

## Local Fourier Analysis of Multigrid Methods with Polynomial Smoothers and Aggressive Coarsening

James Brannick<sup>1</sup>, Xiaozhe Hu<sup>2,\*</sup>, Carmen Rodrigo<sup>3</sup> and Ludmil Zikatanov<sup>1,4</sup>

<sup>1</sup> Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA.

<sup>2</sup> Department of Mathematics, Tufts University, 503 Boston Ave., Medford, MA 02155, USA.

<sup>3</sup> Department of Applied Mathematics, University of Zaragoza, C/ Maria de Luna 3, 50018, Zaragoza, Spain.

<sup>4</sup> Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 8, 1113 Sofia, Bulgaria.

Received 30 October 2013; Accepted 29 July 2014

---

**Abstract.** We focus on the study of multigrid methods with aggressive coarsening and polynomial smoothers for the solution of the linear systems corresponding to finite difference/element discretizations of the Laplace equation. Using local Fourier analysis we determine *automatically* the optimal values for the parameters involved in defining the polynomial smoothers and achieve fast convergence of cycles with aggressive coarsening. We also present numerical tests supporting the theoretical results and the heuristic ideas. The methods we introduce are highly parallelizable and efficient multigrid algorithms on structured and semi-structured grids in two and three spatial dimensions.

**AMS subject classifications:** 65F10, 65N22, 65N55

**Key words:** Multigrid, local Fourier analysis, polynomial smoothers, aggressive coarsening.

---

### 1. Introduction

For emerging many-core parallel architectures it has been observed that visiting the coarser levels of a multilevel hierarchy leads to a loss in performance, as measured by the percentage of peak performance achieved by the multigrid solver on such architectures. Roughly speaking, on the finer levels, computing residuals and smoothing can achieve relatively high performance, whereas on the coarser levels the performance

---

\*Corresponding author. *Email addresses:* brannick@psu.edu (J. Brannick), Xiaozhe.Hu@tufts.edu (X. Hu), carmenr@unizar.es (C. Rodrigo), ludmil@psu.edu (L. Zikatanov)

of multigrid degrades due to the fact that fewer of the active threads are needed for computation there. These observations motivate the further study and development of multigrid methods that apply more smoothing on the finer levels together with aggressive coarsening strategies.

The use of point-wise smoothers (e.g., Jacobi and Gauss Seidel) together with aggressive coarsening in a multigrid solver has been studied using local Fourier analysis (LFA) in [19, 20] for rectangular grids and [10] for triangular grids. In these works, it has been observed that using aggressive coarsening is less efficient in terms of the total number of floating point operations than a more gradual coarsening approach, since these standard smoothing iterations are not able to effectively reduce a sufficiently large subspace of the high frequency components of the error.

However, polynomial smoothers are well suited for aggressive coarsening approach since they can be constructed to achieve a preset convergence rate on a given subspace corresponding, for example, to a subinterval of the high frequency components of the error. As shown in [3, 8, 12, 17] by using a sufficiently large degree in the polynomial approximation it is possible to guarantee prescribed damping on a preset subinterval of high frequency components. These works contain important results and provide efficient algorithms by adjusting the polynomial degree for a given coarsening ratio.

The focus of our work is on determining precisely the parameters of the polynomial smoothers, such as intervals of approximation, damping factors for high frequencies, coarsening ratios which result in best possible convergence rate. This, of course is an ambitious goal, but for semi-structured triangular and also rectangular grids this can be done. Our idea is to use the local Fourier analysis (LFA) to automatically determine the smoother and coarsening parameters which result in best performance. As we show, LFA allows us to obtain quantitative estimates of the performance of multigrid methods with polynomial smoothers of arbitrary degree and aggressive coarsening. As shown in [3, 4] polynomial smoothers result in algorithms with high degree of parallelism and outperform algorithms based on more classical relaxation methods. This is an additional advantage of the algorithms studied here as well.

The paper is organized as follows. We review some basic facts about two-grid and multigrid iterations in Section 2. Next, in Section 3, we introduce the polynomial smoothers of interest — all based on Chebyshev polynomials arising as solutions to different minimization problems — (1) an appropriately shifted and scaled classical Chebyshev polynomial smoother, (2) the so called smoothed aggregation polynomial, used as a smoother in [8], or (3) the best polynomial approximation to  $x^{-1}$  which is proposed as a smoother in [12]. The local Fourier analysis for these polynomial smoothers, together with a two-grid LFA for aggressive coarsening are presented in Section 3.1, as are their extensions to triangular grids. Then, bounds on the smoothing factors for the polynomials are calculated in Section 4. In Section 4, we also show how to utilize the LFA results in choosing the optimal parameters for the corresponding polynomial smoother. In Section 5 we present numerical tests illustrating the findings in the previous sections and also we provide extension to triangular grids in Section 5.3. Finally we draw some conclusions in Section 6.

## 2. Two-grid and multigrid iterations

A variational two-grid (two-level) method with one post smoothing step is defined as follows. Given an approximation  $w \in V$  to the solution  $u$  of the system  $Au = f$ , an update  $v \in V$  is computed in two steps:

$$(1) \quad y = w + PA_H^{-1}P^T(f - Aw).$$

$$(2) \quad v = y + R(f - Ay).$$

Step (1) is the coarse-grid correction iteration and step (2) is called smoothing step and the operator  $R$  is chosen so that it is convergent in  $A$ -norm, namely  $\|I - RA\|_A < 1$ . The corresponding error propagation operator of the two-level method is given by

$$E_{TL} = (I - RA)(I - \pi_A), \quad \pi_A = PA_H^{-1}P^T A,$$

where  $A_H$  denotes the coarse-level operator, obtained by re-discretizing the problem on the coarse mesh, or, more generally, via Galerkin definition  $A_H = P^T A P$  where  $P$  is the prolongation matrix and  $R$  denotes the smoothing operator. Note that if  $\bar{R} = R + R^T - R^T A R$ , then  $(I - \bar{R}A) = (I - R^T A)(I - RA)$ , and if  $\bar{R}$  is SPD, then the smoother is convergent in  $A$ -norm. We will consider smoothers for which  $R = R^T$  and in this case  $\bar{R} = 2R - RAR$ .

The above two-level method can be easily generalized to multilevel methods by simple recursion. Suppose we already have defined the multilevel method on the coarse level, which is denoted by  $B_H$  (on the coarsest level, we solve the linear system exactly, i.e.  $B_H$  on the coarsest level is defined by  $A_H^{-1}$ ), then the multilevel method on the fine level can be defined as follows:

$$(1) \quad y = w + PB_H P^T(f - Aw).$$

$$(2) \quad v = y + R(f - Ay).$$

It is easy to see the corresponding error propagation operator for the multilevel method is given by

$$E_{ML} = I - BA = (I - RA)(I - PB_H P^T A).$$

Note that, the multilevel method, denoted by  $B$ , is defined recursively through the multilevel method  $B_H$  defined on the coarse grid.

## 3. Simple preconditioners and smoothers based on Chebyshev polynomials

To set the terminology, we call any SPD operator  $R_0$  that approximates  $A^{-1}$  a preconditioner. A simple example of a preconditioner  $R_0$  is the inverse of the diagonal

of  $A$ , i.e.  $R_0 = D^{-1}$ . Another example is furnished by the  $\ell_1$ -Jacobi preconditioner introduced in [11]:

$$R_0 = R_{\ell_1} = \text{diag}(r_{11}, r_{22}, \dots), \quad r_{ii} = 1 / \left( a_{ii} + \sum_{j \neq i} |a_{ij}| \right).$$

Note that  $R_0$  in such case is also a convergent smoother in the terminology given above. A weighted version of  $R_{\ell_1}$  can be found in [8].

Next, for any given preconditioner we can construct a convergent smoother in the following way: Given a preconditioner  $R_0$ , let  $\lambda_1$  be a given bound on the spectral radius of  $R_0 A$ , that is,  $\rho(R_0 A) \leq \lambda_1$  and  $q_\nu(x) \in \mathcal{P}_\nu$  be a polynomial of degree  $\nu$  such that

$$|1 - xq_\nu(x)| < 1, \quad \text{for all } x \in [0, \lambda_1]. \quad (3.1)$$

We then set

$$R = q_\nu(R_0 A) R_0 \quad \text{or equivalently} \quad R = R_0^{\frac{1}{2}} q_\nu(R_0^{\frac{1}{2}} A R_0^{\frac{1}{2}}) R_0^{\frac{1}{2}}. \quad (3.2)$$

Note that this is a symmetric, and, by (3.1), a convergent smoother.

Various polynomial approximations and the resulting parallel smoothers have been designed (also in the context of aggressive coarsening), see e.g. [1, 6, 7, 12, 14, 16].

We consider three different polynomial smoothers given in (3.2). As all these are based on Chebyshev polynomials of first kind we recall the definitions of the  $k$ -th such polynomial  $T_k(t)$  and the following recurrence relation

$$T_0(t) = 1, \quad T_1(t) = t, \quad T_k(t) = 2tT_{k-1}(t) - T_{k-2}(t).$$

Equivalently we can compute that

$$T_k(t) = \frac{1}{2} \left( t - \sqrt{t^2 - 1} \right)^k + \frac{1}{2} \left( t + \sqrt{t^2 - 1} \right)^k.$$

Clearly,  $|T_k(t)| \leq 1$ ,  $t \in [-1, 1]$ , and  $T_k(t) = \pm 1$  at  $(k+1)$  different points in  $[-1, 1]$ . It is clear then that  $|T_k(t)|$  is strictly monotone for  $t \notin (-1, 1)$ .

We now define the three sequences of polynomials  $\{q_\nu\}$  that we consider in this study. We fix  $R_0$  and  $\lambda_1 \geq \rho(R_0 A)$ . Generally, we aim to construct a smoother  $q_\nu(x)$  such that the polynomial  $p_{\nu+1}(x) = 1 - xq_\nu(x)$  is small on the interval  $[\lambda_0, \lambda_1]$ , with  $0 \leq \lambda_0$ . Here  $\nu$  denotes the degree of the polynomial smoother,  $\lambda_1$  is an approximation to the largest eigenvalue of  $R_0 A$ , satisfying  $\lambda_1 > \rho(R_0 A)$  and  $\lambda_0$  is a parameter that can be adjusted to define the smoothing interval. In some of the cases we set  $\lambda_0 = \frac{\lambda_1}{\kappa}$ , with  $\kappa > 1$ , and in other cases we take  $\lambda_0 = 0$ . The smoothing factors clearly depend on the choice of  $\lambda_0$ . How to choose this parameter optimally is not obvious and this is discussed and addressed in our study.

The scaled and shifted classical Chebyshev polynomial solves the following minimization problem:

$$p_{\nu+1}(x) = \arg \min \{ \|p\|_{\infty, [\lambda_0, \lambda_1]}, \quad p \in \mathcal{P}_{\nu+1}, \quad \text{such that } p(0) = 1 \}. \quad (3.3)$$

and we have then for  $q_\nu(x)$ :

$$q_\nu(x) = \frac{1 - p_{\nu+1}(x)}{x}, \quad p_{\nu+1}(x) = \frac{T_{\nu+1}\left(\frac{\lambda_0 + \lambda_1 - 2x}{\lambda_1 - \lambda_0}\right)}{T_{\nu+1}\left(\frac{\lambda_0 + \lambda_1}{\lambda_1 - \lambda_0}\right)}. \tag{3.4}$$

Note that  $p_{\nu+1}(x) = 1 - xq_\nu(x)$ , and from the monotonicity of  $T_{\nu+1}(t)$  for  $t \notin (-1, 1)$  we have  $|p_{\nu+1}(x)| < 1$ , for all  $x \in [0, \lambda_1]$ . This inequality holds independently of the choice of  $\lambda_0 \in [0, \lambda_1]$ . We also refer to [2, 18] for more details on these smoothers.

Using the three term recurrence relation of Chebyshev polynomials, it is straightforward to get the following identity

$$q_\nu(x) - q_{\nu-1}(x) = \frac{2\zeta a T_\nu(a)}{T_{\nu+1}(a)}(1 - xq_{\nu-1}(x)) + \frac{T_{\nu-1}(a)}{T_{\nu+1}(a)}(q_{\nu-1}(x) - q_{\nu-2}(x)),$$

where  $\zeta = \frac{2}{\lambda_1 + \lambda_0}$  and  $a = \frac{\lambda_1 + \lambda_0}{\lambda_1 - \lambda_0}$ . This identity leads to the following algorithm which computes

$$Rr = q_\mu(D^{-1}A)D^{-1}r.$$

---

**Algorithm 3.1** Chebyshev Polynomial Preconditioning with  $R = q_\mu(D^{-1}A)D^{-1}$

---

- 1:  $\bar{r} \leftarrow D^{-1}r$ ,
  - 2:  $v_0 \leftarrow a\bar{r}$ ,
  - 3:  $v_1 \leftarrow \frac{4\zeta a^2}{2a^2-1}\bar{r} - \frac{2\zeta^2 a^2}{2a^2-1}D^{-1}A\bar{r}$ ,
  - 4: **for**  $j = 2, 3, \dots, m$  **do**
  - 5:    $\bar{r}_{j-1} \leftarrow D^{-1}(r - Av_{j-1})$ ,
  - 6:    $v_j = v_{j-1} + \frac{T_{j-1}(a)}{T_{j+1}(a)}(v_{j-1} - v_{j-2}) + \frac{2\zeta a T_j(a)}{T_{j+1}(a)}\bar{r}_{j-1}$ ,
  - 7: **end for**
- 

To define the smoothed aggregation polynomial smoother (see [8]), we use a minimization problem to define  $p_{\nu+1}(x)$ :

$$p_{\nu+1} = \arg \min \{ \|p\sqrt{x}\|_{\infty, [0, \lambda_1]}, \quad p \in \mathcal{P}_{\nu+1}, \quad \text{such that } p(0) = 1 \}. \tag{3.5}$$

The smoothed aggregation polynomial  $q_\nu$  is then defined as follows:

$$q_\nu(x) = \frac{1 - p_{\nu+1}(x)}{x}, \quad p_{\nu+1}(x) = (-1)^\nu \left( \frac{1}{2\nu + 3} \right) \left( \frac{\sqrt{\lambda_1}}{\sqrt{x}} \right) T_{2\nu+3} \left( \frac{\sqrt{x}}{\sqrt{\lambda_1}} \right). \tag{3.6}$$

Note that this formulation does not require the parameter  $\lambda_0$ . We refer to [7, 8, 16] for additional properties of this smoother.

Finally, we mention the smoother based on the best polynomial approximation to  $1/x$  in uniform norm from [12]. The minimization problem now directly defines  $q_\nu(x)$  and

$$q_\nu(x) = \arg \min \{ \|x^{-1} - p\|_{\infty, [\lambda_0, \lambda_1]}, \quad p \in \mathcal{P}_\nu \}. \tag{3.7}$$

Introducing  $p_{\nu+1}(x)$  as in the previous cases, we note that this formulation thus biases the approximation to small values of  $x$ . The polynomial  $q_\nu(x)$  is computed using three-term recurrence relation. For details on this polynomial and its implementation we refer to [12].

### 3.1. Local Fourier analysis for polynomial smoothers

We now describe briefly the technique known as Local Fourier Analysis (LFA), introduced by Brandt in [5]. This technique is considered to be a useful tool in providing quantitative convergence estimates for idealized multigrid algorithms. Such estimates can be rigorously justified in cases when the boundary conditions are periodic. It is also known that for structured or semi-structured grids the LFA provides accurate predictions for the asymptotic convergence rates of multigrid methods for problems with other types of boundary conditions as well. The analysis is based on the Discrete Fourier transform, and a good introduction to such analysis is found in the monographs by Trottenberg et al. [15], and Wienands and Joppich [19].

The main idea of the LFA is to formally extend all multigrid components to an infinite grid, neglecting the boundary conditions, and analyze discrete linear operators with constant coefficients. In this way, the eigenfunctions of such operators are the eigenfunctions of the shift operators, namely,  $S_{\pm h}\varphi = \varphi(\cdot \pm h)$ , called Fourier components. If we assume that the error is a linear combination of the Fourier components, then the behavior of a multigrid algorithm can be studied by looking at the reduction on each one of these components. Although this analysis seems to be somewhat heuristic, its practical value has been widely recognized. In general, the LFA does not only provide accurate asymptotic convergence rates, but also provides the means to select optimal components for the multigrid algorithm.

In this section, we present a suitable local Fourier analysis technique to derive quantitative estimates for the convergence of multigrid methods with polynomial smoothers and aggressive coarsening. In particular, a smoothing analysis for polynomial smoothers and a two-level analysis by considering aggressive coarsening from a grid with step-size  $h$  to a coarse-grid of size  $2^k h$ , are introduced next. We begin by setting up the LFA framework. We extend the discrete problem  $A_h u_h = f_h$  to an infinite grid

$$\Omega_h = \{\mathbf{x} = (x_1, x_2) \mid x_i = k_i h_i, k_i \in \mathbb{Z}, i = 1, 2\}, \quad (3.8)$$

where  $\mathbf{h} = (h_1, h_2)$  is the grid spacing. From the definition of the operators on  $\Omega_h$ , the discrete solution, its current approximation and the corresponding error or residual can be represented by formal linear combinations of the Fourier modes:  $\varphi_h(\boldsymbol{\theta}, \mathbf{x}) = e^{i\theta_1 x_1} e^{i\theta_2 x_2}$ , with  $\mathbf{x} \in \Omega_h$ , and  $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \Theta_h = (-\pi/h_1, \pi/h_1] \times (-\pi/h_2, \pi/h_2]$ . These grid functions form a unitary basis for the space of bounded functions on the infinite grid, and define the Fourier space

$$\mathcal{F}(\Omega_h) := \text{span}\{\varphi_h(\boldsymbol{\theta}, \cdot) \mid \boldsymbol{\theta} \in \Theta_h\}.$$

In this way, the behavior of the multigrid method can be analyzed by evaluating the error reduction associated with a particular multigrid component on the Fourier modes. Clearly, the discrete operator, and, also the discrete error transfer operator have simpler form when expressed in the Fourier basis. For instance, the symbol of  $A_h$ , denoted by  $\tilde{A}_h$ , is typically (block) diagonal, and each diagonal element corresponds to a particular “frequency”. In what follows, we denote by  $\tilde{X}$  the Fourier symbol of a given operator  $X$ .

To perform a smoothing or a two-grid analysis, we distinguish high- and low-frequency components on  $\Omega_h$ . The classification of “high” and “low” here is done with respect to the coarse grid, since some Fourier components are not “visible” on the coarse grid. Usually such “invisible” modes are zero at the coarse grid degrees of freedom, or, they are orthogonal (in an appropriate scalar product) to all Fourier modes corresponding to the coarser grid.

Here, we consider aggressive coarsening techniques and the coarse grid is denoted by  $\Omega_{2^k h}$ , with  $k$  characterizing how “aggressive” the coarsening is. Then, the range of frequencies with respect to this coarse grid is defined as

$$\begin{aligned} \text{all frequencies:} & \quad \Theta_h = (-\pi/h_1, \pi/h_1] \times (-\pi/h_2, \pi/h_2], \\ \text{low frequencies:} & \quad \Theta_{2^k h} = (-\pi/(2^k h_1), \pi/(2^k h_1)] \times (-\pi/(2^k h_2), \pi/(2^k h_2)], \\ \text{high frequencies:} & \quad \Theta_h \setminus \Theta_{2^k h}. \end{aligned} \tag{3.9}$$

In order to investigate the action of the smoothing operator on the high-frequency error components we use a technique known as smoothing analysis.

We consider a splitting of the discrete operator  $A_h = A_h^+ + A_h^-$ , where the splitting defines the smoothing iteration: with a given initial guess  $u_{h,0}$  we define  $u_{h,k+1}$  in terms of  $u_{h,k}$  as follows:

$$\mathbf{u}_{h,k+1} = \mathbf{u}_{h,k} + [A_h^+]^{-1}(f - A_h u_{h,k}).$$

Smoothing analysis can be performed for many choices (symmetric or non-symmetric) of  $[A_h^+]^{-1}$ , but to tie this to our earlier discussion, the smoothers we are interested in are given by  $R_0 = [A_h^+]^{-1}$  in (3.2) with SPD  $A_h^+$ . For example, such smoothers are the Jacobi method, i.e.  $A_h^+ = D = \text{diag}(A_h)$  and the  $\ell_1$ -smoother mentioned earlier.

For the polynomial smoothers, which are of interest here, the error propagation operator is

$$S_h = I_h - p_m(X)X, \quad \text{with } X = [A_h^+]^{-1}A_h, \tag{3.10}$$

where  $p_m$  is a polynomial of degree  $m$ , positive on the spectrum of  $A_h$ . Since Fourier modes are eigenfunctions of the smoothing operator, we can estimate the smoothing factor of  $S_h$ , i.e. the error reduction in the space of high-frequencies  $\tilde{S}_h^\nu(\boldsymbol{\theta})$ , as follows

$$\mu = \sup_{\Theta_h \setminus \Theta_{2^k h}} |\tilde{S}_h^\nu(\boldsymbol{\theta})|, \tag{3.11}$$

where  $\nu$  denotes the number of iterations of the relaxation process.

### 3.2. Two-grid LFA for aggressive coarsening

The basis for the efficient performance of a multigrid method is the interplay between the smoothing and the coarse-grid correction parts of the algorithm. Thus, to get more insight in the behavior of a multigrid algorithm, it is convenient to perform at least a two-grid analysis which takes into account the influence of the components involved in the coarse-grid correction. This is even more important when aggressive coarsening strategies are applied, since the smoother should be chosen according to the factor we are coarsening with. With this purpose, in this section we present a two-grid local Fourier analysis which considers an aggressive coarsening strategy from a grid with step-size  $h$  to a grid with step-size  $2^k h$ .

To perform this analysis, we consider the fine- and coarse-grids  $\Omega_h$  and  $\Omega_{2^k h}$ , respectively. It is well known that in the transition from the fine- to the coarse-grid, each low-frequency is coupled with several high-frequencies. In particular, for an arbitrary  $k$ , each low-frequency  $\theta \in \Theta_{2^k h} = (-\pi/(2^k h), \pi/(2^k h)] \times (-\pi/(2^k h), \pi/(2^k h)]$  is coupled with  $4^k - 1$  high-frequencies by the coarse-grid correction operator. Because of this, the Fourier space can be subdivided into the corresponding  $4^k$ -dimensional subspaces  $\mathcal{F}^{2^k h}(\theta)$  which are generated by the Fourier modes associated with these  $4^k$  frequencies, in the way that

$$\mathcal{F}(\Omega_h) = \bigoplus_{\theta \in \Theta_{2^k h}} \mathcal{F}^{2^k h}(\theta).$$

Therefore, a block-matrix representation of the two-grid operator on the Fourier space can be obtained, which simplifies the computation of the spectral radius of the iteration matrix of the method, since the result should be the maximum of the spectral radius of the corresponding blocks.

Let  $B_h^{2^k h}$  be the error propagation matrix of the considered two-grid method, that is,  $e_h^{m+1} = B_h^{2^k h} e_h^m$ , given by

$$B_h^{2^k h} = S_h^{\nu_2} C_h^{2^k h} S_h^{\nu_1} = S_h^{\nu_2} (I_h - P_{2^k h}^h (A_{2^k h})^{-1} R_h^{2^k h} A_h) S_h^{\nu_1}, \tag{3.12}$$

where  $S_h$  is the iteration matrix associated with the smoother,  $\nu_1, \nu_2$  are, respectively, the number of pre- and post-smoothing steps, and  $C_h^{2^k h}$  is the coarse-grid correction operator, mainly composed of the inter-grid transfer operators:  $P_{2^k h}^h, R_h^{2^k h}$  restriction and prolongation, respectively, and the discrete operators  $A_h, A_{2^k h}$  on the fine and coarse grids, respectively. For standard relaxation schemes and, in particular, for the polynomial smoothers used here the two-grid operator  $B_h^{2^k h}$  leaves invariant the subspaces  $\mathcal{F}^{2^k h}(\theta)$ . As a consequence this operator can be represented by a block-diagonal matrix, consisting of  $(4^k \times 4^k)$ -blocks, denoted by

$$\tilde{B}_h^{2^k h}(\theta) = (\tilde{S}_h(\theta))^{\nu_2} \tilde{C}_h^{2^k h}(\theta) (\tilde{S}_h(\theta))^{\nu_1}.$$

Here,  $\theta \in \Theta_{2^k h}$ ,  $\tilde{S}_h(\theta)$ , and  $\tilde{C}_h^{2^k h}(\theta)$  are the block-matrix representations in the subspaces  $\mathcal{F}^{2^k h}(\theta)$  of the smoother and the coarse-grid correction operator. The latter is

computed from the Fourier representations of the coarse-grid correction, namely,

$$\tilde{C}_h^{2^k h} = \tilde{I}_h - \tilde{P}_{2^k h}^h (\tilde{A}_{2^k h})^{-1} \tilde{R}_h^{2^k h} \tilde{A}_h. \quad (3.13)$$

Now, the local Fourier analysis prediction for the asymptotic two-grid convergence factor of the method can be determined as:

$$\rho = \rho(B_h^{2^k h}) = \max_{\boldsymbol{\theta} \in \Theta_{2^k h}} \rho(\tilde{B}_h^{2^k h}(\boldsymbol{\theta})). \quad (3.14)$$

### 3.3. Extension to triangular grids

To analyze the influence of grid-geometry on the behavior of the multigrid method, in this section we extend the ideas to triangular grids. We consider the discretization of the Laplace operator by linear finite elements on a regular triangulation of a general triangle. This triangular grid is characterized by two angles  $\alpha$  and  $\beta$  and a local enumeration with double index is fixed by considering a unitary basis of  $\mathbb{R}^2$ ,  $\{\mathbf{e}'_1, \mathbf{e}'_2\}$ , fitting the geometry of the triangular grid (see Fig. 1). We fix the axis-orientation  $\{\mathbf{e}'_1, \mathbf{e}'_2\}$ , the corresponding enumeration of vertices as in Fig. 1, and we set  $\gamma = \pi - \alpha - \beta$ . The stencil form of the discrete operator then reads (see, e.g. [13, pp. 189–190])

$$A_h = \frac{\cot \alpha + \cot \beta}{h^2} \begin{bmatrix} 0 & -\cot \alpha & -\cot \beta \\ -\cot \gamma & 2(\cot \alpha + \cot \beta + \cot \gamma) & -\cot \gamma \\ -\cot \beta & -\cot \alpha & 0 \end{bmatrix}. \quad (3.15)$$

Following [9], the local Fourier analysis applied to discretizations on rectangular grids, can be extended to discretizations on triangulations. The key to carrying out this generalization is to define a new two-dimensional Fourier transform using non-orthogonal bases. More precisely, we pick a spatial basis that fits the structure of the grid (see Fig. 1) and we chose its reciprocal basis in the frequency space. In this way LFA on triangular grids is performed similarly as on rectangular grids.

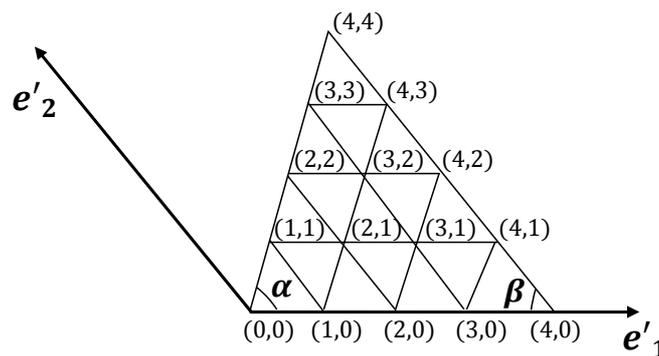


Figure 1: Basis in  $\mathbb{R}^2$  fitting the geometry of a triangular grid characterized by two angles, and local enumeration for the regular grid obtained after two refinement levels.

### 4. Bounds via local Fourier analysis

From the considerations in Section 3 we know that to construct smoothers based on scaled Chebyshev polynomials or based on the best approximation of  $1/x$  we need to specify the interval  $[\lambda_0, \lambda_1]$ . On this interval, as mentioned earlier the polynomials have optimal properties and solve different minimization problems.

In this section, we describe briefly how to optimize the choice of  $\lambda_0$  used in defining  $\mathcal{Q}_m((A_h^+)^{-1}A_h)$ . One simple choice is to use LFA for  $(\tilde{A}_h^+(\boldsymbol{\theta}))^{-1}\tilde{A}_h(\boldsymbol{\theta})$ , and define  $\lambda_0$  and  $\lambda_1$  as

$$\lambda_0 = \min_{\boldsymbol{\theta} \in \Theta_h \setminus \Theta_{2^k h}} |(\tilde{A}_h^+(\boldsymbol{\theta}))^{-1}\tilde{A}_h(\boldsymbol{\theta})|, \quad \lambda_1 = \max_{\boldsymbol{\theta} \in \Theta_h \setminus \Theta_{2^k h}} |(\tilde{A}_h^+(\boldsymbol{\theta}))^{-1}\tilde{A}_h(\boldsymbol{\theta})|. \quad (4.1)$$

This is a good choice, but is not optimal, as is evident from the numerical results presented later. From the properties of the classical Chebyshev polynomial smoothers and the  $BA_{1/x}$  smoother introduced in Section 3 we see that two parameters,  $\lambda_0$  and  $\lambda_1$ , are used to define the smoother. Since we would like to have scaling invariant smoothers we can fix one of these parameters. For reasons which are evident from the analysis provided in [2, 18] for the scaled Chebyshev smoother and [12] for the smoother using the best approximation to  $\frac{1}{x}$ , the parameter that we fix is  $\lambda_1$ . To optimize the performance of the smoother or the two grid method, we try to adjust  $\lambda_0$  in order to achieve one of the following goals:

- get better smoothing properties, that is, minimize  $|\tilde{S}_h(\boldsymbol{\theta})|$  for  $\boldsymbol{\theta} \in \Theta_h \setminus \Theta_{2^k h}$  (high frequency as defined in (3.9)) to compensate for aggressive coarsening;
- get the best possible two grid rate of convergence, namely, minimize  $|\tilde{B}_h^{2^k h}(\boldsymbol{\theta})|$  for  $\boldsymbol{\theta} \in \Theta_h$ .

We find an optimal value of  $\lambda_0$  iteratively, choosing the values given in (4.1) as the initial guess. This can be done as explained in the next section.

#### 4.1. Optimal choice of $\lambda_0$

We now provide several relevant properties of the best polynomial approximation to  $1/x$  and discuss the choice of  $\lambda_0$  for smoothers based on the polynomial of best approximation of  $x^{-1}$ . For the estimates and identities used below, we refer to [12]. Let  $\lambda_0 > 0$  be any estimate of the minimal eigenvalue of  $[A_h^+]^{-1}A_h$ , for example  $\lambda_0 = \lambda_1/\kappa$ , for some  $\kappa > 1$ .

We now discuss the restrictions on the polynomial degree imposed by the requirement that the smoother has certain error reduction and also the requirement that the polynomial is positive on  $(0, \lambda_1]$  (and as a consequence the matrix polynomial will be positive definite). Further, let

$$E_m := \max_{x \in [\lambda_0, \lambda_1]} |1 - xp_m(x)| = \max_{x \in [\lambda_0, \lambda_1]} x \cdot \left| \frac{1}{x} - p_m(x) \right|.$$

As  $\lambda_1$  is a point of Chebyshev alternance, [12, Theorem 2.1 and Equation (2.2)], for the error of approximation  $E_m$  we have

$$E_m = \lambda_1 \left| \frac{1}{\lambda_1} - p_m(\lambda_1) \right| = \left[ \frac{2\lambda_1}{\lambda_1 - \lambda_0} \right] \cdot \left[ \frac{\delta^m}{a^2 - 1} \right] = \frac{2\kappa\delta^m}{(\kappa - 1)(a^2 - 1)}.$$

Here, we have denoted

$$\kappa = \frac{\lambda_1}{\lambda_0}, \quad \delta = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}, \quad a = \frac{\kappa + 1}{\kappa - 1}.$$

Computing the error  $E_m$  then gives

$$E_m = \frac{\delta^m(\kappa - 1)}{2}. \quad (4.2)$$

Regarding the positivity,  $p_m(\lambda_1) > 0$ , a sufficient condition (and also necessary condition in many cases) is that  $\frac{1}{\lambda_1} - E_m > 0$ . Thus, we need to find the smallest  $m$  such that both  $E_m < \rho$ , for a given  $\rho$ , and  $p_m(\lambda_1) > 0$ . We then have that the polynomial is positive if

$$\frac{\delta^m(\kappa - 1)}{2} \leq \frac{1}{\lambda_1} \quad \Rightarrow \quad \delta^m \leq \frac{2}{\lambda_1(\kappa - 1)}.$$

We note that from this it follows that  $R = p_m(A)$  and hence  $\bar{R}$  are symmetric and positive definite, implying that the smoother is convergent in  $A$ -norm.

Also, a straightforward calculation shows that if we want a damping factor less than  $\rho$  on the interval  $[\lambda_0, \lambda_1]$ , we have

$$\frac{\delta^m(\kappa - 1)}{2} \leq \rho \quad \Rightarrow \quad \delta^m \leq \frac{2\rho}{\kappa - 1}.$$

Finally, the minimal  $m$  that will have the desired properties satisfies

$$m \geq \frac{1}{|\log \delta|} \max \left\{ \left| \log \frac{2\rho}{\kappa - 1} \right|, \left| \log \frac{2}{\lambda_1(\kappa - 1)} \right| \right\}. \quad (4.3)$$

Recall that  $\lambda_1$  is fixed, and, hence, both  $\kappa$  and the polynomial degree  $m$  are determined by  $\lambda_0$ . In short, once we choose  $\lambda_0$  we can calculate the degree of the polynomial so that the resulting smoother has a guaranteed convergence rate. This result is of interest in the context of aggressive coarsening since it allows to choose  $\lambda_0$  in accordance with the coarsening ratio and the smoothing rate. Choosing  $\lambda_0$  to optimize these two parameters naturally leads to computationally optimal methods.

Since the convergence rate of the polynomial smoother is determined by  $\max_{x \in [\lambda_0, \lambda_1]} |1 - xp_m(x)|$ , the choice of  $\lambda_0$  using LFA (see (4.1)) does not guarantee the optimal convergence rate of the polynomial smoother. Instead, we consider fixing

$\lambda_1$  and the polynomial degree  $m$ , and finding the best lower bound,  $\lambda_0^*$ , that solves the following min-max problem:

$$\min_{\lambda \in [\lambda_0, \lambda_1]} \max_{x \in [\lambda_0, \lambda_1]} |1 - xp_m(x; \lambda)| \longrightarrow \lambda_0^*, \quad (4.4)$$

where  $p_m(x; \lambda)$  denote the best approximation of  $1/x$  on the interval  $[\lambda, \lambda_1]$ . Next we will explain how to solve this minimization problem and, therefore, obtain the optimal  $\lambda_0^*$ . First, the following lemma shows that the  $|1 - xp_m(x; \lambda)|$ ,  $\lambda \in [\lambda_0, \lambda_1]$  achieves its maximum at the end points  $\lambda_0$  or  $\lambda_1$ .

**Lemma 4.1.** *If  $\lambda \in [\lambda_0, \lambda_1]$ , and  $p_m(x; \lambda)$  is the best polynomial approximation to  $1/x$  on  $[\lambda, \lambda_1]$ , then*

$$\max_{x \in [\lambda_0, \lambda_1]} |1 - xp_m(x; \lambda)| = \max\{|1 - \lambda_1 p_m(\lambda_1; \lambda)|, |1 - \lambda_0 p_m(\lambda_0; \lambda)|\}. \quad (4.5)$$

*Proof.* We first consider  $x \in [\lambda, \lambda_1]$  and since  $p_m(x; \lambda)$  is the best polynomial approximation to  $1/x$  on  $[\lambda, \lambda_1]$ , we have that  $\lambda_1$  is a point of Chebyshev alternance. Therefore,

$$\max_{x \in [\lambda, \lambda_1]} |1 - xp_m(x; \lambda)| = |1 - \lambda_1 p_m(\lambda_1; \lambda)|.$$

On the other hand, for  $x \in [\lambda_0, \lambda]$ , as shown in [12, Lemma 3.1 and Equation (3.5)],  $p_m(x; \lambda)$  is strictly decreasing on  $(0, \lambda]$ , and therefore,

$$\max_{x \in [\lambda_0, \lambda]} |1 - xp_m(x; \lambda)| = |1 - \lambda_0 p_m(\lambda_0; \lambda)|.$$

Then (4.5) follows directly and the proof is complete.  $\square$

Using Lemma 4.1, the minimization problem (4.4) can be simplified as following

$$\min_{\lambda \in [\lambda_0, \lambda_1]} \max\{|1 - \lambda_1 p_m(\lambda_1; \lambda)|, |1 - \lambda_0 p_m(\lambda_0; \lambda)|\}. \quad (4.6)$$

## 4.2. Existence of optimal parameter

In this subsection we provide convincing evidence that an optimal  $\lambda_0^*$  exists. We have not stated this as a theorem because we have used Mathematica to verify the positivity of a derivative in the last step of the proof below. We have the following result.

**Result 4.1.** *If  $\lambda \in [\lambda_0, \lambda_1]$  and  $p_m(x; \lambda)$  is the best polynomial approximation to  $1/x$  on  $[\lambda, \lambda_1]$ , then there is a unique minimizer  $\lambda_0^*$  of the minimization problem (4.4), and*

$$|1 - \lambda_1 p_m(\lambda_1; \lambda_0^*)| = |1 - \lambda_0 p_m(\lambda_0; \lambda_0^*)|. \quad (4.7)$$

Here we outline an argument which at the last step uses a bound verified numerically. As we have observed earlier, since  $\lambda_1$  is a point of Chebyshev alternance, we have

$$|1 - \lambda_1 p_m(\lambda_1; \lambda)| = \frac{(\sqrt{\frac{\lambda_1}{\lambda}} - 1)^m (\frac{\lambda_1}{\lambda} - 1)}{2(\sqrt{\frac{\lambda_1}{\lambda}} + 1)^m}.$$

As it can be easily verified, the right side of this equation is a strictly decreasing function when  $\lambda \in [\lambda_0, \lambda_1]$ .

As  $|1 - \lambda_0 p_m(\lambda_0; \lambda)| = \lambda_0 |1/\lambda_0 - p_m(\lambda_0; \lambda)|$ , we will focus on  $\mathcal{E}_m(x; \lambda) = |1/x - p_m(x; \lambda)|$ . According to the three term recurrence of  $p_m(x; \lambda)$  [12, Equation (2.13)], we have

$$p_{m+1}(x; \lambda) - p_m(x; \lambda) = \delta^2(p_m(x; \lambda) - p_{m-1}(x; \lambda)) + c(1 - xp_m(x; \lambda)),$$

where  $\delta = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$  with  $\kappa = \frac{\lambda_1}{\lambda}$  and  $c = \frac{4\sqrt{\mu_0}\sqrt{\mu_1}}{(\sqrt{\mu_0}+\sqrt{\mu_1})^2}$  with  $\mu_0 = \frac{1}{\lambda_1}$  and  $\mu_1 = \frac{1}{\lambda}$ . Then we have the following three term recurrence of  $\mathcal{E}_m(x; \lambda)$

$$\mathcal{E}_{m+1}(x; \lambda) = 2y\delta\mathcal{E}_m(x; \lambda) - \delta^2\mathcal{E}_{m-1}(x; \lambda),$$

where  $y = \frac{1+\delta^2-cx}{2\delta}$ . Because  $\mathcal{E}_m$  satisfies the homogeneous three term recurrence relationship, we have the following explicit formula

$$\mathcal{E}_m(x; \lambda) = -\delta^m \mathcal{E}_0(x; \lambda) U_{m-2}(y) + \delta^{m-1} \mathcal{E}_1(x; \lambda) U_{m-1}(y),$$

where

$$\mathcal{E}_0(x; \lambda) = \frac{1}{x} - \frac{1}{2}(\mu_0 + \mu_1); \quad \mathcal{E}_1(x; \lambda) = \frac{1}{x} - \left(\frac{1}{2}(\sqrt{\mu_0} + \sqrt{\mu_1})^2 - \mu_0\mu_1 x\right).$$

Using the software package *Mathematica* (see the output below), we can verify that  $\mathcal{E}_m(\lambda_0; \lambda) > 0$  and is a strictly increasing function when  $\lambda \in [\lambda_0, \lambda_1]$ , and therefore, so is

$$|1 - \lambda_0 p_m(\lambda_0; \lambda)| = \lambda_0 \mathcal{E}_m(\lambda_0; \lambda).$$

Therefore, because  $|1 - \lambda_1 p_m(\lambda_1; \lambda)|$  is strictly decreasing and  $|1 - \lambda_0 p_m(\lambda_0; \lambda)|$  is strictly increasing when  $\lambda \in [\lambda_0, \lambda_1]$ ,  $\min_{\lambda \in [\lambda_0, \lambda_1]} \max\{|1 - \lambda_1 p_m(\lambda_1; \lambda)|, |1 - \lambda_0 p_m(\lambda_0; \lambda)|\}$  has a unique minimizer  $\lambda_0^*$  satisfying (4.7). Moreover, according to Lemma 4.1, this minimizer  $\lambda_0^*$  is the unique minimizer of the minimization problem (4.4).

Thus, Result 4.1 and (4.7), show that to find  $\lambda_0^*$  satisfying (4.7) we need to solve a one-dimensional nonlinear equation. This value of  $\lambda_0^*$  results in the best convergence rate for the polynomial smoother for the given range of frequencies.

At the end of this subsection, we provide the *Mathematica* code which is used to verify Result 4.1. Without loss of generality, we assume that  $\lambda_1 = 1$ , introduce  $t = \sqrt{\kappa}$ , and change the formulas of  $y$ ,  $\mathcal{E}_0$ , and  $\mathcal{E}_1$  correspondingly. In the *Mathematica* code, we first input the variables and then compute the derivative of  $\mathcal{E}_m$  with respect to  $t$ .

```

In[1] := delta = (t-1)/(t+1);
          y = (1+t^2 - 2*t*x)/(t^2 - 1);
          E0 = -0.5*t^2 + ((1/x) - 0.5);
          E1 = -0.5*(t+1)^2 + t^2*x + 1/x;
          Em = delta^(m-1)*E1*ChebyshevU[m-1,y]
              - delta^m*E0*ChebyshevU[m-2,y];
          Minimize[{Em, 1<=t<=1/Sqrt[x], 0<x<=1, 3<=m}, {t, x, m}];
Out[2] = {3.51279 × 10-22, {t → 1.03272, x → 0.917826, m → 18.11}}
In[3] := Maximize[{∂x(Em), 1<=t<=1/Sqrt[x], 0<x<=1, 3<=m}, {t, x, m}];
Out[4] = {-3.77826 × 10-48, {t → 1.0027, x → 0.919534, m → 43.6199}}

```

From the output displayed above, we can see that  $\mathcal{E}_m > 0$  and the derivative of  $\mathcal{E}_m$  with respect to  $t$  is negative and, therefore,  $\mathcal{E}_m$  is strictly decreasing, which means  $\mathcal{E}_m$  is strictly increasing with respect to  $\lambda$ . This concludes the heuristic justification of Result 4.1.

## 5. Numerical experiments

In this section we present numerical results on smoothing properties of the polynomial smoothers and results on two- and multi-level methods which use these smoothers in combination with aggressive coarsening strategy. When the polynomial smoother is the best polynomial approximation of  $x^{-1}$  or the Chebyshev polynomial, the tests also include study of the behavior of the smoothers and the multigrid algorithms with respect to the parameters involved in defining these polynomial smoothers (e.g.  $\lambda_0$  as discussed earlier in Section 4).

To fix ideas we consider as our model problem the Poisson equation on a domain  $\Omega$  with homogeneous Dirichlet boundary conditions

$$\mathcal{A}u(\mathbf{x}) = -\Delta u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \Omega \subseteq \mathbb{R}^d, \quad d = 2, 3, \quad (5.1)$$

$$u(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\Omega. \quad (5.2)$$

The domains  $\Omega$  which we consider are the unit cubes in  $\mathbb{R}^d$ ,  $d = 2, 3$  or a triangular domain in  $\mathbb{R}^2$ . We group the numerical tests in this section as follows. In Section 5.1 we present results on the smoothing property of the polynomial smoothers. The two-grid and multigrid results are shown in Section 5.2. As discretization technique on rectangular grids we use the standard 5-point finite difference discretization of problem (5.1)-(5.2). At the end we also show that our results can be easily generalized to continuous linear finite element discretizations on triangular grids (see Section 5.3, and Section 3.3). In the computations, we use bilinear interpolation for rectangular grids and linear interpolation for triangular grids.

One notation that is needed is the corresponding cycling strategy. We call  $k$ -coarsening a coarsening procedure which results in a coarser grid size of  $2^k h$  if the size of the fine grid is  $h$ . For multigrid method with total number of levels  $L$ , this means that the grid size on level  $\ell$  is  $2^{k\ell} h_L$  if  $L$ -th level is the finest grid level. Such coarsening is called aggressive since a coarser grid has  $2^{-kd} n$  degrees of freedom if the spatial dimension is  $d$  and the fine grid has  $n$  degrees of freedom.

### 5.1. Comparison of smoothing rates

The crucial role of the smoothing as one main component in a multigrid process is well known (see the classical paper by A. Brandt [5]). In fact, as shown in this pioneering work on multigrid methods, the smoothing analysis can provide a first, and often accurate, estimate on the convergence factor of the overall method.

In this section, we compare the smoothing properties of the polynomial smoothers introduced in Section 3: (a) the scaled classical Chebyshev polynomial; (b) the smoothed aggregation polynomial, denoted here with (SA); and (c) the best polynomial approximation to  $x^{-1}$ , denoted here with ( $BA_{1/x}$ ). For the latter polynomial we provide two different values for the lower bound of the eigenvalue interval,  $\lambda_0$ : (1) the choice of  $\lambda_0$  given by the LFA (see (4.1)); and (2) the optimal choice of  $\lambda_0$ , given in Section 4.1, and denoted by  $\lambda_0^*$ . The results for the Chebyshev polynomial smoothers are for the interval  $[\lambda_0, \lambda_1]$  with the value of  $\lambda_0$  determined from (4.1). In Table 1, we present the smoothing factors predicted by the local Fourier analysis for these polynomial smoothers in  $2D$ . For the LFA estimates we specify as “low frequencies” the eigenmodes corresponding to a lattice of size  $2^k h$ , where  $h$  is the fine grid mesh size. In this setting, larger values of  $k$  correspond to more aggressive coarsening. For the case of  $BA_{1/x}$  smoother this in turn requires higher degree of the polynomial. We use the same degree for the Chebyshev and also SA polynomials and we observe that the scaled Chebyshev polynomial provides the best smoothing factors for any of the choices of aggressive coarsening. Also we observe that choosing the optimal value of  $\lambda_0^*$  for the smoothing of the  $1/x$  polynomial improves significantly the smoothing rate for this smoother.

Table 1: Smoothing rate comparison of different polynomial smoothers ( $\lambda_1 = 2$ ) in  $2D$ .

	Chebyshev	SA	$BA_{1/x}$	$BA_{1/x}(\lambda_0^*)$	$\lambda_0$	$\lambda_0^*$
$k = 1$ , Degree = 2	0.074	0.233	0.167	0.100	0.500	0.598
$k = 2$ , Degree = 6	0.041	0.221	0.226	0.086	0.146	0.202
$k = 3$ , Degree = 17	0.014	0.172	0.230	0.053	0.038	0.057

Next, similar results are presented in Table 2 for the three-dimensional case. Again, we can draw the same conclusions since Chebyshev polynomial results in the best smoothing rates (even better than in the two-dimensional case) and the optimal choice of  $\lambda_0 = \lambda_0^*$  appears to be crucial to obtain good convergence factors for the  $BA_{1/x}$  smoother.

Table 2: Smoothing rate comparison of different polynomial smoothers ( $\lambda_1 = 2$ ) in 3D.

	Chebyshev	SA	$BA_{1/x}$	$BA_{1/x}(\lambda_0^*)$	$\lambda_0$	$\lambda_0^*$
$k = 1$ , Degree = 3	0.062	0.227	0.185	0.097	0.333	0.419
$k = 2$ , Degree = 9	0.022	0.215	0.171	0.059	0.976	0.134
$k = 3$ , Degree = 22	0.011	0.148	0.268	0.051	0.025	0.039

## 5.2. Multigrid cycles

While the smoothing analysis results presented in previous section give us some insight about the performance of the method, to properly study the behavior of multigrid algorithm we need to involve the coarse grid correction into the tests. We first present a two-grid analysis (LFA) which takes into account the effect of transfer operators and the rest of the components of the coarse-grid correction part of the algorithm. We report the two-grid convergence rates obtained by LFA as well as the performance of a W-cycle solver using the polynomial smoothers. We present the results for two-grid method and W-cycle for 1-coarsening, i.e. all levels are involved in the coarse grid correction and there is no skipping of levels.

Next, we show results on the behavior of multigrid V-cycles. The tests are carried out for  $k$ -coarsening strategies, with  $k = 1, 2, 3$ , namely we study the performance of the algorithms with respect to coarsening strategy with different “aggressiveness”.

For the two- and multi-grid tests we only compare Chebyshev polynomial and  $1/x$  polynomial since from the previous smoothing rate tests, we can assert that these two polynomials outperform the SA polynomials as smoothers. Also all the convergence factors presented in the tables are computed with only one smoothing step.

In the tables below, we denote by  $\rho_{LFA}$  the two-grid convergence factors predicted by LFA. The results for the  $k$ -coarsening strategies are presented in Table 3. In order to validate these results, we also show in this table the corresponding experimentally computed asymptotic two-grid convergence factors  $\rho_W$ . The lower and upper bounds of the eigenvalue interval  $[\lambda_0, \lambda_1]$  provided by the LFA are also given, together with the optimal value  $\lambda_0^*$ , computed as in (4.7), and with the minimal degree of the polynomial smoothers. We observe that according to Table 3 the local Fourier analysis provides sharp estimates for the convergence rates in all the cases.

Finally, we remark on the choice of parameter  $\lambda_0^*$ . As seen this choice improves the convergence factors provided by  $BA_{1/x}$  smoother. The improvement is more significant for  $k$ -coarsening with larger  $k$ . This makes the convergence rates of  $BA_{1/x}$  very close to those given by the best option, namely, the scaled and shifted classical Chebyshev polynomial. It is also interesting to choose a value of  $\lambda_0$  (left end of the high frequency interval) which is optimal with respect to the overall two-grid convergence of the method. Note that, this is a different procedure than what we have done before: choosing  $\lambda_0^*$  such that the smoothing factor is the best possible. This value has been computed by using the local Fourier analysis, for both Chebyshev and  $BA_{1/x}$  polynomials. More precisely, we find the optimal  $\lambda_0^*$  not only by monitoring the convergence of the smoother,

Table 3: Two-grid convergence rate for different polynomial smoothers in 2D.

					Chebyshev		BA <sub>1/x</sub>		BA <sub>1/x</sub> ( $\lambda_0^*$ )	
	$\lambda_0$	$\lambda_0^*$	$\lambda_1$	Degree	$\rho_{LFA}$	$\rho_W$	$\rho_{LFA}$	$\rho_W$	$\rho_{LFA}$	$\rho_W$
$k = 1$	0.5	0.598	2.0	2	0.125	0.126	0.166	0.166	0.134	0.134
$k = 2$	0.146	0.202	2.0	6	0.156	0.155	0.221	0.225	0.166	0.165
$k = 3$	0.038	0.057	2.0	17	0.137	0.137	0.227	0.227	0.148	0.149

Table 4: Two-grid convergence rate for Chebyshev and BA<sub>1/x</sub> polynomial smoothers. The value of  $\lambda_0$  is chosen by applying LFA for the overall convergence factor of the two-grid method.

	Chebyshev( $\lambda_0^{Optimal-TL}$ )					BA <sub>1/x</sub> ( $\lambda_0^{Optimal-TG}$ )		
	$\lambda_1$	Degree	$\lambda_0$	$\rho_{LFA}$	$\rho_W$	$\lambda_0$	$\rho_{LFA}$	$\rho_W$
$k = 1$	2.0	2	0.405	0.111	0.113	0.550	0.128	0.128
$k = 2$	2.0	6	0.095	0.138	0.140	0.167	0.152	0.153
$k = 3$	2.0	17	0.019	0.100	0.100	0.045	0.133	0.133

Table 5: Multigrid convergence rates for a V(1,1)-cycle in 2D. The value of  $\lambda_0$  is chosen by LFA (Eq. (4.1)).

	$\lambda_0$	$\lambda_0^*$	BA <sub>1/x</sub> ( $\lambda_0$ )	BA <sub>1/x</sub> ( $\lambda_0^*$ )	Chebyshev
$k = 1$ , Degree= 2	0.5	0.598	0.103	0.114	0.111
$k = 2$ , Degree= 6	0.146	0.202	0.088	0.103	0.098
$k = 3$ , Degree= 17	0.038	0.057	0.069	0.083	0.076

but also the convergence of the two-grid method. In Table 4 we present two-grid convergence factors for this case as predicted by the LFA and we also show the experimentally computed W-cycle rates. We observe results similar to the ones already discussed in Table 3.

Further, we test the behavior of V-cycle for the different polynomial smoothers and different coarsening strategies with varying coarsening ratios. We consider a V-cycle with one pre- and one post-smoothing step. The corresponding results are shown in Table 5, and we observe that for both polynomials the obtained convergence rates are  $\approx 0.1$ , that is, we have fast convergence reducing the error in energy norm by an order of magnitude per iteration.

In regard to the three-dimensional case the tests provide a solid basis for conclusions analogous to the ones obtained for the two-dimensional case. The results for V(1,1)-cycle in three dimensions are shown in Table 6. Also, as in the 2D case, we include results on the behavior with respect to the choice of the parameter  $\lambda_0$ , computed by LFA, and also its optimal version,  $\lambda_0^*$  given by (4.7). As in the 2D case we observe excellent convergence rates.

**Remark 5.1.** We remark that in Table 5 and Table 6 the value of  $\lambda_0^*$  is calculated so that it optimizes the smoothing factor only. However, since the overall algorithm involves in addition a coarse grid correction, we cannot expect that  $\lambda_0^*$  will yield an optimal convergence rate for the resulting multilevel method (since it does not take

Table 6: Multigrid convergence rates for a V(1,1)-cycle in 3D. The value of  $\lambda_0$  is chosen by LFA (Eq. (4.1)).

	$\lambda_0$	$\lambda_0^*$	$BA_{1/x}(\lambda_0)$	$BA_{1/x}(\lambda_0^*)$	Chebyshev
$k = 1$ , Degree= 3	0.333	0.419	0.101	0.115	0.110
$k = 2$ , Degree= 9	0.098	0.134	0.084	0.099	0.094
$k = 3$ , Degree= 22	0.025	0.039	0.071	0.090	0.079

the coarse grid correction into account). In short, the proposed choice for the value of  $\lambda_0^*$  guarantees a better smoothing rate, but this does not necessarily lead to a better V-cycle convergence rate. This is shown in Table 5 and Table 6. We note that the effects of this observed phenomenon are minimal in practice and in general the convergence rates of the methods for  $\lambda_0$  and for  $\lambda_0^*$  are generally very close.

### 5.3. Extension to triangular grids

All the ideas and techniques we introduced earlier can be extended to the case of triangular grids. This allows us to study how the geometry of the grid influences the convergence rates and what are the optimal parameters for the polynomial smoothers with respect to the geometry of the grid.

We consider continuous linear finite element discretization of the problem (5.1)-(5.2) on a structured triangular grid characterized by two angles  $\alpha$  and  $\beta$ , as explained in Section 3.3. As mentioned in that section, LFA can be applied on triangular grids. This gives us a tool to automatically choose polynomial degrees and suitable values of  $\lambda_0$ , as we did for rectangular grids. As intergrid transfer operator we use the natural inclusion (linear interpolation).

First example is on a grid with equilateral triangles (the domain  $\Omega$  is also such a triangle). In Table 7 we display the two-grid convergence factors provided by the LFA for the multigrid algorithms based on the polynomial smoothers we consider. The results are shown for  $k$ -coarsening for different values of  $k$ . We display also the values of  $\lambda_0$  and  $\lambda_0^*$  in Table 7. Similarly to the observation made earlier, the Chebyshev polynomial smoothers seem to provide the best convergence factors, followed by the  $BA_{1/x}$ -smoother with the optimal choice of  $\lambda_0^*$ . Here the optimal choice of  $\lambda_0^*$  again improves significantly the smoothing properties of the  $BA_{1/x}$ -smoother. Also, we observe that the minimal degrees for the polynomial smoothers are very close to those used in the rectangular case, and even slightly lower.

Next we fix an structured isosceles triangular grid with base angle  $\frac{4}{9}\pi$ . This triangulation has a relative small third angle equal to  $\frac{\pi}{9}$  and this induces some grid anisotropy. We have applied the local Fourier analysis to the resulting discretization in such a grid and the two-grid convergence factors are displayed in Table 8 for the different  $k$ -coarsening strategies and different smoothers. We observe that we are able to obtain good convergence factors at the price of increasing the degree of the polynomial. This shows that much “stronger” smoothers are needed for anisotropic problems, a fact that is known in the multigrid community. In fact, when standard coarsening is considered,

Table 7: Two-grid convergence factors on an equilateral triangle by using linear interpolation. The value of  $\lambda_1$  is  $\frac{3}{2}$  and the value of  $\lambda_0$  is chosen by LFA (Eq. (4.1)).

	$\lambda_0$	$\lambda_0^*$	$BA_{1/x}(\lambda_0)$	$BA_{1/x}(\lambda_0^*)$	Chebyshev
$k = 1$ , Degree= 1	0.529	0.623	0.212	0.138	0.129
$k = 2$ , Degree= 5	0.148	0.195	0.175	0.101	0.102
$k = 3$ , Degree= 14	0.038	0.056	0.236	0.091	0.086

Table 8: Two-grid convergence factors on an isosceles triangle with common angle  $\frac{4}{9}\pi$ , by using linear interpolation. The value of  $\lambda_1$  is  $\frac{17}{9}$  and the value of  $\lambda_0$  is chosen by LFA (Eq. (4.1)).

	$\lambda_0$	$\lambda_0^*$	$BA_{1/x}(\lambda_0)$	$BA_{1/x}(\lambda_0^*)$	Chebyshev
$k = 1$ , Degree= 8	0.112	0.151	0.151	0.079	0.064
$k = 2$ , Degree= 18	0.033	0.049	0.261	0.101	0.092
$k = 3$ , Degree= 43	0.009	0.014	0.616	0.095	0.086

coupled smoothers (line-wise relaxation) are preferred against the standard point-wise relaxations as Jacobi, Gauss-Seidel. We show here that with appropriate polynomial degree, polynomial smoothers also result in good convergence rates and in addition they have advantage in parallel computation (compared to block Gauss-Seidel or block Jacobi method).

## 6. Conclusions

We have devised a simple technique based on the local Fourier analysis which allows us to construct polynomial smoothers with optimal smoothing factors and parameters tied also to the damping of the error on the coarser grids. The theoretical and the numerical results clearly confirm that LFA is a useful tool in designing efficient multigrid algorithms with more aggressive coarsening and polynomial smoothing. Also as it is seen in the numerical examples section the degree of the polynomials grows linearly with respect to the coarsening ratio. This is not surprising since the more we coarsen, the more smoothing is needed to cover the whole high-frequency interval. This technique can be applied on structured as well as semi-structured grids. These methods are suitable for parallelization since they only involve matrix vector multiplications and the local Fourier analysis automates the parameter choice. Studies and comparisons with multicolored Gauss-Seidel and SOR smoothers are part of the planned research.

**Acknowledgments** The work of James Brannick was supported in part by NSF grants DMS-1217142 and DMS-1320608 and by Lawrence Livermore National Laboratory through subcontract B605152. The work of Carmen Rodrigo is supported in part by the Spanish project FEDER/MCYT MTM2010-16917 and the DGA (Grupo consolidado PDIE). The research of Ludmil Zikatanov is supported in part by NSF through DMS-1217142, DMS-1418843 and Lawrence Livermore National Laboratory through subcontract B603526. Carmen Rodrigo gratefully acknowledges the hospitality of the

Center for Computational Mathematics and Applications and the Department of Mathematics of The Pennsylvania State University, where this research was partly carried out.

## References

- [1] Mark Adams, Marian Brezina, Jonathan Hu, and Ray Tuminaro. Parallel multigrid smoothing: polynomial versus Gauss-Seidel. *J. Comp. Phys*, 188:593–610, 2003.
- [2] Owe Axelsson and Panayot S. Vassilevski. Algebraic multilevel preconditioning methods. I. *Numer. Math.*, 56(2-3):157–177, 1989.
- [3] Allison H. Baker, Robert D. Falgout, Tzanio V. Kolev, and Ulrike M. Yang. Multigrid smoothers for ultraparallel computing. *SIAM Journal on Scientific Computing*, 33(5):2864–2887, 2011.
- [4] Allison H. Baker, Robert D. Falgout, Tzanio V. Kolev, and Ulrike M. Yang. Multigrid smoothers for ultraparallel computing: Additional theory and discussion. Technical report, Technical report LLNL-TR-489114, Lawrence Livermore National Laboratory, 2011.
- [5] Achi Brandt. Multi-level adaptive solutions to boundary-value problems. *Math. Comp.*, 31(138):333–390, 1977.
- [6] Marian Brezina, Caroline Heberton, Jan Mandel, and Petr Vaněk. An iterative method with convergence rate chosen a priori. Technical report, Technical report: Center for Computational Mathematics, University of Colorado at Denver, 1999.
- [7] Marian Brezina, Petr Vaněk, and Panayot S. Vassilevski. An improved convergence analysis of smoothed aggregation algebraic multigrid. *Numer. Linear Algebra Appl.*, 19(3):441–469, 2012.
- [8] Marian Brezina and Panayot S. Vassilevski. Smoothed aggregation spectral element agglomeration AMG: SA- $\rho$ AMGe. In *Large-scale scientific computing*, volume 7116 of *Lecture Notes in Comput. Sci.*, pages 3–15. Springer, Heidelberg, 2012.
- [9] Francisco J. Gaspar, José L. Gracia, and Francisco J. Lisbona. Fourier analysis for multigrid methods on triangular grids. *SIAM J. Sci. Comput.*, 31(3):2081–2102, 2009.
- [10] Francisco J. Gaspar, José L. Gracia, Francisco J. Lisbona, and Carmen Rodrigo. On geometric multigrid methods for triangular grids using three-coarsening strategy. *Appl. Numer. Math.*, 59(7):1693–1708, 2009.
- [11] Tzanio V. Kolev and Panayot S. Vassilevski. Parallel auxiliary space AMG for  $H(\text{curl})$  problems. *J. Comput. Math.*, 27(5):604–623, 2009.
- [12] Johannes K. Kraus, Panayot S. Vassilevski, and Ludmil T. Zikatanov. Polynomial of best uniform approximation to  $1/x$  and smoothing in two-level methods. *Computational Methods in Applied Mathematics*, 12(4):448–468, 2012.
- [13] Carmen Rodrigo, Francisco J. Gaspar, and Francisco J. Lisbona. *Geometric multigrid methods on Triangular Grids: Application to semi-structured meshes*. Lambert Academic Publishing, Saarbrücken, 2012.
- [14] Klaus Stüben and Ulrich Trottenberg. Multigrid methods: fundamental algorithms, model problem analysis and applications. In *Multigrid methods (Cologne, 1981)*, volume 960 of *Lecture Notes in Math.*, pages 1–176. Springer, Berlin, 1982.
- [15] Ulrich Trottenberg, Cornelis W. Oosterlee, and Anton Schüller. *Multigrid*. Academic Press Inc., San Diego, CA, 2001. With contributions by A. Brandt, P. Oswald and K. Stüben.
- [16] Petr Vanek, Jan Mandel, and Marian Brezina. Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. *Computing*, 56(3):179–196, 1996.

- [17] Petr Vaněk and Marian Brezina. Nearly optimal convergence result for multigrid with aggressive coarsening and polynomial smoothing. *Applications of Mathematics*, 58(4):369–388, 2013.
- [18] Panayot S. Vassilevski. *Multilevel block factorization preconditioners*. Springer, New York, 2008. Matrix-based analysis and algorithms for solving finite element equations.
- [19] Roman Wienands and Wolfgang Joppich. *Practical Fourier analysis for multigrid methods*, volume 4 of *Numerical Insights*. Chapman & Hall/CRC, Boca Raton, FL, 2005. With 1 CD-ROM (Windows and UNIX).
- [20] Hisham Bin Zubair, Cornelis W. Oosterlee, and Roman Wienands. Multigrid for high-dimensional elliptic partial differential equations on non-equidistant grids. *SIAM J. Sci. Comput.*, 29(4):1613–1636 (electronic), 2007.