INTRODUCTION TO ADAPTIVE FINITE ELEMENT METHODS

LONG CHEN

Adaptive methods are now widely used in the scientific computation to achieve better accuracy with minimum degree of freedom. In this notes, we give a briefly introduction to adaptive finite element methods via adaptive mesh refinements (AMR).

1. Introduction to mesh adaptation

We start with a simple motivation taken from [6] in one dimension for the use of adaptive mesh refinement. Given $\Omega = (0,1)$, a grid $\mathcal{T}_N = \{x_i\}_{i=0}^N$ of Ω

$$0 = x_0 < x_1 < \cdots x_i < \cdots < x_N = 1,$$

and a continuous function $u : \Omega \to \mathbb{R}$, we consider the problem of approximating u by a piecewise constant function u_N over $\mathcal{T}_N$. The approximation error is measured in the maximum norm.

Suppose that u is Lipschitz in $[0,1]$. Consider the approximation

$$u_N(x) := u(x_{i-1}), \quad \text{for } x_{i-1} \le x < x_i, i = 1, \cdots N.$$

If the grid is quasi-uniform in the sense that $h_i = x_i - x_{i-1} \le C/N$ for $i = 1, \cdots N$, then it is easy to show that

$$\|u - u_N\|_\infty \le CN^{-1}\|u'\|_\infty \tag{1}$$

We can achieve the same convergent rate N^{-1} with less smoothness of the function. Suppose $\|u'\|_{L^1} \neq 0$. Let us define a grid distribution function

$$F(x) := \frac{1}{\|u'\|_{L^1}} \int_0^x |u'(t)| \,\mathrm{d}t.$$

Then $F : [0,1] \to [0,1]$ is a non-decreasing function which resembles the cumulative distribution function in probability theory. Let $y_i = i/N, i = 0, \cdots, N$ be a uniform grid of $[0,1]$ along the y-axis. We choose x_i such that $F(x_i) = y_i$, see Fig. 1 for an illustration.

Then

$$\int_{x_{i-1}}^{x_i} |u'(t)| \,\mathrm{d}t = F(y_i) - F(y_{i-1}) = N^{-1}, \tag{2}$$

and

$$|u(x) - u(x_{i-1})| \le \int_{x_{i-1}}^{x_i} |u'(t)| \,\mathrm{d}t \le N^{-1}\|u'\|_{L^1},$$

which leads to the estimate

$$\|u - u_N\|_\infty \le CN^{-1}\|u'\|_{L^1}. \tag{3}$$

To achieve the first order convergence, we still require the function is differentiable but u' is now measured in a much weaker norm.

Date: Latest update: July 6, 2020.

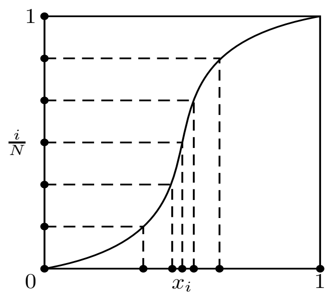


FIGURE 1. A grid distribution function.

We use the following example to illustrate the advantage of (3) over (1). Let us consider the function $u(x) = x^r$ with fixed $r \in (0,1)$. Then $u' \notin L^\infty(\Omega)$ but $u' \in L^1(\Omega)$. Therefore we cannot obtain optimal convergent rate on quasi-uniform grids while we could on a correctly adapted grid. For this simple example, one can easily compute when

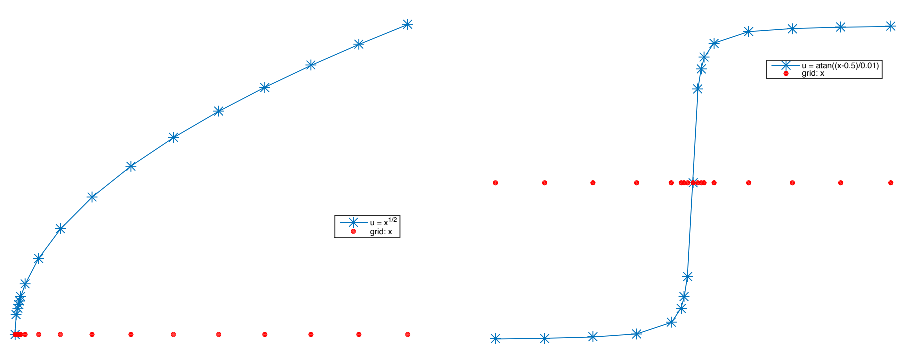
$$x_i = \left(\frac{i}{N}\right)^{1/r}, \quad \text{for all } 0 \le i \le N,$$

estimate (3) will hold on the grid $\mathcal{T}_N = \{x_i\}_{i=0}^N$ which has higher density of grid points near the singularity of u.

A possible MATLAB code is given below, where `M` is a nonnegative function defined on the input grid `x`. The output is a new grid which equidistributes the distribution function $F \propto \int M$.

```
function x = equidistribution(M,x)
h = diff(x);
F = [0; cumsum(h.*M)];
F = F/F(end);
y = (0:1/(length(x)-1):1)';
x = interp1(F,x,y);
```

Examples of adaptive grids for two functions with singularity are plotted below.



(a) $u(x) = x^{1/2}$ (b) $u(x) = \arctan((x - 0.5)/0.01)$

FIGURE 2. Adaptive grids for two functions with singularity.

In (2), we choose a grid such that a upper bound of the error is equidistributed. This is instrumental for adaptive finite element methods on solving PDEs. When applied to numerical solution of PDEs, the function u and its derivatives are unknown. Only an approximated solution to the function u at grid points is available and a good approximation of derivatives or an upper bound of the error should be computed by a post-processing procedure. When generalizing the adaptive procedure to two and higher dimensions, another difficulty is the mesh requirement, coarsening, or movement which is much more complicated in higher dimensions.

Exercise 1.1. We consider the piecewise linear approximation of a second differentiable function in this exercise.

(1) When $|f|$ is monotone decreasing, for a positive integer k, prove

$$\frac{1}{(k-1)!}\int_{x_i}^{x_{i+1}} |f(x)|(x-x_i)^{k-1}\,\mathrm{d}x \le \frac{1}{k!}\left(\int_{x_i}^{x_{i+1}} |f(x)|^{1/k}\,\mathrm{d}x\right)^k .$$

(2) Let u_I be the nodal interpolation of u on a grid, i.e., u_I is piecewise linear and $u_I(x_i) = u(x_i)$ for all i. Prove that if $|u''(x)|$ is monotone decreasing in (x_{i-1}, x_i), then for $x \in (x_{i-1}, x_i)$

$$|(u-u_I)(x)| \le \left(\int_{x_{i-1}}^{x_i} |u''(s)|^{1/2}\,\mathrm{d}s\right)^2 .$$

(3) Give the condition on the grid distribution and the function such that the following second order estimate holds and prove your result.

$$\|u-u_I\|_\infty \le C\|u''\|_{1/2}N^{-2}.$$

2. SINGULARITY

In this section we present several examples to show that the solution of elliptic equation could have singularity when the domain is concave or the coefficient is discontinuous.

In Introduction to Finite Element Methods, we have obtained a first order convergence of the linear finite element approximation to the Poisson equation

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega, \tag{4}$$

provided the solution $u \in H^2(\Omega)$. When the boundary of the domain is smooth or convex and Lipschitz continuous, then $\Delta^{-1}(L^2(\Omega)) \subset H^2(\Omega)$. The requirements of Ω is necessary as shown by the following example.

Consider the domain $\Omega \subset \mathbb{R}^2$ which is defined in a polar coordinate as $\Omega = \{(r,\theta) : 0 < r < 1 \text{ and } 0 < \theta < \pi/\beta\}$ for $1/2 < \beta < \infty$. Obviously if $\beta \ge 1$ then Ω is convex, while if $1/2 < \beta < 1$ then Ω violates the condition of the regularity theory. Set $v = r^\beta \sin(\beta\theta)$ as the imaginary part of the analytic function z^β, i.e., $v = \mathrm{Img}(z^\beta)$. According to the properties of analytic functions, we know $\Delta v = 0$. With this fact, it is easy to verify that $u = (1-r^2)v$ is the solution of the equation (4) with $f = 4(1+\beta)v \in L^2(\Omega)$.

Now we check the regularity of u. The only possible singularity is at the origin. When r is near 0, the second derivative $D^\alpha u \sim r^{\beta-2}$ for any $|\alpha| = 2$. Considering the integral

$$\int_\Omega |D^\alpha u|^2\,\mathrm{d}x\,\mathrm{d}y \lesssim \int_0^1 |D^\alpha u|^2 r\,\mathrm{d}r = \int_0^1 r^{2(\beta-2)+1}\,\mathrm{d}r.$$

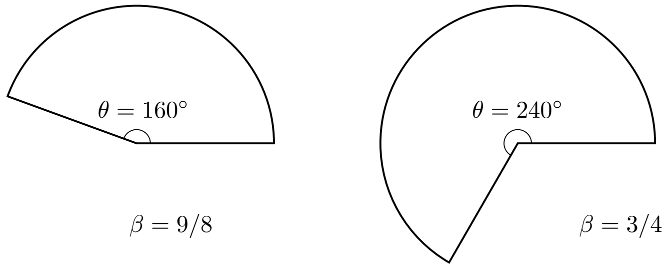


FIGURE 3. Circular domains

Therefore $u \in H^2(\Omega)$ if and only if $2(\beta - 2) + 1 > -1$, i.e., $\beta > 1$. Namely the domain Ω is convex. When β is fixed, by the same calculation, we conclude $u \in H^s(\Omega)$ for any $s < 1 + \beta$. If we look for the smoothness in $W^{2,p}(\Omega)$ instead of $H^2(\Omega)$, similar calculation reveals that $\beta > 2(1 - \frac{1}{p})$. For this example, we conclude u belongs to $H^s(\Omega)$ for $1 \leq s < 3/2$ and $W^{2,p}(\Omega)$ for $1 \leq p < 4/3$.

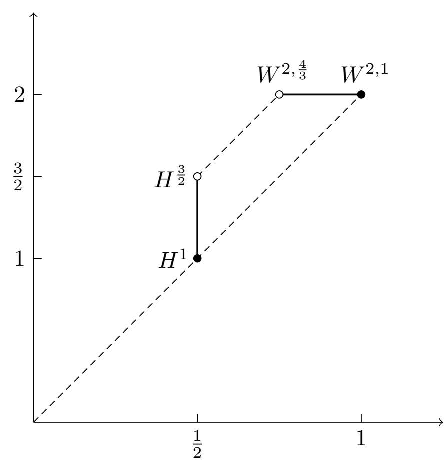


FIGURE 4. Regularity index for Poisson equations on polygons. Spaces on the (embedding) line with slope $d = 2$ are scaling invariant.

In general, following [8], for a polygonal domain Ω with boundary $\partial\Omega$ consisting of a finite number of straight line segments meeting at vertices v_j of interior angles $\alpha_j, j = 1, \cdots, M$, let us introduce the polar coordinates (r, θ) at the vertex v_j so that the interior of the wedge is given by $0 < \theta < \alpha_j$ and set $\beta_j = \pi/\alpha_j$, then near v_j the solution u behaves like

$$u(r, \theta) = k_j r^{\beta_j} |\ln r|^{m_j} \sin(\beta_j \theta) + w,$$

where k_j is a constant called the stress intensity factors, $m_j = 0$ unless $\beta_j = 2, 3, \cdots$, and $w \in H^2(\Omega)$ is a smooth function. Globally it is easy to see that for any $\epsilon > 0$, $u \in H^{1+\min_j \beta_j - \epsilon}(\Omega)$. In particular, $u \in H^{3/2-\epsilon}(\Omega)$ but $u \notin H^2(\Omega)$ for concave polygonal domains.

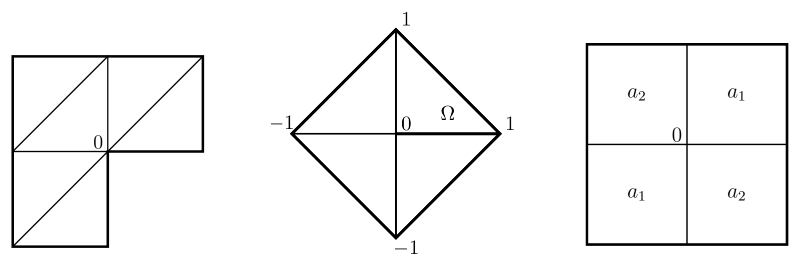
For a general elliptic equation

$$-\operatorname{div}(A\nabla u) = f \text{ in } \Omega, \tag{5}$$

the lack of regularity could also come from the discontinunity of the coefficients of A. See the example designed by Kellogg [9] with discontinuous diffusion coefficients listed in the end of this subsection.

When $u \in H^{1+\epsilon}(\Omega)$ with $\epsilon \in [0,1]$, in view of the approximation theory, we cannot expect the finite element approximation rate $\|u - u_{\mathcal{T}}\|_{1,\Omega}$ better than h^{ϵ} if we insist on quasi-uniform grids. To improve the convergence rate for small ϵ, the element size should be adapted to the behavior of the solution. The element size in areas of the domain where the solution is smooth can stay bounded well away from zero, and thus the maximal element size h of a triangulation $\mathcal{T}$ is not a good measure of the approximation rate. For this reason, $N = \#\mathcal{T}$ the number of elements is used, which is also proportional to the number of degree of freedom. Note that $N = \mathcal{O}(h^{-d})$ for quasi-uniform grids.

We include some typical examples below and refer to iFEM [5] for numerical evidence that finite element methods based on uniform refined grids will not give optimal order of convergence. A correctly adapted grid will recovery the optimal convergent order.



(a) L-shape problem (b) Crack problem (c) Jump coefficients

FIGURE 5. Lack of regularity of elliptic equations

L-shape problem. Let $\Omega := (-1,1)^2 \backslash \{[0,1) \times (-1,0]\}$ be a L-shaped domain with a reentrant corner.

$$-\Delta u = 0, \text{ in } \Omega \quad \text{ and } \; u = g \text{ on } \partial\Omega,$$

We choose the Dirichlet boundary condition g such that the exact solution reads

$$u(r,\theta) = r^{\frac{2}{3}} \sin(\frac{2}{3}\theta),$$

in the polar coordinates.

Crack problem. Let $\Omega = \{|x| + |y| < 1\} \backslash \{0 \le x \le 1, y = 0\}$ with a crack and the solution u satisfies the Poisson equation

$$-\Delta u = f, \text{ in } \Omega \quad \text{ and } \; u = u_D \text{ on } \Gamma_D,$$

where $f = 1$, $\Gamma_D = \partial\Omega$. We choose u_D such that the exact solution u in the polar coordinates is

$$u(r,\theta) = r^{\frac{1}{2}} \sin\frac{\theta}{2} - \frac{1}{4} r^2.$$

Jump coefficients problem. Consider the partial differential equation (5) with $\Omega = (-1,1)^2$ and the coefficient matrix A is piecewise constant: in the first and third quadrants, $A = a_1 I$; in the second and fourth quadrants, $A = a_2 I$. For $f = 0$, the exact solution in the polar coordinates has been chosen to be $u(r,\theta) = r^{\gamma}\mu(\theta)$, where

$$\mu(\theta) = \begin{cases} \cos\left((\frac{\pi}{2}-\sigma)\gamma\right)\cos\left((\theta-\frac{\pi}{2}+\rho)\gamma\right) & \text{if } 0 \le \theta \le \frac{\pi}{2}, \\ \cos(\rho\gamma)\cos\left((\theta-\pi+\sigma)\gamma\right) & \text{if } \frac{\pi}{2} \le \theta \le \pi, \\ \cos(\sigma\gamma)\cos\left((\theta-\pi-\rho)\gamma\right) & \text{if } \pi \le \theta \le \frac{3\pi}{2}, \\ \cos\left((\frac{\pi}{2}-\rho)\gamma\right)\cos\left((\theta-\frac{3\pi}{2}-\sigma)\gamma\right) & \text{if } \frac{3\pi}{2} \le \theta \le 2\pi, \end{cases}$$

and the constants

$$\gamma = 0.1, \quad \rho = \pi/4, \quad \sigma = -14.9225565104455152,$$

and

$$a_1 = 161.4476387975881, \quad a_2 = 1.$$

For this example, we see

$$u \in H^{1+\gamma}(\Omega).$$

One can construct more singular function by choosing arbitrary small γ; see Kellogg [9].

3. EQUIDISTRIBUTION

The equidistribution principle has been widely used in adaptive mesh refinements. But a theoretical justification of this principle is very difficult to be made precise. One early justification of this approach is due to Babuška and Rheinboldt [1] and they provide a heuristic asymptotic analysis.

In this section, we will illustrate the equidistribution principle in a more elementary fashion. Through our simple theoretical analysis, we will see the equidistribution is indeed needed for optimal error control, but on the other hand, we will show optimal convergent rate can still be maintained when equidistribution are much relaxed.

We shall consider a simple elliptic boundary value problem

$$-\Delta u = f \text{ in } \Omega, \qquad u = 0 \text{ on } \partial\Omega, \tag{6}$$

where, for simplicity, we assume Ω is a polygon and is partitioned by a shape regular conforming triangulation $\mathcal{T}_N$ with N number of triangles. Let $\mathcal{V}_N \subset H_0^1(\Omega)$ be the corresponding continuous piecewise linear finite element space associated with this triangulation $\mathcal{T}_N$.

A finite element approximation of the above problem is to find $u_N \in \mathcal{V}_N$ such that

$$a(u_N, v_N) = (f, v_N) \quad \forall v_N \in \mathcal{V}_N, \tag{7}$$

where

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x, \text{ and } (f,v) = \int_\Omega f v \, \mathrm{d}x.$$

For this problem, it is well known that for a fixed finite element space $\mathcal{V}_N$

$$|u - u_N|_{1,\Omega} = \inf_{v_N \in \mathcal{V}_N} |u - v_N|_{1,\Omega}. \tag{8}$$

We then present an H^1 error estimate for linear triangular element interpolation in two dimensions. We note that in two dimensions, the following two embeddings are both valid:

$$W^{2,1}(\Omega) \subset W^{1,2}(\Omega) \equiv H^1(\Omega) \text{ and } W^{2,1}(\Omega) \subset C(\bar{\Omega}). \tag{9}$$

Given $u \in W^{2,1}(\Omega)$, let u_I be the linear nodal value interpolant of u on $\mathcal{T}_N$. For any triangle $\tau \in \mathcal{T}_N$, thanks to (9) and the assumption that τ is shape-regular, we have

$$|u - u_I|_{1,\tau} \lesssim |u|_{2,1,\tau}.$$

As a result,

$$|u - u_I|_{1,\Omega}^2 \lesssim \sum_{\tau \in \mathcal{T}_N} |u|_{2,1,\tau}^2.$$

To minimize the error, we can try to minimize the right hand side. By Cauchy-Schwarz inequality,

$$|u|_{2,1,\Omega} = \sum_{\tau \in \mathcal{T}_N} |u|_{2,1,\tau} \leq (\sum_{\tau \in \mathcal{T}_N} 1)^{1/2} (\sum_{\tau \in \mathcal{T}_N} |u|_{2,1,\tau}^2)^{1/2} = N^{1/2} (\sum_{\tau \in \mathcal{T}_N} |u|_{2,1,\tau}^2)^{1/2}.$$

Thus, we have the following lower bound:

$$(\sum_{\tau \in \mathcal{T}_N} |u|_{2,1,\tau}^2)^{1/2} \geq N^{-1/2} |u|_{2,1,\Omega}. \tag{10}$$

The equality holds if and only if

$$|u|_{2,1,\tau} = \frac{1}{N} |u|_{2,1,\Omega}. \tag{11}$$

The condition (11) is hard to be satisfied in general. But we can considerably relax this condition to ensure the lower bound estimate (10) is still achieved asymptotically. The relaxed condition is as follows:

$$|u|_{2,1,\tau} \leq \kappa_{\tau,N} |u|_{2,1,\Omega} \tag{12}$$

and

$$\sum_{\tau \in \mathcal{T}_N} \kappa_{\tau,N}^2 \leq c_1 N^{-1}. \tag{13}$$

When the above two inequalities hold, we have

$$|u - u_I|_{1,\Omega} \lesssim N^{-1/2} |u|_{2,1,\Omega}.$$

In summary, we have the following theorem.

Theorem 3.1. *If $\mathcal{T}_N$ is a triangulation with at most N triangles and satisfying (12) and (13), then*

$$|u - u_N|_1 \leq |u - u_I|_{1,\Omega} \lesssim N^{-1/2} |u|_{2,1,\Omega}. \tag{14}$$

In the above analysis, we see how equidistribution principle plays an important role in achieving asymptotically optimal accuracy for adaptive grids. We would like to further elaborate that, in the current setting, equidistribution is indeed a sufficient condition for optimal error, but by no means this has to be a necessary condition. Namely the equidistribution principle can be severely violated but asymptoticly optimal error estimates can still be maintained. For example, the following mild violation of this principle is certainly acceptable:

$$|u|_{2,1,\tau} \leq \frac{c}{N} |u|_{2,1,\Omega}. \tag{15}$$

In fact, this condition can be more significantly violated on a finitely many elements $\{\tau\}$

$$|u|_{2,1,\tau} \leq \frac{c}{\sqrt{N}} |u|_{2,1,\Omega}. \tag{16}$$

It is easy to see if a bounded number of elements satisfy (16) and the rest satisfy (15), the estimate (13) is satisfied and hence the optimal error estimate (14) is still valid.

As we can see that the condition (16) is a very serious violation of equidistribution principle, nevertheless, as long as such violations do not occur on too many elements, asymptotically optimal error estimates are still valid. This simple observation is important from both theoretical and practical points of view. The marking strategy proposed by Dörfler [7] may also be interpreted in this way in its relationship with equidistribution principle.

As it turns out, rigorously speaking, we need a slightly stronger assumption on u (namely smoother than $W^{2,1}(\Omega)$), for example, $u \in W^{2,p}(\Omega)$ [1] for some $p > 1$. This assumption is true for most practical domains; see the discussion in the previous section. More precisely, for any $p > 1$, any N, we have a constructive algorithm [2] to find a shape-regular triangulation $\mathcal{T}_N$ with $\mathcal{O}(N)$ elements such that

$$|u|_{2,1,\tau} \leq c_0 N^{-1} |u|_{2,p,\Omega}.$$

As a result, since $|u - u_{\mathcal{T}}|_{1,\Omega} \leq |u - u_I|_{1,\Omega}$, we have the following error estimate

$$|u - u_{\mathcal{T}}|_{1,\Omega} \lesssim N^{-1/2} |u|_{2,p,\Omega}, \tag{17}$$

which is asymptotically best possible for an isotropic triangulation with $\mathcal{O}(N)$ elements. Recent works have shown that the estimate (17) can be practically realized [3, 4, 10, 11] by using *a posteriori* error estimates and will be discussed next in Introduction to Convergence Analysis of Adaptive Finite Element Methods.

References

[1] I. Babuška and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15:736–754, 1978. 6

[2] C. Bacuta, L. Chen, and J. Xu. Equidistribution and optimal approximation class. In M. W. O. X. J. Bank, Randolph Holst, editor, *Domain Decomposition Methods in Science and Engineering XX*, volume 91, pages 3–14, 2013. 8

[3] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004. 8

[4] J. M. Cascón, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008. 8

[5] L. Chen. iFEM: An Integrated Finite Element Methods Package in MATLAB. *Technical Report, University of California at Irvine*, 2009. 5

[6] R. A. DeVore. Nonlinear approximation. *Acta Numer.*, pages 51–150, 1998. 1

[7] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996. 8

[8] R. B. Kellogg. *Higher order singularities for interface problems*, in: The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, Elsevier, 589–602, 1972. 4

[9] R. B. Kellogg. On the Poisson equation with intersecting interface. *Appl. Aanal.*, 4:101–129, 1975. 5, 6

[10] P. Morin, R. H. Nochetto, and K. G. Siebert. Convergence of adaptive finite element methods. *SIAM Rev.*, 44(4):631–658, 2002. 8

[11] R. Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007. 8

[1]it actually suffices if $M(\nabla^2 u) \in L^1(\Omega)$, where $M(f)$ is the Hardy-Littlewood maximal function of f