HOW A NONCONVERGENT RECOVERED HESSIAN WORKS IN MESH ADAPTATION[*]

LENNARD KAMENSKI[†] AND WEIZHANG HUANG[‡]

Abstract. Hessian recovery has been commonly used in mesh adaptation for obtaining the required magnitude and direction information of the solution error. Unfortunately, a recovered Hessian from a linear finite element approximation is nonconvergent in general as the mesh is refined. It has been observed numerically that adaptive meshes based on such a nonconvergent recovered Hessian can nevertheless lead to an optimal error in the finite element approximation. This also explains why Hessian recovery is still widely used despite its nonconvergence. In this paper we develop an error bound for the linear finite element solution of a general boundary value problem under a mild assumption on the closeness of the recovered Hessian to the exact one. Numerical results show that this closeness assumption is satisfied by the recovered Hessian obtained with commonly used Hessian recovery methods. Moreover, it is shown that the finite element error changes gradually with the closeness of the recovered Hessian. This provides an explanation of how a nonconvergent recovered Hessian works in mesh adaptation.

1. Introduction. Gradient and Hessian recovery has been commonly used in mesh adaptation for the numerical solution of partial differential equations (PDEs); e.g., see [3, 4, 13, 19, 24, 26, 27]. The use typically involves the approximation of solution derivatives based on a computed solution defined on the current mesh (recovery), the generation of a new mesh using the recovered derivatives, and the solution of the physical PDE on the new mesh. These steps are often repeated several times until a suitable mesh and a numerical solution defined thereon are obtained. As the mesh is refined, a sequence of adaptive meshes, derivative approximations, and numerical solutions results. A theoretical and also practical question is whether this sequence of numerical solutions converges to the exact solution. Naturally, this question is linked to the convergence of the recovered derivatives used to generate the meshes. It is known that recovered gradient through the least squares fitting [26, 27] or the polynomial preserving technique [24] is convergent for uniform or quasi-uniform meshes [24, 25] and superconvergent for mildly structured meshes [23] as well as for a type of adaptive mesh [22].

For the Hessian, it has been observed that, unfortunately, a convergent recovery cannot be obtained from linear finite element approximations for general nonuniform meshes [2, 15, 18], although Hessian recovery is known to converge when the numerical solution exhibits superconvergence or supercloseness for some special meshes [5, 6, 17].

On the other hand, numerical experiments also show that the numerical solution obtained with an adaptive mesh generated using a nonconvergent recovered Hessian not only is often convergent but also has an error comparable to that obtained with

---

[†]Weierstrass Institute, Berlin 10117, Germany (kamenski@wias-berlin.de).

[‡]Department of Mathematics, University of Kansas, Lawrence, KS 66045-2142 (whuang@ku.edu).

(a) finite element error $|u - u_h|_{H^1(\Omega)}$ (b) recovery error $\max_K \max_{\boldsymbol{x}} \|R_K - H(\boldsymbol{x})\|_\infty$
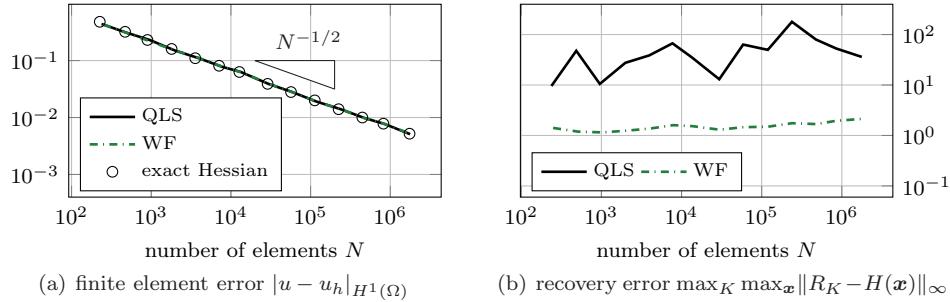
FIG. 1. *Finite element and Hessian recovery errors as a function of N.*

the exact analytical Hessian. To demonstrate this, we consider a Dirichlet boundary value problem (BVP) for the Poisson equation

$$(1.1) \qquad \begin{cases} -\Delta u = f & \text{in } \Omega = (0,1) \times (0,1), \\ u = g & \text{on } \partial\Omega, \end{cases}$$

where f and g are chosen such that the exact solution of the BVP is given by

$$(1.2) \qquad u(x,y) = x^2 + 25y^2.$$

Two Hessian recovery methods, quadratic least squares fitting (QLS) and weak formulation (WF) are used. (See section 4 for the description of these and other Hessian recovery techniques.) Figure 1 shows the error in recovered Hessian and the linear finite element solution with exact and recovered Hessian. One can see that the finite element error is convergent and almost undistinguishable for the exact and approximate Hessian (Figure 1(a)), whereas the error of the Hessian recovery remains $\mathcal{O}(1)$ (Figure 1(b)). Obviously, this indicates that a convergent recovered Hessian is not necessary for the purpose of mesh adaptation. Of course, a badly recovered Hessian does not serve the purpose either.

How accurate should a recovered Hessian be for the purpose of mesh adaptation? This issue has been studied by Agouzal, Lipnikov, and Vassilevski [1] and Vassilevski and Lipnikov [21]. In particular [1, Theorem 3.2], they show that a mesh based on an approximation R of the Hessian H is quasi-optimal if there exist small (with respect to one) positive numbers ε and δ such that

$$(1.3) \qquad \max_{\boldsymbol{x}\in\omega_i} \|H(\boldsymbol{x}) - H_{\omega_i}\|_\infty \le \delta\lambda_{\min}\big(R(\boldsymbol{x}_i)\big),$$

$$(1.4) \qquad \|R(\boldsymbol{x}_i) - H_{\omega_i}\|_\infty \le \varepsilon\lambda_{\min}\big(R(\boldsymbol{x}_i)\big)$$

hold for any mesh vertex $\boldsymbol{x}_i$ and its patch ω_i, where H_{ω_i} is the Hessian at a point in ω_i where $|\det H(\boldsymbol{x})|$ attains its maximum and $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue of a matrix. Notice that (1.4) does not require R to converge to H as the mesh is refined. Instead, it requires the eigenvalues of $R^{-1}H$ to be around one (cf. section 2). Unfortunately, it is still too restrictive to be satisfied by most examples we tested; see section 5. Thus, the work [1, 21] does not give a full explanation for why a nonconvergent recovered Hessian works in mesh adaptation.

The objective of the paper is to present a study on this issue. To be specific, we consider a BVP and its linear finite element solution with adaptive anisotropic meshes

generated from a recovered Hessian. We adopt the M-uniform mesh approach [12, 13] to view any adaptive mesh as a uniform one in some metric depending on the computed solution. An advantage of the approach is that the relation between the recovered Hessian and an adaptive anisotropic mesh generated using it can be fully characterized through the so-called alignment and equidistribution conditions (see (2.3) and (2.4) in section 2). This characterization plays a crucial role in the development of a bound for the H^1 seminorm of the finite element error. The bound converges at a first-order rate in terms of the average element diameter, $N^{-\frac{1}{d}}$, where N is the number of elements and d is the dimension of the physical domain. Moreover, the bound is valid under a condition on the closeness of the recovered Hessian to the exact one; see (2.12) or (3.2) and (3.3). This closeness condition is much weaker than (1.4). Roughly speaking, (1.4) requires the eigenvalues of $R^{-1}H$ to be around one, whereas the new condition only requires them to be bounded below from zero and from above. Numerical results in section 5 show that the new closeness condition is satisfied in all examples for four commonly used Hessian recovery techniques considered in this paper, whereas (1.4) is satisfied only in some examples. Furthermore, the error bound is linearly proportional to the ratio of the maximum (over the physical domain) of the largest eigenvalues of $R^{-1}H$ to the minimum of the smallest eigenvalues. Since the ratio is a measure of the closeness of the recovered Hessian to the exact one, the dependence indicates that the finite element error changes gradually with the closeness of the recovered Hessian. Hence, the error for the linear finite element approximation of the BVP is convergent for the considered Hessian recovery techniques and insensitive to the closeness of the recovered Hessian to the exact one. This provides an explanation of how a nonconvergent recovered Hessian works for mesh adaptation.

An outline of the paper is as follows. Convergence analysis of the linear finite element approximation is given in sections 2 and 3 for the cases with positive definite and general Hessian, respectively. A brief description of four common Hessian recovery techniques is given in section 4, followed by numerical examples in section 5. Finally, section 6 contains conclusions and further comments.

2. Convergence of linear finite element approximation for positive definite Hessian. We consider the BVP

(2.1)
$$\begin{cases} \mathcal{L}u = f & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega, \end{cases}$$

where Ω is a polygonal or polyhedral domain of $\mathbb{R}^d$ ($d \geq 1$), $\mathcal{L}$ is an elliptic second-order differential operator, and f and g are given functions. We are concerned with the adaptive mesh solution of this BVP using the conventional linear finite element method. Denote a family of simplicial meshes for Ω by $\{\mathcal{T}_h\}$ and the corresponding reference element by $\hat{K}$ which is chosen to be unitary in volume. For each mesh $\mathcal{T}_h$, we denote the corresponding finite element solution by u_h. Céa's lemma implies that the finite element error is bounded by the interpolation error, i.e.,

(2.2)
$$|u - u_h|_{H^1(\Omega)} \leq C \, |u - \Pi_h u|_{H^1(\Omega)},$$

where C is a constant independent of u and $\mathcal{T}_h$ and Π_h is the nodal interpolation operator associated with the linear finite element space defined on $\mathcal{T}_h$. Note that (2.2) is valid for any mesh.

2.1. Quasi-M-uniform meshes. In this paper we consider adaptive meshes generated based on a recovered Hessian R and use the M-uniform mesh approach with which any adaptive mesh is viewed as a uniform one in some metric M (defined in terms of R in our current situation). It is known [12, 13] that such an M-uniform mesh satisfies the equidistribution and alignment conditions

$$(2.3) \qquad |K|\det(M_K)^{\frac{1}{2}} = \frac{1}{N}\sum_{\tilde{K}\in\mathcal{T}_h}|\tilde{K}|\det(M_{\tilde{K}})^{\frac{1}{2}} \quad \forall K \in \mathcal{T}_h,$$

$$(2.4) \qquad \frac{1}{d}\operatorname{tr}\left((F_K')^T M_K F_K'\right) = \det\left((F_K')^T M_K F_K'\right)^{\frac{1}{d}} \quad \forall K \in \mathcal{T}_h,$$

where N is the number of mesh elements, M_K is an average of M over K, $F_K\colon \hat{K} \to K$ is the affine mapping from the reference element $\hat{K}$ to a mesh element K, F_K' is the Jacobian matrix of F_K (which is constant on K), and $\det(\cdot)$ and $\operatorname{tr}(\cdot)$ denote the determinant and trace of a matrix, respectively.

In practice, it is more realistic to generate less restrictive quasi-M-uniform meshes which satisfy

$$(2.5) \qquad |K|\det(M_K)^{\frac{1}{2}} \le C_{eq}\frac{1}{N}\sum_{\tilde{K}\in\mathcal{T}_h}|\tilde{K}|\det(M_{\tilde{K}})^{\frac{1}{2}} \quad \forall K \in \mathcal{T}_h,$$

$$(2.6) \qquad \frac{1}{d}\operatorname{tr}\left((F_K')^T M_K F_K'\right) \le C_{ali}\,|K|^{\frac{2}{d}}\det(M_K)^{\frac{1}{d}} \quad \forall K \in \mathcal{T}_h,$$

where $C_{eq}, C_{ali} \ge 1$ are some constants independent of K, N, and $\mathcal{T}_h$. Numerical experiments in [11] and section 5 (Figures 2(a) to 3(b)) show that quasi-M-uniform meshes with relatively small C_{eq} and C_{ali} can be generated in practice. For this reason, we use quasi-M-uniform meshes in our analysis and numerical experiments.

We would like to point out that conditions (2.5) and (2.6) with $C_{eq} = C_{ali} = 1$ imply (2.3) and (2.4). Indeed, the inequality (2.6) with $C_{ali} = 1$ becomes the equality (2.4) because the left-hand side of it (the arithmetic mean of the eigenvalues of $(F_K')^T M_K F_K'$) cannot be smaller than the right-hand side (the geometric mean of the eigenvalues). Further, if $C_{eq} = 1$, then (2.5) becomes

$$|K|\det(M_K)^{\frac{1}{2}} \le \frac{1}{N}\sum_{\tilde{K}\in\mathcal{T}_h}|\tilde{K}|\det(M_{\tilde{K}})^{\frac{1}{2}} \quad \forall K \in \mathcal{T}_h.$$

This implies

$$\max_{K\in\mathcal{T}_h}|K|\det(M_K)^{\frac{1}{2}} \le \frac{1}{N}\sum_{K\in\mathcal{T}_h}|K|\det(M_K)^{\frac{1}{2}}$$

$$\le \frac{1}{N}\left((N-1)\max_{K\in\mathcal{T}_h}|K|\det(M_K)^{\frac{1}{2}} + \min_{K\in\mathcal{T}_h}|K|\det(M_K)^{\frac{1}{2}}\right)$$

and therefore

$$\max_{K\in\mathcal{T}_h}|K|\det(M_K)^{\frac{1}{2}} \le \min_{K\in\mathcal{T}_h}|K|\det(M_K)^{\frac{1}{2}},$$

which can be valid only if all values of $|K|\det(M_K)^{\frac{1}{2}}$ are the same for all K.

(a) elementwise $C_{eq,K}$ for (2.5)

(b) elementwise $C_{ali,K}$ for (2.6)

(c) elementwise ε for (5.1)

(d) elementwise $C_{R+,K}/C_{R-}$

(e) ε for (5.1)

(f) C_{R+}/C_{R-} for (3.2) and (3.3)

(g) finite element error $|u - u_h|_{H^1(\Omega)}$

(h) comparison of α_h and α

Fig. 2. *Numerical results for Example* 5.1.

(a) elementwise $C_{eq,K}$ for (2.5)

(b) elementwise $C_{ali,K}$ for (2.6)

(c) elementwise ε for (5.1)

(d) elementwise $C_{R+,K}/C_{R-}$

(e) ε for (5.1)

(f) C_{R+}/C_{R-} for (3.2) and (3.3)

(g) finite element error $|u - u_h|_{H^1(\Omega)}$

(h) comparison of α_h and α
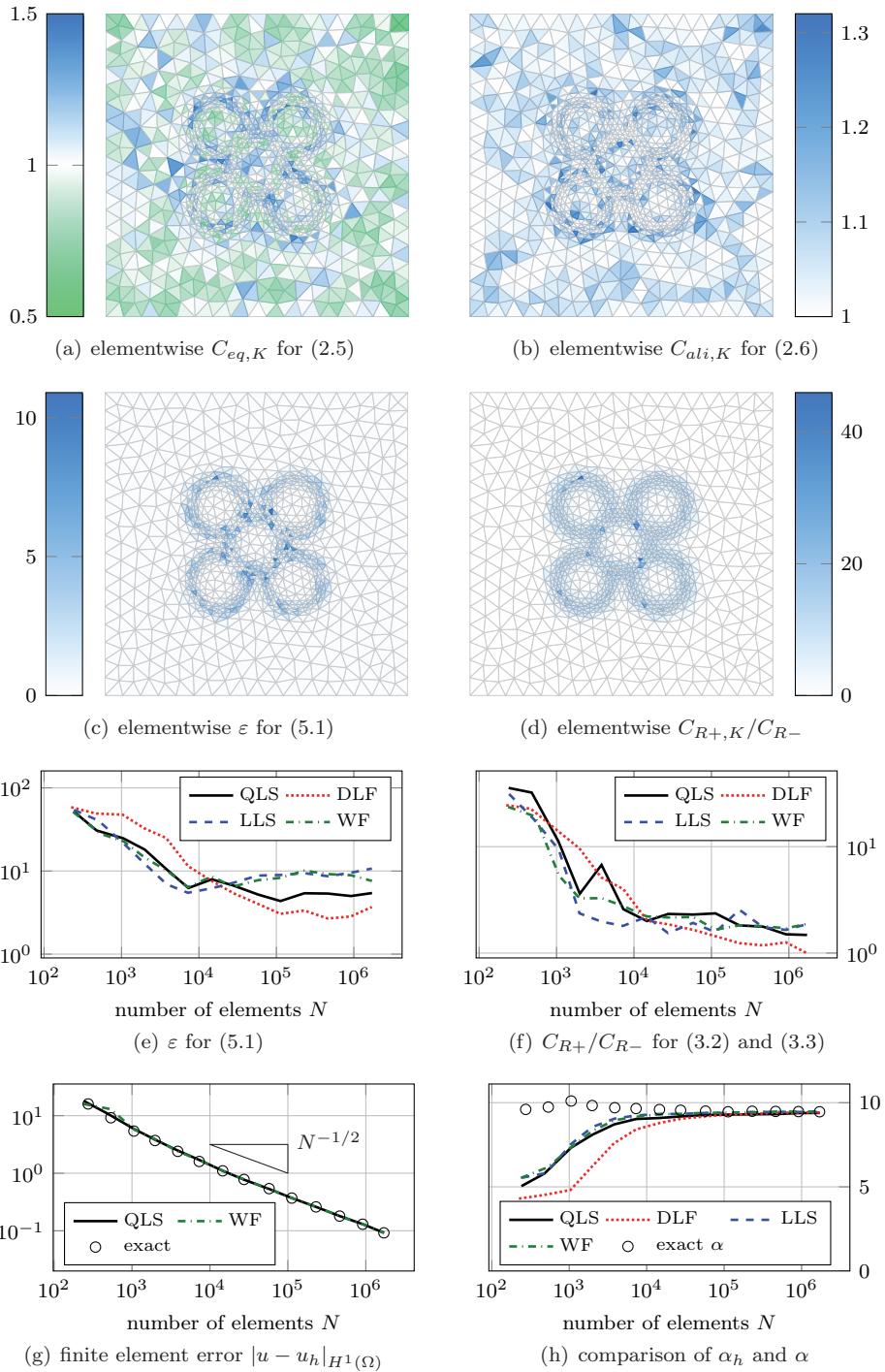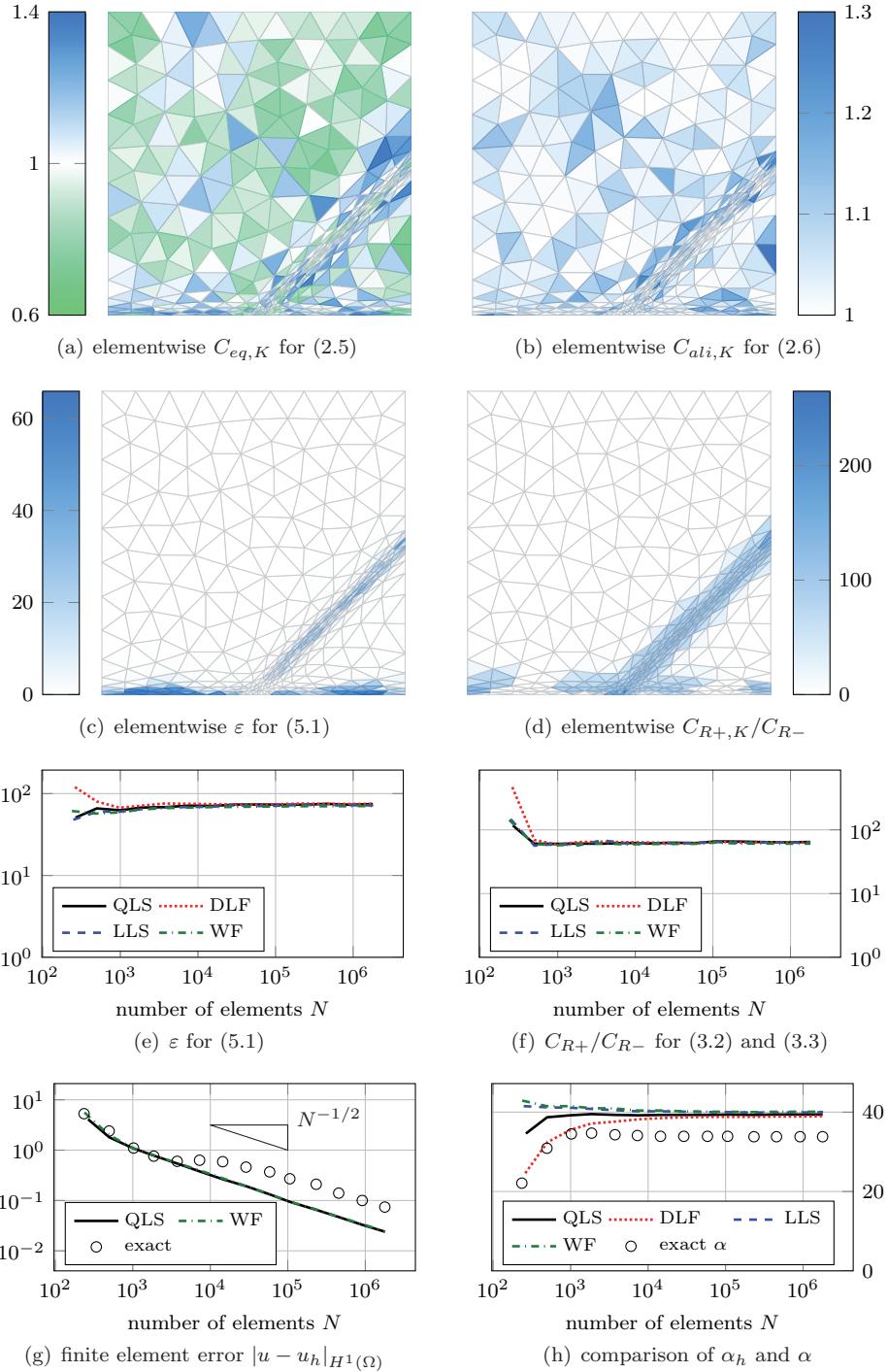
FIG. 3. *Numerical results for Example* 5.2.

2.2. Main result. In this section we consider a special case where the Hessian of the solution is uniformly positive definite in Ω, i.e.,

$$(2.7) \qquad \exists \gamma > 0 \colon H(\boldsymbol{x}) \geq \gamma I \quad \forall \boldsymbol{x} \in \Omega,$$

where the greater-than-or-equal sign means that the difference between the left-hand side and right-hand side terms is positive semidefinite. We also assume that the recovered Hessian R is uniformly positive definite in Ω. This assumption is not essential and will be dropped for the general situation discussed in section 3.

Recall from (2.2) that the finite element error is bounded by the H^1 seminorm of the interpolation error of the exact solution. A metric tensor corresponding to the H^1 seminorm can be defined as

$$(2.8) \qquad M_K = \det(R_K)^{-\frac{1}{d+2}} \, \|R_K\|_2^{\frac{2}{d+2}} \, R_K \quad \forall K \in \mathcal{T}_h,$$

where R_K is an average of R over K [11]. For this metric tensor, mesh conditions (2.5) and (2.6) become

$$(2.9) \quad |K| \det(R_K)^{\frac{1}{d+2}} \, \|R_K\|_2^{\frac{d}{d+2}} \leq C_{eq} \frac{1}{N} \sum_{\tilde{K}} |\tilde{K}| \det(R_{\tilde{K}})^{\frac{1}{d+2}} \, \|R_{\tilde{K}}\|_2^{\frac{d}{d+2}} \quad \forall K \in \mathcal{T}_h,$$

$$(2.10) \qquad \frac{1}{d} \operatorname{tr}\left((F_K')^T R_K F_K' \right) \leq C_{ali} \, |K|^{\frac{2}{d}} \det(R_K)^{\frac{1}{d}} \quad \forall K \in \mathcal{T}_h.$$

Note that the alignment condition (2.10) implies the inverse alignment condition

$$(2.11)$$

$$\frac{1}{d} \operatorname{tr}\left((F_K')^{-1} R_K^{-1} (F_K')^{-T} \right) < \left(\frac{d}{d-1} C_{ali} \right)^{d-1} |K|^{-\frac{2}{d}} \det(R_K)^{-\frac{1}{d}} \quad \forall K \in \mathcal{T}_h.$$

To show this, we denote the eigenvalues of $(F_K')^T R_K F_K'$ by $0 < \lambda_1 \leq \cdots \leq \lambda_d$ and rewrite (2.10) as

$$\sum_i \lambda_i \leq d C_{ali} \left(\prod_i \lambda_i \right)^{\frac{1}{d}}.$$

Then (2.11) follows from

$$\begin{aligned}
\frac{1}{d} \sum_i \lambda_i^{-1} &= \prod_i \lambda_i^{-1} \cdot \frac{1}{d} \sum_i \prod_{j \neq i} \lambda_j \\
&\leq \prod_i \lambda_i^{-1} \cdot \frac{1}{d} \sum_i \left(\frac{\sum_{j \neq i} \lambda_j}{d-1} \right)^{d-1} \\
&< \prod_i \lambda_i^{-1} \cdot \frac{1}{d} \sum_i \left(\frac{\sum_j \lambda_j}{d-1} \right)^{d-1} = \prod_i \lambda_i^{-1} \left(\frac{\sum_j \lambda_j}{d-1} \right)^{d-1} \\
&\leq \left(\frac{d}{d-1} C_{ali} \right)^{d-1} \left(\prod_i \lambda_i \right)^{-\frac{1}{d}}.
\end{aligned}$$

THEOREM 2.1 (positive definite Hessian). *Assume that $H(\boldsymbol{x})$ and the recovered Hessian R are uniformly positive definite in Ω and that R satisfies*

$$(2.12) \qquad C_{R-,K} I \leq R_K^{-1} H(\boldsymbol{x}) \leq C_{R+,K} I \quad \forall \boldsymbol{x} \in K, \quad \forall K \in \mathcal{T}_h,$$

where $C_{R-,K}$ and $C_{R+,K}$ are elementwise constants satisfying

$$(2.13) \qquad C_{R-} \leq \min_{K \in \mathcal{T}_h} C_{R-,K} \quad and \quad \sqrt{\frac{1}{N} \sum_{K \in \mathcal{T}_h} C_{R+,K}^2} \leq C_{R+}$$

with some mesh-independent positive constants C_{R-} and C_{R+}. If the solution of the BVP (2.1) is in $H^2(\Omega)$, then for any quasi-M-uniform mesh associated with the metric tensor (2.8) and satisfying (2.5) and (2.6) the linear finite element error for the BVP is bounded by

$$(2.14) \qquad |u - u_h|_{H^1(\Omega)} \leq C \cdot C_{ali}^{\frac{d+1}{2}} C_{eq}^{\frac{d+2}{2d}} \cdot \frac{C_{R+}}{C_{R-}} \cdot N^{-\frac{1}{d}} \left\| \det(H)^{\frac{1}{d}} H \right\|_{L^{\frac{d}{d+2}}(\Omega)}^{\frac{1}{2}}.$$

Proof. The nodal interpolation error of a function $u \in H^2(\Omega)$ on K is bounded by

$$(2.15) \qquad |u - \Pi_h u|_{H^1(K)} \leq C \left\| (F_K')^{-1} \right\|_2 \left(\int_K \left\| (F_K')^T |H(\boldsymbol{x})| F_K' \right\|_2^2 d\boldsymbol{x} \right)^{\frac{1}{2}},$$

where $|H(\boldsymbol{x})| = \sqrt{H(\boldsymbol{x})^2}$ [13, Theorem 5.1.5]. (The interested reader is referred to, for example, [7, 9, 11, 14, 16] for anisotropic error estimates for interpolation with linear and higher-order finite elements.) Notice that $|H(\boldsymbol{x})| = H(\boldsymbol{x})$ in the current situation (symmetric and positive definite $H(\boldsymbol{x})$).

Further,

$$
\begin{aligned}
\left\| (F_K')^T H(\boldsymbol{x}) F_K' \right\|_2 &= \left\| H(\boldsymbol{x})^{\frac{1}{2}} F_K' \right\|_2^2 = \left\| H(\boldsymbol{x})^{\frac{1}{2}} R_K^{-\frac{1}{2}} R_K^{\frac{1}{2}} F_K' \right\|_2^2 \\
&\leq \left\| H(\boldsymbol{x})^{\frac{1}{2}} R_K^{-\frac{1}{2}} \right\|_2^2 \left\| R_K^{\frac{1}{2}} F_K' \right\|_2^2 \\
&= \left\| R_K^{-\frac{1}{2}} H(\boldsymbol{x}) R_K^{-\frac{1}{2}} \right\|_2 \left\| (F_K')^T R_K F_K' \right\|_2 \\
&= \lambda_{\max}\left(R_K^{-\frac{1}{2}} H(\boldsymbol{x}) R_K^{-\frac{1}{2}} \right) \left\| (F_K')^T R_K F_K' \right\|_2 \\
&= \lambda_{\max}\left(R_K^{-1} H(\boldsymbol{x}) \right) \left\| (F_K')^T R_K F_K' \right\|_2 \\
&\leq \left\| R_K^{-1} H(\boldsymbol{x}) \right\|_2 \left\| (F_K')^T R_K F_K' \right\|_2.
\end{aligned}
$$

Similarly,

$$
\left\| (F_K')^{-1} \right\|_2^2 = \left\| (F_K')^{-1} (F_K')^{-T} \right\|_2 \leq \left\| (F_K')^{-1} R_K^{-1} (F_K')^{-T} \right\|_2 \| R_K \|_2.
$$

Thus, (2.15) yields

$$
\begin{aligned}
&|u - \Pi_h u|_{H^1(K)}^2 \\
&\quad \leq C \left\| (F_K')^{-1} R_K^{-1} (F_K')^{-T} \right\|_2 \| R_K \|_2 \int_K \left\| (F_K')^T R_K F_K' \right\|_2^2 \left\| R_K^{-1} H(\boldsymbol{x}) \right\|_2^2 d\boldsymbol{x}.
\end{aligned}
$$

Using this, (2.6), (2.11), (2.12), the fact that the trace of any $d \times d$ symmetric and positive definite matrix A is equivalent to its l^2 norm, viz., $\|A\|_2 \leq \operatorname{tr}(A) \leq d \|A\|_2$, and absorbing powers of d into the generic constant C, we get

$$
\begin{aligned}
&|u - \Pi_h u|^2_{H^1(\Omega)} \\
&= \sum_K |u - \Pi_h u|^2_{H^1(K)} \\
&\leq C \sum_K C^{d-1}_{ali} |K|^{-\frac{2}{d}} \det(R_K)^{-\frac{1}{d}} \|R_K\|_2 \times |K| C^2_{ali} |K|^{\frac{4}{d}} \det(R_K)^{\frac{2}{d}} C^2_{R+,K} \\
&= C C^{d+1}_{ali} \sum_K |K|^{\frac{d+2}{d}} \det(R_K)^{\frac{1}{d}} \|R_K\|_2 C^2_{R+,K} \\
&= C C^{d+1}_{ali} \sum_K \left(|K| \det(R_K)^{\frac{1}{d+2}} \|R_K\|^{\frac{d}{d+2}}_2 \right)^{\frac{d+2}{d}} C^2_{R+,K}.
\end{aligned}
$$

Applying (2.5) to the above result and using (2.13) gives

$$
\begin{aligned}
&|u - \Pi_h u|^2_{H^1(\Omega)} \\
&\leq C C^{d+1}_{ali} \sum_K \left(\frac{C_{eq}}{N} \sum_{\tilde{K}} |\tilde{K}| \det(R_{\tilde{K}})^{\frac{1}{d+2}} \|R_{\tilde{K}}\|^{\frac{d}{d+2}}_2 \right)^{\frac{d+2}{d}} C^2_{R+,K} \\
&= C C^{d+1}_{ali} C^{\frac{d+2}{d}}_{eq} N^{-\frac{2}{d}} \left(\frac{1}{N} \sum_{K \in \mathcal{T}_h} C^2_{R+,K} \right) \left(\sum_{\tilde{K}} |\tilde{K}| \det(R_{\tilde{K}})^{\frac{1}{d+2}} \|R_{\tilde{K}}\|^{\frac{d}{d+2}}_2 \right)^{\frac{d+2}{d}} \\
&\leq C C^{d+1}_{ali} C^{\frac{d+2}{d}}_{eq} N^{-\frac{2}{d}} C^2_{R+} \left(\sum_K |K| \det(R_K)^{\frac{1}{d+2}} \|R_K\|^{\frac{d}{d+2}}_2 \right)^{\frac{d+2}{d}} \\
&= C C^{d+1}_{ali} C^{\frac{d+2}{d}}_{eq} N^{-\frac{2}{d}} C^2_{R+} \left(\sum_K \int_K \det(R_K)^{\frac{1}{d+2}} \|R_K\|^{\frac{d}{d+2}}_2 \, d\boldsymbol{x} \right)^{\frac{d+2}{d}}.
\end{aligned}
$$

Further, assumption (2.12) implies

$$
(2.16) \qquad \det(R_K) \leq \det\big(H(\boldsymbol{x})\big) \left\| H^{-1}(\boldsymbol{x}) R_K \right\|^d_2 \leq C^{-d}_{R-} \det\big(H(\boldsymbol{x})\big)
$$

and

$$
(2.17) \quad \|R_K\|_2 = \left\| H(\boldsymbol{x}) H^{-1}(\boldsymbol{x}) R_K \right\|_2 \leq \|H(\boldsymbol{x})\|_2 \left\| H^{-1}(\boldsymbol{x}) R_K \right\|_2 \leq C^{-1}_{R-} \|H(\boldsymbol{x})\|_2.
$$

Thus,

$$
\begin{aligned}
&|u - \Pi_h u|^2_{H^1(\Omega)} \\
&\leq C C^{d+1}_{ali} C^2_{R+} C^{\frac{d+2}{d}}_{eq} N^{-\frac{2}{d}} \left(C^{\frac{-2d}{d+2}}_{R-} \int_\Omega \left(\det(H(\boldsymbol{x}))^{\frac{1}{d}} \|H(\boldsymbol{x})\|_2 \right)^{\frac{d+2}{d+2}} d\boldsymbol{x} \right)^{\frac{d+2}{d}} \\
&= C C^{d+1}_{ali} C^{\frac{d+2}{d}}_{eq} \left(\frac{C_{R+}}{C_{R-}} \right)^2 N^{-\frac{2}{d}} \left(\int_\Omega \left\| \det(H(\boldsymbol{x}))^{\frac{1}{d}} H(\boldsymbol{x}) \right\|^{\frac{d}{d+2}}_2 d\boldsymbol{x} \right)^{\frac{d+2}{d}},
\end{aligned}
$$

which, together with (2.2), gives (2.14). $\quad\square$

2.3. Remarks. Theorem 2.1 shows how a nonconvergent recovered Hessian works in mesh adaptation. The error bound (2.14) is linearly proportional to the ratio C_{R+}/C_{R-}, which is a measure for the closeness of R to H. Thus, the finite element error changes gradually with the closeness of the recovered Hessian. If R is a good approximation to H (but not necessarily convergent), then $C_{R+}/C_{R-} = \mathcal{O}(1)$ and the solution-dependent factor in the error bound is

$$(2.18) \qquad \left\| \det(H)^{\frac{1}{d}} H \right\|_{L^{\frac{d}{d+2}}(\Omega)}^{\frac{1}{2}}.$$

On the other hand, if R is not a good approximation to H, solution-dependent factor in the error bound will be larger. For example, consider $R = I$ (the identity matrix), which leads to the uniform mesh refinement. In this case condition (2.12) is satisfied with

$$C_{R+} = C_{R+,K} = \max_{\boldsymbol{x} \in \Omega} \lambda_{\max}\big(H(\boldsymbol{x})\big) \quad \text{and} \quad C_{R-} = \min_{\boldsymbol{x} \in \Omega} \lambda_{\min}\big(H(\boldsymbol{x})\big),$$

where $\lambda_{\max}\big(H(\boldsymbol{x})\big)$ and $\lambda_{\min}\big(H(\boldsymbol{x})\big)$ denote the maximum and minimum eigenvalues of $H(\boldsymbol{x})$, respectively. Thus, for $R = I$ the solution-dependent factor in the bound (2.14) becomes

$$\frac{\max_{\boldsymbol{x} \in \Omega} \lambda_{\max}\big(H(\boldsymbol{x})\big)}{\min_{\boldsymbol{x} \in \Omega} \lambda_{\min}\big(H(\boldsymbol{x})\big)} \left\| \det(H)^{\frac{1}{d}} H \right\|_{L^{\frac{d}{d+2}}(\Omega)}^{\frac{1}{2}},$$

which is obviously larger than (2.18).

Next, we study the relation between (2.12) and (1.3)–(1.4). In practical computation, the Hessian is typically recovered at mesh nodes (see section 4) and a recovered Hessian can be considered on the whole domain as a piecewise linear matrix-valued function. In this case, the average R_K of R over any given element K can be expressed as a linear combination of the nodal values of R. Applying the triangle inequality to (1.3) and (1.4) we get

$$\|R(\boldsymbol{x}_i) - H(\boldsymbol{x})\|_\infty \le (\delta + \varepsilon)\, \lambda_{\min}(R_{\boldsymbol{x}_i}) \quad \forall \boldsymbol{x} \in \omega_i$$

and, since R_K is a linear combination of $R(\boldsymbol{x}_i)$,

$$\|R_K - H(\boldsymbol{x})\|_\infty \le (\delta + \varepsilon)\, \lambda_{\min}(R_K).$$

Since $R_K - H(\boldsymbol{x})$ is symmetric, $\|R_K - H(\boldsymbol{x})\|_2 \le \|R_K - H(\boldsymbol{x})\|_\infty$. Thus, conditions (1.3) and (1.4) imply

$$(2.19) \qquad \|R_K - H(\boldsymbol{x})\|_2 \le (\delta + \varepsilon)\, \lambda_{\min}(R_K) \quad \forall \boldsymbol{x} \in K, \quad \forall K \in \mathcal{T}_h$$

and

$$(2.20) \qquad \left\| R_K^{-1} H(\boldsymbol{x}) - I \right\|_2 \le (\delta + \varepsilon) \quad \forall \boldsymbol{x} \in K, \quad \forall K \in \mathcal{T}_h,$$

which in turn implies (2.12) with $C_{R+,K} = 1 + (\delta + \varepsilon)$ and $C_{R-} = 1 - (\delta + \varepsilon)$, if $(\delta + \varepsilon) < 1$. Condition (2.20) and therefore (1.4) require the eigenvalues of $R_K^{-1} H(\boldsymbol{x})$ to stay closely around one. On the other hand, condition (2.12) only requires the eigenvalues of $R_K^{-1} H(\boldsymbol{x})$ to be bounded from above and below from zero, which is weaker than (1.4). If R converges to $H(\boldsymbol{x})$, both (1.4) and (2.12) can be satisfied.

However, if R does not converge to $H(\boldsymbol{x})$, as is the case for most adaptive computation, the situation is different. As we shall see in section 5, condition (2.12) is satisfied for all the examples tested, whereas condition (1.4) is not satisfied by either of the examples.

We would like to point out that it is unclear if the considered monitor function (2.8) (and the corresponding bound (2.14)) is optimal, although it seems to be the best we can get. For example, if we choose the monitor function to be

$$(2.21) \qquad M_K = \det(R_K)^{-\frac{1}{d+4}} R_K \quad \forall K \in \mathcal{T}_h,$$

which is optimal for the L^2 norm [11], the error bound becomes

$$(2.22) \quad |u - u_h|_{H^1(\Omega)}$$
$$\leq C \cdot C_{ali}^{\frac{d+1}{2}} C_{eq}^{\frac{d+4}{4d}} \cdot \frac{C_{R+}}{C_{R-}} \cdot N^{-\frac{1}{d}} \left\| \det(H)^{\frac{1}{d}} \right\|_{L^{\frac{2d}{d+4}}(\Omega)}^{\frac{2}{d+4}} \left\| \det(H)^{\frac{1}{d+4}} H \right\|_{L^1(\Omega)}^{\frac{1}{2}}.$$

This bound has a larger solution-dependent factor than (2.14) since Hölder's inequality yields

$$\left\| \det(H)^{\frac{1}{d}} H \right\|_{L^{\frac{d}{d+2}}(\Omega)}^{\frac{1}{2}} \leq \left\| \det(H)^{\frac{1}{d}} \right\|_{L^{\frac{2d}{d+4}}(\Omega)}^{\frac{2}{d+4}} \left\| \det(H)^{\frac{1}{d+4}} H \right\|_{L^1(\Omega)}^{\frac{1}{2}}.$$

It is worth mentioning that when the metric tensor (2.21) is used, the L^2 norm of the piecewise linear interpolation error is bounded by

$$(2.23) \qquad \|u - \Pi_h u\|_{L^2(\Omega)} \leq C \cdot C_{ali} C_{eq}^{\frac{d+4}{2d}} \cdot \frac{C_{R+}}{C_{R-}} \cdot N^{-\frac{2}{d}} \left\| \det(H)^{\frac{1}{d}} \right\|_{L^{\frac{2d}{d+4}}(\Omega)},$$

which is optimal in terms of convergence order and solution-dependent factor; e.g., see [7, 14].

Note that Theorem 2.1 holds for $u \in H^2(\Omega)$ although the estimate (2.14) only requires

$$(2.24) \qquad \left\| \det(H)^{\frac{1}{d}} H \right\|_{L^{\frac{d}{d+2}}(\Omega)} < \infty.$$

Since

$$\left\| \det(H)^{\frac{1}{d}} H \right\|_{L^{\frac{d}{d+2}}(\Omega)} \leq \left\| \frac{1}{d} \operatorname{tr}(H) \cdot H \right\|_{L^{\frac{d}{d+2}}(\Omega)},$$

(2.24) can be satisfied when $u \in W^{2, \frac{2d}{d+2}}(\Omega)$. Thus, there is a gap between the sufficient requirement $u \in H^2(\Omega)$ and the necessary requirement $u \in W^{2, \frac{2d}{d+2}}(\Omega)$. The stronger requirement $u \in H^2(\Omega)$ comes from the estimation of the interpolation error in [13, Theorem 5.1.5]. It is unclear to the authors whether this requirement can be weakened.

Note that $u \in H^2(\Omega)$ may not hold when $\partial\Omega$ is not smooth. For example, in two dimensions, if $\partial\Omega$ has a corner with an angle $\omega \in (0, 2\pi)$, the solution of the BVP (2.1) with smooth f and g basically has the following form near the corner:

$$u(r, \theta) = r^{\frac{\pi}{\omega}} u_0(\theta) + u_1(r, \theta),$$

where (r, θ) denote the polar coordinates and $u_0(\theta)$ and $u_1(r, \theta)$ are some smooth functions. Then,

$$|u|^2_{H^2(\Omega)} \sim \int_0^b \left(r^{\frac{\pi}{\omega}-2}\right)^2 r\, dr \sim \left. r^{\frac{2\pi}{\omega}-2}\right|_0^b$$

for some constant $b > 0$. This implies that $u \notin H^2(\Omega)$ if $\omega > \pi$. On the other hand, $W^{2, \frac{2d}{d+2}}(\Omega) = W^{2,1}(\Omega)$ for $d = 2$ and

$$|u|^2_{W^{2,1}(\Omega)} \sim \int_0^b \left(r^{\frac{\pi}{\omega}-2}\right) r\, dr \sim \left. r^{\frac{\pi}{\omega}}\right|_0^b,$$

which indicates that $u \in W^{2,1}(\Omega)$ for all $\omega \in (0, 2\pi)$.

3. Convergence of the linear finite element approximation for a general Hessian. In this section we consider the general situation where $H(\boldsymbol{x})$ is symmetric but not necessarily positive definite. In this case, it is unrealistic to require the recovered Hessian R to be positive definite. Thus, we cannot use R directly to define the metric tensor which is required to be positive definite. A commonly used strategy is to replace R by $|R| = \sqrt{R^2}$ since $|R|$ retains the eigensystem of R. However, $|R|$ can become singular locally. To avoid this difficulty, we regularize $|R|$ with a regularization parameter $\alpha_h > 0$ (to be determined).

From (2.8), we define the regularized metric tensor as

$$(3.1) \qquad M_K = \det(\alpha_h I + |R_K|)^{-\frac{1}{d+2}} \|\alpha_h I + |R_K|\|_2^{\frac{2}{d+2}} (\alpha_h I + |R_K|) \quad \forall K \in \mathcal{T}_h$$

and obtain the following theorem with a proof similar to that of Theorem 2.1.

THEOREM 3.1 (general Hessian). *For a given positive parameter $\alpha_h > 0$, we assume that the recovered Hessian R satisfies*

$$(3.2) \qquad C_{R-,K} I \le (\alpha_h I + |R_K|)^{-1} (\alpha_h I + |H(\boldsymbol{x})|) \quad \forall \boldsymbol{x} \in K, \quad \forall K \in \mathcal{T}_h,$$

$$(3.3) \qquad (\alpha_h I + |R_K|)^{-1} |H(\boldsymbol{x})| \le C_{R+,K} I \quad \forall \boldsymbol{x} \in K, \quad \forall K \in \mathcal{T}_h,$$

where $C_{R-,K}$ and $C_{R+,K}$ are elementwise constants satisfying (2.13). If the solution of the BVP (2.1) is in $H^2(\Omega)$, then for any quasi-M-uniform mesh associated with metric tensor (3.1) and satisfying (2.5) and (2.6) the linear finite element error for the BVP is bounded by

$$(3.4) \qquad |u - u_h|_{H^1(\Omega)}$$
$$\le C \cdot C_{ali}^{\frac{d+1}{2}} C_{eq}^{\frac{d+2}{2d}} \cdot \frac{C_{R+}}{C_{R-}} \cdot N^{-\frac{1}{d}} \left\| \det(\alpha_h I + |H|)^{\frac{1}{d}} (\alpha_h I + |H|) \right\|_{L^{\frac{d}{d+2}}(\Omega)}^{\frac{1}{2}}.$$

From (3.2) to (3.4) we see that the greater α_h is, the easier the recovered Hessian satisfies (3.2) and (3.3); however, the error bound increases as well. For example, consider the extreme case of $\alpha_h \to \infty$. In this case, (3.2) and (3.3) can be satisfied with $C_{R+} = C_{R-} = 1$ for any R. At the same time, the metric tensor defined in (3.1) has an asymptotic behavior $M_K \to \alpha_h^{\frac{4}{d+4}} I$ and the corresponding M-uniform mesh is a uniform mesh. Obviously, the right-hand side of (3.4) is large for this case. Another extreme case is $\alpha_h \to 0$, where (3.4) reduces to (2.14) if both R and $H(\boldsymbol{x})$ are positive definite.

We now consider the choice of α_h. We define a parameter α through the implicit equation

$$(3.5) \qquad \left\| \sqrt[d]{\det(\alpha I + |H|)} \cdot (\alpha I + |H|) \right\|_{L^{\frac{d}{d+2}}(\Omega)} = 2 \left\| \sqrt[d]{\det(|H|)} \cdot H \right\|_{L^{\frac{d}{d+2}}(\Omega)}.$$

The left-hand-side term is an increasing function of α. Moreover, the term is equal to the half of the right-hand-side term when $\alpha = 0$ and tends to infinity as $\alpha \to \infty$. Thus, from the intermediate value theorem we know that (3.5) has a unique solution $\alpha > 0$ if $\| \sqrt[d]{\det(|H|)} \cdot H \|_{L^{\frac{d}{d+2}}(\Omega)} > 0$. If we choose $\alpha_h = \alpha$, then the finite element error is bounded by

$$(3.6) \qquad |u - u_h|_{H^1(\Omega)} \le C \cdot C_{ali}^{\frac{d+1}{2}} C_{eq}^{\frac{d+2}{2d}} \cdot \frac{C_{R+}}{C_{R-}} \cdot 2N^{-\frac{1}{d}} \left\| \sqrt[d]{\det(|H|)} \cdot H \right\|_{L^{\frac{d}{d+2}}(\Omega)}^{\frac{1}{2}},$$

which is essentially the same as (2.14). Note that (3.5) is impractical since it requires the prior knowledge of $H(\boldsymbol{x})$. In practice it can be replaced by

$$(3.7) \qquad \sum_K |K| \det(\alpha_h I + |R_K|)^{\frac{1}{d+2}} \|\alpha_h I + |R_K|\|_2^{\frac{d}{d+2}}$$

$$= 2^{\frac{d}{d+2}} \sum_K |K| \det(|R_K|)^{\frac{1}{d+2}} \|R_K\|_2^{\frac{d}{d+2}}.$$

This equation can be solved effectively using the bisection method. Numerical results show that α_h is close to α (Figure 3(h)).

4. A selection of commonly used Hessian recovery methods. In this section we give a brief description of four commonly used Hessian recovery algorithms for two-dimensional mesh adaptation. The interested reader is referred to [15, 20] for a more detailed description of these Hessian recovery techniques.

Recall that the goal of the Hessian recovery in the current context is to find an approximation of the Hessian in mesh nodes using the linear finite element solution u_h. The approximation of the Hessian on an element is calculated as the average of the nodal approximations of the Hessian at the vertices of the element.

Quadratic least squares fitting to nodal values (QLS). This method involves the fitting of a quadratic polynomial to nodal values of u_h at a selection of neighboring nodes in the least square sense and subsequent differentiation. The original purpose of the QLS was the gradient recovery (e.g., see Zhang and Naga [24]). However, it is easily adopted for the Hessian recovery by simply differentiating the fitting polynomial twice.

More specifically, for a given node (say, $\boldsymbol{x}_0$) at least five neighboring nodes are selected. A quadratic polynomial (denoted by p) is found by least squares fitting to the values of u_h at the selected nodes. The linear system associated with the least squares problem usually has full rank and a unique solution. If it does not, additional nodes from the neighborhood of $\boldsymbol{x}_0$ are added to the selection until the system has full rank. An approximation to the Hessian of the solution u at $\boldsymbol{x}_0$ is defined as the Hessian of p, viz.,

$$R^{QLS}(\boldsymbol{x}_0) = H(p)(\boldsymbol{x}_0).$$

Double linear least squares fitting (DLF). The double linear fitting method computes the Hessian by using linear least squares fitting twice. First, the least squares fitting of the nodal values of u_h in a neighborhood of $\boldsymbol{x}_0$ is employed to find a linear fitting polynomial p. The recovered gradient of function u at $\boldsymbol{x}_0$ is defined as the gradient of p at $\boldsymbol{x}_0$, i.e.,

$$\nabla_h^{DLF} u(\boldsymbol{x}_0) = \nabla p(\boldsymbol{x}_0).$$

Second-order derivatives are then obtained by subsequent application of this linear fitting to the calculated first-order derivatives. Mixed derivatives are averaged in order to obtain a symmetric recovered Hessian.

Linear least squares fitting to first-order derivatives (LLS). This method is similar to DLF except that the first-order derivatives at nodes are calculated in a different way. In this method, the first-order derivatives are first calculated at element centers and then at nodes by linear least squares fitting to their values at element centers.

Weak formulation (WF). This approach recovers the Hessian by means of a variational formulation [8]. More specifically, let ϕ_0 be a canonical piecewise linear basis function at node $\boldsymbol{x}_0$. Then the nodal approximation $u_{xx,h}$ to the second-order derivative u_{xx} at $\boldsymbol{x}_i$ is defined through

$$u_{xx,h}(\boldsymbol{x}_0) \int_\Omega \phi_0(\boldsymbol{x})\,d\boldsymbol{x} = -\int_\Omega \frac{\partial u_h}{\partial x}\frac{\partial \phi_0}{\partial x}\,d\boldsymbol{x}.$$

The same approach is used to compute $u_{xy,h}$ and $u_{yy,h}$. Since ϕ_0 are piecewise linear and vanish outside the patch associated with $\boldsymbol{x}_0$, the involved integrals can be computed efficiently with appropriate quadrature formulas over a single patch.

5. Numerical examples. In this section we present two numerical examples to verify the analysis given in the previous sections. We use `BAMG` [10] to generate adaptive meshes as quasi-M-uniform meshes for the regularized metric tensor (3.1). Special attention will be paid to mesh conditions (2.5) and (2.6) and closeness conditions (1.4), (3.2), and (3.3).

For the recovery closeness condition (1.4) we compare the regularized recovered and exact Hessians, i.e., we compute ε for

$$(5.1)\qquad \|(\alpha_h I + |R_K|) - (\alpha I + |H_K|)\|_\infty \leq \varepsilon\lambda_{\min}(\alpha_h I + |R_K|)\qquad \forall K \in \mathcal{T}_h,$$

where H_K is an average of the exact Hessian on the element K and α_h and α are the regularization parameters for the recovered and the exact Hessians, respectively.

Example 5.1 (see [11, Example 4.3]). The first example is in the form of BVP (1.1) with f and g chosen such that the exact solution is given by

$$\begin{aligned}
u(x,y) = {}& \tanh[30(x^2 + y^2 - 0.125)] \\
& + \tanh[30((x - 0.5)^2 + (y - 0.5)^2 - 0.125)] \\
& + \tanh[30((x - 0.5)^2 + (y + 0.5)^2 - 0.125)] \\
& + \tanh[30((x + 0.5)^2 + (y - 0.5)^2 - 0.125)] \\
& + \tanh[30((x + 0.5)^2 + (y - 0.5)^2 - 0.125)].
\end{aligned}$$

A typical plot of elementwise constants $C_{eq,K}$ and $C_{ali,K}$ in mesh quasi-M-uniformity conditions (2.5) and (2.6) is shown in Figures 2(a) and 2(b), demonstrating that these conditions hold with relatively small C_{eq} and C_{ali}. For the given mesh example we have $0.5 \leq C_{eq,K} \leq 1.5$ and $1 \leq C_{ali,K} \leq 1.3$, which gives $C_{eq} = 1.5$ and $C_{ali} = 1.3$. In fact, we found that $C_{eq} \leq 2.0$ and $C_{ali} \leq 2.1$ for all computations in this paper, indicating that BAMG does a good job in generating quasi-M-uniform meshes for a given metric tensor.

Figures 2(c) and 2(e) show a typical distribution of elementwise values of ε in (5.1) and its values for a sequence of adaptive grids. We observe that for all methods ε is not small with respect to one, which violates condition (1.4).

Typical elementwise values $C_{R+,K}/C_{R-}$ and values of C_{R+}/C_{R-} for a sequence of adaptive grids are shown in Figures 2(d) and 2(f). Notice that C_{R+}/C_{R-} stays relatively small and bounded, thus satisfying the closeness conditions (3.2) and (3.3).

For this example, the finite element error $|u - u_h|_{H^1(\Omega)}$ is almost undistinguishable for meshes obtained by means of the exact and recovered Hessian (Figure 2(g)) and the approximated α_h, computed through (3.7), is very close the value for the exact Hessian (Figure 2(h)).

Example 5.2 (strong anisotropy). The second example is in the form of BVP (1.1) with f and g chosen such that the exact solution is given by

$$u(x,y) = \tanh(60y) - \tanh\big(60(x - y) - 30\big).$$

This solution exhibits a very strong anisotropic behavior and describes the interaction between a boundary layer along the x-axis and a steep shock wave along the line $y = x - 1/2$.

Figures 3(c) and 3(e) show that $\varepsilon \approx 60$ and therefore not small with respect to one, violating condition (1.4) for all meshes in the considered range of N for all four recovery techniques.

On the other hand, Figure 3(f) shows that the ratio C_{R+}/C_{R-} is large ($\approx 10^2$) but, nevertheless, it seems to stay bounded with increasing N, confirming that (3.2) and (3.3) are satisfied by the recovered Hessian. The fact that the ratio C_{R+}/C_{R-} has different values in this and the previous examples indicates that the accuracy or closeness of the four Hessian recovery techniques depends on the behavior and especially the anisotropy of the solution. Fortunately, as shown by Theorems 2.1 and 3.1, the finite element error is insensitive to the closeness of the recovered Hessian. The finite element solution error is shown in Figure 3(g) as a function of N.

Finally, Figure 3(h) shows that α_h, computed through (3.7), is close to the exact value α defined in (3.5).

6. Conclusion and further comments. In the previous sections we have investigated how a nonconvergent recovered Hessian works in mesh adaptation. Our main results are Theorems 2.1 and 3.1, where an error bound for the linear finite element solution of BVP (2.1) is given for quasi-M-uniform meshes corresponding to a metric depending on a recovered Hessian. As conventional error estimates for the H^1 seminorm of the error in linear finite element approximations, our error bound is of first order in terms of the average element diameter, $N^{-\frac{1}{d}}$, where N is the number of elements and d is the dimension of the physical domain. This error bound is valid under the closeness condition (2.12) (or (3.2) and (3.3)), which is weaker than (1.4) used in [1] and [21]. Numerical results in section 5 show that the new closeness condition is satisfied by the recovered Hessian obtained with commonly used Hessian recovery algorithms. The error bound also shows that the finite element error changes gradually

with the closeness of the recovered Hessian to the exact one. These results provide an explanation of how a nonconvergent recovered Hessian works in mesh adaptation.

In this work the closeness conditions (3.2) and (3.3) have been verified only numerically. Developing a theoretical proof of the condition for some Hessian recovery techniques is an interesting topic for further investigations.

REFERENCES

[1] A. AGOUZAL, K. LIPNIKOV, AND YU. VASSILEVSKI, *Adaptive generation of quasi-optimal tetrahedral meshes*, East-West J. Numer. Math., 7 (1999), pp. 223–244.

[2] A. AGOUZAL, K. LIPNIKOV, AND YU. VASSILEVSKI, *Hessian-free metric-based mesh adaptation via geometry of interpolation error*, Zh. Vychisl. Mat. Mat. Fiz., 50 (2010), pp. 131–145.

[3] M. AINSWORTH AND J. T. ODEN, *A Posteriori Error Estimation in Finite Element Analysis*, Pure Appl. Math. (N.Y.), Wiley-Interscience, New York, 2000.

[4] I. BABUŠKA AND T. STROUBOULIS, *The finite element method and its reliability*, in Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2001.

[5] R. E. BANK AND J. XU, *Asymptotically exact a posteriori error estimators.* I. *Grids with superconvergence*, SIAM J. Numer. Anal., 41 (2003), pp. 2294–2312.

[6] R. E. BANK AND J. XU, *Asymptotically exact a posteriori error estimators.* II. *General unstructured grids*, SIAM J. Numer. Anal., 41 (2003), pp. 2313–2332.

[7] L. CHEN, P. SUN, AND J. XU, *Optimal anisotropic meshes for minimizing interpolation errors in L^p-norm*, Math. Comp., 76 (2007), pp. 179–204.

[8] V. DOLEJŠÍ, *Anisotropic mesh adaptation for finite volume and finite element methods on triangular meshes*, Comput. Vis. Sci., 1 (1998), pp. 165–178.

[9] L. FORMAGGIA AND S. PEROTTO, *New anisotropic a priori error estimates*, Numer. Math., 89 (2001), pp. 641–667.

[10] F. HECHT, *BAMG: Bidimensional Anisotropic Mesh Generator.* http://www.ann.jussieu.fr/hecht/ftp/bamg/.

[11] W. HUANG, *Metric tensors for anisotropic mesh generation*, J. Comput. Phys., 204 (2005), pp. 633–665.

[12] W. HUANG, *Anisotropic mesh adaptation and movement*, in Adaptive computations: Theory and Algorithms, T. Tang and J. Xu, eds., Math. Monogr. Ser. 6, Science Press, Beijing, 2007, pp. 68–158.

[13] W. HUANG AND R. D. RUSSELL, *Adaptive Moving Mesh Methods*, Appl. Math. Sci. 174, Springer, New York, 2011.

[14] W. HUANG AND W. SUN, *Variational mesh adaptation.* II. *Error estimates and monitor functions*, J. Comput. Phys., 184 (2003), pp. 619–648.

[15] L. KAMENSKI, *Anisotropic Mesh Adaptation Based on Hessian Recovery and A Posteriori Error Estimates*, Dissertation, Technische Universität Darmstadt, 2009.

[16] J.-M. MIREBEAU, *Optimally adapted meshes for finite elements of arbitrary order and $W^{1,p}$ norms*, Numer. Math., 120 (2012), pp. 271–305.

[17] J. S. OVALL, *Function, gradient, and Hessian recovery using quadratic edge-bump functions*, SIAM J. Numer. Anal., 45 (2007), pp. 1064–1080.

[18] M. PICASSO, F. ALAUZET, H. BOROUCHAKI, AND P.-L. GEORGE, *A numerical study of some Hessian recovery techniques on isotropic and anisotropic meshes*, SIAM J. Sci. Comput., 33 (2011), pp. 1058–1076.

[19] T. TANG AND J. XU, EDS., *Adaptive Computations: Theory and Algorithms*, Math. Monogr. Ser. 6, Science Press, Beijing, 2007.

[20] M.-G. VALLET, C.-M. MANOLE, J. DOMPIERRE, S. DUFOUR, AND F. GUIBAULT, *Numerical comparison of some Hessian recovery techniques*, Internat. J. Numer. Methods Engrg., 72 (2007), pp. 987–1007.

[21] YU. VASSILEVSKI AND K. LIPNIKOV, *An adaptive algorithm for constructing quasi-optimal grids*, Zh. Vychisl. Mat. Mat. Fiz., 39 (1999), pp. 1532–1551.

[22] H. WU AND Z. ZHANG, *Can we have superconvergent gradient recovery under adaptive meshes?*, SIAM J. Numer. Anal., 45 (2007), pp. 1701–1722.

[23] J. Xu AND Z. Zhang, *Analysis of recovery type a posteriori error estimators for mildly structured grids*, Math. Comp., 73 (2004), pp. 1139–1152.

[24] Z. Zhang AND A. Naga, *A new finite element gradient recovery method: superconvergence property*, SIAM J. Sci. Comput., 26 (2005), pp. 1192–1213.

[25] Z. Zhang AND J. Zhu, *Analysis of the superconvergent patch recovery technique and a posteriori error estimator in the finite element method.* I, Comput. Methods Appl. Mech. Engrg., 123 (1995), pp. 173–187.

[26] O. C. Zienkiewicz AND J. Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates.* I. *The recovery technique*, Internat. J. Numer. Methods Engrg., 33 (1992), pp. 1331–1364.

[27] O. C. Zienkiewicz AND J. Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates.* II. *Error estimates and adaptivity*, Internat. J. Numer. Methods Engrg., 33 (1992), pp. 1365–1382.