TWO-LEVEL FOURIER ANALYSIS OF MULTIGRID FOR HIGHER-ORDER FINITE-ELEMENT METHODS*

YUNHUI HE† AND SCOTT MACLACHLAN†

Abstract. In this paper, we employ local Fourier analysis (LFA) to analyze the convergence properties of multigrid methods for higher-order finite-element approximations to the Laplacian problem. We find that the LFA smoothing factor fails to accurately predict the observed multigrid performance. This failure of the LFA smoothing factor is explained, and we propose a modification to the analysis that yields a reasonable prediction to help choose the correct damping parameters for relaxation. Finally, we present two-grid and multigrid experiments, and the corrected parameter choice is shown to yield a significant improvement in the resulting two-grid and multigrid convergence factors.

Key words. Finite-element method, higher-order elements, Jacobi iteration, local Fourier analysis, multigrid

AMS subject classifications. 65M55, 65N30, 65Txx

1. Introduction.

Multigrid methods [2, 7, 19, 23, 24] are very popular to solve the linear systems that arise from the discretization of many PDEs. The choice of the multigrid components, such as grid transfer operators and the relaxation scheme, has a great influence on the performance of these algorithms. In this paper, we focus on the Laplace problem,

$$\begin{cases} -\Delta u(x) = f, & x \in \Omega, \\ u(x) = g, & x \in \partial\Omega, \end{cases} \tag{1.1}$$

discretized using higher-order finite elements. In the literature, there are many efficient multigrid methods for problem (1.1), see [9, 21]. It is worthwhile, however, to understand how these methods work efficiently. LFA [21, 24] has proven a good tool for theoretical investigation and multigrid method design, including for the curl-curl equation [1, 15], parabolic partial differential equations [6, 22], the Stokes equations [10, 14, 15], and the Poisson equation [8, 17, 21].

Recently, some studies have reported that LFA fails to accurately predict some multigrid results, see [5, 6]. In [6], LFA does not offer its usual predictivity of the convergence behavior of the space-time diffusion equation and its generalizations. However, in [5], the authors develop new tools to make up for the failure of standard LFA to provide insight into the asymptotic convergence behaviour of multigrid methods for these problem. In [15], an LFA is presented for general problems, focusing on analyzing the complementarity between relaxation and coarse-grid correction (CGC) within multigrid solvers for systems of PDEs with finite-element discretizations. In that paper, the smoothing factor of LFA overestimates the two-grid convergence factor for the Taylor-Hood ($Q_2 - Q_1$) discretization of the Stokes equations. However, no further explanation is given. We show here that the failure might be related to the Q_2 approximation used for the velocity unknowns.

To our knowledge, the vast majority of existing LFA for the Poisson problem focuses on discretization using finite differences or linear finite elements [19, 21, 24].

*Submitted to the editors DATE.

Funding: The work of S.M. was partially funded by an NSERC Discovery Grant.

†Department of Mathematics and Statistics, Memorial University of Newfoundland, St. Johns, NL A1C 5S7, Canada (yunhui.he@mun.ca, smaclachlan@mun.ca).

In contrast, [8] studies the convergence of a multigrid method for the solution of a linear second-order elliptic equation by discontinuous Galerkin methods. In [17], the cell-centered finite-difference discretization on triangular grids is considered. A variant of LFA is applied to discretization matrices arising from Galerkin B-spline isogeometric analysis in [4], focusing on 2-level analysis in place of classical smoothing analysis. Here, we focus on standard higher-order finite-element discretizations of Poisson's equation with weighted Jacobi relaxation, and use LFA to understand performance. In contrast to the cases of standard finite-difference or (bi)linear finite-element discretizations, we will see that the LFA smoothing factor does not offer a good prediction of performance in the higher-order case.

In the literature, there are many studies about higher-order methods for different types of PDEs. The spectral element method for second-order problems was studied both numerically and theoretically in [16, 18], showing good smoothing properties of simple Jacobi relaxation for the Laplace problem. The impact of different higher-order finite-element discretizations for the Laplace problem on multigrid convergence, with Richardson and Jacobi relaxation, was considered in [13]. Comparison of different multigrid methods for higher-order finite-element discretizations, either as direct solvers or preconditioners, was reported in [20]. There, the convergence behaviour was seen to strongly depend on the polynomial order when multigrid is used as a preconditioner, but not for multigrid as a solver. Other studies of higher-order finite-element methods and multigrid include those for nonlinear problems [3] and the incompressible Navier-Stokes equations [12, 11].

Supporting numerical results demonstrate some key conclusions of our analysis. First, there is a notable gap between the classical LFA smoothing factor and the two-grid convergence factor for these elements. The standard LFA assumption of an "ideal" coarse-grid correction operator, which annihilates the low-frequency error components and leaves the high-frequency components unchanged is not true for higher-order finite-element discretizations, where our results show that the CGC reduces some high-frequency error quickly. Furthermore, minimizing the classical smoothing factor does not minimize the corresponding convergence factor.

The outline of the paper is as follows. In Section 2, we recall the standard definitions of LFA. In Section 3, we analyse the weighted Jacobi relaxation scheme for the Q_2 finite-element approximation in one dimension (1D) and show how to obtain optimal parameters to minimize the convergence factor. We extend this analysis to higher-order finite-elements in Section 4. In Section 5, two-grid LFA is presented for biquadratic Lagrangian elements in two dimensions (2D), and we discuss the optimal parameter choice. Conclusions are presented in Section 6.

2. Definitions and notations. In order to describe LFA for finite-element methods, we first introduce some terminology. More details can be found, for example, in [21]. We first consider one-dimensional infinite uniform grids, G_h. Let L_h be a scalar Toeplitz operator acting on G_h

$$L_h \triangleq [s_\kappa]_h \ (\kappa \in V);\ L_h w_h(x) = \sum_{\kappa \in V} s_\kappa w_h(x + \kappa h), \tag{2.1}$$

with constant coefficients $s_\kappa \in \mathbb{R}$ (or $\mathbb{C}$), where $w_h(x)$ is a function in $l^2(G_h)$. Here, V is taken to be a finite index set of integers, $V \subset \mathbb{Z}$. Note that since L_h is Toeplitz, it is diagonalized by the standard Fourier modes $\psi(\theta, x) = e^{\iota\theta \cdot x/h}$, where $\iota^2 = -1$.

DEFINITION 2.1. *If for all grid functions* $\psi(\theta, x)$,

$$L_h\psi(\theta, x) = \widetilde{L}_h(\theta)\psi(\theta, x),$$

we call $\widetilde{L}_h(\theta) = \sum_{\kappa\in V} s_\kappa e^{\iota\theta\kappa}$ *the symbol of* L_h.

Here, we consider multigrid methods for finite-element discretizations with standard geometric grid coarsening; that is, we construct a sequence of coarse grids by doubling the mesh size in each spatial direction. High and low frequencies for standard coarsening are given by

$$\theta \in T^{\text{low}} = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right), \theta \in T^{\text{high}} = \left[-\frac{\pi}{2}, \frac{3\pi}{2}\right) \setminus \left[-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

The error-propagation operator for a relaxation scheme, represented similarly by a Toeplitz operator M_h, applied to a finite-element approximation is

$$\mathcal{S}_h(\omega, \theta) = I - \omega M_h^{-1} L_h,$$

where ω is an overall weighting factor.

DEFINITION 2.2. *The error-propagation symbol,* $\widetilde{\mathcal{S}}_h(\theta)$, *for smoother* $\mathcal{S}_h$ *on the infinite grid* G_h *satisfies*

$$\mathcal{S}_h\psi(\theta, x) = \widetilde{\mathcal{S}}_h\psi(\theta, x),\ \theta \in \big[-\frac{\pi}{2}, \frac{3\pi}{2}\big),$$

for all $\psi(\theta, x)$, *and the corresponding smoothing factor for* $\mathcal{S}_h$ *is given by*

$$\mu_{\text{loc}} := \mu_{\text{loc}}(\mathcal{S}_h) = \max_{\theta\in T^{\text{high}}} \big\{\big|\widetilde{\mathcal{S}}_h(\theta)\big|\ \big\}. \tag{2.2}$$

DEFINITION 2.3. *Because the smoothing factor is a function of some parameters, let* $\mathbf{D}$ *be the set of allowable parameters and define the optimal smoothing factor over* $\mathbf{D}$ *as*

$$\mu_{\text{opt}} = \min_{\mathbf{D}} \mu_{\text{loc}}.$$

In what follows, we consider $(q \times q)$ linear systems of operators, which read

$$\mathbf{L}_h = \begin{pmatrix} L_h^{1,1} & \cdots & L_h^{1,q} \\ \vdots & \cdots & \vdots \\ L_h^{q,1} & \cdots & L_h^{q,q} \end{pmatrix}.$$

The $L_h^{i,j}(i, j = 1, 2, \ldots, q)$ are scalar Toeplitz operators. Each entry in $\widetilde{\mathbf{L}}_h$ is computed as the (scalar) symbol of the corresponding block of $L_h^{i,j}$, following Definition 2.1. For simplicity, we reuse the notation in (2.2) for the case of block symbols as described in the following.

On a collocated mesh, all blocks in $\mathbf{L}_h$ are diagonalized by the same transformation. However, in our setting, we consider $G_h = G_{h,N} \bigcup G_{h,C}$, for quadratic Lagrangian elements, with

$$G_{h,N} = \big\{x_{k,N} := kh, k \in \mathbb{Z}\big\}, \text{and } G_{h,C} = \big\{x_{k,C} := kh + h/2, k \in \mathbb{Z}\big\}. \tag{2.3}$$

Here G_h contains two types of meshpoints, the nodes of the mesh and the cell centres. The coarse grid, G_{2h}, is defined similarly. Each block $L_h^{i,j}$ in $\mathbf{L}_h$ for $i,j=1,2$ is defined as in (2.1), with V taken to be either a finite index set of integer (V_N) or half-integer (V_C) values, with $V_N \subset \mathbb{Z}$ and $V_C \subset \left\{z+\frac{1}{2} | z \in \mathbb{Z}\right\}$. The operators discussed later are naturally treated as block operators, and the Fourier representation of each block can be calculated based on Definition 2.1, with Fourier bases adapted to account for the staggering of the mesh points. In Definition 2.2, the symbol $\widetilde{\mathcal{S}}_h(\theta)$ will be a matrix, thus, $\left|\widetilde{\mathcal{S}}_h(\theta)\right|$ is replaced by $\left|\lambda(\widetilde{\mathcal{S}}_h(\theta))\right|$, the absolute value of the eigenvalues of $\widetilde{\mathcal{S}}_h(\theta)$, in (2.2).

The resulting Fourier functions are $\varphi(\theta,x_k) \in \operatorname{span}\left\{\varphi_N(\theta,x_k), \varphi_C(\theta,x_k)\right\}$ on G_h, in which

$$\varphi_N(\theta,x_k)= \begin{pmatrix} e^{\iota\theta\cdot x_{k,N}/h} & 0 \end{pmatrix}^T,\ \varphi_C(\theta,x_k) = \begin{pmatrix} 0 & e^{\iota\theta\cdot x_{k,C}/h} \end{pmatrix}^T,$$

where T denotes the (non-conjugate) transpose of the row vectors. Because $\varphi(\theta,x_k)$ is periodic in θ with period 2π, we consider the domain $\theta \in \left[-\frac{\pi}{2}, \frac{3\pi}{2}\right)$.

3. LFA for quadratics in 1D. Here, we consider the discretization of problem (1.1) in 1D, using quadratic (Q_2) finite elements, and nodal basis functions defined at the nodes of the mesh and cell centres (but the analysis could be modified for other bases), and will focus on weighted Jacobi relaxation.

3.1. Quadratic Lagrangian Elements. For these quadratic Lagrangian elements, the elementary contributions to the stiffness and mass matrices as 3×3 symmetric matrices are

$$EK = \frac{1}{3h}\begin{pmatrix} 7 & -8 & 1 \\ -8 & 16 & -8 \\ 1 & -8 & 7 \end{pmatrix},\ EM = \frac{h}{30}\begin{pmatrix} 4 & 2 & -1 \\ 2 & 16 & 2 \\ -1 & 2 & 4 \end{pmatrix},$$

respectively. We can decompose the resulting stencils into connections among and between the degrees of freedom (DOFs) located at the nodes of the mesh and those located at cell centres. The node-to-node connections yield the stencils

$$\frac{1}{3h}\begin{bmatrix} 1 & 14 & 1 \end{bmatrix} \text{ and } \frac{h}{30}\begin{bmatrix} -1 & 8 & -1 \end{bmatrix}.$$

The node-to-centre stencils are given by

$$\frac{1}{3h}\begin{bmatrix} -8 & \star & -8 \end{bmatrix} \text{ and } \frac{h}{30}\begin{bmatrix} 2 & \star & 2 \end{bmatrix},$$

with transposed connections between centres and nodes, where $\star$ stands for the degree-of-freedom position in the off-diagonal blocks. The centre-to-centre stencils are diagonal,

$$\frac{1}{3h}\begin{bmatrix} 16 \end{bmatrix} \text{ and } \frac{h}{30}\begin{bmatrix} 16 \end{bmatrix}.$$

On the infinite grid G_h, each of these stencils defines a Toeplitz operator on $\ell_2(G_{h,*})$ and, so, the block systems can be block diagonalized by considering the invariant subspace given by linear combinations of $\varphi_N(\theta,x)$ and $\varphi_C(\theta,x)$. The resulting block symbols of the stiffness and mass operators are

$$\widetilde{A}_h(\theta) = \frac{1}{3h}\begin{pmatrix} 14+2\cos\theta & -16\cos\frac{\theta}{2} \\ -16\cos\frac{\theta}{2} & 16 \end{pmatrix},\ \widetilde{B}_h(\theta) = \frac{h}{30}\begin{pmatrix} 8-2\cos\theta & 4\cos\frac{\theta}{2} \\ 4\cos\frac{\theta}{2} & 16 \end{pmatrix}, \tag{3.1}$$

respectively. The error-propagation symbol of weighted Jacobi relaxation is given by

$$\widetilde{\mathcal{S}}_h(\theta) = I - \omega \widetilde{M}_h^{-1}(\theta)\widetilde{A}_h(\theta), \tag{3.2}$$

where $\widetilde{M}_h(\theta)$ is the symbol of the diagonal operator,

$$M_h = \frac{1}{3h}\begin{pmatrix} 14I & 0 \\ 0 & 16I \end{pmatrix}. \tag{3.3}$$

Using (3.1) and (3.3), we plot the distribution of eigenvalues of $\widetilde{M}_h^{-1}(\theta)\widetilde{A}_h(\theta)$, at the left of Figure 1. Note that as a block symbol, $\widetilde{M}_h^{-1}(\theta)\widetilde{A}_h(\theta)$ has 2 eigenvalues, each of which can be seen to be a continuous function of θ/π.

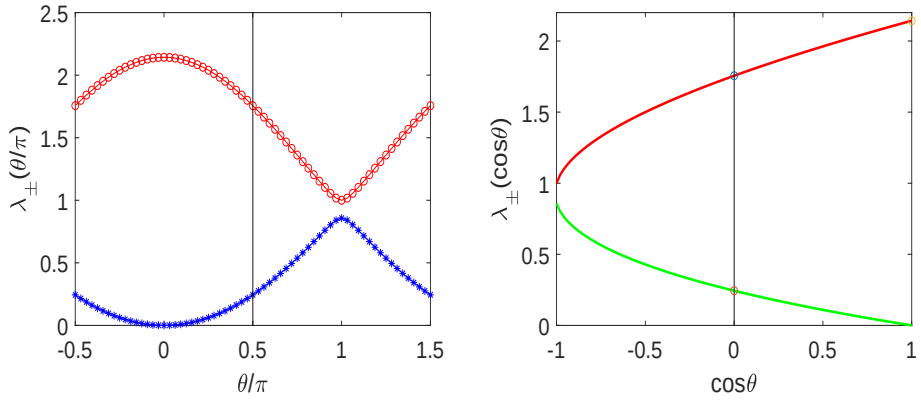

FIG. 3.1. *At left, the distribution of the two eigenvalues of $\widetilde{M}_h^{-1}(\theta)\widetilde{A}_h(\theta)$ as a function of θ/π. At right, the distribution of the two eigenvalues of $\widetilde{M}_h^{-1}(\theta)\widetilde{A}_h(\theta)$, as a function of $\cos\theta$.*

To derive an analytical expression for the eigenvalues of $\widetilde{M}_h^{-1}(\theta)\widetilde{A}_h(\theta)$, we note that the determinant of $\widetilde{M}_h^{-1}(\theta)\widetilde{A}_h(\theta) - \lambda I$ is

$$(\lambda - 1)(\lambda - 1 - \frac{\cos\theta}{7}) - \frac{4}{7}(1 + \cos\theta).$$

Let λ_+ and λ_- be the eigenvalues of $\widetilde{M}_h^{-1}(\theta)\widetilde{A}_h(\theta)$; from above, we have

$$\lambda_\pm = \frac{14 + \cos\theta \pm \sqrt{\cos^2(\theta) + 112\cos\theta + 112}}{14}.$$

Taking $x = \cos\theta$, then we can write

$$\lambda_+(x) = \frac{14 + x + \sqrt{x^2 + 112x + 112}}{14},\ \lambda_-(x) = \frac{14 + x - \sqrt{x^2 + 112x + 112}}{14}.$$

It is easy to check that

$$\lambda_+(x)_{\max} = \lambda_+(1) = \frac{15}{7},\ \lambda_+(x)_{\min} = \lambda_+(-1) = 1,$$
$$\lambda_-(x)_{\max} = \lambda_+(-1) = \frac{6}{7},\ \lambda_-(x)_{\min} = \lambda_-(1) = 0.$$

We plot $\lambda_+(x), \lambda_-(x)$ at the right of Figure 1.

Throughout this paper, we denote $\lambda_{\max,\mathrm{H}}$ and $\lambda_{\min,\mathrm{H}}$ as the biggest and smallest eigenvalues over only the high frequency range, respectively. Since $\lambda_-(x) < \lambda_+(x)$, for high frequencies ($x \in [-1, 0]$), we have

$$\lambda_{\max,\mathrm{H}} = \lambda_+(0) = \frac{7+2\sqrt{7}}{7}, \quad \lambda_{\min,\mathrm{H}} = \lambda_-(0) = \frac{7-2\sqrt{7}}{7}.$$

Thus, the classical optimal choice of ω that minimizes the resulting smoothing factor for relaxation scheme (3.2) is given by

$$\omega^* = \frac{2}{\lambda_{\min,\mathrm{H}} + \lambda_{\max,\mathrm{H}}} = 1, \tag{3.4}$$

and the corresponding smoothing factor is

$$\mu_2^* = \min_{\omega} \max_{\theta \in T^{\mathrm{high}}} \left|\lambda(\widetilde{\mathcal{S}}_h(\omega, \theta))\right| = \frac{2\sqrt{7}}{7} \approx 0.760.$$

Note, however, that this choice of ω^* leads to a diverging relaxation scheme, as $|1 - \omega^*\lambda_+(1)| > 1$. While this might be acceptable assuming ideal CGC, it is worrisome from the perspective of robustness of the resulting multilevel algorithm. Thus, we consider another relaxation weight,

$$\omega^{**} = \frac{2}{\lambda^*_{\max} + \lambda_{\min,H}} = \frac{14}{22-2\sqrt{7}} \approx 0.838, \tag{3.5}$$

where $\lambda^*_{\max}$ is the biggest of all eigenvalues; that is $\lambda^*_{\max} = \lambda_+(1) = \frac{15}{7}$. For this choice, the corresponding smoothing factor is

$$\mu_2^{**} = \max_{\theta \in T^{\mathrm{high}}} \left|\lambda(\widetilde{\mathcal{S}}_h(\omega^{**}, \theta))\right| = \frac{4+\sqrt{7}}{11-\sqrt{7}} \approx 0.795.$$

To understand and compare these choices, we now consider two-grid LFA and measured two-grid performance. We use the notation $TG(\nu_1, \nu_2)$ and $V(\nu_1, \nu_2)$ to indicate the cycle type and the number of pre- and postsmoothing steps employed. Here, we use the defects $d_h^{(k)} (k = 1, 2, \cdots$, with $d_h^{(k)} = b - A_h x_h^{(k)})$ to experimentally measure the convergence factor as $\hat{\rho}_h^{(k)} = \sqrt[k]{\frac{\|d_h^{(k)}\|_2}{\|d_h^{(0)}\|_2}}$ (see [21]), with $k = 100$. We consider the homogeneous problem, $A_h x_h = b = 0$, with discrete solution $x_h \equiv 0$, and start with a random initial guess, $x_h^{(0)}$, to test the multigrid convergence factor. The coarsest grid is a mesh with 4 elements. Rediscretization is used to define the coarse-grid operator (CGO). For comparison, we present the LFA-predicted convergence factors, ρ_h, for two-grid cycles with ν_1 prerelaxation and ν_2 postrelaxation steps (see (3.14)). We consider periodic boundary conditions.

In Table 1, we use ω^* as the weight. Note that the LFA convergence factor is larger than the smoothing factor. As noted earlier, while we see convergence for $\nu_1 + \nu_2 < 3$, we see divergence when $\nu_1 + \nu_2 = 3, 4$ for the two-grid method. Furthermore, even though the smoothing factor fails to predict the convergence factor, we see that the measured convergence factor matches well with the LFA-predicted two-grid convergence factor. For $\omega = \omega^{**}$, Table 2 shows a good improvement in the convergence factor compared with the choice of ω^*. We again see a good agreement between the

measured convergence factor and the LFA-predicted convergence factor, but now the two-grid convergence factor is smaller than the smoothing factor, in contrast to the case of ω^*. Moreover, while the smoothing factor for the choice of ω^{**} is larger than that of ω^*, the two-grid factor is much better.

TABLE 1
Two-grid convergence factors for the Q_2 approximation with ω^* in 1D

Cycle / $\hat{\rho}_h$	$TG(0,1)$	$TG(1,0)$	$TG(1,1)$	$TG(1,2)$	$TG(2,1)$	$TG(2,2)$
$\omega = \omega^* = 1.000, \mu^* = 0.760$						
$\rho_{h=1/128}$	0.821	0.821	0.985	1.118	1.119	1.279
$\hat{\rho}^{(100)}_{h=1/128}$	0.813	0.815	0.974	1.096	1.102	1.255
$\hat{\rho}^{(100)}_{h=1/256}$	0.814	0.814	0.972	1.104	1.100	1.263

TABLE 2
Two-grid convergence factors for the Q_2 approximation with ω^{**} in 1D

Cycle / $\hat{\rho}_h$	$TG(0,1)$	$TG(1,0)$	$TG(1,1)$	$TG(1,2)$	$TG(2,1)$	$TG(2,2)$
$\omega = \omega^{**} = \frac{14}{22-2\sqrt{7}} \approx 0.838, \mu^{**} = 0.796$						
$\rho_{h=1/128}$	0.526	0.526	0.495	0.372	0.372	0.302
$\hat{\rho}^{(100)}_{h=1/128}$	0.522	0.521	0.491	0.365	0.366	0.296
$\hat{\rho}^{(100)}_{h=1/256}$	0.521	0.522	0.491	0.366	0.366	0.298

3.2. Two-grid LFA in 1D. Two natural questions are raised by these results. First, why is the LFA smoothing factor such a bad predictor of performance? Secondly, is ω^{**} the best choice for a weight, in terms of two-grid performance? To answer these questions, we consider two-grid LFA in more details.

DEFINITION 3.1. *The $2h$-harmonics, $\mathcal{F}_{2h}(\theta)$, are given by*

$$\mathcal{F}_{2h}(\theta) = \text{span}\{\varphi_h(\theta^0, x), \varphi_h(\theta^1, x)\},$$

with $\theta = \theta^0 \in T^{\text{low}} := \Theta_{2h}$, and $\theta^\alpha = \theta + \alpha\pi$, where $\alpha = 0, 1$.

To apply LFA to the two-grid operator,

$$\mathcal{M}_h^{\text{TGM}} = S_h^{\nu_2} \mathcal{M}_h^{\text{CGC}} S_h^{\nu_1}, \tag{3.6}$$

we require the representation of the CGC operator,

$$\mathcal{M}_h^{\text{CGC}} = I - P A_{2h}^{-1} R A_h.$$

Inserting the representations of S_h, A_h, A_{2h}, R, P into (3.6), we obtain the Fourier representation of two-grid error-propagation operator as

$$\hat{\mathcal{M}}_h^{\text{TGM}}(\theta) = \hat{S}_h^{\nu_2}(\theta)\big(I - \hat{P}(\theta)(\widetilde{A}_{2h}(2\theta))^{-1}\hat{R}(\theta)\hat{A}_h(\theta)\big)\hat{S}_h^{\nu_1}(\theta),$$

where

$$\hat{A}_h(\theta) = \text{diag}\left\{\widetilde{A}_h(\theta), \widetilde{A}_h(\theta+\pi)\right\}, \hat{S}_h(\theta) = \text{diag}\left\{\widetilde{\mathcal{S}}_h(\theta), \widetilde{\mathcal{S}}_h(\theta+\pi)\right\},$$
$$\hat{P}_h(\theta) = \left(\widetilde{P}_h(\theta); \widetilde{P}_h(\theta+\pi)\right), \hat{R}_h(\theta) = \left(\widetilde{R}_h(\theta), \widetilde{R}_h(\theta+\pi)\right),$$

and

$$\widetilde{A}_{2h}(2\theta) = \frac{1}{6h}\begin{pmatrix} 14+2\cos(2\theta) & -16\cos\theta \\ -16\cos\theta & 16 \end{pmatrix},$$

in which $\mathrm{diag}\{A, B\}$ stands for the block diagonal matrix with diagonal blocks, A and B.

The symbols $\widetilde{A}_h(\theta)$ and $\widetilde{A}_h(\theta+\pi)$ are as given above, while the symbols for relaxation are

$$\widetilde{\mathcal{S}}_h(\theta) = I - \omega\widetilde{M}_h^{-1}(\theta)\widetilde{A}_h(\theta),\ \widetilde{\mathcal{S}}_h(\theta+\pi) = I - \omega\widetilde{M}_h^{-1}(\theta+\pi)\widetilde{A}_h(\theta+\pi).$$

To derive symbols for the grid-transfer operators, we first consider an arbitrary restriction operator characterized by a constant coefficient stencil $R \triangleq [r_\kappa]_h^{2h}$. Then, an infinite grid function $w_h : G_h \to \mathbb{R}\,(\text{or } \mathbb{C})$ is transferred to the coarse grid, G_{2h}, in the following way:

$$(Rw_h)(x) = \sum_{\kappa\in V} r_\kappa w_h(x+\kappa h)\ (x \in G_{2h}).$$

In our case, we have two types of grid points on the fine and coarse grids, so the restriction operator can also be decomposed based on the partitioning of DOFs associated with nodes of the mesh and cell centres.

Let $\varphi_h(\theta^\alpha, x) = e^{\iota\theta^\alpha x/h}$. We have the following equality

$$\varphi_h(\theta^\alpha, x) = e^{\iota\alpha\pi x/h}\varphi_{2h}(2\theta^0, x),\ \text{for all } x \in G_{2h}. \tag{3.7}$$

Note that $\varphi_h(\theta^\alpha, x)$ coincides on $G_{2h,N}$ with the respective grid function $\varphi_{2h}(2\theta^0, x)$, since $e^{\iota\alpha\pi x/h} \equiv 1$ in (3.7), when $x = 2jh$ for $j \in \mathbb{Z}$. However, $e^{\iota\alpha\pi x/h} = (-1)^\alpha$ when $x = 2(j+\frac{1}{2})h$ coincides with a point in $G_{2h,C}$.

Using this for $x \in G_{2h}$, we have

$$(R\varphi_h)(\theta^\alpha, \cdot)(x) = \sum_{\kappa\in V} r_\kappa e^{\iota(x+\kappa h)\theta^\alpha/h} = \sum_{\kappa\in V} r_\kappa e^{\iota\kappa\theta^\alpha} e^{\iota\alpha\pi x/h}\varphi_{2h}(2\theta^0, x).$$

Definition 3.2. *We call* $\widetilde{R}(\theta^\alpha) = \sum_{\kappa\in V} r_\kappa e^{\iota\kappa\theta^\alpha} e^{\iota\alpha\pi x/h} := \sum_{\kappa\in V} \widetilde{r}_\kappa$ *the restriction symbol of* R.

Remark 3.3. If the restriction operator is defined on a collocated mesh, we have only $G_{2h,N}$, and $e^{\iota\alpha\pi x/h} \equiv 1$ in Definition 3.2, which coincides with the definition of the classical restriction symbol [24, Section 6.2.3].

We consider biquadratic interpolation, and the corresponding adjoint operator for the restriction of the corrections. In stencil notation, the restriction operators are given by

$$R_N \triangleq [(r_N)_\kappa]_h = \begin{bmatrix} 0 & -\frac{1}{8} & 0 & \frac{3}{8} & 1(\star) & \frac{3}{8} & 0 & -\frac{1}{8} & 0 \end{bmatrix}_h, \tag{3.8}$$

and

$$R_C \triangleq [(r_C)_\kappa]_h = \begin{bmatrix} 0 & \frac{3}{4} & 1(\star) & \frac{3}{4} & 0 \end{bmatrix}_h, \tag{3.9}$$

where N, C stand for the node and centre points, respectively, and the $\star$ denotes the position (on the coarse grid) at which the discrete operator is applied. Note that these stencils include contributions from both find-grid nodes and centers to the coarse-grid quantities. We illustrate these in Figure 2.

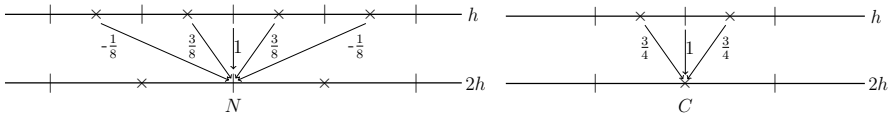

FIG. 3.2. *At left, R_N-restriction operator. At right, R_C-restriction operator.*

As with the fine-grid matrix, both R_N and R_C require values from nodes and centres on the fine grid. We decompose R_N as $[R_N(N), R_N(C)]$ and R_C as $[R_C(N), R_C(C)]$ defined in the following

$$R_N(N) = [1],\ R_N(C) = [-\frac{1}{8} \quad \frac{3}{8} \ \star\ \frac{3}{8} \quad -\frac{1}{8}], \tag{3.10}$$

$$R_C(N) = [1],\ R_C(C) = [\frac{3}{4} \ \star\ \frac{3}{4}], \tag{3.11}$$

then apply Definition 3.2 to each piece separately to obtain the symbol of the restriction operator.

THEOREM 3.4. *Define R as in (3.8) and (3.9). Then the Fourier representation of R is given by the (2×4)-matrix*

$$\begin{aligned}\hat{R}(\theta) &= \begin{pmatrix}\widetilde{R}(\theta^0) & \widetilde{R}(\theta^1)\end{pmatrix} \\ &= \begin{pmatrix} 1 & \frac{3\cos(\frac{\theta}{2})-\cos(\frac{3\theta}{2})}{4} & 1 & \frac{-3\sin(\frac{\theta}{2})-\sin(\frac{3\theta}{2})}{4} \\ 1 & \frac{3\cos(\frac{\theta}{2})}{2} & -1 & \frac{3\sin(\frac{\theta}{2})}{2} \end{pmatrix}.\end{aligned}$$

Proof. Let $x \in G_{2h}$ and consider a fine-grid mode $\varphi(\theta^\alpha, y) = \beta_N \varphi_N(\theta^\alpha, y) + \beta_C \varphi_C(\theta^\alpha, y)$ for $y = x + \kappa h \in G_h$. Clearly the value of $[R\varphi(\theta^\alpha), \cdot](x)$ depends on whether x is a node on the coarse grid (and (3.8) is used) or x is a cell centre on the coarse grid (and (3.9) is used). From (3.10) and (3.11), we write the symbol for R in matrix form,

$$\widetilde{R}(\theta^\alpha) = \begin{pmatrix} \widetilde{R}_N(N, \theta^\alpha) & \widetilde{R}_N(C, \theta^\alpha) \\ \widetilde{R}_C(N, \theta^\alpha) & \widetilde{R}_C(C, \theta^\alpha) \end{pmatrix}, \tag{3.12}$$

acting on the vector $\begin{pmatrix}\beta_N & \beta_C\end{pmatrix}^T$, where T denotes the (non-conjugate) transpose of the row vectors.

From (3.10), (3.11), and Definition 3.2, we obtain the symbols

$$\begin{aligned}&\widetilde{R}_N(N, \theta^\alpha) = 1,\ \widetilde{R}_N(C, \theta^\alpha) = \frac{3}{4}\cos\big(\frac{\theta^\alpha}{2}\big) - \frac{1}{4}\cos\big(\frac{3\theta^\alpha}{2}\big), \\ &\widetilde{R}_C(N, \theta^\alpha) = (-1)^\alpha,\ \widetilde{R}_C(C, \theta^\alpha) = \frac{3}{2}\cos\big(\frac{\theta^\alpha}{2}\big)(-1)^\alpha.\end{aligned}$$

Concatenating $\hat{R}(\theta) = \begin{pmatrix}\widetilde{R}(\theta^0) & \widetilde{R}(\theta^1)\end{pmatrix}$ gives the symbol in the statement of the theorem. ∎

A similar calculation (see [15]) gives the symbol of biquadratic interpolation as

$$\hat{P}(\theta) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3\cos(\frac{\theta}{2})-\cos(\frac{3\theta}{2})}{8} & \frac{3\cos(\frac{\theta}{2})}{4} \\ \frac{1}{2} & -\frac{1}{2} \\ \frac{-3\sin(\frac{\theta}{2})-\sin(\frac{3\theta}{2})}{8} & \frac{3\sin(\frac{\theta}{2})}{4} \end{pmatrix}, \tag{3.13}$$

satisfying the usual relationship that $\hat{P}(\theta) = \frac{1}{2}(\hat{R}(\theta))^H$, where H denotes the conjugate transpose.

We again use rediscretization for the CGO, which matches the Galerkin CGO. The asymptotic two-grid convergence factor, ρ_{asp}, is defined as

$$\rho_{\text{asp}} = \sup\{\rho(\hat{\mathcal{M}}(\theta)^{\text{TGM}}) : \theta \in \Theta_{2h}\}. \tag{3.14}$$

In what follows, we consider a discrete form of ρ_{asp}, denoted by ρ_h, resulting from sampling ρ_{asp} over only finite set of frequencies. We consider only the case of a single relaxation; that is $\nu_1 + \nu_2 = 1$. Without loss of generality, let $\nu_1 = 1$, giving the two-grid representation as

$$\hat{\mathcal{M}}_h^{\text{TGM}}(\theta) = \big(I - \hat{P}(\theta)(\widetilde{A}_{2h}(2\theta))^{-1}\hat{R}(\theta)\hat{A}_h(\theta)\big)\hat{S}_h(\theta). \tag{3.15}$$

3.3. A lower bound on convergence in 1D. To gain some insight and a lower bound on convergence, we consider now the limiting behavior when $\theta \to 0$. When $\theta = 0$, the two eigenvalues of

$$\widetilde{\mathcal{S}}_h(\theta+\pi) = I - \omega\widetilde{M}_h^{-1}(\theta+\pi)\widetilde{A}_h(\theta+\pi),$$

are $1-\omega, 1-\frac{6}{7}\omega$ and the eigenvector corresponding to $1-\omega$ is $v_1 = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$.

From (3.13), when $\theta = 0$, we have the representation of interpolation

$$\hat{P}(0) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 \end{pmatrix},$$

and vector $\hat{v}_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}^T$ is not in the range of interpolation. Taken together, this tells us $\hat{v}_1$ is an eigenvector of $\hat{\mathcal{M}}_h^{\text{TGM}}(\theta)$ in the limit as $\theta \to 0$, allowing us to establish a lower bound on convergence.

THEOREM 3.5. *For $\hat{\mathcal{M}}_h^{\text{TGM}}(\theta)$ defined as in (3.15),*

$$\text{trace}\big(\lim_{\theta\to 0}\hat{\mathcal{M}}_h^{\text{TGM}}(\theta)\big) = 2 - \frac{79}{28}\omega.$$

Proof. By standard calculation, we have

$$\lim_{\theta\to 0}\hat{\mathcal{M}}_h^{\text{TGM}}(\theta) = \begin{pmatrix} \frac{7-15\omega}{14} & \frac{-7+15\omega}{14} & \frac{-7+6\omega}{28} & 0 \\ -\frac{7-15\omega}{28} & -\frac{-7+15\omega}{28} & -\frac{-7+6\omega}{56} & 0 \\ -\frac{7-15\omega}{14} & -\frac{-7+15\omega}{14} & -\frac{-7+6\omega}{28} & 0 \\ 0 & 0 & 0 & 1-\omega \end{pmatrix}.$$

Thus, $\text{trace}\big(\lim_{\theta\to 0}\hat{\mathcal{M}}_h^{\text{TGM}}(\theta)\big) = \frac{7-15\omega}{14} - \frac{-7+15\omega}{28} - \frac{-7+6\omega}{28} + 1 - \omega = 2 - \frac{79}{28}\omega.$ ∎

Note that $\widetilde{P}(0)$ is full-rank, so there must be two zero eigenvalues of $\lim_{\theta\to 0}\hat{\mathcal{M}}_h^{\mathrm{TGM}}(\theta)$. As $1-\omega$ is also an eigenvalue of $\lim_{\theta\to 0}\hat{\mathcal{M}}_h^{\mathrm{TGM}}(\theta)$, Theorem 3.5 tells us that the other eigenvalue is $2-\frac{79}{28}\omega-(1-\omega)=1-\frac{51}{28}\omega$. In order to minimize the spectral radius of $\lim_{\theta\to 0}\hat{\mathcal{M}}_h^{\mathrm{TGM}}(\theta)$, we have the following result.

LEMMA 3.6.

$$\min_{\omega}\left\{\max\{|\lambda^*|\}:\lambda^*\in\lambda\big(\lim_{\theta\to 0}\hat{\mathcal{M}}_h^{\mathrm{TGM}}(\theta)\big)\right\}=\frac{23}{79}\approx 0.291, \tag{3.16}$$

and only $\omega=\omega^{***}=\frac{56}{79}$ *achieves the minimum.*

Proof. Note that the four eigenvalues of $\lim_{\theta\to 0}\hat{\mathcal{M}}_h^{\mathrm{TGM}}(\theta)$ are $0,0,1-\omega$, and $1-\frac{51}{28}\omega$. Setting $|1-\omega|=|1-\frac{51}{28}\omega|$, gives $\omega=\frac{56}{79}$. ∎

COROLLARY 3.7. *For any* ω*, the optimal two-grid convergence factor for a single relaxation (i.e.,* $\nu_1+\nu_2=1$*) is not less than* $\frac{23}{79}$*, and this factor can be achieved if and only if* $\omega=\omega^{***}$*.*

Corollary 3.7 only tells us that the two-grid convergence factor has a lower bound, but we do not know whether it can be achieved or not. We show this numerically. For the remaining part of this paper, let μ and ρ be the LFA-predicted smoothing and two-grid convergence factors, respectively, computed with $h=\frac{1}{64}$. For ρ, we consider only one step of pre-smoothing (which gives the same results as one step of post-smoothing). We plot the predicted smoothing and convergence factors as a function of ω in 1D. The left of Figure 3 indicates that when the classical smoothing factor achieves its optimal value, the corresponding ω does not minimize the two-grid convergence factor. The choices of ω^* and ω^{**} in (3.4) and (3.5) both are clearly not the best choice. The left of Figure 3 shows that the optimal ω is $\omega^{***}=\frac{56}{79}\approx 0.709$, as proposed in Corollary 3.7. We explore the reasons for this below.

To see that the prediction of Lemma 3.6 is not a coincidence, we plot the two-grid convergence factor and $\max\left\{|1-\omega|,|1-\frac{51}{28}\omega|\right\}$ as a function of ω. Comparing the left and right of Figure 3 indicates that, for all ω, the two-grid convergence factor is given by $\max\left\{|1-\omega|,|1-\frac{51}{28}\omega|\right\}$.

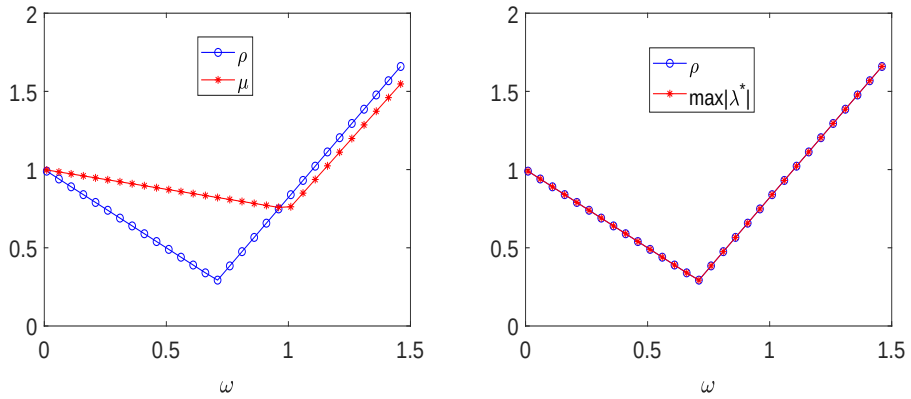

FIG. 3.3. *At left, LFA-predicted two-grid convergence and smoothing factors as a function of* ω*. At right, LFA-predicted two-grid convergence factor and* $\max\{|\lambda^*|\}$ *as a function of* ω *for the* Q_2 *approximation in 1D.*

3.3.1. Two-grid and multigrid performance in 1D. Table 3 confirms that ω^{***} provides the best observed convergence factor, compared with the choices ω^{*} and ω^{**}, shown in Tables 1 and 2. Table 3 also confirms that a single pre- or post-relaxation offers the most cost-effective cycle. Table 4 shows that similar convergence factors are obtained for full V-cycles.

TABLE 3
Two-grid convergence factors for the Q_2 approximation with ω^{***} in 1D

Cycle / $\hat{\rho}_h$	$TG(0,1)$	$TG(1,0)$	$TG(1,1)$	$TG(1,2)$	$TG(2,1)$	$TG(2,2)$
$\omega = \omega^{***} = \frac{56}{79} \approx 0.709, \mu = 0.822$						
$\rho_{h=1/128}$	0.291	0.291	0.249	0.090	0.090	0.064
$\hat{\rho}^{(100)}_{h=1/128}$	0.289	0.290	0.245	0.088	0.088	0.063
$\hat{\rho}^{(100)}_{h=1/256}$	0.289	0.289	0.246	0.088	0.088	0.063

TABLE 4
Multigrid convergence factors for the Q_2 approximation with ω^{***} in 1D

Cycle / $\hat{\rho}_h$	$V(0,1)$	$V(1,0)$	$V(1,1)$	$V(1,2)$	$V(2,1)$	$V(2,2)$
$\omega = \omega^{***} = \frac{56}{79} \approx 0.709, \mu = 0.822$						
$\rho_{h=1/128}$	0.291	0.291	0.249	0.090	0.090	0.064
$\hat{\rho}^{(100)}_{h=1/128}$	0.281	0.282	0.246	0.080	0.081	0.068
$\hat{\rho}^{(100)}_{h=1/256}$	0.284	0.280	0.246	0.083	0.082	0.068

3.4. A modified two-grid analysis. To better understand the failure of classical smoothing analysis for the Q_2 approximation, we first consider why the smoothing factor is a good predictor of performance for the Q_1 approximation. In the Q_1 case, we denote the CGC operator as $\hat{\mathcal{M}}^{\mathrm{CGC}}_{1,h}(\theta)$, and the symbol of the relaxation scheme as $\hat{S}_{1,h}(\theta)$, which are both 2×2 matrices. Here we use linear interpolation for P and $R = P^H$. By standard calculation, we have

$$\hat{\mathcal{M}}^{\mathrm{CGC}}_{1,h}(\theta) = \begin{pmatrix} \sin^2(\frac{\theta}{2}) & \cos^2(\frac{\theta}{2}) \\ \sin^2(\frac{\theta}{2}) & \cos^2(\frac{\theta}{2}) \end{pmatrix}.$$

In the standard LFA smoothing analysis, we assume an "ideal" CGC operator, $\mathcal{Q}_h$, in place of the true CGC, $\hat{\mathcal{M}}^{\mathrm{CGC}}_{1,h}(\theta)$, that annihilates the low-frequency error components and leaves the high-frequency components unchanged, see [21]. A natural choice for $\mathcal{Q}_h$ is as a projection operator,

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

To compute the convergence factor, we replace the CGC operator in (3.14) by $\mathcal{Q}_h$, giving

$$\sup\{\rho(\mathcal{Q}_h \hat{S}_{1,h}(\theta)) : \theta \in \Theta_{2h}\}. \tag{3.17}$$

Remark 3.8. Note that (3.17) is equivalent to form (2.2).

From the form of $\mathcal{Q}_h$ we can consider optimizing the smoothing factor by working only over the high frequencies as in Definition 2.3. In Figure 4, we plot the LFA-predicted two-grid convergence factor (3.14) and the smoothing factor as a function of ω and see that the smoothing factor perfectly captures the LFA-predicted two-grid convergence behavior.

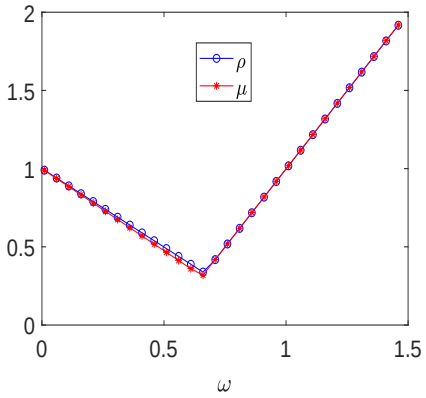

FIG. 3.4. *LFA-predicted two-grid convergence and smoothing factors as a function of ω for the Q_1 approximation in 1D.*

However, as shown above in Subsection 3.1, generalizing $\mathcal{Q}_h$ to

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

does not give a good prediction of the two-grid convergence factor for the Q_2 approximation. Instead, we note that for the Q_1 case,

$$\lim_{\theta \to 0} \hat{\mathcal{M}}_{1,h}^{\mathrm{CGC}}(\theta) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$

and, if we replace $\mathcal{Q}_h$ by this limit, then the eigenvalues of $\mathcal{Q}_h \hat{S}_{1,h}(\theta)$ do not change. This suggests that using $\lim_{\theta \to 0} \hat{\mathcal{M}}_{1,h}^{\mathrm{CGC}}(\theta)$ as the ideal CGC operator may improve the robustness of the smoothing factor. We now extend this approximation for two-grid analysis of the Q_2 approximation.

Define

$$\mathcal{Q}_0 := \lim_{\theta \to 0} \left(I - \hat{P}(\theta)(\widetilde{A}_{2h}(2\theta))^{-1} \hat{R}(\theta) \hat{A}_h(\theta) \right). \tag{3.18}$$

By standard calculation,

$$\mathcal{Q}_0 = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{4} & 0 \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{8} & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

To see how well $\mathcal{Q}_0$ works as an idealized CGC operator when predicting the two-grid convergence factor, let

$$\rho_0 = \rho_0(\omega) = \sup\{\rho(\mathcal{Q}_0\hat{S}_h(\theta)) : \theta \in \Theta_{2h}\}. \tag{3.19}$$

We plot ρ as a function of ω, compared with the LFA-predicted two-grid convergence factor ρ. Figure 5 shows that ρ_0 provides a much better prediction than the classical smoothing factor. Note that for smaller values of ω, ρ_0 slightly overpredicts the convergence factor, as $\mathcal{Q}_0$ captures poorly the true effects of CGC for values of θ near $\pm\frac{\pi}{2}$. We see that the optimal parameter of ρ_0 is very close to the optimal parameter for the two-grid convergence factor, ρ. Whether further improvement is possible is an open question.

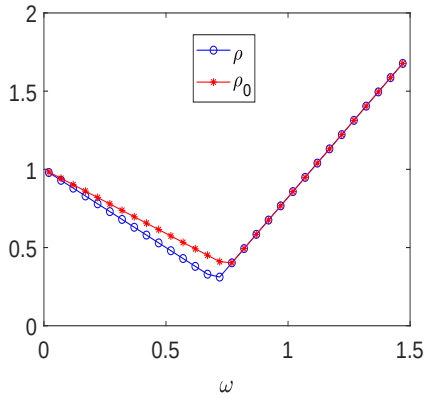

FIG. 3.5. *ρ and ρ_0, as a function of ω for the Q_2 approximation in 1D.*

We now consider a modified two-grid error-propagation operator,

$$\hat{\mathcal{M}}^{\mathrm{MTGM}}(\theta) := \mathcal{Q}_0\hat{S}(\theta), \quad \theta \in \Theta_{2h},$$

which gives a good prediction for the convergence of multigrid for the Q_2 approximation. Now, we consider minimizing the spectral radius of $\hat{\mathcal{M}}^{\mathrm{MTGM}}(\theta)$; that is, to minimize ρ_0.

By standard calculation, we have

$$\hat{S}(\theta) = \begin{pmatrix} 1-\omega(1+\frac{\cos(\theta)}{7}) & \frac{8}{7}\cos(\frac{\theta}{2})\omega & 0 & 0 \\ \cos(\frac{\theta}{2})\omega & 1-\omega & 0 & 0 \\ 0 & 0 & 1-\omega(1-\frac{\cos(\theta)}{7}) & -\frac{8}{7}\sin(\frac{\theta}{2})\omega \\ 0 & 0 & -\sin(\frac{\theta}{2})\omega & 1-\omega \end{pmatrix}.$$

Because $\mathcal{Q}_0$ has rank 2, $\hat{\mathcal{M}}^{\mathrm{MTGM}}(\theta)$ has at most rank 2. By a straightforward calculation (done using a computer algebra system), the four eigenvalues of $\mathcal{Q}_0\hat{S}(\theta)$ are given by

$$\lambda(\theta) = 1 - g_{\pm}(\theta)\omega, \ 0, \ 0,$$

where $g_{\pm}(\theta)$ is

$$\frac{112+44\cos(\frac{\theta}{2})+2\cos(\theta)\pm\sqrt{2\big(1381+44(\cos(\frac{\theta}{2})+\cos(\frac{3\theta}{2}))-412\cos(\theta)+\cos(2\theta)\big)}}{112}.$$

We can check that $g_{\pm}(\theta)$ is an increasing function over $[-\frac{\pi}{2},0]$ and a decreasing function over $[0,\frac{\pi}{2}]$. We plot $g_{\pm}(\theta)$ as a function of θ over $[-\frac{\pi}{2},\frac{\pi}{2}]$ in Figure 6.

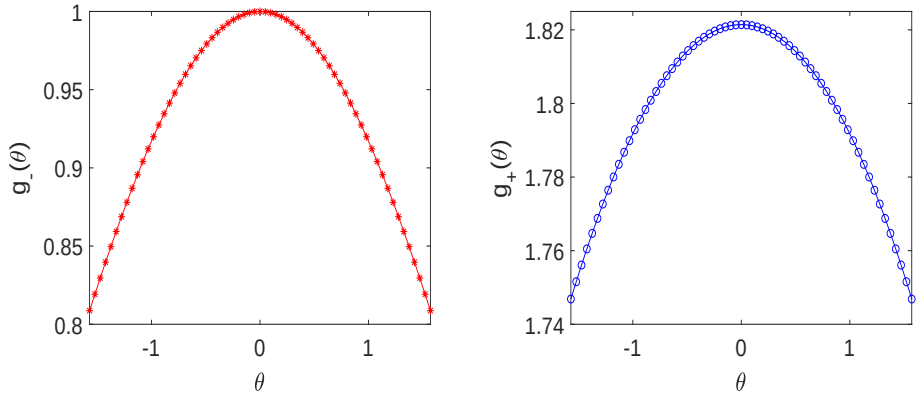

FIG. 3.6. *At left, $g_-(\theta)$ as a function of θ. At right, $g_+(\theta)$ as a function of θ.*

The extreme values of $g_{\pm}(\theta)$ are obtained at $\theta=0$ and $\theta=\pm\frac{\pi}{2}$; that is,

$$\begin{aligned} g_+(0)&=\frac{51}{28},\quad g_-(0)=1,\\ g_+(\pm\frac{\pi}{2})&=\frac{56+11\sqrt{2}+\sqrt{690}}{56}<\frac{51}{28},\\ g_-(\pm\frac{\pi}{2})&=\frac{56+11\sqrt{2}-\sqrt{690}}{56}<1. \end{aligned}$$

Thus,

$$\rho_0=\sup\{\rho(\mathcal{Q}_0\hat{S}_h(\theta)):\theta\in\Theta_{2h}\}=\max\Big\{\big|1-\frac{51}{28}\omega\big|,\big|1-g_-(\pm\frac{\pi}{2})\omega\big|\Big\}.$$

Then, the optimal parameter minimizing ρ_0 is given by

$$\omega_{0,\text{opt}}=\frac{2}{\frac{51}{28}+\frac{56+11\sqrt{2}-\sqrt{690}}{56}}\approx 0.760,$$

and the corresponding predicted smoothing factor is

$$\rho_{0,\text{opt}}=\frac{\frac{51}{28}-\frac{56+11\sqrt{2}-\sqrt{690}}{56}}{\frac{51}{28}+\frac{56+11\sqrt{2}-\sqrt{690}}{56}}\approx 0.385.$$

Recall the optimal parameter and the true two-grid convergence factor are $\omega^{***}=0.709$, $\rho=0.291$, respectively. Compared with the true two-grid convergence, ρ_0 over-predicts the convergence factor based on the mode $\theta=\pm\frac{\pi}{2}$. However, this modified

$\hat{\mathcal{M}}^{\mathrm{MTGM}}(\theta)$ still offers useful information and a reasonable predictor of performance. Whether this "ideal" predictor can be used for other higher-order finite-element approximations will be explored in the following sections.

Remark 3.9. Improved two-grid behavior can be achieved by considering different weights for the DOFs at nodes and those at cell centres for Jacobi relaxation; that is, putting distinct parameters in each diagonal block in the diagonal operator in (3.3). Then, the LFA shown above can be extended to this relaxation scheme to optimize the two-grid convergence factor, resulting in somewhat better convergence.

4. Higher-order finite-element methods. In this section, we consider the finite-element spaces Q_p for $p = 3, 4$ and again examine the relationship between the LFA smoothing and two-grid convergence factors. In order to distinguish the block symbols for different p, we use superscripts in the matrices and block symbols in this section.

4.1. Cubic Lagrangian Elements. For cubic Lagrangian elements (Q_3), using nodal finite-element basis functions defined at the mesh nodes and the 1/3 and 2/3 points of the element, the elementary contributions to the stiffness matrix can be written as

$$EK_h^{(3)} = \frac{1}{40h}\begin{pmatrix} 296 & -189 & 54 & -13 \\ -189 & 432 & -297 & 54 \\ 54 & -297 & 432 & -189 \\ -13 & 54 & -189 & 296 \end{pmatrix}.$$

The corresponding symbol of stiffness operator is

$$\widetilde{A}_h^{(3)}(\theta) = \frac{1}{h}\begin{pmatrix} \frac{148-13\cos\theta}{20} & \frac{54e^{-\frac{2}{3}\iota\theta}-189e^{\frac{1}{3}\iota\theta}}{40} & \frac{54e^{\frac{2}{3}\iota\theta}-189e^{-\frac{1}{3}\iota\theta}}{40} \\ \frac{54e^{\frac{2}{3}\iota\theta}-189e^{-\frac{1}{3}\iota\theta}}{40} & \frac{54}{5} & -\frac{297e^{\frac{1}{3}\iota\theta}}{40} \\ \frac{54e^{-\frac{2}{3}\iota\theta}-189e^{\frac{1}{3}\iota\theta}}{40} & -\frac{297e^{-\frac{1}{3}\iota\theta}}{40} & \frac{54}{5} \end{pmatrix},$$

ordered as mesh nodes, then the 1/3 points and 2/3 points, respectively. The error-propagation symbol of weighted Jacobi relaxation is given by

$$\widetilde{\mathcal{S}}_h^{(3)}(\theta) = I - \omega\big(\widetilde{M}_h^{(3)}(\theta)\big)^{-1}\widetilde{A}_h^{(3)}(\theta), \tag{4.1}$$

where

$$\widetilde{M}_h^{(3)}(\theta) = \frac{1}{h}\begin{pmatrix} \frac{37}{5} & 0 & 0 \\ 0 & \frac{54}{5} & 0 \\ 0 & 0 & \frac{54}{5} \end{pmatrix}.$$

In Figure 7, we plot the eigenvalues of $\big(\widetilde{M}_h^{(3)}(\theta)\big)^{-1}\widetilde{A}_h^{(3)}(\theta)$. Considering the high frequencies, we see $\lambda_{\min,\mathrm{H}} = 0.085$ is obtained at $\theta = \frac{\pi}{2}$, and $\lambda_{\max,\mathrm{H}} = 2.394$ is obtained at $\theta = \pi$.

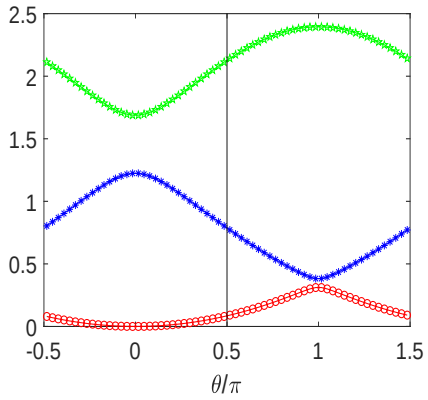

FIG. 4.1. *The distribution of eigenvalues of* $\left(\widetilde{M}_h^{(3)}(\theta)\right)^{-1}\widetilde{A}_h^{(3)}(\theta)$ *as a function of* θ/π

Thus, the classical optimal choice of ω for (4.1) is given by

$$\omega_3^* = \frac{2}{\lambda_{\min,\mathrm{H}} + \lambda_{\max,\mathrm{H}}} = 0.807,$$

and

$$\mu_3^* = \min_{\omega} \max_{\theta\in T^{\mathrm{high}}} \left|\lambda(\widetilde{\mathcal{S}}_h^{(3)}(\omega,\theta))\right| = \frac{\lambda_{\max,\mathrm{H}} - \lambda_{\min,\mathrm{H}}}{\lambda_{\max,\mathrm{H}} + \lambda_{\min,\mathrm{H}}} \approx 0.931.$$

Denote the cubic finite-element interpolation operator as $R^{(3)}$ and the corresponding symbol as $\widetilde{R}^{(3)}$. Similarly to Theorem 3.4, we can write the symbol of restriction, $R^{(3)}(\theta^\alpha)$, as

$$\widetilde{R}^{(3)}(\theta^\alpha) = \begin{pmatrix} 1 - \frac{e^{\iota\theta^\alpha}}{16} - \frac{e^{-\iota\theta^\alpha}}{16} & \frac{5}{16}e^{\frac{1}{3}\iota\theta^\alpha} + \frac{1}{16}e^{-\frac{5}{3}\iota\theta^\alpha} & \frac{5}{16}e^{-\frac{1}{3}\iota\theta^\alpha} + \frac{1}{16}e^{\frac{5}{3}\iota\theta^\alpha} \\ \frac{9}{16}e^{\frac{1}{3}\iota\theta^\alpha}\beta & \frac{15}{16}e^{-\frac{1}{3}\iota\theta^\alpha}\beta & (1 - \frac{5}{16}e^{\iota\theta^\alpha})\beta \\ \frac{9}{16}e^{-\frac{1}{3}\iota\theta^\alpha}\beta^2 & (1 - \frac{5}{16}e^{-\iota\theta^\alpha})\beta^2 & \frac{15}{16}e^{\frac{1}{3}\iota\theta^\alpha}\beta^2 \end{pmatrix},$$

where $\beta = (e^{\frac{2}{3}\iota\pi})^\alpha$. Thus, the symbol of $R^{(3)}$ is the 3×6 matrix

$$\hat{R}^{(3)}(\theta) = \begin{pmatrix} \widetilde{R}^{(3)}(\theta^0) & \widetilde{R}^{(3)}(\theta^1) \end{pmatrix}, \text{ where } \theta = \theta^0 \in \Theta_{2h}.$$

The Fourier representation of $P^{(3)}$ is given by the 6×3 matrix,

$$\hat{P}^{(3)}(\theta) = \frac{1}{2}\left(\hat{R}^{(3)}(\theta)\right)^H.$$

We plot the smoothing factor and LFA-predicted two-grid convergence factor as a function of ω for cubic elements in 1D. Figure 8 indicates that when the smoothing factor achieves its optimal value, the corresponding ω does not minimize the two-grid convergence factor. From Figure 8, note that the optimal convergence factor, ρ, is 0.491 with $\omega = 0.650$, but the corresponding smoothing factor is 0.943, which is larger than the smoothing factor of 0.931 for $\omega_3^* = 0.807$ given above.

As the LFA smoothing factor again fails to predict the convergence factor, we extend the modification above to yield a new prediction based on $\hat{\mathcal{M}}^{\mathrm{MTGM}}(\theta)$, calculating $\mathcal{Q}_0$ again using the limit in (3.18). We plot ρ_0, compared with the true

convergence factor at the right of Figure 8, and see that using $\mathcal{Q}_0$ accurately predicts the true convergence factor, except for a small overestimate for ω less than 0.65, as $\mathcal{Q}_0$ captures poorly the true effects of CGC for values of θ near $\pm\frac{\pi}{2}$. We observe that when $\theta = 0$, ρ_0 underestimates the true two-grid convergence factor. However, the optimal parameter of $\hat{\mathcal{M}}^{\mathrm{MTGM}}(\theta)$ is very close to the true optimal parameter for the two-grid convergence factor.

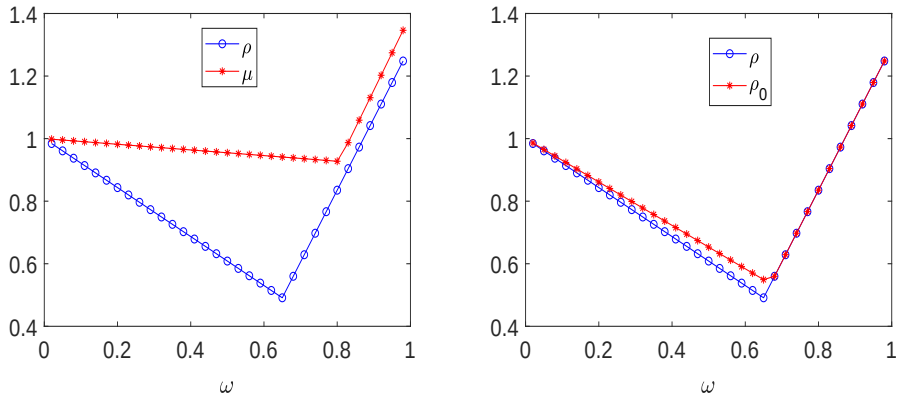

FIG. 4.2. *At left, the LFA-predicted two-grid convergence and smoothing factors as a function of ω. At right, ρ and ρ_0 as a function of ω for the Q_3 approximation in 1D.*

4.2. Quartic Lagrangian Elements. For quartic Lagrangian elements (Q_4), using nodal finite-element basis functions defined at the mesh nodes and the $1/4, 1/2$, and $3/4$ points of the element, the elementary contributions to the stiffness matrix can be written as

$$EK_h^{(4)} = \frac{1}{945h}\begin{pmatrix} 9850 & -6848 & 3048 & -1472 & 347 \\ -6848 & 16640 & -14208 & 5888 & -1472 \\ 3048 & -14208 & 22320 & -14208 & 3048 \\ -1472 & 5888 & -14208 & 16640 & -6848 \\ 347 & -1472 & 3048 & -6848 & 9850 \end{pmatrix},$$

and the corresponding symbol of stiffness operator is

$$\widetilde{A}_h^{(4)}(\theta) = \frac{1}{h}\begin{pmatrix} \frac{9850+347(\eta^{-4}+\eta^4)}{945} & -\frac{6848\eta+1472\eta^{-3}}{945} & \frac{1016\eta^{-2}+1016\eta^2}{315} & -\frac{6848\eta^{-1}+1472\eta^3}{945} \\ -\frac{6848\eta^{-1}+1472\eta^3}{945} & \frac{3328}{189} & -\frac{4736\eta}{315} & \frac{5888\eta^2}{945} \\ \frac{1016\eta^2+1016\eta^{-2}}{315} & -\frac{4736\eta^{-1}}{315} & \frac{496}{21} & -\frac{4736\eta}{315} \\ -\frac{6848\eta+1472\eta^{-3}}{945} & \frac{5888\eta^{-2}}{945} & -\frac{4736\eta^{-1}}{315} & \frac{3328}{189} \end{pmatrix},$$

where $\eta = e^{\frac{\iota\theta}{4}}$, with both ordered as mesh nodes, then the $1/4$, $1/2$, and $3/4$ points of the mesh (followed by the right-hand node in $EK_h^{(4)}$).

The error-propagation symbol of weighted Jacobi relaxation is

$$\widetilde{\mathcal{S}}_h^{(4)}(\theta) = I - \omega\big(\widetilde{M}_h^{(4)}(\theta)\big)^{-1}\widetilde{A}_h^{(4)}(\theta),$$

where

$$\widetilde{M}_h^{(4)}(\theta) = \frac{1}{h}\begin{pmatrix} \frac{1970}{189} & 0 & 0 & 0 \\ 0 & \frac{3328}{189} & 0 & 0 \\ 0 & 0 & \frac{496}{21} & 0 \\ 0 & 0 & 0 & \frac{3328}{189} \end{pmatrix},$$

Using these symbols, we plot the distribution of eigenvalues of $\left(\widetilde{M}_h^{(4)}(\theta)\right)^{-1}\widetilde{A}_h^{(4)}(\theta)$ in Figure 9. From Figure 9, we see that the smallest eigenvalue over the high frequencies, $\lambda_{\min,\mathrm{H}} = 0.036$ is obtained at $\theta = \frac{\pi}{2}$ or $\frac{3\pi}{2}$. Similarly, $\lambda_{\max,\mathrm{H}} = 2.557$ is achieved with $\theta = \frac{\pi}{2}$ or $\frac{3\pi}{2}$.

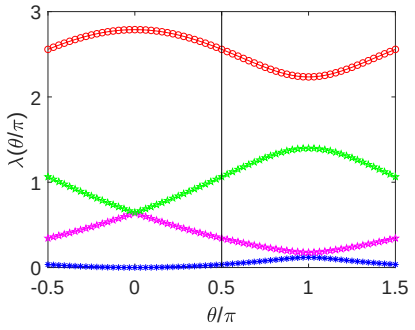

FIG. 4.3. *The distribution of eigenvalues of* $\left(\widetilde{M}_h^{(4)}(\theta)\right)^{-1}\widetilde{A}_h^{(4)}(\theta)$ *as a function of* θ/π

Thus, the optimal ω for the classical smoothing factor and the corresponding smoothing factor are

$$\omega_4^* = \frac{2}{\lambda_{\min,\mathrm{H}} + \lambda_{\max,\mathrm{H}}} = 0.772, \quad \mu_4^* = 0.973, \tag{4.2}$$

respectively.

As in the Q_2 case, the biggest eigenvalue over all frequencies is $\lambda_{\max}^* = 2.789 > \lambda_{\max,\mathrm{H}}$, obtained at $\theta = 0$. We, thus, consider the case of

$$\omega_4^{**} = \frac{2}{\lambda_{\min,\mathrm{H}} + \lambda_{\max}^*} = 0.708.$$

Then, the corresponding smoothing factor is

$$\mu_4^{**} = \max_{\theta\in T^{\mathrm{high}}} \left|\lambda(\widetilde{\mathcal{S}}_h^{(4)}(\omega^{**},\theta))\right| = \frac{\lambda_{\max}^* - \lambda_{\min,\mathrm{H}}}{\lambda_{\max}^* + \lambda_{\min,\mathrm{H}}} = 0.975. \tag{4.3}$$

Denote the quartic interpolation operator as $R^{(4)}$ and the corresponding symbol as $\widetilde{R}^{(4)}$. Similarly to Theorem 3.4, we can write the symbol of restriction, $R^{(4)}(\theta^\alpha)$, as

$$\widetilde{R}^{(4)}(\theta^\alpha) = \begin{pmatrix} 1 & \frac{35}{128}\xi + \frac{3}{128}\xi^5 - \frac{5}{128}\xi^{-7} - \frac{5}{128}\xi^{-3} & 0 & \frac{35}{128}\xi^{-1} + \frac{3}{128}\xi^{-5} - \frac{5}{128}\xi^7 - \frac{5}{128}\xi^3 \\ 0 & (\frac{35}{32}\xi^{-1} - \frac{5}{32}\xi^3)\gamma & \gamma & (\frac{15}{32}\xi + \frac{7}{32}\xi^5)\gamma \\ \gamma^2 & (-\frac{35}{64}\xi^{-3} + \frac{45}{64}\xi)\gamma^2 & 0 & (\frac{45}{64}\xi^{-1} - \frac{35}{64}\xi^3)\gamma^2 \\ 0 & (\frac{7}{32}\xi^{-5} + \frac{15}{32}\xi^{-1})\gamma^3 & \gamma^3 & (-\frac{5}{32}\xi^{-3} + \frac{35}{32}\xi)\gamma^3 \end{pmatrix},$$

where $\xi = e^{\frac{\iota\theta^\alpha}{4}}, \gamma = (e^{\frac{1}{2}\iota\pi})^\alpha$. Thus, the symbol of $R^{(4)}$ is the 4×8 matrix

$$\hat{R}^{(4)}(\theta) = \begin{pmatrix} \widetilde{R}^{(4)}(\theta^0) & \widetilde{R}^{(4)}(\theta^1) \end{pmatrix}, \text{ where } \theta = \theta^0 \in \Theta_{2h}.$$

The Fourier representation of $P^{(4)}$ is given by the 8×4 matrix,

$$\hat{P}^{(4)}(\theta) = \frac{1}{2}\big(\hat{R}^{(4)}(\theta)\big)^H.$$

We plot the LFA smoothing and two-grid convergence factors as a function of ω for this algorithm. At the left of Figure 10, we see that the LFA smoothing factor again fails to predict the two-grid convergence factor, and that the optimal convergence factor ρ is 0.608 with $\omega = 0.640$. The choices of ω in (4.2) and (4.3) both fail.

We present the results of the modified prediction using $\hat{\mathcal{M}}^{\mathrm{MTGM}}(\theta)$ here again defining Q_0 following (3.18). At the right of Figure 10, we compare ρ_0 with ρ, as a function of the relaxation parameter, ω, seeing that ρ_0 matches well with the true convergence, except for a small overestimation for small ω, as $\mathcal{Q}_0$ captures poorly the true effects of CGC for values of θ near $\pm\frac{\pi}{2}$. We also observe that when $\theta = 0$, ρ_0 is exactly the true two-grid convergence factor, which is the same as in the case of the Q_2 approximation.

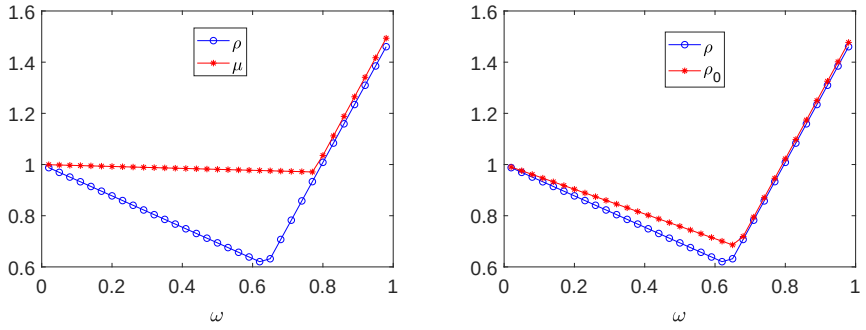

FIG. 4.4. *At right, LFA-predicted two-grid convergence and smoothing factors as a function of ω. At right, ρ and ρ_0 as a function of ω for the Q_4 approximation in 1D.*

5. LFA for the Q_2 approximation in 2D. In this section, we consider LFA for problem (1.1) in 2D, using biquadratic finite elements and the nodal basis functions defined at the mesh nodes, edge midpoints and element centres. We order the DOFs of the Q_2 approximation as nodes first, then midpoints of the edges parallel to the x-axis (the "x-edges"), followed by the midpoints of the edges parallel to the y-axis (the "y-edges"), and then the element centres. In this way, the grids in 2D are defined as

$$\boldsymbol{G_h} = G_{h_x} \bigoplus G_{h_y},$$

where

$$\boldsymbol{x} := (x, y) \in \boldsymbol{G_h} \text{ if and only if } x \in G_{h_x} \text{ and } y \in G_{h_y},$$

where G_{h_x} and G_{h_y} are defined as in 1D, see (2.3). Here, we consider $h_x = h_y = h$.

Thus, $\boldsymbol{G}_h$ can be rewritten as $\boldsymbol{G}_h = \boldsymbol{G}_h^1 \bigcup \boldsymbol{G}_h^2 \bigcup \boldsymbol{G}_h^3 \bigcup \boldsymbol{G}_h^4$ with

$$\boldsymbol{G}_h^j = \begin{cases} G_{h,N} \bigoplus G_{h,N} & \text{if} \quad j = 1, \\ G_{h,C} \bigoplus G_{h,N} & \text{if} \quad j = 2, \\ G_{h,N} \bigoplus G_{h,C} & \text{if} \quad j = 3, \\ G_{h,C} \bigoplus G_{h,C} & \text{if} \quad j = 4. \end{cases}$$

We refer to $\boldsymbol{G}_h^1, \boldsymbol{G}_h^2, \boldsymbol{G}_h^3$, and $\boldsymbol{G}_h^4$ as the NN-, CN-, NC-, and CC-type points on the grid $\boldsymbol{G}_h$, respectively.

5.1. Representation of the stiffness and mass operators. It is known that the stiffness and mass matrices for the Q_1 approximation in 2D can be written using tensor products of their 1D analogues. However, for the Q_2 approximation in 2D, we must carefully consider the ordering of the DOFs and the block structure of the resulting system. Assume that the stiffness and mass matrices in 1D are ordered by nodes and centres in 2×2-block matrices, given by

$$\mathcal{A}^{(2)} = \begin{pmatrix} A_{nn} & A_{nc} \\ A_{cn} & A_{cc} \end{pmatrix}, \quad \mathcal{B}^{(2)} = \begin{pmatrix} B_{nn} & B_{nc} \\ B_{cn} & B_{cc} \end{pmatrix},$$

respectively. For the 2D case, we use the Tracy-Singh product to preserve block structuring in the product. Let $\mathbf{A}$ be an $(s \times t)$-block matrix, whose (i,j)-block is denoted by A_{ij}, and $\mathbf{B}$ be a $(p \times q)$-block matrix, whose (i,j)-block is denoted by B_{ij}. The Tracy-Singh product of $\mathbf{A}$ and $\mathbf{B}$ is defined by the pairwise Kronecker product for each pair of blocks in matrices $\mathbf{A}$ and $\mathbf{B}$, that is,

$$\mathbf{A} \circ \mathbf{B} = \begin{pmatrix} A_{11} \bar{\otimes} \mathbf{B} & \cdots & A_{1t} \bar{\otimes} \mathbf{B} \\ \vdots & \ddots & \vdots \\ A_{s1} \bar{\otimes} \mathbf{B} & \cdots & A_{st} \bar{\otimes} \mathbf{B} \end{pmatrix}, \text{where } A_{i,j} \bar{\otimes} \mathbf{B} = \begin{pmatrix} A_{ij} \otimes B_{11} & \cdots & A_{ij} \otimes B_{1q} \\ \vdots & \ddots & \vdots \\ A_{ij} \otimes B_{p1} & \cdots & A_{ij} \otimes B_{pq} \end{pmatrix},$$

where $\otimes$ is the standard Kronecker product. Then, the stiffness and mass matrices in 2D are given by

$$\mathcal{A}_2 = \mathcal{A}^{(2)} \circ \mathcal{B}^{(2)} + \mathcal{B}^{(2)} \circ \mathcal{A}^{(2)}, \quad \mathcal{B}_2 = \mathcal{B}^{(2)} \circ \mathcal{B}^{(2)},$$

respectively, and the ordering of the 4×4 block system corresponds to the indexing of the $\boldsymbol{G}_h^j$ given above. Similarly, if the biquadratic restriction matrix in 1D is given in block form as

$$\mathcal{R}^{(2)} = \begin{pmatrix} R_{nn} & R_{nc} \\ R_{cn} & R_{cc} \end{pmatrix},$$

then the corresponding restriction matrix in 2D is given by

$$\mathcal{R}_2 = \mathcal{R}^{(2)} \circ \mathcal{R}^{(2)},$$

with the same block ordering as the blocks in $\mathcal{A}_2$.

Using the Tracy-Singh product for the discretized operators allows us to compute symbols using standard Kronecker products. Given the symbols of the stiffness and

mass operators for the Q_2 approximation in 1D, $\widetilde{A}_h(\theta)$ and $\widetilde{B}_h(\theta)$, respectively, the symbols of the stiffness and mass matrices in 2D are given by

$$\widetilde{A}_2(\theta_1,\theta_2) = \widetilde{A}_h(\theta_2)\otimes\widetilde{B}_h(\theta_1) + \widetilde{B}_h(\theta_2)\otimes\widetilde{A}_h(\theta_1),$$
$$\widetilde{B}_2(\theta_1,\theta_2) = \widetilde{B}_h(\theta_2)\otimes\widetilde{B}_h(\theta_1),$$

respectively.

The above discussion is not limited to Q_2, and extends to Q_k as follows.

Remark 5.1. The stiffness and mass matrices for the Q_k discretization in 2D can be written as

$$\mathcal{A}_k = \mathcal{A}^{(k)}\circ\mathcal{B}^{(k)} + \mathcal{B}^{(k)}\circ\mathcal{A}^{(k)},\ \ \mathcal{B}_k = \mathcal{B}^{(k)}\circ\mathcal{B}^{(k)},$$

respectively, where $\mathcal{A}^{(k)}$ and $\mathcal{B}^{(k)}$ are stiffness and mass matrices for the Q_k discretization in 1D, respectively.

Remark 5.2. The symbols of the stiffness and mass matrices for the Q_k discretization in 2D are as follows

$$\widetilde{A}_k(\theta_1,\theta_2) = \widetilde{A}_h^{(k)}(\theta_2)\otimes\widetilde{B}_h^{(k)}(\theta_1) + \widetilde{B}_h^{(k)}(\theta_2)\otimes\widetilde{A}_h^{(k)}(\theta_1),$$
$$\widetilde{B}_k(\theta_1,\theta_2) = \widetilde{B}_h^{(k)}(\theta_2)\otimes\widetilde{B}_h^{(k)}(\theta_1),$$

respectively, where $\widetilde{A}_h^{(k)}$ and $\widetilde{B}_h^{(k)}$ are the stiffness and mass symbols for the Q_k discretization in 1D, respectively.

Remark 5.3. The restriction matrix corresponding to the Q_k approximation in 2D is given by

$$\mathcal{R}_k = \mathcal{R}^{(k)}\circ\mathcal{R}^{(k)},$$

with the same block ordering as $\mathcal{A}_k$ if $\mathcal{R}^{(k)}$ is ordered consistently with $\mathcal{A}^{(k)}$.

5.2. Fourier representation of grid transfer operators. Now we turn to the representation of biquadratic interpolation and its adjoint operator, restriction, in 2D. The extension of the restriction operator given in (3.8) and (3.9) from 1D to 2D with blocks ordered as mesh nodes, x-edge midpoints, y-edge midpoints, and cell centres can be written as $\boldsymbol{R} = \{\boldsymbol{R}_{NN}, \boldsymbol{R}_{CN}, \boldsymbol{R}_{NC}, \boldsymbol{R}_{CC}\}$, respectively. Let $\widetilde{\boldsymbol{R}}_{NN}, \widetilde{\boldsymbol{R}}_{CN}, \widetilde{\boldsymbol{R}}_{NC}$, and $\widetilde{\boldsymbol{R}}_{CC}$ be their Fourier representations. We show the representation of transfer operators is given by tensor products of their symbols in 1D.

Let

$$\boldsymbol{\alpha} = (\alpha_1,\alpha_2) \in \big\{(0,0),(1,0),(0,1),(1,1)\big\},$$
$$\boldsymbol{\theta^{\alpha}} = (\theta_1^{\alpha_1},\theta_2^{\alpha_2}) = (\theta_1+\alpha_1\pi,\ \theta_2+\alpha_2\pi),\ \boldsymbol{\theta} := \boldsymbol{\theta^{(0,0)}}.$$

We use the ordering of $\boldsymbol{\alpha} = (0,0),(1,0),(0,1),(1,1)$ for the four harmonics.

DEFINITION 5.4. *Assume that $T = [t_{\kappa_1}]$ and $S = [s_{\kappa_2}]$ are two stencil operators in 1D. The 2D stencil $S\bigotimes T$ is given by*

$$S\bigotimes T := [\boldsymbol{r_\kappa}]_{\boldsymbol{h}},\ \textit{with}\ \boldsymbol{r_\kappa} = t_{\kappa_1}s_{\kappa_2},\ \textit{and}\ \boldsymbol{\kappa} = (\kappa_1,\kappa_2),$$

so that R is the outer product of S and T.

We use this outer-product notation to simplify the computation of the symbol of the restriction operator in block form. Rewrite (3.8) and (3.9) as

$$R_N = \begin{bmatrix} -\frac{1}{8} & 0 & \frac{3}{8} & 1(\star) & \frac{3}{8} & 0 & -\frac{1}{8} \end{bmatrix}, \tag{5.1}$$

and

$$R_C = \begin{bmatrix} \frac{3}{4} & 1(\star) & \frac{3}{4} \end{bmatrix}, \tag{5.2}$$

respectively, by discarding the points outside the stencil of restriction. Then, the four restriction stencils in 2D for the Q_2 approximation can be denoted by

$$\boldsymbol{R}_{I_xI_y} = R_{I_y} \bigotimes R_{I_x} := [\boldsymbol{r}_{\boldsymbol{\kappa}}]_{I_xI_y}, \tag{5.3}$$

where $I_x, I_y \in \{N, C\}$.

We can extend Definition 3.2 to a "standard" restriction operator in 2D as follows.

DEFINITION 5.5. *Let $\boldsymbol{T}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) = [\boldsymbol{t}_{\boldsymbol{\kappa}}]$ be a restriction stencil in 2D given as $\boldsymbol{T} = \mathcal{T}_2 \bigotimes \mathcal{T}_1$. We call*

$$\widetilde{\boldsymbol{T}}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) = \sum_{\boldsymbol{\kappa}\in\boldsymbol{V}} \boldsymbol{t}_{\boldsymbol{\kappa}} e^{\iota\boldsymbol{\kappa}\cdot\boldsymbol{\theta}^{\boldsymbol{\alpha}}} e^{\iota\pi\boldsymbol{\alpha}\cdot\boldsymbol{x}/h} := \sum_{\boldsymbol{\kappa}\in\boldsymbol{V}} \widetilde{\boldsymbol{t}}_{\boldsymbol{\kappa}} = \sum_{(\kappa_1,\kappa_2)\in\boldsymbol{V}} \widetilde{t}_{\kappa_1}\widetilde{t}_{\kappa_2}, \tag{5.4}$$

the restriction symbol of $\boldsymbol{T}$.

Here, by "standard", we mean the restriction operator is associated with only one type of meshpoint.

Remark 5.6. It is easy to check that in (5.4),

$$\widetilde{\boldsymbol{T}}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) = \sum_{(\kappa_1,\kappa_2)\in\boldsymbol{V}} \widetilde{t}_{\kappa_1}\widetilde{t}_{\kappa_2} = \sum_{\kappa_1}\sum_{\kappa_2} \widetilde{t}_{\kappa_1}\widetilde{t}_{\kappa_2} = \widetilde{\mathcal{T}}_1(\theta_1^{\alpha_1})\widetilde{\mathcal{T}}_2(\theta_2^{\alpha_2}),$$

where $\widetilde{\mathcal{T}}_1(\theta_1^{\alpha_1})$ and $\widetilde{\mathcal{T}}_2(\theta_2^{\alpha_2})$ are the restriction symbols for $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively, due to the tensor product of $\mathcal{T}_2 \otimes \mathcal{T}_1$.

Note that $\boldsymbol{R}_{I_xI_y}$ draws values from four types of meshpoints on the fine grid. Similarly to 1D, the stencil $\boldsymbol{R}_{I_xI_y}$ can be split into 4 types of substencils, and the Fourier representation of $\boldsymbol{R}_{I_xI_y}$ can be written as a (1×4)-matrix as follows,

$$\widetilde{\boldsymbol{R}}_{I_xI_y}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) = \begin{pmatrix} \widetilde{R}_{I_xI_y,NN}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) & \widetilde{R}_{I_xI_y,CN}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) & \widetilde{R}_{I_xI_y,NC}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) & \widetilde{R}_{I_xI_y,CC}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) \end{pmatrix}. \tag{5.5}$$

The subscript J_xJ_y of $\widetilde{R}_{I_xI_y,J_xJ_y}(\boldsymbol{\theta}^{\boldsymbol{\alpha}})$ $(J_x, J_y \in \{N, C\})$ denotes the contributions of the J_xJ_y-type points on the fine grid to the I_xI_y points on the coarse grid.

Thus, we can use Definition 5.5 to calculate $\widetilde{R}_{I_xI_y,J_xJ_y}(\boldsymbol{\theta}^{\boldsymbol{\alpha}})$.

THEOREM 5.7. *The entries in $\widetilde{\boldsymbol{R}}_{I_xI_y}(\boldsymbol{\theta}^{\boldsymbol{\alpha}})$ in (5.5) are given by,*

$$\widetilde{R}_{I_xI_y,J_xJ_y}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) = \widetilde{R}_{I_y}(J_y, \theta_2^{\alpha_2})\widetilde{R}_{I_x}(J_x, \theta_1^{\alpha_1}) \tag{5.6}$$

where $I_x, I_y, J_x, J_y \in \{N, C\}$. Note that the notation for the right-hand side of (5.6) is defined in the proof of Theorem 3.4.

Proof. Consider a 2D Fourier mode with frequency with $\boldsymbol{\theta}^{\boldsymbol{\alpha}}$, restricted to the coarse grid by the tensor product restriction operators given in (5.3). Because $\boldsymbol{R}_{I_xI_y} = R_{I_y} \bigotimes R_{I_x}$, $\boldsymbol{R}_{I_xI_y}$ can be split into four substencils $R_{I_xI_y,J_xJ_y}$, where $J_x, J_y \in \{N, C\}$, with corresponding symbol $\widetilde{R}_{I_xI_y,J_xJ_y}$. Since the tensor product preserves the stencil structure, $R_{I_xI_y,J_xJ_y} = R_{I_y}(J_y) \otimes R_{I_x}(J_x)$, where $R_{I_y}(J_y)$ stands for the substencil of R_{I_y} corresponding to the contributions from J_y-type points on the find grid, see (3.10) and (3.11). Thus, $\widetilde{R}_{I_xI_y,J_xJ_y}$ can be calculated based on Definition 5.5. According to Remark 5.6, $\widetilde{R}_{I_xI_y,J_xJ_y} = \widetilde{R}_{I_x}(J_x, \theta_1^{\alpha_1})\widetilde{R}_{I_y}(J_y, \theta_2^{\alpha_2})$. ∎

COROLLARY 5.8. *The symbol of restriction in 2D can be written as a tensor product of the restriction symbols in 1D, that is, $\widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{\boldsymbol{\alpha}})$ is the 4×4-matrix given by*

$$\widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) = \widetilde{R}(\theta_2^{\alpha_2}) \otimes \widetilde{R}(\theta_1^{\alpha_1}),$$

ordered as mesh nodes, x-edge midpoints, y-edge midpoints, and cell centres.

Furthermore, the Fourier representation of $\boldsymbol{R}$ is given by the (1×4)-block-matrix

$$\hat{\boldsymbol{R}}(\boldsymbol{\theta}) = \begin{pmatrix} \widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{(0,0)}) & \widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{(1,0)}) & \widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{(0,1)}) & \widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{(1,1)}) \end{pmatrix}.$$

The Fourier representation of $\boldsymbol{P}$ is given by a (16×4)-matrix and

$$\hat{\boldsymbol{P}}(\boldsymbol{\theta}) = \frac{1}{4}\big(\hat{\boldsymbol{R}}(\boldsymbol{\theta})\big)^H.$$

This approach can be extended to Q_k or any other nodal basis for Q_2 as long as the 2D node points are given as a tensor-product of 1D meshes.

COROLLARY 5.9. *The restriction symbol for the Q_k discretization in 2D can be written as a tensor product of the corresponding restriction symbols in 1D. That is, $\widetilde{\boldsymbol{R}}^{(k)}(\boldsymbol{\theta}^{\boldsymbol{\alpha}})$ is the $k^2 \times k^2$-matrix given by*

$$\widetilde{\boldsymbol{R}}^{(k)}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) = \widetilde{R}^{(k)}(\theta_2^{\alpha_2}) \otimes \widetilde{R}^{(k)}(\theta_1^{\alpha_1}),$$

ordered correspondingly to the order of $\widetilde{R}^{(k)}(\theta_1^{\alpha_1})$. Furthermore,

$$\hat{\boldsymbol{P}}^{(k)}(\boldsymbol{\theta}) = \frac{1}{4}\big(\hat{\boldsymbol{R}}^{(k)}(\boldsymbol{\theta})\big)^H.$$

5.3. A lower bound on convergence in 2D. Here, we also discuss the weighted Jacobi relaxation for the Q_2 approximation in 2D. The symbol of the two-grid error propagation operator is

$$\hat{\mathcal{M}}_h^{\mathrm{TGM}}(\boldsymbol{\theta}) = \Big(I - \hat{\boldsymbol{P}}(\boldsymbol{\theta})\hat{A}_{2h}(2\boldsymbol{\theta})^{-1}\hat{\boldsymbol{R}}(\boldsymbol{\theta})\hat{\boldsymbol{A}}_{\boldsymbol{2}}(\boldsymbol{\theta})\Big)\hat{\boldsymbol{S}}_{\boldsymbol{2}}(\boldsymbol{\theta}),$$

where

$$\begin{aligned}
\hat{A}_{2h}(2\boldsymbol{\theta}) &= \widetilde{A}_{2h}(2\theta_2) \otimes \widetilde{B}_{2h}(2\theta_1) + \widetilde{B}_{2h}(2\theta_2) \otimes \widetilde{A}_{2h}(2\theta_1),\\
\hat{\boldsymbol{A}}_{\boldsymbol{2}}(\boldsymbol{\theta}) &= \operatorname{diag}\Big\{\widetilde{A}_2(\boldsymbol{\theta}^{(0,0)}), \widetilde{A}_2(\boldsymbol{\theta}^{(1,0)}), \widetilde{A}_2(\boldsymbol{\theta}^{(0,1)}), \widetilde{A}_2(\boldsymbol{\theta}^{(1,1)})\Big\},\\
\hat{\boldsymbol{S}}_2(\boldsymbol{\theta}) &= \operatorname{diag}\Big\{\widetilde{\mathcal{S}}(\boldsymbol{\theta}^{(0,0)}), \widetilde{\mathcal{S}}(\boldsymbol{\theta}^{(1,0)}), \widetilde{\mathcal{S}}(\boldsymbol{\theta}^{(0,1)}), \widetilde{\mathcal{S}}(\boldsymbol{\theta}^{(1,1)})\Big\},\\
\hat{\boldsymbol{R}}(\boldsymbol{\theta}) &= \Big(\widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{(0,0)}), \widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{(1,0)}), \widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{(0,1)}), \widetilde{\boldsymbol{R}}(\boldsymbol{\theta}^{(1,1)})\Big),\\
\hat{\boldsymbol{P}}(\boldsymbol{\theta}) &= \frac{1}{4}\big(\hat{\boldsymbol{R}}(\boldsymbol{\theta})\big)^H,
\end{aligned}$$

in which

$$\widetilde{\mathcal{S}}(\boldsymbol{\theta}^{\boldsymbol{\alpha}}) = I - \omega \widetilde{M}_2^{-1} \widetilde{A}_2(\boldsymbol{\theta}^{\boldsymbol{\alpha}}), \text{ with}$$

$$\widetilde{M}_2 = \begin{pmatrix} \frac{112}{45} & 0 & 0 & 0 \\ 0 & \frac{176}{45} & 0 & 0 \\ 0 & 0 & \frac{176}{45} & 0 \\ 0 & 0 & 0 & \frac{256}{45} \end{pmatrix}.$$

First, we take a look at the eigenvalues of $\widetilde{M}_2^{-1}\widetilde{A}_2(\boldsymbol{\theta})$. The left of Figure 11 shows the eigenvalue distribution of $\widetilde{M}_2^{-1}\widetilde{A}_2(\boldsymbol{\theta})$ over $[-\frac{\pi}{2}, \frac{3\pi}{2}]^2$. Note that both the smallest and the biggest eigenvalues are achieved over the low frequencies, $[-\frac{\pi}{2}, \frac{\pi}{2}]^2$. As shown at the right of Figure 11 and discussed in more detail below, the standard smoothing analysis fails to predict the two-grid convergence factor in this case as well.

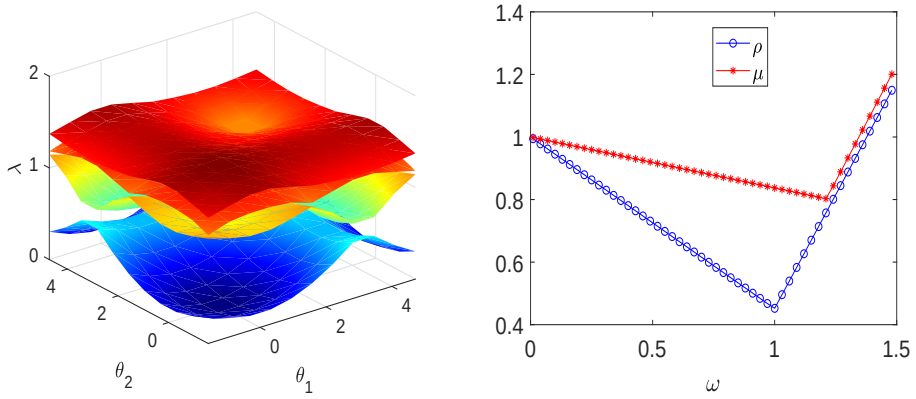

FIG. 5.1. *At left, the distribution of eigenvalues, λ, of* $\widetilde{M}_2^{-1}\widetilde{A}_2(\boldsymbol{\theta})$ *as a function of* $\boldsymbol{\theta} = (\theta_1, \theta_2)$. *At right, LFA-predicted two-grid convergence and smoothing factors as a function of ω for the* Q_2 *approximation in 2D.*

Motivated by the analysis in Subsection 3.3, we consider the limiting behavior of $\hat{\mathcal{M}}_h^{\mathrm{TGM}}(\boldsymbol{\theta})$ when $\theta \to 0$. We first look at the range of the restriction operator when $\boldsymbol{\theta} = (0, 0)$. From Corollary 5.8, we can calculate $\hat{\boldsymbol{R}}(\mathbf{0})$, given by

$$\widetilde{\boldsymbol{R}}(0,0) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \\ 1 & \frac{3}{2} & \frac{1}{2} & \frac{3}{4} \\ 1 & \frac{1}{2} & \frac{3}{2} & \frac{3}{4} \\ 1 & \frac{3}{2} & \frac{3}{2} & \frac{9}{4} \end{pmatrix}, \quad \widetilde{\boldsymbol{R}}(\pi,0) = \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 \\ -1 & 0 & -\frac{1}{2} & 0 \\ 1 & 0 & \frac{3}{2} & 0 \\ -1 & 0 & -\frac{3}{2} & 0 \end{pmatrix},$$

$$\widetilde{\boldsymbol{R}}(0,\pi) = \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 \\ 1 & \frac{3}{2} & 0 & 0 \\ -1 & -\frac{1}{2} & 0 & 0 \\ -1 & -\frac{3}{2} & 0 & 0 \end{pmatrix}, \quad \widetilde{\boldsymbol{R}}(\pi,\pi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Note that the dimensions of the null spaces of $\widetilde{\boldsymbol{R}}(\pi,0)$, $\widetilde{\boldsymbol{R}}(0,\pi)$ and $\widetilde{\boldsymbol{R}}(\pi,\pi)$ are 2, 2, and 3, respectively. Because $\hat{\boldsymbol{P}}(\mathbf{0}) = \frac{1}{4}\hat{\boldsymbol{R}}(\mathbf{0})^H$, we can easily identify seven vectors

that are not treated by coarse-grid correction, and provide a lower bound on the two-grid convergence behavior.

To find the seven vectors (and the associated eigenvalues of $\lim_{\boldsymbol{\theta}\to\mathbf{0}} \hat{\mathcal{M}}_h^{\mathrm{TGM}}(\boldsymbol{\theta})$), we consider the high frequencies corresponding to $(\theta_1^0, \theta_2^0) = (0, 0)$. Let $T_2 = \widetilde{M}_2^{-1}\widetilde{A}_2(\pi, 0)$, $T_3 = \widetilde{M}_2^{-1}\widetilde{A}_2(0, \pi)$, and $T_4 = \widetilde{M}_2^{-1}\widetilde{A}_2(\pi, \pi)$. By standard calculation, we have

$$T_2 = \begin{pmatrix} \frac{29}{28} & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{6}{11} \\ -\frac{7}{22} & 0 & 1 & 0 \\ 0 & -\frac{3}{8} & 0 & 1 \end{pmatrix}, \; T_3 = \begin{pmatrix} \frac{29}{28} & -\frac{1}{2} & 0 & 0 \\ -\frac{7}{22} & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{6}{11} \\ 0 & 0 & -\frac{3}{8} & 1 \end{pmatrix}, \; T_4 = \begin{pmatrix} \frac{15}{14} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Standard calculation shows that T_2 has two eigenvalues, $\hat{\lambda}_{1,2} = 1 \pm \sqrt{\frac{9}{44}}$, with the corresponding eigenvectors $x_{1,2} = \begin{pmatrix} 0 & 1 & 0 & \pm\sqrt{\frac{11}{16}} \end{pmatrix}$, which are the in the null space of $\widetilde{\boldsymbol{R}}(\pi, 0)^H$. Denote $\hat{x}_{1,2} = \begin{pmatrix} \mathrm{z} & x_{1,2} & \mathrm{z} & \mathrm{z} \end{pmatrix}^T$, where z stands for a zero vector with size 1×4. Similarly, it is easy to check that $\hat{\lambda}_{3,4} = 1 \pm \sqrt{\frac{9}{44}}$ are the two eigenvalues of T_3 corresponding to eigenvectors $x_{3,4} = \begin{pmatrix} 0 & 0 & 1 & \pm\sqrt{\frac{11}{16}} \end{pmatrix}$. Denote $\hat{x}_{3,4} = \begin{pmatrix} \mathrm{z} & \mathrm{z} & x_{3,4} & \mathrm{z} \end{pmatrix}^T$.

Finally, the structure of T_3 tells us that it has three eigenvalues: $\hat{\lambda}_{5,6,7} = 1$ and the corresponding eigenvectors are $x_5 = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}$, $x_6 = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}$, $x_7 = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}$, which are in the null space of $\widetilde{\boldsymbol{R}}(\pi, \pi)^H$. Denote $\hat{x}_5 = \begin{pmatrix} \mathrm{z} & \mathrm{z} & \mathrm{z} & x_5 \end{pmatrix}^T$, $\hat{x}_6 = \begin{pmatrix} \mathrm{z} & \mathrm{z} & \mathrm{z} & x_6 \end{pmatrix}^T$, $\hat{x}_7 = \begin{pmatrix} \mathrm{z} & \mathrm{z} & \mathrm{z} & x_7 \end{pmatrix}^T$.

The above discussion gives seven eigenvalues of the two-grid operator $\lim_{\boldsymbol{\theta}\to\mathbf{0}} \hat{\mathcal{M}}_h^{\mathrm{TGM}}(\boldsymbol{\theta})$, leading to the following results.

LEMMA 5.10.

$$\min_{\omega} \left\{ \max \left\{ |\lambda^{**}| \right\} : \lambda^{**} = 1 - \omega\hat{\lambda}_j, 1 \le j \le 7 \right\} = \sqrt{\frac{9}{44}} \approx 0.453, \tag{5.7}$$

and only $\omega = \omega_2^* = 1$ achieves the minimum.

Proof. Since the smallest and largest values of $\hat{\lambda}_j (j = 1, 2, \cdots, 7)$ are $1 - \sqrt{\frac{9}{44}}$ and $1 + \sqrt{\frac{9}{44}}$, respectively, the optimal ω for (5.7) is $\omega_2^* = \frac{2}{1+\sqrt{\frac{9}{44}}+1-\sqrt{\frac{9}{44}}} = 1$. It follows $1 - \omega_2^*\left(1 - \sqrt{\frac{9}{44}}\right) = \sqrt{\frac{9}{44}}$. ∎

COROLLARY 5.11. *For any ω, the optimal convergence factor for the two-grid algorithm using a single weighted Jacobi relaxation (i.e., $\nu_1 + \nu_2 = 1$) on the Q_2 discretization in 2D, is not less than $\sqrt{\frac{9}{44}}$, and this factor can be achieved if and only if $\omega = \omega_2^*$.*

5.3.1. Two-grid and multigrid performance in 2D. In order to see how the parameter ω_2^* performs in practice in a multigrid method, we present two-grid and multigrid results. Table 5 shows that ω_2^* achieves the best possible results, with measured multigrid convergence factors that coincide with the LFA-predicted convergence factors. The same convergence factor is also obtained using full V-cycles, shown in Table 6.

TABLE 5
Two-grid convergence factors for the Q_2 approximation in 2D

Cycle / $\hat{\rho}_h$	$TG(0,1)$	$TG(1,0)$	$TG(1,1)$	$TG(1,2)$	$TG(2,1)$	$TG(2,2)$
$\omega = \omega_2^* = 1.000, \mu = 0.842$						
$\rho_{h=1/128}$	0.452	0.452	0.288	0.123	0.123	0.091
$\hat{\rho}^{(100)}_{h=1/128}$	0.442	0.442	0.280	0.119	0.119	0.088
$\hat{\rho}^{(100)}_{h=1/256}$	0.442	0.442	0.280	0.119	0.119	0.088

TABLE 6
Multigrid convergence factors for the Q_2 approximation in 2D

Cycle / $\hat{\rho}_h$	$V(0,1)$	$V(1,0)$	$V(1,1)$	$V(1,2)$	$V(2,1)$	$V(2,2)$
$\omega = \omega_2^* = 1.000, \mu = 0.842$						
$\rho_{h=1/128}$	0.452	0.452	0.288	0.123	0.123	0.091
$\hat{\rho}^{(100)}_{h=1/128}$	0.442	0.442	0.280	0.117	0.117	0.097
$\hat{\rho}^{(100)}_{h=1/256}$	0.442	0.442	0.281	0.116	0.117	0.097

5.4. A modified two-grid analysis for the Q_2 approximation in 2D. Considering the classical LFA smoothing and convergence factors, The right of Figure 11 indicates that the optimal ω minimizing the two-grid convergence factor is 1, and that the LFA smoothing factor fails to predict the two-grid convergence factor for the Q_2 finite-element approximation in 2D.

In contrast, we plot the LFA-predicted two-grid convergence factor and $\max\{|\lambda^{**}|\}$ as defined in (5.7) as a function of ω, at the left of Figure 12. This shows that for all ω, the two-grid convergence factor is given by $\max\{|\lambda^{**}|\}$, and that convergence is dominated by the harmonic space associated with $\boldsymbol{\theta} = (0,0)$.

The modified prediction given by defining $\mathcal{Q}_0$ using the limit in (3.18) and ρ_0 as in (3.19) can also be extended to this case. We plot ρ_0, compared with the true convergence factor at the right of Figure 12. We see that ρ_0 again overpredicts the convergence factor, as $\mathcal{Q}_0$ captures poorly the true effects of CGC for values of (θ_1, θ_2) near $(\pm\frac{\pi}{2}, \pm\frac{\pi}{2})$. However, ρ_0 still offers a reasonable prediction of convergence and of the optimal relaxation parameter.

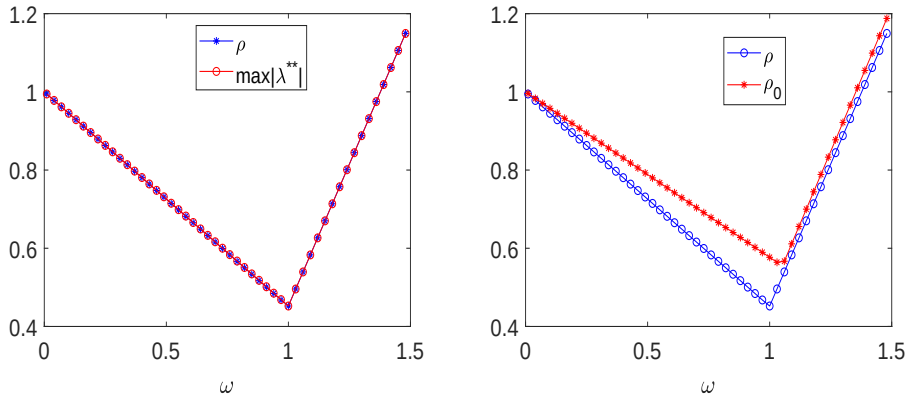

FIG. 5.2. *At left, LFA-predicted two-grid convergence factor and* $\max\{|\lambda^{**}|\}$ *as a function of* ω. *At right, LFA-predicted two-grid convergence factor and* ρ_0, *for the* Q_2 *approximation in 2D.*

6. Conclusion. In this paper, we apply LFA to analyse and optimize the two-grid convergence factor for multigrid methods with higher-order finite-element approximations, especially focusing on optimal parameter choice for quadratic Lagrange elements in 1D and 2D. We find that minimizing the classical LFA smoothing factor fails to accurately predict the two-grid convergence factor. Ideal CGC operators are provided to overcome this failure, and optimal parameters that minimize the two-grid convergence factor are chosen based on the LFA results. With these parameters, we see good agreement between the measured convergence factor and predicted LFA convergence factor with periodic boundary conditions. Compared with the traditional parameter choice, based on minimizing the smoothing factor, we note a big improvement in performance with the corrected parameters. This may also explain why the LFA smoothing factor cannot predict the two-grid convergence factor for higher-order finite-element approximations to other types of PDEs, such as the $Q_2 - Q_1$ approximation to the Stokes equations, which was observed in [10].

REFERENCES

[1] T. BOONEN, J. VAN LENT, AND S. VANDEWALLE, *Local Fourier analysis of multigrid for the curl-curl equation*, SIAM Journal on Scientific Computing, 30 (2008), pp. 1730–1755.
[2] W. L. BRIGGS, V. E. HENSON, AND S. F. MCCORMICK, *A multigrid tutorial*, SIAM, 2000.
[3] J. BROWN, *Efficient nonlinear solvers for nodal high-order finite elements in 3D*, J. Sci. Comput., 45 (2010), pp. 48–63.
[4] M. DONATELLI, C. GARONI, C. MANNI, S. SERRA-CAPIZZANO, AND H. SPELEERS, *Symbol-based multigrid methods for Galerkin B-spline isogeometric analysis*, SIAM J. Numer. Anal., 55 (2017), pp. 31–62.
[5] S. FRIEDHOFF AND S. MACLACHLAN, *A generalized predictive analysis tool for multigrid methods*, Numerical Linear Algebra with Applications, 22 (2015), pp. 618–647.
[6] S. FRIEDHOFF, S. MACLACHLAN, AND C. BORGERS, *Local Fourier analysis of space-time relaxation and multigrid schemes*, SIAM Journal on Scientific Computing, 35 (2013), pp. S250–S276.
[7] W. HACKBUSCH, *Multi-grid methods and applications*, vol. 4, Springer Science & Business Media, 2013.
[8] P. W. HEMKER, W. HOFFMANN, AND M. VAN RAALTE, *Fourier two-level analysis for discontinuous Galerkin discretization with linear elements*, Numerical Linear Algebra with Applications, 11 (2004), pp. 473–491.
[9] J. HEYS, T. MANTEUFFEL, S. MCCORMICK, AND L. OLSON, *Algebraic multigrid for higher-order finite elements*, Journal of Computational Physics, 204 (2005), pp. 520–532.

[10] L. John, U. Rüde, B. Wohlmuth, and W. Zulehner, *On the analysis of block smoothers for saddle point problems*, arXiv preprint arXiv:1612.01333, (2016).
[11] V. John, *Higher order finite element methods and multigrid solvers in a benchmark problem for the 3D Navier-Stokes equations*, Internat. J. Numer. Methods Fluids, 40 (2002), pp. 775–798.
[12] V. John and G. Matthies, *Higher-order finite element discretizations in a benchmark problem for incompressible flows*, International Journal for Numerical Methods in Fluids, 37 (2001), pp. 885–903.
[13] M. Köster and S. Turek, *The influence of higher order FEM discretisations on multigrid convergence*, Comput. Methods Appl. Math., 6 (2006), pp. 221–232.
[14] P. Luo, C. Rodrigo, F. J. Gaspar, and C. W. Oosterlee, *On an Uzawa smoother in multigrid for poroelasticity equations*, Numer. Linear Algebra Appl., 24 (2017). e2074.
[15] S. P. MacLachlan and C. W. Oosterlee, *Local Fourier analysis for multigrid with overlapping smoothers applied to systems of PDEs*, Numer. Linear Algebra Appl., 18 (2011), pp. 751–774.
[16] Y. Maday and R. Muñoz, *Spectral element multigrid. II. Theoretical justification*, J. Sci. Comput., 3 (1988), pp. 323–353.
[17] C. Rodrigo, P. Salinas, F. J. Gaspar, and F. J. Lisbona, *Local Fourier analysis for cell-centered multigrid methods on triangular grids*, Journal of Computational and Applied Mathematics, 259 (2014), pp. 35–47.
[18] E. M. Rønquist and A. T. Patera, *Spectral element multigrid. I. Formulation and numerical results*, J. Sci. Comput., 2 (1987), pp. 389–406.
[19] K. Stüben and U. Trottenberg, *Multigrid methods: Fundamental algorithms, model problem analysis and applications*, Multigrid Methods, (1982), pp. 1–176.
[20] H. Sundar, G. Stadler, and G. Biros, *Comparison of multigrid algorithms for high-order continuous finite element discretizations*, Numer. Linear Algebra Appl., 22 (2015), pp. 664–680.
[21] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, Inc., San Diego, CA, 2001. With contributions by A. Brandt, P. Oswald and K. Stüben.
[22] S. Vandewalle and G. Horton, *Fourier mode analysis of the multigrid waveform relaxation and time-parallel multigrid methods*, Computing, 54 (1995), pp. 317–330.
[23] P. Wesseling, *An introduction to multigrid methods*, Pure and Applied Mathematics (New York), John Wiley & Sons, Ltd., Chichester, 1992.
[24] R. Wienands and W. Joppich, *Practical Fourier analysis for multigrid methods*, CRC press, 2004.