A MULTISCALE IMPLEMENTATION BASED ON ADAPTIVE MESH REFINEMENT FOR THE NONLOCAL PERIDYNAMICS MODEL IN ONE DIMENSION[*]

FEIFEI XU[†], MAX GUNZBURGER[†], JOHN BURKARDT[†], AND QIANG DU[‡]

Abstract. Peridynamics models for solid mechanics feature a horizon parameter δ that specifies the maximum extent of nonlocal interactions. In this paper, a multiscale implementation of peridynamics models is proposed. In regions in which the displacement field is smooth, grid sizes are large relative to δ, leading to a local behavior of the models, whereas in regions containing defects, e.g., cracks, δ is larger than the grid size. Discontinuous (continuous) Galerkin finite element discretizations are used in regions where defects do (do not) occur. Moreover, in regions where no defects occur, the multiscale implementation seamlessly transitions to the use of a standard finite element discretization of a corresponding PDE model. Here, we demonstrate the multiscale implementation in a simple one-dimensional setting. An adaptive strategy is incorporated to detect discontinuities and effect grid refinement, resulting in a highly accurate and efficient implementation of peridynamics.

Key words. peridynamics, multiscale methods, nonlocal models, solid mechanics, finite element methods, discontinuous displacements

AMS subject classifications. 74A99, 65N30, 74A45, 45B05, 45E05, 65R99

DOI. 10.1137/15M1010300

1. Introduction. The nucleation and propagation of defects (e.g., cracks) in solids that have suffered damage in the form of discontinuities in the displacement field cannot be modeled by classical approaches such as the equations of linear elasticity. As a result, there have been many efforts, too numerous to list here, directed at the development of models that can handle defects in solids. Many of these efforts involve the development of different (often nonlinear) systems of partial differential equations (PDEs). On the other hand, peridynamics (PD) models [27, 30], which were also developed to treat defects in solids, do not involve spatial derivatives of the displacement field; hence, they are especially adept at resolving defects. PD models are continuum models in which the internal forces at a point are determined through an integral, as opposed to a differential, term. In this sense, PD models are *nonlocal;* i.e., the mechanical response at a point $\mathbf{x}$ in the material depends on the displacement field at points that are a finite distance away from $\mathbf{x}$. In contrast, for PDE models, that response depends only on points which are within an infinitesimal distance from $\mathbf{x}$. Successful PD modeling of fracture and other defects in solids has been demonstrated in several settings [5, 6, 15, 27, 29, 30].

Classical PDE models for solid mechanics do not possess a length scale other than that determined by material inhomogeneities. On the other hand, PD models involve a (possibly phenomenological) parameter δ, referred to as the *horizon*, which limits the

extent of nonlocal interactions; i.e., two points in the material separated by a distance greater than δ do not interact. As such, PD models have a "mechanical" length scale, even if the material is homogeneous. It has been shown that, as $\delta \to 0$, solutions of PD models converge to those of classical PDE models; see, e.g., [13, 31] for smooth solutions and [19, 20] for solutions with minimal regularity. On the other hand, when δ is fixed, straightforward discretizations of the PD model using atomistically spaced grids have been shown to reduce to molecular dynamics models; see, e.g., [25, 26]. This is the *multiscale* nature of PD that we exploit in the development of a multiscale implementation of PD.

Variational formulations of PD not only allow for the analysis of the well-posedness and other properties of PD models [9, 19, 20], but also naturally suggest a discretization using finite element methods (FEMs) [7, 11, 12, 18, 32, 33, 35]. It is important to note that *for PD models, discontinuous Galerkin (DG) methods are conforming for a large class of nonlocal interaction kernels,* which is the case considered in this work.[1] By including discontinuous functions in the finite element space, one naturally hopes that jump discontinuities in the displacement field can be more easily resolved. However, discontinuities in finite element basis functions occur at element boundaries, whereas discontinuities in the displacement field, whose locations are not known a priori, in general occur within an element. Such a mismatch of the locations of discontinuities can be a major cause of the loss of accuracy in naive, e.g., quasi-uniform grid, implementations of DG methods for PD.

In [7], piecewise-linear polynomial FEMs for one-dimensional PD were considered, and a strategy was proposed to mitigate the loss of accuracy caused by the mismatch in the locations of discontinuities of the solution of the PD model and of its DG approximation. The mitigation strategy involved the insertion, within a grid of nominal size h, of elements of length h^4, each of which contains one of the points at which the solution is discontinuous. Also proposed in [7] was a hybrid strategy in which the DG method was applied in elements that contain solution discontinuities, and a continuous Galerkin (CG) method was used elsewhere. Such a hybrid strategy mitigates the greater cost, i.e., the larger number of degrees of freedom, attendant to DG methods compared to CG methods. We note that in [7], the locations of the discontinuities in the solution of the one-dimensional PD model were assumed to be known; of course, in practice, this is generally not the case.

The multiscale nature of PD models, resulting from the horizon δ and the FEM discretizations studied in [7], motivates the *multiscale implementation of PD models* that is the focus of this paper. Suppose h is a typical grid size in a discretization of a PDE model for solid mechanics in cases where the solution (i.e., the displacement) is known to be smooth; for example, if one is interested in achieving an L^2 error of $O(\varepsilon)$ and one uses continuous piecewise-linear finite element approximations, then h would be chosen to be of $O(\varepsilon^{1/2})$. Even in cases where the solution has discontinuities, one can safely use the PDE model in regions where the displacement is smooth. However, near the location of discontinuities in the displacement, the PDE model breaks down, and one instead uses a PD model. However, as discussed above, to preserve accuracy, one also has to refine the grid near the location of discontinuities. Of course, because the locations of those discontinuities are, in general, not known a priori, one has to be able to detect where they occur. One also has to devise a strategy for coupling the PD and PDE models.

---

[1]For more singular kernels that do not allow conforming DG approximations, nonconforming DG methods for nonlocal problems were developed in [34].

Based on the above discussions, a multiscale implementation of the PD model would start with a choice for the bulk grid size h and a horizon parameter δ and then include the following components:

1. detection of elements that contain a discontinuity in the displacement,
2. refinement of the grid as necessary near the discontinuities,
3. use of DG for PD in regions containing the discontinuity,
4. use of CG for PD in regions neighboring the discontinuity,
5. use of CG for PDE if sufficiently far away from the discontinuity,
6. use of quadrature rules that can be applied for any combination of h and δ.

Generically, one would choose $\delta \ll h$ because then the discretized PD model essentially reduces to a local model—e.g., only nearest neighbor interactions occur—whenever the local grid size is of $O(h)$. This effect greatly facilitates the coupling of a PD model to a PDE model. On the other hand, we would also choose $\delta > \underline{h}$, where $\underline{h}$ denotes the grid size one would use for elements in which solution discontinuities occur. With $\delta > \underline{h}$, the PD model is nonlocal. Recall that, to preserve accuracy, $\underline{h}$ depends on h and, in fact, $\underline{h} \ll h$, so that we have $\underline{h} < \delta \ll h$; for example, for piecewise-linear finite element approximations ideally one would choose $\underline{h} = O(h^4)$.

For simplicity, in this paper we apply the different models based on nodal locations; this is akin to using a "force"-based approach for choosing which model to apply where. The alternative is to use an "energy"-based approach for which one applies the different models based on element locations. Although the former approach leads to some anomalies in the discretized system, the results obtained using the two approaches are very much the same. Figure 1 suggests how, in a one-dimensional setting, such a multiscale, force-based implementation would subdivide a region into models of different types and into FEM discretizations of different types.
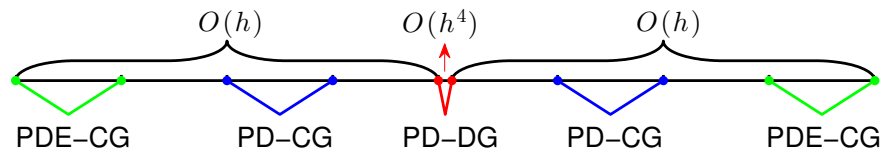

Fig. 1. *A multiscale implementation of PD in one dimension. The discontinuity in the displacement occurs in the element of size $O(h^4)$.*

Obviously, our multiscale implementation is also a *multimodel* one because it not only involves a nonlocal PD model, but also a local PDE model. Moreover, the implementation is also a *multiapproximation* model because it involves both CG and DG FEMs. All these components are judiciously used to detect and resolve jump discontinuities in the solution while achieving the same accuracy with only a slight increase in cost compared to a standard finite element discretization of a PDE having smooth solutions.

In this paper, we demonstrate how to carry out the components of the multiscale implementation of PD as listed above. Regarding the first component, an adaptive strategy based on posterior error estimators has to be adopted to detect the unknown locations of the discontinuities in the solution. Regarding the second component, during the adaptive process we not only refine near points at which the solution is discontinuous, but also coarsen previously refined regions where now the solution is verified as being smooth. Doing so results in abrupt transitions from refined to coarse grids and in small increases in the number of degrees of freedom compared to those

for a discretization of a problem for which the solution is smooth everywhere.

Our goal for this paper is to present the general principles and demonstrate the potential usefulness of our multiscale implementation of PD. To do so, we consider only the one-dimensional case. Doing so enables us to avoid several serious issues that arise in any FEM discretization of PD in higher dimensions and to thus concentrate on only those issues that are specific to the multiscale implementation and to discuss those issues in a simple, easy to explain and understand, setting. For the same reasons, we consider only time-independent problems.

Other papers, e.g., [1, 16, 17, 21, 23, 24, 28], can be viewed as being concerned with multiscale implementations of PD models. However, all these papers lack at least one, and most often several, of the components of our complete multiscale implementation, components such as the use of FEM, automatic detection of discontinuities, adaptive mesh refinement based on posterior error estimation resulting in meshes with abrupt size transitions, the use of both continuous and discontinuous Galerkin methods to seamlessly couple local and nonlocal models and to economize on the cost of PD implementation, and being able to differentiate between true discontinuities and steep gradients.

The paper is organized as follows. In section 2, a nonlocal problem and its specialization to one-dimensional PD are introduced. A review of some of the results given in [7] that are germane to our purposes is given in section 3. Two important ingredients in our multiscale implementation are discussed in sections 4 and 5, namely the posterior error estimators and local-nonlocal coupling strategies that we use. The multiscale implementation strategy is then given in section 6. Extensive numerical results that explore several features of the implementation are provided in section 7. In section 8, we briefly consider solutions having steep gradients and discuss a variant of our algorithm that differentiates between and can handle true discontinuities and steep gradients occurring in the same solution. Finally, in section 9, we make some concluding remarks, including a brief discussion of the extension of our strategy to higher dimensions.

2. Nonlocal model problem. The action of the operator $\mathcal{L}$ on a function $u(\mathbf{x})$ is defined as

$$\mathcal{L}u(\mathbf{x}) = -2 \int_{\mathbb{R}^n} \big(u(\mathbf{x}') - u(\mathbf{x})\big)\gamma(\mathbf{x}, \mathbf{x}')\,d\mathbf{x}' \quad \forall\,\mathbf{x} \in \mathbb{R}^n,$$

where γ denotes a symmetric kernel, i.e., $\gamma(\mathbf{x}, \mathbf{x}') = \gamma(\mathbf{x}', \mathbf{x})$ for all $\mathbf{x}, \mathbf{x}'$. The *interaction domain* corresponding to a given open subset $\Omega \subset \mathbb{R}^n$ is defined as

$$\Omega_{\mathcal{I}} = \{\mathbf{x}' \in \mathbb{R}^n \backslash \Omega \,:\, \gamma(\mathbf{x}, \mathbf{x}') \neq 0 \text{ for some } \mathbf{x} \in \Omega\}; \tag{1}$$

i.e., $\Omega_{\mathcal{I}}$ consists of all points in $\mathbb{R}^n$ that interact with points in Ω but are located outside of Ω. In general, $\Omega_{\mathcal{I}}$ has nonzero volume in $\mathbb{R}^n$. With this notation, the *nonlocal volume-constrained problem* we consider is given by

$$\begin{cases} \mathcal{L}u(\mathbf{x}) = -2 \displaystyle\int_{\Omega \cup \Omega_{\mathcal{I}}} \big(u(\mathbf{x}') - u(\mathbf{x})\big)\gamma(\mathbf{x}, \mathbf{x}')\,d\mathbf{x}' = f(\mathbf{x}) \quad \forall\,\mathbf{x} \in \Omega, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad u(\mathbf{x}) = g(\mathbf{x}) \quad \forall\,\mathbf{x} \in \Omega_{\mathcal{I}}, \end{cases} \tag{2}$$

where $f(\mathbf{x}) : \Omega \to \mathbb{R}$ and $g(\mathbf{x}) : \Omega_{\mathcal{I}} \to \mathbb{R}$ are given functions. Note that the constraint $u = g$ is applied on the set $\Omega_{\mathcal{I}}$ having finite volume and not on the boundary of Ω; hence, the terminology *volume-constrained problem*.

The problem (2), which involves a scalar-valued unknown function u, is used as a model for anomalous diffusion in $\mathbb{R}^n$ (see, e.g., [9]) and, in case $n = 1$, is also of the form of a linear PD model for mechanics in one spatial dimension. Although the latter case is the focus of this paper, in this section we present the more general case to highlight the fact that many of the discussions in this paper apply to the diffusion case as well.

We define the bilinear form

$$(3) \qquad A(u, v) = \int_{\Omega \cup \Omega_\mathcal{I}} \int_{\Omega \cup \Omega_\mathcal{I}} \big(u(\mathbf{x}') - u(\mathbf{x})\big)\big(v(\mathbf{x}') - v(\mathbf{x})\big)\gamma(\mathbf{x}, \mathbf{x}') \, d\mathbf{x}' d\mathbf{x},$$

the "energy" norm $|||u||| = (A(u, u))^{1/2}$, the energy space

$$V(\Omega \cup \Omega_\mathcal{I}) = \{u \in L^2(\Omega \cup \Omega_\mathcal{I}) \ : \ |||u||| < \infty\},$$

and the constrained energy space

$$V_c(\Omega \cup \Omega_\mathcal{I}) = \{v \in V \ : \ v(\mathbf{x}) = 0 \ \forall \mathbf{x} \in \Omega_\mathcal{I}\}.$$

It is shown in, e.g., [9, 10], that (2) has the variational formulation

$$(4) \qquad \begin{cases} given \ f \in L^2(\Omega) \ and \ g \in L^2(\Omega_\mathcal{I}), \ seek \ u \in V(\Omega \cup \Omega_\mathcal{I}) \ such \ that \\ u(\mathbf{x}) = g(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega_\mathcal{I} \ and \\ A(u, v) = \int_\Omega fv \, d\mathbf{x} \quad \forall v \in V_c(\Omega \cup \Omega_\mathcal{I}), \end{cases}$$

and that the problem (4) is well posed with respect to the energy space $V(\Omega \cup \Omega_\mathcal{I})$.

Many kernels, both integrable and nonintegrable, have been proposed in the literature; see, e.g., [9, 10]. For the sake of concreteness, we consider kernels of the form

$$(5) \qquad \gamma(\mathbf{x}, \mathbf{x}') = \frac{1 - 1s}{\delta^{2-2s}} \frac{1}{|\mathbf{x} - \mathbf{x}'|^{n+2s}} \mathbb{1}_{H_{\mathbf{x}, \delta}},$$

where n denotes the spatial dimension, s a constant, $\mathbb{1}$ the indicator function, $H_{\mathbf{x}, \delta} = \{\mathbf{x}' \in \mathbb{R}^n \ : \ |\mathbf{x} - \mathbf{x}'| \leq \delta\}$, and $\delta > 0$ the horizon. Note that $\gamma(\mathbf{x}, \mathbf{x}')$ given by (5) is positive and depends only on the relative position of $\mathbf{x}$ and $\mathbf{x}'$ in the reference configuration, i.e., $\gamma(\mathbf{x}, \mathbf{x}') = \gamma(\mathbf{x} - \mathbf{x}')$. Also, because the indicator function appears in (5), we have that $\Omega_\mathcal{I} = \{\mathbf{x}' \in \mathbb{R}^n \backslash \Omega \ : \ |\mathbf{x}' - \mathbf{x}| \leq \delta$ for some $\mathbf{x} \in \Omega\}$; i.e., $\Omega_\mathcal{I}$ is a layer of thickness δ surrounding Ω.

If $s < 0$, the kernel (5) is integrable; i.e., we have that, for some constant $c^*(\delta) > 0$ whose value depends on δ,

$$(6) \qquad \int_{\mathbb{R}^n} \gamma(\mathbf{x}, \mathbf{x}') d\mathbf{x}' \leq c^*(\delta) < \infty \quad \forall \mathbf{x} \in \Omega.$$

If $s \geq 0$, the kernel (5) is not integrable. It is known that for integrable kernels, i.e., for kernels that satisfy (6), the energy space $V(\Omega \cup \Omega_\mathcal{I}) = L^2(\Omega \cup \Omega_\mathcal{I})$. For $s \in (0, 1)$, it is known that $V(\Omega \cup \Omega_\mathcal{I}) = H^s(\Omega \cup \Omega_\mathcal{I})$, where $H^s(\cdot)$ denotes the fractional Sobolev space of order s. For $s = 0$, there is no known characterization of the energy space in terms of well-known spaces; what is known is that in this case $V(\Omega \cup \Omega_\mathcal{I})$ is a Hilbert space and is a strict subspace of $L^2(\Omega \cup \Omega_\mathcal{I})$. Note also that if $n + 2s > 0$, the kernel (5) is singular. Detailed discussions are given in, e.g., [9, 10].

If $n = 1$ and $s = 0$ in (5), the nonlocal problem (2) is a typical example of a linear PD model in one dimension. In higher dimensions, integrable kernels are also of interest in the PD setting. Thus, in the one-dimensional setting, we also consider cases for which $n = 1$ and $s < 0$ in (5). We leave to a later study the case $s \in (0, 1)$, which is of greater interest in the anomalous diffusion setting.

3. FEMs for PD in one dimension. In this section, we review some results from [7] for FEMs for the problem (2) for the kernel (5) with $n = 1$ and $s = 0$, i.e., for one-dimensional PD. In this case, we have $\Omega = (a, b)$ and $\Omega_\mathcal{I} = (a - \delta, a] \cup [b, b + \delta)$. Three finite element spaces were considered: continuous piecewise-linear (CL), discontinuous piecewise-constant (DC), and discontinuous piecewise-linear (DL). In all cases, two of the nodes of the grid for $\Omega \cup \Omega_\mathcal{I}$ were located at $x = a$ and $x = b$, i.e., at the end points of Ω. With N denoting the number of elements, i.e., subintervals, in Ω, the number of unknowns associated with each choice of finite element space is $N - 1$, N, and $2N$, respectively.

Two different means for choosing the horizon δ were used: δ proportional to the grid size h and δ fixed independent of h. Because the latter choice is relevant to our multiscale implementation of PD, we now quote some of the results of [7] for that case where two different manufactured solutions were considered, one being smooth and one containing a jump discontinuity within Ω. Two different sequences of uniform grids were considered, resulting in three cases in total:

1. *smooth case*: the solution is smooth;
2. *fortuitous case*: for all grids used the solution has a jump discontinuity at a point that coincides with a grid point;
3. *general case*: for all grids used the solution has a jump discontinuity at a point that lies strictly within a grid interval.

The $L^2(\Omega \cup \Omega_\mathcal{I})$ and $L^\infty(\Omega \cup \Omega_\mathcal{I})$ norms of the error as reported in [7] are given in Table 1 for the three solution types and for the three FEMs.

TABLE 1
Convergence behavior, as provided in [7], of $L^2(\Omega \cup \Omega_\mathcal{I})$ and $L^\infty(\Omega \cup \Omega_\mathcal{I})$ errors for the continuous linear (CL), discontinuous constant (DC), and discontinuous linear (DL) FEMs with fixed δ and uniform grids of size h for different types of solutions.

Error type	$L^2(\Omega \cup \Omega_\mathcal{I})$			$L^\infty(\Omega \cup \Omega_\mathcal{I})$		
FEM type	CL	DC ($h < \delta$)	DL	CL	DC ($h < \delta$)	DL
Smooth	$O(h^2)$	$O(h)$	$O(h^2)$	$O(h^2)$	$O(h)$	$O(h^2)$
Fortuitous	$O(h^{\frac{1}{2}})$	$O(h)$	$O(h^2)$	$O(h^0)$	$O(h)$	$O(h^2)$
General	–	–	$O(h^{\frac{1}{2}})$	–	–	$O(h^0)$

From the table, we see that for smooth solutions, error behaviors for both norms and for all three FEMs are in accordance with expectations, i.e., second-order convergence for linear approximations and first-order convergence for the constant approximation. Note, however, that results for the DC approximations were reported only for $h < \delta$ because the study in [7] found that, even for smooth solutions, DC approximations are not robust with respect to the relative sizes of δ and h due to the loss of asymptotical compatibility; see [32, 33] for detailed explanations.

For the "fortuitous" case, the CL method results in severely degraded error convergence rates. In contrast, for the DL method (and the DC method with $h < \delta$), optimal convergence rates are again obtained. This can be attributed to the fact that methods with discontinuous elements allow for discontinuities at grid points, so that

they can resolve the discontinuity in the solution occurring at a grid point. Of course, this cannot be achieved using only continuous elements; hence the deterioration in the convergence rates for that case. Because the CL method leads to poor performance when dealing with discontinuities, even for the fortuitous case, in what follows we consider the use of such methods only in regions in which the solution is smooth. We also do not further consider the use of the DC method because it is not robust for $h > \delta$.

For the "general" case, the DL method exhibits the same deterioration in accuracy that occurred for the CL method for the fortuitous case. Within an element, DL basis functions are continuous, so they cannot resolve a discontinuity in the solution that occurs in the interior of an element.

A strategy was proposed in [7] to mitigate, for the DL method in the general case, the loss of accuracy with respect to the $L^2(\Omega \cup \Omega_\mathcal{I})$ norm, indeed to in some sense recover the optimal convergence rate. Within a uniform grid having grid size h, a single element of length $O(h^4)$ was inserted that contained the point at which the solution is discontinuous. This strategy results in an overall error of $O(h^2)$ with respect to the $L^2(\Omega \cup \Omega_\mathcal{I})$ norm. This result was illustrated in [7] using numerical examples. Note that the transition from the element of size $O(h^4)$ to those of size h can be abrupt; i.e., smooth transition layers of elements are not necessary. Of course, to implement this strategy one must know where the discontinuity in the solution is located. In general, to do so, one would have to adopt some sort of adaptive refinement strategy.

Unfortunately, the mitigation strategy could not do anything about the loss of accuracy with respect to the $L^\infty(\Omega \cup \Omega_\mathcal{I})$ norm, because regardless of how small the element containing the discontinuity in the solution may be, the $L^\infty(\Omega \cup \Omega_\mathcal{I})$-norm error remains of $O(1)$. However, in [7], it was shown through numerical experiments that if the element containing the discontinuity is omitted when computing the $L^\infty(\Omega \cup \Omega_\mathcal{I})$ norm of the error, the $O(h^2)$ rate is recovered; i.e., there is no pollution to other elements from the large error caused by the element containing the solution discontinuity.

Because the DL method generates twice as many unknowns as the CL method, it makes sense to try a hybrid approach in which the DL method is applied in regions where discontinuities are encountered, whereas the less expensive CL method is used elsewhere, i.e., where the solution is smooth. This is the *hybrid* approach proposed in [7], which continues to exhibit optimal convergence rates while reducing the cost relative to applying the DL method everywhere. In fact, if one uses the CL method in all elements of size h, the cost of the hybrid approach is essentially the same as that for using the CL method everywhere.

Guided by the results of [7, 11] as well as those in [32, 33], the subsequent discussion focuses on continuous linear and discontinuous linear finite element approximations as candidates for use in the multiscale implementation. Those results also lead us to consider adaptive refinement strategies to detect where discontinuities in the solution occur and to use grids with abrupt size transitions to efficiently recover optimal convergence rates.

4. Posterior error estimator. Adaptive grid refinement and discontinuity detection are important steps within our multiscale implementation strategy. In this section, we discuss the means we use to effect these steps.

Posterior error analyses for nonlocal models were investigated in [11, 12]. A residual-based error estimator was introduced, and its equivalence with the energy

norm $|||e^h|||$ of the solution error was proved. We base the error estimator in our multiscale implementation on that developed in [11, 12].

For any element K in a mesh, the residual error is defined as

$$R^h(\mathbf{x}) = f(\mathbf{x}) - \mathcal{L}\big(u^h(\mathbf{x})\big) \quad \forall\, \mathbf{x} \in K.$$

If γ satisfies (5) with $s \in [0, 1)$, the posterior error is defined by

(7) $$\widetilde{\eta}(u^h, K) = h^{2s}\|R^h\|_{L^2(K)},$$

where s is defined in (5). Recall that for the one-dimensional PD setting we take $n = 1$ and $s = 0$, though this is a case, technically speaking, that is not covered by the theory given in [11] (with integrable kernels) and [12] (with $s > 0$). Nevertheless, we did not find any inconsistency with the theory in the practical implementation.

The total error estimator over $\Omega = (a, b)$ is then given by

$$\widetilde{\eta}(u^h, \Omega) = \left(\sum_K \widetilde{\eta}^2(u^h, K)\right)^{1/2}.$$

For simplicity, we will henceforth simply write $\widetilde{\eta}(K)$ and $\widetilde{\eta}(\Omega)$ for $\widetilde{\eta}(u^h, K)$ and $\widetilde{\eta}(u^h, \Omega)$, respectively.

The following convergent adaptive refinement algorithm is proposed in [12].

Algorithm 1: Standard mesh refinement.

Input: Dörfler marking parameter (see [8]) $\theta \in (0, 1]$, and an initial mesh $\mathcal{T}_0$.

1 Initialization: $k = 0$;

2 Solve the discretized nonlocal problem to obtain the finite element approximation $u_k^h(x)$ with respect to the mesh $\mathcal{T}_k$;

3 Evaluate the posterior error estimator $\widetilde{\eta}(K)$ for every element $K \in \mathcal{T}_k$;

4 Determine the subset $\mathcal{M}_k \subset \mathcal{T}_k$ with a minimal cardinality such that

$$\widetilde{\eta}^2(u_k^h, \mathcal{M}_k) \geq \theta\widetilde{\eta}^2(\Omega);$$

refine the elements in $\mathcal{M}_k$ to obtain the mesh $\mathcal{T}_{k+1}$;

5 Set $k := k + 1$ and return to line 2.

Note that in step 4 one would simply order the elements by their errors, where the first element is the one with maximum error and the last one is the one with minimum error, and then $\mathcal{M}_k$ is composed of the first few elements in such a list.

Algorithm 1 is efficient with respect to error reduction and is provably convergent [12] in the context considered there. However, it is not aimed at detecting discontinuities in the solution. To see this, consider the one-dimensional case, and let $\widehat{K} \in \mathcal{T}_k$ denote an element that contains a point at which the solution has a jump discontinuity. Given the discussion in section 3, we have that $\widetilde{\eta}(\widehat{K}) = O(h_{\widehat{K}}^{1/2})$, whereas $\widetilde{\eta}(K) = O(h_K^2)$ for elements K within which the solution is smooth. Again, given the discussion in section 3, as $k \to \infty$ we presumably want to drive the length of the element $\widehat{K}$ to $O(h^4)$, whereas the elements K within which the solution is smooth remain of size $O(h)$. We then have that

$$\widetilde{\eta}^2(\widehat{K}) = O(h_{\widehat{K}}) = O(h^4) \quad \text{and} \quad \widetilde{\eta}^2(K) = O(h_K^4) = O(h^4) \qquad \text{as } k \to \infty;$$

i.e., the estimator cannot tell the difference between an element $\widehat{K}$ that contains a jump discontinuity in the solution and the elements K within which the solution is smooth.

To be able to detect the elements containing points at which the solution is discontinuous, we instead use the *grid size–weighted* posterior estimators

$$(8) \qquad \eta^2(K) = \frac{\widetilde{\eta}^2(K)}{|K|} \quad \forall\, K \in \mathcal{T}_k \qquad \text{and} \qquad \eta^2(\Omega) = \sum_{K \in \mathcal{T}_k} \eta^2(K),$$

where $|K|$ denotes the length of the element K; in two and three dimensions, $|K|$ would denote the area and volume, respectively, of the element K. We now have

$$\eta^2(\widehat{K}) = \frac{\widetilde{\eta}^2(\widehat{K})}{|\widehat{K}|} = \frac{O(h_{\widehat{K}})}{h_{\widehat{K}}} = \frac{O(h^4)}{h^4} = O(1) \qquad \text{as } k \to \infty$$

and

$$\eta^2(K) = \frac{\widetilde{\eta}^2(K)}{|K|} = \frac{O(h^4)}{h} = O(h^3) \qquad \text{as } k \to \infty;$$

i.e., the estimator can now tell the difference between an element $\widehat{K}$ that contains a jump discontinuity in the solution and the elements K within which the solution is smooth. In our multiscale implementation, the weighted posterior error estimators (8) are used to detect the location of discontinuities in the solution.

It will be useful to sort and normalize the elementwise errors because, as is seen below, a few elements near the discontinuity may account for a substantial portion of the overall error. Therefore, suppose that the elementwise posterior error estimators have been sorted so that $\widetilde{\eta}(K_{i_1}) \geq \widetilde{\eta}(K_{i_2}) \geq \cdots \geq \widetilde{\eta}(K_{i_M})$, $\eta(K_{j_1}) \geq \eta(K_{j_2}) \geq \cdots \geq \eta(K_{j_M})$, and define the sequence of cumulative relative errors $\widetilde{E}^n$ and E^n by

$$(9) \qquad \widetilde{E}^n = \sum_{s=1}^{n} \frac{\widetilde{\eta}^2(K_{i_s})}{\widetilde{\eta}^2(\Omega)} \quad \text{and} \quad E^n = \sum_{s=1}^{n} \frac{\eta^2(K_{j_s})}{\eta^2(\Omega)}.$$

Thus, e.g., $\widetilde{E}^1$ denotes the maximum relative error over any of the elements, and $\widetilde{E}^2$ denotes the sum of the two largest relative errors.

Integrable kernels, i.e., kernels satisfying (6), are of interest for PD in two and three dimensions, so we also consider integrable kernels in our one-dimensional numerical experiments. In this case, the posterior error estimator (7) now takes the form

$$\widetilde{\eta}(u^h, K) = \frac{\|R^h\|_{L^2(K)}}{\sqrt{c^*(\delta)}},$$

where $c^*(\delta)$ is defined in (6).

5. Local-nonlocal coupling. The final ingredient to be specified before we present our multiscale implementation strategy is how we effect local-nonlocal coupling within that strategy. Local-nonlocal coupling should occur only in regions where the local model is a valid model, i.e., in regions where the solution does not have discontinuities. A local-nonlocal coupling approach similar to ours was presented in [22] where patch test and wave propagation results are also included.

TABLE 2
Nonzero entries in typical rows of the stiffness matrices for the local and nonlocal models for several values of the grid size h. For the nonlocal model, the horizon $\delta = 0.02$ for all cases.

h	1/4	1/8
Local	$(-4, 8, -4)$	$(-8, 16, -8)$
Nonlocal	$(-0.04, -3.84, 7.76, -3.84, -0.04)$	$(-0.16, -7.36, 15.04, -7.36, -0.16)$

h	1/16	1/32
Local	$(-16, 32, -16)$	$(-32, 64, -32)$
Nonlocal	$(-0.64, -13.44, 28.16, -13.44, -0.64)$	$(-2.56, -21.76, 48.64, -21.76, -2.56)$

h	1/64
Local	$(-64, 128, -64)$
Nonlocal	$(-0.09, -5.27, -18.93, -3.32, 55.24, -3.32, -18.93, -5.27, -0.09)$

Let $\Omega_n \subset \Omega$ and $\Omega_l = \Omega\backslash\Omega_n$ denote the regions in which the nonlocal and local models, respectively, are applied. The local-nonlocal model we consider is given by

$$(10) \qquad \begin{cases} 2\displaystyle\int_{x-\delta}^{x+\delta} \big(u(x) - u(x')\big)\gamma(x, x')\, dx' = f(x) & \text{for } x \in \Omega_n, \\[2mm] \qquad\qquad -\dfrac{d^2u(x)}{dx^2} = f(x) & \text{for } x \in \Omega_l, \end{cases}$$

where the first equation is the nonlocal model in (2) restricted to one dimension. The system (10) should be supplemented with constraints. For simplicity, we consider only constraints applied to the solution $u(x)$, i.e., Dirichlet-type constraints. Wherever defects occur near the boundary $\partial\Omega$ of Ω, i.e., within a δ-neighborhood *interior* to the boundary, the nonlocal model has to be used, so we apply a volume constraint $u(x) = g(x)$ in a δ-neighborhood *exterior* to the boundary. If there are no defects near the boundary, one can use a local model there, and the condition $u(x) = g(x)$ needs be applied only at boundary points.

To motivate our strategy for the coupling of local and nonlocal models, we first provide some information about the stiffness matrix for the nonlocal model. A similar study is given in [22], including discussion about bandwidth changes with respect to different combinations of h and δ. For simplicity, we consider the integrable and nonsingular kernel (5) with $n = 1$ and $s = -1/2$; however, the results given in this subsection are qualitatively valid for the other kernels considered elsewhere in this paper. The entries of the stiffness matrix are determined using Legendre–Gauss quadrature rules. Table 2 shows, for several values of the grid size h, typical rows of the stiffness matrices for the local and nonlocal models, where for the latter the horizon is set to $\delta = 0.02$. Of course, in the one-dimensional setting, the local model stiffness matrix is the tridiagonal matrix $\frac{1}{h}(-1, 2, -1)$. We see that for $h > \delta$ the nonlocal force-based stiffness matrix (see section 1) is penta-diagonal, whereas for the one value of h smaller than δ, there are nine nonzero diagonals.

In Table 3, we again list the nonzero entries of the nonlocal stiffness matrix, but now do so for several values of δ with h fixed. We see that for $\delta > h$ the matrix is septa-diagonal, whereas for $\delta < h$, the matrix is penta-diagonal. Of more interest is to observe what occurs for $\delta \ll h$. We see that the entries in the outermost diagonals decrease as δ decreases and, in fact, those entries seem to decrease linearly with δ. The conclusion is that, as $\delta \to 0$, the entries of the nonlocal stiffness matrix converge to the entries of the local stiffness matrix. Of course, this is not surprising because

Nonzero entries in typical rows of the stiffness matrix for the nonlocal model for $h = 1/8$ and several values of the horizon δ. The typical row for the local model has entries $(-8, 16, -8)$.

δ	Nonzero entries in a typical row
1	$(-0.0469, -0.0469, 0.0781, 0.4531, 0.0781, -0.0469, -0.0469)$
0.5	$(-0.3594, -0.3750, 0.1250, 1.6250, 0.1250, -0.3750, -0.3594)$
0.2	$(-0.0316, -1.4734, -1.8215, 6.6531, -1.8215, -1.4734, -0.0316)$
0.1	$(-0.8000, -4.8000, 11.2000, -4.8000, -0.8000)$
0.01	$(-0.0800, -7.6800, 15.5200, -7.6800, -0.0800)$
0.001	$(-0.0080, -7.9680, 15.9520, -7.9680, -0.0080)$
0.0001	$(-0.0008, -7.9968, 15.9952, -7.9968, -0.0008)$

we have already mentioned that, as $\delta \to 0$, the nonlocal model (the first equation in (10)) converges to the local model (the second equation in (10)) and, moreover, the CL element is asymptotically compatible, as shown in [32, 33], so one expects that the nonlocal stiffness matrix, i.e., the discretized nonlocal operator, converges to the local one in that limit.

Our local-nonlocal coupling strategy is based on the following requirements: transitions from the nonlocal model to the local model occur only in regions where

1. the solution is smooth;
2. $h \gg \delta$, i.e., the local grid size is much greater than δ;
3. continuous finite element spaces are used for both the local and nonlocal models.

The first requirement is needed to ensure that the local model is valid in the transition region. The other two requirements ensure that the stiffness matrices for the local and nonlocal models are essentially the same; i.e., the differences in the entries of the two matrices are negligible with respect to the overall error incurred in discretization. Specifically, with respect to the second requirement, we see from, e.g., Table 3, that the difference between the nonlocal and local stiffness matrices seems to be of $O(\delta)$ as $\delta \to 0$. Thus, in a practical setting, if $\widetilde{h}$ denotes a desired grid size in regions where the solution is smooth, one should choose $\delta = O(\widetilde{h}^2)$ to ensure that the overall approximation error is of $O(\widetilde{h}^2)$.

The coupling strategy we use is very simple. Let $G_N = \{x_j\}_{j=1}^N$ denote the set of interior nodes of the grid in Ω. Then, let $G_{loc} \subset G_N$ denote the subset of "local" nodes; these should be chosen in regions of Ω in which the solution is smooth. Then, $G_{nonlocal} = G \backslash G_{loc}$ denotes the subset of "nonlocal" nodes. The row in the global stiffness matrix corresponding to a local (respectively, nonlocal) node is determined from the weak form of the local (respectively, nonlocal) model. For example, suppose $\Omega = (0, 1)$, $h = 1/8$, and $\delta = 0.001$; note that $\delta \ll h$ and that there are seven interior nodes in Ω. Let G_{loc} consist of the leftmost and rightmost pairs of interior nodes so that $G_{nonlocal}$ consists of the three central interior nodes. Using our coupling strategy, the resulting global stiffness matrix is then given by

$$\begin{pmatrix} 16 & -8 & 0 & 0 & 0 & 0 & 0 \\ -8 & 16 & -8 & 0 & 0 & 0 & 0 \\ -0.080 & -7.968 & 15.952 & -7.968 & -0.080 & 0 & 0 \\ 0 & -0.080 & -7.968 & 15.952 & -7.968 & -0.0800 & 0 \\ 0 & 0 & -0.080 & -7.968 & 15.952 & -7.968 & -0.080 \\ 0 & 0 & 0 & 0 & -8 & 16 & -8 \\ 0 & 0 & 0 & 0 & 0 & -8 & 16 \end{pmatrix}.$$

A crime is committed by our strategy, namely the loss of symmetry of the stiffness matrix. However, within our strategy, in particular because $\delta \ll h$, this crime does not appreciably affect the overall error of the approximations. For example, if we apply the patch test to our coupling scheme, i.e., we try to reproduce the solution $u = x$, the errors incurred as measured in any of the L^2, L^∞, or H^1 norms and for a range of grid sizes from 1/16 to 1/512 are on the level of machine precision. The loss of symmetry does engender increases in the cost of solving the discretized problem. One can devise several strategies for symmetrizing the discrete problem, but we do not concern ourselves with this issue because our central goal is to show the effectiveness of the multiscale implementation.

6. Multiscale implementation with discontinuity detection. We now have in hand all the ingredients needed to define the *multiscale implementation* of the one-dimensional PD model. We assume that the solution, i.e., the displacement, has a finite number of isolated jump discontinuities in the interior of the solution domain Ω. We also assume that an initial grid, possibly uniform, of size $O(h)$ is selected, as is a horizon δ satisfying $h^4 < \delta \ll h$.

The posterior error estimator enables one to identify elements in which discontinuities in the solution may occur and then provides a means for reducing the influence of discontinuities by drastically shrinking the size of the elements that contain them. The process is straightforward. If, for a particular element, the estimator is large (so that that element is suspected of containing a discontinuity) and the element is not of sufficiently small size, then it is refined, and the error indicator can be recalculated based on the newly created subelements. Those with small errors can be absorbed into a neighboring region of low-error elements in a coarsening procedure. Subelements with large error can be further refined as many times as necessary, until they become of size of $O(h^4)$. In this way, a mesh can be constructed which consists of relatively large $O(h)$ elements away from discontinuities and very small $O(h^4)$ elements covering the discontinuities. With the proper formulation of the solution procedure, the L^2 norm errors arising from the discontinuities will now be of the same order as that arising from smooth regions. Note that, at each step of the iterative grid refinement process, the current set of elements K is divided into two groups:

$\text{Group}(K) = 1 \quad \Leftarrow$ elements K for which the posterior error estimator is large;

$\text{Group}(K) = 2 \quad \Leftarrow$ elements K for which the posterior error estimator is small.

The elements of the initial grid can all be assigned to the first group. The refinement-coarsening process results in the elements in the second group having size $O(h)$, whereas the elements in the first group end up having size $O(h^4)$.

The decomposition of the computational grids into PD and PDE nodes, and into CG and DG nodes, is also straightforward. Nodes associated with the end points of an element which is suspected of containing a discontinuity in the solution, i.e., an element having a large posterior error estimator, are PD-DG nodes; i.e., the PD model with a DG discretization is used. Nodes which are a distance less than δ from a PD-DG node are PD-CG nodes; i.e., the PD model with a CG discretization is used. The remaining nodes are PDE-CG nodes; i.e., the PDE model with a CG discretization is used. Coupled with the grid refinement strategy described in the previous paragraph, at the end the elements consisting of PD-DG nodes will be of size $O(h^4)$ and contain all the discontinuities, whereas the remaining elements will be of size $O(h)$; i.e., we obtain a grid and a discretization distribution as sketched in Figure 1. In this way, the multiscale implementation can not only recover the asymptotic $O(h^2)$ L^2 norm error decay rate that would occur if the solution were smooth everywhere, but also

do so at a cost which is minimally greater than the cost that would be incurred if a piecewise-linear FEM for the PDE were applied throughout the region.

The meshing and discretization processes just outlined are incorporated into Algorithm 2, which is a modification of Algorithm 1. The grid size–weighted error estimator is used to detect discontinuities, which is the main feature with which we are concerned. In line 20 of Algorithm 2, refinement is effected by dividing each Group 1 element into two elements of equal length. In line 16 of the algorithm, coarsening is effected as follows: The Group 2 elements form subintervals in Ω separated by subintervals containing Group 1 elements; each subinterval containing Group 2 elements is divided into a uniform mesh of $\lceil L/h \rceil$ elements, where L denotes the length of the subinterval.

Algorithm 2: Multiscale mesh refinement.

    Input: Dörfler marking parameter $\theta \in (0,1]$, a grid size h, a uniform
               triangulation $\mathcal{T}_0$ of Ω into elements of size h, and a set $\mathcal{P}_0$ of nodes
               consisting of $\mathcal{T}_0$.

1 Initialization: $k = 0$;

2 **foreach** *element K in $\mathcal{T}_0$* **do** set Group$(K) = 1$;

3 **while** *true* **do**

4   **foreach** *node p in $\mathcal{P}_k$, the set of nodes consisting of $\mathcal{T}_k$* **do**

5   **if** *p is an end point of a Group 1 element K,* **then** set p as a PD-DG node;

6   **else if** *p is an end point of an element within the δ-neighborhood of any Group 1 element,* **then** set p as a PD-CG node;

7   **else** set p as a PDE-CG node;

8   **end**

9   using the triangulation $\mathcal{T}_k$, solve the multiscale problem for $u_k^h(x)$;

10   **foreach** *element K in $\mathcal{T}_k$* **do** compute the error indicator $\eta(u_k^h, K)$;

11   define the set $\mathcal{M}_k$ of elements contributing the proportion θ of the total error, i.e., $\eta^2(u_k^h, \mathcal{M}_k) \geq \theta\eta^2(u_k^h, \Omega)$, with a minimal cardinality;

12   **foreach** *element K in $\mathcal{T}_k$,* **do**

13   **if** $K \in \mathcal{M}_k$, **then** Group$(K) = 1$;

14   **else** $Group(K) = 2$;

15   **end**

16   create triangulation $\mathcal{T}_{k+1}$ of the Group 2 elements with coarsened $O(h)$ mesh;

17   **if** *all Group 1 elements have size $O(h^4)$,* **then**

18   do line 3–9 again with the coarsened final grid and then break;

19   **end**

20   modify triangulation $\mathcal{T}_{k+1}$ by refining Group 1 elements;

21   set $k := k + 1$;

22 **end**

---

7. Numerical experiments. In this section, we use a series of computational experiments to study several aspects of the multiscale implementation described in section 6. Continuous (CG) and discontinuous (DG) Galerkin methods are compared, first using uniform grid refinement and then the adaptive refinement method. Sequences of decreasing grid sizes are used, allowing convergence rates to be estimated.

We consider kernels of the form (5) with $n = 1$ and for three choices of s, as given in Table 4. The scaling constants that appear in the kernels and which depend on the horizon δ are such that, in each case, as $\delta \to 0$, the nonlocal model (the first equation in (10)) reduces to the local one (the second equation in (10)).

TABLE 4
The values of s and the corresponding kernels $\gamma(x, x')$ used in the numerical examples.

Case	s	$\gamma(x, x')$	Kernel properties
I	$-1/2$	$\dfrac{3}{2\delta^3} \mathbb{1}_{H_{\mathbf{x},\delta}}$	Integrable, nonsingular
II	$-1/4$	$\dfrac{5}{4\delta^{5/2}} \dfrac{1}{|x - x'|^{1/2}} \mathbb{1}_{H_{\mathbf{x},\delta}}$	Integrable, singular
III	0	$\dfrac{1}{\delta^2} \dfrac{1}{|x - x'|} \mathbb{1}_{H_{\mathbf{x},\delta}}$	Nonintegrable

7.1. Uniform grid refinement for a solution having a discontinuity. In this subsection, we use uniform grids and apply the nonlocal model and either the CG or DG discretization throughout. It is natural to expect that the performance of the posterior error estimator may differ, depending on whether the CG or DG method is being used. We consider the nonlocal problem given by (2) with $\gamma(x, x')$ given by Case I in Table 4. The manufactured solution used is given by

(11)
$$u(x) = \begin{cases} x & \text{if } x < \widetilde{x}, \\ x^2 & \text{otherwise,} \end{cases}$$

where the parameter $\widetilde{x} \in (0, 1)$ allows the location of the discontinuity to be varied for different experiments. From this exact solution, the volume constraint data is given by $g(x) = u(x)$, and the right-hand side $f(x)$ is given by

$$f(x) = \begin{cases} 0 & \text{if } x \leq \widetilde{x} - \delta, \\ \dfrac{3}{\delta^3}\left(2\delta x + \dfrac{\widetilde{x}^3}{3} - \dfrac{\widetilde{x}^2}{2} + \dfrac{(x - \delta)^2}{2} - \dfrac{(x + \delta)^3}{3}\right) & \text{if } x \in (\widetilde{x} - \delta, \widetilde{x}), \\ \dfrac{3}{\delta^3}\left(2\delta x^2 + \dfrac{\widetilde{x}^3}{3} - \dfrac{\widetilde{x}^2}{2} + \dfrac{(x - \delta)^2}{2} - \dfrac{(x + \delta)^3}{3}\right) & \text{if } x \in (\widetilde{x}, \widetilde{x} + \delta), \\ -2 & \text{if } x \geq \widetilde{x} + \delta. \end{cases}$$

The error behavior is quite different depending on whether the point $\widetilde{x}$ at which the discontinuity in the solution occurs coincides or does not coincide with a grid point, i.e., at the interface between two elements. Therefore, two sequences of element refinements are considered for a problem in which the discontinuity occurs at $\widetilde{x} = 0.5$. For the sequence of an even number of elements, the discontinuity point occurs at an element interface, whereas for the odd sequence, the discontinuity point occurs at the center of an element. A third study again uses a sequence of an even number of elements, but with the discontinuity now located at $\widetilde{x} = 0.503$, so that it now occurs within a single element but not at the center of that element; the last case is what one would expect in practice where one does not know beforehand where the discontinuity in the solution occurs.

Table 5 and Figure 2 correspond to the CG method and to $\delta = 0.02$. Table 5 considers the errors and convergence rates with respect to the L^2 and energy norms, the posterior error estimate and its convergence rate, as well as several of the relative posterior error measures as defined in (9). For both cases for which the discontinuity

TABLE 5

For the exact solution (11) *with* $\delta = 0.02$, *the nonsingular kernel of Case* I *in Table* 4, *and CG approximations with different uniform grid sizes h, errors and convergence rates with respect to the* L^2 *and energy norms, the posterior error estimate and its convergence rate (CR), and relative error measures as defined in* (9).

Even number of elements, discontinuity located at a grid point									
h	$\|e^h\|_{L^2}$	CR	$\|\|e^h\|\|$	CR	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$	$\widetilde{E}^2$	$\widetilde{E}^4$
1/4	5.44e-2	–	2.15e+0	–	1.77e+0	–	0.506	0.999	1.000
1/8	3.63e-2	0.583	2.11e+0	0.026	1.77e+0	-0.001	0.502	0.999	1.000
1/16	2.51e-2	0.532	2.04e+0	0.047	1.78e+0	-0.005	0.499	0.995	1.000
1/32	1.73e-2	0.534	1.90e+0	0.108	1.80e+0	-0.022	0.489	0.978	1.000
1/64	1.19e-2	0.544	1.50e+0	0.337	1.56e+0	0.213	0.448	0.896	0.992
1/128	8.40e-3	0.502	1.03e+0	0.550	1.02e+0	0.603	0.464	0.929	0.995
1/256	5.94e-3	0.500	7.27e-1	0.496	7.27e-1	0.495	0.464	0.928	0.994
1/512	4.20e-3	0.500	5.14e-1	0.500	5.14e-1	0.500	0.464	0.928	0.995

Odd number of elements, discontinuity located at the midpoint of an element								
h	$\|e^h\|_{L^2}$	CR	$\|\|e^h\|\|$	CR	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$	$\widetilde{E}^3$
1/5	3.25e-2	–	2.10e+0	–	1.77e+0	–	0.998	1.000
1/9	2.36e-2	0.541	2.03e+0	0.059	1.78e+0	-0.005	0.994	1.000
1/17	1.69e-2	0.529	1.88e+0	0.121	1.80e+0	-0.020	0.978	1.000
1/33	1.20e-2	0.514	1.50e+0	0.340	1.54e+0	0.236	0.916	0.999
1/65	8.53e-3	0.504	1.04e+0	0.535	1.04e+0	0.576	0.934	0.990
1/129	6.06e-3	0.500	7.42e-1	0.496	7.43e-1	0.495	0.935	0.993
1/257	4.29e-3	0.500	5.26e-1	0.501	5.26e-1	0.502	0.936	0.995
1/513	3.04e-3	0.500	3.72e-1	0.500	3.72e-1	0.500	0.936	0.995

Even number of elements, discontinuity located within an element									
h	$\|e^h\|_{L^2}$	CR	$\|\|e^h\|\|$	CR	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$	$\widetilde{E}^2$	$\widetilde{E}^4$
1/4	5.83e-2	–	2.15e+0	–	1.77e+0	–	0.696	0.999	1.000
1/8	3.83e-2	0.607	2.11e+0	0.028	1.76e+0	0.006	0.705	0.998	1.000
1/16	2.59e-2	0.567	2.02e+0	0.059	1.75e+0	0.007	0.721	0.994	1.000
1/32	1.71e-2	0.595	1.84e+0	0.133	1.76e+0	-0.003	0.749	0.972	1.000
1/64	1.07e-2	0.673	1.35e+0	0.454	1.39e+0	0.340	0.773	0.884	0.989
1/128	6.38e-3	0.752	7.81e-1	0.785	7.81e-1	0.828	0.921	0.993	0.999
1/256	5.18e-3	0.299	6.35e-1	0.299	6.35e-1	0.200	0.847	0.928	0.994
1/512	3.05e-3	0.764	3.74e-1	0.764	3.74e-1	0.764	0.935	0.979	0.998

point is located at $\widetilde{x} = 0.5$, we see that the expected convergence rate tends to $1/2$ for the L^2 norm and for the energy norm; for the energy norm, the asymptotic convergence rate is not achieved until $h < \delta$. When the discontinuity is located at $\widetilde{x} = 0.503$, the convergence behavior is much more erratic because of the use of a uniform grid for the discontinuous solution, thus making the error highly dependent on the position of the discontinuity point relative to the ends of the element in which it is located. However, even in this case, the convergence rate tends, in some average sense, to $1/2$. Also note that for all three cases the error estimator convergence rates mimic the exact error convergence rates.

In Figure 2, we show how the error is distributed across the elements in $\Omega = (0, 1)$
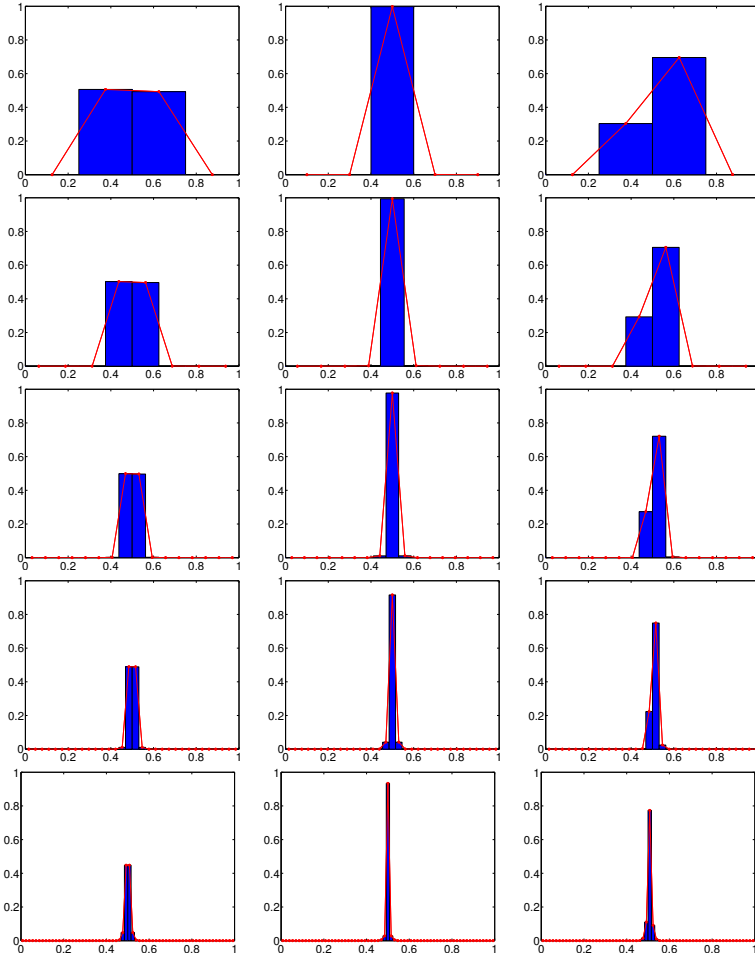
FIG. 2. *For the exact solution* (11) *with* $\delta = 0.02$, *the nonsingular kernel of Case* I *in Table 4, and CG approximations with different uniform grid sizes h, the relative posterior error measures* $\widetilde{\eta}^2(K)/\widetilde{\eta}^2(\Omega)$ *for each element K. Left:* $\widetilde{x} = 0.5$, *and the number of elements is even with the grid size ranging from* $1/4$ *to* $1/64$ *(top to bottom), in which case the discontinuity in the solution is located at a grid point. Center:* $\widetilde{x} = 0.5$, *and the number of elements is odd with the grid size ranging from* $1/5$ *to* $1/65$ *(top to bottom), in which case the discontinuity is located at the midpoint of an element. Right:* $\widetilde{x} = 0.503$, *and the number of elements is even with the grid size ranging from* $1/4$ *to* $1/64$ *(top to bottom), in which case the discontinuity is located within an element but not at the midpoint.*

by plotting relative posterior error measures $\widetilde{\eta}^2(K)/\widetilde{\eta}^2(\Omega)$ for each element. The left column corresponds to the case for which the discontinuity in the solution is located at a grid point; we see that the two elements meeting at this node account for the majority ($> 90\%$) of the error, and the four nearest elements account for more than 99% of the total squared error. The middle column corresponds to the case for which the discontinuity occurs at the midpoint of an element, and that element alone accounts for more than 93% of error, whereas 99% of the error occurs in that element and its two immediate neighbors. The right column corresponds to the case for which the discontinuity occurs within an element but not at its center; the error shows a similar concentration near where the discontinuity occurs, but with a noticeable

TABLE 6
Same information as for Table 5, but now for DG approximations.

Even number of elements, discontinuity located at a grid point							
h	$\|e^h\|_{L^2}$	CR	$\|\|\|e^h\|\|\|$	CR	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$
1/4	7.84e-3	–	9.57e-2	–	3.64e-2	–	0.515
1/8	1.92e-3	2.030	4.48e-2	1.094	2.46e-2	0.564	0.255
1/16	4.62e-4	2.056	1.94e-2	1.207	1.55e-2	0.669	0.126
1/32	1.10e-4	2.065	6.81e-3	1.512	7.36e-3	1.071	0.063
1/64	2.69e-5	2.040	1.49e-3	2.194	1.41e-3	2.389	0.035
1/128	6.70e-6	2.004	3.98e-4	1.903	4.02e-4	1.807	0.016
1/256	1.67e-6	2.002	9.75e-5	2.029	9.65e-5	2.058	0.008
1/512	4.18e-7	1.999	2.45e-5	1.995	2.43e-5	1.990	0.004

Odd number of elements, discontinuity located at the midpoint of an element							
h	$\|e^h\|_{L^2}$	CR	$\|\|\|e^h\|\|\|$	CR	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$
1/5	3.20e-2	–	2.10e+0	–	1.78e+0	–	0.999
1/9	2.29e-2	0.570	2.02e+0	0.065	1.79e+0	-0.011	0.999
1/17	1.58e-2	0.581	1.83e+0	0.150	1.82e+0	-0.032	0.998
1/33	1.09e-2	0.561	1.33e+0	0.484	1.33e+0	0.473	0.995
1/65	7.76e-3	0.504	9.49e-1	0.499	9.50e-1	0.501	0.998
1/129	5.50e-3	0.501	6.74e-1	0.500	6.74e-1	0.500	0.999
1/257	3.90e-3	0.500	4.78e-1	0.500	4.78e-1	0.500	1.000
1/513	2.76e-3	0.500	3.38e-1	0.500	3.38e-1	0.500	1.000

Even number of elements, discontinuity located within an element							
h	$\|e^h\|_{L^2}$	CR	$\|\|\|e^h\|\|\|$	CR	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$
1/4	3.05e-2	–	1.52e+0	–	1.48e+0	–	0.990
1/8	1.99e-2	0.620	1.48e+0	0.040	1.44e+0	0.042	0.991
1/16	1.49e-2	0.419	1.46e+0	0.021	1.43e+0	0.010	0.992
1/32	1.17e-2	0.348	1.37e+0	0.088	1.38e+0	0.048	0.994
1/64	9.03e-3	0.373	1.10e+0	0.318	1.10e+0	0.324	0.997
1/128	5.80e-3	0.637	7.11e-1	0.633	7.10e-1	0.634	1.000
1/256	4.50e-3	0.366	5.51e-1	0.366	5.51e-1	0.366	1.000
1/512	2.77e-3	0.699	3.40e-1	0.698	3.40e-1	0.698	1.000

asymmetry. In all cases, however, we see that whereas the CG method is believed to "spread" errors, the contamination is mainly confined to the immediate neighbors of the element containing the discontinuity.

Table 6 and Figure 3 correspond to the same information as in Table 5 and Figure 2, respectively, but now for DG approximations. We now see, from Table 6, that if the discontinuity of the solution is located at a grid point, i.e., at an end point of an element, that the optimal convergence rates $O(h^2)$ are achieved. However, for the two cases for which the discontinuity occurs in the interior of an element, the convergence rate drops to $O(h^{1/2})$ (in an average sense for the latter). These behaviors are expected; see, e.g., [7]. However, Figure 3 and the last column of Table 6 illustrate an important feature of DG methods for the nonlocal problem we consider, namely that *when the discontinuity in the solution is located within an*
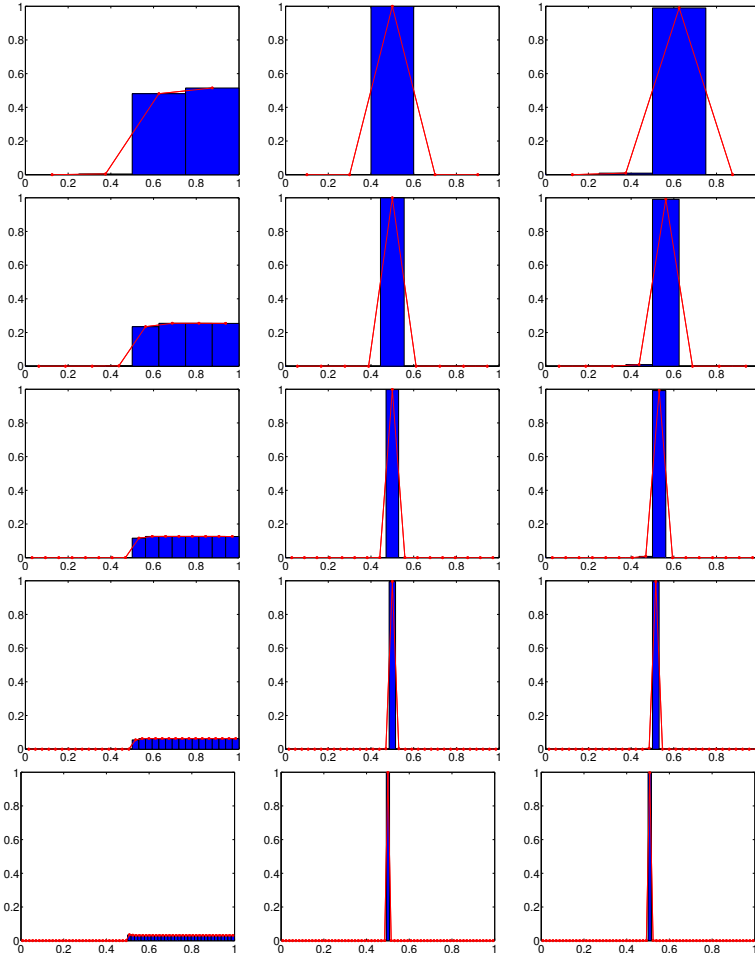
FIG. 3. *Same information as for Figure* 2*, but now for DG approximations.*

element, almost all the error will occur within that single element, with almost no pollution effects to neighboring elements. The multiscale implementation uses the DG method in elements that contain a discontinuity in the solution precisely because of its ability to localize discontinuity errors in this manner.

It is worth mentioning that, from Table 6, the convergence rates of $1/2$ or 2, as the case may be, of the L^2 error are observed for all values of h, whereas for the energy error they are observed only for $h < \delta$. Even for $h > \delta$, the L^2 convergence behavior can be explained by the theoretical results in [33] that established the convergence of the numerical solution of the nonlocal problem to that of the local problem in the L^2 norm regardless of the relative sizes of h and δ, so long as both are getting smaller.

7.2. Adaptive refinement for a nonsingular kernel. We now consider the adaptive Algorithm 2, applied to the problem in (11), with the discontinuity in the solution located at $\widetilde{x} = 0.503$ and with $\delta = 0.02$. The kernel $\gamma(x, x')$ is again given by Case I in Table 4. Note that initial uniform grids with sizes h_0 greater and smaller than δ are considered.

In Table 7, information is provided about the grid obtained at the end of the

For the exact solution (11) *with* $\delta = 0.02$, $\widetilde{x} = 0.503$, *and the nonsingular kernel of Case* I *in Table* 4 *and for several initial grids, the distribution of node types and the maximal and minimal grid sizes of the final grid resulting from the application of the adaptive Algorithm* 2.

N_0^n	N_f^n	N_{dg}^{pd}	N_{cg}^{pd}	N_{cg}^{pde}	h_0	$\overline{h}$	$\underline{h}$	$\overline{h}/h_0$	$\underline{h}/h_0^4$	$\underline{h}/\overline{h}^4$
5	6	2	2	2	2.50e-1	2.50e-1	3.91e-03	1.000	1.000	1.000
9	11	2	2	7	1.25e-1	1.24e-1	2.44e-04	0.994	1.000	1.026
17	19	2	2	15	6.25e-2	6.21e-2	1.53e-05	0.994	1.000	1.024
33	35	2	2	31	3.13e-2	3.11e-2	9.54e-07	0.994	1.000	1.024
65	67	2	4	61	1.56e-2	1.55e-2	5.96e-08	0.994	1.000	1.024
129	131	2	6	123	7.81e-3	7.77e-3	3.73e-09	0.994	1.000	1.024
257	259	2	12	245	3.91e-3	3.90e-3	2.33e-10	0.998	1.000	1.007

adaptive Algorithm 2. In the table, N_0^n and N_f^n refer to the number of nodes in the initial uniform mesh and final nonuniform mesh, respectively; h_0 refers to the initial uniform grid size, $\overline{h}$ the size of the longest element in the final grid, and $\underline{h}$ the size of the shortest element in the final grid; N_{dg}^{pd}, N_{cg}^{pd}, and N_{cg}^{pde} refer to the number of nodes at the end of the algorithm at which the nonlocal model and the DG method, the nonlocal model and the CG method, and the local model and the CG method, respectively, are applied. From the table, we see that, for all initial grids, the final grid contains at most two more nodes than does the initial grid. We also see that the nonlocal model with the DG approximation is applied at only two nodes (which turn out to be the two nodes of the element containing the discontinuity in the solution); the nonlocal model with the CG approximation is applied at only a few nodes (which turn out to be nodes of the elements within a δ-neighborhood of the element containing the discontinuity); at all but a few nodes, the local model with the CG approximation is applied. Thus, we see the efficiency resulting from the application of Algorithm 2: Very little cost is incurred over that of using a local model with a CG discretization everywhere. From Table 7, we also see that, to the number of significant figures used in the table, $\overline{h}/h_0 \in [0.994, 1.000]$, $\underline{h}/h_0^4 \approx 1.000$, and $\underline{h}/\overline{h}^4 \in [1.000, 1.026]$. Thus, we also see the effectiveness of Algorithm 2 in isolating the discontinuity in the solution into a single element of size $\approx h_0^4$ while effecting an abrupt transition to the remaining elements, all of which are of size $\approx h_0$.

In Table 8, for different initial number of nodes N_0^n, results for errors and convergence rates with respect to the L^2, L^∞, and energy norms are provided, as is the posterior error estimate and the values of $\widetilde{E}^1$ and $\overline{E}^1$, the maximum value of the elementwise relative error and its weighted version, respectively. The second-order convergence behavior with respect to the L^2 and energy norms as well as of the posterior error is evident from the table. However, no convergence is achieved with respect to the L^∞ norm of the error. This is all expected from the discussion of section 3. Because the DG method is expected to isolate a substantial portion of the error into the element containing the discontinuity, a separate set of results are reported in Table 8 which omit the errors associated with that element. The L^∞ and H^1 norms of the error now exhibit the optimal $O(h^2)$ and $O(h)$ convergence behaviors, respectively. Note that because the CG method is used outside of the element containing the discontinuity, the H^1 norm is well defined when errors are calculated omitting that element.

In Figure 4, the approximate solution, as determined using Algorithm 2, is com-

TABLE 8

For the exact solution (11) *with* $\delta = 0.02$, $\widetilde{x} = 0.503$, *and the nonsingular kernel, errors, posterior error indicators, maximal elementwise relative errors, and convergence rates (CR), resulting from the application of the adaptive Algorithm 2.*

Errors and convergence rates including all elements											
N_0^n	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$\|\|\|e^h\|\|\|$	CR	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$	E^1
5	8.74e-3	–	1.29e-1	–	5.59e-1	–	5.53e-1	–	0.99	1.00
9	2.19e-3	1.98	1.37e-1	-0.08	1.41e-1	1.97	1.37e-1	2.00	0.97	1.00
17	5.24e-4	2.06	1.35e-1	0.02	3.66e-2	1.95	3.47e-2	1.98	0.81	1.00
33	1.29e-4	2.01	1.35e-1	0.00	1.08e-2	1.76	1.12e-2	1.64	0.58	1.00
65	3.12e-5	2.05	1.38e-1	-0.03	2.49e-3	2.12	2.44e-3	2.19	0.68	1.00
129	7.87e-6	1.99	1.37e-1	0.01	6.32e-4	1.98	6.35e-4	1.94	0.61	1.00
257	1.92e-6	2.05	1.64e-1	-0.26	1.56e-4	2.03	1.55e-4	2.04	0.63	1.00

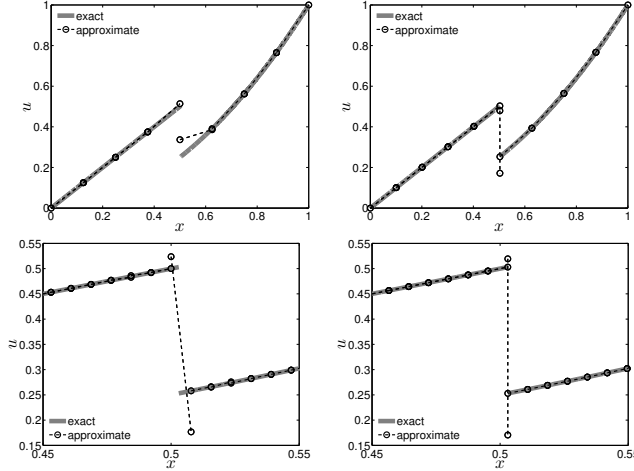| Errors excluding the element containing the discontinuity						
N_0^n	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$|e^h|_{H^1}$	CR
5	7.49e-3	–	1.50e-2	–	1.01e-1	–
9	1.89e-3	1.97	3.82e-3	1.96	5.06e-2	0.99
17	4.58e-4	2.05	9.57e-4	2.00	2.53e-2	1.00
33	1.09e-4	2.07	2.39e-4	2.00	1.26e-2	1.00
65	2.66e-5	2.04	6.00e-5	2.00	6.32e-3	1.00
129	6.75e-6	1.98	1.50e-5	2.00	3.16e-3	1.00
257	1.64e-6	2.06	3.76e-6	2.01	1.58e-3	1.01

FIG. 4. *For the exact solution* (11) *with* $\delta = 0.02$, $\widetilde{x} = 0.503$, *the initial grid size* $h_0 = 1/8$ *(top) and* $h_0 = 1/128$ *(bottom), and the nonsingular kernel, the approximate solution at the initial (left) and final (right) step for the adaptive multiscale implementation of Algorithm 2. The plots for* $h_0 = 1/128$ *are zoomed into the vicinity of the discontinuity.*

pared to the exact solution (11) at the initial and final steps of that algorithm; the initial grid sizes $h_0 = 1/8$ and $h_0 = 1/128$ are considered. We observe the substantial improved accuracy of the approximation at the final step and also the confinement of the error to the element containing the discontinuity.
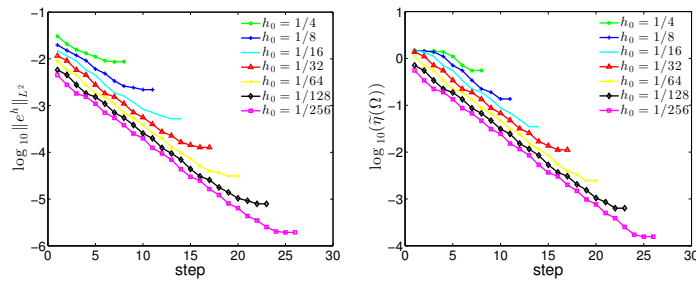
Fig. 5. *For the exact solution* (11) *with* $\delta = 0.02$, $\widetilde{x} = 0.503$, *and the nonsingular kernel, the* L^2 *error (left) and posterior error indicator* $\widetilde{\eta}$ *(right) at each step in the adaptive multiscale implementation of Algorithm* 2.

Figure 5 illustrates the steady decrease of the L^2 error and the posterior error indicator as the multiscale implementation of Algorithm 2 proceeds, as well as the decrease in error with decreasing initial grid size h_0.

7.3. Adaptive refinement for a singular but integrable kernel. We now consider the singular but integrable kernel $\gamma(x, x')$ given by Case II in Table 4, for which the results are given in Table 9. For this case, the entries in Table 7 carry over almost verbatim, so that we do not bother providing that information here. Just as for the case of the nonsingular kernel (Case I), optimal convergence rates are achieved, although for the L^∞ and H^1 norm errors it is again necessary to omit the element containing the discontinuity in the solution. Also, as in the case of the nonsingular kernel, it is clear that by examining the results for E^1, the single element containing the discontinuity again accounts for almost the entire error during most of the refinement process, so that only that single element is refined. For this case, the figure providing the information analogous to Figure 5 would be entirely similar so that again, for the sake of brevity, it is omitted here.

7.4. Adaptive refinement for a nonintegrable kernel. We next consider the nonintegrable kernel $\gamma(x, x')$ given by Case III in Table 4, for which the results are given in Table 10. Recall that this kernel corresponds to one-dimensional PD. The convergence rates with respect L^2 and L^∞ norms are somewhat more erratic than those for Cases I and II. This is probably due to the stronger singularity of the kernel for Case III, which requires, in order to preserve full accuracy, alternative quadrature rules to those we use to assemble the stiffness matrix for the other cases. However, our results seem to follow the optimal rates in an average sense, at least for the L^2 norm.

7.5. Adaptive refinement for a solution having two discontinuities. We next consider the manufactured solution,

$$(12) \qquad u(x) = \begin{cases} x & \text{for } x < \widetilde{x}_1, \\ x^2 & \text{for } x \in (\widetilde{x}_1, \widetilde{x}_2), \\ x & \text{for } x > \widetilde{x}_2, \end{cases}$$

that is discontinuous at two distinct points in $\Omega = (0, 1)$. Specifically, we set $\widetilde{x}_1 = 0.253$ and $\widetilde{x}_2 = 0.753$ so that, in general, neither point will coincide with a grid point. Case I kernel in Table 4 is used, and the Dörfler marking parameter is set to be $\theta = 0.9$. Solutions are computed for a sequence of initial grid sizes h_0, with results

TABLE 9
For the exact solution (11) *with* $\delta = 0.02$, $\widetilde{x} = 0.503$ *and the singular but integrable kernel of Case* II *in Table* 4 *and for several initial grids, errors, convergence rates (CR), and grid information.*

Errors and convergence rates including all elements							
N_0^n	N_f^n	$\underline{h}$	$\overline{h}$	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$	E^1
5	6	3.91e-03	2.50e-1	6.51e-1	–	1.00	1.00
9	11	2.44e-04	1.24e-1	1.71e-1	1.91	0.99	1.00
17	19	1.53e-05	6.21e-2	4.18e-2	2.03	0.91	1.00
33	35	9.54e-07	3.11e-2	1.26e-2	1.73	0.77	1.00
65	67	5.96e-08	1.55e-2	2.98e-3	2.07	0.77	1.00
129	131	3.73e-09	7.77e-3	7.70e-4	1.95	0.69	1.00
257	259	2.33e-10	3.90e-3	1.95e-4	1.99	0.68	1.00

N_0^n	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$\|\|e^h\|\|$	CR
5	8.79e-3	–	1.28e-1	–	6.86e-1	–
9	2.20e-3	1.98	1.35e-1	-0.08	1.77e-1	1.94
17	5.25e-4	2.06	1.33e-1	0.02	4.45e-2	1.99
33	1.29e-4	2.03	1.35e-1	-0.02	1.30e-2	1.77
65	3.13e-5	2.04	1.38e-1	-0.03	3.11e-3	2.07
129	7.87e-6	1.99	1.35e-1	0.03	7.89e-4	1.98
257	1.94e-6	2.02	1.64e-1	-0.28	1.99e-4	2.00

Errors excluding the element containing the discontinuity						
N_0^n	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$|e^h|_{H^1}$	CR
5	7.55e-3	–	1.50e-2	–	1.01e-1	–
9	1.90e-3	1.97	3.83e-3	1.96	5.06e-2	0.99
17	4.59e-4	2.05	9.57e-4	2.00	2.53e-2	1.00
33	1.08e-4	2.09	2.39e-4	2.00	1.26e-2	1.00
65	2.67e-5	2.02	6.00e-5	2.00	6.32e-3	1.00
129	6.76e-6	1.98	1.50e-5	2.00	3.16e-3	1.00
257	1.66e-6	2.03	3.76e-6	2.00	1.58e-3	1.00

reported in Tables 11 and 12. Similar to the previous cases, optimal convergence rates are obtained. Because of the presence of two discontinuities, the values of cumulative elementwise relative errors are of interest. In particular, the values for E^1 and E^2 listed in the table indicate that almost all the error is confined to the two elements that contain a discontinuity in the solution. Of course, because we now have two discontinuities, we have two or three grid points added by the refinement process, in contrast to the case of a single discontinuity, for which only one or two points are added.

In Figure 6, the exact solution is displayed and compared to the first and last approximate solutions obtained with initial grid sizes of $h_0 = 1/8$ and $h_0 = 1/128$. As was the case for Figure 4, we observe the substantial improved accuracy of the approximation at the final step and also the confinement of the error to the two elements containing the two discontinuities.

Figure 7 illustrates the refinement process for the initial grid size $h_0 = 1/4$. Points are added to shrink the size of the elements that contain the discontinuities until the algorithm accepts them as being tiny enough; also, some grid points are removed in

TABLE 10

For the exact solution (11) with $\delta = 0.02$, $\widetilde{x} = 0.503$, and the singular kernel of Case III in Table 4 and for several initial grids, errors, convergence rates (CR), and grid information.

Errors and convergence rates including all elements							
N_0^n	N_f^n	$\underline{h}$	$\overline{h}$	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$	E_1
5	6	3.91e-03	2.50e-1	2.19e+2	–	1.00	1.00
9	11	2.44e-04	1.24e-1	8.27e+1	1.39	1.00	1.00
17	19	1.53e-05	6.21e-2	2.66e+1	1.64	0.99	1.00
33	35	9.54e-07	3.11e-2	9.03e+0	1.56	0.98	1.00
65	67	5.96e-08	1.55e-2	2.62e+0	1.79	0.98	1.00
129	131	3.73e-09	7.77e-3	7.90e-1	1.73	0.91	1.00
257	259	2.33e-10	3.90e-3	2.65e-1	1.59	0.67	1.00

N_0^n	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$\|\|e^h\|\|$	CR
5	9.00e-3	–	1.32e-1	–	9.84e-1	–
9	2.20e-3	2.01	1.37e-1	-0.05	3.03e-1	1.68
17	5.27e-4	2.06	1.33e-1	0.04	8.43e-2	1.85
33	1.51e-4	1.81	1.33e-1	0.00	2.60e-2	1.70
65	3.78e-5	2.00	1.37e-1	-0.05	6.83e-3	1.93
129	1.14e-5	1.73	1.80e-1	-0.39	1.83e-3	1.88
257	2.18e-6	2.41	2.03e-1	-0.17	4.97e-4	1.89

Errors excluding the element containing the discontinuity						
N_0^n	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$|e^h|_{H^1}$	CR
5	7.79e-3	–	1.53e-2	–	1.01e-1	–
9	1.91e-3	2.01	3.83e-3	1.98	5.06e-2	1.00
17	4.62e-4	2.05	9.57e-4	2.00	2.53e-2	1.00
33	1.34e-4	1.79	2.47e-4	1.96	1.26e-2	1.00
65	3.38e-5	1.99	6.03e-5	2.04	6.32e-3	1.00
129	1.03e-5	1.71	1.78e-5	1.65	3.16e-3	1.00
257	1.56e-6	2.73	5.72e-6	1.65	1.59e-3	1.00

TABLE 11

For the exact solution (12) with $\delta = 0.02$, $\widetilde{x}_1 = 0.253$, $\widetilde{x}_2 = 0.753$, and the nonsingular kernel of Case I in Table 4 and for several initial grids, the distribution of node types and the maximal and minimal grid sizes of the final grid resulting from the application of the adaptive Algorithm 2.

N_0^n	N_f^n	N_{dg}^{pd}	N_{cg}^{pd}	N_{cg}^{pde}	h_0	$\overline{h}$	$\underline{h}$	$\overline{h}/h_0$	$\underline{h}/h_0^4$	$\underline{h}/\overline{h}^4$
5	7	4	3	0	2.50e-1	2.50e-1	3.91e-03	1.000	1.000	1.000
9	12	4	4	4	1.25e-1	1.25e-1	2.44e-04	1.000	1.000	1.000
17	20	4	4	12	6.25e-2	6.25e-2	1.53e-05	1.000	1.000	1.000
33	36	4	4	28	3.13e-2	3.13e-2	9.54e-07	1.000	1.000	1.000
65	68	4	8	56	1.56e-2	1.56e-2	5.96e-08	1.000	1.000	1.000
129	132	4	12	116	7.81e-3	7.81e-3	3.73e-09	1.000	1.000	1.000
257	260	4	24	232	3.91e-3	3.90e-3	2.33e-10	1.000	1.000	1.000

later steps to coarsen the elements not containing the discontinuities so that, away from solution discontinuities, the grid size remains very close to that of the initial grid.

TABLE 12
For the exact solution (12) *with $\delta = 0.02$, $\widetilde{x}_1 = 0.253$, $\widetilde{x}_2 = 0.753$, and the nonsingular kernel of Case* I *in Table 4, errors, posterior error indicators, maximal elementwise relative errors, and convergence rates (CR) resulting from the application of the adaptive Algorithm 2.*

Errors and convergence rates including all elements						
N_0^n	$\widehat{\eta}(\Omega)$	CR	$\widetilde{E}^1$	$\widetilde{E}^2$	E^1	E^2
5	5.85e-1	–	0.51	1.00	0.51	1.00
9	1.44e-1	2.02	0.49	0.97	0.51	1.00
17	3.64e-2	1.99	0.42	0.82	0.51	1.00
33	1.16e-2	1.65	0.31	0.60	0.51	1.00
65	2.56e-3	2.18	0.35	0.70	0.51	1.00
129	6.60e-4	1.95	0.32	0.63	0.51	1.00
257	1.62e-4	2.03	0.33	0.64	0.51	1.00
N_0^n	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$\|\|\|e^h\|\|\|$	CR
5	8.53e-3	–	9.56e-2	–	5.92e-1	–
9	2.03e-3	2.07	1.02e-1	-0.09	1.49e-1	1.99
17	4.25e-4	2.26	1.01e-1	0.02	3.82e-2	1.96
33	9.87e-5	2.11	1.02e-1	-0.02	1.13e-2	1.76
65	2.17e-5	2.19	1.04e-1	-0.03	2.60e-3	2.12
129	5.44e-6	2.00	1.02e-1	0.03	6.58e-4	1.98
257	1.33e-6	2.03	1.23e-1	-0.28	1.62e-4	2.02

Errors excluding the two elements containing the discontinuities						
N_0^n	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$|e^h|_{H^1}$	CR
5	7.07e-3	–	1.41e-2	–	1.01e-1	–
9	1.67e-3	2.08	3.38e-3	2.06	5.10e-2	0.98
17	3.29e-4	2.34	6.92e-4	2.29	2.55e-2	1.00
33	6.58e-5	2.32	1.41e-4	2.30	1.28e-2	1.00
65	1.29e-5	2.35	4.09e-5	1.78	6.38e-3	1.00
129	3.35e-6	1.95	1.17e-5	1.81	3.19e-3	1.00
257	8.10e-7	2.05	2.69e-6	2.12	1.59e-3	1.00

8. Distinguishing between steep gradients and jump discontinuities. Besides the occurrence of discontinuities, it is possible that the solution also has locally steep gradients, which also result in large errors when coarse grid approximations are used. Thus, the multiscale implementation of PD should also be able to detect regions in which the solution has a steep gradient. On a coarse grid, i.e., a grid size larger than the regions in which the gradient is steep, one cannot tell the difference between a jump discontinuity and a steep gradient. However, as the grid is refined, there is a fundamental difference in how the errors due to a DG discretization are distributed in the two cases. For jump discontinuities, the error is mostly confined to the elements which contain the discontinuity, whereas for steep gradients, once the grid becomes sufficiently small, the error will spread out over several elements.

8.1. Algorithmic modifications for solutions containing discontinuities and steep gradients. A multiscale implementation of PD should expect and detect discontinuities and steep gradients, to differentiate between the two cases, and to
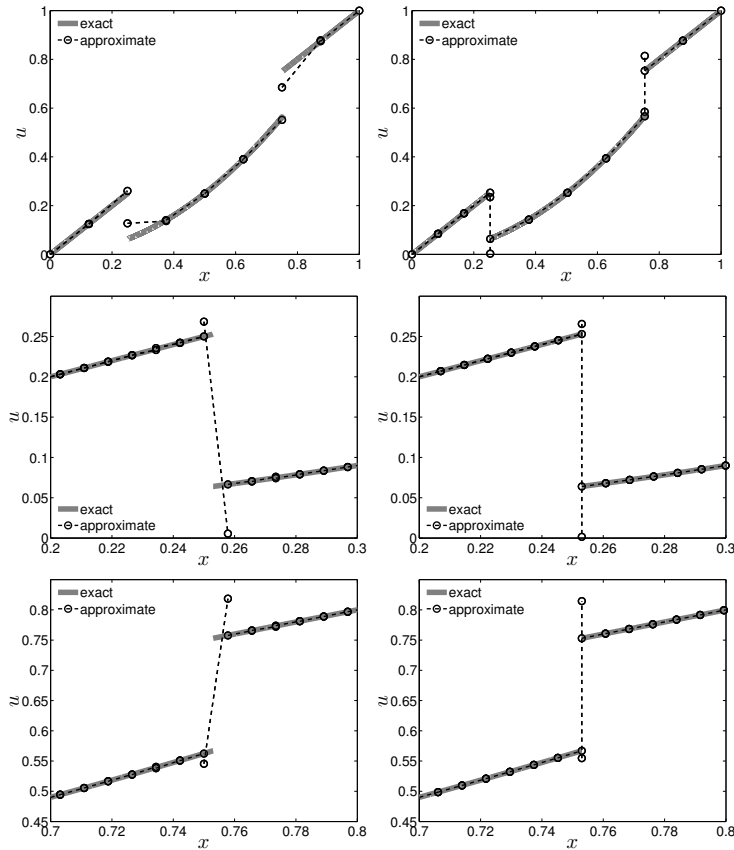
FIG. 6. *For the setting of Table* 12 *and the initial grid size* $h_0 = 1/8$ *(top) and* $h_0 = 1/128$ *(middle and bottom), the approximate solution at the initial (left) and final (right) step for the adaptive multiscale implementation of Algorithm* 2. *For the finer initial grid, the middle and bottom rows are zoom-ins near the left and right discontinuities, respectively.*

handle each case in an appropriate fashion. For discontinuities, during the refinement process, the weighted error estimator will have large values in isolated elements; e.g., the elements abutting an element containing a discontinuity will have relatively minute errors. On the other hand, once the region in which the solution has a steep gradient is covered by two or more elements, the errors in those elements will all be relatively the same. In this way, discontinuities and steep gradients can both be detected and differentiated from each other.

To put this idea in a concrete algorithm, we need to modify Algorithm 2 by setting a parameter $\beta \in (0, 1)$ as an indicator to distinguish the steep gradient from the discontinuity. Suppose we are currently in the kth iteration; then we look at the marked set $\mathcal{M}_{k-1}$ in the previous iteration. Each element $K \in \mathcal{M}_{k-1}$ has been split into two small elements K_1 and K_2, and the weighted relative errors $r = \eta(u_h^{k-1}, K)/\eta(u_h^{k-1}, \Omega)$ and $r_1 = \eta(u_h^k, K_1)/\eta(u_h^k, \Omega), r_2 = \eta(u_h^k, K_2)/\eta(u_h^k, \Omega)$ are calculated in iterations $k-1$ and k, respectively. If $r_1 \geq \beta r$ and $r_2 < \beta r$, then we claim that K_1 contains the discontinuity. If $\max\{r_1, r_2\} < \beta r$, which means the error of element K is reduced and spread into elements K_1, K_2, then we claim that there is a steep gradient within element K, and if K_1 and/or K_2 are not in $\mathcal{M}_k$, they should not be coarsened.
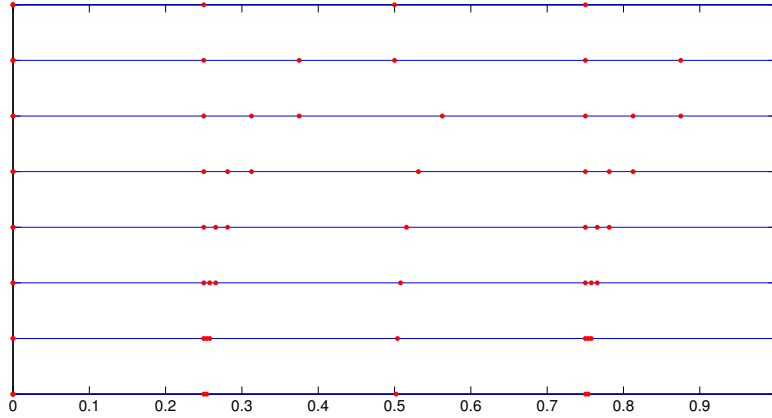
FIG. 7. *For the setting of Table* 12, *the grids at each refinement step (top to bottom) of Algorithm* 2. *The topmost grid is the initial uniform grid of size* $h_0 = 1/4$; *the bottommost grid is the final refined grid. All grids except the two at the top and the one at the bottom have nine grid points, with three points clustered near each of the two discontinuities in the solution. The bottommost one has two points clustered near each of the two discontinuities in the solution.*

Another modification is the criterion for selecting the marked set $\mathcal{M}_k$. Instead of using the weighted posterior error, we go back to the standard error estimator $\widetilde{\eta}$. Otherwise, once the weighted errors corresponding to the elements containing the true discontinuities account for a sufficiently large portion of the total error relative to the threshold specified by the marking parameter θ, the elements where steep gradients are located will not be further refined, even though the latter may contain a large error with respect to the unweighted measure. Moreover, as an indicator of the L^2 norm of the solution error, the standard error estimator is more suitable for choosing where refinement is to take place.

There is another difference between how steep gradients and discontinuities should be discretized. Across a steep gradient, the solution is still smooth, and in particular is smooth within the elements which cover the region of the steep gradient. Thus, one should be able to apply the PDE (or PD if in the vicinity of the DG nodes) and use CG discretizations at nodes of those elements.

There is also seemingly a big difference in how the iteration should be terminated in the two cases. When considering discontinuities, the iteration is stopped when the element size is of $O(h^4)$ in $\mathcal{M}_k$, which includes the elements containing the discontinuities; the $O(h^4)$ criterion arises because the L^2 norm error in an element containing the discontinuity is of $O(h_{disc}^{1/2})$, where h_{disc} denotes the mesh size of that element. On the other hand, with h_{steep} denoting the size of the elements that cover the step gradients, the error for those elements is $O(h_{steep}^2)$, so that it seems that one should be able to stop the refinement process much sooner in this case. Whether this is true or not depends on how large the gradient of the solution is, because the constant in the order relation depends on the size of gradient. For example, suppose that, in a region in which the gradient is not large, we have that the error is given by $C_{nonsteep}h^2$, whereas in a region in which the gradient is large, it is given by $C_{steep}h_{steep}^2$, where $C_{steep} \gg C_{nonsteep}$. Then, the iteration should stop when $h_{steep} \approx (\frac{C_{nonsteep}}{C_{steep}})^{1/2}h$.

8.2. Adaptive refinement for a solution having a discontinuity and a steep gradient. The final numerical example involves the modified adaptive algorithm described in section 8.1. We use the Case I kernel in Table 4 and, for simplicity and because our focus here is to show that steep gradients can be detected and differentiated from discontinuities, we use only the nonlocal equation everywhere. For the same reasons, we stop the refinement process for both steep gradients and discontinuities when the element size is of $O(h^4)$.

The manufactured solution is now chosen as

$$(13) \qquad u(x) = \begin{cases} \dfrac{1}{1 + e^{\alpha(x - \widetilde{x}_2)}} & \text{for } x < \widetilde{x}_1, \\ x^2 & \text{otherwise.} \end{cases}$$

The parameter values are set to $\alpha = 400$, $\widetilde{x}_1 = 0.753$, and $\widetilde{x}_2 = 0.3$. The solution (13) is a piecewise-smooth function with a jump discontinuity at $\widetilde{x}_1 = 0.753$ and a steep gradient in the vicinity of the point $\widetilde{x}_2 = 0.3$. For the algorithm parameters, we set $\theta = \beta = 0.8$.

Table 13 lists the various error measures associated with a sequence of decreasing grid sizes which are consistent with our expectation. However, the elementwise relative errors given in Table 13, as well as Table 14 and Figures 8, 9, and 10, point out the differences between adaptive refinement for resolving discontinuities and steep gradients in the solution. In Table 14, information is provided about the grid obtained by the algorithm described in section 8.1. One can observe the decreasing trend of the maximal relative error $\widetilde{E}^1$ in Table 13, whereas $\theta = 0.8$ is bigger than $\widetilde{E}^1$ except for the case $h_0 = 1/4$. This indicates that many of the elements around the steep gradient are also refined until $O(h_0^4)$ has been achieved (see Figure 8). That is why there is a large increase in N_f^n (see Table 14). However, it is seen in Table 14 that only two PD-DG nodes are involved because there is only one discontinuity, which also hints that we have succeeded in distinguishing the steep gradient from discontinuity; e.g., the steep gradient is handled using CG, whereas the discontinuity is handled using DG.

Figure 8 displays the exact solution and compares the initial and final approximate solutions for the initial grid sizes $h_0 = 1/8$ and $h_0 = 1/128$. The figure shows that the refinement algorithm was able to detect the problematic regions and adjust the grid to isolate them and control the error. For the discontinuity, only one element of small size contains the discontinuity, whereas the surrounding elements are still of large size. For the steep gradient, several elements are of small size.

Figure 9 provides a more detailed picture of how the mesh is gradually adjusted to bracket the point at which the solution discontinuity occurs and the thin region across which the solution gradient is large. In the vicinity of the point $\widetilde{x}_2 = 0.3$ across which the gradient of the solution is steep, several grid points are added, and the nearby elements are shrunk. The discontinuity occurring at $\widetilde{x}_1 = 0.753$ is handled more economically, in that only a single small element is used to isolate it and the neighboring elements are of the nominal mesh size $O(h_0)$.

Figure 10 displays the posterior errors. At the initial step of the algorithm, both the steep gradient and the discontinuity cause large relative errors in the elements that contain them, with the steep gradient error predominating because of the huge derivatives involved, i.e., because of the large constant in the error relation Ch^2; see section 8.1. However, the refinement algorithm is able to drive down the error associated with the steep gradient, and this remains true even if the weighted elementwise relative error measure is used. Meanwhile, the single element containing the discon-

TABLE 13

For the exact solution (13) with $\delta = 0.02$, $\widetilde{x}_1 = 0.753$, $\widetilde{x}_2 = 0.3$, and the nonsingular kernel of Case I in Table 4, errors and convergence rates (CR) resulting from the application of the modified adaptive algorithm described in section 8.1.

Errors and convergence rates including all elements						
N_0^e	$\widetilde{\eta}(\Omega)$	CR	$\widetilde{E}^1$	$\widetilde{E}^2$	E^1	E^2
4	1.31e+0	–	0.910	0.948	0.954	0.974
8	3.61e-1	1.861	0.712	0.871	0.983	0.990
16	8.38e-2	2.110	0.708	0.771	0.998	0.998
32	2.34e-2	1.841	0.678	0.726	1.000	1.000
64	5.79e-3	2.013	0.620	0.657	1.000	1.000
128	1.46e-3	1.989	0.589	0.614	1.000	1.000
256	3.52e-4	2.049	0.643	0.652	1.000	1.000
N_0^e	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$\|\|e^h\|\|$	CR
4	1.20e-2	–	2.90e-1	–	1.31e+0	–
8	3.25e-3	1.880	3.04e-1	-0.068	3.62e-1	1.857
16	7.94e-4	2.035	3.03e-1	0.005	8.45e-2	2.098
32	2.07e-4	1.942	3.06e-1	-0.016	2.33e-2	1.858
64	5.24e-5	1.979	3.13e-1	-0.030	5.80e-3	2.006
128	1.32e-5	1.985	3.05e-1	0.034	1.46e-3	1.992
256	3.24e-6	2.027	3.71e-1	-0.281	3.52e-4	2.048

| Errors excluding the element containing the discontinuities						
N_0^e	$\|e^h\|_{L^2}$	CR	$\|e^h\|_{L^\infty}$	CR	$|e^h|_{H^1}$	CR
4	6.27e-3	–	4.73e-2	–	2.62e+0	–
8	2.10e-3	1.576	3.45e-2	0.458	1.96e+0	0.417
16	5.47e-4	1.941	5.16e-3	2.739	8.93e-1	1.137
32	1.34e-4	2.027	1.76e-3	1.551	4.74e-1	0.915
64	3.69e-5	1.862	5.43e-4	1.698	2.43e-1	0.967
128	9.58e-6	1.946	1.82e-4	1.580	1.27e-1	0.933
256	2.28e-6	2.066	3.93e-5	2.203	6.00e-2	1.079

TABLE 14

For the exact solution (13) with $\delta = 0.02$, $\widetilde{x}_1 = 0.753$, $\widetilde{x}_2 = 0.3$, and the nonsingular kernel of Case I in Table 4 and for several initial grid sizes, the distribution of node types and the maximal and minimal grid sizes of the final grid resulting from the application of the modified adaptive algorithm described in section 8.1. Note that in this section, only the nonlocal model is applied.

N_0^n	N_f^n	N_{dg}^{pd}	N_{cg}^{pd}	h_0	$\overline{h}$	$\underline{h}$	$\overline{h}/h_0$	$\underline{h}/h_0^4$	$\underline{h}/\overline{h}^4$
5	11	2	9	2.50e-1	2.50e-1	3.91e-3	1.00e+0	1.00e+0	1.00e+0
9	17	2	15	1.25e-1	1.25e-1	2.44e-4	1.00e+0	1.00e+0	1.00e+0
17	30	2	28	6.25e-2	6.25e-2	1.53e-5	1.00e+0	1.00e+0	1.00e+0
33	57	2	55	3.13e-2	3.12e-2	9.54e-7	1.00e+0	1.00e+0	1.00e+0
65	108	2	106	1.56e-2	1.56e-2	5.96e-8	1.00e+0	1.00e+0	1.00e+0
129	210	2	208	7.81e-3	7.81e-3	3.73e-9	1.00e+0	1.00e+0	1.00e+0
257	423	2	421	3.91e-3	3.90e-3	2.33e-10	9.99e-1	1.00e+0	1.00e+0
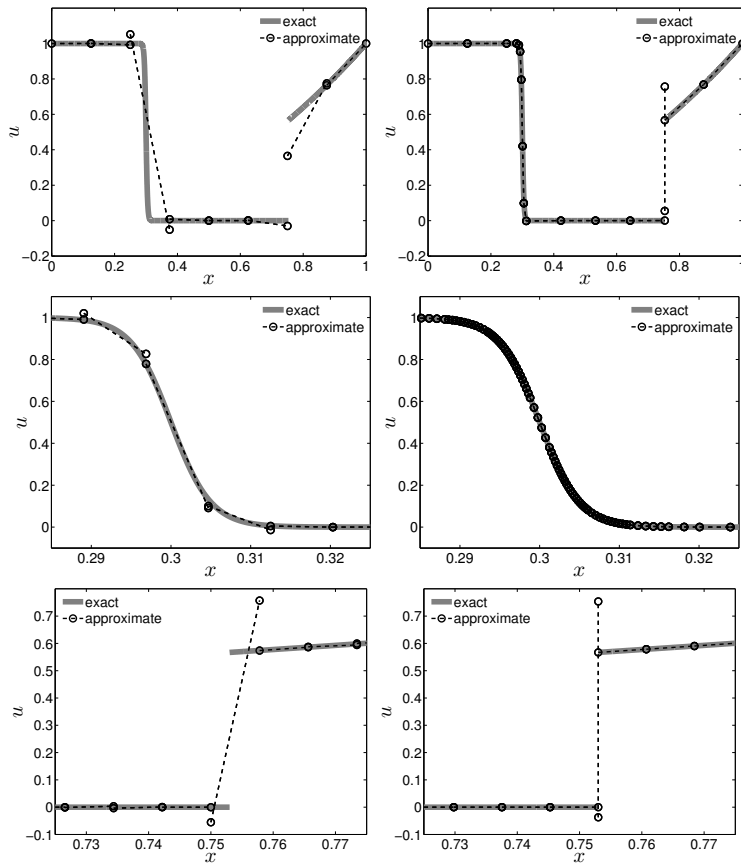
Fig. 8. *For the exact solution* (13) *with* $\delta = 0.02$, $\widetilde{x}_1 = 0.753$, $\widetilde{x}_2 = 0.3$, *and the nonsingular kernel of Case* I *in Table* 4 *and the initial grid sizes* $h_0 = 1/8$ *(top) and* $h_0 = 1/128$ *(middle and bottom), the approximate solution at the initial (left) and final (right) step for the adaptive multiscale implementation of the modified adaptive algorithm described in section* 8.1. *For the finer initial grid, the middle and bottom rows are zoom-ins near the steep gradient and discontinuity, respectively.*

tinuity shrinks in size and, as the adaptive iterations proceed, gathers almost all of the weighted relative error; for further evidence, see the value given for E^1 and E^2 in Table 13. Thus, the discontinuity and steep gradient can be effectively isolated, each in its own way, which is consistent with the discussions in sections 4 and 8.1.

9. Conclusion. In this work, we provide a multiscale finite element implementation of a linear one-dimensional peridynamics (PD) model for solid mechanics which also features coupling of the nonlocal PD model to local PDE models and the use of both discontinuous and continuous Galerkin FEMs. An adaptive algorithm is defined that controls not only the local mesh size, but also which of the two FEM discretizations is used at each point in the mesh. Adaptivity is guided by a posterior error estimator, which is developed from a standard residual error estimator that can detect discontinuities during adaptive refinement. Several numerical examples show how our adaptive algorithm can detect the location of a discontinuity in the solution and bracket that discontinuity within a single element. Moreover, by using DG methods in that element, CG methods elsewhere, and switching to a local PDE model away
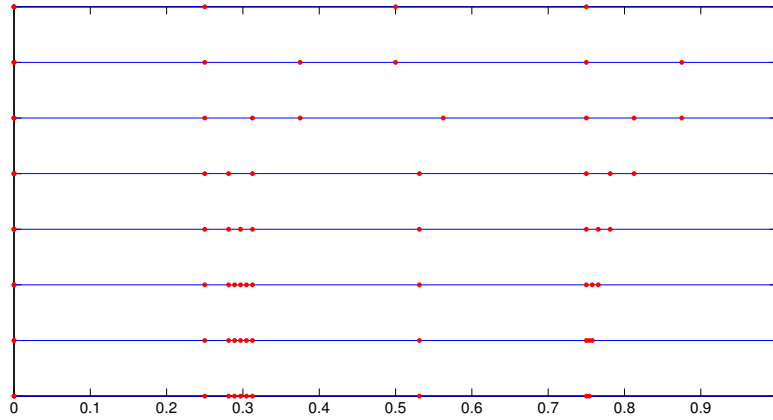
FIG. 9. *For the exact solution* (13) *with* $\delta = 0.02$, $\widetilde{x}_1 = 0.753$, $\widetilde{x}_2 = 0.3$, *and the nonsingular kernel of Case* I *in Table* 4, *the grids at each refinement step (top to bottom, i.e., starting with the initial grid at the top) of the modified adaptive algorithm described in section* 8.1. *The topmost grid is the initial uniform grid of size* $h_0 = 1/4$; *the bottommost grid is the final refined grid.*
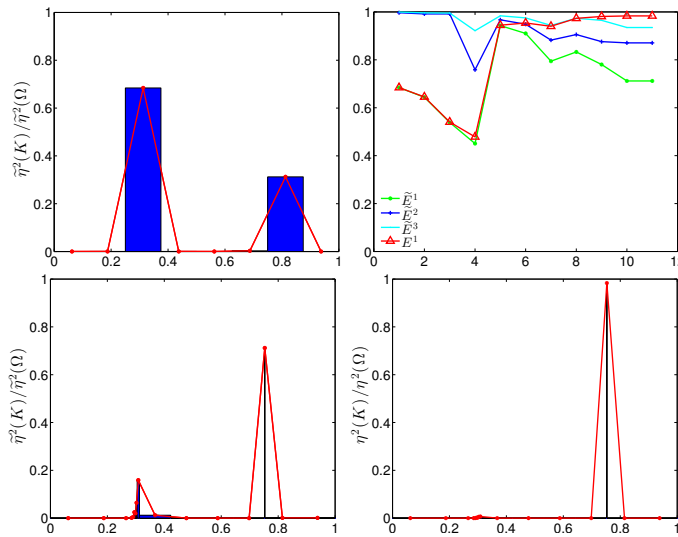

FIG. 10. *For the setting of Table* 13 *with an initial uniform grid of size* $h_0 = 1/8$, *top-left: the error fraction distribution* $\widetilde{\eta}^2(K)/\widetilde{\eta}^2(\Omega)$ *across the elements* K *of the initial uniform grid; top-right:* $\widetilde{E}^i$ *for* $i = 1, 2, 3$ *and* E^1 *for each step in the adaptive multiscale implementation of the modified adaptive algorithm described in section* 8.1; *and the elementwise relative errors* $\widetilde{\eta}^2(K)/\widetilde{\eta}^2(\Omega)$ *(bottom-left) and* $\eta^2(K)/\eta^2(\Omega)$ *(bottom-right) for the final grid obtained using the modified adaptive algorithm described in section* 8.1.

from the elements containing solution discontinuities, we can recover the quadratic convergence of the approximation without incurring a large increase in costs due to the use of DG elements and grid refinement. We also briefly discuss how modifications of our algorithm can be used to detect and differentiate between regions with steep

gradients and true discontinuities in the solution.

For the most part, the extension of our algorithm to two and three dimensions shares the same difficulties that one encounters with any finite element discretization of the nonlocal PD model, the most troublesome being defining good quadrature rules for regions defined as the intersections of balls of radius δ and triangular or quadrilateral elements used in two dimensions and tetrahedral or hexahedral elements used in three dimensions.

The one difficulty somewhat specific to our algorithm is the refinement to elements of size $O(h^4)$ across discontinuities, but of size h along discontinuities; such highly anisotropic refinement is needed in order to avoid incurring large additional costs. On the other hand, for discontinuities occurring along curves and curved surfaces, it is not possible to contain the discontinuity within elements of width $O(h^4)$ without the use of isoparametric elements [2, 3, 4]. Clearly, although there are clear paths to extending our methodology to higher dimensions, actually effecting such extensions is not a trivial matter.

REFERENCES

[1] Y. AZDOUD, F. HAN, AND G. LUBINEAU, *The morphing method as a flexible tool for adaptive local/non-local simulation of fracture*, Comput. Mech., 54 (2014), pp. 711–722.

[2] R. BARSOUM, *On the use of isoparametric finite elements in linear fracture mechanics*, Internat. J. Numer. Methods Engrg., 10 (1976), pp. 25–37.

[3] K.-J. BATHE AND E. WILSON, *Numerical Methods in Finite Element Analysis*, Prentice–Hall, Englewood Cliffs, NJ, 1976.

[4] S. BENZLEY, *Representation of singularities with isoparametric finite elements*, Internat. J. Numer. Methods Engrg., 8 (1974), pp. 537–545.

[5] F. BOBARU, J. FOSTER, P. GEUBELLE, AND S. SILLING, EDS., *Handbook of Peridynamic Modeling*, CRC Press, Boca Raton, FL, 2015.

[6] M. BREITENFELD, P. GEUBELLE, O. WECKNER, AND S. SILLING, *Non-ordinary state-based peridynamic analysis of stationary crack problems*, Comput. Methods Appl. Mech. Engrg., 272 (2014), pp. 233–250.

[7] X. CHEN AND M. GUNZBURGER, *Continuous and discontinuous finite element methods for a peridynamics model of mechanics*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 1237–1250.

[8] W. DÖRFLER, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124, doi:10.1137/0733054.

[9] Q. DU, M. GUNZBURGER, R. B. LEHOUCQ, AND K. ZHOU, *Analysis and approximation of nonlocal diffusion problems with volume constraints*, SIAM Rev., 54 (2012), pp. 667–696, doi:10.1137/110833294.

[10] Q. DU, M. GUNZBURGER, R. LEHOUCQ, AND K. ZHOU, *A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws*, Math. Models. Methods Appl. Sci., 23 (2013), pp. 493–540.

[11] Q. DU, L. JU, L. TIAN, AND K. ZHOU, *A posteriori error analysis of finite element method for linear nonlocal diffusion and peridynamic models*, Math. Comp., 82 (2013), pp. 1889–1922.

[12] Q. DU, L. TIAN, AND X. ZHAO, *A convergent adaptive finite element algorithm for nonlocal diffusion and peridynamic models*, SIAM J. Numer. Anal., 51 (2013), pp. 1211–1234, doi:10.1137/120871638.

[13] E. EMMRICH AND O. WECKNER, *On the well-posedness of the linear peridynamic model and its convergence towards the Navier equation of linear elasticity*, Commun. Math. Sci., 5 (2007), pp. 851–864.

[14] M. GUNZBURGER AND R. B. LEHOUCQ, *A nonlocal vector calculus with application to nonlocal boundary value problems*, Multiscale Model. Simul., 8 (2010), pp. 1581–1598, doi:10.1137/090766607.

[15] Y. HA AND F. BOBARU, *Characteristics of dynamic brittle fracture captured with peridynamics*, Engrg. Fracture Mech., 78 (2011), pp. 1156–1168.

[16] F. HAN AND G. LUBINEAU, *Coupling of nonlocal and local continuum models by the Arlequin approach*, Internat. J. Numer. Methods Engrg., 89 (2012), pp. 671–685.

[17] G. LUBINEAU, Y. AZDOUD, F. HAN, C. REY, AND A. ASKARI, *A morphing strategy to couple*

nonlocal to local continuum mechanics, J. Mech. Phys. Solids, 60 (2012), pp. 1088–1102.

[18] R. Macek and S. Silling, *Peridynamics via finite element analysis*, Finite Elem. Anal. Des., 43 (2007), pp. 1169–1178.

[19] T. Mengesha and Q. Du, *The bond-based peridynamic system with Dirichlet-type volume constraint*, Proc. Roy. Soc. Edinburgh Sect. A, 144 (2014), pp. 161–186.

[20] T. Mengesha and Q. Du, *Nonlocal constrained value problems for a linear peridynamic Navier equation*, J. Elasticity, 116 (2014), pp. 27–51.

[21] R. Rahman, J. Foster, and A. Haque, *A multiscale modeling scheme based on peridynamic theory*, Internat. J. Multiscale Comput. Engrg., 12 (2014), pp. 223–248.

[22] P. Seleson, *Peridynamic Multiscale Models for the Mechanics of Materials: Constitutive Relations, Upscaling from Atomistic Systems, and Interface Problems*, Ph.D. thesis, Scientific Computing, Florida State University, 2010 (Electronic Theses, Treatises and Dissertations, Paper 273).

[23] P. Seleson, S. Beneddine, and S. Prudhomme, *A force-based coupling scheme for peridynamics and classical elasticity*, Comput. Mater. Sci., 66 (2013), pp. 34–49.

[24] P. Seleson, Y. Ha, and S. Beneddine, *Concurrent coupling of bond-based peridynamics and the Navier equation of classical elasticity by blending*, Internat. J. Multiscale Comput. Engrg., 13 (2015), pp. 91–113.

[25] P. Seleson, M. Parks, and M. Gunzburger, *Peridynamic state-based models and the embedded-atom model*, Commun. Comput. Phys., 15 (2014), pp. 179–205.

[26] P. Seleson, M. L. Parks, M. Gunzburger, and R. B. Lehoucq, *Peridynamics as an upscaling of molecular dynamics*, Multiscale Model. Simul., 8 (2009), pp. 204–227, doi:10.1137/09074807X.

[27] S. Silling, *Reformulation of elasticity theory for discontinuities and long-range forces*, J. Mech. Phys. Solids, 48 (2000), pp. 175–209.

[28] S. Silling, *A coarsening method for linear peridynamics*, Internat. J. Multiscale Comput. Engrg., 9 (2011), pp. 609–622.

[29] S. Silling and F. Bobaru, *Peridynamic modeling of membranes and fibers*, Internat. J. Non-Linear Mech., 40 (2005), pp. 395–409.

[30] S. Silling, M. Epton, O. Weckner, J. Xu, and E. Askari, *Peridynamic states and constitutive modeling*, J. Elasticity, 88 (2007), pp. 151–184.

[31] S. Silling and R. Lehoucq, *Convergence of peridynamics to classical elasticity theory*, J. Elasticity, 93 (2008), pp. 13–37.

[32] X. Tian and Q. Du, *Analysis and comparison of different approximations to nonlocal diffusion and linear peridynamic equations*, SIAM J. Numer. Anal., 51 (2013), pp. 3458–3482, doi:10.1137/13091631X.

[33] X. Tian and Q. Du, *Asymptotically compatible schemes and applications to robust discretization of nonlocal models*, SIAM J. Numer. Anal., 52 (2014), pp. 1641–1665, doi:10.1137/130942644.

[34] X. Tian and Q. Du, *Nonconforming discontinuous Galerkin methods for nonlocal variational problems*, SIAM J. Numer. Anal., 53 (2015), pp. 762–781, doi:10.1137/140978831.

[35] K. Zhou and Q. Du, *Mathematical and numerical analysis of linear peridynamic models with nonlocal boundary conditions*, SIAM J. Numer. Anal., 48 (2010), pp. 1759–1780, doi:10.1137/090781267.