# ON THE MONOTONICITY OF HIGH ORDER DISCRETE LAPLACIAN \*

#### 3

#### LOGAN J. CROSS <sup>†</sup> AND XIANGXIONG ZHANG <sup>‡</sup>

**Abstract.** The monotonicity of discrete Laplacian, i.e., inverse positivity of stiffness matrix, implies discrete maximum principle, which is in general not true for high order schemes on unstructured meshes. But on structured meshes, it is possible to have high order accurate monotone schemes. We first review previously known high order accurate inverse positive schemes, all of which are fourth order accurate with proven monotonicity on uniform meshes. Then we discuss the monotonicity of a fourth order variational difference scheme on quasi-uniform meshes and prove the inverse positivity of a fifth order accurate variational difference scheme on a uniform mesh.

Key words. Inverse positivity, discrete maximum principle, high order accuracy, monotonicity,
 discrete Laplacian, quasi uniform meshes

#### 13 AMS subject classifications. 65N30, 65N06, 65N12

14 **1.** Introduction. In many applications, monotone discrete Laplacian operators are desired and useful for ensuring stability such as discrete maximum principle [8] or 15 positivity-preserving of physically positive quantities. Let  $\Delta_h$  denote the matrix repre-16 sentation of a discrete Laplacian operator, then it is called *monotone* if  $(-\Delta_h)^{-1} \ge 0$ , 17i.e., the matrix  $(-\Delta_h)^{-1}$  has nonnegative entries. In this paper, all inequalities for 18 matrices are entry-wise inequalities. The simplest second order accurate centered finite difference  $u''(x_i) \approx \frac{u(x_{i-1})-2u(x_i)+u(x_{i+1})}{\Delta x^2}$  is monotone because the corresponding matrix  $(-\Delta_h)^{-1}$  is an M-matrix thus inverse positive. The most general extension of 1920 21 this result is to state that linear finite element method under a mild mesh constraint 22 forms an M-matrix thus monotone on unstructured triangular meshes [25]. 23

In general, the discrete maximum principle is not true for high order finite element 24methods on unstructured meshes [13]. On the other hand, there exist a few high order 26 accurate inverse positive schemes on structured meshes. To the best of our knowledge, the followings schemes for solving a Poisson equation are the only ones proven to be 27monotone beyond the second order accuracy and all of them are fourth order accurate: 28 1. Fourth order compact finite difference schemes, including the classical 9-point 29scheme [15, 10, 2] are monotone because the stiffness matrix is an M-matrix. 30 2. In [3, 5], a fourth order accurate finite difference scheme was constructed. 31 The stiffness matrix is a product of two M-matrices thus monotone. 32 3. The Lagrangian  $P^2$  finite element method on a regular triangular mesh [24] 33 has a monotone stiffness matrix [20]. On an equilateral triangular mesh, the 34 discrete maximum principle can also be proven [13]. It can be regarded as a 35 finite difference scheme at vertices and edge centers, on which superconver-36 gence of fourth order accuracy holds. 37

4. Monotonicity was proven in the simplest finite difference implementation of Lagrangian  $Q^2$  finite element scheme on an uniform rectangular mesh for a variable coefficient Poisson equation under suitable mesh constraints [18].

All schemes above can be written in the form  $S\mathbf{u} = M\mathbf{f}$  with  $S^{-1} \ge 0$  and  $M \ge 0$ ,  $(-\Lambda_{-1})^{-1} = C^{-1}M \ge 0$ , where M is the form  $S\mathbf{u} = M\mathbf{f}$  with  $S^{-1} \ge 0$  and  $M \ge 0$ .

<sup>42</sup> thus  $(-\Delta_h)^{-1} = S^{-1}M \ge 0$ , where *M* denotes the mass matrix. The last two methods

<sup>\*</sup>L. Cross and X. Zhang were supported by the NSF grant DMS-1913120.

<sup>&</sup>lt;sup>†</sup>One Allison Way, Indianapolis, IN 46222 (logancross68@gmail.com).

<sup>&</sup>lt;sup>‡</sup>Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067 zhan1966@purdue.edu

are variational finite difference schemes, i.e., finite difference schemes constructed from
the variational formulation, thus they do not suffer from the drawbacks of the first
two conventional finite difference schemes, such as loss of accuracy on quasi-uniform
meshes, difficulty with other types of boundary conditions, etc.

For proving inverse positivity, the main viable tool in the literature is to use M-47 matrices which are inverse positive. All off-diagonal entries of M-matrices must be 48 non-positive. Except the fourth order compact finite difference, all high order accurate 49 schemes induce positive off-diagonal entries, destroying M-matrix structure, which is 50a major challenge of proving monotonicity. In [5] and [1], and also the appendix in [18], M-matrix factorizations of the form  $(-\Delta_h)^{-1} = M_1 M_2$  were shown for special high order schemes but these M-matrix factorizations seem ad hoc and do not apply 53 to other schemes or other equations. In [20], Lorenz proposed some matrix entry-wise 54inequality for ensuring a matrix to be a product of two M-matrices and applied it to  $P^2$  finite element method on uniform regular triangular meshes. In [18], Lorenz's 56 condition was applied to  $Q^2$  variational difference scheme on uniform meshes.

The main focus of this paper is to discuss Lorenz's condition for a fourth order scheme on nonuniform meshes and higher order accurate schemes. We discuss mesh constraints to preserve monotonicity of  $Q^2$  variational finite difference scheme on a nonuniform mesh. One can of course also discuss  $P^2$  variational difference scheme on a nonuniform regular triangular mesh, but there does not seem to be any advantage of using  $P^2$ . The scheme by  $Q^2$  is easier to implement, see Section 7 in [19].

For higher order scheme, it does not seem possible to apply Lorenz's condition 64 65 directly. Instead, we will demonstrate that Lorenz's condition can be applied to a few auxiliary matrices to establish the monotonicity in  $Q^3$  variational difference scheme. 66 To the best of our knowledge, this is the first time that monotonicity can be proven 67 for a fifth order accurate scheme in two dimensions. For one-dimensional Laplacian, 68 discrete maximum principle was proven for high order finite element methods [22]. 69 We are able to show the fifth order  $Q^3$  variational difference scheme in two dimen-70 71sions can be factored into a product of four M-matrices, whereas existing M-matrix factorizations for high order schemes involved products of two M-matrices. 72

The rest of the paper is organized as follows. In Section 2, we briefly review the 73 conventional monotone high order finite difference schemes. In Section 3, we review 74the fourth order  $P^2$  and  $Q^2$  variational finite difference schemes. In Section 4, we 75 review the Lorenz's condition for proving monotonicity and propose a relaxed version 76 of Lorenz's condition. In Section 5, we discuss the monotonicity of  $Q^2$  variational finite 77 difference scheme on a quasi-uniform mesh. In Section 6, we prove the monotonicity 78of  $Q^3$  variational finite difference scheme on a uniform mesh. Accuracy tests of these 79 schemes are given in Section 7. Section 8 are concluding remarks. 80

# 2. Classical finite difference schemes.

2.1. 9-point scheme. The 9-point scheme was somewhat suggested already in [12] and discussed in details in [10, 15]. It can be extended to higher dimensions [2, 4]. Consider solving the two-dimensional Poisson equations  $-u_{xx} - u_{yy} = f$  with homogeneous Dirichlet boundary conditions on a rectangular domain  $\Omega = (0, 1) \times$ (0, 1). Let  $u_{i,j}$  denote the numerical solutions at a uniform grid  $(x_i, y_j) = (\frac{i}{Nx}, \frac{j}{Ny})$ , and  $f_{i,j} = f(x_i, y_j)$ . For convenience, we introduce two matrices,

88 
$$U = \begin{pmatrix} u_{i-1,j+1} & u_{i,j+1} & u_{i+1,j+1} \\ u_{i-1,j} & u_{i,j} & u_{i+1,j} \\ u_{i-1,j-1} & u_{i,j-1} & u_{i+1,j-1} \end{pmatrix}, \quad F = \begin{pmatrix} f_{i-1,j+1} & f_{i,j+1} & f_{i+1,j+1} \\ f_{i-1,j} & f_{i,j} & f_{i+1,j} \\ f_{i-1,j-1} & f_{i,j-1} & f_{i+1,j-1} \end{pmatrix}.$$

Then the 9-point discrete Laplacian for the Poisson equation at a grid point  $(x_i, y_j)$ ocan be written as

(2.1)

91 
$$\frac{1}{12\Delta x^2} \begin{pmatrix} -1 & 2 & -1 \\ -10 & 20 & -10 \\ -1 & 2 & -1 \end{pmatrix} : U + \frac{1}{12\Delta y^2} \begin{pmatrix} -1 & -10 & -1 \\ 2 & 20 & 2 \\ -1 & -10 & -1 \end{pmatrix} : U = \frac{1}{12} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 8 & 1 \\ 0 & 1 & 0 \end{pmatrix} : F.$$

92 where : denotes the sum of all entry-wise products in two matrices of the same size. 93 Under the assumption  $\Delta x = \Delta y = h$ , it reduces to the following:

94 (2.2) 
$$\frac{1}{6h^2} \begin{pmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{pmatrix} : U = \frac{1}{12} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 8 & 1 \\ 0 & 1 & 0 \end{pmatrix} : F.$$

The 9-point scheme can also be regarded as a compact finite difference scheme [11]. There can exist a few or many different compact finite difference approximations of the same order [16]. For instance, with the fourth order compact finite difference approximation to Laplacian used in [17], we get the following scheme: (2.3)

99 
$$\frac{1}{12\Delta x^2} \begin{pmatrix} -1 & 2 & -1 \\ -10 & 20 & -10 \\ -1 & 2 & -1 \end{pmatrix} : U + \frac{1}{12\Delta y^2} \begin{pmatrix} -1 & -10 & -1 \\ 2 & 20 & 2 \\ -1 & -10 & -1 \end{pmatrix} : U = \frac{1}{144} \begin{pmatrix} 1 & 10 & 1 \\ 10 & 100 & 10 \\ 1 & 10 & 1 \end{pmatrix} : F.$$

Both schemes (2.1) and (2.3) are fourth order accurate and they have the same stencil and the same stiffness matrix in the left hand side. We have not observed any significant difference in numerical performances between these two schemes.

103 REMARK 1. For solving 2D Laplace equation  $-\Delta u = 0$  with Dirichlet boundary 104 conditions, the 9-point scheme becomes sixth order accurate [11].

105 Nonsingular M-matrices are inverse-positive matrices. There are many equivalent 106 definitions or characterizations of M-matrices, see [21]. The following is a convenient 107 sufficient but not necessary characterization of nonsingular M-matrices [18]:

108 THEOREM 2.1. For a real square matrix A with positive diagonal entries and non-109 positive off-diagonal entries, A is a nonsingular M-matrix if all the row sums of A 110 are non-negative and at least one row sum is positive.

111 By condition  $K_{35}$  in [21], a sufficient and necessary characterization is,

112 THEOREM 2.2. For a real square matrix A with positive diagonal entries and non-113 positive off-diagonal entries, A is a nonsingular M-matrix if and only if that there 114 exists a positive diagonal matrix D such that AD has all positive row sums.

REMARK 2. Non-negative row sum is not a necessary condition for M-matrices. For instance, the following matrix A is an M-matrix by Theorem 2.2:

$$A = \begin{bmatrix} 10 & 0 & 0 \\ -10 & 2 & -10 \\ 0 & 0 & 10 \end{bmatrix}, D = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}, AD = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 4 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

The stiffness matrix in the scheme (2.2) has diagonal entries  $\frac{20}{6h^2}$  and offdiagonal entries  $-\frac{1}{6h^2}$ ,  $-\frac{4}{6h^2}$  and 0, thus by Theorem 2.1 it is an M-matrix and the scheme is monotone. In order for the stiffness matrix in (2.1) and (2.3) to be an M-matrix, we need all the off-diagonal entries to be nonnegative, which is true under the mesh constraints  $\frac{1}{\sqrt{5}} \leq \frac{\Delta x}{\Delta y} \leq \sqrt{5}$ .

120 **2.2. The Bramble and Hubbard's scheme.** In [5], a fourth order accurate 121 monotone scheme was constructed. Consider solving a one-dimensional problem

122 (2.4) 
$$-u'' = f, \quad x \in [0,1], \quad u(0) = \sigma_0, u(1) = \sigma_1,$$

123 on a uniform grid  $x_i = \frac{i}{n+1}$   $(i = 0, 1, \dots, n+1)$ . The scheme can be written as

124 
$$\frac{-\sigma_0 + 2u_1 - u_2}{\Delta x^2} = f_1, \quad \frac{-u_{n-1} + 2u_n - \sigma_1}{\Delta x^2} = f_n$$

125

126 
$$\frac{\frac{1}{12}u_{i-2} - \frac{4}{3}u_{i-1} + \frac{5}{2}u_i - \frac{4}{3}u_{i+1} + \frac{1}{12}u_{i+2}}{\Delta x^2} = f_i, \quad i = 2, 3, \cdots, n-1.$$

127 The matrix vector form of the scheme is  $\frac{1}{\Delta x^2} H \mathbf{u} = \tilde{\mathbf{f}}$  where

$$128 \quad H = \begin{pmatrix} 2 & -1 & & & \\ -\frac{4}{3} & \frac{5}{2} & -\frac{4}{3} & \frac{1}{12} & & & \\ \frac{1}{12} & -\frac{4}{3} & \frac{5}{2} & -\frac{4}{3} & \frac{1}{12} & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & \frac{1}{12} & -\frac{4}{3} & \frac{5}{2} & -\frac{4}{3} & \frac{1}{12} \\ & & & \frac{1}{12} & -\frac{4}{3} & \frac{5}{2} & -\frac{4}{3} & \frac{1}{12} \\ & & & & & -1 & 2 \end{pmatrix}, \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{pmatrix}, \tilde{\mathbf{f}} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix} + \begin{pmatrix} \frac{\sigma_0}{\Delta x^2} \\ \frac{\sigma_0}{12\Delta x^2} \\ 0 \\ -\frac{\sigma_1}{12\Delta x^2} \\ \frac{\sigma_1}{\Delta x^2} \\ \frac{\sigma_1}{\Delta x^2} \end{pmatrix}.$$

For two-dimensional Laplacian, the scheme is defined similarly. In particular, assume  $\Delta x = \Delta y = h$  for a square domain, the stiffness matrix can be written as  $\frac{1}{h^2}(H \otimes I + I \otimes H)$  where I is the identity matrix and  $\otimes$  is the Kronecker product. Its monotonicity was proven in [5].

#### **3.** Variational finite difference schemes.

134 **3.1. Finite element method with the simplest quadrature.** Consider an 135 elliptic equation on  $\Omega = (0, 1) \times (0, 1)$  with Dirichlet boundary conditions:

136 (3.1) 
$$\mathcal{L}u \equiv -\nabla \cdot (a\nabla u) + cu = f \text{ on } \Omega, \quad u = g \text{ on } \partial\Omega.$$

Assume there is a function  $\bar{g} \in H^1(\Omega)$  as an extension of g so that  $\bar{g}|_{\partial\Omega} = g$ . The variational form of (3.1) is to find  $\tilde{u} = u - \bar{g} \in H^1_0(\Omega)$  satisfying

139 (3.2) 
$$\mathcal{A}(\tilde{u}, v) = (f, v) - \mathcal{A}(\bar{g}, v), \quad \forall v \in H_0^1(\Omega),$$

140 where  $\mathcal{A}(u, v) = \iint_{\Omega} a \nabla u \cdot \nabla v dx dy + \iint_{\Omega} cuv dx dy$ ,  $(f, v) = \iint_{\Omega} fv dx dy$ . 141 Let *h* be quadrature point spacing of a regular triangular mesh shown in Figure

141 Let *h* be quadrature point spacing of a regular triangular mesh shown in Figure 142 1 (or a rectangular mesh shown in Figure 2) and  $V_0^h \subseteq H_0^1(\Omega)$  be the continuous 143 finite element space consisting of piecewise  $P^2$  polynomials (or  $Q^2$  polynomials), then 144 the most convenient implementation of finite element method is to use the simple 145 quadrature consisting of vertices and edge centers with equal weights (or  $3 \times 3$  Gauss-146 Lobatto quadrature rule) for all the integrals, see Figure 1 for  $P^2$  method (or Figure 2 147 for  $Q^2$  method). Such a numerical scheme can be defined as: find  $u_h \in V_0^h$  satisfying

148 (3.3) 
$$\mathcal{A}_h(u_h, v_h) = \langle f, v_h \rangle_h - \mathcal{A}_h(g_I, v_h), \quad \forall v_h \in V_0^h,$$

where  $\mathcal{A}_h(u_h, v_h)$  and  $\langle f, v_h \rangle_h$  denote using simple quadrature for integrals  $\mathcal{A}(u_h, v_h)$ and  $(f, v_h)$  respectively, and  $g_I$  is the piecewise  $P^2$  (or  $Q^2$ ) Lagrangian interpolation



FIG. 1. An illustration of Lagrangian  $P^2$  element and the simple quadrature using vertices and edge centers.



FIG. 2. An illustration of Lagrangian  $Q^2$  element and the  $3 \times 3$  Gauss-Lobatto quadrature.

polynomial at the quadrature points shown in Figure 1 for  $P^2$  method (or Figure 2 for  $Q^2$  method) of the following function:

153 
$$g(x,y) = \begin{cases} 0, & \text{if } (x,y) \in (0,1) \times (0,1), \\ g(x,y), & \text{if } (x,y) \in \partial \Omega. \end{cases}$$

Then  $\bar{u}_h = u_h + g_I$  is the numerical solution for the problem (3.1). Notice that (3.3) is not a straightforward approximation to (3.2) since  $\bar{g}$  is never used. When the numerical solution is represented by a linear combination of Lagrangian interpolation polynomials at the grid points, it can be rewritten as a finite difference scheme. We also call it a variational difference scheme since it is derived from the variational form.

159 **3.2.** The  $P^2$  variational difference scheme derived. For Laplacian  $\mathcal{L}u = 160 -\Delta u$ , the scheme (3.3) on a uniform regular triangular mesh can be given as [24]:

161 (3.4a) 
$$\frac{1}{h^2} \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} : U = f_{i,j}, \quad \text{if } (x_i, y_j) \text{ is an edge center;}$$

162 (3.4b) 
$$\frac{1}{9h^2} \begin{pmatrix} 1 & -4 & 1\\ -4 & 12 & -4\\ 1 & -4 & 1 \end{pmatrix} : U = 0, \text{ if } (x_i, y_j) \text{ is a vertex.}$$

163 Notice that the stiffness matrix is not an M-matrix due to the positive off-diagonal 164 entries in (3.4b) and its inverse positivity was proven in [20].

165 Since the simple quadrature is exact for integrating only quadratic polynomials 166 on triangles, it is not obvious why the variational difference scheme (3.4) is fourth

order accurate. With such a quadrature on two adjacent triangles forming a rectangle
in a regular triangular mesh, we obtain a quadrature on the rectangle, see Figure 3.

For a reference square  $[-1,1] \times [-1,1]$ , the quadrature weights are  $\frac{2}{3}$  and  $\frac{4}{3}$  for an edge center and the cell center respectively.



FIG. 3. The simple quadrature on two triangles give a quadrature on a square.

171 LEMMA 3.1. The quadrature on a square  $[-1,1] \times [-1,1]$  using only four edge 172 centers with weight  $\frac{2}{3}$  and one cell center with weight  $\frac{4}{3}$  is exact for  $P^3$  polynomials. 173 Proof. Since the quadrature is exact for integrating  $P^2$  polynomials on either 174 triangle in Figure 3, it suffices to show that it is exact for integrating basis polynomials 175 of degree three, i.e.,  $x^2y$ ,  $xy^2$ ,  $x^3$  and  $y^3$ . It is straightforward to verify that both 176 exact integrals and quadrature of these four polynomials on the square are zero.  $\Box$ 

Therefore, with Bramble-Hilbert Lemma (see Exercise 3.1.1 and Theorem 4.1.3 in [9]), we can show that the quadrature rule is fourth order accurate if we regard the regular triangular mesh in Figure 3 (a) as a rectangular mesh.

The standard  $L^2(\Omega)$ -norm estimate for the finite element method with quadrature 180 (3.3) using Lagrangian  $P^2$  elements is third order accurate for smooth exact solutions 181 [9]. On the other hand, superconvergence of function values in finite element method 182without quadrature can be proven [6, 23], e.g., the errors at vertices and edge centers 183 are fourth order accurate on triangular meshes for function values if using  $P^2$  basis, 184see also [14]. It can be shown that using such fourth order accurate quadrature will 185 not affect the fourth order superconvergence even for a general variable coefficient 186 elliptic problem, see [19]. Notice that the scheme can also be given on a nonuniform 187 mesh and its fourth order accuracy still holds on a quasi uniform mesh since it is also 188 a finite element method. 189

190 **3.3.**  $Q^2$  variational difference scheme. The scheme (3.3) with Lagrangian  $Q^2$ 191 basis is fourth order accurate [19] and monotone on a uniform mesh under suitable 192 mesh constraints [18]. In the next section, we will discuss its monotonicity for the 193 Laplacian operator on quasi-uniform meshes.

Consider a uniform grid  $(x_i, y_j)$  for a rectangular domain  $[0, 1] \times [0, 1]$  where  $x_i = ih, i = 0, 1, ..., n + 1$  and  $y_j = jh, j = 0, 1, ..., n + 1, h = \frac{1}{n+1}$ , where n must be odd. Let  $u_{ij}$  denote the numerical solution at  $(x_i, y_j)$ . Let **u** denote an abstract vector consisting of  $u_{ij}$  for i, j = 1, 2, ..., n. Let  $\bar{\mathbf{u}}$  denote an abstract vector consisting of  $u_{ij}$  for i, j = 0, 1, 2, ..., n, n + 1. Let  $\bar{\mathbf{f}}$  denote an abstract vector consisting of  $f_{ij}$  for i, j = 1, 2, ..., n and the boundary condition g at the boundary grid points. Then the matrix vector representation of (3.3) is  $S\bar{\mathbf{u}} = M\mathbf{f}$  where S is the stiffness matrix and M is the lumped mass matrix. For convenience, after inverting the mass matrix, 202 with the boundary conditions, the whole scheme can be represented in a matrix vector

203 form 
$$\bar{L}_h \bar{\mathbf{u}} = \bar{\mathbf{f}}$$
. For Laplacian  $\mathcal{L}u = -\Delta u$ ,  $\bar{L}_h \bar{\mathbf{u}} = \bar{\mathbf{f}}$  on a uniform mesh is given as (3.5)

$$\begin{aligned} (\bar{L}_{h}\bar{\mathbf{u}})_{i,j} &\coloneqq \frac{-u_{i-1,j} - u_{i+1,j} + 4u_{i,j} - u_{i,j+1} - u_{i+1,j}}{h^{2}} = f_{i,j}, & \text{if } (x_{i}, y_{j}) \text{ is a cell center,} \\ (\bar{L}_{h}\bar{\mathbf{u}})_{i,j} &\coloneqq \frac{-u_{i-1,j} + 2u_{i,j} - u_{i+1,j}}{h^{2}} + \frac{u_{i,j-2} - 8u_{i,j-1} + 14u_{i,j} - 8u_{i,j+1} + u_{i,j+2}}{4h^{2}} = f_{i,j}, \\ & \text{if } (x_{i}, y_{j}) \text{ is an edge center for an edge parallel to the y-axis,} \\ (\bar{L}_{h}\bar{\mathbf{u}})_{i,j} &\coloneqq \frac{u_{i-2,j} - 8u_{i-1,j} + 14u_{i,j} - 8u_{i+1,j} + u_{i+2,j}}{4h^{2}} + \frac{-u_{i,j-1} + 2u_{i,j} - u_{i,j+1}}{h^{2}} = f_{i,j}, \\ & \text{if } (x_{i}, y_{j}) \text{ is an edge center for an edge parallel to the x-axis,} \\ (\bar{L}_{h}\bar{\mathbf{u}})_{i,j} &\coloneqq \frac{u_{i-2,j} - 8u_{i-1,j} + 14u_{i,j} - 8u_{i+1,j} + u_{i+2,j}}{4h^{2}} + \frac{u_{i,j-2} - 8u_{i,j-1} + 14u_{i,j} - 8u_{i,j+1} + u_{i,j+2}}{4h^{2}} = f_{i,j} \end{aligned}$$

$$(\bar{L}_h \bar{\mathbf{u}})_{i,j} := u_{i,j} = g_{i,j}$$
 if  $(x_i, y_j)$  is a boundary point.

205 If ignoring the denominator  $h^2$ , then the stencil can be represented as:

206 cell center 
$$-1$$
  $4$   $-1$  knots  $\frac{1}{4}$   $-2$   $7$   $-2$   $\frac{1}{4}$   
 $-1$   $-2$   $-2$   $\frac{1}{4}$ 

207

208 edge center (edge parallel to y-axis) 
$$\frac{1}{4}$$
 -2  $\frac{11}{2}$  -2  $\frac{1}{4}$   
209

210 edge center (edge parallel to x-axis) 
$$-1$$
  $\frac{1}{4}$   $-2$   
 $-2$   $-1$   $\frac{11}{2}$   $-1$   $-2$ 



FIG. 4. An illustration of a mesh for  $Q^3$  element and the  $4 \times 4$  Gauss-Lobatto quadrature.

3.4.  $Q^3$  variational difference scheme. In (3.3), if using Lagrangian  $Q^3$  basis with 4 × 4 Gauss-Lobatto quadrature, we get a fifth order accurate scheme [19]. The

FIG. 5. Three adjacent 1D cells for  $P^3$  elements using 4-point Gauss-Lobatto quadrature.

4-point Gauss-Lobatto quadrature for the reference interval [-1,1] has quadrature 213 points  $\left[-1 - \frac{\sqrt{5}}{5} \frac{\sqrt{5}}{5} 1\right]$ . Thus on an uniform rectangular mesh, the corresponding finite 214difference grid consisting of quadrature points is not exactly uniform, see Figure 4. 215

Now consider a uniform mesh for a one-dimensional problem and assume each 216 cell has length h, see Figure 5. There are two quadrature points inside each interval, 217 and we refer to them as the left interior point and the right interior point. The  $Q^3$ 218variational difference scheme for one-dimension problem (2.4) is given as  $\bar{L}_h \bar{\mathbf{u}} = \bar{\mathbf{f}}$ : 219(3.6)

$$(\bar{L}_h \bar{\mathbf{u}})_i := \frac{4}{h^2} \left[ 13u_i - \frac{15\sqrt{5} + 25}{8} (u_{i-1} + u_{i+1}) + \frac{15\sqrt{5} - 25}{8} (u_{i-2} + u_{i+2}) - \frac{1}{4} (u_{i-3} + u_{i+3}) \right] = f_i, x_i \text{ is a knot};$$
  
$$(\bar{L}_h \bar{\mathbf{u}})_i := \frac{4}{h^2} \left[ -\frac{3\sqrt{5} + 5}{4} u_{i-1} + 5u_i + \frac{-5}{2} u_{i+1} + \frac{15\sqrt{5} - 25}{8} u_{i+2} \right] = f_i, \quad x_i \text{ is the left interior point};$$

220

$$(\bar{L}_h \bar{\mathbf{u}})_i := \frac{4}{h^2} \left[ \frac{15\sqrt{5} - 25}{8} u_{i-2} - \frac{5}{2} u_{i-1} + 5u_i - \frac{3\sqrt{5} + 5}{4} u_{i+1} \right] = f_i, \quad \text{if } x_i \text{ is the right interior point.}$$
$$(\bar{L}_h \bar{\mathbf{u}})_0 := u_0 = \sigma_0, \qquad (\bar{L}_h \bar{\mathbf{u}})_{n+1} := u_{n+1} = \sigma_1.$$

The explicit scheme in two dimensions will be given in Section 6. 221

#### 4. Lorenz's condition for monotonicity. 2.2.2

4.1. Discrete maximum principle. For a finite difference scheme, assume there are N grid points in the domain  $\Omega$  and  $N^{\partial}$  boundary grid points on  $\partial \Omega$ . Define

$$\mathbf{u} = \begin{pmatrix} u_1 & \cdots & u_N \end{pmatrix}^T, \mathbf{u}^{\partial} = \begin{pmatrix} u_1^{\partial} & \cdots & u_{N^{\partial}}^{\partial} \end{pmatrix}^T, \tilde{\mathbf{u}} = \begin{pmatrix} u_1 & \cdots & u_N & u_1^{\partial} & \cdots & u_{N^{\partial}}^{\partial} \end{pmatrix}^T.$$

A finite difference scheme can be written as 223

224 
$$\mathcal{L}_{h}(\tilde{\mathbf{u}})_{i} = \sum_{j=1}^{N} b_{ij} u_{j} + \sum_{j=1}^{N^{\partial}} b_{ij}^{\partial} u_{j}^{\partial} = f_{i}, \quad 1 \le i \le N,$$
225 
$$u_{i}^{\partial} = g_{i}, \quad 1 \le i \le N^{\partial}.$$

225

232

The matrix form is 227

228 
$$\tilde{L}_h \tilde{\mathbf{u}} = \tilde{\mathbf{f}}, \tilde{L}_h = \begin{pmatrix} L_h & B^{\partial} \\ 0 & I \end{pmatrix}, \tilde{\mathbf{u}} = \begin{pmatrix} \mathbf{u} \\ \mathbf{u}^{\partial} \end{pmatrix}, \tilde{\mathbf{f}} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix}.$$

229 The discrete maximum principle is

230 (4.1) 
$$\mathcal{L}_h(\tilde{\mathbf{u}})_i \le 0, 1 \le i \le N \Longrightarrow \max_i u_i \le \max\{0, \max_i u_i^\partial\}$$

which implies 231

$$\mathcal{L}_h(\tilde{\mathbf{u}})_i = 0, 1 \le i \le N \Longrightarrow |u_i| \le \max_i |u_i^{\partial}|$$

The following result was proven in [8]: 233

THEOREM 4.1. A finite difference operator  $\mathcal{L}_h$  satisfies the discrete maximum 234principle (4.1) if  $\tilde{L}_h^{-1} \geq 0$  and all row sums of  $\tilde{L}_h$  are non-negative. 235

With the same  $\bar{L}_h$  as defined in the previous section, it suffices to have  $\bar{L}_h^{-1} \ge 0$ , see [18]:

THEOREM 4.2. If  $\bar{L}_h^{-1} \ge 0$ , then  $\tilde{L}_h^{-1} \ge 0$  thus  $L_h^{-1} \ge 0$ . Moreover, if row sums of  $\bar{L}_h$  are non-negative, then the finite difference operator  $\mathcal{L}_h$  satisfies the discrete maximum principle.

Let **1** be an abstract vector of the same shape as  $\bar{\mathbf{u}}$  with all ones. For the  $Q^2$ or  $Q^3$  variational difference scheme, we have that  $(\bar{L}_h \mathbf{1})_{i,j} = 1$  if  $(x_i, y_j) \in \partial\Omega$  and  $(\bar{L}_h \mathbf{1})_{i,j} = 0$  if  $(x_i, y_j) \in \Omega$ , which implies the row sums of  $\bar{L}_h$  are non-negative. Thus from now on, we only need to discuss the monotonicity of the matrix  $\bar{L}_h$ .

#### 4.2. Lorenz's sufficient condition for monotonicity.

246 DEFINITION 1. Let  $\mathcal{N} = \{1, 2, ..., n\}$ . For  $\mathcal{N}_1, \mathcal{N}_2 \subset \mathcal{N}$ , we say a matrix A of 247 size  $n \times n$  connects  $\mathcal{N}_1$  with  $\mathcal{N}_2$  if

248 (4.2) 
$$\forall i_0 \in \mathcal{N}_1, \exists i_r \in \mathcal{N}_2, \exists i_1, \dots, i_{r-1} \in \mathcal{N} \quad s.t. \quad a_{i_{k-1}i_k} \neq 0, \quad k = 1, \cdots, r.$$

If perceiving A as a directed graph adjacency matrix of vertices labeled by  $\mathcal{N}$ , then (4.2) simply means that there exists a directed path from any vertex in  $\mathcal{N}_1$  to at least one vertex in  $\mathcal{N}_2$ . In particular, if  $\mathcal{N}_1 = \emptyset$ , then any matrix A connects  $\mathcal{N}_1$  with  $\mathcal{N}_2$ .

Given a square matrix A and a column vector  $\mathbf{x}$ , we define

253 
$$\mathcal{N}^0(A\mathbf{x}) = \{i : (A\mathbf{x})_i = 0\}, \quad \mathcal{N}^+(A\mathbf{x}) = \{i : (A\mathbf{x})_i > 0\}.$$

Given a matrix  $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ , define its diagonal, off-diagonal, positive and negative off-diagonal parts as  $n \times n$  matrices  $A_d$ ,  $A_a$ ,  $A_a^+$ ,  $A_a^-$ :

256 
$$(A_d)_{ij} = \begin{cases} a_{ii}, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}, \quad A_a = A - A_d,$$

257

258 
$$(A_a^+)_{ij} = \begin{cases} a_{ij}, & \text{if } a_{ij} > 0, \quad i \neq j \\ 0, & \text{otherwise.} \end{cases}, \quad A_a^- = A_a - A_a^+.$$

259 The following two results were proven in [20]. See also [18] for a detailed proof.

THEOREM 4.3. If  $A \leq M_1 M_2 \cdots M_k L$  where  $M_1, \cdots, M_k$  are nonsingular Mmatrices and  $L_a \leq 0$ , and there exists a nonzero vector  $\mathbf{e} \geq 0$  such that one of the matrices  $M_1, \cdots, M_k, L$  connects  $\mathcal{N}^0(A\mathbf{e})$  with  $\mathcal{N}^+(A\mathbf{e})$ . Then  $M_k^{-1}M_{k-1}^{-1}\cdots M_1^{-1}A$ is an M-matrix, thus A is a product of k + 1 nonsingular M-matrices and  $A^{-1} \geq 0$ .

THEOREM 4.4 (Lorenz's condition). If  $A_a^-$  has a decomposition:  $A_a^- = A^z + A^s = 265$   $(a_{ij}^z) + (a_{ij}^s)$  with  $A^s \leq 0$  and  $A^z \leq 0$ , such that

(4.3a)

266  $A_d + A^z$  is a nonsingular M-matrix, (4.3b)

267 
$$A_a^+ \le A^z A_d^{-1} A^s \text{ or equivalently } \forall a_{ij} > 0 \text{ with } i \ne j, a_{ij} \le \sum_{k=1}^n a_{ik}^z a_{kk}^{-1} a_{kj}^s,$$

$$(4.3c)$$

- $\exists \mathbf{e} \in \mathbb{R}^n \setminus \{\mathbf{0}\}, \mathbf{e} \ge 0 \text{ with } A\mathbf{e} \ge 0 \text{ s.t. } A^z \text{ or } A^s \text{ connects } \mathcal{N}^0(A\mathbf{e}) \text{ with } \mathcal{N}^+(A\mathbf{e}).$
- 270 Then A is a product of two nonsingular M-matrices thus  $A^{-1} \ge 0$ .

# 271 COROLLARY 4.5. The matrix L in Theorem 4.3 must be an M-matrix.

272 Proof. Let  $M^{-1} = M_k^{-1} M_{k-1}^{-1} \dots M_1^{-1}$ , following the proof of Theorem 7 in [18], 273 then  $M^{-1}A\mathbf{e} \ge cA\mathbf{e}$  for some positive number c. Then  $A\mathbf{e} \ge 0 \Rightarrow M^{-1}A\mathbf{e} \ge 0$ . Now 274 since  $\mathbf{e} \ge 0$ ,  $M^{-1}A \le L \Rightarrow 0 \le (L - M^{-1}A)\mathbf{e} \Rightarrow M^{-1}A\mathbf{e} \le L\mathbf{e}$  thus  $L\mathbf{e} \ge 0$ .

Assume *L* connects  $\mathcal{N}^0(A\mathbf{e})$  with  $\mathcal{N}^+(A\mathbf{e})$ . Since  $M^{-1}A\mathbf{e} \leq L\mathbf{e}$ ,  $\mathcal{N}^0(L\mathbf{e}) \subseteq$ 276  $\mathcal{N}^0(A\mathbf{e})$  and  $\mathcal{N}^+(A\mathbf{e}) \subseteq \mathcal{N}^+(L\mathbf{e})$ , so *L* also connects  $\mathcal{N}^0(L\mathbf{e})$  with  $\mathcal{N}^+(L\mathbf{e})$ .

Assume  $M_i$  connects  $\mathcal{N}^0(A\mathbf{e})$  with  $\mathcal{N}^+(A\mathbf{e})$ , following the proof of Theorem 7 in [18], we have  $M^{-1}A\mathbf{e} > 0$ . Now L trivially connects  $\mathcal{N}^0(L\mathbf{e})$  with  $\mathcal{N}^+(L\mathbf{e})$  since  $L\mathbf{e} \ge M^{-1}A\mathbf{e} \Rightarrow L\mathbf{e} > 0$  and  $\mathcal{N}^0(L\mathbf{e}) = \emptyset$ .

280 Then Theorem 6 in [18] applies to show L is an M-matrix.

In practice, the condition (4.3c) can be difficult to verify. For variational difference schemes, the vector **e** can be taken as **1** consisting of all ones, then the condition (4.3c)can be simplified. The following theorem was proven in [18].

THEOREM 4.6. Let A denote the matrix representation of the variational difference scheme (3.3) with  $Q^2$  basis solving  $-\nabla \cdot (a\nabla)u + cu = f$ . Assume  $A_a^-$  has a decomposition  $A_a^- = A^z + A^s$  with  $A^s \leq 0$  and  $A^z \leq 0$ . Then  $A^{-1} \geq 0$  if the following are satisfied:

288 1.  $(A_d + A^z)\mathbf{1} \neq \mathbf{0} \text{ and } (A_d + A^z)\mathbf{1} \geq 0;$ 

289 2.  $A_a^+ \leq A^z A_d^{-1} A^s;$ 

290 3. For  $c(x,y) \ge 0$ , either  $A^z$  or  $A^s$  has the same sparsity pattern as  $A_a^-$ . If 291 c(x,y) > 0, then this condition can be removed.

4.3. A relaxed Lorenz's condition. In practice, both (4.3a) and (4.3b) impose mesh constraints for the  $Q^2$  variational difference scheme on non-uniform meshes. The condition (4.3a) can be relaxed as the following:

THEOREM 4.7 (A relaxed Lorenz's condition). If  $A_a^-$  has a decomposition:  $A_a^- = A^z + A^s = (a_{ij}^z) + (a_{ij}^s)$  with  $A^s \leq 0$  and  $A^z \leq 0$ , and there exists a diagonal matrix  $A_{d^*} \geq A_d$  such that

(4.4a)

298  $A_d^* + A^z$  is a nonsingular M-matrix,

(4.4b)

299  $A_a^+ \leq A^z A_{d^*}^{-1} A^s,$ 

(4.4c)

 $\exists \mathbf{e} \in \mathbb{R}^n \setminus \{\mathbf{0}\}, \mathbf{e} \ge 0 \text{ with } A\mathbf{e} \ge 0 \text{ s.t. } A^z \text{ or } A^s \text{ connects } \mathcal{N}^0(A\mathbf{e}) \text{ with } \mathcal{N}^+(A\mathbf{e}).$ 

302 Then A is a product of two nonsingular M-matrices thus  $A^{-1} \ge 0$ .

303 Proof. It is straightforward that  $A = A_d + A_a^+ + A^z + A^s \leq A_{d^*} + A^z + A^s + A^{3/2} + A^{3/2}A_{d^*}^{-1}A^s = (A_{d^*} + A^z)(I + A_{d^*}^{-1}A^s)$ . By (4.4c), either  $A_{d^*} + A^z$  or  $I + A_{d^*}^{-1}A^s$  connects 305  $\mathcal{N}^0(A\mathbf{e})$  with  $\mathcal{N}^+(A\mathbf{e})$ . By applying Theorem 4.3 for the case  $k = 1, M_1 = A_{d^*} + A^z$ 306 and  $L = I + A_{d^*}^{-1}A^s$ , we get  $A^{-1} \geq 0$ .

307 REMARK 3. Since  $A_d \leq A_{d^*}$ , only (4.4a) is more relaxed than (4.3a), and (4.4b) 308 is more stringent than (4.3b). However, we will show in next section that it is possible 309 to construct  $A_{d^*}$  such that (4.3b) and (4.4b) impose identical mesh constraints.

310 With Theorem 2.1, combining Theorem 4.7 and Theorem 4.6, we have:

THEOREM 4.8. Let A denote the matrix representation of the variational difference scheme (3.3) with  $Q^2$  basis solving  $-\nabla \cdot (a\nabla)u + cu = f$ . Assume  $A_a^-$  has a

decomposition  $A_a^- = A^z + A^s$  with  $A^s \leq 0$  and  $A^z \leq 0$  and there exists a diagonal matrix  $A_{d^*} \geq A_d$ . Then  $A^{-1} \geq 0$  if the following are satisfied: 313 314

- 1.  $(A_{d^*} + A^z)\mathbf{1} \neq \mathbf{0}$  and  $(A_{d^*} + A^z)\mathbf{1} \geq 0$ ; 315
- 2.  $A_a^+ \leq A^z A_{d^*}^{-1} A^s;$ 316

3. For  $c(x,y) \geq 0$ , either  $A^z$  or  $A^s$  has the same sparsity pattern as  $A_a^-$ . If 317 c(x, y) > 0, then this condition can be removed. 318

5. Monotonicity of  $Q^2$  variational difference scheme on quasi-uniform 319 meshes. The discussion in this section can be easily extended to more general cases such as  $\mathcal{L}u = -\Delta u + cu$  and Neumann boundary conditions. For simplicity, we only 321 discuss the Laplacian case  $\mathcal{L}u = -\Delta u$  and Dirichlet boundary conditions. 322

323 Consider a grid  $(x_i, y_j)$  (i, j = 0, 1, ..., n+1) for a rectangular domain  $[0, 1] \times [0, 1]$ where n must be odd and i, j = 0, n+1 correspond to boundary points. Let  $u_{ij}$  denote 324 the numerical solution at  $(x_i, y_j)$ . Let  $\bar{\mathbf{u}}$  denote an abstract vector consisting of  $u_{ij}$ 325 for  $i, j = 0, 1, 2, \dots, n, n + 1$ . Let **f** denote an abstract vector consisting of  $f_{ij}$  for 326  $i, j = 1, 2, \cdots, n$  and the boundary condition g at the boundary grid points. Then 327 the matrix vector representation of (3.3) with  $Q^2$  basis is  $\bar{L}_h \bar{\mathbf{u}} = \bar{\mathbf{f}}$ . 328

The focus of this section is to show  $L_h^{-1} \ge 0$  under suitable mesh constraints for 329 quasi-uniform meshes. Moreover, it is straightforward to verify that  $(\bar{L}_h \mathbf{1})_{i,j} = 0$  for 330 interior points  $(x_i, y_j)$  and  $(\bar{L}_h \mathbf{1})_{i,j} = 1$  for boundary points  $(x_i, y_j)$ . Thus by Section 331 4.1, the scheme also satisfies the discrete maximum principle. 332

For simplicity, in the rest of this section we use A to denote the matrix  $L_h$  and let 333 334  $\mathcal{A}$  be the linear operator corresponding to the matrix A. For convenience, we can also regard the abstract vector  $\bar{\mathbf{u}}$  as a matrix of size  $(n+2) \times (n+2)$ . Then by our notation, 335 the mapping  $\mathcal{A}: \mathbb{R}^{(n+2)\times(n+2)} \to \mathbb{R}^{(n+2)\times(n+2)}$  is given as  $\mathcal{A}(\bar{\mathbf{u}})_{i,j} := (\bar{L}_h \bar{\mathbf{u}})_{i,j}$ . 336



four adjacent  $Q^2$  elements.

FIG. 6. A non-uniform mesh for  $Q^2$  variational difference scheme. Each edge in a cell has length 2h.

5.1. The scheme in two dimensions. For boundary points  $(x_i, y_j) \in \partial \Omega$ , the 337 scheme is  $\mathcal{A}(\bar{\mathbf{u}})_{i,j} := u_{i,j} = g_{i,j}$ . The scheme for interior grid points  $(x_i, y_j) \in \Omega$  on 338 a non-uniform mesh can be given on four distinct types of points shown in Figure 6 339 (b). For simplicity, from now on, we will use *edge center* (2) to denote an interior 340 edge center for an edge parallel to the y-axis, and edge center (3) to denote an interior 341edge center for an edge parallel to the x-axis. The scheme at an interior grid point is 342

given as  $\mathcal{A}(\bar{\mathbf{u}})_{i,j} = f_{i,j}$  with 343

(5.1)  
344 
$$\mathcal{A}(\bar{\mathbf{u}})_{i,j} := \frac{2h_a^2 + 2h_b^2}{h_a^2 h_b^2} u_{i,j} - \left(\frac{1}{h_a^2} u_{i+1,j} + \frac{1}{h_a^2} u_{i-1,j} + \frac{1}{h_b^2} u_{i,j+1} + \frac{1}{h_b^2} u_{i,j-1}\right)$$

345 if  $(x_i, y_j)$  is a cell center;

346 
$$\mathcal{A}(\bar{\mathbf{u}})_{i,j} := \frac{7h_b^2 + 4h_ah_{a-1}}{2h_ah_{a-1}h_b^2}u_{i,j} - \frac{4}{h_a(h_a + h_{a-1})}u_{i+1,j} - \frac{4}{h_{a-1}(h_a + h_{a-1})}u_{i-1,j}$$
347 
$$-\frac{1}{12}u_{i,j+1} - \frac{1}{12}u_{i,j-1} + \frac{1}{212}u_{i+1,j} - \frac{1}{12}u_{i+2,j} + \frac{1}{212}u_{i+2,j} + \frac{1}{212}u_{i+2,j}$$

47 
$$-\frac{1}{h_b^2}u_{i,j+1} - \frac{1}{h_b^2}u_{i,j-1} + \frac{1}{2h_a(h_a + h_{a-1})}u_{i+2,j} + \frac{1}{2h_{a-1}(h_a + h_{a-1})}u_{i-2,j},$$

348 if 
$$(x_i, y_j)$$
 is edge center (2);

349 
$$\mathcal{A}(\bar{\mathbf{u}})_{i,j} := \frac{7h_a^2 + 4h_bh_{b-1}}{2h_bh_{b-1}h_a^2} u_{i,j} - \frac{4}{h_b(h_b + h_{b-1})} u_{i,j+1} - \frac{4}{h_{b-1}(h_b + h_{b-1})} u_{i,j-1}$$

350 
$$-\frac{1}{h_a^2}u_{i+1,j} - \frac{1}{h_a^2}u_{i-1,j} + \frac{1}{2h_b(h_b + h_{b-1})}u_{i,j+2} + \frac{1}{2h_{b-1}(h_b + h_{b-1})}u_{i,j-2}$$

351 if 
$$(x_i, y_j)$$
 is edge center (3);

352 
$$\mathcal{A}(\bar{\mathbf{u}})_{i,j} := \frac{7h_a h_{a-1} + 7h_b h_{b-1}}{2h_a h_{a-1} h_b h_{b-1}} u_{i,j} - \left[\frac{4}{h_a (h_a + h_{a-1})} u_{i+1,j} + \frac{4}{h_{a-1} (h_a + h_{a-1})} u_{i-1,j} + \frac{4}{h_{a-1} (h_{a-1} (h_{a-1})} u_{i-1,j} +$$

353 
$$+ \frac{4}{h_b(h_b + h_{b-1})} u_{i,j+1} + \frac{4}{h_{b-1}(h_b + h_{b-1})} u_{i,j-1} \right] + \frac{1}{2h_a(h_a + h_{a-1})} u_{i+2,j}$$
354 
$$+ \frac{1}{2h_{a-1}(h_a + h_{a-1})} u_{i-2,j} + \frac{1}{2h_b(h_b + h_{b-1})} u_{i,j+2} + \frac{1}{2h_{b-1}(h_b + h_{b-1})} u_{i,j-2},$$

354 
$$+ \frac{1}{2h_{a-1}(h_a + h_{a-1})} u_{i-2,j} + \frac{1}{2h_b(h_b + h_{b-1})} u_{i,j+2} + \frac{1}{2h_b} u_{i,j+2} + \frac{1$$

interior knot.

356 if 
$$(x_i, y_j)$$
 is an

For a uniform mesh 
$$h_a = h_{a-1} = h_b = h_{b-1} = h$$
, the scheme reduces to (3.5).

5.2. The Decomposition of  $A_a^-$ . Next, by the same notations defined in Sec-358tion 4.2, we will decompose the matrix  $A = A_d + A_a^- + A_a^+$  and  $A_a^- = A^z + A^s$  to verify Theorem 4.6. We will use  $\mathcal{A}_a^-$ ,  $\mathcal{A}_a^+$ ,  $\mathcal{A}^z$  and  $\mathcal{A}^s$  to denote linear operators for 359360 corresponding matrices. First, for the diagonal part we have 361

362 
$$\mathcal{A}_d(\bar{\mathbf{u}})_{i,j} = u_{i,j}, \text{ if } (x_i, y_j) \text{ is a boundary point;}$$

363 
$$\mathcal{A}_d(\bar{\mathbf{u}})_{i,j} = \frac{2h_a^2 + 2h_b^2}{h_a^2 h_b^2} u_{i,j}, \quad \text{if } (x_i, y_j) \text{ is a cell center;}$$

364 
$$\mathcal{A}_d(\bar{\mathbf{u}})_{i,j} = \frac{7}{4}$$

$$\mathcal{A}_{d}(\bar{\mathbf{u}})_{i,j} = \frac{7h_{b}^{2} + 4h_{a}h_{a-1}}{2h_{a}h_{a-1}h_{b}^{2}}u_{i,j}, \quad \text{if } (x_{i}, y_{j}) \text{ is edge center } (2);$$

365 
$$\mathcal{A}_d(\bar{\mathbf{u}})_{i,j} = \frac{7h_a^2 + 4h_b h_{b-1}}{2h_b h_{b-1} h_a^2} u_{i,j}, \quad \text{if } (x_i, y_j) \text{ is edge center } (3);$$

$$\mathcal{A}_{d}(\bar{\mathbf{u}})_{i,j} = \frac{7h_bh_{b-1} + 7h_ah_{a-1}}{2h_ah_{a-1}h_bh_{b-1}}u_{i,j}, \quad \text{if } (x_i, y_j) \text{ is an interior knot.}$$

Notice that for a boundary point  $(x_i, y_j) \in \partial \Omega$  we have  $\mathcal{A}(\bar{\mathbf{u}})_{i,j} = \mathcal{A}_d(\bar{\mathbf{u}})_{i,j} = u_{i,j}$ , thus 368369 for off-diagonal parts, we only need to look at the interior grid points. For positive

#### 370 off-diagonal entries, we have

371  $\mathcal{A}_{a}^{+}(\bar{\mathbf{u}})_{i,j} = 0$ , if  $(x_{i}, y_{j})$  is a cell center; 372  $\mathcal{A}_{a}^{+}(\bar{\mathbf{u}})_{i,j} = \frac{1}{2h_{a}(h_{a} + h_{a-1})}u_{i+2,j} + \frac{1}{2h_{a-1}(h_{a} + h_{a-1})}u_{i-2,j}$ , edge center (2);

373 
$$\mathcal{A}_{a}^{+}(\bar{\mathbf{u}})_{i,j} = \frac{1}{2h_{b}(h_{b}+h_{b-1})}u_{i,j+2} + \frac{1}{2h_{b-1}(h_{b}+h_{b-1})}u_{i,j-2}, \text{ edge center (3);}$$

374 
$$\mathcal{A}_{a}^{+}(\bar{\mathbf{u}})_{i,j} = \frac{1}{2h_{a}(h_{a} + h_{a-1})} u_{i+2,j} + \frac{1}{2h_{a-1}(h_{a} + h_{a-1})} u_{i-2,j} + \frac{1}{2h_{b}(h_{b} + h_{b-1})} u_{i,j+2}$$
375 
$$+ \frac{1}{2h_{a}(h_{a} + h_{a-1})} u_{i,j-2}, \quad \text{if } (x_{i}, y_{j}) \text{ is an interior knot.}$$

 $+ \frac{1}{2h_{b-1}(h_b + h_{b-1})} u_{i,j-2}, \quad \text{if } (x_i, y_j) \text{ is an interior knot.}$ 

Then we perform a decomposition  $A_a^- = A^z + A^s$ , which depends on two constants 0 <  $\epsilon_1 \le 1$  and 0 <  $\epsilon_2 \le 1$ .

379 
$$\mathcal{A}^{z}(\bar{\mathbf{u}})_{i,j} = -\epsilon_{1} \left( \frac{1}{h_{a}^{2}} u_{i+1,j} + \frac{1}{h_{a}^{2}} u_{i-1,j} + \frac{1}{h_{b}^{2}} u_{i,j+1} + \frac{1}{h_{b}^{2}} u_{i,j-1} \right), \quad \text{if } (x_{i}, y_{j}) \text{ is a cell center;}$$
  
380  $\mathcal{A}^{z}(\bar{\mathbf{u}})_{i,j} = -\epsilon_{1} \left( \frac{1}{k} u_{i,j+1} + \frac{1}{k} u_{i,j-1} \right) - \epsilon_{2} \left[ \frac{4}{k} u_{i,j+1,j} + \frac{4}{k} u_{i,j+1,j} \right]$ 

380 
$$\mathcal{A}(\mathbf{u})_{i,j} = -\epsilon_1 \left( \frac{1}{h_b^2} u_{i,j+1} + \frac{1}{h_b^2} u_{i,j-1} \right) - \epsilon_2 \left[ \frac{1}{h_a(h_a + h_{a-1})} u_{i+1,j} + \frac{1}{h_{a-1}(h_a + h_{a-1})} u_{i-1,j} \right]$$
  
381 if  $(x_i, y_j)$  is edge center (2);

382 
$$\mathcal{A}^{z}(\bar{\mathbf{u}})_{i,j} = -\epsilon_{1} \left( \frac{1}{h_{a}^{2}} u_{i+1,j} + \frac{1}{h_{a}^{2}} u_{i-1,j} \right) - \epsilon_{2} \left[ \frac{4}{h_{b}(h_{b} + h_{b-1})} u_{i,j+1} + \frac{4}{h_{b-1}(h_{b} + h_{b-1})} u_{i,j-1} \right],$$
383 if  $(x, u_{b})$  is edge center (3):

if  $(x_i, y_j)$  is edge center (3);

384 
$$\mathcal{A}^{z}(\bar{\mathbf{u}})_{i,j} = -\epsilon_{2} \left[ \frac{4}{h_{a}(h_{a}+h_{a-1})} u_{i+1,j} + \frac{4}{h_{a-1}(h_{a}+h_{a-1})} u_{i-1,j} \right]_{i=1,j}$$

$$+\frac{4}{h_b(h_b+h_{b-1})}u_{i,j+1} + \frac{4}{h_{b-1}(h_b+h_{b-1})}u_{i,j-1} \right], \quad \text{if } (x_i, y_j) \text{ is an interior knot.}$$

Notice that  $A^z$  defined above has exactly the same sparsity pattern as  $A_a^-$  for  $0 < \epsilon_1 \le 1$  and  $0 < \epsilon_2 \le 1$ . Let  $A^s = A_a^- - A^z$  then  $A^s \le 0$ .

5.3. Mesh constraints for  $A^{z}A_{d}^{-1}A^{s} \ge A_{a}^{+}$ . In order to verify  $A^{z}A_{d}^{-1}A^{s} \ge A_{a}^{+}$ , we only need to discuss nonzero entries in the output of  $\mathcal{A}_{a}^{+}(\bar{\mathbf{u}})$  since  $A^{z}A_{d}^{-1}A^{s} \ge 0$ . First consider the case that  $(x_{i}, y_{j})$  is an interior knot. Figure 7 (a) shows the positive coefficients in the output of  $\mathcal{A}_{a}^{+}(\bar{\mathbf{u}})_{ij}$  at a knot  $(x_{i}, y_{j})$ . Figure 7 (b) shows the stencil of  $\mathcal{A}^{z}(\bar{\mathbf{u}})_{ij}$ . Thus  $\mathcal{A}^{z}(\bar{\mathbf{u}})$  acting as an operator on  $[\mathcal{A}_{d}^{-1}\mathcal{A}^{s}](\bar{\mathbf{u}})$  at a knot is:

$$[\mathcal{A}^{z}\mathcal{A}_{d}^{-1}\mathcal{A}^{s}](\bar{\mathbf{u}})_{i,j} = -4\epsilon_{2}\left[\frac{1}{h_{a}(h_{a-1}+h_{a})}[\mathcal{A}_{d}^{-1}\mathcal{A}^{s}](\bar{\mathbf{u}})_{i+1,j} + \frac{1}{h_{a-1}(h_{a-1}+h_{a})}[\mathcal{A}_{d}^{-1}\mathcal{A}^{s}](\bar{\mathbf{u}})_{i-1,j} + \frac{1}{h_{b}(h_{b-1}+h_{b})}[\mathcal{A}_{d}^{-1}\mathcal{A}^{s}](\bar{\mathbf{u}})_{i,j+1} + \frac{1}{h_{b-1}(h_{b-1}+h_{b})}[\mathcal{A}_{d}^{-1}\mathcal{A}^{s}](\bar{\mathbf{u}})_{i,j-1}\right].$$

In the expression above, the output of the operator  $\mathcal{A}^{z}(\bar{\mathbf{u}})_{ij}$  are at interior edge centers as shown in Figure 7 (b). Hence  $[\mathcal{A}_{d}^{-1}\mathcal{A}^{s}]$  will act on these edge centers with the mesh lengths corresponding to Figure 6. Carefully considering the mesh lengths



FIG. 7. Stencil of operators at an interior knot  $(x_i, y_j)$ . The four red dots are the locations/entries where  $\mathcal{A}_{a}^{+}(\bar{\mathbf{u}})_{i,j}$  are nonzero. Gray nodes in (c) represent positive entries that can be discarded for the purposes of verifying (4.4b). The mesh is illustrated as a uniform one only for simplicity.

and operations of  $\mathcal{A}_d^{-1}$  at these points gives: 398

$$\begin{aligned} \left[\mathcal{A}^{z}\mathcal{A}_{d}^{-1}\mathcal{A}^{s}\right](\bar{\mathbf{u}})_{i,j} &= -4\epsilon_{2} \left[\frac{1}{h_{a}(h_{a-1}+h_{a})}\frac{2h_{b}h_{b-1}h_{a}^{2}}{7h_{a}^{2}+4h_{b}h_{b-1}}\mathcal{A}^{s}(\bar{\mathbf{u}})_{i+1,j} \right. \\ \left. +\frac{1}{h_{a-1}(h_{a-1}+h_{a})}\frac{2h_{b}h_{b-1}h_{a-1}^{2}}{7h_{a-1}^{2}+4h_{b}h_{b-1}}\mathcal{A}^{s}(\bar{\mathbf{u}})_{i-1,j} + \frac{1}{h_{b}(h_{b-1}+h_{b})}\frac{2h_{a}h_{a-1}h_{b}^{2}}{7h_{b}^{2}+4h_{a}h_{a-1}}\mathcal{A}^{s}(\bar{\mathbf{u}})_{i,j+1} \right. \\ \left. +\frac{1}{h_{b-1}(h_{b-1}+h_{b})}\frac{2h_{a}h_{a-1}h_{b-1}^{2}}{7h_{b-1}^{2}+4h_{a}h_{a-1}}\mathcal{A}^{s}(\bar{\mathbf{u}})_{i,j-1} \right], \quad \text{if } (x_{i},y_{j}) \text{ is an interior knot.} \end{aligned}$$

Next consider the effect of  $\mathcal{A}^{s}(\bar{\mathbf{u}})$  operator which has the same sparsity pattern as 400 A<sup>z</sup>( $\mathbf{\bar{u}}$ ). Figure 7 (c) shows the stencil of  $[\mathcal{A}^{z}\mathcal{A}_{d}^{-1}\mathcal{A}^{s}](\mathbf{\bar{u}})_{i,j}$  for an interior knot. Recall that  $A^{z} \leq 0$ ,  $A^{s} \leq 0$ , and  $A_{d}^{-1} \geq 0$ , thus we have  $A^{z}A_{d}^{-1}A^{s} \geq 0$ . So we only need to compare the outputs of  $[\mathcal{A}^{z}\mathcal{A}_{d}^{-1}\mathcal{A}^{s}](\mathbf{\bar{u}})_{i,j}$  and  $\mathcal{A}_{a}^{+}(\mathbf{\bar{u}})_{i,j}$  at nonzero entries of  $\mathcal{A}_{a}^{+}(\mathbf{\bar{u}})_{i,j}$ , 401 402 403 i.e., the four red dots in Figure 7 (a) and Figure 7 (c). 404

Thus we only need coefficients of  $u_{i+2,j}, u_{i-2,j}, u_{i,j+2}$ , and  $u_{i,j-2}$  in the final 405 expression of  $[\mathcal{A}^z \mathcal{A}_d^{-1} \mathcal{A}^s](\mathbf{u})_{i,j}$ , which are found to be 406

407 
$$u_{i+2,j}: \quad 4\epsilon_2(1-\epsilon_1)\frac{1}{h_a(h_{a-1}+h_a)}\frac{2h_bh_{b-1}h_a^2}{7h_a^2+4h_bh_{b-1}}\frac{1}{h_a^2}$$

408 
$$u_{i-2,j}: \quad 4\epsilon_2(1-\epsilon_1)\frac{1}{h_{a-1}(h_{a-1}+h_a)}\frac{2h_bh_{b-1}h_{a-1}^2}{7h_{a-1}^2+4h_bh_{b-1}}\frac{1}{h_{a-1}^2}$$
409 
$$u_{i,j+2}: \quad 4\epsilon_2(1-\epsilon_1)\frac{1}{h_b(h_{b-1}+h_b)}\frac{2h_ah_{a-1}h_b^2}{7h_b^2+4h_ah_{a-1}}\frac{1}{h_b^2}$$

409 
$$u_{i,j+2}: 4\epsilon_2(1-\epsilon_1)\frac{1}{h_b(h_{b-1}+h_b)}\frac{2h_ah_{a-1}h_b^2}{7h_i^2+4h_ah_{a-1}h_b^2}$$

410 
$$u_{i,j-2}: \quad 4\epsilon_2(1-\epsilon_1)\frac{1}{h_{b-1}(h_{b-1}+h_b)}\frac{2h_ah_{a-1}h_{b-1}^2}{7h_{b-1}^2+4h_ah_{a-1}}\frac{1}{h_{b-1}^2}$$
411

In order to maintain  $A_a^+ \leq A^z A_d^{-1} A^s$ , by comparing to the coefficients of  $u_{i+2,j}$  for  $\mathcal{A}_a^+(\bar{\mathbf{u}})$ , we obtain a mesh constraint  $4\epsilon_2(1-\epsilon_1)\frac{2h_bh_{b-1}}{7h_a^2+4h_bh_{b-1}} \geq \frac{1}{2}$ . Similar constraints are obtained by comparing other coefficients at  $u_{i,j\pm 2}$  and  $u_{i-2,j}$ . Define

$$\ell(\epsilon_1, \epsilon_2) = 4\epsilon_2(1 - \epsilon_1).$$

Then the following constraints are sufficient for  $\mathcal{A}_a^+(\bar{\mathbf{u}})$  to be controlled by  $\mathcal{A}^z \mathcal{A}_d^{-1} \mathcal{A}^s(\bar{\mathbf{u}})$ 412at an interior knot: 413

414 (5.2a) 
$$h_a h_{a-1} \ge \frac{7}{4\ell - 4} \max\{h_b^2, h_{b-1}^2\}, \quad h_b h_{b-1} \ge \frac{7}{4\ell - 4} \max\{h_a^2, h_{a-1}^2\}.$$

415 Second, we need to discuss the case when  $(x_i, y_j)$  is an interior edge center. With-416 out loss of generality, assume  $(x_i, y_j)$  is an interior edge center of an edge paral-417 lel to the y-axis. Then similar to the interior knot case, the output coefficients of 418  $[\mathcal{A}^z \mathcal{A}_d^{-1} \mathcal{A}^s](\bar{\mathbf{u}})_{i,j}$  at the relevant non-zero entries of  $\mathcal{A}_a^+(\bar{\mathbf{u}})_{i,j}$  are:

419 
$$u_{i+2,j}: \quad 4\epsilon_2(1-\epsilon_1)\frac{1}{h_a(h_{a-1}+h_a)}\frac{h_a^2h_b^2}{2h_a^2+2h_b^2}\frac{1}{h_a^2}$$
420 
$$u_{i-2,j}: \quad 4\epsilon_2(1-\epsilon_1)\frac{1}{h_a(h_{a-1}+h_a)}\frac{h_{a-1}^2h_b^2}{h_{a-1}^2h_b^2}$$

20 
$$u_{i-2,j}: \quad 4\epsilon_2(1-\epsilon_1)\frac{1}{h_{a-1}(h_{a-1}+h_a)}\frac{h_{a-1}^2h_b^2}{2h_{a-1}^2+2h_b^2}\frac{1}{h_{a-1}^2}$$

421 By comparing with coefficients of  $\mathcal{A}_{a}^{+}(\bar{\mathbf{u}})_{i,j}$ , we get  $\frac{h_{b}^{2}}{h_{a}^{2}+h_{b}^{2}} \geq \frac{1}{\ell}$ ,  $\frac{h_{b}^{2}}{h_{a-1}^{2}+h_{b}^{2}} \geq \frac{1}{\ell}$ . 422 To ensure  $\mathcal{A}_{a}^{+}(\bar{\mathbf{u}})$  is controlled by  $\mathcal{A}^{z}\mathcal{A}_{d}^{-1}\mathcal{A}^{s}(\bar{\mathbf{u}})$  at edge centers, it suffices to have:

422 To ensure  $\mathcal{A}_a^+(\bar{\mathbf{u}})$  is controlled by  $\mathcal{A}^z \mathcal{A}_d^{-1} \mathcal{A}^s(\bar{\mathbf{u}})$  at edge centers, it suffices to have: (5.2b) \_\_\_\_\_\_

423 
$$min\{h_a, h_{a-1}\} \ge \sqrt{\frac{1}{\ell-1}}max\{h_b, h_{b-1}\}, min\{h_b, h_{b-1}\} \ge \sqrt{\frac{1}{\ell-1}}max\{h_a, h_{a-1}\}.$$

Note that  $\mathcal{A}_a^+(\bar{\mathbf{u}})_{i,j} = 0$  if  $(x_i, y_j)$  is a cell center. Since  $\mathcal{A}^z \mathcal{A}_d^{-1} \mathcal{A}^s(\bar{\mathbf{u}}) \ge 0$ , there is no mesh constraint to enforce the inequality at cell centers.

426 **5.4.** Mesh constraints for  $A_d + A^z$  being an M-matrix. Let  $\mathcal{B} = \mathcal{A}_d + \mathcal{A}^z$ . 427 Then  $\mathcal{B}(\mathbf{1})_{i,j} = 1$  for a boundary point  $(x_i, y_j)$ . For interior points, we have:

428 
$$\mathcal{B}(\mathbf{1})_{i,j} = -\epsilon_1 \left( \frac{1}{h_a^2} + \frac{1}{h_a^2} + \frac{1}{h_b^2} + \frac{1}{h_b^2} \right) + \frac{2h_a^2 + 2h_b^2}{h_a^2 h_b^2} = (1 - \epsilon_1) \frac{2h_a^2 + 2h_b^2}{h_a^2 h_b^2}, \quad \text{cell center;}$$
429 
$$\mathcal{B}(\mathbf{1})_{i,j} = -\epsilon_1 \left( \frac{1}{h_b^2} + \frac{1}{h_b^2} \right) - \epsilon_2 \left[ \frac{4}{h_a(h_a + h_{a-1})} + \frac{4}{h_{a-1}(h_a + h_{a-1})} \right] + \frac{7h_b^2 + 4h_a h_{a-1}}{2h_a h_{a-1} h_b^2}$$

430 = 
$$(1 - \epsilon_1)\frac{2}{h_b^2} + (1 - \frac{8}{7}\epsilon_2)\frac{7}{2h_ah_{a-1}}$$
, edge center (2);

431 
$$\mathcal{B}(\mathbf{1})_{i,j} = -\epsilon_1 \left( \frac{1}{h_a^2} + \frac{1}{h_a^2} \right) - \epsilon_2 \left[ \frac{4}{h_b(h_b + h_{b-1})} + \frac{4}{h_{b-1}(h_b + h_{b-1})} \right] + \frac{7h_a^2 + 4h_bh_{b-1}}{2h_bh_{b-1}h_a^2}$$
  
432  $= (1 - \epsilon_1)\frac{2}{h_a^2} + (1 - \frac{8}{7}\epsilon_2)\frac{7}{2h_bh_{b-1}}, \quad \text{edge center (3);}$ 

432 = 
$$(1 - \epsilon_1)\frac{2}{h_a^2} + (1 - \frac{3}{7}\epsilon_2)\frac{1}{2h_bh_{b-1}}$$
, edge center

$$\begin{array}{ll} 433 \quad \mathcal{B}(\mathbf{1})_{i,j} = -\epsilon_2 \left[ \frac{4}{h_a(h_a + h_{a-1})} + \frac{4}{h_{a-1}(h_a + h_{a-1})} + \frac{4}{h_b(h_b + h_{b-1})} + \frac{4}{h_{b-1}(h_b + h_{b-1})} \right] \\ \\ 434 \\ 435 \\ \qquad + \frac{7h_bh_{b-1} + 7h_ah_{a-1}}{2h_ah_{a-1}h_bh_{b-1}} = (1 - \frac{8}{7}\epsilon_2)\frac{7h_bh_{b-1} + 7h_ah_{a-1}}{2h_ah_{a-1}h_bh_{b-1}}, \quad \text{interior knot.} \end{array}$$

A36 Notice that larger values of  $\ell$  give better mesh constraints in (5.2). And we have 437  $\sup_{0 < \epsilon_1, \epsilon_2 \le 1} \ell(\epsilon_1, \epsilon_2) = \sup_{0 < \epsilon_1, \epsilon_2 \le 1} 4\epsilon_2(1 - \epsilon_1) = 4$ . In order to apply Theorem 2.1 438 for  $A_d + A^z$  be an M-matrix, we need  $[\mathcal{A}_d + \mathcal{A}^z](1) \ge 0$ . This is true if and only if 439  $\epsilon_1 \le 1$  and  $\epsilon_2 \le \frac{7}{8}$ , which only give  $\sup_{0 < \epsilon_1 \le 1, 0 < \epsilon_2 \le \frac{7}{8}} \ell(\epsilon_1, \epsilon_2) = 3.5$ .

440 **5.5. Improved mesh constraints by the relaxed Lorenz's condition.** To 441 get a better mesh constraint, the constraint on  $\epsilon_2$  can be relaxed so that the value 442 of  $\ell(\epsilon_1, \epsilon_2)$  can be improved. One observation from Section 5.3 is that the value of 443  $\mathcal{A}_d(\bar{\mathbf{u}})_{i,j}$  for  $(x_i, y_j)$  being a knot is not used for verifying  $A_a^+ \leq A^z A_d^{-1} A^s$  (for both 444 interior knots and edge centers). To this end, we define a new diagonal matrix  $A_{d^*}$ ,

which is different from  $A_d$  only at the interior knots. 445

446 
$$\mathcal{A}_{d^*}(\bar{\mathbf{u}})_{i,j} = u_{i,j} = \mathcal{A}_d(\bar{\mathbf{u}})_{i,j}, \text{ if } (x_i, y_j) \text{ is a boundary point;}$$

447 
$$\mathcal{A}_{d^*}(\bar{\mathbf{u}})_{i,j} = \frac{2h_a^2 + 2h_b^2}{h_a^2 h_b^2} u_{i,j} = \mathcal{A}_d(\bar{\mathbf{u}})_{i,j}, \quad \text{if } (x_i, y_j) \text{ is a cell center;}$$

448 
$$\mathcal{A}_{d^*}(\bar{\mathbf{u}})_{i,j} = \frac{7h_b^2 + 4h_ah_{a-1}}{2h_ah_{a-1}h_b^2}u_{i,j} = \mathcal{A}_d(\bar{\mathbf{u}})_{i,j}, \quad \text{edge center (2)};$$

449 
$$\mathcal{A}_{d^*}(\bar{\mathbf{u}})_{i,j} = \frac{7h_a^2 + 4h_bh_{b-1}}{2h_bh_{b-1}h_a^2}u_{i,j} = \mathcal{A}_d(\bar{\mathbf{u}})_{i,j}, \quad \text{edge center (3);}$$

450 
$$\mathcal{A}_{d^*}(\bar{\mathbf{u}})_{i,j} = \frac{8h_bh_{b-1} + 8h_ah_{a-1}}{2h_ah_{a-1}h_bh_{b-1}}u_{i,j} \neq \mathcal{A}_d(\bar{\mathbf{u}})_{i,j}, \text{ if } (x_i, y_j) \text{ is an interior knot.}$$

Since the values of  $\mathcal{A}_d(\bar{\mathbf{u}})_{i,j}$  for  $(x_i, y_j)$  being a knot is not involved in Section 5.3, 452the same discussion in Section 5.3 also holds for verifying  $A_a^+ \leq A^z A_{d^*}^{-1} A^s$ . Namely, under mesh constraints (5.2), we also have  $A_a^+ \leq A^z A_{d^*}^{-1} A^s$ . Let  $B^* = A_{d^*} + A^z$ , then the row sums of  $B^*$  are: 453454455

456 
$$\mathcal{B}^*(\mathbf{1})_{i,j} = 1$$
, if  $(x_i, y_j)$  is a boundary point;  
457  $\mathcal{B}^*(\mathbf{1})_{i,j} = -\epsilon_1 \left( \frac{1}{h_a^2} + \frac{1}{h_a^2} + \frac{1}{h_b^2} + \frac{1}{h_b^2} \right) + \frac{2h_a^2 + 2h_b^2}{h_a^2 h_b^2} = (1 - \epsilon_1) \frac{2h_a^2 + 2h_b^2}{h_a^2 h_b^2}$ , cell center;

458 
$$\mathcal{B}^{*}(\mathbf{1})_{i,j} = -\epsilon_{1} \left( \frac{1}{h_{b}^{2}} + \frac{1}{h_{b}^{2}} \right) - \epsilon_{2} \left[ \frac{4}{h_{a}(h_{a} + h_{a-1})} + \frac{4}{h_{a-1}(h_{a} + h_{a-1})} \right] + \frac{7h_{b}^{2} + 4h_{a}h_{a-1}}{2h_{a}h_{a-1}h_{b}^{2}}$$
459 
$$= (1 - \epsilon_{1})\frac{2}{L^{2}} + (1 - \frac{8}{7}\epsilon_{2})\frac{7}{2L}, \quad \text{edge center (2);}$$

459 
$$= (1-\epsilon_1)\frac{2}{h_b^2} + (1-\frac{8}{7}\epsilon_2)\frac{7}{2h_ah_{a-1}}, \quad \text{edge center} ($$

$$460 \quad \mathcal{B}^{*}(\mathbf{1})_{i,j} = -\epsilon_{1} \left( \frac{1}{h_{a}^{2}} + \frac{1}{h_{a}^{2}} \right) - \epsilon_{2} \left[ \frac{4}{h_{b}(h_{b} + h_{b-1})} + \frac{4}{h_{b-1}(h_{b} + h_{b-1})} \right] + \frac{7h_{a}^{2} + 4h_{b}h_{b-1}}{2h_{b}h_{b-1}h_{a}^{2}}$$

$$461 \qquad = (1 - \epsilon_{1})\frac{2}{L^{2}} + (1 - \frac{8}{7}\epsilon_{2})\frac{7}{2L}, \quad \text{edge center (3);}$$

461 
$$= (1-\epsilon_1)\frac{2}{h_a^2} + (1-\frac{8}{7}\epsilon_2)\frac{7}{2h_bh_{b-1}}, \quad \text{edge center}$$

$$462 \quad \mathcal{B}^{*}(\mathbf{1})_{i,j} = -\epsilon_{2} \left[ \frac{4}{h_{a}(h_{a} + h_{a-1})} + \frac{4}{h_{a-1}(h_{a} + h_{a-1})} + \frac{4}{h_{b}(h_{b} + h_{b-1})} + \frac{4}{h_{b-1}(h_{b} + h_{b-1})} \right]$$

$$463 \qquad + \frac{8h_{b}h_{b-1} + 8h_{a}h_{a-1}}{2h_{a}h_{a-1}h_{b}h_{b-1}} = (1 - \epsilon_{2})\frac{8h_{b}h_{b-1} + 8h_{a}h_{a-1}}{2h_{a}h_{a-1}h_{b}h_{b-1}}, \quad \text{interior knot.}$$

Now  $[\mathcal{A}_{d^*} + \mathcal{A}^z](\mathbf{1})_{i,j} \geq 0$  at cell centers and knots is true if and only if  $\epsilon_1 \leq 1$ 465 and  $\epsilon_2 \leq 1$ . 466

Next, we will show that the mesh constraints (5.2) with  $0 < \epsilon_1 \leq \frac{1}{2}$  and  $\epsilon_2 = 1$  are sufficient to ensure  $[\mathcal{A}_{d^*} + \mathcal{A}^z](\mathbf{1})_{i,j} \geq 0$  at edge centers. We have  $0 < \epsilon_1 \leq \frac{1}{2}, \epsilon_2 = 1 \implies 2 \leq \ell < 4 \implies \frac{7}{4\ell-4} \geq \frac{1}{\ell}$ . The mesh constraints (5.2) imply that 467 468 469  $h_a h_{a-1} \geq \frac{7}{4\ell-4} h_b^2 \geq \frac{1}{\ell} h_b^2$ , thus 470

$$\begin{array}{l} {}_{471} \quad (1-\epsilon_1)\frac{2}{h_b^2} + (1-\frac{8}{7}\epsilon_2)\frac{7}{2h_ah_{a-1}} = (1-\epsilon_1)\frac{2}{h_b^2} - \frac{1}{2}\frac{1}{h_ah_{a-1}} = \frac{1}{2}\left[\frac{\ell}{h_b^2} - \frac{1}{h_ah_{a-1}}\right] \ge 0. \end{array}$$

Similarly,  $(1-\epsilon_1)\frac{2}{h_a^2} + (1-\frac{8}{7}\epsilon_2)\frac{7}{2h_bh_{b-1}} \ge 0$  also holds. 473

Therefore, for constants  $0 < \epsilon_1 \leq \frac{1}{2}$  and  $\epsilon_2 = 1$ , we have  $[\mathcal{A}_{d^*} + \mathcal{A}^z](\mathbf{1}) \geq \mathbf{0}$ . In 474particular, we have a larger  $\ell$  compared to constraints from  $\mathbf{A}_d$ . 475

This manuscript is for review purposes only.

476 **5.6. The main result.** We have shown that for two constants  $0 < \epsilon_1 \leq \frac{1}{2}$  and 477  $\epsilon_2 = 1$ , under mesh constraints (5.2), the matrices  $A_{d^*}$ ,  $A^z$ ,  $A^s$  constructed above 478 satisfy  $(A_{d^*} + A^z)\mathbf{1} \geq \mathbf{0}$  and  $A_a^+ \leq A^z A_{d^*}^{-1} A^s$ .

For any fixed  $\epsilon_1 > 0$  and  $\epsilon_2 = 1$ ,  $\tilde{A}^z$  also has the same sparsity pattern as A. Thus if  $\ell$  in (5.2) is replaced by  $\sup_{0 < \epsilon_1 \le \frac{1}{2}, \epsilon_2 = 1} \ell(\epsilon_1, \epsilon_2) = 4$ , Theorem 4.8 still applies to conclude that  $A^{-1} \ge 0$ .

482 THEOREM 5.1. The  $Q^2$  variational difference scheme (5.1) has a monotone ma-483 trix  $\bar{L}_h$  thus satisfies discrete maximum principle under the following mesh constraints:

$$h_{a}h_{a-1} \geq \frac{7}{12}max\{h_{b}^{2}, h_{b-1}^{2}\}, \quad h_{b}h_{b-1} \geq \frac{7}{12}max\{h_{a}^{2}, h_{a-1}^{2}\},$$

$$(5.3) \quad \min\{h_{a}, h_{a-1}\} \geq \sqrt{\frac{1}{3}}max\{h_{b}, h_{b-1}\}, \quad \min\{h_{b}, h_{b-1}\} \geq \sqrt{\frac{1}{3}}max\{h_{a}, h_{a-1}, \}$$

where  $h_a, h_{a-1}$  are mesh sizes for x-axis and  $h_b, h_{b-1}$  are mesh sizes for y-variable in four adjacent rectangular cells as shown in Figure 6.

487 REMARK 4. The following global constraint is sufficient to ensure (5.3):

$$\begin{array}{l} 488 \\ 489 \end{array} (5.4) \qquad \qquad \frac{25}{32} \le \frac{h_m}{h_n} \le \frac{32}{25}, \end{array}$$

484

490 where  $h_m$  and  $h_n$  are any two grid spacings in a non-uniform grid generated from a 491 non-uniform rectangular mesh for  $Q^2$  elements.

492 REMARK 5. For  $Q^1$  finite element method solving  $-\Delta u = f$  to satisfy discrete 493 maximum principle on non-uniform rectangular meshes [7], the mesh constraints are

494 (5.5) 
$$h_a h_{a-1} \ge \frac{1}{2} max\{h_b^2, h_{b-1}^2\} \quad h_b h_{b-1} \ge \frac{1}{2} max\{h_a^2, h_{a-1}^2\}.$$

5.7. Necessity of Mesh Constraints. Even though the mesh constraints de-495rived above are only sufficient conditions, in practice a mesh constraint is still neces-496sary for the inverse positivity to hold. Consider a non-uniform  $Q^2$  mesh with  $5 \times 5$ 497 cells on the domain  $[0,1] \times [0,1]$ , which has a  $9 \times 9$  grid for the interior of the domain. 498Let the mesh on both axes be the same and let the four outer-most cells for each 499dimension be identical with length 2h. Then the middle cell has size  $2h' \times 2h'$  with 500  $h' = \frac{1}{2} - 2h$ . Let the ratio h'/h increase gradually from h'/h = 1 (a uniform mesh) 501until the minimum value of the inverse of the matrix becomes negative. Increasing by 502 values of 0.05, we obtain the first negative entry of  $\bar{L}_h^{-1}$  at h'/h = 5.35 with h = 0.0535and h' = 0.2861 shown in Figure 8 (a). Figure 8 (b) shows how the minimum entry 503 504of  $\bar{L}_h^{-1}$  decreases as h'/h increases. 505

6. Monotonicity of  $Q^3$  variational difference scheme on a uniform mesh. Even though Lorenz's condition can be nicely verified for the  $Q^2$  scheme, it is very difficult to apply Lorenz's condition to higher order schemes due to their much more complicated structure. In particular, even for  $Q^3$  scheme, simple decomposition of  $A_a^- = A^z + A^s$  such that  $A_a^+ \leq A^z A_d^{-1} A^s$  is difficult to show. Instead, we propose to apply Lorenz's theorems to a few simpler intermediate matrices. To be specific, let  $A = A_3$  be the matrix representation of the scheme, and let  $A_0 = M_1$  be an M-matrix. Then we construct matrices  $A_i$  and  $L_i$  such that

514 
$$A_1 \le A_0 L_0, \quad A_2 \le A_1 L_1, \quad A_3 \le A_2 L_2,$$





(a) A non-uniform mesh with 5 × 5 cells on which the  $C^0 - Q^2$  scheme is not inverse positive. The minimum value of  $\bar{L}_h^{-1}$  is -6.14E-8.

(b) A plot of the minimum value of  $\bar{L}_h^{-1}$  as h'/h increases.

FIG. 8. Necessity of mesh constraints for inverse positivity  $\bar{L}_h^{-1} \ge 0$  where  $\bar{L}_h$  is the matrix in  $Q^2$  variational difference scheme on non-uniform meshes.

with the constraints that  $A_i \mathbf{1} \ge 0$  and  $A_0 = M_1$  connects  $\mathcal{N}^0(A_i \mathbf{1})$  with  $\mathcal{N}^+(A_i \mathbf{1})$  for all  $A_i$ . By Theorem 4.3, then we have

$$A_1 \leq A_0 L_0 = M_1 L_0 \Rightarrow A_1 = M_1 M_2 \Rightarrow A_2 \leq M_1 M_2 L_1 \Rightarrow A_2 = M_1 M_2 M_3$$
$$\Rightarrow A_3 \leq M_1 M_2 M_3 L_2 \Rightarrow A = A_3 = M_1 M_2 M_3 M_4.$$

# 515 The matrices $A_i$ and $L_i$ satisfying constraints above are not unique.

**6.1. One-dimensional scheme.** We first demonstrate the main idea for the one-dimensional case, for which we only need to construct matrices such that  $A_1 \leq A_0 L_0, A \leq A_1 L_1$ .

519 Let  $\bar{L}_h$  denote the coefficient matrix in (3.6), then consider  $A = \frac{\hbar^2}{4}\bar{L}_h$ . For 520 convenience, we will perceive the matrix A as a linear operator  $\mathcal{A}$ . Notice that the 521 coefficients for two interior points are symmetric in (3.6), thus we will only show 522 stencil for the left interior point for simplicity:

523 
$$\mathcal{A}$$
 at boundary point  $x_0$  or  $x_{n+1} : \frac{\mathbf{h}^2}{4}$   
524  $\mathcal{A}$  at knot  $: -\frac{1}{4}$   $\frac{15\sqrt{5}-25}{8}$   $\frac{-15\sqrt{5}-25}{8}$  **13**  $\frac{-15\sqrt{5}-25}{8}$   $\frac{15\sqrt{5}-25}{8}$   $-\frac{1}{4}$   
525  $\mathcal{A}$  at interior point  $:$   $\frac{-3\sqrt{5}-5}{4}$  **5**  $-\frac{5}{2}$   $\frac{3\sqrt{5}-5}{4}$ ,

527 where bolded entries indicate the coefficient for the operator output location  $x_i$ .

For all the matrices defined below, they will have symmetric structure at two interior points, thus for simplicity we will only show the stencil of the corresponding linear operators for the left interior point. We first define three matrices  $A_1$ ,  $A_0$ , and 531  $Z_0$ .

532 
$$A_1$$
 at boundary :  $\frac{h^2}{4}$   
533  $A_1$  at knot :  $0 \quad \frac{15\sqrt{5}-25}{8} \quad -7 \quad 13 \quad -7 \quad \frac{15\sqrt{5}-25}{8}$   
534  $A_1$  at interior point:  $-\frac{1}{2} \quad 4.8 \quad -2 \quad 0$   
535  $A_0$  at boundary :  $\frac{h^2}{4}$   
536  $A_0$  at knot:  $0 \quad 0 \quad -7 \quad 15 \quad -7 \quad 0 \quad 0$   
537  $A_0$  at interior point:  $-\frac{1}{2} \quad 4.8 \quad -\frac{1}{2} \quad 0$   
538  $Z_0$  at boundary :  $0$   
539  $Z_0$  at knot:  $0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0$   
540  $Z_0$  at interior point:  $0 \quad 0 \quad -2 + \frac{1}{2} \quad 0$ 

Then we define  $L_0 = I + (A_0)_d^{-1} Z_0$  where *I* is the identity matrix and  $(A_0)_d$  denotes the diagonal part of  $A_0$ . By considering composition of two operators  $\mathcal{A}_0$  and  $\mathcal{L}_0$ , we get the matrix product  $A_0 L_0$ . Due to the definition of  $Z_0$ ,  $\mathcal{A}_0 \mathcal{L}_0$  still has the same stencil as above:

546 
$$\mathcal{A}_0 \mathcal{L}_0$$
 at boundary :  $\frac{\mathbf{h}^2}{4}$   
547  $\mathcal{A}_0 \mathcal{L}_0$  at knot:  $0 \quad \frac{35}{16} \quad -7 \quad \mathbf{15} \quad -7 \quad \frac{35}{16}$ 

$$\mathcal{A}_0 \mathcal{L}_0 \text{ at interior point:} \quad -\frac{1}{2} \quad \mathbf{4.8} + \frac{\mathbf{5}}{\mathbf{32}} \quad -2 \quad 0$$

It is straightforward to see  $A_1 \leq A_0 L_0$ . By Theorem 2.1,  $A_0$  is an M-matrix, thus we set  $M_1 = A_0$ . Also it is easy to see that  $\mathcal{A}_1(\mathbf{1}) > 0$  thus  $\mathcal{N}^0(A_1\mathbf{1})$  is an empty set. So  $A_0$  trivially connects  $\mathcal{N}^0(A_1\mathbf{1})$  with  $\mathcal{N}^+(A_1\mathbf{1})$ . By Theorem 4.3, we have  $A_1 \leq A_0 L_0 = M_1 L_0 \Rightarrow A_1 = M_1 M_2$  where  $M_2$  is an M-matrix.

Let  $(A_1)_d$  denote the diagonal part of  $A_1$ . Then define  $L_1 = I + (A_1)_d^{-1} Z_1$  using the following  $Z_1$ :

556  $\mathcal{Z}_1$  at boundary: **0** 

557
 
$$Z_1$$
 at knot:
 0
 0
 0
 0
 0
 0

 558
  $Z_1$  at interior point:
  $-\frac{11}{10}$ 
 $\mathbf{0}$ 
 $-\frac{1}{2}$ 
 $0$ 

560 And the matrix  $A_1L_1$  still have the same stencil and symmetry:

561 
$$\mathcal{A}_{1}\mathcal{L}_{1}$$
 at boundary:  $\frac{\mathbf{h}^{2}}{4}$   
562  $\mathcal{A}_{1}\mathcal{L}_{1}$  at knot:  $\frac{-165\sqrt{5}+275}{384}$   $\frac{15\sqrt{5}-25}{8} + \frac{35}{48}$   $-7 + \frac{-75\sqrt{5}+125}{384}$   $\mathbf{13} + 2(\frac{77}{48})$   $-7 + \frac{-75\sqrt{5}+125}{384}$   $\frac{15\sqrt{5}-25}{8} + \frac{35}{48}$   $\frac{-165\sqrt{5}+275}{384}$   
563  $\mathcal{A}_{1}\mathcal{L}_{1}$  at interior point:  $-\frac{8}{5}$   $\mathbf{4.8} + \frac{\mathbf{5}}{\mathbf{24}}$   $-\frac{5}{2}$   $\frac{11}{24}$ 

A direct comparison verifies that  $A \leq A_1L_1 = M_1M_2L_1$ . Also it is easy to see that  $\mathcal{A}(\mathbf{1})_i = 0$  if  $x_i$  is not a boundary point. The operator  $\mathcal{A}_0$  has a three-point

0

stencil at interior grid points, thus the directed graph defined by the adjacency matrix  $A_0$  has a directed path starting from any interior grid point to any other point, see Figure 9. So  $M_1 = A_0$  connects  $\mathcal{N}^0(A\mathbf{1})$  with  $\mathcal{N}^+(A\mathbf{1})$ . By Theorem 4.3, we have  $A \leq A_1L_1 = M_1M_2L_1 \Rightarrow A = M_1M_2M_3$  where  $M_3$  is an M-matrix. Therefore,  $A^{-1} = M_3^{-1}M_2^{-1}M_1^{-1} \geq 0.$ 



FIG. 9. The directed graph defined by matrix  $M_1$  for the finite difference grid shown in Figure 5.



(a) Three point types defining the stench. knot (black), edge point (blue), interior point (green).

(b) The directed graph defined by the matrix  $M_1$ .

FIG. 10. An illustration of a  $Q^3$  mesh with  $2 \times 2$  cells.

**6.2.** Two-dimensional case. Due to symmetry, the stencil of the scheme can be defined at three different types of points, see Figure 10 (a). Let each rectangular cell have size  $h \times h$  and denote  $Q^3$  scheme by  $\bar{L}_h \bar{\mathbf{u}} = \bar{\mathbf{f}}$ . Let  $A = \frac{h^2}{4}\bar{L}_h$ . Then for a boundary point  $(x_i, y_j) \in \partial\Omega$ ,  $\mathcal{A}(\bar{\mathbf{u}})_{ij} = \frac{h^2}{4}u_{ij}$ . And the stencil of  $\mathcal{A}$  at interior grid points is given as

$$-\frac{1}{4}$$

$$\frac{15\sqrt{5}-25}{8}$$

$$-\frac{15\sqrt{5}-25}{8}$$

$$\frac{-15\sqrt{5}-25}{8}$$

$$\frac{26}{8}$$

$$\frac{-15\sqrt{5}-25}{8}$$

$$\frac{15\sqrt{5}-25}{8}$$

$$\frac{15\sqrt{5}-25}{8}$$

$$\frac{15\sqrt{5}-25}{8}$$

$$\frac{15\sqrt{5}-25}{8}$$

$$-\frac{1}{4}$$

$$-\frac{1}{4}$$

$$\frac{15\sqrt{5}-25}{8}$$

$$-\frac{1}{4}$$

$$\frac{15\sqrt{5}-25}{8}$$

$$-\frac{1}{4}$$

$$\frac{15\sqrt{5}-25}{8}$$

$$-\frac{5}{2}$$

$$\frac{3\sqrt{5}-5}{4}$$

$$-\frac{5}{2}$$

$$\frac{18}{8}$$

$$\frac{-3\sqrt{5}-5}{4}$$

 $\mathcal{A}$ 

Next we list the definition of matrices  $A_i$  and  $Z_i$  by the corresponding linear operators  $\mathcal{A}_i$  and  $\mathcal{Z}_i$ . For convenience, we will only list the stencil at interior grid points. For the domain boundary points  $(x_i, y_j) \in \partial\Omega$ , all  $A_i$  matrices will have the same value as A:  $\mathcal{A}_i(\bar{\mathbf{u}})_{ij} = \frac{h^2}{4}u_{ij}$ . And  $\mathcal{Z}_i(\bar{\mathbf{u}})_{ij} = 0$  for  $(x_i, y_j) \in \partial\Omega$ . The matrix  $L_i$  is defined as  $L_i = I + (A_i)_d^{-1}Z_i$ , i = 0, 1, 2. The matrices and their products are given by:

> 0  $\frac{15\sqrt{5}-25}{2}$  $\frac{-15\sqrt{5}-25}{8}$  $A_1$  at knot:  $0 \quad \frac{15\sqrt{5}-25}{8} \quad \frac{-15\sqrt{5}-25}{8}$  $\mathbf{26}$  $\frac{-15\sqrt{5}-25}{8}$   $\frac{15\sqrt{5}-25}{8}$  0  $\frac{-15\sqrt{5}-25}{2}$  $\frac{15\sqrt{5}-25}{8}$ 0 00 0 -7 $-\frac{1}{2}$  $A_1$  at edge point:  $0 -\frac{5}{2}$ 17 $-rac{1}{100}$  $\mathcal{A}_1$  at interior point:  $0 -\frac{1}{2}$  $\mathbf{10}$  $-\frac{1}{2}$  $-\frac{1}{2}$ -700





This manuscript is for review purposes only.



By Theorem 2.1,  $A_0$  is an M-matrix, thus we set  $M_1 = A_0$ . Notice that the matrix  $M_1 = A_0$  has a 5-point stencil and the directed graph defined by  $M_1$  is given in Figure 10 (b), in which there is a directed path starting from any interior grid point to any other point. For convenience, let  $A_3 = A$ . Then we have  $\mathcal{A}_k(\mathbf{1}) \ge 0$  (k = 0, 1, 2, 3). Moreover,  $\mathcal{A}_k(\mathbf{1})_{ij} > 0$  (k = 0, 1, 2, 3) for domain boundary point ( $x_i, y_j$ )  $\in \partial\Omega$ . The directed graph defined by  $M_1$  easily implies that  $M_1$  connects  $\mathcal{N}^0(A_i\mathbf{1})$  with  $\mathcal{N}^+(A_i\mathbf{1})$ for all i = 0, 1, 2, 3.

By straightforward comparison, we can verify that  $A_1 \leq A_0 L_0, A_2 \leq A_1 L_1, A \leq A_2 L_2$ . By Theorem 4.3, we have

$$A_1 \le A_0 L_0 = M_1 L_0 \Rightarrow A_1 = M_1 M_2 \Rightarrow A_2 \le M_1 M_2 L_1 \Rightarrow A_2 = M_1 M_2 M_3$$
$$\Rightarrow A \le M_1 M_2 M_3 L_2 \Rightarrow A = M_1 M_2 M_3 M_4 \Rightarrow A^{-1} \ge 0.$$

590 7. Numerical Tests.

7.1. Efficient implementation. For all schemes discussed in this paper, except 591 the  $P^2$  variational difference scheme, the stiffness matrix can be efficiently inverted 592by an eigenvector method, very similar to the inversion of 5-point discrete Laplacian by Fast Fourier Transform (FFT). We demonstrate it for the 9-point scheme on a 594 $N_x \times N_y$  grid. For instance, the stiffness matrix in the scheme (2.2) can be written as 595

596 
$$-\frac{1}{6h^2}[H_x \otimes H_y - 36I_x \otimes I_y]$$

where  $\otimes$  is the Kronecker product,  $I_x$  is the identity matrix of size  $N_x \times N_x$  and  $H_x$ 597 and  $H_y$  are symmetric tridiagonal matricies of size  $N_x \times N_x$  and  $N_y \times N_y$  respectively, 598 with H defined as 599

600

$$H = \begin{pmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & & \\ & 1 & 4 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 4 & 1 \\ & & & & 1 & 4 \end{pmatrix}$$

Let  $H = S\Lambda S^{-1}$  be the eigenvalue decomposition of H, then we also have the eigen-601 value decomposition of stiffness matrix 602

$$603 \qquad -\frac{1}{6h^2}[H_x \otimes H_y - 36I_x \otimes I_y] = -\frac{1}{6h^2}(S_x \otimes S_y)(\Lambda_x \otimes \Lambda_y - 36I_x \otimes I_y)(S_x^{-1} \otimes S_y^{-1}).$$

Therefore, the stiffness matrix can be efficiently inverted by the eigenvector method, 604 e.g., Section 7.4 in [19]. Moreover, for a matrix H of size  $n \times n$ , its m-th eigenvector is 605  $\left[\sin(m\pi\frac{1}{n+1}) \cdots \sin(m\pi\frac{n}{n+1})\right]^T$  and corresponding eigenvalue is  $4+2\cos(m\pi\frac{1}{n+1})$ . Thus multiplication of S and  $S^{-1}$  can be implemented through FFT. 606 607

The stiffness matrix in the  $Q^k$  variational difference scheme also has a kron struc-608 ture [19]. But such a kron structure does not seem possible for (3.4). For  $Q^2/Q^3$  and 609 Bramble-Hubbard schemes, the eigenvectors for the small matrices  $H_x$  and  $H_y$  can 610 be computed numerically. 611

7.2. Accuracy tests. We show some accuracy tests of the schemes mentioned 612 in this paper for solving  $-\Delta u = f$  on a square  $(0,1) \times (0,1)$  with Dirichlet boundary 613 conditions. Quasi-uniform meshes were generated by setting each pair of consecutive 614finite element cells along the axis to have a fixed ratio  $\frac{h_k}{h_{k-1}} = 1.01$ . We will simply 615 refer to the classical 9-point scheme (2.1) as 9-point scheme, and refer to its variant 616 (2.3) as compact finite difference. The schemes are tested for the following very 617 618 smooth solutions:

1. The Laplace equation  $-\Delta u = 0$  with Dirichlet boundary conditions and 619  $u(x,y) = log((x+1)^2 + (y+1)^2) + sin(y)e^x.$ 620

621 2. Poisson equation 
$$-\Delta u = f$$
 with homogeneous Dirichlet boundary condition:

622 (7.1) 
$$f(x,y) = 13\pi^2 \sin(3\pi y)\sin(2\pi x) + 2y(1-y) + 2x(1-x) u(x,y) = \sin(3\pi y)\sin(2\pi x) + xy(1-x)(1-y)$$

3. Poisson equation  $-\Delta u = f$  with nonhomogeneous Dirichlet boundary condi-623 tion: 624

- . 2

625 
$$f = 74\pi^2 \cos(5\pi x)\cos(7\pi y) - 8$$

$$u = \cos(5\pi x)\cos(7\pi y) + x^2 + y^2$$

The errors of fourth order accurate schemes on uniform grids are listed in Table 1, Table 2 and Table 3. The errors of  $Q^2$  and  $P^2$  variational difference scheme on quasi uniform rectangular meshes are listed in Table 4. The errors of  $Q^3$  variational difference scheme on uniform rectangular meshes are listed in Table 5. For the Laplace equation, 9-point scheme (2.1) and compact finite difference (2.3) are the same scheme and they are indeed sixth order accurate, see Remark 1.

Finite Difference Coid	$Q^2$	variation	al diff	ferenc	e	$P^2$	$P^2$ variational difference					9-point scheme (2.1)			
Finite Difference Grid	l <sup>2</sup> error	order	$l^{\infty}$ e	error	order	l <sup>2</sup> error	order	$l^{\infty}$	error	order	$l^2$ error	order	$l^{\infty}$ erro	r order	
7 × 7	1.04E-5	-	- 2.50E-5		-	2.05E-5	-	3.8	89E-5	-	1.50E-9	-	3.52E-9	) -	
$15 \times 15$	6.91E-7	3.92	3.92 1.81E-6		3.78	1.38E-6	3.89	2.8	83E-6	3.78	2.35E-11	5.99	5.51E-1	1 6.00	
$31 \times 31$	4.42E-08	3 3.96	1.26	6E-7	3.83	8.93E-08	3.95	2.0	)5E-7	3.78	3.98E-13	5.88	8.89E-1	3 5.95	
$63 \times 63$	2.79E-9	3.98	8.56	E-9	3.88	5.65E-9	3.98	1.4	41E-8	3.85	1.32E-13	1.58	2.37E-1	3 1.90	
Finite Difference	Crid	compact finite difference $(2.3)$							Bramble-Hubbard scheme						
Finite Difference Grid		$l^2  \mathrm{err}$	ror order		der	$l^{\infty}$ error	ore	order		error	order	$l^{\infty}$	error	order	
$7 \times 7$		1.50E	1.50E-9 -		-	3.52E-9	-	- 5.04		)4E-5	-	6.9	7E-5	-	
$15 \times 15$		2.35E	E-11 5.		99	5.51E-11	6.	00	3.7	75E-6	3.74	5.34	IE-06	3.70	
$31 \times 31$		3.98E	-13	-13 5.8		8.89E-13	5.9	5.95		52E-7	3.89	3.8	6E-7	3.78	
$63 \times 63$		1.32E	-13	1.	58	2.37E-13	1.	1.90		3E-08	3.95	2.7	7E-8	3.80	

TABLE 1 Accuracy test on uniform meshes for  $-\Delta u = 0$ .

TABLE 2Accuracy test on uniform meshes for (7.1).

Finite Difference Crid	$Q^2$	variation	nal dif	ferenc	ce	$P^2$ variational difference					9-point scheme (2.1)				
Finite Difference Grid	$l^2$ error	order	r $l^{\infty}$ error		order	$l^2$ error	order	$l^{\infty}$	error	order	$l^2$ error	order	$l^{\infty}$ erro	or order	
$7 \times 7$	2.22e-02	-	4.90e-02		-	4.50e-02	-	1.6	7e-01	-	2.22e-04	-	4.45e-0	4 -	
$15 \times 15$	1.31e-03	4.08	3.03e-03		4.01	2.49e-03	4.17	9.4	2e-03	4.15	5.63e-06	5.30	1.12e-0	5 5.30	
$31 \times 31$	8.04e-05	4.02	1.88e-04		4.01	1.50e-04	4.05	5.6	9e-04	4.04	2.32e-07	4.59	4.65e-0	4.59	
$63 \times 63$	5.00e-06	4.00	1.17	e-05	4.00	9.30e-06	4.01	3.5	2e-05	4.01	1.27e-08	4.19	2.54e-0	4.19	
Finite Difference Grid		COI	mpa	ct fi	nite o	ifference (2.3) Brar					mble-Hubbard scheme				
		$l^2  \mathrm{err}$	or order		ler	$l^{\infty}$ error	or order		$l^2$	error	order	$l^{\infty}$	error	order	
$7 \times 7$		3.18I	2-3 -		-	6.36E-3	3 -		3.74E-2		-	8.65	2E-2	-	
$15 \times 15$		1.91I	E-4	2-4 4.0		3.82E-4	4	4.05 2		6E-3	3.98	5.28	8E-3	4.02	
$31 \times 31$		1.18I	E-5	4.	01	2.36E-5	4	01	1.0	1E-4	4.54	2.1	1E-4	4.64	
$63 \times 63$		7.38I	E-7	4.	00	1.47E-6	4	00	4.1	7E-6	4.60	7.89	9E-6	4.74	

Finite Difference Grid	$Q^2$	variatio	nal dif	feren	ce	$P^2$ variational difference					9-point scheme (2.1)				
Third Difference Offic	$l^2$ error	order	$l^{\infty}$ e	error	order	$l^2$ error	order	$l^{\infty}$	error	order	$l^2$ error	order	$l^{\infty}$ error	order	
$7 \times 7$	3.62E-1	-	1.10	1.10E-0		9.68E-1	-	2.5	9E-0	-	2.48E-2	-	5.69E-2	-	
$15 \times 15$	3.75E-2	3.26	9.68E-2		3.50	7.81E-2	3.63	3.0	0E-1	3.11	2.61E-4	6.56	6.46E-4	6.45	
$31 \times 31$	2.44E-3	3.94	7.18	7.18E-3		4.70E-3	4.05	1.8	4E-2	4.02	3.65E-5	2.84	8.97E-5	2.85	
63  imes 63	1.54E-4	3.98	5.50	E-4	3.70	2.89E-4	4.02	1.1	1E-3	4.04	2.55E-6	3.83	6.57E-6	3.77	
Finite Difference Grid		coi	mpa	et fi	nite o	ifference (2.3) Bran					mble-Hubbard scheme				
		$l^2 \mathrm{err}$	ror	or orde		$l^{\infty}$ error	error ord		$l^2$	error	order	$l^{\infty}$	error	order	
7  imes 7		9.881	88E-2 -			2.26E-1		3.14E-1		4E-1	- 8.2		3E-1	-	
$15 \times 15$		5.401	E-3 4.1		19	1.33E-2	4.	4.08		6E-2	4.15	6.1	6E-2	3.73	
$31 \times 31$		3.22I	E-4	4.0	06	7.91E-4	4.	07	3.3	8E-3	2.37	1.1	5E-2	2.41	
63  imes 63		1.98H	E-5	4.0	01	5.11E-5	3.	95 3.0		4E-4	3.47 1.2		0E-3	3.32	

TABLE 3Accuracy test on uniform meshes for (7.2).

8. Concluding remarks. We reviewed four existing high order monotone discrete Laplacian. By verifying a relaxed Lorenz's condition, we have discussed suitable mesh constraints, under which the fourth order accurate  $Q^2$  variational difference on quasi-uniform meshes is monotone. The fifth order accurate  $Q^3$  variational difference scheme on a uniform mesh is proven be a product of four M-matrices thus inverse positive.

```
This manuscript is for review purposes only.
```

Finita Difference Chid	Patio hi	$Q^2$ variati	ional difference	$P^2$ variational difference					
Finite Difference Grid	fratio $\overline{h_{i-1}}$	$l^{\infty}$ error	order	$l^\infty$ error	order				
test on $-\Delta u = 0$									
$7 \times 7$	1.01	2.66E-5	-	3.98E-5	-				
$15 \times 15$	1.01	1.97E-6	3.74	3.17E-6	3.65				
$31 \times 31$	1.01	1.54E-7	3.67	2.57E-7	3.62				
63  imes 63	1.01	1.37E-8	3.49	2.74E-8	3.22				
test on (7.1)									
$7 \times 7$	1.01	4.92E-2	-	1.69E-1	-				
$15 \times 15$	1.01	3.19E-3	3.94	9.90E-3	4.10				
$31 \times 31$	1.01	2.29E-4	3.79	6.72E-4	3.87				
63  imes 63	1.01	1.80E-5	3.67	5.34E-5	3.65				
test on (7.2)									
$7 \times 7$	1.01	1.20E-0	-	2.95E-0	-				
$15 \times 15$	1.01	1.03E-1	3.54	3.56E-1	3.05				
$31 \times 31$	1.01	9.10E-3	3.50	2.48E-2	3.84				
$63 \times 63$	1.01	9.64E-4	3.23	1.80E-3	3.77				

TABLE 4Accuracy test on quasi-uniform meshes.

# TABLE 5 Accuracy test of $Q^3$ variational difference scheme on uniform meshes.

$Q^3$ Finite Element Mesh	Finite Difference Grid	$l^2$ error	order	$l^{\infty}$ error	order					
test on $-\Delta u = f$										
$2 \times 2$	$5 \times 5$	1.89E-4	-	4.71E-4	-					
$4 \times 4$	$11 \times 11$	6.88E-8	4.78	2.46E-7	4.26					
$8 \times 8$	$23 \times 23$	2.23E-9	4.88	9.90E-9	4.64					
$16 \times 16$	$47 \times 47$	7.61E-11	4.94	3.98E-10	4.64					
$32 \times 32$	95  imes 95	2.44E-12	4.96	1.41E-11	4.82					
test on (7.1)										
$2 \times 2$	$5 \times 5$	3.28E-2	-	5.53E-2	-					
$4 \times 4$	$11 \times 11$	1.58E-3	4.38	3.51E-3	3.98					
8 × 8	$23 \times 23$	4.81E-5	5.03	1.13E-4	4.96					
$16 \times 16$	$47 \times 47$	1.48E-6	5.03	3.52E-6	5.00					
	test on (7.2)									
$2 \times 2$	$5 \times 5$	1.18E0	-	2.61E0	-					
$4 \times 4$	$11 \times 11$	6.08E-2	4.28	1.45E-1	4.17					
$8 \times 8$	$23 \times 23$	2.87E-3	4.40	7.10E-3	4.35					
$16 \times 16$	$47 \times 47$	9.82E-5	4.87	2.41E-4	4.88					
$32 \times 32$	$95 \times 95$	3.12E-6	4.97	7.60E-6	4.99					

0	0	0
n	- <	~
U	U	O.

# REFERENCES

639	[1]	E. BOHL AND J. LORENZ, Inverse monotonicity and difference schemes of higher order. a	
640		summary for two-point boundary value problems, acquationes mathematicae, 19 (1979),	
641		pp. 1–36.	
642	[2]	J. BRAMBLE AND B. HUBBARD, On the formulation of finite difference analogues of the Dirichlet	
643		problem for Poisson's equation. Numerische Mathematik, 4 (1962), pp. 313–327.	
644	[3]	J BRAMBLE AND B. HUBBARD. On a finite difference analogue of an elliptic boundary problem	
645	[0]	which is neither diagonally dominant nor of non-negative type Journal of Mathematics	
646		and Divise 42 (1064) pp 117 122	
647	[4]	and i hysics, 45 (1904), pp. 117-152.	
047	[4]	5. II. BRAMBLE, FOULD-older finite afference analogues of the Directice protecting for Forson's	
048		equation in three and four aimensions, Mathematics of Computation, 17 (1963), pp. 217-	
649	[=1		
650	[5]	J. H. BRAMBLE AND B. E. HUBBARD, New monotone type approximations for elliptic problems,	
651		Mathematics of Computation, 18 (1964), pp. 349–367.	
652	[6]	C. CHEN, Structure theory of superconvergence of finite elements (In Chinese), Hunan Science	
653		and Technology Press, Changsha, 2001.	
654	[7]	I. CHRISTIE AND C. HALL, The maximum principle for bilinear elements, International Journal	
655		for Numerical Methods in Engineering, 20 (1984), pp. 549–553.	
656	[8]	P. G. CIARLET, Discrete maximum principle for finite-difference operators, Aequationes math-	
657		ematicae, 4 (1970), pp. 338–352.	
658	[9]	P. G. CIARLET, The Finite Element Method for Elliptic Problems, Society for Industrial and	
659		Applied Mathematics, 2002.	
660	[10]	L. COLLATZ, The numerical treatment of differential equations, Springer-Verlag, Berlin, 1960.	
661	[11]	B FORNBERG AND N FLYER. A mimer on radial basis functions with amhications to the	
662	[++]	aeosciences SIAM 2015	
663	[12]	L Fox Some improvements in the use of relaxation methods for the solution of ordinary	
664	[12]	and martial differential anations. Providence of the Royal Society of London Society	
665		Mathematical and Division Sciences 100 (1047), pp. 21–20.	
666	[19]	W HOW AND H D. MUTTER MANN. Some new order on the discrete maximum principle for finite	
667	[13]	W. HOHN AND IL. D. MILLELMANN, Some remains on the discrete matrixmin-principle for fittle	
007	[1.4]	elements of nighter order, Computing, 27 (1961), pp. 145–154.	
000	[14]	1. HUANG AND J. AU, Superconvergence of quatratic finite elements on mular structurea grias,	
669	[ 1 2 ]	Mathematics of computation, (7 (2008), pp. 1253–1268.	
670	[15]	V. I. KRYLOV AND L. V. KANTOROVITCH, Approximate methods of higher analysis, P. Noord-	
671	[4.0]	hoff, 1958.	
672	[16]	S. K. LELE, Compact finite difference schemes with spectral-like resolution, Journal of compu-	
673		tational physics, 103 (1992), pp. 16–42.	
674	[17]	H. LI, S. XIE, AND X. ZHANG, A high order accurate bound-preserving compact finite difference	
675		scheme for scalar convection diffusion equations, SIAM Journal on Numerical Analysis,	
676		56 (2018), pp. 3308–3345.	
677	[18]	H. LI AND X. ZHANG, On the monotonicity and discrete maximum principle of the finite dif-	
678		ference implementation of $C^0$ - $Q^2$ finite element method, Numerische Mathematik, (2020),	
679		pp. 1–36.	
680	[19]	H. LI AND X. ZHANG, Superconvergence of high order finite difference schemes based on varia-	
681		tional formulation for elliptic equations, Journal of Scientific Computing, 82 (2020), p. 36.	
682	[20]	J. LORENZ, Zur inversionotonie diskreter probleme, Numerische Mathematik, 27 (1977),	
683		pp. 227–238.	
684	[21]	R. J. PLEMMONS, M-matrix characterizations, I—-nonsingular M-matrices, Linear Algebra and	
685		its Applications, 18 (1977), pp. 175–188.	
686	[22]	T. VELCHODSKY AND P. SOLÍN. Discrete maximum principle for higher-order finite elements	
687	[]	in 1D. Mathematics of Computation, 76 (2007) pp. 1833–1846	
688	[23]	L. WAHLBIN Superconvergence in Galerkin finite element methods. Springer 2006	
680	[20] [94]	L. WAILERAN, Depresentergence in Generation Induce contents methods, Springer, 2000.	
600	[24]	in Numerische Behandlung von Differentieleleichungen Springer 1075 pp. 221-225	
601	[05]	I YU AND I ZWATANOV A monotone finite element scheme for convection difference	
091	[20]	J. AU AND L. ZIKATANOV, A monotone junite element scheme for confection-alifusion equa-	
092		nons, Mathematics of Computation of the American Mathematical Society, 68 (1999),	
093		pp. 1429–1440.	