Learning to Optimize Multigrid PDE Solvers -Supplementary Material-

Daniel Greenfeld [1] **Meirav Galun** [1] **Ron Kimmel** [2] **Irad Yavneh** [2] **Ronen Basri** [1]

Below we prove Theorem 1. The proof is based on two supporting lemmas. We begin with some mathematical terms.

Consider the $n \times n$ block-circulant matrix of the following form, where $n = kb$ and all numbering of rows, columns, blocks, etc., starts from 0 for convenience

$$A = \begin{pmatrix} A^{(0)} \\ A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(b-1)} \end{pmatrix},$$

where the blocks $A^{(m)}, m = 0, \ldots, b-1$ are $k \times n$ real or complex matrices whose elements satisfy

$$A^{(m)}_{l,j} = A^{(m-1)}_{l,\ \bmod (j-k,n)},\ m = 1, \ldots, b-1 \tag{1}$$

and hence $A_{l,j} = A_{\bmod (l-k,n),\ \bmod (j-k,n)}$. Here, we are adopting the MATLAB form $\mathrm{mod}(x, y) =$ "x modulo y", i.e., the remainder obtained when dividing integer x by integer y. Below, we continue to use l and j to denote row and column numbers, respectively, and apply the decomposition:

$$l = l_0 + tk, \quad j = j_0 + sk\,, \tag{2}$$

where $l_0 = \mathrm{mod}(l, k)$, $t = \lfloor \frac{l}{k} \rfloor$, $j_0 = \mathrm{mod}(j, k)$, $s = \lfloor \frac{j}{k} \rfloor$. Note that $l, j \in \{0, ..., n-1\}$; $l_0, j_0 \in \{0, ..., k-1\}$; $t, s \in \{0, ..., b-1\}$.

Let the column vector

$$v_m = \left[1, e^{i\frac{2\pi m}{n}}, \ldots, e^{i\frac{2\pi mj}{n}}, \ldots, e^{i\frac{2\pi m(n-1)}{n}}\right]^*$$

denote the unnormalized mth Fourier component of dimension n, for $m = 0, \ldots, n-1$. Let W denote the $n \times n$ matrix whose nonzero values are comprised of the elements of the first b Fourier components as follows:

$$W_{l,j} = \delta_{l_0,j_0} v_s(l)\,, \tag{3}$$

[1]Weizmann Institute of Science, Rehovot, Israel [2]Technion, Israel Institute of Technology, Haifa, Israel. Correspondence to: Daniel Greenfeld <daniel.greenfeld@weizmann.ac.il>.

Proceedings of the 36^{th} International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.

where $v_s(l)$ denotes the lth element of v_s, and δ is the Kronecker delta. An example for W, with $k = 3$ and $b = 4$, is given in Fig. 1.

Lemma 1. $\frac{1}{\sqrt{b}}W$ *is a unitary matrix.*

Proof. Let W_j and W_m denote the jth and mth columns of W. Consider the inner product $W_j^\star W_m = \sum_{q=0}^{n-1} W_j^\star(q) W_m(q)$. For $\bmod (j-m, k) \neq 0$, the product evidently vanishes because in each term of the sum at least one of the factors is zero. For $j = m$, the terms where $\bmod (q, k) = j_0$ are equal to 1, while the rest are equal to zero, and therefore the product is b. Finally, for $j \neq m$ but $\bmod (j-m, k) = 0$, we can write $m = j + rk$ for some integer r s.t. $0 < |r| < b$. Summing up the non-zero terms, we obtain:

$$\sum_{p=0}^{n-1} W_j^\star(p)\, W_m(p) = \sum_{q=0}^{b-1} v_s^\star(j_0 + qk)\, v_{s+r}(j_0 + qk)$$
$$= \sum_{q=0}^{b-1} e^{-i\frac{2\pi r(j_0+qk)}{n}} = e^{-i\frac{2\pi r j_0}{n}} \sum_{q=0}^{b-1} \left(e^{-i\frac{2\pi r}{b}}\right)^q$$
$$= e^{-i\frac{2\pi r j_0}{n}} \frac{1 - e^{-i\frac{2\pi rb}{b}}}{1 - e^{-i\frac{2\pi r}{b}}} = 0.$$

We conclude that $\frac{1}{\sqrt{b}}W^\star \frac{1}{\sqrt{b}}W = I_n$, the $n \times n$ identity matrix. ☐

Lemma 2. *The similarity transformation,* $\hat{A} = \frac{1}{\sqrt{b}}W^\star A \frac{1}{\sqrt{b}}W$, *yields a block-diagonal matrix* $\hat{A}$ *with* b *blocks of size* $k \times k$.

Proof. Denote the lth row of A by A^l. Then, the product $A^l W_j$ reads

$$A^l W_j = \sum_{p=0}^{n-1} A^l(p)\, W_j(p) = \sum_{q=0}^{b-1} A^l(j_0 + qk)\, v_s(j_0 + qk)$$
$$= \sum_{q=0}^{b-1} A^{l_0+tk}(j_0 + qk)\, e^{-is\frac{2\pi(j_0+qk)}{n}}.$$

$$W=\begin{bmatrix}
1&0&0&1&0&0&1&0&0&1&0&0\\
0&1&0&0&e^{-i\frac{2\pi}{12}}&0&0&e^{-i\frac{4\pi}{12}}&0&0&e^{-i\frac{6\pi}{12}}&0\\
0&0&1&0&0&e^{-i\frac{4\pi}{12}}&0&0&e^{-i\frac{8\pi}{12}}&0&0&e^{-i\frac{12\pi}{12}}\\
1&0&0&e^{-i\frac{6\pi}{12}}&0&0&e^{-i\frac{12\pi}{12}}&0&0&e^{-i\frac{18\pi}{12}}&0&0\\
0&1&0&0&e^{-i\frac{8\pi}{12}}&0&0&e^{-i\frac{16\pi}{12}}&0&0&e^{-i\frac{24\pi}{12}}&0\\
0&0&1&0&0&e^{-i\frac{10\pi}{12}}&0&0&e^{-i\frac{20\pi}{12}}&0&0&e^{-i\frac{30\pi}{12}}\\
1&0&0&e^{-i\frac{12\pi}{12}}&0&0&e^{-i\frac{24\pi}{12}}&0&0&e^{-i\frac{36\pi}{12}}&0&0\\
0&1&0&0&e^{-i\frac{14\pi}{12}}&0&0&e^{-i\frac{28\pi}{12}}&0&0&e^{-i\frac{42\pi}{12}}&0\\
0&0&1&0&0&e^{-i\frac{16\pi}{12}}&0&0&e^{-i\frac{32\pi}{12}}&0&0&e^{-i\frac{48\pi}{12}}\\
1&0&0&e^{-i\frac{18\pi}{12}}&0&0&e^{-i\frac{36\pi}{12}}&0&0&e^{-i\frac{54\pi}{12}}&0&0\\
0&1&0&0&e^{-i\frac{20\pi}{12}}&0&0&e^{-i\frac{40\pi}{12}}&0&0&e^{-i\frac{60\pi}{12}}&0\\
0&0&1&0&0&e^{-i\frac{22\pi}{12}}&0&0&e^{-i\frac{44\pi}{12}}&0&0&e^{-i\frac{66\pi}{12}}
\end{bmatrix}$$

Figure 1. An example for W with $k = 3$ and $b = 4$.

By repeated use of (1), this yields

$$A^l W_j = \sum_{q=0}^{b-1} A^{l_0}\left(j_0 + \operatorname{mod}(q-t,b)\,k\right) e^{-is\frac{2\pi(j_0+qk)}{n}}$$
$$= e^{-is\frac{2\pi tk}{n}} \sum_{q=0}^{b-1} A^{l_0}\left(j_0 + \operatorname{mod}(q-t,b)\,k\right) e^{-is\frac{2\pi(j_0+(q-t)k)}{n}}$$
$$= e^{-is\frac{2\pi t}{b}} \sum_{q=0}^{b-1} A^{l_0}\left(j_0 + \operatorname{mod}(q-t,b)\,k\right) e^{-is\frac{2\pi(j_0+\operatorname{mod}(q-t,b)k)}{n}}$$
$$= e^{-is\frac{2\pi t}{b}} A^{l_0} W_j.$$

Denoting $u_j = A^{(0)} W_j$, we thus obtain

$$AW_j = \begin{pmatrix} u_j \\ e^{-i\frac{2\pi s}{b}} u_j \\ e^{-2i\frac{2\pi s}{b}} u_j \\ \vdots \\ e^{-(b-1)i\frac{2\pi s}{b}} u_j \end{pmatrix},$$

where the qth element of u_j, $q = 0, \ldots, k-1$, is given by

$$u_j(q) = A^q W j = \sum_{q=0}^{b-1} A^q\left(j_0 + qk\right) v_s\left(j_0 + qk\right).$$

Multiplying on the left by $W_l^\star$ for any $l = l_0 + tk$ with $l_0 \in \{0, \ldots, k-1\}$ and $t \in \{0, \ldots, b-1\}$, yields

$$W_l^\star AW_j = \sum_{q=0}^{b-1} W_l^\star\left(l_0 + qk\right) e^{-qi\frac{2\pi s}{b}} u_j\left(l_0\right)$$
$$= \sum_{q=0}^{b-1} v_t^\star\left(l_0 + qk\right) e^{-qi\frac{2\pi s}{b}} u_j\left(l_0\right)$$
$$= \sum_{q=0}^{b-1} e^{i\frac{2\pi t(l_0+qk)}{n}} e^{-qi\frac{2\pi s}{b}} u_j\left(l_0\right)$$
$$= e^{i\frac{2\pi t l_0}{n}} u_j\left(l_0\right) \sum_{q=0}^{b-1} e^{qi\frac{2\pi(t-s)}{b}}.$$

For $t \neq s$, the final sum yields

$$\sum_{q=0}^{b-1} e^{qi\frac{2\pi(t-s)}{b}} = \sum_{q=0}^{b-1} \left(e^{qi\frac{2\pi(t-s)}{b}}\right)^q = \frac{1 - e^{i2\pi(t-s)}}{1 - e^{i\frac{2\pi(t-s)}{b}}} = 0.$$

We conclude that $W_l^\star AW_j$ vanishes unless $t = s$, which implies the block-periodic form stated in the proposition. □

For $t = s$, all the terms in the final sum in the proof are equal to 1, and therefore the sum is equal to b. This yields the following.

Theorem 1. *Let W be the matrix defined in (3). Then, $\hat{A} = \frac{1}{\sqrt{b}} W^\star A \frac{1}{\sqrt{b}} W = \operatorname{blockdiag}(B^{(0)}, \ldots, B^{(b-1)})$, where the elements of the $k \times k$ blocks $B^{(s)}$, $s = 0, \ldots, b-1$, are* *given by*

$$
\begin{aligned}
B^{(s)}_{l_0,j_0} &= e^{i\frac{2\pi s l_0}{n}} u_j\left(l_0\right) \\
&= e^{i\frac{2\pi s l_0}{n}} \sum_{q=0}^{b-1} A^{l_0}\left(j_0+qk\right) v_s\left(j_0+qk\right) \\
&= e^{i\frac{2\pi s l_0}{n}} \sum_{q=0}^{b-1} A^{l_0}\left(j_0+qk\right) e^{-i\frac{2\pi s (j_0+qk)}{n}} \\
&= e^{i\frac{2\pi s (l_0-j_0)}{n}} \sum_{q=0}^{b-1} A^{l_0}\left(j_0+qk\right) e^{-i\frac{2\pi s q}{b}}.
\end{aligned}
$$

Remark 1. *In the context of multigrid analysis, particularly in the present paper, the block size is chosen to be at least as large as the stencil. In such cases, the coefficients* $A^{l_0}(j_0+qk)$ *are equal to zero for all non-zero* q*, yielding the simple form:*

$$
B^{(s)}_{l_0,j_0} = e^{i\frac{2\pi s (l_0-j_0)}{n}} A^{l_0}\left(j_0\right) = e^{i\frac{2\pi s (l_0-j_0)}{n}} A_{l_0,j_0}.
$$

Remark 2. *The block Fourier analysis is applicable to discretized partial differential equations of any dimension* d *by recursion. That is, for* $d>1$ *the blocks of* A *are themselves block-circulant, and so on. Remark 1 also generalizes to any dimension. That is, if the diameter of the discretization stencil is at most* k *then each element of* B *is easily computed from a single element of* A*.*