - \int_{\Omega_0} \boldsymbol{B} \cdot \boldsymbol{\varphi} \, \mathrm{d}\Omega - \int_{\partial \Omega_0^N} \overline{\boldsymbol{T}} \cdot \boldsymbol{\varphi} \, \mathrm{d}\Gamma.$$

We denote by $\mathscr{C}$ the space of admissible deformation mapping satisfying the essential boundary conditions $\boldsymbol{\varphi} = \overline{\boldsymbol{\varphi}}$ on $\partial \Omega_0^D$. According to the principle of minimal potential energy, the stable equilibrium deformations $\boldsymbol{\varphi}^*$ minimize the total potential energy functional

$$\boldsymbol{\varphi}^* = \arg \inf_{\boldsymbol{\varphi} \in \mathscr{C}} I[\boldsymbol{\varphi}]. \tag{3}$$

It takes a simple calculation to show that the Euler–Lagrange equations of this variational problem are the strong form of the governing equations of non-linear elastostatics. Additionally, the solutions of Equation (3) are stable equilibria.

The Ritz view of Galerkin methods for partial differential equations stemming from minimization principles proceeds as follows. The physical field, here the deformation mapping, is discretized as

$$\boldsymbol{\varphi}_h(\boldsymbol{X}) = \sum_{a=1}^N \boldsymbol{\varphi}_a p_a(\boldsymbol{X}),$$

where $\boldsymbol{\varphi}_a$ denote the nodal values and $p_a(\boldsymbol{X})$ are the basis functions defined in the undeformed body. We denote by $\boldsymbol{\Phi} = (\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots, \boldsymbol{\varphi}_N)$ the array containing all the nodal values. Plugging this expression into the potential energy functional, we obtain the discrete potential energy function $I_h(\boldsymbol{\Phi}) = I[\boldsymbol{\varphi}_h]$, which also involves numerical quadrature. The solution of the discretized problem is simply

$$\boldsymbol{\Phi}^* = \arg \min_{\boldsymbol{\Phi} \in \mathscr{C}_h} I_h(\boldsymbol{\Phi}), \tag{4}$$

where $\mathscr{C}_h$ is the set of nodal values such that $\boldsymbol{\varphi}_h$ is an admissible discrete deformation mapping consistent with the boundary conditions. Since the set of deformation mappings of the discrete scheme is a subset of $\mathscr{C}$, the minimization in Equation (4) can be viewed as a constrained version of the minimization problem in Equation (3), hence yields higher optimal values of the potential energy.

Let us explicitly denote the dependence of the basis functions $p_a(X; P)$ on the set of discretization parameters P, which in FE methods is the location of the nodes and the mesh connectivity, and in the present setting is the location of the nodes and the locality parameters. Then, we can view the discrete potential energy as a function of P too, and seek for minima with respect to both the nodal values of the physical field and the discretization parameters

$$(\mathbf{\Phi}; P)^* = \arg \min_{\mathbf{\Phi} \in \mathscr{C}_h, P \in \mathscr{P}} I_h(\mathbf{\Phi}; P). \tag{5}$$

Here, $\mathscr{P}$ denotes some suitable admissible set of discretization parameters. In this way, the minimum principle yields the equilibrium solutions and also relaxes as much as possible the frustration introduced by the discretization scheme. This approach has received considerable attention in the FE context in the recent years [13, 15, 21, 22]. These references show the robustness and effectivity of variational adaptivity, and highlight its relation with configurational mechanics. Related approaches in structural optimization [14, 23] and elastodynamics [24] have been proposed.

For LME approximants, the set of discretization parameters is $P = (X, \boldsymbol{\beta})$, i.e. the node set and the nodal values of the locality parameter. In the present paper, only variational adaption with respect to $\boldsymbol{\beta}$ is considered, of interest by its own as argued in the introduction. The full optimization of the discretization parameters is the object of current research. Thus, we will consider the shape functions as explicitly depending on the locality parameters $p_a(X; \boldsymbol{\beta})$, and will solve the minimization problem

$$(\mathbf{\Phi}; \boldsymbol{\beta})^* = \arg \min_{\mathbf{\Phi} \in \mathscr{C}_h, \boldsymbol{\beta} \in \mathscr{P}} I_h(\mathbf{\Phi}; \boldsymbol{\beta}). \tag{6}$$

The nature of the space of admissible locality parameters $\mathscr{P}$ is discussed in the following section. This minimization equilibrates the standard forces conjugate to the physical fields, as well as the configurational forces conjugate to the locality of the shape functions. By the chain rule, these forces, required in the numerical optimization by gradient methods, at the bth node, can be written as

$$\begin{aligned}
\frac{\partial I_h}{\partial \boldsymbol{\varphi}_b} &= \int_{\Omega_0} \boldsymbol{P} \nabla_0 p_b \, d\Omega - \int_{\Omega_0} p_b \boldsymbol{B} \, d\Omega - \int_{\partial \Omega_0^N} p_b \overline{\boldsymbol{T}} \, d\Gamma, \\
\frac{\partial I_h}{\partial \beta_b} &= \int_{\Omega_0} \sum_{a=1}^{N} \left(\boldsymbol{\varphi}_a^{\mathrm{T}} \boldsymbol{P} \frac{\partial (\nabla_0 p_a)}{\partial \beta_b} \right) d\Omega - \int_{\Omega_0} \boldsymbol{B} \cdot \left(\sum_{a=1}^{N} \frac{\partial p_a}{\partial \beta_b} \boldsymbol{\varphi}_a \right) d\Omega \\
&\quad - \int_{\partial \Omega_0^N} \overline{\boldsymbol{T}} \cdot \left(\sum_{a=1}^{N} \frac{\partial p_a}{\partial \beta_b} \boldsymbol{\varphi}_a \right) d\Gamma.
\end{aligned} \tag{7}$$

Here, ∇_0 is the gradient with respect to the material coordinates of the undeformed body, where the basis functions are defined. Although in these expressions the sums run over all the nodes, in practice only the nodes neighboring the bth node contribute to the sums. A key technical fact behind the proposed method is that the sensitivities of the shape functions with respect to the locality parameters can be readily computed for LME approximants, as shown in Appendix A.

The problem presented in Equation (5) is, in general, highly non-convex and difficult. Figure 4 illustrates this fact. The energy can be thought of as dependent on two sets of variables, one associated with the discretization of the physical field $\mathbf{\Phi}$ and the other associated with the discretization parameters $P = (X, \boldsymbol{\beta})$. For fixed discretization parameters, the minimization problem with respect
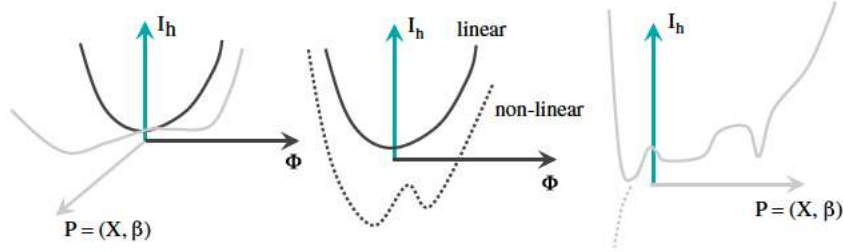
Figure 4. Energy I_h thought as dependent on two sets of variables, one representing the physical field Φ and the other related to the discretization parameters $P = (X, \beta)$.

to the physical nodal values can be a nice quadratic function (linear problems) or it can be non-convex for non-linear problems, such as finite elasticity close to buckling. In any case, the complexity of this minimization problem is directly related to the physics of the problem at hand. However, the minimization with respect to the discretization parameters is massively non-convex in general. Consider, for instance, a boundary value problem with an affine solution in part of the domain. In this region, the energy is independent of the discretization parameters, since all LME schemes satisfy the linear consistency conditions. It is only the features of the solution that endow the energy landscape in the right-hand side of Figure 4 with (possibly multiple) minimizers. In practice, the energy as a function of the discretization parameters may even present spurious minima or even become unbounded from below, for instance if nodes move significantly and the quadrature is not adapted accordingly. For these reasons, optimization with respect to P must be performed carefully and conservatively.

4. NUMERICAL EXAMPLES

The numerical strategy to solve the problem posed in Equation (6) is described in this section. Next, the numerical performance of the locality variational adaption methodology is illustrated by way of six numerical examples. The first two are Poisson problems, the third is a linear elasticity problem with a non-convex domain, and the other three are non-linear elasticity problems with a compressible neo-Hookean material.

4.1. Numerical strategy

Two sets of equations have to be solved in the problem posed in Equation (6), namely the physical and the configurational equilibria. These equations are coupled in that the respective unknowns appear in both sets of equations. Different numerical strategies can be followed to approach this problem. One possibility is to solve simultaneously the two sets of equations in a monolithic scheme. It is also possible to solve the coupled system with a staggered scheme, see [22] for a detailed explanation. This reference suggests that monolithic schemes are more effective, on the basis of simple low-dimensional tests.

Given the radical difference of the expected energy landscapes in each of the two sets of variables discussed before, we avoid solving for both types of unknowns at once. Yet, we consider a nested algorithm that tightly couples the physical and the discretization parameters. The idea is, for a fixed set of discretization parameters, to find the physical equilibrium; hence, formally view Φ as

a function of $\boldsymbol{\beta}$:

$$\boldsymbol{\Phi}^*(\boldsymbol{\beta}) = \arg\min_{\boldsymbol{\beta} \in \mathscr{P}} I_h(\boldsymbol{\Phi}; \boldsymbol{\beta}).$$

Numerically, this involves the solution of a system of linear equations for Poisson and linear elasticity problems, or a non-linear optimization problem for non-linear elasticity. For the latter, we use a limited memory BFGS algorithm [25]. Algorithmically, this can be understood as an inner loop. This allows us to formally define a discrete energy function depending only on the locality parameters

$$\widehat{I_h}(\boldsymbol{\beta}) = I_h(\boldsymbol{\Phi}^*(\boldsymbol{\beta}); \boldsymbol{\beta}),$$

which is optimized in an outer loop. In this optimization, the set of admissible locality parameters $\mathscr{P}$ is defined as follows. As illustrated before, values of the non-dimensional locality γ higher than 4 lead to basis functions very close to the Delaunay limit. For this reason, we limit the maximum value of β_a to $4/h_a^2$, where h_a is the typical nodal spacing around node a. On the other hand, low values of the locality parameter lead to widespread shape functions that can increase the computational cost of the solution considerably, by enlarging the band-width of the system and by requiring more quadrature points for accurate results. For this reason, we impose the constraints $\beta_a \geqslant \gamma_{\min}/h_a^2$, where depending on the application $\gamma_{\min}$ takes values around 0.8. Thus, we select $\mathscr{P} = [\gamma_{\min}/h_1^2, 4/h_1^2] \times [\gamma_{\min}/h_2^2, 4/h_2^2] \times \cdots \times [\gamma_{\min}/h_N^2, 4/h_N^2]$. The numerical optimization of $\widehat{I_h}(\boldsymbol{\beta})$ within $\mathscr{P}$ is carried out with a limited memory BFGS implementation for bound-constrained problems [26].

The algorithm can be summarized as follows for a non-linear problem:

1. Set $\boldsymbol{\beta}^0 = \boldsymbol{\beta}^{\text{guess}}$ and $n = 0$
2. Set $\boldsymbol{\beta}_{\min}$ and $\boldsymbol{\beta}_{\max}$
3. Outer loop: L-BFGS-B algorithm (iterations run over n)
   (a) Inner loop: L-BFGS

      (i) $\boldsymbol{\Phi}^n = \boldsymbol{\Phi}^*(\boldsymbol{\beta}^n)$
      (ii) $I_h^n = I_h(\boldsymbol{\Phi}^n; \boldsymbol{\beta}^n)$
      (iii) $\frac{\partial I_h}{\partial \boldsymbol{\beta}}(\boldsymbol{\Phi}^n; \boldsymbol{\beta}^n)$

   (b) Exit if $\left|\boldsymbol{\beta}^n - \boldsymbol{\beta}^{n-1}\right| \leqslant Tol_\beta$ and $\left|I_h^n - I_h^{n-1}\right| \leqslant Tol_I$ and $\left|\frac{\partial I_h}{\partial \boldsymbol{\beta}}\right| \leqslant Tol_{DerI}$
   (c) Update to $\boldsymbol{\beta}^{n+1}$
   (d) Reset n to $n+1$, and go to 3.

For a linear problem, the inner loop is replaced by a solver for linear systems of equations. Despite the robust and efficient performance of the proposed algorithm, we believe that there is a room for improvement in the numerical optimization strategy for these kinds of variational adaptivity programs.

4.2. Example 1: 2D heat conduction with Gaussian solution

To illustrate the proposed method, we consider a Poisson boundary value problem

$$-\nabla \cdot (k\nabla u) = s \quad \text{in } \Omega = [0, 1] \times [0, 1],$$
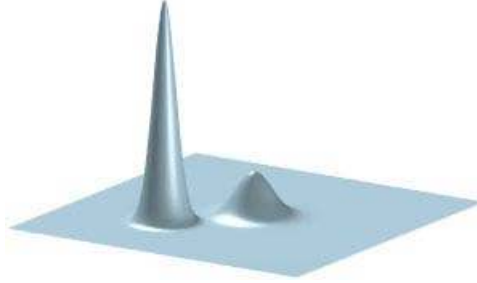
$$u = \bar{u} \quad \text{in } \partial\Omega.$$

Figure 5. Analytical solution of Example 1.

We consider in this example $k = 1$, and chose the source s and the boundary data $\bar{u}$ such that the exact solution has the following expression (illustrated in Figure 5):

$$u(x, y) = 10\,\mathrm{e}^{-180[(x-0.51)^2+(y-0.52)^2]} + 50\,\mathrm{e}^{-450[(x-0.31)^2+(y-0.34)^2]}$$

The strong form of the boundary value problem arises as the Euler–Lagrange equations of the minimization of the functional

$$I[u] = \frac{1}{2} \int_\Omega k\,|\nabla u|^2 \,\mathrm{d}\Omega - \int_\Omega us\,\mathrm{d}\Omega$$

for admissible functions u satisfying the essential boundary conditions. Thus, an energy minimization problem equivalent to that of Equation (6) can be posed, and similar forces to those of Equation (7) can be computed.

The relative energy error E_{er} and the L_2 error are plotted in Figure 6 for different fixed support sizes of the shape functions and for optimized locality parameters, as well as for different levels of uniform refinement. The relative energy error is defined as

$$E_{\mathrm{er}} = \frac{|I_{\mathrm{exact}} - I_h|}{|I_{\mathrm{exact}}|}$$

where I_{exact} is the exact value of the energy, and I_h is the numerical value of the energy at the optimal physical parameters and possibly the optimal locality parameters.

Figure 6 (left) shows that, for this example, the accuracy strongly depends on the choice of locality parameters, and that the lower the value of γ, i.e. the more widespread the basis functions are, the more accurate the solution is in terms of the total energy error. As it has been mentioned before, this comes at a cost, since for $\gamma = 0.8$, the band-width of the system is larger and more quadrature points are needed. In this example, we consider this as the lowest acceptable value from an efficiency viewpoint. The figure also shows that numerical solutions with the optimized locality parameters constrained to $0.8/h^2 \leqslant \beta_a \leqslant 4/h^2$ exhibit more accurate values of the total potential energy, particularly for coarse discretizations. Figure 6 (right) shows the corresponding convergence plots in the L_2 norm, which is not the target quantity of the variational adaptivity method. Still, it can be observed that the adaptive strategy performs better than the most accurate solution with uniform $\gamma = 0.8$. For instance, the number of nodes required by the non-adaptive LME scheme with $\gamma = 0.8$ to achieve an L_2 error of 10^{-3} is twice the number of nodes needed with the adaptive strategy, and at least 10 times higher when $\gamma \geqslant 1.6$.
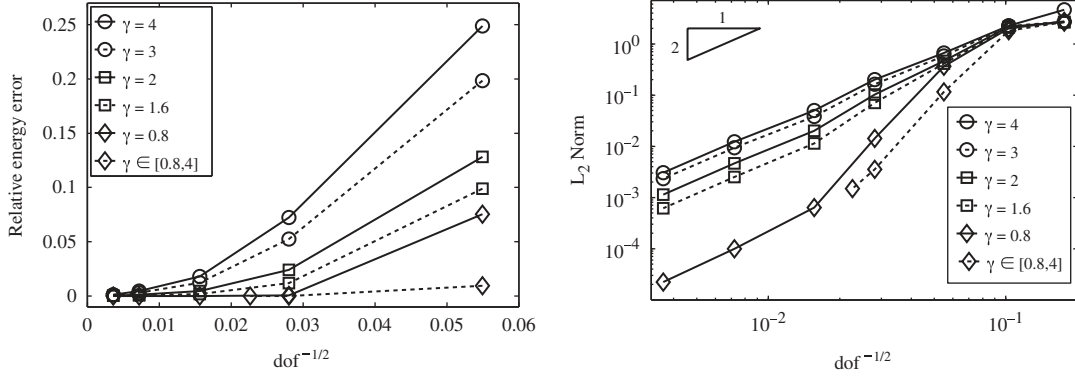
Figure 6. Relative energy error (left) and L_2 error (right) for Example 1. The variational adaptive scheme with $\gamma_a \in [0.8, 4]$ improves the accuracy of the solution as compared with non-adaptive local *max-ent* approximations with uniform γ_a.
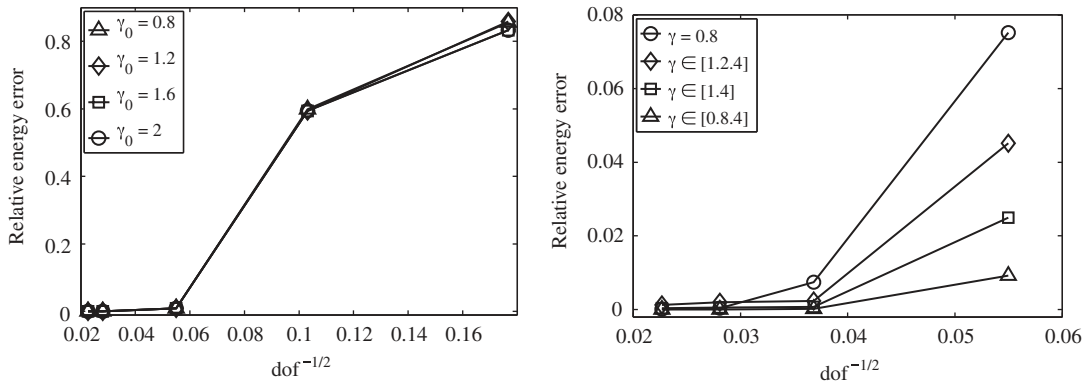

Figure 7. Relative energy error dependence on the initial guess for γ_0 (left) and the admissible interval (right) for Example 1.

Figure 7 (left) shows the relative energy error for the variational adaptive strategy with different values of the initial guess for the locality γ_0 in the numerical optimization. It is interesting to appreciate that the result of the energy I_h is essentially independent on this initial guess, which suggests that the optimization is not trapped in the local minima. The dependence of the relative energy error on the admissible interval for the support size is illustrated in Figure 7 (right). It can be noticed that the dependence is more important when coarse meshes are used, and as expected, the smaller the interval, the less efficient the optimization is. The interval $\gamma \in [0.8, 4]$ is used in the remainder of the paper.

We now analyze the optimal locality distributions given by the algorithm. For this, it is convenient to define for each node a the numerically effective support size R_a of the corresponding shape function, related to the locality parameter β_a through the equation $R_a = \sqrt{(-\log(\text{Tol}_0)/\beta_a)}$. This formula comes from equating Equation (2) to Tol_0, and assuming that Z and λ^* are bounded,
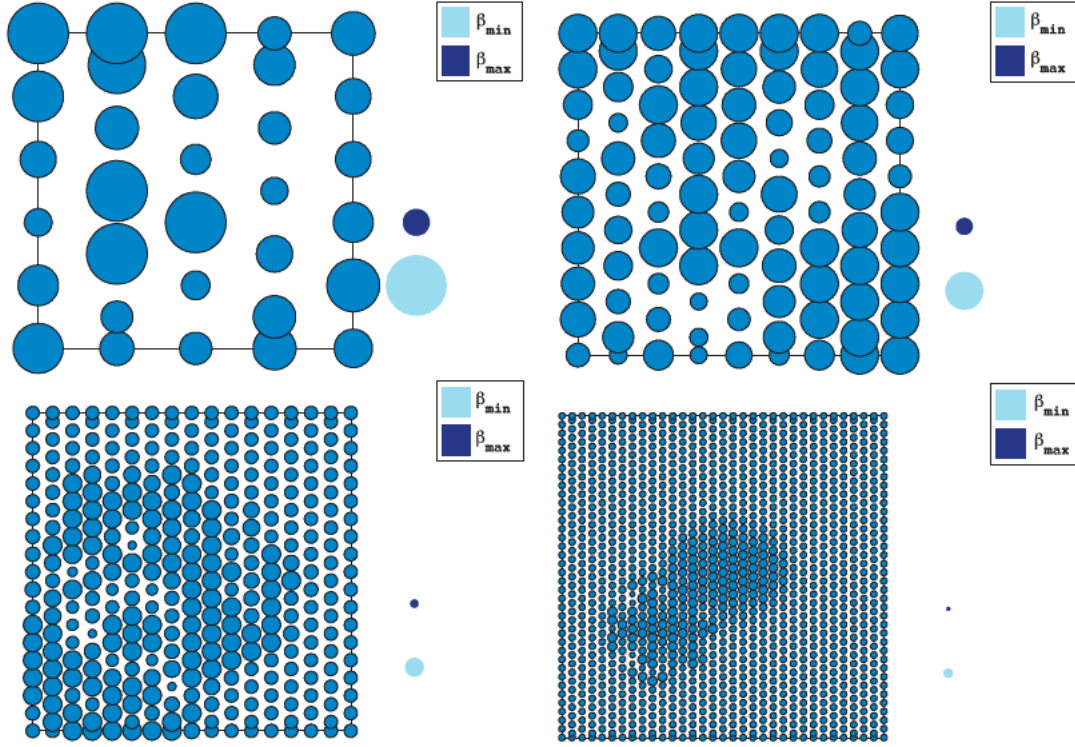
Figure 8. Optimum support size (represented as disks centered at each node and proportional to R_a, see text) for different spatial discretizations used to solve Example 1. The maximum and minimum radii of the disks, corresponding to $\gamma = 0.8$ and $\gamma = 4$, respectively, are also shown for illustration.

nearly constant functions of position, which is true at least for uniform node distributions. The numerical support can be interpreted as a ball of radius R_a centered at the nodal position x_a, outside of which the shape functions are zero within a tolerance Tol_0; a value of $\mathrm{Tol}_0 = 10^{-8}$ is adopted here.

Figure 8 shows the optimum support size for different levels of refinement. For fine node sets, the features of the exact solution (see Figure 5) can be identified in the optimum support size map. Note that in this and subsequent figures, the radius of the circle centered at x_a and represented in the plots is proportional to R_a. Otherwise, the considerable overlapping of the supports would make the visualization meaningless. This figure highlights the strong dependence of the optimal locality pattern on the level of refinement, which stresses the difficulty of using heuristic rules to adapt the locality parameters in meshfree methods. The figure also shows that the optimal distributions hit the boundary of the feasible region $\beta_a = 0.8/h^2$ only in some regions of the domain. Hence, the adaptive strategy grants very accurate solutions, while limiting the extra computational cost of widespread shape functions to specific regions, where these are really needed. This method is particularly effective in combination with a quadrature scheme adapted to the value of the locality parameters, with large densities of quadrature points only where low values of γ are present.
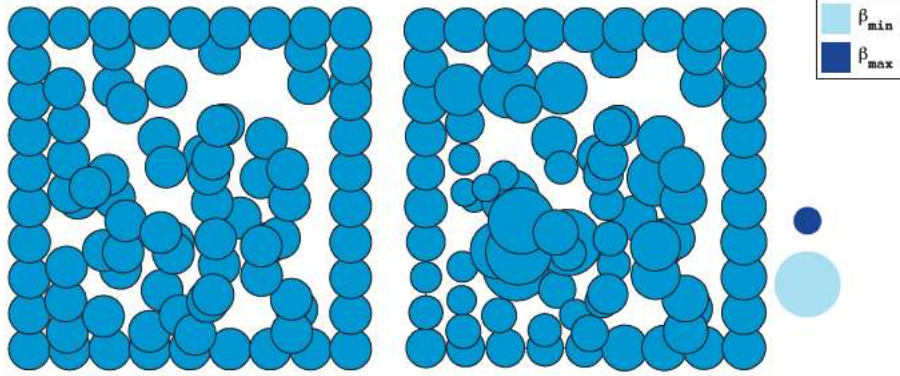
Figure 9. Initial nominal support size representation for constant $\{\beta_a\}_{a=1,...,N}$ (left), and optimized nominal support size representation (right) for Example 1.

We next explore utility of the proposed adaptive strategy to identify good values of β for highly non-uniform node sets. For non-uniform node sets, as illustrated before, a uniform value of β_a leads to basis functions of variable locality. In practice, it is important to control the locality of the shape functions, since it has a direct impact on the smoothness of the approximation or the required quadrature scheme. If the node set is very irregular, it becomes difficult to define a meaningful nodal spacing, and consequently it is difficult to identify reasonable values for β. We have considered a highly non-uniform node set with an initially uniform value of β_a, and have optimized the locality with the adaptive scheme. The results are shown in Figure 9. In this example, an improvement greater than 40% in the total energy is achieved.

4.3. Example 2: 2D heat conduction with discontinuous conductivity

This examples studies a Poisson problem in a domain $\Omega = [0, 1] \times [0, 1]$ made out of two materials with conductivities k_1 and k_2, respectively. The interface is located at $x = x_I$. The difference between the conductivities generates a discontinuity at the interface in the normal derivative of the solution. The analytical solution, illustrated in Figure 10 with a black line to highlight the interface, is given by

$$u(x, y) = 100x(x-1)\cos(5x)y(y-1)(y^3 - 0.2)[1 + mH(x-x_I)(x-x_I)],$$

where $H(x-x_I)$ is the Heaviside step function with a discontinuity at x_I, $m = (k_1/k_2 - 1)$ $[(2x_I - 1)/(x_I(x_I - 1)) - 5\tan(5x_I)]$, $k_1/k_2 = 10$, and $x_I = 0.7$.

This kind of problem, with discontinuities in the normal derivatives to internal interfaces, is naturally treated by C^0 FEs as long as element faces cover the interface. Meshfree methods based on smooth basis functions, on the contrary, do not reproduce this feature of the exact solution, and exhibit in general sub-optimal convergence rates. With the LME approximants, it is easy to overcome this difficulty by increasing the parameter γ to reach the Delaunay limit. This can be done throughout the domain or locally around the interface, with a strip of nodes with high values of γ. Note that for this to work, the Delaunay triangulation of the node set must conform with the discontinuity. The node set we have considered has this property.
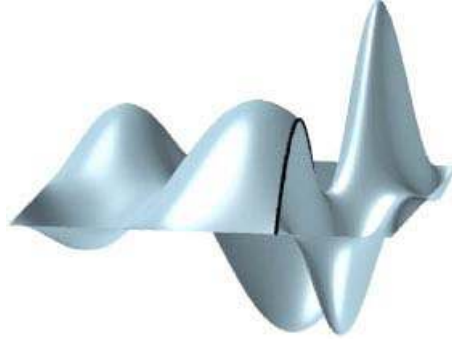
Figure 10. Analytical solution of Example 2. A black line highlights the interface between the two materials.
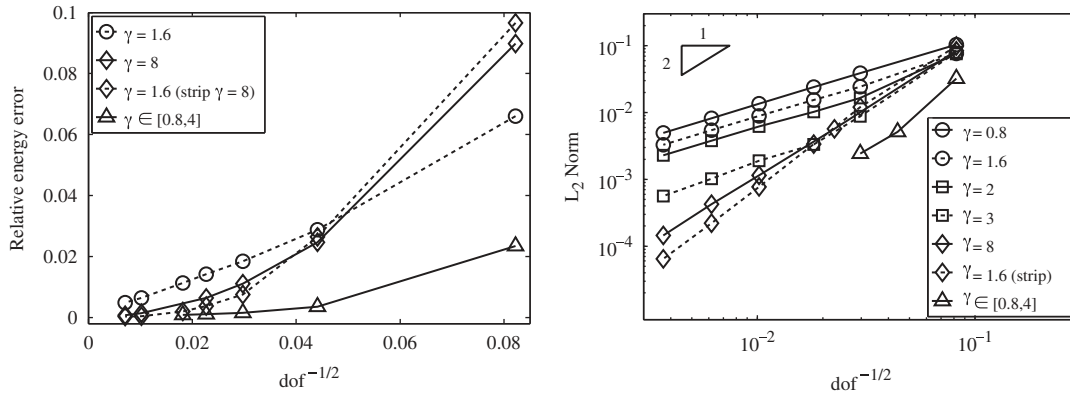

Figure 11. Relative energy error (left) and L_2 error (right) for Example 2.

The behavior of the relative energy error as a function of refinement for this example is shown in Figure 11 (left). It is apparent from this figure that a uniform refinement with smooth basis functions of uniform locality ($\gamma = 1.6$) exhibits much slower convergence than the Delaunay FE solution (uniform $\gamma = 8$), or than smooth basis functions with a local transition to the Delaunay shape functions around the interface ($\gamma_a = 1.6$ with strip of nodes with $\gamma_a = 8$). The left plot in this figure also shows the much superior performance of the proposed adaptive method for this example, with almost converged energies for very coarse discretizations. The convergence of L_2 error norm is illustrated in Figure 11 (right). It can be observed that when smooth shape functions are used (values of $\gamma \leqslant 3$), the optimal rate of convergence is lost, as expected. This rate of convergence is recovered when the shape functions close to the interface reach the Delaunay limit ($\gamma = 8$ in a strip or everywhere). The remarkable accuracy of the adaptive strategy is illustrated in this L_2 error plot.

We now turn to the locality patterns given by the variational adaptive method. Figure 12 (left) qualitatively illustrates the extent of the support of the basis functions when a strip of high values of γ is placed around the interface, with a value of $\gamma = 1.6$ in the rest of the domain where a smooth solution is expected. This is a good educated guess for the nodal locality parameters, as argued
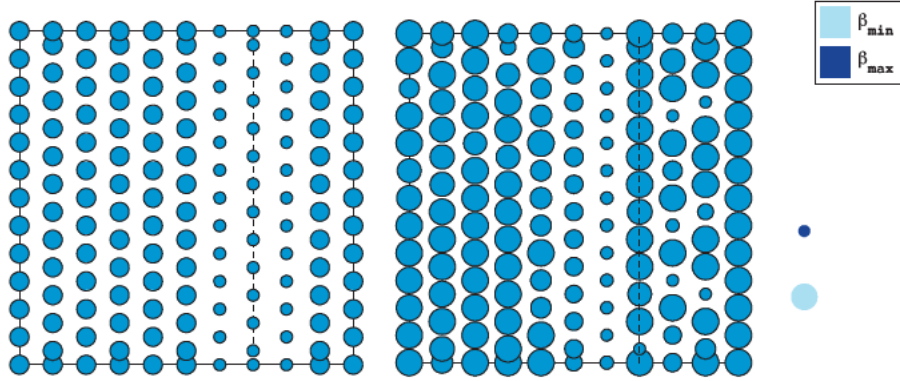
Figure 12. Support size (represented as disks centered at each node and proportional to R_a, see text): locality parameter locally set to $\gamma = 8$ in the region of the discontinuity and $\gamma = 1.6$ in the rest of the domain (left), and optimum support size (right) for a given node set and Example 2.
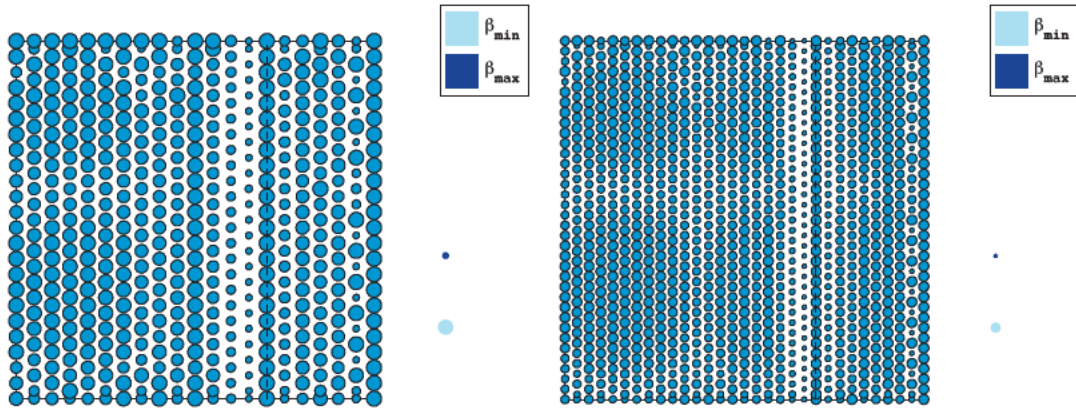

Figure 13. Optimum support size (represented as disks centered at each node and proportional to R_a, see text) for several levels of refinement and Example 2.

in the previous paragraph. The right plot shows the optimal solution provided by the adaptive strategy, which divides the energy error of the solution in the left by four. The adaptive method produces a very different pattern near the discontinuity. The nodes located on it have small values of the nodal locality parameter β_a, hence large nominal support sizes, whereas adjacent nodes on the left of the interface exhibit very local basis functions. Interestingly, this high contrast pattern in the vicinity of the interface is observed as well in the optimal locality parameters for finer node sets (see Figure 13). For these finer node sets, the features away from the interface can also be noticed from the locality pattern.

To investigate the nature of the adapted approximation scheme, in Figure 14 we plot the basis function for a node located on the interface. Interestingly, the locality parameters are optimized to capture the discontinuity of the gradient in the direction normal to the interface, but the smoothness of the basis function is not spoiled in the direction of the interface. This observation may explain the superior accuracy of the adaptive solution. To further illustrate this fact, Figure 15 shows
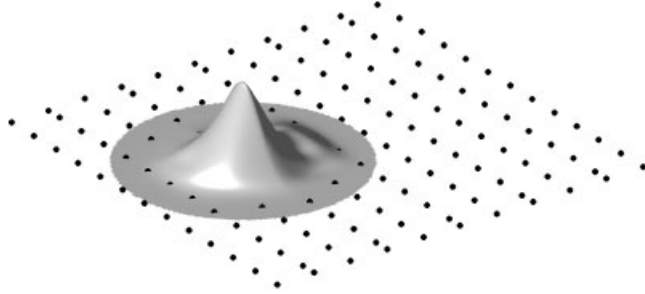
17

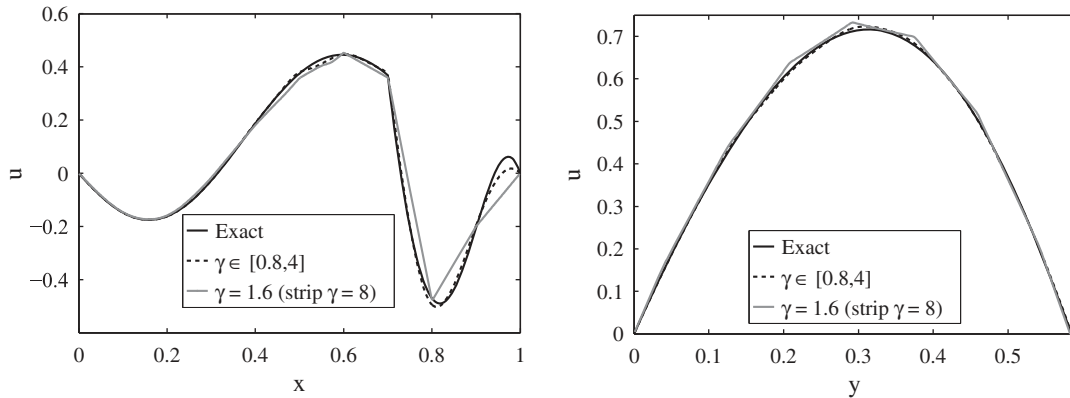Figure 14. Shape function for a node located on the interface (Example 2).

Figure 15. Solution of Example 2 on a line perpendicular to the interface (left) and on the interface (right).

cross-sections of the exact (solid black line), the adapted (dashed line), and the numerical solution for fixed locality parameters with a high locality strip near the interface (solid gray line), in a line perpendicular to the interface (left) and on the interface (right). It can be observed that the adaptive solution introduces the discontinuity in the gradient precisely where needed, whereas the numerical solution with *ad hoc* adapted locality parameters introduces too many discontinuities associated with the Delaunay faces in the interface region, and results in a poorer approximation. It should be emphasized that strictly speaking, these are not discontinuities in the gradients, but rather sharp changes in the gradients. In order to describe mathematical discontinuities on the gradients of local *max-ent* shape functions, the Delaunay limit much be reached, i.e. $\gamma = +\infty$.

4.4. Example 3: infinite plate with a hole

In this example we test the behavior of the proposed methodology in the standard benchmark problem of an infinite plate with a hole subject to a far-field uniaxial traction σ_0 in the x direction (see Figure 16). The exact solution of this linear elasticity boundary value problem can be looked up in reference [27]. By the symmetry of the problem, only one quarter of the plate needs to be considered in the computation. The exact tractions are applied at the boundary of the non-convex numerical domain (gray area in Figure 16). A Young's Modulus of $E = 1$ and a Poisson ratio of $v = 0.495$ are considered.
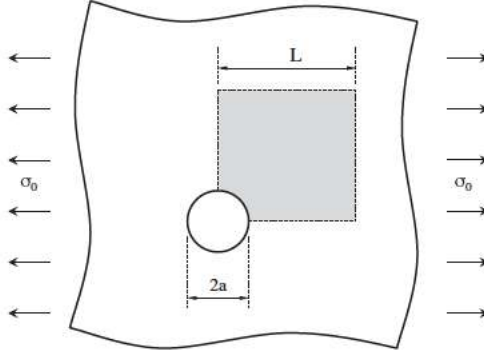
Figure 16. Geometry of the problem of an infinite plate with a hole of radius $a=1$ subject to a far-field uniaxial traction σ_0 in the x direction (Example 3). The gray color represents the non-convex domain analyzed in the numerical simulations, where $L=5$.

Table I. I_h^{FE}, I_h^{LME} and I_h^{VA} are the energies of the non-adaptive FE and LME schemes, and the VA scheme, respectively, for Example 3.

# nodes	I_h^{FE}	I_h^{LME}	I_h^{VA}	$\left\|1-\dfrac{I_h^{\mathrm{FE}}}{I_h^{\mathrm{VA}}}\right\|$ (%)	$\left\|1-\dfrac{I_h^{\mathrm{LME}}}{I_h^{\mathrm{VA}}}\right\|$ (%)
112	−9.5124	−9.5544	−9.5693	0.59	0.16
189	−9.6352	−9.6615	−9.6752	0.41	0.14
286	−9.6777	−9.7444	−9.7699	0.94	0.26
403	−9.7085	−9.7726	−9.8111	1.05	0.39
1288	−9.7836	−9.8120			
2059	−9.8106				

Table I shows that for a number of nodes equal to 403, the variational adaptive LME scheme decreases the energy by 1.05 and 0.39% as compared with non-adaptive FE and LME schemes (uniform $\gamma=1.6$). That small percentage of reduction in the computed total energy in comparison with the LME method is not surprising. Indeed, the exact solution is relatively easy to represent numerically, compared with the previous problems with more localized features. Note that the change of the computed total energy between a coarse mesh (112 nodes) and a fine mesh (50 193 nodes) is less than 3% with the LME scheme ($\gamma=1.6$). Nevertheless, the benefits of the adaptive scheme can be appreciated in that the number of nodes required by non-adaptive FE and LME schemes to get a similar energy is 5 and 3 times higher, respectively.

The optimum support size of each basis function is illustrated in Figure 17 for two levels of refinement (112 and 403 nodes). In this example and for coarse meshes, the degree of locality of the shape functions on the hole boundary is constrained to avoid unwanted correlations between far apart nodes across the non-convex part of the domain [28]. The figure shows that the variational adaptive method adjusts the support size in the vicinity of the hole, where the solution presents more features, only for coarse meshes. We observe that for finer meshes, the optimal support is nearly uniform, and close to the largest allowed support size. This is expected for an over-resolved node distribution, and also explains the marginal gain in terms of the global energy in this example.
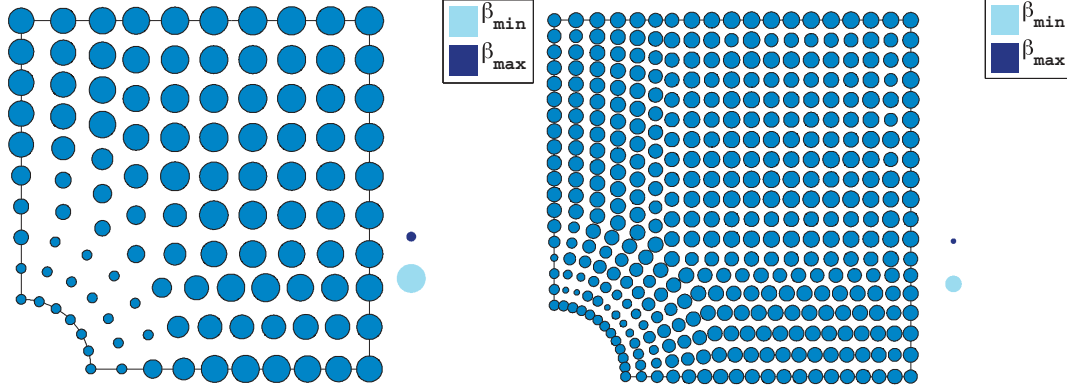
Figure 17. Optimum support size (represented as disks centered at each node and proportional to R_a, see text) for two levels of refinement and Example 3.

4.5. Example 4: stretching of a neo-Hookean hyperelastic slab

We now consider a slab given by $\Omega_0 = [0, 1] \times [0, 1] \times [0, 0.25]$ clamped on two opposite sides and subject to prescribed extensional deformation. The nominal stretch ratio is 1.5 and the entire deformation is applied in one step. The slab is made out of a compressible neo-Hookean material with strain–energy density

$$W(\boldsymbol{F}) = \tfrac{1}{2}\lambda \ln^2(J) + \tfrac{1}{2}\mu \operatorname{tr}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}) - \tfrac{3}{2}\mu - \mu \ln(J)$$

where $J = \det(\boldsymbol{F})$, and λ and μ are the Lamé constants. In calculations these constants are set to $\lambda = 12115.38$ and $\mu = 8071.92$, which corresponds to an initial Poisson ratio of $\nu = 0.3$. A similar problem was analyzed in [15].

The undeformed and deformed configurations of the slab are shown in Figure 18. The radius of the spheres illustrated in the reference configuration is proportional to the optimum support size of the shape functions. In Table II, it can be noticed that variational adaptive LME scheme (VA) present a 3.05 and 0.81% of energy reduction in comparison with non-adaptive FEs and non-adaptive LME schemes (uniform $\gamma = 1.6$), respectively. Although these figures seem to be a little disappointing, they are similar to other results obtained with the mesh-based variational adaption schemes [15, 29] for this simple example.

4.6. Example 5: upsetting of a neo-Hookean hyperelastic block

A hyperelastic block clamped on two opposite sides and subjected to prescribed compressive deformations is studied in this example. The nominal stretch ratio is 0.5 and the entire deformation is applied incrementally in 10 steps. The material is compressible neo-Hookean with a relation of constants $\lambda/\mu = 10$, corresponding to an initial Poisson ratio of $\nu = 0.4545$. Although the dimensions of the block are $[0, 16] \times [0, 16] \times [0, 16]$, only an eighth is analyzed by the symmetry of the problem. A similar problem was presented in [6].

The undeformed and deformed configurations of the block are shown in Figure 19. As in the previous example, the radius of the spheres shown in the reference configuration is proportional to the optimum support size of the shape functions. It can be observed that the optimum locality parameters are nearly uniform throughout the domain, with the exception of the nodes in the
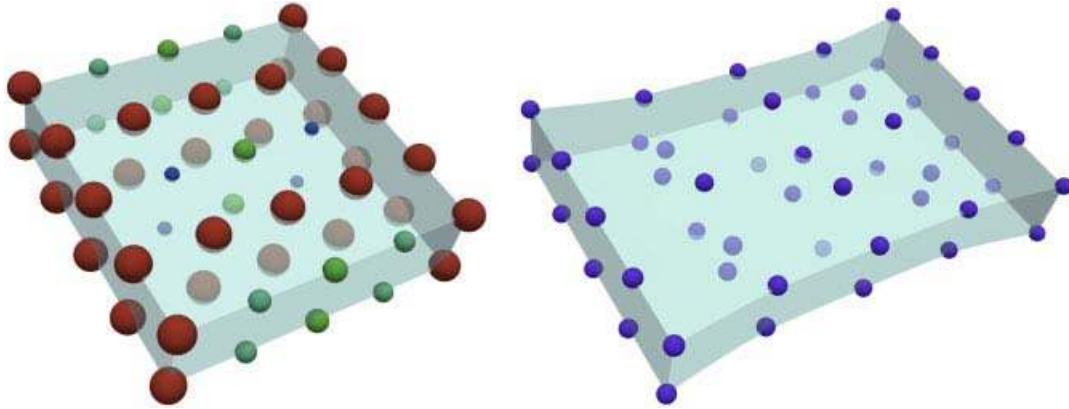
Figure 18. Stretching of a hyperelastic slab (Example 4): reference (left) and deformed (right) configurations. The optimum support size for each shape function is shown in the reference configuration, by spheres of radius proportional to R_a, see text.

Table II. I_h^{FE}, I_h^{LME}, and I_h^{VA} are the energies of the non-adaptive FE and LME schemes, and the VA scheme, respectively, for Example 4.

# nodes	I_h^{FE}	I_h^{LME}	I_h^{VA}	$\left\|1 - \dfrac{I_h^{\text{FE}}}{I_h^{\text{VA}}}\right\|$ (%)	$\left\|1 - \dfrac{I_h^{\text{LME}}}{I_h^{\text{VA}}}\right\|$ (%)
50	577.859	565.253	560.729	3.05	0.81

vicinity of the edge where the largest deformations occur. As in Example 2, a large contrast of nominal supports sizes can be observed in this region. The quadrature effort can be concentrated in the vicinity of these edges.

Table III shows that the variational adaptive LME scheme presents a 12.35 and 3.85% of energy reduction in comparison with non-adaptive FE and LME schemes (uniform $\gamma = 1.6$), respectively, when equal number of nodes is used. It can be also observed that the number of nodes required by non-adaptive FE and LME methods to get a similar energy is 230 and 15 times higher, respectively. Furthermore, this accuracy is achieved with a moderate computational cost. Figure 20 illustrates the convergence of iterative method for the optimization of $\boldsymbol{\beta}$ (outer loop), for a node set with 1147 nodes. Note that more than 90% of the energy reduction is achieved in less than 10 iterations. Thus, with a few iterations on the locality parameters, very accurate solutions can be obtained with coarse discretizations.

4.7. Example 6: buckling of a neo-Hookean hyperelastic beam

Here, we compress of a hyperelastic slender body $\Omega_0 = [0, 1] \times [0, 1] \times [0, 10]$. The specimen is clamped on two opposite sides and is subject to prescribed compressive deformations. The nominal stretch ratio is 0.35 and the entire deformation is applied incrementally in 20 steps. The material parameters are those of the previous example.

The undeformed and deformed configurations of the body, together with the optimum support sizes, are shown in Figure 21. In this example, it can be observed that for the optimum discretization,
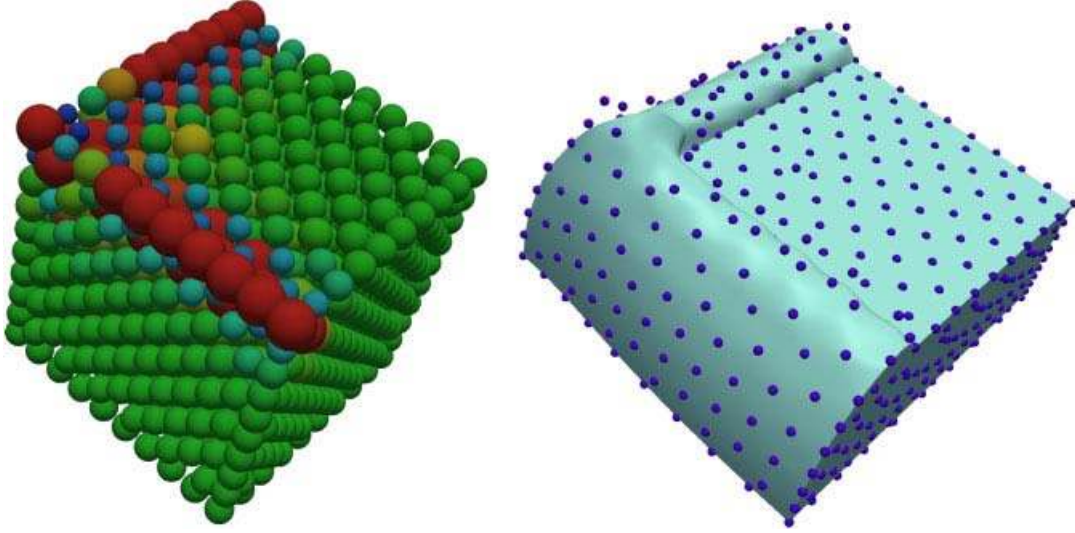
Figure 19. Compression of a hyperelastic block (Example 5): reference (left) and deformed (right) configurations. The optimum support size for each shape function is shown in the reference configuration, by spheres of radius proportional to R_a, see text.

Table III. I_h^{FE}, I_h^{LME}, and I_h^{VA} are the energies of the non-adaptive FE and LME schemes, and the VA scheme, respectively, for Example 5.

# nodes	I_h^{FE}	I_h^{LME}	I_h^{VA}	$\left\|1 - \frac{I_h^{FE}}{I_h^{VA}}\right\|$ (%)	$\left\|1 - \frac{I_h^{LME}}{I_h^{VA}}\right\|$ (%)
401	357.810	322.729	307.311	16.43	5.02
1147	341.262	315.427	303.743	12.35	3.85
18 065	319.135	303.180			
265 937	303.457				

most shape functions at the free surface of the body exhibit large support sizes, whereas the inner shape functions adopt variable support sizes depending on the features of the deformation. Table IV shows that the variational adaptive LME scheme decreases the energy by 76.3 and 5.3% as compared with non-adaptive FE and LME schemes (uniform $\gamma = 1.6$), respectively, for an equal number of nodes. It can be also observed that the number of nodes required by non-adaptive FE and LME to get a similar energy is 178 and 6 times higher, respectively.

5. CONCLUSIONS

We have proposed a variational adaptive approach to optimize the support size of meshfree shape functions in the numerical approximation of boundary value problems stemming from a minimum principle. The flexibility in scattering nodes in the domain and assigning different support sizes
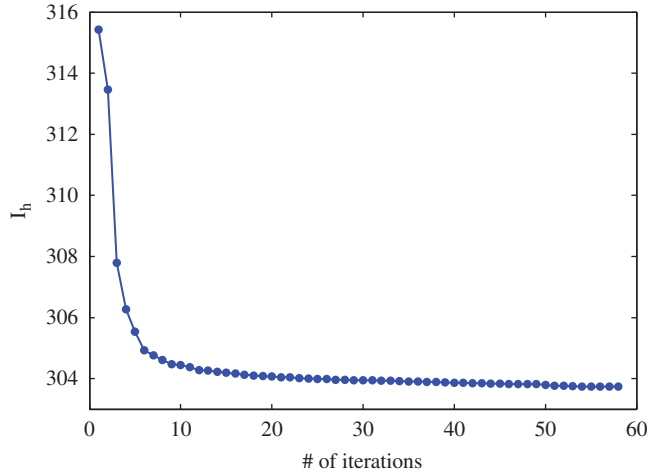
Figure 20. Energy vs number of iterations in $\boldsymbol{\beta}$ for Example 5.

of the shape functions has often been presented as an attractive feature of meshfree methods. However, the accuracy of the numerical approximation to partial differential equations strongly depends on the choice of support size, referred to here as locality, and the current theories (*a priori* estimates) do not help in selecting an adequate value, which is found to depend on the problem at hand and the level of refinement. We have presented a rational method to select the locality of the basis functions, which optimizes the total potential energy functions not only with respect to the nodal physical values, but also with respect to the support sizes. The resulting method has been shown to be very effective in achieving very accurate solutions with very coarse node sets in the Poisson and non-linear elasticity problems. This accuracy has been observed not only in terms of the target quantity of the method, i.e. the total potential energy, but also in terms of the L_2 norm. The variational adaptivity of the support size has been presented in the context of LME approximation schemes, similar in some respect to MLS approximants. The former are particularly well suited for our purposes because their evaluation is robust when the locality parameters are varied arbitrarily, and because the derivatives of the basis functions with respect to the locality parameters can be readily computed in closed form.

For problems with smooth solutions, meshfree basis functions with larger support sizes typically lead to more accurate solutions, at the expense of extra computational cost due to the larger bandwidth of the stiffness matrix and the requirements in the numerical quadrature. The optimized locality parameters allow us to use such widespread shape functions only in some areas of the domain; hence, concentrating the computational effort where really needed. The produced patterns in the locality distribution are difficult to guess *a priori*, particularly in problems involving material interfaces, hence discontinuities in the normal derivative to the interface.

Finally, we note that in the present work, we have considered a scalar locality parameter β_a. However, as shown in [6], it is possible to consider a locality metric tensor (symmetric, positive-definite) $\boldsymbol{\beta}_a$ at each node, and then minimize the function $\sum_{a=1}^{N} p_a (\boldsymbol{x} - \boldsymbol{x}_a) \cdot \boldsymbol{\beta}_a (\boldsymbol{x} - \boldsymbol{x}_a) + \sum_{a=1}^{N} p_a \ln p_a$ to obtain the basis functions. A tensorial locality gives more flexibility to generate anisotropic shape functions, which could be useful to capture different physical features of a
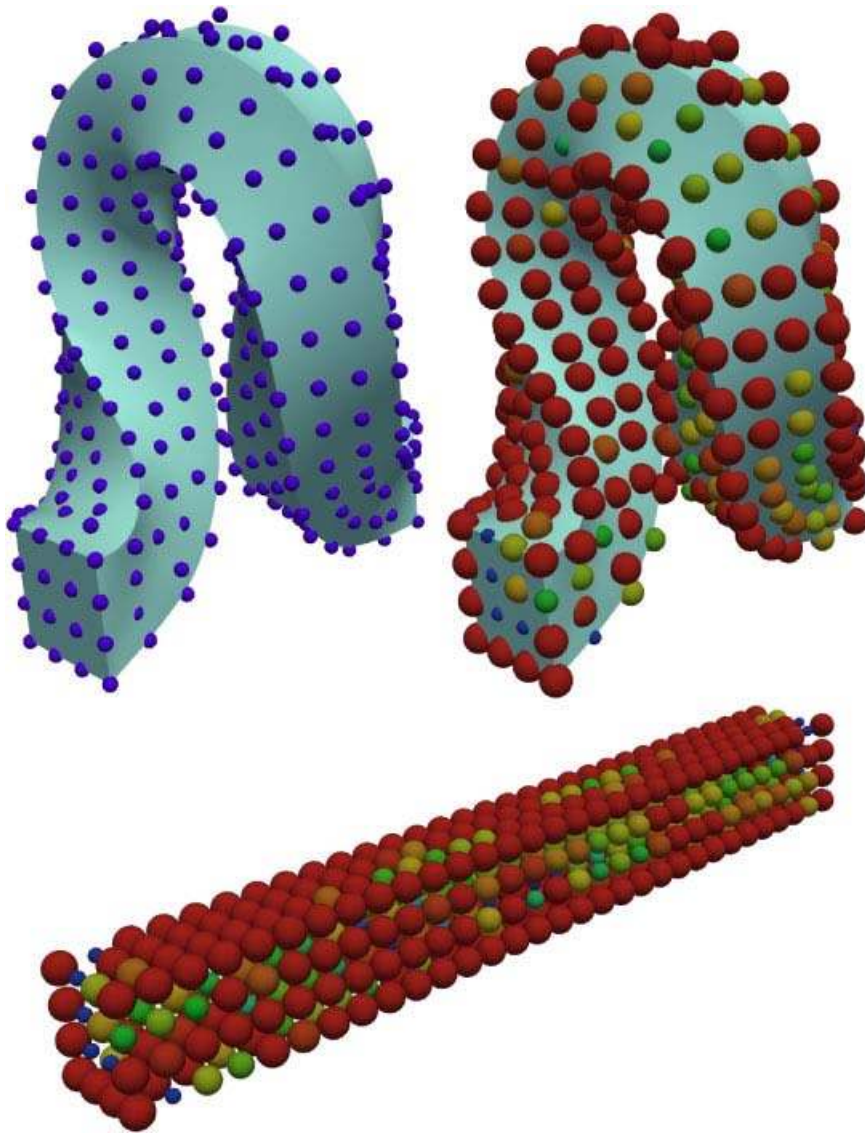
Figure 21. Buckling of a hyperelastic slender body (Example 6): deformed (top) and reference (bottom) configurations. For a better visualization, the optimum support sizes are illustrated in the reference configuration, and in the deformed configuration for half of the nodes.

problem. Such an approach seems to be appropriate for examples exhibiting localization or bi-material interfaces such as that in Section 4.3. This, together with the variational adaptivity with respect to the node set, i.e. variational r adaption for the LME approximants, is the topic of the current work.

Table IV. I_h^{FE}, I_h^{LME}, and I_h^{VA} are the energies corresponding to the non-adaptive FE and LME schemes, and the VA scheme, respectively, for Example 6.

# nodes	I_h^{FE}	I_h^{LME}	I_h^{VA}	$\left\|1-\frac{I_h^{FE}}{I_h^{VA}}\right\|$ (%)	$\left\|1-\frac{I_h^{LME}}{I_h^{VA}}\right\|$ (%)
189	1.547	0.764	0.711	117.6	7.4
496	1.166	0.697	0.662	76.3	5.3
2989	0.818	0.661			
88 641	0.664				

APPENDIX A: DERIVATIVES OF THE SHAPE FUNCTIONS

In this appendix, we detail the procedure to compute the various derivatives of the shape functions needed in the paper. We denote spatial gradients of scalar functions by ∇, whereas for vector-valued functions we denote by $D\boldsymbol{y}(\boldsymbol{x})$ the matrix of partial derivatives with respect to position. The symbol ∂ denotes partial differentiation. The subindexes a, b and c refer to nodes. Within the scope of the appendix, we define the following functions:

$$f_a(\boldsymbol{x}, \lambda, \beta_a) = -\beta_a |\boldsymbol{x} - \boldsymbol{x}_a|^2 + \lambda \cdot (\boldsymbol{x} - \boldsymbol{x}_a), \tag{A1}$$

$$p_a(\boldsymbol{x}, \lambda, \boldsymbol{\beta}) = \frac{\exp[f_a(\boldsymbol{x}, \lambda, \beta_a)]}{\sum_b \exp[f_b(\boldsymbol{x}, \lambda, \beta_b)]} = \frac{\exp[f_a(\boldsymbol{x}, \lambda, \beta_a)]}{Z(\boldsymbol{x}, \lambda, \boldsymbol{\beta})}, \tag{A2}$$

$$\boldsymbol{r}(\boldsymbol{x}, \lambda, \boldsymbol{\beta}) = \sum_a p_a(\boldsymbol{x}, \lambda, \boldsymbol{\beta})(\boldsymbol{x} - \boldsymbol{x}_a), \tag{A3}$$

$$\boldsymbol{J}(\boldsymbol{x}, \lambda, \boldsymbol{\beta}) = \frac{\partial \boldsymbol{r}}{\partial \lambda} = \sum_a p_a(\boldsymbol{x}, \lambda, \boldsymbol{\beta})(\boldsymbol{x} - \boldsymbol{x}_a) \otimes (\boldsymbol{x} - \boldsymbol{x}_a) - \boldsymbol{r}(\boldsymbol{x}, \lambda, \boldsymbol{\beta}) \otimes \boldsymbol{r}(\boldsymbol{x}, \lambda, \boldsymbol{\beta}). \tag{A4}$$

The dependence on the node set of locality parameters $\boldsymbol{\beta}$, on the evaluation point $\boldsymbol{x}$, and on the Lagrange multiplier λ is dropped for notational simplicity. The symbol * is used to denote that a function is evaluated in $\lambda^*(\boldsymbol{x}, \boldsymbol{\beta}) = \arg\min_{\lambda \in \mathbb{R}^d} \ln Z(\boldsymbol{x}, \lambda, \boldsymbol{\beta})$. This introduces explicit and implicit dependencies on $\boldsymbol{x}$ and $\boldsymbol{\beta}$ on all functions with *. Note that what has been denoted by p_a in the remainder of the paper is denoted by p_a^* in the appendix. No implied sum is assumed for repeated node indices.

A.1. Spatial derivatives

The first spatial derivative of the shape functions will be referred as ∇p_a^*. It is readily verified [6] that

$$\nabla p_a^* = p_a^* \left(\nabla f_a^* - \sum_c p_c^* \nabla f_c^* \right). \tag{A5}$$

Applying the chain rule, we have

$$\nabla f_a^* = \left(\frac{\partial f_a}{\partial \boldsymbol{x}} \right)^* + D\lambda^* \left(\frac{\partial f_a}{\partial \lambda} \right)^*, \tag{A6}$$

where

$$\left(\frac{\partial f_a}{\partial \boldsymbol{x}}\right)^* = -2\beta_a(\boldsymbol{x} - \boldsymbol{x}_a) + \lambda^*, \quad \left(\frac{\partial f_a}{\partial \lambda}\right)^* = (\boldsymbol{x} - \boldsymbol{x}_a).$$

The only term that is not available explicitly in Equation (A6) is $D\lambda^*$. In order to compute it we note that, since $\boldsymbol{r}^*$ is identically zero,

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

where

$$\left(\frac{\partial \boldsymbol{r}}{\partial \lambda}\right)^* = \boldsymbol{J}^*, \quad \left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{x}}\right)^* = -\boldsymbol{J}_\beta + \boldsymbol{I}, \quad \boldsymbol{J}_\beta = 2\sum_a \beta_a p_a^* (\boldsymbol{x} - \boldsymbol{x}_a) \otimes (\boldsymbol{x} - \boldsymbol{x}_a).$$

It follows that

$$D\lambda^* = (\boldsymbol{J}_\beta - \boldsymbol{I})(\boldsymbol{J}^*)^{-1}.$$

Rearranging terms, we finally obtain the spacial gradients of the shape functions as

$$\nabla p_a^* = p_a^* (\boldsymbol{r}_\beta - \boldsymbol{M}_a(\boldsymbol{x} - \boldsymbol{x}_a)),$$

where

$$\boldsymbol{r}_\beta = 2\sum_a \beta_a \, p_a^* (\boldsymbol{x} - \boldsymbol{x}_a), \quad \boldsymbol{M}_a = 2\beta_a \boldsymbol{I} - D\lambda^*.$$

A.2. Derivatives of p_a^* with respect to β_b

The goal is to obtain an explicit expression for the derivative of the shape function of node a with respect to the locality parameter of the node b. It is readily verified that

$$\frac{\partial p_a^*}{\partial \beta_b} = p_a^* \left(\frac{\partial f_a^*}{\partial \beta_b} - \sum_c p_c^* \frac{\partial f_c^*}{\partial \beta_b}\right), \tag{A7}$$

where

$$\frac{\partial f_a^*}{\partial \beta_b} = \left(\frac{\partial f_a}{\partial \beta_b}\right)^* + \left(\frac{\partial f_a}{\partial \lambda}\right)^* \left(\frac{\partial \lambda}{\partial \beta_b}\right)^*.$$

Note the difference between taking the partial derivative after evaluation at the optimal Lagrange multiplier λ^* (total derivative with respect to β_b) and taking the partial derivative of the explicit dependence of f_a on β_b, and then evaluating at the optimal Lagrange multiplier. In the above equation,

$$\left(\frac{\partial f_a}{\partial \beta_b}\right)^* = -\delta_{ab}|\boldsymbol{x} - \boldsymbol{x}_a|^2, \quad \left(\frac{\partial f_a}{\partial \lambda}\right)^* = (\boldsymbol{x} - \boldsymbol{x}_a),$$

where δ_{ab} is the Kronecker delta. Using the fact that $\boldsymbol{r}^* = 0$, we have

$$\left(\frac{\partial \lambda}{\partial \beta_b}\right)^* = p_b^* |\boldsymbol{x} - \boldsymbol{x}_b|^2 (\boldsymbol{J}^*)^{-1} (\boldsymbol{x} - \boldsymbol{x}_b).$$

Replacing all these expressions in Equation (A7), we find the derivative of the shape functions with respect to the locality parameters

$$\frac{\partial p_a^*}{\partial \beta_b} = p_a^* |x - x_b|^2 [p_b^*(1 + \Delta_{ab}) - \delta_{ab}],$$

where

$$\Delta_{ab} = (x - x_b) \cdot (J^*)^{-1} (x - x_a).$$

A.3. Derivatives of ∇p_a^* with respect to β_b

From Equation (A5) it follows that

$$\frac{\partial(\nabla p_a^*)}{\partial \beta_b} = \underbrace{\frac{\partial p_a^*}{\partial \beta_b} \left(\nabla f_a^* - \sum_c p_c^* \nabla f_c^* \right)}_{A} + \underbrace{p_a^* \frac{\partial(\nabla f_a^*)}{\partial \beta_b}}_{B} - \underbrace{p_a^* \sum_c \left[\frac{\partial p_c^*}{\partial \beta_b} \nabla f_c^* + p_c^* \frac{\partial(\nabla f_c^*)}{\partial \beta_b} \right]}_{C}.$$

Replacing the expressions of the different derivatives and defining

$$K_b = \sum_c \Delta_{bc} p_c^* M_c (x - x_c) \otimes (x - x_c), \quad j_a = (J^*)^{-1}(x - x_a),$$

it follows that

$$A = \nabla p_a^* |x - x_b|^2 [-\delta_{ab} + p_b^*(1 + \Delta_{ab})],$$

$$B = -2\delta_{ab}(x - x_a) + p_b^* \{|x - x_b|^2 [j_b + (I + K_b)j_a - \Delta_{ab} M_b (x - x_b)] + 2\Delta_{ab}(x - x_b)\},$$

$$C = -2p_b^*(x - x_b) - |x - x_b|^2 \nabla p_b^*.$$

Finally, the derivatives of the gradient of the shape functions with respect to the locality parameters can be rewritten as

$$\frac{\partial(\nabla p_a^*)}{\partial \beta_b} = p_a^*[-2(x - x_b)\delta_{ab} + (1 + \Delta_{ab})|x - x_b|^2 \nabla p_b^*] + \nabla p_a^* |x - x_b|^2 [-\delta_{ab} + (1 + \Delta_{ab})p_b^*]$$

$$+ p_a^* p_b^* \{2(1 + \Delta_{ab})(x - x_b) + |x - x_b|^2 [j_b + (I + K_b)j_a - \Delta_{ab} r_\beta]\}.$$

REFERENCES

1. Belytschko T, Krongauz Y, Organ D, Fleming M, Krysl P. Meshless methods: an overview and recent developments. *Computer Methods in Applied Mechanics and Engineering* 1996; **139**:3–47.
2. Li S, Liu W. Meshfree and particle methods and their applications. *Applied Mechanics Reviews* 2002; **55**(1):1–34.

3. Huerta A, Belytscko T, Fernández-Méndez S, Rabczuk T. Meshfree methods, vol. 1. *Encyclopedia of Computational Mechanics*. Wiley: Chichester, 2004; 279–309.
4. Sukumar N, Wright R. Overview and construction of meshfree basis functions: from moving least squares to entropy approximants. *International Journal for Numerical Methods in Engineering* 2007; **70**(2):181–205.
5. Lancaster P, Salkauskas K. Surfaces generated by Moving Least-Squares methods. *Mathematics of Computation* 1981; **37**(155):141–158.
6. Arroyo M, Ortiz M. Local maximum-entropy approximation schemes: a seamless bridge between finite elements and meshfree methods. *International Journal for Numerical Methods in Engineering* 2006; **65**:2167–2202.
7. Cyron CJ, Arroyo M, Ortiz M. Smooth, second order, non-negative meshfree approximants selected by maximum entropy. *International Journal for Numerical Methods in Engineering* 2009; **79**(13):1605–1632.
8. Sukumar N. Construction of polygonal interpolants: a maximum entropy approach. *International Journal for Numerical Methods in Engineering* 2004; **61**(12):2159–2181.
9. Ma Z, Chen H, Zhou C. A study of point moving adaptivity in gridless method. *Computer Methods in Applied Mechanics and Engineering* 2008; **197**:1926–1937.
10. Vidal Y, Parés N, Díez P, Huerta A. Bounds for quantities of interest and adaptivity in the element free Galerkin method. *International Journal for Numerical Methods in Engineering* 2008; **76**(77):1782–1818.
11. Liu W, Li S, Belytschko T. Moving least square reproducing kernel methods. Part i: methodology and convergence. *Computer Methods in Applied Mechanics and Engineering* 1997; **143**(1–2):113–154.
12. Fries T, Matthies H. Classification and overview of meshfree methods. *Technical Report*, Institute of Scientific Computing, Technical University Braunschweig, Brunswick, Germany, July 2004.
13. Kuhl E, Askes H, Steinmann P. An ALE formulation based on spatial and material settings of continuum mechanics. Part 1: generic hyperelastic formulation. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**(39–41):4207–4222.
14. Thoutireddy P, Ortiz M. A variational r-adaption and shape-optimization method for finite-deformation elasticity. *International Journal for Numerical Methods in Engineering* 2004; **61**:1–21.
15. Mosler J, Ortiz M. On the numerical implementation of variational arbitrary Lagrangian–Eulerian (VALE) formulations. *International Journal for Numerical Methods in Engineering* 2006; **67**:1272–1289.
16. Sukumar N, Moran B, Belytschko T. The natural element method in solid mechanics. *International Journal for Numerical Methods in Engineering* 1998; **43**(5):839–887.
17. Cirak F, Ortiz M, Schröder P. Subdivision surfaces: a new paradigm for thin-shell finite-element analysis. *International Journal for Numerical Methods in Engineering* 2000; **47**(12):2039–2072.
18. Hughes T, Cottrell J, Bazilevs Y. Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Computer Methods in Applied Mechanics and Engineering* 2005; **194**:4135–4195.
19. Rajan V. Optimality of the Delaunay triangulation in R^d. *Discrete and Computational Geometry* 1994; **12**(2): 189–202.
20. Parés N, Díez P, Huerta A. Exact bounds for linear outputs of the advection–diffusion-reaction equation using flux-free error estimates. *SIAM Journal on Scientific Computing* 2009; **31**(4):3064–3089.
21. Thoutireddy P. Variational arbitrary Lagrangian–Eulerian method. *Ph.D. Thesis*, California Institute of Technology 2003.
22. Askes H, Kuhl E, Steinmann P. An ALE formulation based on spatial and material settings of continuum mechanics. Part 2: classification and applications. *Computer Methods in Applied Mechanics and Engineering* 2004; **193**(39–41):4223–4245.
23. Askes H, Bargmann S, Kuhl E, Steinmann P. Structural optimization by simultaneous equilibration of spatial and material forces. *Communications in Numerical Methods in Engineering* 2005; **21**:433–442.
24. Zielonka M, Ortiz M, Marsden J. Variational R-adaption in elastodynamics. *International Journal for Numerical Methods in Engineering* 2008; **74**:1162–1197.
25. Nocedal J. Updating quasi-Newton matrices with limited storage. *Mathematics of Computation* 1980; **35**:773–782.
26. Byrd R, Lu P, Nocedal J. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing* 1995; **16**(5):1190–1208.
27. Timoshenko S, Goodier J. *Theory of Elasticity*. McGraw-Hill: New York, 1951.
28. Krysl P, Belytschko T. Element-free Galerkin method: convergence of the continuous and discontinuous shape functions. *Computer Methods in Applied Mechanics and Engineering* 1997; **148**:257–277.
29. Mosler J, Ortiz M. An error-estimate-free and remapping-free variational mesh refinement and coarsening method for dissipative solids at finite strains. *International Journal for Numerical Methods in Engineering* 2009; **77**(3):437–450.