# Closed-Form Multigrid Smoothing Factors for Lexicographic Gauss-Seidel*

L. Robert Hocking† and Chen Greif†

August 11, 2011

**Abstract**

This paper aims to present a unified framework for deriving analytical formulas for smoothing factors in arbitrary dimensions, under certain simplifying assumptions. To derive these expressions we rely on complex analysis and geometric considerations, using the maximum modulus principle and Möbius transformations. We restrict our attention to pointwise and block lexicographic Gauss-Seidel smoothers on a $d$-dimensional uniform mesh, where the computational molecule of the associated discrete operator forms a $2d + 1$ point star. In the pointwise case, the effect of a relaxation parameter is analyzed. Our results apply to any number of spatial dimensions, and are applicable to high-dimensional versions of a few common model problems with constant coefficients, including the Poisson and anisotropic diffusion equations, as well as a special case of the convection-diffusion equation. We show that in most cases our formulas, exact under the simplifying assumptions of Local Fourier Analysis, form tight upper bounds for the asymptotic convergence of geometric multigrid in practice. We also show that there are asymmetric cases where lexicographic Gauss-Seidel smoothing outperforms red-black Gauss-Seidel smoothing; this occurs for certain model convection-diffusion equations with high mesh Reynolds numbers.

**Keywords**: multigrid, smoothing factor, local Fourier analysis, elliptic partial differential equations, Gauss-Seidel

## 1 Introduction

In this paper we revisit a problem that is as old as the days of geometric multigrid [5, 6, 9]: the (analytical) computation of smoothing factors on a uniform mesh. To that end, consider the $d$-dimensional linear elliptic PDE

$$\mathcal{L}u = f,$$

*This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

†Department of Computer Science, The University of British Columbia, Vancouver BC, Canada, V6T 1Z4, rob.l.hocking@gmail.com, greif@cs.ubc.ca

with prescribed boundary conditions, discretized on a rectangular grid with uniform mesh spacing $h$. This yields a linear system of the form

$$\sum_{\vec{J}} \mathcal{L}^h_{\vec{I},\vec{J}} u^h_{\vec{J}} = f^h_{\vec{I}}, \tag{1.1}$$

where $\vec{I}$ and $\vec{J}$ vary over $\mathbb{Z}^d$. We will assume the associated computational molecule forms a $2d+1$ point star, or equivalently that $\mathcal{L}^h_{\vec{I},\vec{J}}$ is of the form

$$\mathcal{L}^h_{\vec{I},\vec{J}} = \begin{cases} a & \text{if } \vec{I} - \vec{J} = \vec{0} \\ -b_k^+ & \text{if } \vec{I} - \vec{J} = e_k \\ -b_k^- & \text{if } \vec{I} - \vec{J} = -e_k \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

The connection between $\mathcal{L}^h_{\vec{I},\vec{J}}$ and the computational molecule is illustrated in Figure 1. We assume throughout that $\mathcal{L}^h_{\vec{I},\vec{J}}$ satisfies

$$a \geq \sum_{k=1}^{d} (b_k^+ + b_k^-) \quad \text{and} \quad b_k^+, b_k^- > 0 \quad \text{for all} \quad k \in \{1, \ldots, d\}, \tag{1.3}$$

so that $[\mathcal{L}^h_{\vec{I},\vec{J}}]$ is a diagonally dominant M-matrix [8, p. 85]. Note that in the context of multigrid, there are cases (for example convection-diffusion) where this property may hold on the finest grid but not on coarser grids, potentially leading to difficulties with convergence.

The eigenfunctions of certain simple relaxation schemes form a complete set of Fourier modes, and the eigenvalues can be split into high frequencies and low frequencies according to their wave numbers. The smoothing factor is then defined as the maximal absolute value of the high frequency eigenvalues [1], [6, p. 104]. In particular, the smoothing factor of lexicographic Gauss-Seidel (GS-LEX), which we focus on in this paper, falls into this category provided one makes the simplifying assumption of an infinite grid, as in Local Fourier Analysis (LFA).

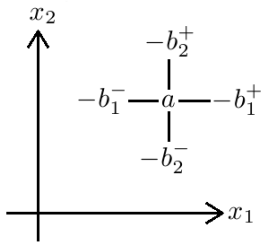


Figure 1: The computational molecule associated with the operator $\mathcal{L}^h$ $(d = 2)$.

For $d \leq 2$, smoothing factors of GS-LEX can be computed using standard techniques from multivariate calculus, and several results are available in closed form. On the other hand, for $d > 2$ the resulting system of equations is typically intractable, and the literature is much sparser with regard to closed-form formulas. In this case the smoothing factor may be computed numerically by scanning over a dense set of high frequencies, but this may be computationally expensive, especially if one seeks to identify a

trend with respect to parameters in the underlying PDE. An example here is the convection-diffusion equation, where one may wish to determine how the smoothing factor changes as a function of the mesh Reynolds numbers.

For red-black Gauss-Seidel (GS-RB), smoothing factors have been obtained for a broad class of *symmetric* operators of the form (1.2) in arbitrary dimensions and for both pointwise and block relaxations; see Yavneh [10]. In this paper we offer complementary analysis for GS-LEX in arbitrary dimensions, and include the asymmetric case, under certain simplifying assumptions. We show that in the strongly asymmetric setting GS-LEX can be a better smoother than GS-RB. This is in contrast to the symmetric case, where the latter is superior. We also analyze the effect of a relaxation parameter.

The remainder of this paper is organized as follows. Section 2 is devoted to the derivation of our main results. In section 3 we demonstrate the generality of our approach by applying it to a few examples and comparing against measured asymptotic convergence rates. We also present in this section a few comparisons of GS-LEX with GS-RB. Finally, in section 4 we draw some conclusions.

# 2 Smoothing Analysis

We will assume an ordering in which grid points are ordered according to the following rule: along dimension $k$, unknowns are ordered from $-e_k$ to $+e_k$ if $b_k^+ \leq b_k^-$, and from $+e_k$ to $-e_k$ otherwise. For convenience, let us define:

$$\begin{aligned} c_k &\equiv \max(b_k^+, b_k^-)/a \\ d_k &\equiv \min(b_k^+, b_k^-)/a. \end{aligned}$$

It is straightforward to show that the smoothing factor of pointwise lexicographic Gauss-Seidel predicted by Local Fourier Analysis (LFA) is given by

$$\mu^{pt} = \max_{\vec{\theta} \in \Theta^d} \left| \frac{\sum_{k=1}^d d_k e^{\imath \theta_k}}{1 - \sum_{k=1}^d c_k e^{-\imath \theta_k}} \right|, \tag{2.1}$$

where $\Theta^d = [-\pi, \pi]^d \backslash (-\pi/2, \pi/2)^d$ is the set of rough modes in $d$ dimensions and $\imath \equiv \sqrt{-1}$.

We restrict our analysis to the case

$$d_k = \alpha c_k, \tag{2.2}$$

where $\alpha \in (0, 1]$ is a constant independent of $k$. This restriction includes all cases where $[\mathcal{L}^h_{\vec{I},\vec{J}}]$ is symmetric, and other cases, for example the case of the convection-diffusion equation with all mesh Reynolds numbers equal.

It is well-known that pointwise smoothing is ineffective for highly anisotropic problems [6, p. 131]; one way to resolve this is the use of block smoothers [6, p. 134]. Such smoothers are defined by a partitioning of $\{1, \ldots, d\}$ into disjoint sets

$$\mathcal{I}_b \subset \{1, \ldots, d\} \quad \text{and} \quad \mathcal{I}_p = \{1, \ldots, d\} \backslash \mathcal{I}_b,$$

where coordinates belonging to $\mathcal{I}_b$ are relaxed simultaneously. To avoid degenerate cases, we assume both $\mathcal{I}_p$ and $\mathcal{I}_b$ are non-empty. In this case the smoothing factor is given by

$$\mu^{block} = \max_{\vec{\theta}\in\Theta^d} \left| \frac{\sum_{k\in\mathcal{I}_p} d_k e^{\imath\theta_k}}{1 - \sum_{k\in\mathcal{I}_p} c_k e^{-\imath\theta_k} - \sum_{k\in\mathcal{I}_b} (c_k e^{-\imath\theta_k} + d_k e^{\imath\theta_k})} \right|. \tag{2.3}$$

It is worth noting that the computational cost per iteration of a block smoother is higher than that of a point smoother. However, for highly anisotropic cases this overhead is typically small compared to the gains in convergence rates.

Our analysis shows that $\mu^{pt}$ depends only on $\alpha$ and any two of the quantities

$$c \equiv c_1 + c_2 + \ldots + c_d \qquad c_m \equiv \min(c_1, \ldots, c_d) \qquad c_r \equiv c - c_m, \tag{2.4}$$

regardless of the number of dimensions, $d$.

Similarly, in the case that $\mathcal{L}^h$ is symmetric, our analysis shows that the block smoothing factor $\mu^{block}$ depends on any two of the quantities

$$C^p = \sum_{k\in\mathcal{I}_p} c_k, \qquad C^p_m = \min_{k\in\mathcal{I}_p} c_k, \qquad C^p_r = C^p - C^p_m \tag{2.5}$$

as well as any two of $C^b$, $C^b_m$ and $C^b_r$ (which are defined analoguously).

Expressions for $\mu^{pt}$ and $\mu^{block}$ are given in Theorems 2.5 and 2.6 respectively.

## 2.1 Pointwise Relaxation

Starting from (2.1), we complex-conjugate the denominator and apply (2.2) to obtain

$$\mu^{pt} = \max_{\vec{\theta}\in\Theta^d} \left| \frac{\sum_{k=1}^d d_k e^{\imath\theta_k}}{1 - \sum_{k=1}^d c_k e^{-\imath\theta_k}} \right| = \max_{\vec{\theta}\in\Theta^d} \left| \frac{\alpha \sum_{k=1}^d c_k e^{\imath\theta_k}}{1 - \sum_{k=1}^d c_k e^{\imath\theta_k}} \right| \equiv \max_{\vec{\theta}\in\Theta^d} \left| \frac{\alpha g(\vec{\theta})}{1 - g(\vec{\theta})} \right|,$$

where $g(\vec{\theta}) \equiv \sum_{k=1}^d c_k e^{\imath\theta_k}$.

The key to our analysis lies in the observation that this may be rewritten as

$$\mu^{pt} = \max_{z\in g(\Theta^d)} |f(z)|,$$

where

$$f(z) \equiv \frac{\alpha z}{1 - z}$$

is analytic in the punctured plane $\mathbb{C}\backslash\{1\}$.

If the set $g(\Theta^d)$ is sufficiently well-behaved (a connected open set or the closure of one) and does not contain the point $z = 1$, the maximum principle [4, p. 88] applies and we can write

$$\mu^{pt} = \max_{z\in\partial g(\Theta^d)} |f(z)|,$$

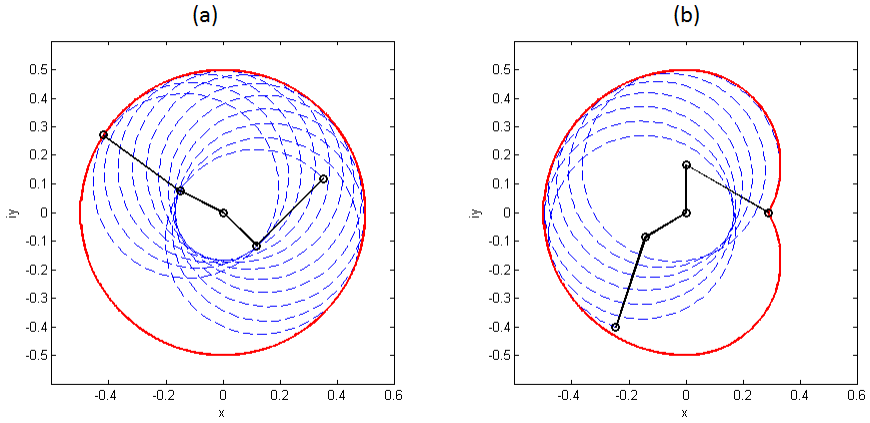


Figure 2: An illustration of Example 2.3: (a) If a double pendulum with arms of length $\frac{1}{6}$ and $\frac{1}{3}$ is allowed to swing freely, it remains inside a disk of radius 0.5. (b) If one arm is constrained to lie in the left half plane, the pendulum swings in the heart-shaped region shown. In both cases the pendulum is unable to reach a set of points surrounding the origin.

where $\partial g(\Theta^d)$ denotes the boundary of $g(\Theta^d)$.

However, computing $g(\Theta^d)$ explicitly is challenging because for certain values of $\{c_k\}$ it contains a hole in the vicinity of the origin - see Figure 2. Keeping track of this hole is difficult - we avoid the issue by proving we can replace $g(\Theta^d)$ with a simply-connected set $D$ equivalent to it in the following sense:

**Definition 2.1** *Let $A, B \subset \mathbb{C}$. We say $A \sim B$ if*

$$\sup_A |f| = \sup_B |f|$$

*for all functions $f$ analytic on $A \cup B$.*

The following observation establishes an important sufficient condition for two sets to be equivalent.

**Observation 2.2** *Let $A, B \subset \mathbb{C}$, and suppose $B$ is a connected open set, or the closure thereof. If $\partial B \subseteq A \subseteq B$, then $A \sim B$.*

To see this, let $f$ be any function analytic on $A \cup B = B$. We clearly have $\sup_{\partial B} |f| \leq \sup_A |f| \leq \sup_B |f|$, but the maximum modulus principle implies the equality of the leftmost and rightmost terms.

In light of Observation 2.2, our goal is to find a simply-connected set $D$ such that $\partial D \subseteq g(\Theta^d) \subseteq D$. This is done in Lemma 2.4, but first we build some intuition with a simple example.

**Example 2.3** *Suppose $d = 2$ and $\mathcal{L}^h$ arises from a centered differences discretization of the Laplacian-like operator $\mathcal{L} = 0.5\partial_{xx} + \partial_{yy}$, giving $a = 3$, $b_1^+ = b_1^- = 0.5$ and $b_2^+ = b_2^- = 1$. Then $c_1 = 1/6$, $c_2 = 1/3$, and*

$$g(\theta_1, \theta_2) = \frac{1}{6}e^{\imath\theta_1} + \frac{1}{3}e^{\imath\theta_2}.$$

*Geometrically, $g(\theta_1, \theta_2)$ may be viewed as a double pendulum with arms of length $1/6$ and $1/3$, making angles $\theta_k$ (k=1,2) with the x-axis. As $(\theta_1, \theta_2)$ varies over $[-\pi, \pi]^2$, this pendulum (and therefore the range of g) remains confined to the disk of radius $0.5$ shown in Figure 2(a). Furthermore, the boundary of the disk is swept out as the double pendulum completes a full revolution with a constant angle of $180^\circ$ between the arms. By Observation 2.2, $g([-\pi, \pi]^2)$ is equivalent to this disk in the sense of definition 2.1.*

*However, if $(\theta_1, \theta_2)$ is constrained to lie in $\Theta^2$, then $|\theta_k| \geq \pi/2$ for at least one $k \in \{1, 2\}$. In other words, at least one arm is constrained to lie in the left half plane. With this constraint, $g(\theta_1, \theta_2)$ now lives in the set shown in Figure 2(b) - the union of two disks and a half disk, as demonstrated in Figure 3. It is simple to show the boundary of this set is once again in the range of g, so this heart-shaped set provides us with the $D$ we are looking for.*

*Note that while $g(\Theta^2) \sim D$ it is not true that $g(\Theta^2) = D$, since $0 \in D$ but $|g| \geq \frac{1}{3} - \frac{1}{6} > 0$.*

**Lemma 2.4** *Let $B(r, z)$ denote the closed ball of radius $r \geq 0$ centered at $z \in \mathbb{C}$. Let $\mathbb{C}^-$ denote the set of all complex numbers with negative real part, and let $\sim$ be the relation in definition 2.1. Then*

$$\begin{aligned} g(\Theta^d) &\sim [B(c, 0) \cap \mathbb{C}^-] \cup B(c_r, \imath c_m) \cup B(c_r, -\imath c_m) \equiv D, \\ g([-\pi, \pi]^d) &\sim B(c, 0), \end{aligned}$$

*where $c$, $c_m$, and $c_r$ are defined in (2.4).*

**Proof.** By Observation 2.2 we clearly have $g([-\pi, \pi]^d) \sim B(c, 0)$, since $|g(\theta_1, \ldots, \theta_d)| \leq c$ and $g(\theta, \ldots, \theta) = ce^{\imath\theta}$. On the other hand, if $\vec{\theta} \in \Theta^d$ then there is a $j$ such that $|\theta_j| \geq \pi/2$. Assuming we have made some fixed choice of $j$ and set $\theta_j$ to a fixed value $\alpha$, the remaining variables $\{\theta_k\}_{k \neq j}$ range over $[-\pi, \pi]^{d-1}$. It follows that $g$ is confined to $B(c - c_j, c_j e^{\imath\alpha})$, and therefore

$$g(\Theta^d) \subseteq \bigcup_{|\alpha| \geq \frac{\pi}{2}} \bigcup_{j=1}^{d} B(c - c_j, c_j e^{\imath\alpha}).$$

But $B(c - c_j, c_j e^{\imath\alpha}) \subseteq B(c_r, c_m e^{\imath\alpha})$ for all $j$ since $|z - c_j e^{\imath\alpha}| \leq c - c_j$ implies

$$|z - c_m e^{\imath\alpha}| \leq |z - c_j e^{\imath\alpha}| + |c_j e^{\imath\alpha} - c_m e^{\imath\alpha}| \leq (c - c_j) + (c_j - c_m) = c_r.$$

Therefore

$$g(\Theta^d) \subseteq \bigcup_{|\alpha| \geq \frac{\pi}{2}} B(c_r, c_m e^{\imath\alpha}).$$

It can be shown that any $z \in B(c_r, c_m e^{\imath\alpha})$ obeying $\mathrm{Re}(z) \geq 0$ lies in $B(c_r, \imath c_m)$ if $\mathrm{Im}(z) \geq 0$ and $B(c_r, -\imath c_m)$ if $\mathrm{Im}(z) \leq 0$. Similarly, since $|z| \leq c$ we have $z \in B(c, 0) \cap \mathbb{C}^-$ if $\mathrm{Re}(z) < 0$. It follows that

$$g(\Theta^d) \subseteq D.$$

It is trivial to show $\partial D \subseteq g(\Theta^d)$, which completes the proof. □

Lemma 2.4 and Definition 2.1 allow us to conclude

$$\mu^{pt} = \max_{\partial D} |f| \tag{2.6}$$

provided $f$ is analytic in $D$, that is, provided we can show $1 \notin D$. To that end, we note that the diagonal dominance of $\mathcal{L}^h$ implies

$$\begin{aligned} c &= \frac{\sum_{k=1}^d \max(b_k^+, b_k^-)}{a} \\ &\leq \frac{\sum_{k=1}^d \max(b_k^+, b_k^-)}{\sum_{k=1}^d (b_k^+ + b_k^-)} < 1. \end{aligned}$$

The desired result follows from the observation

$$D \cap \mathbb{R} = [-c, \sqrt{c_r^2 - c_m^2}] \subseteq [-c, c).$$

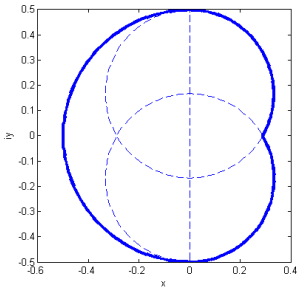


Figure 3: The heart-shaped set $D$ is the union of two disks and a half disk.

**Theorem 2.5** *The smoothing factor of pointwise GS-LEX applied to discrete operators of the form* (1.2) *obeying the constraints* (1.3) *and* (2.2) *is given by*

$$\mu^{pt}(c_m, c_r, \alpha) = \frac{\alpha c_r + \alpha\sqrt{c_m^2 + (c_m^2 - c_r^2)^2}}{1 + c_m^2 - c_r^2}, \tag{2.7}$$

*where $c_m$ and $c_r$ are defined in* (2.4) *and $\alpha$ is defined in* (2.2).

**Proof.** $\partial D$ is the union of the three semicircular arcs

$$\begin{aligned} S_1 &= \left\{ce^{\imath\theta} : \theta \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]\right\}, \quad S_2 = \left\{\imath c_m + c_r e^{\imath\theta} : \theta \in \left[-\sin^{-1}\left(\frac{c_m}{c_r}\right), \frac{\pi}{2}\right]\right\}, \\ S_3 &= \left\{z \in \mathbb{C} : \bar{z} \in S_2\right\}. \end{aligned}$$

By (2.6), it suffices to compute the maximum of the maxima of $|f|$ over $S_1$, $S_2$, and $S_3$. However, since $|f(\bar{z})| = |f(z)|$, the maxima over $S_2$ and $S_3$ are the same and hence we omit $S_3$ from consideration. Furthermore, it is trivial to show that the maximum of $|f|$ over $S_1$ is achieved at $\imath c$, the point of intersection

of $S_1$ and $S_2$. Therefore, the maximum of $|f|$ over $S_2$ is at least as large as the maximum over $S_1$ - hence we may omit $S_1$ as well.

It follows that the maximum of $|f|$ over $D$ is attained on the semicircle $S_2$. However, it is also true that the maximum is attained on the full circle $|z - \imath c_m| = c_r$, as this circle contains $S_2$ and is contained in $D$. It is more convenient to work with the full circle, so we conclude

$$\mu^{pt} = \max_{|z-\imath c_m|=c_r} |f(z)|. \tag{2.8}$$

Now, $f(z) = \frac{\alpha z}{1-z}$ is a *möbius transform*, and hence the image of a circle is either a circle or a line [4, p. 65]. Since the circle $|z-\imath c_m| = c_r$ does not contain any poles of $f$, it follows that $f(\{|z - \imath c_m| = c_r\})$ is a circle.

If we let $z_c \in \mathbb{C}$ and $r \geq 0$ denote the center and radius respectively of this circle, then from (2.8) we have

$$\mu^{pt} = \max_{z \in f(\{|z-\imath c_m|=c_r\})} |z| = |z_c| + r. \tag{2.9}$$

To find the parameters $z_c$ and $r$, we first write $f$ as a sequence of elementary Möbius transformations: $f = f_4 \circ f_3 \circ f_2 \circ f_1$, where $f_1(z) = z - 1$, $f_2(z) = z^{-1}$, $f_3(z) = -\alpha z$, and $f_4(z) = z - \alpha$. Then, starting with the circle $|z - \imath c_m| = c_r$, we track the changes in its center and radius as we compute its image under $f_1$, then compute the image of the result under $f_2$, and so on.

Since $f_1$, $f_3$, and $f_4$ are all either translations or dilations, the steps involving them are straightforward. For $f_2$, we make the observation that the image of a circle with radius $R$ and center $z_0$ under inversion is the circle with radius $R'$ and center $z_0'$ given by

$$z_0' = \frac{\bar{z}_0}{|z_0|^2 - R^2} \text{ and } R' = \frac{R}{\left||z_0|^2 - R^2\right|};$$

this fact can easily be derived from the argument found in [4, p. 66].

In the end we find

$$r = \frac{\alpha c_r}{1 + c_m^2 - c_r^2}$$

and

$$z_c = \alpha \frac{c_r^2 - c_m^2}{1 + c_m^2 - c_r^2} + \imath\alpha \frac{c_m}{1 + c_m^2 - c_r^2},$$

which completes the proof. □

Notice that the smoothing factor $\mu^{pt}$ is proportional to $\alpha$. Since the latter is small when the operator $\mathcal{L}^h$ is strongly asymmetric, this suggests that GS-LEX may be particularly effective in this regime. In particular, as $\alpha \to 0$ the part of $\mathcal{L}^h$ above its diagonal is converging to the zero matrix, and for $\alpha = 0$ we have that $\mathcal{L}^h$ is lower triangular and GS-LEX becomes a (direct) solver. See also Section 3 for further illustrations of this behavior.

## 2.2 SOR Relaxation

It is well-known that the smoothing factors of GS-RB, Jacobi-RB[1], and Jacobi can be greatly improved by incorporating a relaxation parameter [12, 6, 11]. For example, for the standard seven-point centered difference discretization of the three-dimensional Poisson problem, overrelaxation with $\omega = 1.15$ improves the smoothing factor of GS-RB (equivalent to Jacobi-RB in this specific case) from $\mu \approx 0.44$ to $\mu \approx 0.23$ [11]. For the same problem, the smoothing factor of Jacobi is improved from $\mu = 1$ (no convergence) to $\mu = 5/7$ by underrelaxation with $\omega = 6/7$ [6, p. 73].

We can analyze the effect of a relaxation parameter $\omega$ in GS-LEX smoothing by repeating the steps of Theorem 2.5 with the Möbius transformation

$$f_\omega(z) = \frac{\alpha z + \frac{1}{\omega} - 1}{\frac{1}{\omega} - z}$$

used in place of $f$. Note that we must assume $\omega < \frac{1}{\sqrt{c_r^2 - c_m^2}}$ in order to ensure that $f_\omega$ has no poles in $D$.

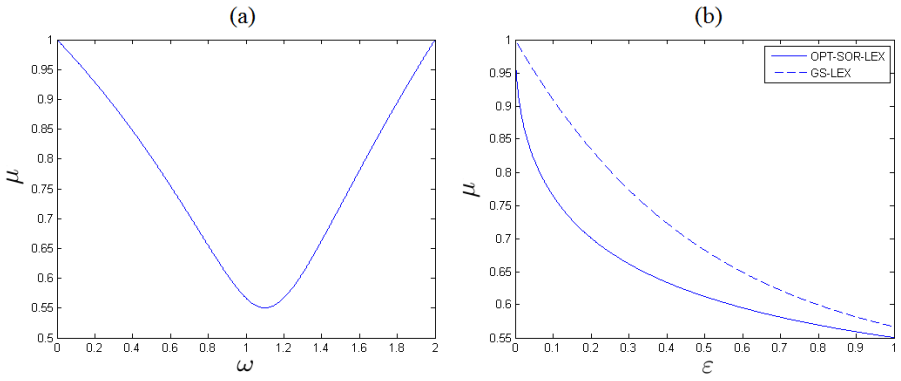


Figure 4: Left: Smoothing factor of SOR-LEX with relaxation parameter $\omega$, as a function of $\omega$, for the standard seven-point centered differences discretization of the 3D Poisson problem. Right: Smoothing factors of GS-LEX and optimal SOR-LEX applied to the anisotropic diffusion problem $-\varepsilon u_{xx} - u_{yy} - u_{zz} = f$, as a function of $\varepsilon$, assuming seven-point centered differences discretization.

Making use of the factorization $f_\omega = f_4 \circ f_3 \circ f_2 \circ f_1$, where $f_1(z) = z - \frac{1}{\omega}$, $f_2(z) = z^{-1}$, $f_3(z) = \left(1 - \frac{1+\alpha}{\omega}\right) z$, $f_4(z) = z - \alpha$, we obtain

$$\mu_{SOR}^{pt}(\omega) = \frac{c_r \left|1 - \frac{1+\alpha}{\omega}\right| + \sqrt{c_m^2 \left(1 - \frac{1+\alpha}{\omega}\right)^2 + \left(\frac{1}{\omega}\left(1 - \frac{1}{\omega}\right) + \alpha(c_m^2 - c_r^2)\right)^2}}{\frac{1}{\omega^2} + c_m^2 - c_r^2}. \tag{2.10}$$

[1] Jacobi-RB consists of a Jacobi sweep over the red points, followed by a Jacobi sweep over the black points using the updated values at red points; see for example [6, p 173].

Figure 4(a) shows a plot of $\mu^{pt}_{SOR}$ as a function of $\omega \in [0,2]$ for the 3D Poisson problem mentioned above ($c_m = 1/6$, $c_r = 1/3$, $\alpha = 1$). The optimal smoothing factor $\mu \approx 0.551$ is attained at $\omega \approx 1.1$, a modest improvement over the smoothing factor $\mu \approx 0.567$ obtained when $\omega = 1$. This finding supports the conclusion [6, p. 105] that for GS-LEX, the inclusion of a relaxation parameter is not necessarily worth the extra work per iteration (two operations per point per relaxation sweep).

In Figure 4(b), the anisotropic problem $-\varepsilon u_{xx} - u_{yy} - u_{zz} = f$ is discretized using standard seven-point centered differences, and the theoretical smoothing factors of both GS-LEX and optimal SOR-LEX (obtained by numerical minimization of (2.10)) are plotted against $\varepsilon$.

## 2.3 Block Relaxation

In this section we analyze the smoothing properties of block GS-LEX relaxation. To keep the algebra manageable, we restrict the scope of our analysis to the case that $[\mathcal{L}^h_{\vec{I},\vec{J}}]$ is symmetric ($\alpha = 1$). Our main result is Theorem 2.6, which establishs a connection between block and pointwise smoothing factors.

Since $\alpha = 1$, we have $c_k = d_k$ for all $k$ - substituting this into equation (2.3), complex-conjugating the denominator and then simplifying, we obtain

$$\begin{aligned} \mu^{block} &= \max_{\vec{\theta}\in\Theta^d} \left| \frac{\sum_{k\in\mathcal{I}_p} c_k e^{\imath\theta_k}}{1 - 2\sum_{k\in\mathcal{I}_b} c_k \cos\theta_k - \sum_{k\in\mathcal{I}_p} c_k e^{\imath\theta_k}} \right| \\ &= \max_{(z,x)\in G(\Theta^d)} \left| \frac{z}{1-x-z} \right|, \end{aligned}$$

where $G : \Theta^d \to \mathbb{C}\times\mathbb{R}$ is given by $G(\vec{\theta}) = (G_1(\vec{\theta}), G_2(\vec{\theta}))$ with

$$G_1(\vec{\theta}) = \sum_{k\in\mathcal{I}_p} c_k e^{\imath\theta_k} \text{ and } G_2(\vec{\theta}) = 2\sum_{k\in\mathcal{I}_b} c_k \cos\theta_k.$$

Defining $z' = z/(1-x)$, this becomes

$$\mu^{block} = \max_{z'\in\Lambda} \left| \frac{z'}{1-z'} \right| = \max_{z'\in\Lambda} |f(z')|, \tag{2.11}$$

where $f$ is defined in section 2.1 and

$$\Lambda \equiv \{z/(1-x) : (z,x) \in G(\Theta^d)\}.$$

We will show that $\Lambda$ is $\sim$ (in the sense of Definition 2.1) to the union of a scaled copy of the set $D$ from Lemma 2.4 and a particular ball centered at the origin. This fact will allow us to express $\mu^{block}$ in terms of $\mu^{pt}$.

If $\vec{\theta} \in \Theta^d$, then there is a $j$ such that $|\theta_j| \geq \pi/2$. If $j \in \mathcal{I}_p$, it follows that

$$\{\theta_k\}_{k\in\mathcal{I}_p} \in \Theta^{|\mathcal{I}_p|} \text{ and } \{\theta_k\}_{k\in\mathcal{I}_b} \in [-\pi,\pi]^{|\mathcal{I}_b|},$$

while if $j \in \mathcal{I}_b$, then

$$\{\theta_k\}_{k\in\mathcal{I}_p} \in [-\pi,\pi]^{|\mathcal{I}_p|} \text{ and } \{\theta_k\}_{k\in\mathcal{I}_b} \in \Theta^{|\mathcal{I}_b|}.$$

The function $G_1$ has the same form as $g$ of Section 2.1 with the variables $\{\theta_k\}_{k=1}^d$ replaced by $\{\theta_k\}_{k\in\mathcal{I}_p}$. Thus, Lemma 2.4 applies with $C^p$, $C^p_m$ and $C^p_r$ playing the roles of $c$, $c_m$, $c_r$:

$$G_1(\Theta^d) \sim \begin{cases} D(C^p_m, C^p_r) & \text{if } k \in \mathcal{I}_p, \\ B(C^p, 0) & \text{otherwise.} \end{cases}$$

At the same time we have

$$G_2(\Theta^d) = \begin{cases} [-2C^b, 2C^b] & \text{if } k \in \mathcal{I}_p, \\ [-2C^b, 2C^b_r] & \text{otherwise.} \end{cases}$$

Assuming $j \in \mathcal{I}_p$, since $G_1$ and $G_2$ depend on disjoint variables we have

$$\Lambda\Big|_{j\in\mathcal{I}_p} = \bigcup_{x\in G_2(\Theta^d)\Big|_{j\in\mathcal{I}_p}} \frac{1}{1-x} G_1(\Theta^d)\Big|_{j\in\mathcal{I}_p} \sim \bigcup_{x\in[-2C^b,2C^b]} \frac{1}{1-x} D(C^p_m, C^p_r).$$

Note that if $2C^b < 1$, then for all $x \in [-2C^b, 2C^b]$ we have

$$\frac{1}{1-x} D(C^p_m, C^p_r) \subseteq \frac{1}{1-2C^b} D(C^p_m, C^p_r).$$

It follows from the diagonal dominance of $\mathcal{L}^h$ and the condition $\alpha = 1$ that $c \le 0.5$. Since $C^b < c$, we have $2C^b < 1$ as desired, and therefore

$$\Lambda\Big|_{j\in\mathcal{I}_p} \sim \frac{1}{1-2C^b} D(C^p_m, C^p_r) = D\Big(\frac{C^p_m}{1-2C^b}, \frac{C^p_r}{1-2C^b}\Big). \tag{2.12}$$

A similar analysis with $j \in \mathcal{I}_b$ shows

$$\Lambda\Big|_{j\in\mathcal{I}_b} \sim B\left(\frac{C^p}{1-2C^b_r}, 0\right).$$

**Theorem 2.6** *The smoothing factor of block GS-LEX applied to* symmetric *discrete operators of the form* (1.2) *obeying the constraints* (1.3) *and* (2.2) *is given by*

$$\mu^{block} = \max\left(\mu^{pt}(\tilde{c}_m, \tilde{c}_r, \alpha = 1), \frac{C^p}{1 - C^p - 2C^b_r}\right), \tag{2.13}$$

*where*

$$\tilde{c}_m = \frac{C^p_m}{1-2C^b}, \;\; \tilde{c}_r = \frac{C^p_r}{1-2C^b},$$

*and* $C^p, C^p_m, C^p_r, C^b, C^b_r$ *are defined in* (2.5).

**Proof.** Clearly the maximum of $|f|$ over $\Lambda$ is the maximum of the separate maxima over $\Lambda\Big|_{j\in\mathcal{I}_p}$ and $\Lambda\Big|_{j\in\mathcal{I}_b}$. By (2.12) and Theorem 2.5 we have

$$\max_{z\in\Lambda|_{j\in\mathcal{I}_p}} |f(z)| = \mu^{pt}\left(\tilde{c}_m, \tilde{c}_r, 1\right),$$

provided $f$ is analytic in $D(\tilde{c}_m, \tilde{c}_r)$. From the inequalities $C^p, C^b < c \leq 1/2$ it is easy to show $\tilde{c} \equiv \tilde{c}_m + \tilde{c}_r \leq 1$, and the desired result follows in the same way as in Section 2.1.

Similarly, since $C_r^b, C^p < 1/2$ we have

$$1 \notin B\left(\frac{C^p}{1-2C_r^b}, 0\right) \sim \Lambda\Big|_{j\in\mathcal{I}_b}.$$

It follows from the maximum modulus principle that $f(z)$ attains its maximum modulus over $\Lambda\Big|_{j\in\mathcal{I}_b}$ at $z = C^p/(1-2C_r^b)$, and therefore

$$\max_{z\in\Lambda|_{j\in\mathcal{I}_b}} |f(z)| = \frac{C^p}{1-C^p-2C_r^b},$$

which completes the proof. □

Theorem 2.6 shows that applying a block smoother can be equivalent to applying a point-smoother to a related problem.

# 3 Examples

In this section we apply Theorems 2.5 and 2.6 to a few examples, obtaining formulas for the smoothing factors of standard discretizations of several common PDEs; these are then compared with the measured asymptotic convergence rate of multigrid with GS-LEX smoothing. We also include measured convergence rates with GS-RB smoothing, and show that while the latter has better smoothing properties for the symmetric case [10], lexicographic smoothing is more effective in the strongly asymmetric setting of the convection-dominated convection-diffusion equation. All experiments are done on a rectangular domain with homogeneous Dirichlet boundary conditions. Unless stated otherwise, we use the Galerkin coarse grid operator $\mathcal{L}^H = I_h^{2h}\mathcal{L}^h I_{2h}^h$ in our experiments, where $I_{2h}^h$ is the prolongation operator with linear interpolation, and $I_h^{2h} = 2^d(I_{2h}^h)^T$ is the restriction operator with full weighting. The asymptotic convergence rate $\rho$ is estimated from the sequence of residuals $\{r^{(k)}\}_{k=0}^m$ using the identity

$$\rho = \lim_{m\to\infty} \sqrt[m]{\frac{\|r^{(m)}\|_2}{\|r^{(0)}\|_2}}$$

and approximating the limit with a sufficiently large value of $m$ (see [6, p. 54] for justification).

| relaxation/dimension | $d=1$ | $d=2$ | $d=3$ |
|---|---|---|---|
| point relaxation | 0.447 | 0.5 | 0.567 |
| line relaxation | | 0.447 | 0.5 |
| plane relaxation | | | 0.447 |

Table 1: Smoothing factors of GS-LEX with point, line, and plane relaxation for the Poisson problem in 1D, 2D, and 3D. Moving from point to line or from line to plane relaxation is equivalent to reducing the dimension by one.

**Example 3.1** *Suppose the d-dimensional Poisson problem*

$$-\Delta u = f$$

*is discretized using centered differences on a uniform grid. Theorem 2.5 gives a general formula for the smoothing factor of pointwise GS-LEX, namely*

$$\mu^{pt}_{poisson}(d) = \frac{2(d-1)+\sqrt{d^2-4d+8}}{3d+2}.$$

*It can be verified that this reduces to the known values $1/\sqrt{5}$ and $1/2$ for $d=1$ and $d=2$ respectively [1], but also yields the exact expression*

$$\mu^{pt}_{poisson}(3) = \frac{4+\sqrt{5}}{11} \approx 0.5669,$$

*which as far as we know has not appeared in the literature in closed form.*

*Similarly, Theorem 2.6 provides a general formula for smoothing factors of block GS-LEX applied to the d-dimensional poisson problem. If each block is a k-dimensional subproblem, then we have*

$$\mu^{block}_{poisson}(d,k) = \begin{cases} \mu^{pt}_{poisson}(d-k) & \text{if } d-k<3, \\ \frac{d-k}{2+d-k} & \text{otherwise;} \end{cases}$$

*see Table 1.*

**Example 3.2** *Suppose the anisotropic steady-state diffusion problem*

$$-\rho_1 u_{x_1x_1} - \rho_2 u_{x_2x_2} - \ldots - \rho_d u_{x_dx_d} = f \tag{3.1}$$

*is discretized over $\mathbb{R}^d$ with standard second-order centered differences. We then obtain a linear system with $\mathcal{L}^h$ of the form (1.2) where $b_k^+ = b_k^- = \rho_k$, $a = 2\sum_{k=1}^d \rho_k$.*

*In particular, if $d=3$ and $\rho_1=\rho_2=1$ while $\rho_3=\varepsilon\in(0,1]$, then*

$$\mu^{pt} = \frac{4+\sqrt{5\varepsilon^2-4\varepsilon+4}}{6+5\varepsilon}.$$

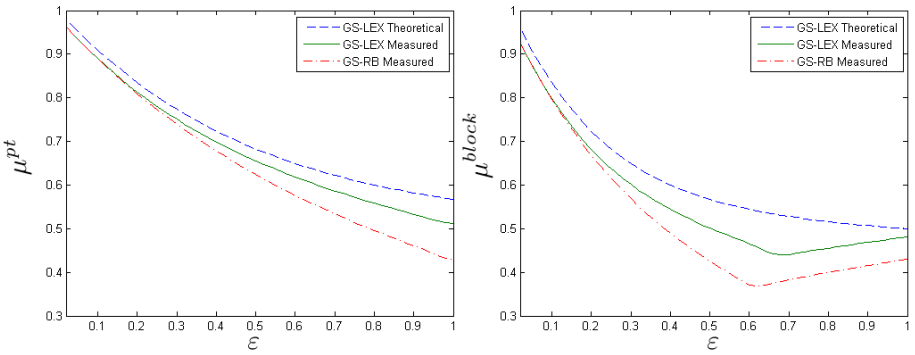


Figure 5: Theoretical smoothing factors of GS-LEX vs. experimental asymptotic convergence rates of two-level multigrid with GS-LEX and GS-RB smoothers, applied to the steady-state diffusion equation $-u_{xx} - u_{yy} - \varepsilon u_{zz} = f$. Pointwise relaxation appears on the left, while $x$-oriented line relaxation is depicted on the right. The discretization is on a $23 \times 23 \times 23$ grid with centered differences.

| $\varepsilon$ | 1.00 | 0.50 | 0.10 | 0.05 |
|---|---|---|---|---|
| $\mu^2$ | 0.25 | 0.32 | 0.70 | 0.83 |
| $V(1,1)$ | 0.12 | 0.27 | 0.68 | 0.81 |

Table 2: Comparison of predicted smoothing factors of GS-LEX with measured convergence rates of multigrid $V(1,1)$-cycles, for the PDE $-\varepsilon u_{xx} - u_{yy} = f$. The finest grid is $1023 \times 1023$, while the coarsest is $1 \times 1$. Various values of $\varepsilon$ are considered.

*Note that $\mu^{pt} \to 1$ as $\varepsilon \to 0$. If line relaxation is performed in which unknowns along the $x$ or $y$ directions are relaxed simultaneously, one obtains*

$$\mu^{line} = \frac{2 + \sqrt{5\varepsilon^2 - 2\varepsilon + 1}}{3 + 5\varepsilon}.$$

*It is interesting to note that $\mu^{line}$ above is identical to $\mu^{pt}$ in the lower dimensional case $d = 2$ and $\rho_1 = 1$, $\rho_2 = \varepsilon$. Figure 5 provides a comparison between the formulas for $\mu^{pt}$ and $\mu^{line}$ and measured asymptotic converge rates of two-level multigrid on a $23 \times 23 \times 23$ grid, for both GS-LEX and GS-RB. The graphs illustrate that GS-RB is superior to GS-LEX in this symmetric case. We also consider multigrid $V$-cycles with one pre and one post GS-LEX smoothing step, applied to model problem (3.1) with $d = 2$, $\rho_1 = 1$, $\rho_2 = \varepsilon$ on a $1023 \times 1023$ grid. These are compared against theoretical smoothing factors in Table 2.*

**Example 3.3** *Suppose the constant coefficient convection-diffusion problem*

$$-\Delta u + \vec{w} \cdot \nabla u = f$$

| | centered differences, $\gamma \in (-1,1)$ | upwinding, $\gamma \geq 0$ |
|---|---|---|
| $d=1$ | $\frac{1-\lvert\gamma\rvert}{\sqrt{4+(1+\lvert\gamma\rvert)^2}}$ | $\frac{1}{\sqrt{8\gamma^2+12\gamma+5}}$ |
| $d=2$ | $\frac{1-\lvert\gamma\rvert}{2}$ | $\frac{1}{2(1+\gamma)}$ |
| $d=3$ | $\frac{(1-\lvert\gamma\rvert)\left(2+\sqrt{1+\left(\frac{1+\lvert\gamma\rvert}{2}\right)^2}\right)}{6-\frac{1}{2}(1+\lvert\gamma\rvert)^2}$ | $\frac{2+\sqrt{1+\frac{1}{4}\left(1+\frac{\gamma}{1+\gamma}\right)^2}}{(1+\gamma)\left(6-\frac{1}{2}\left(1+\frac{\gamma}{1+\gamma}\right)^2\right)}$ |

Table 3: Smoothing factors of pointwise GS-LEX applied to two common discretizations of the convection-diffusion equation with all mesh Reynolds numbers equal to $\gamma$, for $d = 1, 2, 3$.

*is discretized on a uniform rectangular grid with mesh spacing $h$. We define the $d$ mesh Reynolds numbers by $\gamma_k = w_k h/2$, and assume they are all equal, that is*

$$\gamma_1 = \gamma_2 = \ldots = \gamma_d = \gamma.$$

*If we discretize this PDE using centered differences, we obtain $a = 2d$, $b_k^+ = 1-\gamma$, $b_k^- = 1+\gamma$ for all $k$. Provided $|\gamma| < 1$, Equation* (1.3) *is satisfied and Theorem 2.5 gives*

$$\mu^{pt} = (1-|\gamma|)\left[\frac{(d-1)+\sqrt{1+(1+|\gamma|)^2(1-\frac{d}{2})^2}}{2d+(1+|\gamma|)^2(1-\frac{d}{2})}\right].$$

*On the other hand, if we use first order upwinding, then (assuming for simplicity that $\gamma \geq 0$) we obtain $a = 2d(1+\gamma)$, $b_k^+ = 1$, $b_k^- = 1+2\gamma$ for all $k$. The constraints* (1.3) *are satisfied for all $\gamma \geq 0$ and*

$$\mu^{pt} = \frac{1}{1+\gamma}\left[\frac{(d-1)+\sqrt{1+\left(1+\frac{\gamma}{1+\gamma}\right)^2\left(1-\frac{d}{2}\right)^2}}{2d+\left(1+\frac{\gamma}{1+\gamma}\right)^2\left(1-\frac{d}{2}\right)}\right].$$

*Table 3 lists the above smoothing factors for the cases of interest $d = 1, 2, 3$.*

*For asymmetric problems such as convection-diffusion, Galerkin coarse grid operators obtained by linear interpolation and full weighting may cause unstable coarse grid discretizations, leading to multigrid convergence problems, especially as the number of levels increases [7]. Indeed, in Table 4 we observe divergence for multigrid with three or more levels for upwinding in the convection-dominated regime; for two levels, however, we see a good agreement between measured convergence rates and the predicted smoothing factor. We note that there are ways of dealing with these instabilities within the Galerkin framework; see for example [3].*

*The coarse grid operator $\mathcal{L}^H$ may also be defined by direct discretization of the underlying PDE on the coarse mesh. For upwinding with constant coefficients, this yields a stable discretization on all grid levels, and we can expect*

| $\gamma$ | 0.00 | 0.33 | 0.66 | 1.00 | 3.00 | 5.00 |
|---|---|---|---|---|---|---|
| $\mu$ | 0.50 | 0.38 | 0.30 | 0.25 | 0.13 | 0.08 |
| $2L$ | 0.39 | 0.36 | 0.30 | 0.25 | 0.13 | 0.08 |
| $3L$ | 0.41 | 0.36 | 0.31 | NC | NC | NC |
| $4L$ | 0.41 | NC | NC | NC | NC | NC |

Table 4: Comparison of the smoothing factor $\mu$ of GS-LEX with the measured asymptotic convergence rate of multigrid applied to the PDE $-u_{xx}-u_{yy}+\sigma u_x+\sigma u_y = f$, discretized uniformly. The problem is discretized using upwinding on a $1023\times1023$ grid. The Galerkin coarse grid operator is used. $W(1,0)$-cycles are used, and various values of $\gamma = \sigma h/2$, where $h$ is the grid spacing on the finest mesh, are considered. 'NC' stands for no convergence. '$nL$' (with $n = 2, 3, 4$) signifies the number of levels.

*multigrid methods to converge. However, it has been shown that in this setting the coarse grid correction becomes less effective as convection becomes more dominant. In particular, the reduction factor of certain smooth modes increases to a non-negligible constant, which grows as the number of levels increases [2]. In the convection-dominated regime where the smoothing factor becomes smaller than this constant, it is these smooth modes, rather than the rough modes, that dominate the error asymptotically. Consequently, it is this constant, rather than the smoothing factor, that determines the convergence rate. Therefore, we expect that the asymptotic convergence rate of multigrid will follow the trend indicated by our smoothing analysis in the diffusion-dominated regime, while tending to a constant in the convection-dominated regime; see Table 5. We also include GS-RB W-cycles in the table, to illustrate the superiority of GS-LEX in this case.*

| $\gamma$ | 0.00 | 1.00 | 2.00 | 3.00 | 4.00 | 5.00 |
|---|---|---|---|---|---|---|
| $\mu$ | 0.50 | 0.25 | 0.17 | 0.13 | 0.10 | 0.08 |
| $2L$ | 0.37 | 0.25 | 0.24 | 0.27 | 0.29 | 0.30 |
| $V(1,0)$ | 0.42 | 0.56 | 0.54 | 0.52 | 0.50 | 0.48 |
| $W(1,0)_{LEX}$ | 0.37 | 0.29 | 0.34 | 0.35 | 0.36 | 0.36 |
| $W(1,0)_{RB}$ | 0.23 | 0.37 | 0.46 | 0.59 | 0.71 | 0.79 |

Table 5: Comparison of the smoothing factor $\mu$ of GS-LEX with measured convergence rates of multigrid $V(1,0)$ and $W(1,0)$-cycles, for the PDE $-u_{xx} - u_{yy} + \sigma(u_x + u_y) = f$, discretized uniformly. The finest grid is $1023 \times 1023$, while the coarsest is $1 \times 1$. The coarse grid operator is obtained by direct discretization of the PDE on the coarse mesh. Various values of $\gamma = \sigma h/2$ are considered, where $h$ is the grid spacing of the finest mesh. '$2L$' stands for a two-level scheme, whereas $V(1,0)$ and $W(1,0)$ signify $V$ and $W$-cycles, respectively, with one pre-smoothing sweep and no post-smoothing.

*For the centered difference discretization of the convection-diffusion problem,*

| $\gamma$ | 0.00 | 0.20 | 0.40 | 0.60 | 0.80 | 0.99 |
|---|---|---|---|---|---|---|
| $\mu$ | 0.50 | 0.40 | 0.30 | 0.20 | 0.10 | 0.005 |
| $2L$ | 0.37 | 0.36 | 0.30 | 0.21 | 0.10 | 0.005 |
| $3L$ | 0.37 | 0.36 | 0.30 | NC | NC | NC |
| $4L$ | 0.37 | 0.36 | NC | NC | NC | NC |
| $5L$ | 0.37 | NC | NC | NC | NC | NC |

Table 6: Comparison of the smoothing factor $\mu$ of GS-LEX with the measured asymptotic convergence rate of two-level and multi-level multigrid applied to the PDE $-u_{xx} - u_{yy} + \sigma(u_x + u_y) = f$, discretized uniformly. The problem is discretized using centered differences on a $1023 \times 1023$ grid. The coarse grid operator is obtained by direct discretization of the PDE on the coarse mesh. $W(1,0)$-cycles are used, and various values of $\gamma = \sigma h/2$, where $h$ is the grid spacing on the finest mesh, are considered. 'NC' stands for no convergence.

*our smoothing analysis applies to the case where the mesh Reynolds number $\gamma$ is less than one in magnitude. But in a multigrid setup, we have different mesh Reynolds numbers on different grids - for our analysis to apply, they must all obey this constraint. Therefore, if $L$ denotes the number of levels in our algorithm and $\gamma$ denotes the mesh Reynolds number on the finest grid, what we really require is $|\gamma| < 2^{2-L}$.*

*In Table 6, we compare our predictions with measured convergence rates for various values of $L$; $W(1,0)$-cycles are applied. In practice, we can get away with a mesh Reynolds number moderately larger than 1 on the coarsest grid and still maintain expected convergence rates.*

**Example 3.4** *Suppose the time-dependent diffusion problem*

$$u_t = \rho_1 u_{x_1 x_1} + \rho_2 u_{x_2 x_2} + \ldots + \rho_d u_{x_d x_d}$$

*is discretized on a uniform mesh using backward Euler in time and centered differences in space, with multigrid used to solve the resulting linear system at each time step. If $\tau$ denotes the timestep size, then $\mu^{pt}$ follows from Theorem 2.5 by setting*

$$c_r = \frac{\rho - \rho_m}{2\rho + h^2/\tau}, \qquad c_m = \frac{\rho_m}{2\rho + h^2/\tau}, \qquad \alpha = 1.$$

*where $\rho = \sum_{k=1}^{d} \rho_k$ and $\rho_m = \min(\{\rho_k\})$.*

*In particular, in the isotropic case $\rho_1 = \rho_2 = \ldots = \rho_d = 1$, we obtain*

$$\mu^{pt} = \frac{(d-1)(2d + h^2/\tau) + \sqrt{(2d + h^2/\tau)^2 + d^2(2-d)^2}}{(2d + h^2/\tau)^2 + d(2-d)}.$$

*When $d = 2$ this reduces to the particularly simple expression*

$$\mu^{pt} = \frac{1}{2 + \frac{h^2}{2\tau}}.$$

*Note the (expected) behavior of the smoothing factor as a function of $\tau$: when $h$ is fixed and $\tau$ gets smaller, the smoothing factor becomes smaller too.*

# 4 Conclusions

Using results from complex analysis, primarily the maximum modulus principle and properties of Möbius transformations, we have derived closed-form expressions for the smoothing factors of lexicographic pointwise and block Gauss-Seidel (Theorems 2.5 and 2.6). An extension of Theorem 2.5 that incorporates a relaxation parameter is provided in (2.10). In the pointwise case our results are valid for general operators of the form (1.2) satisfying the constraints (1.3) and (2.2), whereas in the block case we require the additional assumption of symmetry. In some cases, block smoothing on a high dimensional problem has the same smoothing factor as pointwise smoothing on a related lower dimensional problem.

Our analysis provides smoothing factors for, among other equations, the following:

- Pointwise GS-LEX applied to the $d$-dimensional *anisotropic steady-state diffusion* equation $-\sum_{k=1}^{d} \rho_k u_{x_k x_k} = f$, discretized with centered differences.
- Pointwise and block GS-LEX applied to the $d$-dimensional *Poisson* equation $-\Delta u = f$, discretized with centered differences.
- Pointwise GS-LEX applied to the $d$-dimensional constant coefficient *convection-diffusion* equation $-\Delta u + \vec{w} \cdot \nabla u = f$, discretized with centered differences or upwinding, and all mesh Reynolds numbers equal.
- Pointwise GS-LEX applied to the linear systems arising in each timestep of the solution of the $d$-dimensional *time-dependent diffusion* equation $u_t = \sum_{k=1}^{d} \rho_k u_{x_k x_k}$, discretized with centered differences in space and backward Euler in time.

We have also observed that lexicographic Gauss-Seidel smoothing is effective for equations with strong asymmetry. In particular, for the constant coefficient convection-diffusion equation with equal mesh Reynolds numbers we have shown for upwind discretizations that GS-LEX has a smaller smoothing factor than GS-RB in the convection-dominated regime.

# Acknowledgments

We are grateful to two anonymous referees for their constructive comments and suggestions.

# References


[1] Achi Brandt. Multi-level adaptive solutions to boundary-value problems. *Mathematics of Computation*, 31(138):333–390, 1977.

[2] Achi Brandt and Irad Yavneh. Accelerated multigrid convergence and high-reynolds recirculating flows. *SIAM J. Sci. Comput.*, 14(3):607–626, 1993.

[3] P.M. de Zeeuw. Matrix-dependent prolongations and restrictions in a black-box multigrid solver. *Journal of computational and applied mathematics*, 33(1):1–27, 1990.

[4] Theodore W. Gamelin. *Complex Analysis*. Springer, 2001.

[5] Wolfgang Hackbusch. *Multigrid methods and applications*, volume 4 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1985.

[6] U. Trottenberg, C. W. Oosterlee, and A. Schüller. *Multigrid*. Academic Press Inc., San Diego, CA, 2001.

[7] E. J van Asselt and P.M. de Zeeuw. The convergence rate of multi-level algorithms applied to the convection-diffusion equation. *SIAM journal on scientific and statistical computing*, 6(2):492–503, 1985.

[8] Richard S. Varga. *Matrix iterative analysis*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1962.

[9] Pieter Wesseling. *An introduction to multigrid methods*. Pure and Applied Mathematics (New York). John Wiley & Sons Ltd., Chichester, 1992.

[10] Irad Yavneh. Multigrid smoothing factors of red-black gauss-seidel applied to a class of elliptic operators. *SIAM J. Numer. Anal.*, 32(6):1126–1138, 1995.

[11] Irad Yavneh. On red-black SOR smoothing in multigrid. *SIAM J. Sci. Comput.*, 17(1):180–192, 1996.

[12] H. Bin Zubair, C.W. Oosterlee, and R. Wienands. Multigrid for high-dimensional elliptic partial differential equations on non-equidistant grids. *SIAM J. Sci. Comput.*, 29(4):1613–1636, 2007.