SIAM J. MATRIX ANAL. APPL.
Vol. 40, No. 2, pp. 564–608




# NECESSARY CONDITIONS AND TIGHT TWO-LEVEL CONVERGENCE BOUNDS FOR PARAREAL AND MULTIGRID REDUCTION IN TIME*

BEN S. SOUTHWORTH†

**Abstract.** Parareal and multigrid reduction in time (MGRiT) are two of the most popular parallel-in-time methods. The basic idea is to treat time integration in a parallel context by using a multigrid method in time. If $\Phi$ is the (fine-grid) time-stepping scheme of interest, such as any Runge–Kutta scheme, then let $\Psi$ denote a "coarse-grid" time-stepping scheme chosen to approximate $k$ steps of $\Phi$, where $k \geq 1$. In particular, $\Psi$ defines the coarse-grid correction, and evaluating $\Psi$ should be (significantly) cheaper than evaluating $\Phi^k$. Parareal is a two-level method with a fixed relaxation scheme, and MGRiT is a generalization to the multilevel setting, with the additional option of a modified, stronger relaxation scheme. A number of papers have studied the convergence of Parareal and MGRiT. However, general conditions on the convergence of Parareal or MGRiT that answer the following simple questions have yet to be developed: (i) For a given $\Phi$ and $k$, what is the best $\Psi$? (ii) Can Parareal/MGRiT converge for my problem? This work derives necessary and sufficient conditions for the convergence of Parareal and MGRiT applied to linear problems, along with tight two-level convergence bounds, under minimal additional assumptions on $\Phi$ and $\Psi$. Results all rest on the introduction of a *temporal approximation property* (TAP) that indicates how $\Phi^k$ must approximate the action of $\Psi$ on different vectors. Loosely, for unitarily diagonalizable operators, the TAP indicates that the fine-grid and coarse-grid time integration schemes must integrate geometrically smooth spatial components similarly, and less so for geometrically high frequency. In the (nonunitarily) diagonalizable setting, the conditioning of each eigenvector, $\mathbf{v}_i$, must also be reflected in how well $\Psi\mathbf{v}_i \sim \Phi^k\mathbf{v}_i$. In general, worst-case convergence bounds are exactly given by $\min \varphi < 1$ such that an inequality along the lines of $\|(\Psi - \Phi^k)\mathbf{v}\| \leq \varphi\|(I - \Psi)\mathbf{v}\|$ holds for all $\mathbf{v}$. Such inequalities are formalized as different realizations of the TAP in section 2 and form the basis for convergence of MGRiT and Parareal applied to linear problems.

**Key words.** parallel-in-time, Parareal, MGRiT, convergence, multigrid, reduction

**AMS subject classifications.** 65F10, 65M55, 65M12, 15B05

**DOI.** 10.1137/18M1226208

## 1. Introduction.

Efficiently distributing computational work over many processors, or parallelizing, is fundamental to running large-scale numerical simulations. In the case of partial differential equations (PDEs) in space and time, problems are at least three to four dimensional, with as many as seven dimensions or more for problems such as radiative transport. Additionally accounting for multiple variables that may have to be solved for, even a moderate number of points in each dimension requires a massive number of unknowns, as well as a high level of parallelism, to obtain an accurate solution. Furthermore, computational power is largely increasing in the number of processors available and less in the power of individual processors, making increased parallelism an important area of research.

Steady state PDEs (non–time-dependent) are typically posed as boundary value problems (BVPs), which provide a natural mechanism to parallelize in space. When

*Received by the editors November 12, 2018; accepted for publication (in revised form) by S. Le Borne March 4, 2019; published electronically May 9, 2019.

http://www.siam.org/journals/simax/40-2/M122620.html

**Funding:** This work was performed under the auspices of the U.S. Department of Energy under grant (NNSA) DE-NA0002376 and Lawrence Livermore National Laboratory under contracts B614452 and B627942.

†Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309 (ben.s.southworth@gmail.com).

time derivates are introduced, adding parallelism in the time dimension is more complicated. In particular, for time-dependent PDEs, it is often the case that only an initial value in time is given. To that end, propagating information through the temporal domain appears to be an inherently sequential process because the initial information can only be propagated in one direction, namely forward in time. This is how most time-dependent PDEs are solved—given some initial value in time, a BVP is formulated and discretized in the spatial domain. The initial value problem (IVP) is then propagated forward one time step by solving the BVP and applying some time integration routine, and the process repeats based on a new "initial value" in time. However, as the number of processors available to run numerical simulations has increased, so has the interest in so-called parallel-in-time methods, which are designed to parallelize the process of integrating forward in time.

Because time integration typically involves solving for a solution at a set of discrete "time points," it can be represented in block matrix form, where

$$A\mathbf{u} = \begin{bmatrix} I & & & \\ -\Phi & I & & \\ & -\Phi & I & \\ & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \end{bmatrix} = \mathbf{f}. \tag{1}$$

Here, $\mathbf{u}_i$ is the solution at the $i$th time point, $\mathbf{f}$ the right-hand side, and $\Phi$ some invertible operator that advances the solution from time $t_i$ to $t_{i+1}$. In this setting, classical (sequential) time integration can be seen as a direct (forward) solve of (1). Parallel-in-time methods can typically be posed as some form of preconditioner or iterative method to solve (1). This introduces new questions on the convergence of such iterations, which do not arise in the sequential setting. Interestingly, although a lower bidiagonal matrix is one of the easiest linear systems to solve in serial through a forward solve, solving lower triangular matrices remains difficult in the parallel setting.

Many parallel-in-time methods have been proposed, with varying levels of success. Some of the prominent methods include full space-time multigrid [7, 20], parallel full approximation scheme in space and time (PFASST) [6], Parareal [23], and multigrid reduction in time (MGRiT) [8]. Parareal is perhaps the most well known and one of the original ideas for parallel-in-time integration. Parareal is effectively a two-level multigrid method. Time points are partitioned into C-points and F-points, and relaxation consists of integrating each C-point $k-1$ time steps forward, based on its current solution (that is, to the final F-point preceding the next C-point); here $k$ denotes the coarsening factor. This is followed by a coarse-grid correction, which approximately inverts the Schur complement of (1). In particular, $k$ steps on the fine grid, $\Phi^k$, are approximated by some operator $\Psi$ that is cheaper to evaluate. The simplest example is letting $\Psi$ be the same time-integration scheme as $\Phi$, using time steps that are $k$ times larger. The MGRiT algorithm generalizes this to the multilevel setting by recursively coarsening the temporal grid until it is sufficiently small to solve directly at minimal cost.

Analysis of Parareal dates back to [1], where Parareal is analyzed from a time-integration perspective, looking at the stability and error of Parareal as a time-integration routine with respect to the continuous problem. In [15], the connection between Parareal and a two-level multigrid algorithm with F-relaxation is discussed, and initial bounds developed for Parareal that are, in some cases, sharp. An analysis of the nonlinear case is developed in [13], largely demonstrating that Parareal

is applicable to nonlinear problems, and the convergence of Parareal applied to elasticity and plasma simulations is discussed, respectively, in [18, 30]. More recently, [40, 39] analyze several specific time-integration schemes applied to problems of the form $\mathbf{u}_t = \mathcal{L}\mathbf{u} + \mathbf{g}$, where $\mathcal{L}$ is symmetric negative definite. Some of the results are tight but are indeed limited to specific time-integration schemes and symmetric negative definite spatial discretizations. A detailed analysis of error propagation of two-level MGRiT and Parareal is developed in [5], under the assumption that fine- and coarse-grid time-stepping operators commute and are diagonalizable. Results in [5] are, to some extent, a generalization of [15] and also introduce FCF-relaxation to the analysis, a variation in MGRiT that is not used in Parareal but that can be important for convergence on difficult problems. Numerical results in [5] demonstrate that the derived bounds appear to be tight when applied to a number of model problems. In fact, the framework developed in this paper is a substantial generalization of that in [5]. One corollary proven here is that, for a certain class of problems, the bounds developed in [5] are indeed exact to $O(1/N_c)$, where $N_c$ is the number of time points on the coarse grid. Relaxation is generalized further in the recent paper [14] using an algebraic perspective, similar to the framework used in this paper. The framework developed in [5] is also extended to the multilevel setting in [19], under the assumption of diagonalizable operators.[1] Finally, in [11], local mode analysis techniques are generalized and applied to parabolic parallel-in-time problems, providing more accurate estimates of convergence than traditional techniques.

Despite a number of works analyzing Parareal and MGRiT, there remains a gap in the literature in answering the following fundamental question: For a general problem, what must a coarse-grid time-stepping scheme satisfy to see convergent behavior? The main contribution of this paper is the development of, under minimal assumptions, necessary and sufficient conditions for the convergence in norm of linear Parareal and two-level MGRiT. A simple *temporal approximation property* (TAP) is introduced that measures how accurately the fine-grid time-integration scheme approximates the coarse-grid integration scheme. This leads to necessary and sufficient conditions for convergence of error propagation in the $\ell^2$- and $A^*A$-norms, including tight bounds on convergence. Additional results are established under further assumptions on the time-integration operators being diagonalizable and unitarily diagonalizable. Indeed, if we assume that the spatial discretization is symmetric and definite as in [39, 40], results here provide exact bounds on convergence for arbitrary time-integration schemes. For the most general results, the only assumption is that the problem is linear. Most results also require that the same operator $\Phi$ integrates the solution at all time-points (for example, there can be no time-dependent differential spatial coefficients). In all cases, some variation on the TAP provides a simple and relatively intuitive explanation of exactly how the coarse-grid operator must approximate the fine-grid operator for convergence. Given a time-integration scheme of interest, it can easily be plugged into results here for a more problem-specific description of necessary and sufficient conditions for convergence. The theory is based on building error-propagation operators and appealing to block-Toeplitz matrix theory; it is interesting to point out that a similar approach to that used here likely provides a general convergence framework for the recent idea of using circulant preconditioners in time to solve the space-time system in parallel [17, 27].

The rest of this paper proceeds as follows. Section 2 presents the main theoreti-

[1] Results and the framework developed in this paper are also extended in [19] to discuss the difficulties that a multilevel method presents over the two-level setting.

cal contributions in a concise and accessible manner. Proofs of these results are then established in the following sections. Section 3 discusses convergence of reduction-based multigrid-type methods and derives analytic formulas for two-level error- and residual-propagation operators of MGRiT. The most general theorems are then derived in section 4, and further analysis based on additional assumptions is given in section 5. Some of the analysis can be extended to the time-dependent case, and a discussion on that topic is given in section 6. There are many applications for the new theorems, such as finding the "best" coarse-grid time-stepping scheme for a given problem, better understanding why hyperbolic problems tend to be difficult for parallel-in-time solvers [33], and understanding the effect of spatial coarsening on the convergence of Parareal/MGRiT [32, 21], among others. A brief discussion on implications of results derived here is given in section 7, and a detailed study is the topic of a forthcoming paper.

**2. Statement of results.** This section presents the theoretical contributions of this paper; proofs are derived in the sections that follow. The underlying idea is that Parareal and MGRiT are iterative methods to solve a discrete linear system (1), of which the exact solution is simply the space-time vector achieved through sequential time-stepping. Here, convergence theory is derived to provide, under certain assumptions, necessary conditions and sufficient conditions for two-level MGRiT/Parareal to converge to the sequential solution in norm.

**2.1. The framework.** Let $\Phi$ be an $N_x \times N_x$ invertible fine-grid time-stepping operator and $\Psi$ an $N_x \times N_x$ invertible coarse-grid time-stepping operator, and suppose we coarsen in time by a factor of $k$. The primary results rest on three assumptions:

1. $\Phi$ and $\Psi$ are linear.
2. $\|\Phi\|, \|\Psi\|$ are stable; that is, $\|\Phi^p\|, \|\Psi^p\| < 1$ for some $p$.
3. $\Phi$ and $\Psi$ are independent of time; that is, the same operator propagates the solution from time $t_i$ to $t_{i+1}$ and from $t_j$ to $t_{j+1}$, for all $i, j$.

Assumption 1 restricts our attention to the linear case, which, as in many problems, allows for a more detailed analysis. The second assumption is an algebraic requirement for a stable time-stepping scheme, which is a also a natural and reasonable thing to assume. Note that a stronger assumption is $\|\Phi\|, \|\Psi\| < 1$. LeVeque refers to this as "strong stability" [22, Chapter 9.5], but results here hold for the more general case as well. The third assumption is the strongest one, yet it still encompasses all problems for which variables do not have time-dependent coefficients, which consists of a large class of space-time PDEs, among other problems. Some of the theory developed here applies to the time-dependent case as well. In particular, sufficient conditions can be derived (see Theorem 35) for convergence if $\Phi$ and $\Psi$ are time-dependent but simultaneously diagonalizable for all time steps. This occurs, for example, in the case of time-dependent reaction terms, or for adaptive time-stepping. Some other results hold in the time-dependent case as well, which are discussed in section 6.

Further assumptions that come up (yielding stronger convergence results) are as follows:

4. Assume that $(\Psi - \Phi^k)$ is invertible.
5. Assume that $\Phi$ and $\Psi$ commute.
   (a) Assume that $\Phi$ and $\Psi$ are diagonalizable.
   (b) Assume that $\Phi$ and $\Psi$ are normal (unitarily diagonalizable).

A discussion on these assumptions is provided in section 2.5; however, it is worth noting that all of them are plausible assumptions for many problems of interest. Also, it is believed that assumption 4 is not actually necessary for any of the presented

theoretical results. However, without assumption 4, some of the analysis becomes significantly more complicated and is not pursued in this work.

To consider convergence of Parareal and MGRiT, let $\hat{\mathbf{u}}$ be the exact solution to (1). Then the error and residual for an approximate solution at the $i$th iteration, $\mathbf{u}_i$, are given, respectively, by

$$\begin{aligned}\mathbf{e}_i &= \hat{\mathbf{u}} - \mathbf{u}_i,\\ \mathbf{r}_i &= \mathbf{b} - A\mathbf{u}_i = A(\hat{\mathbf{u}} - \mathbf{u}_i) = A\mathbf{e}_i.\end{aligned}$$

Here, we seek bounds on how Parareal and MGRiT propagate the error and residual corresponding to the linear system in (1). To measure this propagation, we use the discrete $\ell^2$-norm ($\|\cdot\|$) and $A^*A$-norm ($\|\cdot\|_{A^*A}$),[2] defined by

$$\|\mathbf{e}_i\|^2 = \langle \mathbf{e}_i, \mathbf{e}\rangle, \qquad \|\mathbf{e}_i\|_{A^*A}^2 = \langle A^*A\mathbf{e}_i, \mathbf{e}_i\rangle = \|\mathbf{r}_i\|^2.$$

Note that assuming $A$ is nonsingular, $\mathbf{e}_i = \mathbf{0}$ if and only if $\mathbf{r}_i = \mathbf{0}$, and as $\mathbf{e}, \mathbf{r} \to \mathbf{0}$, we converge to the discrete solution obtained through sequential time stepping. Although in practice users typically want the error to be small, the error cannot be easily measured in practice, while the residual can be, making error and residual propagation both of interest.

Moving forward, let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, for all time points, with coarsening factor $k$, and $N_c$ time-points on the coarse grid. Let $\mathbf{e}_i^{(F)}$ denote error associated with Parareal or MGRiT with F-relaxation after $i$ iterations, $\mathbf{e}_i^{(FCF)}$ denote error associated with MGRiT with FCF-relaxation after $i$ iterations, and similarly for residual vectors. Most results here are asymptotic in the sense that certain approximation properties and bounds are given up to order $O(1/N_c)$. However, the leading constants in the $O(1/N_c)$ terms are also generally quite small, and positive in all cases. Furthermore, because parallel-in-time is most often used when the number of time steps is relatively large, in practice these terms can often be considered negligible. Note that the operators $\Phi$ and $\Psi$ correspond to one time step and are independent of $N$ and $N_c$.

The most general results and remarks are presented in section 2.2. Several extensions for specific cases are given in sections 2.3 and 2.4, and proofs for all results are provided in the sections that follow.

**2.2. Necessary and sufficient conditions.** This section introduces necessary and sufficient conditions for convergence of residual in the $\ell^2$-norm and, equivalently, error in the $A^*A$-norm, including tight bounds in norm. In the case that $\Phi$ and $\Psi$ commute, these results hold for error in the $\ell^2$-norm as well. To start, a new TAP is introduced, which is the fundamental assumption leading to convergence.

DEFINITION 1 (Temporal approximation property). *Let $\Phi$ denote a fine-grid time-stepping operator and $\Psi$ denote a coarse-grid time-stepping operator, for all time points, with coarsening factor $k$. Then, $\Phi$ satisfies an F-relaxation temporal approximation property with power $p$ (F-TAP$_p$), with respect to $\Psi$, with constant $\varphi_{F,p}$, if, for all vectors $\mathbf{v}$,*

$$\|(\Psi - \Phi^k)^p\mathbf{v}\| \leq \varphi_{F,p}\left[\min_{x\in[0,2\pi]} \left\|(I - e^{\mathrm{i}x}\Psi)^p\mathbf{v}\right\|\right]. \tag{2}$$

[2]The $\ell^2$-norm and $A^*A$-norm are generally the most common norms used for nonsymmetric problems, where the latter corresponds to a normal-equation formulation.

*Similarly, $\Phi$ satisfies an FCF-relaxation temporal approximation property with power $p$ (FCF-TAP$_p$), with respect to $\Psi$, with constant $\varphi_{FCF,p}$, if, for all vectors $\mathbf{v}$,*

$$\|(\Psi - \Phi^k)^p \mathbf{v}\| \leq \varphi_{FCF,p} \left[ \min_{x \in [0,2\pi]} \left\| (\Phi^{-k}(I - e^{\mathrm{i}x}\Psi))^p \mathbf{v} \right\| \right]. \tag{3}$$

Necessary and sufficient conditions for convergence of MGRiT and Parareal under various further assumptions are all based on satisfying one of the above approximation properties with a nicely bounded constant, typically less than one. The two variations on a TAP are conceptually simple and can be presented in a more intuitive manner as follows. An F-TAP requires that $\Phi^k$ approximates the action of $\Psi$ very accurately for vectors $\mathbf{v} \approx \Psi\mathbf{v}$ and less accurately for $\mathbf{v}$ that differs significantly from $\Psi\mathbf{v}$. If $\Phi$ and $\Psi$ commute and have an orthogonal eigenvector basis, then $\Phi^k$ must approximate $\Psi$ very accurately for eigenvectors of $\Psi$ with associated eigenvalue close to one in magnitude, and less accurately for smaller eigenvalues. In the context of PDEs, order-one eigenmodes of $\Psi$ typically correspond to the smallest eigenvalues of the spatial discretization. To that end, the fine-grid and coarse-grid time-stepping operators must propagate "smooth" modes in the spatial domain (corresponding to small eigenvalues) very similarly. In the case of an FCF-TAP, the additional term $\Phi^{-k}$ often makes the TAP easier to satisfy. If $\|\Phi^k\| < 1$, that is, $\Phi$ is strongly stable [22, Chapter 9.5], then $\|\Phi^{-k}\mathbf{v}\| > \|\mathbf{v}\|$ for all $\mathbf{v}$. Think of this as an extra fudge factor to help convergence (at the added expense of FCF-relaxation). When $\Psi$ is not diagonalizable, the eigenvectors do not form a basis; when $\|(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\| \approx 0$ is a more complicated question. Further analysis of that case, particularly for hyperbolic problems, is ongoing work.

Necessary and sufficient conditions for convergence of MGRiT and Parareal are now presented with respect to the TAP.

THEOREM 2 (Necessary and sufficient conditions: Error in the $A^*A$-norm). *Suppose that assumptions* 1, 2, *and* 3 *hold and that $\Phi$ satisfies an F-TAP$_1$ with respect to $\Psi$, with constant $\varphi_F$, and $\Phi$ satisfies an FCF-TAP$_1$ with respect to $\Psi$, with constant $\varphi_{FCF}$. Then,*

$$\frac{\|\mathbf{r}_{i+1}^{(F)}\|}{\|\mathbf{r}_i^{(F)}\|} = \frac{\|\mathbf{e}_{i+1}^{(F)}\|_{A^*A}}{\|\mathbf{e}_i^{(F)}\|_{A^*A}} < \varphi_F \left(1 + \|\Psi^{N_c}\|\right), \tag{4}$$

$$\frac{\|\mathbf{r}_{i+1}^{(FCF)}\|}{\|\mathbf{r}_i^{(FCF)}\|} = \frac{\|\mathbf{e}_{i+1}^{(FCF)}\|_{A^*A}}{\|\mathbf{e}_i^{(FCF)}\|_{A^*A}} < \varphi_{FCF} \left(1 + \|\Phi^{-k}\Psi^{N_c}\Phi^k\|\right) \tag{5}$$

*for iterations $i > 1$. Thus, satisfying $\varphi_F \left(1 + \|\Psi^{N_c}\|\right) < 1$ and $\varphi_{FCF} \left(1 + \|\Phi^{-k}\Psi^{N_c}\Phi^k\|\right) < 1$ are sufficient conditions for convergence of MGRiT with F-relaxation and FCF-relaxation, respectively, on every iteration but one, with respect to error in the $A^*A$-norm.*

*Additionally, assume that $(\Psi - \Phi^k)$ is invertible (assumption* 4*). Then,*

$$\frac{\varphi_F}{1 + O(1/\sqrt{N_c})} \leq \frac{\|\mathbf{r}_{i+1}^{(F)}\|}{\|\mathbf{r}_i^{(F)}\|} = \frac{\|\mathbf{e}_{i+1}^{(F)}\|_{A^*A}}{\|\mathbf{e}_i^{(F)}\|_{A^*A}}, \tag{6}$$

$$\frac{\varphi_{FCF}}{1 + O(1/\sqrt{N_c})} \leq \frac{\|\mathbf{r}_{i+1}^{(FCF)}\|}{\|\mathbf{r}_i^{(FCF)}\|} = \frac{\|\mathbf{e}_{i+1}^{(FCF)}\|_{A^*A}}{\|\mathbf{e}_i^{(FCF)}\|_{A^*A}} \tag{7}$$

*for iterations $i > 1$. Thus, satisfying $\varphi_F < 1 + O(1/\sqrt{N_c})$ and $\varphi_{FCF} < 1 + O(1/\sqrt{N_c})$ are necessary conditions for convergence of MGRiT with F-relaxation and*

*FCF-relaxation, respectively, on every iteration but one, with respect to error in the $A^*A$-norm.*

*Finally, assume that $\Phi$ and $\Psi$ commute and either* (i) *$(\Psi - \Phi^k)$ is invertible (assumptions* 4 *and* 5*) or* (ii) *$\Phi$ and $\Psi$ are diagonalizable (assumptions* 5 *and* 5(a)/5(b)*). Then, $\Phi$ satisfying an F-TAP$_p$, for power $p \geq 1$, with respect to $\Psi$, with $\varphi_{F,p} < (1 + O(1/\sqrt{N_c}))$, is a necessary condition to see convergent behavior of Parareal and two-level MGRiT with F-relaxation, after $p$ iterations, with respect to error in the $A^*A$-norm. Similarly, $\Phi$ satisfying an FCF-TAP$_p$, for power $p \geq 1$, with respect to $\Psi$, with $\varphi_{FCF,p} < (1+O(1/\sqrt{N_c}))$, is a necessary condition to see convergent behavior of two-level MGRiT with FCF-relaxation, after $p$ iterations, with respect to error in the $A^*A$-norm.*

Theorem 2 presents necessary and sufficient conditions for convergence of Parareal and MGRiT with minimal assumptions. The first statements provide necessary and sufficient conditions that every iteration is convergent in the $\ell^2$-norm for residual and $A^*A$-norm for error. Note that the bounds are tight. That is, as $N_c$ increases, $\|\Psi^{N_c}\| \approx 0$ and the worst-case ratio of successive error vectors in the $A^*A$-norm converges exactly to $\varphi_F$ or $\varphi_{FCF}$.

However, this only considers worst-case convergence for a single iteration. It is possible that convergence after $p$ iterations is $\ll \varphi^p$. In theory, it is possible to see divergent behavior on initial iterations but eventual convergence (because for some non-Hermitian operators, $\|M^p\| \ll \|M\|^p$). Under the additional assumptions that $\Phi$ and $\Psi$ commute (which holds, for example, in using arbitrary single-step multistage integration schemes for $\Phi$ and $\Psi$ (section 2.5)) and either assumption 4 or assumption 5(a)/5(b), the final statement in each theorem provides necessary conditions to see convergence of residual in the $\ell^2$-norm and error in the $A^*A$-norm after an arbitrary number of iterations.

COROLLARY 3 (Extension to error in the $\ell^2$-norm). *If $\Phi$ and $\Psi$ commute, then identical conditions and bounds as in Theorem* 2 *hold for convergence of error in the $\ell^2$-norm on all iterations except the last (as opposed to the first).*

If $\Phi$ and $\Psi$ do not commute, similar results as in Theorem 2 hold but require a modified version of the TAP, which is introduced in section 2.3. If they do commute *and* are diagonalizable, section 2.4 introduces additional results in a modified norm.

*Remark* 4 (Real-valued operators). Suppose $\Psi$ is real-valued and we only consider real-valued $\mathbf{v}$. Expanding $\|(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\|$ as an inner product yields

$$\|(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\|^2 = \|\mathbf{v}\|^2 + \|\Psi\mathbf{v}\|^2 - 2\cos(\theta_x)\langle\Psi\mathbf{v}, \mathbf{v}\rangle,$$

where $e^{\mathrm{i}x} = \cos(\theta_x) + i\sin(\theta_x)$. Then,

$$\min_{x\in[0,2\pi]} \|(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\| = \begin{cases} \|(I+\Psi)\mathbf{v}\| & \text{if } \langle\Psi\mathbf{v}, \mathbf{v}\rangle \leq 0, \\ \|(I-\Psi)\mathbf{v}\| & \text{if } \langle\Psi\mathbf{v}, \mathbf{v}\rangle > 0. \end{cases}$$

Similarly, if $\Phi$ is also real-valued, then

$$\min_{x\in[0,2\pi]} \|\Phi^{-k}(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\| = \begin{cases} \|\Phi^{-k}(I+\Psi)\mathbf{v}\| & \text{if } \langle\Phi^{-k}\Psi\mathbf{v}, \Phi^{-k}\mathbf{v}\rangle \leq 0, \\ \|\Phi^{-k}(I-\Psi)\mathbf{v}\| & \text{if } \langle\Phi^{-k}\Psi\mathbf{v}, \Phi^{-k}\mathbf{v}\rangle > 0. \end{cases}$$

*Remark* 5 (F(CF)$^\rho$-relaxation). Here, we only consider the cases of F- and FCF-relaxation. However, results generalize naturally to arbitrary F(CF)$^\rho$-relaxation (as

considered in [14]), where the CF-steps are repeated $\rho$ times, by simply adding the term $\Phi^{-k\rho}$ to the right-hand side of the F-TAP, analogous to the $\Phi^{-k}$ in the FCF-TAP.

*Remark* 6 (Error tolerance, $\delta t$, and superlinear convergence). It is important to see Parareal and MGRiT as iterative solvers of a discrete linear system (1), rather than an integration scheme to solve the continuous problem. In general, the solution obtained through Parareal or MGRiT should be no more accurate than that obtained through sequential time stepping, which is exactly defined by the choice of $\Phi$. An important question in discrete linear systems is how accurately to solve them. Suppose $\Phi$ is a time-integration scheme with global accuracy $\mathcal{O}(\delta t^s)$. Then it is generally only necessary to solve the discrete linear system (1) (for example, using Parareal or MGRiT) to accuracy $\mathcal{O}(\delta t^s)$.

When superlinear convergence of Parareal to the continuous solution is observed (for example, see [16]), this corresponds to Parareal iterations converging faster (in the discrete sense) than the integration accuracy of $\Phi$. From Theorem 2, we see that this is likely a result of satisfying the F-TAP with constant $O(\delta t^\ell)$, where $\ell$ is greater than the integration accuracy of $\Phi$. For example, if the F-TAP is satisfied with constant $\varphi_F = \delta t^2$, for given $\Phi$ and $\Psi$, independent of $\delta t$, then Parareal will converge like $\delta t^2$, even as $\delta t \to 0$.

*Remark* 7 (Self-consistency of $\Psi$). One of the most surprising results of this theory is how convergence of Parareal depends on the coarse-grid time stepper, $\Psi$. It is natural to assume that $\Psi$ must approximate $k$ steps on the fine grid, $\Phi^k$, with accuracy that somehow depends on $\Phi$. However, this is not the case. Indeed, the TAP illustrates that $\Psi$ must approximate $\Phi^k$ with accuracy based on $I-\Psi$, indicating that there must be some self-consistency in terms of which vectors $\Psi$ approximates the action of $\Phi^k$ on well.

*Remark* 8 (Computing TAP constants). The constants in the TAP are exactly defined as the maximum generalized singular value of the pair $\{\Psi - \Phi^k, I - e^{\mathrm{i}\hat{x}}\Psi\}$ for some $\hat{x} \in [0, 2\pi]$. If we consider real-valued operators, then (from Remark 4) we seek the maximum generalized singular value of $\{\Psi - \Phi^k, I - \Psi\}$ and $\{\Psi - \Phi^k, I + \Psi\}$. For sparse matrices, such as those that arise with explicit time stepping of differential discretizations, iterative methods have been developed to compute these values and vectors for relatively cheap and without forming $(I - \Psi)^{-1}$ (for example, see [45]). Even if $\Phi$ and $\Psi$ are implicit and thus contain inverses, iterative methods to compute the largest generalized singular value are typically applicable if the actions of $\Phi$ and $\Psi$ are available.

**2.3. Tight convergence of $\ell^2$-error.** Section 2.2 developed necessary and sufficient conditions for convergence of error in the $A^*A$-norm, and Corollary 3 states that if $\Phi$ and $\Psi$ commute, results equivalent to those of Theorem 2 follow immediately for error in the $\ell^2$-norm. If $\Phi$ and $\Psi$ do not commute, we need to introduce a modified inverse TAP (ITAP) to study convergence of error in the $\ell^2$-norm.

Definition 9 (Inverse temporal approximation property). *Let $\Phi$ denote a fine-grid time-stepping operator and $\Psi$ denote a coarse-grid time-stepping operator, for all time points, with coarsening factor $k$, such that $(I - e^{\mathrm{i}x}\Psi)$ is invertible. Then, $\Phi$ satisfies an F-relaxation inverse temporal approximation property (F-ITAP), with respect to $\Psi$, with constant $\tilde{\varphi}_F$, if, for all vectors $\mathbf{v}$,*

$$\max_{x\in[0,2\pi]} \left\|(I - e^{\mathrm{i}x}\Psi)^{-1}(\Psi - \Phi^k)\mathbf{v}\right\| \leq \tilde{\varphi}_F\|\mathbf{v}\|. \tag{8}$$

*Similarly, $\Phi$ satisfies an FCF-relaxation inverse temporal approximation property (FCF-ITAP), with respect to $\Psi$, with constant $\tilde{\varphi}_{FCF}$, if, for all vectors $\mathbf{v}$,*

$$\max_{x\in[0,2\pi]} \left\|(I - e^{\mathrm{i}x}\Psi)^{-1}(\Psi - \Phi^k)\mathbf{v}\right\| \leq \tilde{\varphi}_{FCF}\|\Phi^{-k}\mathbf{v}\|. \tag{9}$$

In the case that $(\Psi - \Phi^k)$ is invertible, $(\Psi - \Phi^k)$ can be moved to the right-hand side. For example, the F-ITAP can be expressed as

$$\max_{x\in[0,2\pi]} \left\|(I - e^{\mathrm{i}x}\Psi)^{-1}\mathbf{v}\right\| \leq \tilde{\varphi}_F\|(\Psi - \Phi^k)^{-1}\mathbf{v}\|$$

for all $\mathbf{v}$. Note that the ITAP is not considered with respect to powers $p$. This is because derived results based on powers also assume that $\Phi$ and $\Psi$ commute (see Theorem 2), in which case the results from section 2.2 hold for error in the $\ell^2$-norm, and the ITAP is not necessary. Also, the assumption that $I - e^{\mathrm{i}x}\Psi$ is invertible is equivalent to assuming that $\Psi$ does not have an eigenvalue of magnitude exactly one.[3]

THEOREM 10 (Necessary and sufficient conditions: Error in the $\ell^2$-norm). *Suppose that assumptions* 1, 2, *and* 3 *hold and that $\Phi$ satisfies an F-ITAP with respect to $\Psi$, with constant $\tilde{\varphi}_F$, and $\Phi$ satisfies an FCF-ITAP with respect to $\Psi$, with constant $\tilde{\varphi}_{FCF}$. Then, with $n$ total iterations,*

$$\frac{\|\mathbf{e}_{i+1}^{(F)}\|}{\|\mathbf{e}_i^{(F)}\|} < \tilde{\varphi}_F\left(1 + \|\Psi^{N_c}\|\right), \tag{10}$$

$$\frac{\|\mathbf{e}_{i+1}^{(FCF)}\|}{\|\mathbf{e}_i^{(FCF)}\|} < \tilde{\varphi}_{FCF}\left(1 + \|\Psi^{N_c}\|\right) \tag{11}$$

*for iterations $i = 0, \ldots, n-2$. Thus, satisfying $\tilde{\varphi}_F\left(1 + \|\Psi^{N_c}\|\right) < 1$ and $\tilde{\varphi}_{FCF}(1 + \|\Phi^{-k}\Psi^{N_c}\Phi^k\|) < 1$ are sufficient conditions for convergence of MGRiT with F-relaxation and FCF-relaxation, respectively, on every iteration but one, with respect to error in the $\ell^2$-norm.*

*Additionally, assume that $(\Psi - \Phi^k)$ is invertible (assumption 4). Then, with $n$ total iterations,*

$$\frac{\tilde{\varphi}_F}{1 + O(1/\sqrt{N_c})} \leq \frac{\|\mathbf{e}_{i+1}^{(F)}\|}{\|\mathbf{e}_i^{(F)}\|}, \tag{12}$$

$$\frac{\tilde{\varphi}_{FCF}}{1 + O(1/\sqrt{N_c})} \leq \frac{\|\mathbf{e}_{i+1}^{(FCF)}\|}{\|\mathbf{e}_i^{(FCF)}\|} \tag{13}$$

*for iterations $i = 0, \ldots, n-2$. Thus, satisfying $\tilde{\varphi}_F < 1 + O(1/\sqrt{N_c})$ and $\tilde{\varphi}_{FCF} < 1 + O(1/\sqrt{N_c})$ are necessary conditions for convergence of MGRiT with F-relaxation and FCF-relaxation, respectively, on every iteration but one, with respect to error in the $\ell^2$-norm.*

As in section 2.2, worst-case convergence in the $\ell^2$-norm is given by constants $\tilde{\varphi}_F$ and $\tilde{\varphi}_{FCF}$ to $O(1/N_c)$.

[3] This assumption is likely a flaw in our line of proof and not actually necessary.

**2.4. Additional results for commuting diagonalizable operators.** As it turns out, results above can be strengthened in some sense under the additional assumption that $\Phi$ and $\Psi$ commute and are diagonalizable. This leads to exact bounds on convergence in a modified norm that are fairly tight for a large number of iterations, $p$, as well. By norm equivalence in finite-dimensional spaces, convergence in the modified norm is also necessary and sufficient for (asymptotic) convergence in the $\ell^2$- and $A^*A$-norms. The constants in norm equivalence depend on the conditioning of the eigenvectors. First, we introduce a less general approximation property based on the assumption that $\Phi$ and $\Psi$ commute and are diagonalizable.

DEFINITION 11 (Temporal eigenvalue approximation property). *Let $\Phi$ denote a fine-grid time-stepping operator and $\Psi$ denote a coarse-grid time-stepping operator, of size $N_x \times N_x$, for all time points, with coarsening factor $k$. Suppose that $\Phi$ and $\Psi$ commute and are diagonalizable, with eigenvalues given by $\{\lambda_\ell\}$ and $\{\mu_\ell\}$, respectively. Then, $\Phi$ satisfies an F-relaxation temporal eigenvalue approximation property (F-TEAP), with respect to $\Psi$, with constant $\varphi_F$, if, for $\ell = 0, \ldots, N_x - 1$,*

$$|\mu_\ell - \lambda_\ell^k| \leq \varphi_F(1 - |\mu_\ell|). \tag{14}$$

*Similarly, $\Phi$ satisfies an FCF-relaxation temporal eigenvalue approximation property (FCF-TEAP), with respect to $\Psi$, with constant $\varphi_{FCF}$, if, for $\ell = 0, \ldots, N_x - 1$,*

$$|\mu_\ell - \lambda_\ell^k| \leq \varphi_{FCF} \frac{1 - |\mu_\ell|}{|\lambda_\ell^k|}. \tag{15}$$

Note that for the temporal eigenvalue approximation property (TEAP), there is no distinction between powers, because scalars commute. Furthermore, assumption 2 implies that all eigenvalues $|\mu_i|, |\lambda_i| < 1$.

THEOREM 12 (The diagonalizable case: F-relaxation). *Let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, for all time points, with coarsening factor $k$, and $N_c$ time points on the coarse grid. Assume that $\Phi$ and $\Psi$ commute and are diagonalizable, with eigenvectors given as columns of $U$, and that $\Phi$ satisfies an F-TEAP with respect to $\Psi$, with constant $\varphi_F < 1$. Let $\mathbf{e}_{p+1}$ denote the error vector of Parareal/MGRiT with F-relaxation after $p+1$ iterations. Then, $\|\mathbf{e}_1\|^2_{(UU^*)^{-1}} \leq k\|\mathbf{e}_0\|^2_{(UU^*)^{-1}}$, and*

$$\|\mathbf{e}_{p+1}\|^2_{(UU^*)^{-1}} \leq \left(\varphi_F^{2p} - O(1/N_c^2)\right) \|\mathbf{e}_1\|^2_{(UU^*)^{-1}}. \tag{16}$$

*Furthermore, this bound is tight; that is, there exists an initial error $\mathbf{e}_0$ such that (16) holds with equality, to $O(1/N_c^2)$.*

*This also provides necessary and sufficient (asymptotic) conditions for convergence in the $\ell^2$- and $A^*A$-norms. That is, iterations may diverge at first but will eventually converge in the $\ell^2$- and $A^*A$-norms.*

THEOREM 13 (The diagonalizable case: FCF-relaxation). *Let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, for all time points, with coarsening factor $k$, and $N_c$ time points on the coarse grid. Assume that $\Phi$ and $\Psi$ commute and are diagonalizable, with eigenvectors given as columns of $U$, and that $\Phi$ satisfies an FCF-TEAP with respect to $\Psi$, with constant $\varphi_{FCF} < 1$. Let $\mathbf{e}_{p+1}$ denote the error vector of MGRiT with FCF-relaxation after $p+1$ iterations. Then, $\|\mathbf{e}_1\|^2_{(UU^*)^{-1}} \leq k\|\mathbf{e}_0\|^2_{(UU^*)^{-1}}$, and*

$$\|\mathbf{e}_{p+1}\|^2_{(UU^*)^{-1}} \leq \left(\varphi_{FCF}^{2p} - O(1/N_c^2)\right) \|\mathbf{e}_1\|^2_{(UU^*)^{-1}}. \tag{17}$$

*Furthermore, this bound is tight; that is, there exists an initial error* $\mathbf{e}_0$ *such that* (17) *holds with equality, to* $O(1/N_c^2)$.

*This also provides necessary and sufficient (asymptotic) conditions for convergence in the* $\ell^2$*- and* $A^*A$*-norms. That is, iterations may diverge at first but will eventually converge in the* $\ell^2$*- and* $A^*A$*-norms.*

Note that in the case of normal matrices, $U^{-1} = U^*$, and we have that the $(UU^*)^{-1}$-norm is exactly equal to the $\ell^2$-norm. In that case, the TEAP and TAP are equivalent, and we have an exact bound on Parareal and two-level MGRiT convergence of residual in the $\ell^2$-norm and error in the $\ell^2$- and $A^*A$-norms. In the commuting and diagonalizable case, these results are an extension of the upper bounds developed in [5].

Some of these results can be extended to the time-dependent case as well, such as when there are time-dependent reaction terms in a PDE or variable time-step size. Such scenarios are discussion in section 6.

*Remark* 14 (Convergence bounds observed in practice). It is worth pointing out that not only do the TAP and Theorems 2, 10, 12, and 13 define worst-observable convergence factors of Parareal and two-level MGRiT, but also such convergence factors are likely to be observed in practice. In the theoretical derivations that follow, convergence bounds are derived based on minimum and maximum eigenvalues or singular values of block-Toeplitz matrices. In many cases, it can be shown that there are clusters of singular modes or eigenmodes near these upper or lower bounds [29, 34, 35, 37], making them likely to be observed in practice. Numerical results confirming this for diagonalizable model problems can be found in [5], where the proposed upper bounds match the exact bounds of Theorems 12 and 13.

**2.5. Discussion on assumptions.** To remark on assumptions 4 and 5 from section 2.2, note that almost all time-integration routines (including all single-step Runge–Kutta-type methods) are rational functions of some invertible operator $\mathcal{L}$, where, for example, $\mathcal{L}$ is a scalar in the case of a standard ODE, or a spatial discretization operator in the case of a space-time PDE. Starting with assumption 5, if $\Phi$ and $\Psi$ are both functions of $\mathcal{L}$, assuming that they commute is a mild assumption, because any rational function of $\mathcal{L}$ commutes. This includes most time-integration schemes, including all single-step multistage Runge–Kutta-type schemes. Assumptions 5(a) and 5(b) then follow if, in addition, $\mathcal{L}$ is diagonalizable and normal, respectively. These are stronger assumptions that are satisfied, for example, in the case of many parabolic PDEs.

Returning to assumption 4, note again that it is believed that assumption 4 is not necessary and is rather a flaw in our line of proof. Nevertheless, here we use an example to show that assuming $\Psi - \Phi^k$ is invertible is reasonable anyway. Of course, because we want $\Psi \approx \Phi^k$, we *do not want* $\Psi - \Phi^k$ to be invertible. If $\Psi\mathbf{v} = \Phi^k\mathbf{v}$ for any vector $\mathbf{v}$, then indeed it is not invertible. However, in practice it is unlikely for $\Psi$ to *exactly* preserve a mode of $\Phi^k$ in this manner.

*Example* 15 (RK4 and $\Psi - \Phi^k$). Consider RK4 time integration with coarsening by a factor of two, where $\Phi$ corresponds to RK4 with time step $\delta t$, and $\Psi$ corresponds to RK4 with time step $2\delta t$ (the standard approach used in Parareal and MGRiT to approximate $\Phi^2$). Assume that $\Phi$ and $\Psi$ are stable. Then,

$$\Phi = I + \delta t\mathcal{L} + \tfrac{\delta t^2}{2}\mathcal{L}^2 + \tfrac{\delta t^3}{6}\mathcal{L}^3 + \tfrac{\delta t^4}{24}\mathcal{L}^4,$$
$$\Psi = I + 2\delta t\mathcal{L} + 2\delta t^2\mathcal{L}^2 + \tfrac{4\delta t^3}{3}\mathcal{L}^3 + \tfrac{2\delta t^4}{3}\mathcal{L}^4$$

$$
\begin{aligned}
&= \Phi^2 - \left[\tfrac{\delta t^5}{4}\mathcal{L}^5 + \tfrac{5\delta t^6}{72}\mathcal{L}^6 + \tfrac{\delta t^7}{72}\mathcal{L}^7 + \tfrac{\delta t^8}{576}\mathcal{L}^8\right] \\
&= \Phi^2 - \tfrac{\delta t^5}{4}\mathcal{L}^5\left[I + \tfrac{5\delta t}{18}\mathcal{L} + \tfrac{\delta t^2}{18}\mathcal{L}^2 + \tfrac{\delta t^3}{144}\mathcal{L}^3\right], \\
\Psi - \Phi^2 &= -\tfrac{\delta t^5}{4}\mathcal{L}^5\left[I + \tfrac{5\delta t}{18}\mathcal{L} + \tfrac{\delta t^2}{18}\mathcal{L}^2 + \tfrac{\delta t^3}{144}\mathcal{L}^3\right].
\end{aligned}
$$

Assuming $\mathcal{L}$ is nonsingular (which it should be for a well-posed problem), $\Psi - \Phi^2$ is only singular (noninvertible) if an eigenvalue $\lambda_i$ of $\mathcal{L}$ is exactly one of the three roots of

$$p(\lambda) = 1 + \tfrac{5\delta t}{18}\lambda + \tfrac{\delta t^2}{18}\lambda^2 + \tfrac{\delta t^3}{144}\lambda^3.$$

If such an eigenvalue does not exist, then $\Psi - \Phi^k$ is invertible. Working through the closed form for cubic roots, one can show that the roots of $p(\lambda)$ are approximately given by

$$\lambda_0 \approx -\frac{5.5}{\delta t}, \quad \lambda_1 \approx \frac{4.96\mathrm{i} - 0.2}{\delta t}, \quad \lambda_2 \approx \frac{-4.96\mathrm{i} - 0.2}{\delta t}.$$

Returning to the assumption that $\Phi$ is stable, a necessary condition for this is that all eigenvalues of $\Phi$ are less than one in magnitude. Thus suppose $\hat{\lambda} \in \{\lambda_0, \lambda_1, \lambda_2\}$ is an eigenvalue of $\mathcal{L}$ that is also a root of $p(\lambda)$. Applying $\Phi$ to the corresponding eigenvector $\hat{\mathbf{v}}$ shows that $\Phi$ must have an eigenvalue $\gg 1$. By contradiction, $\Psi - \Phi^k$ must be invertible if $\Phi$ is stable.

Example 15 proves that for RK4 with coarsening by a factor of two, $\Psi - \Phi^k$ must be invertible if $\Phi$ is stable. In general, if $\Phi$ and $\Psi$ are rational functions of some operator $\mathcal{L}$ (the same operator on each level), then $\Psi - \Phi^k$ is invertible as long as one of the eigenvalues of $\mathcal{L}$ is not a root of the difference of the two characteristic polynomials. If $\mathcal{L}$ has nonnegative eigenvalues, then this holds for *all* explicit Runge–Kutta schemes.

The analysis is more complicated for larger $k$ or if $\Phi$ and $\Psi$ are based on different operators (for example, if spatial coarsening is used [21]). However, the general heuristic stands that it is unlikely for $\Phi^k$ to exactly preserve a mode of $\Psi$. If, in fact, it does for a specific problem, a small perturbation to $\delta t$ would likely nullify that property, and the assumption that $\Psi - \Phi^k$ is invertible will stand.

## 3. Convergence theory framework.

**3.1. Error and residual propagation.** Let $\mathcal{E}$ denote the error-propagation operator and $\mathcal{R}$ the residual-propagation operator of Parareal or two-level MGRiT. These operators are derived analytically in this section. Note that for fixed-point iterative methods, error propagation takes the form $\mathcal{E} = I - M^{-1}A$, where $M$ is some approximation of $A$. From section 2.1, observe that

$$\mathbf{e}_i = \mathcal{E}^i\mathbf{e}_0 \quad \Longleftrightarrow \quad A^{-1}\mathbf{r}_i = \mathcal{E}^i A^{-1}\mathbf{r}_0 \quad \Longleftrightarrow \quad \mathbf{r}_i = (A\mathcal{E}A^{-1})^i\mathbf{r}_0.$$

Thus, residual propagation is formally similar to error propagation, where $\mathcal{R} = A\mathcal{E}A^{-1} = I - AM^{-1}$. In this form, error propagation is a measure of $M$ as a left approximate inverse of $A$ and residual propagation a measure of $M$ as a right approximate inverse of $A$. Then, observe that

(18)
$$\|\mathcal{R}\|^2 = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\langle\mathcal{R}\mathbf{x}, \mathcal{R}\mathbf{x}\rangle}{\langle\mathbf{x},\mathbf{x}\rangle} = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\langle A\mathcal{E}A^{-1}\mathbf{x}, A\mathcal{E}A^{-1}\mathbf{x}\rangle}{\langle\mathbf{x},\mathbf{x}\rangle} = \sup_{\mathbf{y}\neq\mathbf{0}} \frac{\langle A^*A\mathcal{E}\mathbf{y}, \mathcal{E}\mathbf{y}\rangle}{\langle A^*A\mathbf{y},\mathbf{y}\rangle} = \|\mathcal{E}\|^2_{A^*A};$$

that is, the norm of the residual-propagation operator in the $\ell^2$-norm is equivalent to that of the error-propagation operator in the $A^*A$-norm.[4] Note that this is consistent with the previous noted relation, $\|\mathbf{e}\|_{A^*A} = \|\mathbf{r}\|$.

**3.2. Reduction-based multigrid.** MGRiT and Parareal are both reduction-based multigrid algorithms. Multigrid methods are a class of iterative methods based on two parts: relaxation and coarse-grid correction, which are designed to be complementary in the sense that they each reduce different, complementary, error modes [3]. Error propagation of relaxation typically takes the form $I - M^{-1}A$, where $M$ is some easy to compute approximation of $A$, such as the diagonal or lower-triangular block. Coarse-grid correction is a subspace projection, for which error propagation takes the form $I - P(RAP)^{-1}RA$. Here, $R$ is a restriction operator, which restricts residuals on the fine grid to a coarse-grid problem; $RAP$ defines the coarse-grid problem to be solved; and $P$ is an interpolation operator to interpolate a correction back to the fine grid. Moving forward, at times we will just refer to MGRiT but will imply Parareal as well.

For most multigrid methods, the purpose of relaxation is to reduce error associated with highly oscillatory modes (large eigenvalues/singular values in the algebraic case). Coarse-grid correction is then complementary by reducing error associated with geometrically smooth modes, or small eigenvalues/singular values. In a reduction-based multigrid method, relaxation and coarse-grid correction instead reduce error associated with different degrees of freedom (DOFs), or, equivalently, blocks in the matrix. To this end, suppose all DOFs are partitioned into a disjoint covering of C-points and F-points, and matrices $A, P$, and $R$ take the following block forms:

$$A = \begin{bmatrix} A_{ff} & A_{fc} \\ A_{cf} & A_{cc} \end{bmatrix}, \quad P = \begin{bmatrix} W \\ I \end{bmatrix}, \quad R = \begin{bmatrix} Z & I \end{bmatrix}, \tag{19}$$

where the identity blocks in $P$ and $R$ are on the $n_c \times n_c$ C-point block. A simple two-level reduction-based multigrid method is given by letting $Z = -A_{cf}A_{ff}^{-1}$ and $W = \mathbf{0}$. In this case, coarse-grid correction yields zero error at C-points [25]. The restriction operator defined through $Z = -A_{cf}A_{ff}^{-1}$ is referred to as "ideal restriction," denoted $R_{\text{ideal}}$, where it is ideal in being the unique restriction operator that yields an exact coarse-grid correction at C-points. Following this with an exact solve on F-points as a relaxation scheme then yields an exact solution at F-points, without modifying the solution at C-points [26, 25]. Thus, the solution is exact and we have a two-grid reduction, where solving $A\mathbf{x} = \mathbf{b}$ is reduced to solving one system based on $A_{ff}$ and one system based on $RAP$.

MGRiT is also a reduction-based multigrid method, instead using the so-called *ideal interpolation* operator. Ideal interpolation, denoted $P_{\text{ideal}}$, is defined through $W = -A_{ff}^{-1}A_{fc}$. For symmetric positive definite matrices, ideal interpolation is ideal in a certain theoretical sense [10, 43]. In the nonsymmetric setting, ideal interpolation is ideal as the unique interpolation operator that eliminates the contribution of F-point residuals to the coarse-grid right-hand side [26, 8]. When coupled with $R = \begin{bmatrix} \mathbf{0} & I \end{bmatrix}$, referred to as restriction by injection, coarse-grid correction then yields zero residual at C-points. Note that an exact solve on F-points yields zero residual at F-points. Thus, coarse-grid correction with $P_{\text{ideal}}$ and restriction by injection, preceded by an

[4] In general, the $A^*A$-norm is considered a stronger norm, which is also consistent with the result that one can have an arbitrarily accurate left approximate inverse that makes for a poor right approximate inverse [28].

exact solve on F-points, also yields an exact two-level reduction [26].[5]

In the algebraic setting, $A_{ff}^{-1}$ is often not easily computed, so approximations are made, such as in AMG methods based on an approximate ideal restriction (AIR) [25, 26]. MGRiT and the system in (1) are unique in that the action of $A_{ff}^{-1}$ can be computed, so ideal interpolation and exact F-relaxation are feasible choices. In this case, assuming a block form as in (19), $RAP_{\text{ideal}} = S_A$, where $S_A$ is the Schur complement, independently of $R$ [26]. Although we can express a closed form for $S_A$ (see (31)), $S_A$ is not amenable to a recursive multilevel algorithm. In particular, one time step on the Schur-complement coarse grid simply consists of taking $k$ steps on the fine grid, which is no cheaper than solving the fine-grid problem directly. Because of this, MGRiT is based on a non-Galerkin coarse grid, where we approximate $\Phi^k$ with some other operator $\Psi$. Usually, this is accomplished by approximating $k$ steps of size $\delta t$ with one step of size $k\delta t$.

The following section derives error- and residual-propagation operators for MGRiT. Further details on reduction-based multigrid methods can be found in [25, 26, 24, 31], and further details on the MGRiT algorithm can be found in, for example, [5, 8, 9, 12]. The reduction properties rely on the ideal interpolation and restriction operators,

$$R_{\text{ideal}} = \begin{bmatrix} -A_{cf}A_{ff}^{-1} & I \end{bmatrix}, \qquad P_{\text{ideal}} = \begin{bmatrix} -A_{ff}^{-1}A_{fc} \\ I \end{bmatrix}.$$

**3.3. Error- and residual-propagation operators.** Consider residual and error propagation of two-level MGRiT, with a non-Galerkin coarse-grid operator, $B_\Delta^{-1}$, to approximate the Schur complement, $A_\Delta := S_A^{-1}$ ($A_\Delta$ is used to be consistent with previous works [5, 8]). Because MGRiT is based on ideal interpolation, here we use a prerelaxation scheme of F-relaxation or FCF-relaxation [26]. It is important to note that, in the case of the MGRiT algorithm, F-relaxation and C-relaxation refer to an exact solve on F- and C-points, respectively. This is how parallelization in time is achieved—an exact solve on F-points corresponds to using the current (spatial) solution at each C-point (in time) and integrating the spatial solution forward in time over the proceeding $k-1$ F-points. This local time integration is coupled with a global time integration on a coarse grid that is cheaper to evaluate ($B_\Delta^{-1}$) for a two-level parallel-in-time iterative method.

Recall that error propagation of relaxation and coarse-grid correction each take the form of a classic fixed-point method, $I - M^{-1}A$. The approximate inverses for an exact solve on F-points, an exact solve on C-points, and coarse-grid correction are given, respectively, by

$$M_F^{-1} = \begin{bmatrix} A_{ff}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad M_C^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_{cc}^{-1} \end{bmatrix},$$
$$M_{cgc}^{-1} = \begin{bmatrix} -A_{ff}^{-1}A_{fc} \\ I \end{bmatrix} B_\Delta^{-1} \begin{bmatrix} \mathbf{0} & I \end{bmatrix} = \begin{bmatrix} \mathbf{0} & -A_{ff}^{-1}A_{fc}B_\Delta^{-1} \\ \mathbf{0} & B_\Delta^{-1} \end{bmatrix}.$$

Then, error and residual propagation of two-level MGRiT with pre F-relaxation are given by $\mathcal{E}_F = I - (M_F^{-1} + M_{cgc}^{-1} - M_{cgc}^{-1}AM_F^{-1})A$ and $\mathcal{R}_F = I - A(M_F^{-1} + M_{cgc}^{-1} - M_{cgc}^{-1}AM_F^{-1})$, respectively. Note that

$$M_{cgc}^{-1}AM_F^{-1} = \begin{bmatrix} \mathbf{0} & -A_{ff}^{-1}A_{fc}B_\Delta^{-1} \\ \mathbf{0} & B_\Delta^{-1} \end{bmatrix} \begin{bmatrix} A_{ff} & A_{fc} \\ A_{cf} & A_{cc} \end{bmatrix} \begin{bmatrix} A_{ff}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

[5] Note that the ordering is important: Coarse-grid correction with $P_{\text{ideal}}$ and restriction by injection, followed by an exact solve on F-points, does not yield a two-grid reduction [26].

$$= \begin{bmatrix} -A_{ff}^{-1}A_{fc}B_{\Delta}^{-1}A_{cf}A_{ff}^{-1} & \mathbf{0} \\ B_{\Delta}^{-1}A_{cf}A_{ff}^{-1} & \mathbf{0} \end{bmatrix}.$$

Combining, we get

$$\mathcal{R}_F = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} - \begin{bmatrix} A_{ff} & A_{fc} \\ A_{cf} & A_{cc} \end{bmatrix} \begin{bmatrix} A_{ff}^{-1} + A_{ff}^{-1}A_{fc}B_{\Delta}^{-1}A_{cf}A_{ff}^{-1} & -A_{ff}^{-1}A_{fc}B_{\Delta}^{-1} \\ -B_{\Delta}^{-1}A_{cf}A_{ff}^{-1} & B_{\Delta}^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ -(I - A_{\Delta}B_{\Delta}^{-1})A_{cf}A_{ff}^{-1} & I - A_{\Delta}B_{\Delta}^{-1} \end{bmatrix} \tag{20}$$

$$= \begin{bmatrix} \mathbf{0} \\ I - A_{\Delta}B_{\Delta}^{-1} \end{bmatrix} R_{\text{ideal}}, \tag{21}$$

$$\mathcal{E}_F = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} - \begin{bmatrix} A_{ff}^{-1} + A_{ff}^{-1}A_{fc}B_{\Delta}^{-1}A_{cf}A_{ff}^{-1} & -A_{ff}^{-1}A_{fc}B_{\Delta}^{-1} \\ -B_{\Delta}^{-1}A_{cf}A_{ff}^{-1} & B_{\Delta}^{-1} \end{bmatrix} \begin{bmatrix} A_{ff} & A_{fc} \\ A_{cf} & A_{cc} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} & -A_{ff}^{-1}A_{fc}(I - B_{\Delta}^{-1}A_{\Delta}) \\ \mathbf{0} & I - B_{\Delta}^{-1}A_{\Delta} \end{bmatrix} \tag{22}$$

$$= P_{\text{ideal}} \begin{bmatrix} \mathbf{0} & I - B_{\Delta}^{-1}A_{\Delta} \end{bmatrix}. \tag{23}$$

To consider FCF-relaxation, note that MGRiT residual and error propagation for FCF-relaxation are equivalent to multiplying $\mathcal{R}_F$ and $\mathcal{E}_F$ by residual and error propagation for FC-relaxation, which are respectively given by

$$I - A(M_F^{-1} + M_C^{-1} - M_C^{-1}AM_F^{-1}) = I - \begin{bmatrix} A_{ff} & A_{fc} \\ A_{cf} & A_{cc} \end{bmatrix} \begin{bmatrix} A_{ff}^{-1} & \mathbf{0} \\ -A_{cc}^{-1}A_{cf}A_{ff}^{-1} & A_{cc}^{-1} \end{bmatrix}$$
$$= \begin{bmatrix} A_{fc}A_{cc}^{-1}A_{cf}A_{ff}^{-1} & -A_{fc}A_{cc}^{-1} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$
$$I - (M_F^{-1} + M_C^{-1} - M_C^{-1}AM_F^{-1})A = I - \begin{bmatrix} A_{ff}^{-1} & \mathbf{0} \\ -A_{cc}^{-1}A_{cf}A_{ff}^{-1} & A_{cc}^{-1} \end{bmatrix} \begin{bmatrix} A_{ff} & A_{fc} \\ A_{cf} & A_{cc} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{0} & -A_{ff}^{-1}A_{fc} \\ \mathbf{0} & A_{cc}^{-1}A_{cf}A_{ff}^{-1}A_{fc} \end{bmatrix}.$$

It follows that residual and error propagation of two-level MGRiT with pre FCF-relaxation are given by

$$\mathcal{R}_{FCF} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ -(I - A_{\Delta}B_{\Delta}^{-1})A_{cf}A_{ff}^{-1} & I - A_{\Delta}B_{\Delta}^{-1} \end{bmatrix} \begin{bmatrix} A_{fc}A_{cc}^{-1}A_{cf}A_{ff}^{-1} & -A_{fc}A_{cc}^{-1} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ -(I - A_{\Delta}B_{\Delta}^{-1})A_{cf}A_{ff}^{-1}A_{fc}A_{cc}^{-1}A_{cf}A_{ff}^{-1} & (I - A_{\Delta}B_{\Delta}^{-1})A_{cf}A_{ff}^{-1}A_{fc}A_{cc}^{-1} \end{bmatrix} \tag{24}$$

$$= \begin{bmatrix} \mathbf{0} \\ (I - A_{\Delta}B_{\Delta}^{-1})A_{cf}A_{ff}^{-1}A_{fc}A_{cc}^{-1} \end{bmatrix} R_{\text{ideal}}, \tag{25}$$

$$\mathcal{E}_{FCF} = \begin{bmatrix} \mathbf{0} & -A_{ff}^{-1}A_{fc}(I - B_{\Delta}^{-1}A_{\Delta}) \\ \mathbf{0} & I - B_{\Delta}^{-1}A_{\Delta} \end{bmatrix} \begin{bmatrix} \mathbf{0} & -A_{ff}^{-1}A_{fc} \\ \mathbf{0} & A_{cc}^{-1}A_{cf}A_{ff}^{-1}A_{fc} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} & -A_{ff}^{-1}A_{fc}(I - B_{\Delta}^{-1}A_{\Delta})A_{cc}^{-1}A_{cf}A_{ff}^{-1}A_{fc} \\ \mathbf{0} & (I - B_{\Delta}^{-1}A_{\Delta})A_{cc}^{-1}A_{cf}A_{ff}^{-1}A_{fc} \end{bmatrix} \tag{26}$$

$$= P_{\text{ideal}} \begin{bmatrix} \mathbf{0} & (I - B_{\Delta}^{-1}A_{\Delta})A_{cc}^{-1}A_{cf}A_{ff}^{-1}A_{fc} \end{bmatrix}. \tag{27}$$

Note from (18) that $\|\mathcal{E}_F^p\|_{A^*A} = \|\mathcal{R}_F^p\|$ and $\|\mathcal{E}_{FCF}^p\|_{A^*A} = \|\mathcal{R}_{FCF}^p\|$ for $p \geq 1$.

**3.4. MGRiT matrices.** So far, derivations have been purely algebraic and assumed no structure of the linear system. Focusing on the MGRiT framework, consider the MGRiT system (1) and suppose we coarsen in time by a factor of $k$. This corresponds to partitioning time-points into C-points and F-points, such that for every $k$ points, $k-1$ are F-points. For convenience, assume that the first and last time-points are C-points, in which case the total number of C-points is given by $N_c = 1 + \frac{N-1}{k}$, where $N$ is the total number of time-points.[6] Then, blocks in an FC-partitioning of the matrix $A$ (19) take the following form:

$$A_{ff} = \begin{bmatrix} I & & & & & & \\ -\Phi & I & & & & & \\ & \ddots & \ddots & & & & \\ & & -\Phi & I & & & \\ & & & & \ddots & & \\ & & & & I & & \\ & & & & -\Phi & I & \\ & & & & & \ddots & \ddots \\ & & & & & & -\Phi \quad I \end{bmatrix}, \quad A_{fc} = \begin{bmatrix} -\Phi & & & \mathbf{0} \\ \mathbf{0} & & & \vdots \\ \vdots & & & \vdots \\ \mathbf{0} & & & \mathbf{0} \\ & \ddots & & \vdots \\ & & -\Phi & \mathbf{0} \\ & & \mathbf{0} & \vdots \\ & & \vdots & \vdots \\ & & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$A_{cf} = \begin{bmatrix} \mathbf{0} & \dots & & & & & & \\ \mathbf{0} & \dots & \mathbf{0} & -\Phi & & & & \\ & & & & \ddots & & & \\ & & & & & \mathbf{0} & \dots & \mathbf{0} \quad -\Phi \end{bmatrix}, \qquad A_{cc} = \begin{bmatrix} I & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix}.$$

Dotted lines are used to highlight the block nature, where the $k-1$ F-points in each group are adjacent in the time domain, while all C-points are disconnected. Next, further matrix forms that arise in residual and error propagation are derived:

(28)

$$(A_{ff})^{-1} = \begin{bmatrix} I & & & & & & \\ \Phi & I & & & & & \\ \vdots & \ddots & \ddots & & & & \\ \Phi^{k-2} & \dots & \Phi & I & & & \\ & & & & \ddots & & \\ & & & & I & & \\ & & & & \Phi & I & \\ & & & & \vdots & \ddots & \ddots \\ & & & & \Phi^{k-2} & \dots & \Phi \quad I \end{bmatrix},$$

(29)

$$-A_{cf}(A_{ff})^{-1} = \begin{bmatrix} \mathbf{0} & \dots & & & & & & \\ \Phi^{k-1} & \Phi^{k-2} & \dots & \Phi & & & & \\ & & & & \ddots & & & \\ & & & & & \Phi^{k-1} & \Phi^{k-2} & \dots \quad \Phi \end{bmatrix},$$

[6] This is slightly different notation from that used in [5]. There, it is assumed that $N_c = N/k$; however, the coarse grid then has $N_c + 1$ points. Here, $N_c$ exactly denotes the number of coarse-grid time points, at the expense of a slightly more complicated relation between $N$ and $N_c$.

(30)
$$-(A_{ff})^{-1}A_{fc} = \left[\begin{array}{c|c|c|c} \Phi & & & \mathbf{0} \\ \Phi^2 & & & \vdots \\ \vdots & & & \vdots \\ \Phi^{k-1} & & & \mathbf{0} \\ \hline & \ddots & & \vdots \\ \hline & & \Phi & \mathbf{0} \\ & & \Phi^2 & \vdots \\ & & \vdots & \vdots \\ & & \Phi^{k-1} & \mathbf{0} \end{array}\right], \qquad A_{cf}(A_{ff})^{-1}A_{fc} = \left[\begin{array}{c|c|c|c} \mathbf{0} & & & \\ \hline \Phi^k & \mathbf{0} & & \\ \hline & \ddots & \ddots & \\ \hline & & \Phi^k & \mathbf{0} \end{array}\right].$$

Note from (30) that the action of $XA_{cf}(A_{ff})^{-1}A_{fc}$ shifts all columns of $X$ to the left and scales all entries by $\Phi^k$. This will be useful in future derivations.

Recall that $RAP_{\text{ideal}}$ is given by the Schur complement of $A$, independently of $R$. To be consistent with [5, 8, 11], denote $A_\Delta := S_A = RAP_{\text{ideal}}$. From above, it follows that

$$A_\Delta = A_{cc} - A_{cf}(A_{ff})^{-1}A_{fc} = \begin{bmatrix} I & & & \\ -\Phi^k & I & & \\ & \ddots & \ddots & \\ & & -\Phi^k & I \end{bmatrix}. \tag{31}$$

Observe that the coarse-grid operator consists of taking $k$ time steps with the time-stepping function $\Phi$. Because this is no cheaper to evaluate than $k$ individual steps of $\Phi$—that is, propagating $k$ steps on the fine grid—a non-Galerkin coarse grid is used. Instead of taking $k$ time steps of size $\delta t$, corresponding to $\Phi^k$, $k$ steps are approximated by some operator $\Psi$,

$$\begin{aligned} A_\Delta B_\Delta^{-1} &= \begin{bmatrix} I & & & \\ -\Phi^k & I & & \\ & \ddots & \ddots & \\ & & -\Phi^k & I \end{bmatrix} \begin{bmatrix} I & & & & \\ \Psi & I & & & \\ \Psi^2 & \Psi & I & & \\ \vdots & \vdots & \ddots & I & \\ \Psi^{N_c-1} & \Psi^{N_c-2} & \dots & \Psi & I \end{bmatrix} \\ &= \begin{bmatrix} I & & & & \\ \Psi - \Phi^k & I & & & \\ (\Psi - \Phi^k)\Psi & \Psi - \Phi^k & I & & \\ \vdots & \vdots & \ddots & \ddots & \\ (\Psi - \Phi^k)\Psi^{N_c-2} & (\Psi - \Phi^k)\Psi^{N_c-3} & \dots & \Psi - \Phi^k & I \end{bmatrix}, \end{aligned} \tag{32}$$

$$I - A_\Delta B_\Delta^{-1} = \operatorname{diag}(\Psi - \Phi^k) \begin{bmatrix} \mathbf{0} & & & & \\ I & \mathbf{0} & & & \\ \Psi & I & \mathbf{0} & & \\ \vdots & \vdots & \ddots & \ddots & \\ \Psi^{N_c-2} & \Psi^{N_c-3} & \dots & I & \mathbf{0} \end{bmatrix}, \tag{33}$$

$$I - B_\Delta^{-1}A_\Delta = \begin{bmatrix} \mathbf{0} & & & & \\ I & \mathbf{0} & & & \\ \Psi & I & \mathbf{0} & & \\ \vdots & \vdots & \ddots & \ddots & \\ \Psi^{N_c-2} & \Psi^{N_c-3} & \dots & I & \mathbf{0} \end{bmatrix} \operatorname{diag}(\Psi - \Phi^k). \tag{34}$$

Note that if $\Phi$ and $\Psi$ commute, then $I - B_\Delta^{-1}A_\Delta = I - A_\Delta B_\Delta^{-1}$.

**4. The general case.** This section derives a sequence of linear algebra lemmas, which are then used to present and prove a more precise version of Theorem 2. The underlying idea is that the $\ell^2$-norm of an operator is given by the largest singular value, which is also equivalent to one divided by the smallest nonzero singular value of the respective pseudoinverse. Here, we rely on block-Toeplitz matrix theory to place tight bounds on the maximum and minimum singular values of operators related to error and residual propagation.

From section 3.3 and (21), (23), (25), and (27), error- and residual-propagation operators for $p$ iterations of two-level MGRiT, with F-relaxation and FCF-relaxation, take the following forms:

$$\mathcal{E}_F^p = P_{\text{ideal}}(I - B_\Delta^{-1}A_\Delta)^p, \qquad \mathcal{E}_{FCF}^p = P_{\text{ideal}}\left((I - B_\Delta^{-1}A_\Delta)A_{cf}A_{ff}^{-1}A_{fc}\right)^p, \tag{35}$$

$$\mathcal{R}_F^p = (I - A_\Delta B_\Delta^{-1})^p R_{\text{ideal}}, \qquad \mathcal{R}_{FCF}^p = \left((I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right)^p R_{\text{ideal}}, \tag{36}$$

where matrices are as in section 3.4. Notice that convergence over $p > 1$ iterations in all cases is determined by bounding either $\|(I - A_\Delta B_\Delta^{-1})^p\| < 1$ for F-relaxation or $\|((I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc})^p\| < 1$ for FCF-relaxation. The leading (trailing) factor of $R_{\text{ideal}}$ ($P_{\text{ideal}}$) accounts for the single iteration in Theorem 2 on which convergence may not be observed. The following lemma proves that $\|R_{\text{ideal}}\|, \|P_{\text{ideal}}\| < \sqrt{k}$ if $\Phi$ is strongly stable.

LEMMA 16 (Bounds on $\|R_{\text{ideal}}\|, \|P_{\text{ideal}}\|$). *Let* $\|\Phi\| < 1$. *Then,*

$$\|R_{\text{ideal}}\| = \|P_{\text{ideal}}\| < \sqrt{k}.$$

*Proof.* From (29), note that $\|R_{\text{ideal}}\| = \sigma_{max}(R_{\text{ideal}}) = \sqrt{\lambda_{max}(R_{\text{ideal}}R_{\text{ideal}}^*)}$, where $R_{\text{ideal}}R_{\text{ideal}}^*$ is block diagonal, with an identity in the first block, and the rest given by $\sum_{\ell=0}^{k-1}\Phi^\ell(\Phi^\ell)^*$. Then,

$$\|R_{\text{ideal}}\| = \sqrt{\left\|\sum_{\ell=0}^{k-1}\Phi^\ell(\Phi^\ell)^*\right\|} \le \sqrt{\sum_{\ell=0}^{k-1}\|\Phi^\ell(\Phi^\ell)^*\|} = \sqrt{\sum_{\ell=0}^{k-1}\|\Phi^\ell\|^2} < \sqrt{k}.$$

An analogous derivation confirms that $\|P_{\text{ideal}}\| < \sqrt{k}$. ∎

Note that Lemma 16 is not necessarily tight, but it is sufficient to prove that error cannot diverge significantly in the $A^*A$- or $\ell^2$-norm in the first/last iteration.

**4.1. Residual propagation and $I - A_\Delta B_\Delta^{-1}$.** Now, we consider the maximum singular value of $I - A_\Delta B_\Delta^{-1}$ and $(I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}$, which arises in residual propagation. From (30) and (33), it is clear that both of these operators are block-Toeplitz matrices. Appealing to block-Toeplitz matrix theory, asymptotically (in $N_c$) tight bounds can be placed on the maximum singular value by way of considering the

operator's generator function. Let $\alpha_i$ denote the (potentially matrix-valued) Toeplitz coefficient for the $i$th diagonal of a (block) Toeplitz matrix, where $\alpha_0$ is the diagonal, $\alpha_{-1}$ the first subdiagonal, and so on. Then the Toeplitz matrix corresponds with a Fourier generator function,

$$F(x) = \sum_{i=-\infty}^{\infty} \alpha_i e^{-\mathrm{i}x}.$$

The following theorems introduce specific results from the field of block-Toeplitz operator theory, which prove to be important in further derivations.

THEOREM 17 (Minimum eigenvalue of Hermitian block-Toeplitz operators [29, 34, 36]). *Let $T_N(F)$ be an $N \times N$ Hermitian block-Toeplitz matrix, with self-adjoint, continuous generating function $F(x) : [0, 2\pi] \to \mathbb{C}^{m\times m}$, where $F(x) = F(x)^*$, and the minimum eigenvalue of $F(x)$ is not constant. Then, the smallest eigenvalue of $T_N(F)$ is given by*

$$\lambda_{\min}(T_N) = \min_{x\in[0,2\pi]} \lambda_{\min}(F(x)) + O(N^{-\alpha}),$$

*where $\alpha > 0$ is the order of the highest-order zero in $x$ of*

$$\lambda_{\min}(F(x)) - \left[\min_{y\in[0,2\pi]} \lambda_{\min}(F(y))\right].$$

THEOREM 18 (Maximum singular value of block-Toeplitz operators [38]). *Let $T_N(F)$ be an $N \times N$ block-Toeplitz matrix, with continuous generating function $F(x) : [0, 2\pi] \to \mathbb{C}^{m\times m}$. Then, the maximum singular value is bounded above by*

$$\sigma_{max}(T_N(F)) \leq \max_{x\in[0,2\pi]} \sigma_{max}(F(x))$$

*for all $N \in \mathbb{N}$.*

Theorem 18 is now used in the following theorem to derive upper bounds on the maximum singular values of interest.

THEOREM 19 (Sufficient conditions). *Let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, with coarsening factor $k$ and $N_c$ coarse-grid time points. Assume that $\Phi$ satisfies an F-TAP$_1$ with respect to $\Psi$, with constant $\varphi_{F,1}$. Then,*

$$\left\|I - A_\Delta B_\Delta^{-1}\right\| \leq \varphi_{F,1}\left(1 + \|\Psi^{N_c}\|\right).$$

*Similarly, assume that $\Phi$ satisfies an FCF-TAP$_1$ with respect to $\Psi$, with constant $\varphi_{FCF,1}$. Then,*

$$\left\|(I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right\| \leq \varphi_{FCF,1}\left(1 + \|\Phi^{-k}\Psi^{N_c}\Phi^k\|\right).$$

*Proof.* Notice from (33) that $I - A_\Delta B_\Delta^{-1}$ is a block-Toeplitz matrix with generating coefficients $\alpha_i = (\Psi - \Phi^k)\Psi^{-(1+i)}$ for $i = -1, \ldots, -N_c$ and $\alpha_i = \mathbf{0}$ for $i \geq 0$. Let $F_F(x)$ denote this generating function. First, note that if $\|\Psi^p\| < 1$ for some $p$, all eigenvalues of $\Psi$ must be less than one in magnitude. It then follows that $I - e^{\mathrm{i}x}\Psi$ is invertible for all $x$. If it were singular, then for some $x$ and $\mathbf{v}$, we would have

$e^{-\mathrm{i}x}\mathbf{v} = \Psi\mathbf{v}$, which contradicts that all eigenvalues of $\Psi$ must be less than one in magnitude. Then,

$$\begin{aligned} F_F(x) &= (\Psi - \Phi^k)\sum_{\ell=1}^{N_c} \Psi^{\ell-1} e^{\mathrm{i}\ell x} \\ &= (\Psi - \Phi^k)e^{\mathrm{i}x}\sum_{\ell=0}^{N_c-1} \Psi^{\ell} e^{\mathrm{i}\ell x} \\ &= e^{\mathrm{i}x}(\Psi - \Phi^k)(I - e^{\mathrm{i}N_c x}\Psi^{N_c})(I - e^{ix}\Psi)^{-1}. \end{aligned}$$

Recall that, under the assumption of an F-TAP$_1$, $\|(\Psi-\Phi^k)\mathbf{v}\| \le \varphi_{F,1}[\min_{x\in[0,2\pi]} \|(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\|]$ for all $\mathbf{v}$. Theorem 18 then yields

$$\begin{aligned} \|I - A_\Delta B_\Delta^{-1}\| &\le \max_{x\in[0,2\pi]} \sigma_{max}(F_F(x)) \\ &= \max_{\substack{x\in[0,2\pi],\\ \mathbf{v}\neq\mathbf{0}}} \frac{\left\|(\Psi - \Phi^k)(I - e^{\mathrm{i}N_c x}\Psi_c^N)(I - e^{\mathrm{i}x}\Psi)^{-1}\mathbf{v}\right\|}{\|\mathbf{v}\|} \\ &\le \max_{\mathbf{v}\neq\mathbf{0}} \frac{\left\|(\Psi - \Phi^k)\mathbf{v}\right\| + \left\|(\Psi - \Phi^k)\Psi^{N_c}\mathbf{v}\right\|}{\min_{x\in[0,2\pi]} \|(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\|} \\ &\le \max_{\mathbf{v}\neq\mathbf{0}} \varphi_{F,1} + \varphi_{F,1}\frac{\min_{x\in[0,2\pi]} \left\|(I - e^{\mathrm{i}x}\Psi)\Psi^{N_c}\mathbf{v}\right\|}{\min_{x\in[0,2\pi]} \|(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\|} \\ &\le \max_{\mathbf{v}\neq\mathbf{0}} \varphi_{F,1} + \varphi_{F,1}\frac{\min_{x\in[0,2\pi]} \left\|\Psi^{N_c}(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\right\|}{\min_{x\in[0,2\pi]} \|(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\|} \\ &= \varphi_{F,1}(1 + \|\Psi^{N_c}\|). \end{aligned} \tag{37}$$

A similar proof follows for the case of FCF-relaxation, where the generator function, $F_{FCF}$, has coefficients $\alpha_i = (\Psi - \Phi^k)\Psi^{-(1+i)}\Phi^k$ for $i = -1, \ldots, -N_c$ and $\alpha_i = \mathbf{0}$ for $i \ge 0$. Then, by assumption of an FCF-TAP$_1$ with constant $\varphi_{FCF,1}$,

$$\begin{aligned} \left\|(I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right\| &\le \max_{x\in[0,2\pi]} \sigma_{max}(F_{FCF}(x)) \\ &= \max_{\substack{x\in[0,2\pi],\\ \mathbf{v}\neq\mathbf{0}}} \frac{\left\|(\Psi - \Phi^k)(I - e^{\mathrm{i}Nx}\Psi^{N_c})(I - e^{\mathrm{i}x}\Psi)^{-1}\Phi^k\mathbf{v}\right\|}{\|\mathbf{v}\|} \\ &\le \max_{\mathbf{v}\neq\mathbf{0}} \frac{\left\|(\Psi - \Phi^k)\mathbf{v}\right\| + \left\|(\Psi - \Phi^k)\Psi^{N_c}\mathbf{v}\right\|}{\min_{x\in[0,2\pi]} \|\Phi^{-k}(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\|} \\ &\le \max_{\mathbf{v}\neq\mathbf{0}} \varphi_{FCF,1} + \varphi_{FCF,1}\frac{\min_{x\in[0,2\pi]} \left\|\Phi^{-k}(I - e^{\mathrm{i}x}\Psi)\Psi^{N_c}\mathbf{v}\right\|}{\min_{x\in[0,2\pi]} \|\Phi^{-k}(I - e^{\mathrm{i}x}\Psi)\mathbf{v}\|} \\ &= \max_{\mathbf{v}\neq\mathbf{0}} \varphi_{FCF,1} + \varphi_{FCF,1}\frac{\left\|\Phi^{-k}\Psi^{N_c}\Phi^k\mathbf{v}\right\|}{\|\mathbf{v}\|} \\ &= \varphi_{FCF,1}(1 + \|\Phi^{-k}\Psi^{N_c}\Phi^k\|). \end{aligned}$$

∎

Next, a more technical path is pursued, where the maximum singular values of $(I - A_\Delta B_\Delta^{-1})^p$ and $((I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc})^p$ are analyzed by means of the minimum singular value of the respective pseudoinverses. First, a pseudoinverse is derived for operators of the form $(I - A_\Delta B_\Delta^{-1})^p$ and $((I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc})^p$ for $p \ge 1$. These

pseudodinverses almost take the form of finite block Toeplitz matrices, and we appeal again to Toeplitz matrix theory to bound the smallest nonzero singular value from above.

First, some general pseudoinverses and their properties are derived. Let $f, g$, and $h$ be invertible scalars or operators, and define

$$A_0 = \begin{bmatrix} g & & & & \\ & g & & & \\ & & g & & \\ & & & g & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{0} & & & & & \\ I & \mathbf{0} & & & & \\ f & I & \mathbf{0} & & & \\ f^2 & f & I & \mathbf{0} & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \end{bmatrix} \begin{bmatrix} h & & & & \\ & h & & & \\ & & h & & \\ & & & h & \\ & & & & \ddots \end{bmatrix} \tag{38}$$

$$= \begin{bmatrix} \mathbf{0} & & & & & \\ gh & \mathbf{0} & & & & \\ gfh & gh & \mathbf{0} & & & \\ gf^2h & gfh & gh & \mathbf{0} & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \end{bmatrix},$$

$$A_1 = \begin{bmatrix} g & & & & \\ & g & & & \\ & & g & & \\ & & & g & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{0} & & & & & \\ \mathbf{0} & \mathbf{0} & & & & \\ I & \mathbf{0} & \mathbf{0} & & & \\ f & I & \mathbf{0} & \mathbf{0} & & \\ f^2 & f & I & \mathbf{0} & \mathbf{0} & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} h & & & & \\ & h & & & \\ & & h & & \\ & & & h & \\ & & & & \ddots \end{bmatrix} \tag{39}$$

$$= \begin{bmatrix} \mathbf{0} & & & & & \\ \mathbf{0} & \mathbf{0} & & & & \\ gh & \mathbf{0} & \mathbf{0} & & & \\ gfh & gh & \mathbf{0} & \mathbf{0} & & \\ gf^2h & gfh & gh & \mathbf{0} & \mathbf{0} & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

Note that (38) and (39) are general matrix forms, which encompass the coarse-grid correction and diagonal blocks of two-grid residual and error propagation for MGRiT, with F- and FCF-relaxation; in particular, $I - A_\Delta B_\Delta^{-1}$ (33) takes the form of (38) and $(I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}$ (30), (33) takes the form of (39), with an additional zero row due to FCF-relaxation (30). To construct pseudoinverses for operators of these forms, recall the four properties that define a pseudoinverse: $AA^\dagger A = A$, $A^\dagger AA^\dagger = A^\dagger$, $(AA^\dagger)^* = AA^\dagger$, and $(A^\dagger A)^* = A^\dagger A$. The subtle part of constructing a pseudoinverse is preserving the (not full rank) image and kernel of $A$. However, matrices in (38) and (39) have the nice property that they are full rank in the first $n-1$ or $n-2$ rows and columns, respectively. In the case of (38), note that

$$\begin{bmatrix} \mathbf{0} & & & & \\ gh & \mathbf{0} & & & \\ gfh & gh & \mathbf{0} & & \\ gf^2h & gfh & gh & \mathbf{0} & \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} = \begin{bmatrix} \mathbf{0} & & & & \\ & I & & & \\ & & I & & \\ & & & I & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{0} & & & & \\ gh & \mathbf{0} & & & \\ gfh & gh & \mathbf{0} & & \\ gf^2h & gfh & gh & \mathbf{0} & \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} & & & & \\ gh & \mathbf{0} & & & \\ gfh & gh & \mathbf{0} & & \\ gf^2h & gfh & gh & \mathbf{0} & \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I & & & & \\ & I & & & \\ & & \ddots & & \\ & & & I & \\ & & & & \mathbf{0} \end{bmatrix}.$$

Then, if we can build a matrix $\hat{A}_0^\dagger$ such that

(40)
$$A_0\hat{A}_0^\dagger = \begin{bmatrix} \mathbf{0} & & & & \\ & I & & & \\ & & I & & \\ & & & I & \\ & & & & \ddots \end{bmatrix}, \quad \hat{A}_0^\dagger A_0 = \begin{bmatrix} I & & & & \\ & I & & & \\ & & \ddots & & \\ & & & I & \\ & & & & \mathbf{0} \end{bmatrix}, \quad \hat{A}_0^\dagger \begin{bmatrix} \mathbf{0} & & & & \\ & I & & & \\ & & I & & \\ & & & I & \\ & & & & \ddots \end{bmatrix} = \hat{A}_0^\dagger,$$

it follows that all four properties of a pseudoinverse are satisfied. A similar result holds for $A_1$. We now have all the tools needed to construct a pseudoinverse of $A_0$ and $A_1$, which is summarized in the following lemma.

LEMMA 20. *Let $f$, $g$, and $h$ be invertible operators*[7] *and $A_0$ and $A_1$ be matrices defined as in* (38) *and* (39)*, respectively. Then, the unique pseudoinverses of $A_0$ and $A_1$ are given by*

(41)
$$A_0^\dagger = \begin{bmatrix} \mathbf{0} & & & & \\ gh & \mathbf{0} & & & \\ gfh & gh & \mathbf{0} & & \\ gf^2h & gfh & gh & \mathbf{0} & \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}^\dagger = \begin{bmatrix} \mathbf{0} & h^{-1}g^{-1} & & & \\ \mathbf{0} & -h^{-1}fg^{-1} & h^{-1}g^{-1} & & \\ \vdots & & \ddots & \ddots & \\ & & & -h^{-1}fg^{-1} & h^{-1}g^{-1} \\ & & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

(42)
$$A_1^\dagger = \begin{bmatrix} \mathbf{0} & & & & & \\ \mathbf{0} & \mathbf{0} & & & & \\ gh & \mathbf{0} & \mathbf{0} & & & \\ gfh & gh & \mathbf{0} & \mathbf{0} & & \\ gf^2h & gfh & gh & \mathbf{0} & \mathbf{0} & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}^\dagger = \begin{bmatrix} \mathbf{0} & \mathbf{0} & h^{-1}g^{-1} & & & \\ \mathbf{0} & \mathbf{0} & -h^{-1}fg^{-1} & h^{-1}g^{-1} & & \\ \vdots & \vdots & & \ddots & \ddots & \\ & & & & -h^{-1}fg^{-1} & h^{-1}g^{-1} \\ & & & \cdots & \mathbf{0} & \mathbf{0} \\ & & & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

*Proof.* The third relation in (40) simply requires that the first column of $A_0^\dagger$ is zero. Working through the system of equations established by the first two relations (and similar relations for $A_1$) yields the operators in (41) and (42). ∎

Notice that the pseudoinverse of $A_1$ is effectively equivalent to that of $A_0$, except with an additional zero row and column. The only difference this leads to in the final results is an $O(1/\sqrt{N_c})$ perturbation versus an $O(1/\sqrt{N_c - 1})$ perturbation. For moderate to large $N_c$, this difference is arbitrary, and, for simplicity's sake, F-relaxation and FCF-relaxation are both treated in the form $A_0$ moving forward.

The following lemma generalizes this result, deriving the pseudoinverse for operators of the form in (38) raised to powers. In the context of MGRiT, this corresponds

[7] This is where assumption 4, that $\Psi - \Phi^k$ is invertible, comes in. If a pseudoinverse must be used for $g$ or $h$ instead of a formal inverse, the resulting pseudoinverses of $A_0$ and $A_1$ do not take on such simple forms. For example, simply replacing $g^{-1}$ with $g^\dagger$ in (41) fails to satisfy the self-adjoint property $(A_0^\dagger A_0)^* = A_0^\dagger A_0$, although it does satisfy the other three.

to powers of error and residual propagation, which define how error and residual are propagated over multiple iterations. For nonnormal operators $\Phi$ and $\Psi$, it is possible that, for example, $\|\mathcal{E}^p\| < \|\mathcal{E}\|^p$. In fact, it is possible that $\|\mathcal{E}\| > 1$ appears divergent, but raising to powers results in a convergent method.

LEMMA 21 (Pseudoinverse for matrix powers). *Let $A_0$ be as in* (38)*, and define the Toeplitz matrix*

$$\mathcal{T}_0 =: \begin{bmatrix} -h^{-1}fg^{-1} & h^{-1}g^{-1} & & & \\ & -h^{-1}fg^{-1} & h^{-1}g^{-1} & & \\ & & \ddots & \ddots & \\ & & & -h^{-1}fg^{-1} & h^{-1}g^{-1} \\ & & & & -h^{-1}fg^{-1} \end{bmatrix}.$$

*Then, for $p \geq 1$,*

$$(A_0^p)^\dagger = \begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix} \mathcal{T}_0^p \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix}.$$

*Proof.* The case of $p = 1$ was proven in Lemma 20, with the pseudoinverse denoted by $A_0^\dagger$. Now let $A_l^\dagger$ and $A_r^\dagger$ be tentative left and right pseudoinverses for $A_0^p$, $p > 1$. We start the proof by building $A_l^\dagger$ and $A_r^\dagger$ to satisfy certain properties of the pseudoinverse, and we conclude by merging them in a certain way to derive $(A_0^p)^\dagger$.

First, note that $A_0^p$ is a strictly lower triangular matrix, with zeros on the diagonal and first $p-1$ subdiagonals. Building on the proof of Lemma 20, let us build $A_l^\dagger$ such that $A_l^\dagger A_0^p$ is diagonal, with zeros on the last $p$ entries, and the identity elsewhere (similar to (40)). This immediately satisfies two properties of a pseudoinverse, $(A_l^\dagger A_0^p)^* = A_l^\dagger A_0^p$ and $A_0^p A_l^\dagger A_0^p = A_0^p$. To do so, let us start by considering $(A_0^\dagger)^p$ as a naive attempt at a pseudoinverse for $A_0^p$ and observing that

$$(A_0^\dagger)^p A_0^p = (A_0^\dagger)^{p-1}(A_0^\dagger A_0)A_0^{p-1} = (A_0^\dagger)^{p-1} \begin{bmatrix} I & \\ & 0 \end{bmatrix} A_0 A_0^{p-2}.$$

Here, $[\begin{smallmatrix} I & 0 \end{smallmatrix}]A_0$ simply eliminates the final row of $A_0$. Note from (41) that when forming the product $A_0^\dagger M$ for some matrix $M$, only the second-to-last row of $A_0^\dagger$ depends on the final row of $M$. Thus, if we consider $A_0^\dagger[\begin{smallmatrix} I & 0 \end{smallmatrix}]A_0$, $A_0^\dagger$ will act as a (left) pseudoinverse on all rows but the last one. Repeating a similar process by applying one more power of $A_0^\dagger$ within the product $(A_0^\dagger)^p A_0^p$ yields

$$(A_0^\dagger)^p A_0^p = (A_0^\dagger)^{p-2}\left(A_0^\dagger \begin{bmatrix} I & \\ & 0 \end{bmatrix} A_0\right) A_0^{p-2} = (A_0^\dagger)^{p-2} \begin{bmatrix} I & & \\ \mathbf{e}_0 & e_1 & \\ \mathbf{0} & 0 & 0 \end{bmatrix} A_0 A_0^{p-3},$$

where $\mathbf{e}_0$ and $e_1$ are an error vector and scalar. Now, only the second-to-last and third-to-last rows of $A_0^\dagger$ depend on the final two rows of $A_0$. Continuing this process to the power of $p$, $(A_0^\dagger)^p A_0^p$ is given by an identity in the upper left block, zeros in the upper right, and $p$ rows of error. To that end, define

$$A_l^\dagger A_0^p := \left(\begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix} (A_0^\dagger)^p\right) A_0^p = \begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix}. \tag{43}$$

Note that we have defined $A_l^\dagger$ by forming $(A_0^\dagger)^p$ and eliminating the last $p$ rows.

In an analogous process, define $A_r^\dagger$ as $(A_0^\dagger)^p$ with the *first p columns* set to zero. Following steps similar to those above, we arrive at

$$A_0^p A_r^\dagger := A_0^p \left( (A_0^\dagger)^p \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix} \right) = \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix}. \tag{44}$$

Now, recalling that $A_0^p$ is zero in the first $p$ rows and last $p$ columns, (43) and (44) yield

$$A_l^\dagger A_0^p = A_l^\dagger \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix} A_0^p = \left( \begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix} (A_0^\dagger)^p \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix} \right) A_0^p = \begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix},$$
$$A_0^p A_r^\dagger = A_0^p \begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix} A_r^\dagger = A_0^p \left( \begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix} (A_0^\dagger)^p \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix} \right) = \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix}.$$

Defining

$$(A_0^p)^\dagger := \begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix} (A_0^\dagger)^p \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix} = \begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix} \mathcal{T}_0^p \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix},$$

it immediately follows that $(A_0^p)^\dagger$ satisfies the four properties of a pseudoinverse. ∎

Now, we introduce three lemmas on Toeplitz matrices related to the pseudoinverse derived in Lemma 21. These lemmas derive the appropriate Toeplitz generating functions and then provide a framework to bound the smallest nonzero singular value from above.

LEMMA 22. *Consider a matrix of the form*

$$\mathcal{T} = \begin{bmatrix} -a & b & & \\ & -a & \ddots & \\ & & \ddots & b \\ & & & -a \end{bmatrix}, \tag{45}$$

*where $\mathcal{T}$ is $n \times n$ and $a$ and $b$ some invertible coefficients or operators. Then, $\mathcal{T}^p$ is (block) Toeplitz, for $p \in \mathbb{N}$, $p < \lfloor n/2 \rfloor$, with Fourier generating function given by*

$$F_*(x) = (-a + be^{\mathrm{i}x})^p. \tag{46}$$

*Proof.* Because $\mathcal{T}$ is upper triangular, $\mathcal{T}^p$ is upper triangular for $p \geq 0$; furthermore, the stencil expands one superdiagonal with each matrix multiplication, so $\mathcal{T}^p$ has exactly $p+1$ nonzero diagonals. Then, the defining Toeplitz coefficients of $\mathcal{T}^p$ are given by the $p+1$ nonzero entries in the first row of $\mathcal{T}^p$, or, equivalently, the $p+1$ nonzero elements of $\mathbf{e}_0\mathcal{T}^p$, where $\mathbf{e}_0$ is the first canonical (block) basis vector, $(I, \mathbf{0}, \ldots, \mathbf{0})$.

Now, consider a linear algebra framework to represent polynomials, where columns of $\mathcal{T}$ represent powers of $x$. Then, given some coefficient vector $\mathbf{v}$, $\mathbf{v}\mathcal{T} = \mathbf{w}$, where $\mathbf{w}$ represents some polynomial $p(x) \sim \mathbf{w}$, and $\mathbf{w}_i$ corresponds to the polynomial coefficient of $x^i$. Then, for example, $\mathbf{e}_0\mathcal{T} \sim -a + bx \mapsto -a\mathbf{e}_0 + b\mathbf{e}_1$. Continuing,

$$\begin{aligned} \mathbf{e}_0\mathcal{T}^2 &= (-a\mathbf{e}_0 + b\mathbf{e}_1)\mathcal{T} \\ &\sim -a(-a + bx) + b(-ax + bx^2) \\ &= (-a + bx)^2 \end{aligned}$$

$$\mapsto a^2\mathbf{e}_0 - (ab+ba)\mathbf{e}_1 + b^2\mathbf{e}_2.$$

Here, the $\ell$th polynomial coefficient in $(-a+bx)^2$, $x^\ell$, corresponds to the $\ell$th basis vector, $\mathbf{e}_\ell$, for $\ell = 0,1,2$. By an inductive argument, this process continues, with the $\ell$th element in the first row of $\mathcal{T}^p$ being given by the $\ell$th polynomial coefficient of $(-a+bx)^p$. Recalling that the Fourier generating function is given by $F_*(x) = \sum_\ell \alpha_\ell e^{\mathrm{i}\ell x}$, where $\alpha_\ell$ is the $\ell$th Toeplitz coefficient, we can simply replace $x$ with $e^{\mathrm{i}x}$ to get

$$F_*(x) = (-a+be^{\mathrm{i}x})^p.$$

∎

LEMMA 23. *Consider an $n \times n$ matrix $\mathcal{T}$ as in* (45)*, and define*

$$\widehat{\mathcal{T}}_p := \begin{bmatrix} I_{(n-p)\times(n-p)} & \\ & \mathbf{0}_{p\times p} \end{bmatrix} \mathcal{T}^p \tag{47}$$

*for $p < \lfloor n/2 \rfloor$. That is, $\widehat{\mathcal{T}}_p$ corresponds to the last $p$ rows eliminated from $\mathcal{T}^p$. Then, $\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^*$ is Toeplitz in the upper $(n-p)\times(n-p)$ block and zero elsewhere, with real-valued Fourier generating function for the nonzero Toeplitz block given by*

$$F_p(x) = (-a+be^{\mathrm{i}x})^p \left[(-a+be^{\mathrm{i}x})^p\right]^*. \tag{48}$$

*Proof.* The proof proceeds by first deriving the Toeplitz coefficients for $\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^*$ based on those of $\mathcal{T}$ and showing that these coefficients correspond to the generating function $F_p(x) = (-a+be^{\mathrm{i}x})^p \left[(-a+be^{\mathrm{i}x})^p\right]^*$.

From Lemma 22, $\mathcal{T}^p$ is Toeplitz and upper triangular with $p$ nonzero superdiagonals. Eliminating the final $p$ rows of $\mathcal{T}^p$, it is straightforward to confirm that $\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^*$ is self-adjoint, Toeplitz in the upper left $(n-p)\times(n-p)$ block, and zero in the final $p$ rows and columns. By self-adjointness, the generating coefficients of $\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^*$ can be found by considering the $p+1$ nonzero entries in the first row of $\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^*$ (and their adjoints will be coefficients for the first $p+1$ rows). Let $\mathbf{e}_\ell$ denote the $\ell$th canonical basis vector and $\{\hat{\alpha}_\ell\}$ be the Fourier generating coefficients for $\mathcal{T}_p$. Then, for $\ell = 0,\dots,p$, the Fourier generating coefficients for the nonzero Toeplitz block in $\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^*$ are given by

$$\begin{aligned}\hat{\alpha}_\ell &= \left[\mathbf{e}_0\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^*\right]_\ell = \sum_{j=0}^{p}\left[\widehat{\mathcal{T}}_p\right]_{0,j}\left[\widehat{\mathcal{T}}_p^*\right]_{j,\ell} = \sum_{j=0}^{p}\left[\widehat{\mathcal{T}}_p\right]_{0,j}\left[\overline{\widehat{\mathcal{T}}_p}\right]_{\ell,j} \\ &= \sum_{j=0}^{p-\ell}\left[\widehat{\mathcal{T}}_p\right]_{0,j+\ell}\left[\overline{\widehat{\mathcal{T}}_p}\right]_{0,j} = \sum_{j=0}^{p-\ell}\alpha_{j+\ell}\alpha_j^*,\end{aligned} \tag{49}$$

where $\overline{\widehat{\mathcal{T}}_p}$ denotes the adjoint of operator entries, either the conjugate in the case of a scalar matrix, or block adjoint in the case of a block matrix. The second-to-last equality follows from the fact that in the $\ell$th column of $\widehat{\mathcal{T}}_p$, the first $\ell-1$ rows are zero.

Recall that the Fourier generating function for $\mathcal{T}_p$ is given by $F_*(x) = -a + be^{\mathrm{i}x}$ (Lemma 22), with coefficients $\{\alpha_\ell\}$. Then consider the block-Toeplitz matrix associated with generating function

$$F_p(x) = (-a+be^{\mathrm{i}x})^p \left[(-a+be^{\mathrm{i}x})^p\right]^*$$

$$= \left(\sum_{j=0}^{p} \alpha_j e^{\mathrm{i}x}\right)\left(\sum_{j=0}^{p} \alpha_j^* e^{-\mathrm{i}x}\right)$$
$$= \left(\alpha_0 + \alpha_1 e^{\mathrm{i}x} + \alpha_2 e^{2\mathrm{i}x} + \cdots\right)\left(\alpha_0^* + \alpha_1^* e^{-\mathrm{i}x} + \alpha_2^* e^{-2\mathrm{i}x} + \cdots\right).$$

The corresponding coefficients can be obtained by gathering terms of $e^{\mathrm{i}jx}$, $j = -p, -(p-1), \ldots, 0, \ldots, p$. This leads to coefficients

$$\hat{\alpha}_\ell = \sum_{\substack{j_0 - j_1 = \ell, \\ j_0, j_1 \leq p}} \alpha_{j_0} \alpha_{j_1}^* = \sum_{j=0}^{p-\ell} \alpha_{j+\ell} \alpha_j^*.$$

Indeed, these are exactly the Toeplitz coefficients obtained by directly computing $\widehat{\mathcal{T}}_p \widehat{\mathcal{T}}_p^*$ in (49), which completes the proof. ∎

*Remark* 24 (Generating coefficients). If $a$ and $b$ in (45) commute, the Binomial Theorem gives a closed form for Toeplitz generating coefficients $\{\alpha_\ell\}$ and $\{\hat{\alpha}_\ell\}$ by expanding $(a + be^{\mathrm{i}x})^p$. If $a$ and $b$ do not commute, there is a generalization of the Binomial Theorem that takes the form

$$(a + be^{\mathrm{i}x})^p = \sum_{\ell=0}^{p} \binom{p}{\ell} \left[(a + d_b)^\ell \mathbf{1}\right] (be^{\mathrm{i}x})^{p-\ell},$$

where $\mathbf{1}$ denotes the identity on the underlying associative algebra, and $d_b$ is a derivation defined by

$$d_b(z) = be^{\mathrm{i}x} z - zbe^{\mathrm{i}x}$$

for linear transformation $z$ [41, 42].

LEMMA 25. *Define $\widehat{\mathcal{T}}_p$ as in Lemma* 23 (47)*, and define $\widehat{\mathcal{A}}_p$ similarly, in the form of the pseudoinverse from Lemma* 21*,*

$$\widehat{\mathcal{A}}_p = \begin{bmatrix} I_{(n-p)\times(n-p)} & \\ & \mathbf{0}_{p\times p} \end{bmatrix} \mathcal{T}^p \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I_{(n-p)\times(n-p)} \end{bmatrix},$$

*that is, by setting the first $p$ columns and last $p$ rows of $\mathcal{T}^p$ equal to zero. Then,*

$$\sigma_{\min}(\widehat{\mathcal{A}}_p) \leq \sigma_{\min}(\widehat{\mathcal{T}}_p),$$

*where $\sigma_{\min}$ denotes the minimum nonzero singular value.*

*Proof.* Recall that $\widehat{\mathcal{T}}_p$ is upper triangular and zero in the last $p$ rows. Then, consider expressing $\widehat{\mathcal{T}}_p \widehat{\mathcal{T}}_p^*$ and $\widehat{\mathcal{A}}_p \widehat{\mathcal{A}}_p^*$ in block form, with $\varepsilon$, $M_0$, and $M_1$ chosen to denote the nonzero blocks in $\widehat{\mathcal{T}}_p$:

$$\widehat{\mathcal{T}}_p \widehat{\mathcal{T}}_p^* = \begin{bmatrix} \varepsilon & M_0 \\ \mathbf{0}_{(n-2p)\times p} & M_1 \\ \mathbf{0}_{p\times p} & \mathbf{0}_{p\times(n-p)} \end{bmatrix} \begin{bmatrix} \varepsilon^* & \mathbf{0}_{p\times(n-2p)} & \mathbf{0}_{p\times p} \\ M_0^* & M_1^* & \mathbf{0}_{(n-p)\times p} \end{bmatrix}$$
$$= \begin{bmatrix} \varepsilon\varepsilon^* + M_0 M_0^* & M_0 M_1^* & \mathbf{0}_{p\times p} \\ M_1 M_0^* & M_1 M_1^* & \mathbf{0}_{(n-2p)\times p} \\ \mathbf{0}_{p\times p} & \mathbf{0}_{p\times(n-2p)} & \mathbf{0}_{p\times p} \end{bmatrix},$$

$$\widehat{\mathcal{A}}_p\widehat{\mathcal{A}}_p^* = \begin{bmatrix} M_0M_0^* & M_0M_1^* & \mathbf{0}_{p\times p} \\ M_1M_0^* & M_1M_1^* & \mathbf{0}_{(n-2p)\times p} \\ \mathbf{0}_{p\times p} & \mathbf{0}_{p\times(n-2p)} & \mathbf{0}_{p\times p} \end{bmatrix}.$$

Because we are interested in the minimum *nonzero* singular value of $\widehat{\mathcal{T}}_p$, consider the nonzero block in $\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^* - \widehat{\mathcal{A}}_p\widehat{\mathcal{A}}_p^*$, given by

(50)
$$\begin{bmatrix} \varepsilon\varepsilon^* + M_0M_0^* & M_0M_1^* \\ M_1M_0^* & M_1M_1^* \end{bmatrix} - \begin{bmatrix} M_0M_0^* & M_0M_1^* \\ M_1M_0^* & M_1M_1^* \end{bmatrix} = \begin{bmatrix} \varepsilon\varepsilon^* & \mathbf{0}_{p\times(n-2p)} \\ \mathbf{0}_{(n-2p)\times p} & \mathbf{0}_{(n-p)\times(n-p)} \end{bmatrix} \geq 0,$$

in a positive semidefinite sense.

The proof then follows from a generalization of the monotonicity theorem or Weyl's inequality [2, Theorem 10.4.11]. In particular, let $A$ and $B$ be Hermitian matrices. Then,

$$\lambda_{\min}(A) + \lambda_{\min}(B) \leq \lambda_{\min}(A+B) \leq \lambda_{\min}(A) + \lambda_{max}(B). \tag{51}$$

Applying (51) to (50) yields

$$\begin{aligned} \lambda_{\min}\left(\begin{bmatrix} M_0M_0^* & M_0M_1^* \\ M_1M_0^* & M_1M_1^* \end{bmatrix}\right) &\leq \lambda_{\min}\left(\begin{bmatrix} \varepsilon\varepsilon^* + M_0M_0^* & M_0M_1^* \\ M_1M_0^* & M_1M_1^* \end{bmatrix}\right) \\ &\quad - \lambda_{\min}\left(\begin{bmatrix} \varepsilon\varepsilon^* & \mathbf{0}_{p\times(n-2p)} \\ \mathbf{0}_{(n-2p)\times p} & \mathbf{0}_{(n-p)\times(n-p)} \end{bmatrix}\right) \\ &\leq \lambda_{\min}\left(\begin{bmatrix} \varepsilon\varepsilon^* + M_0M_0^* & M_0M_1^* \\ M_1M_0^* & M_1M_1^* \end{bmatrix}\right). \end{aligned}$$

To that end, the minimum nonzero eigenvalue of $\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^*$ provides an upper bound on the minimum nonzero eigenvalue of $\widehat{\mathcal{A}}_p\widehat{\mathcal{A}}_p^*$, and the result follows because the singular values of a matrix $M$ are given by the square root of the eigenvalues of $MM^*$. ∎

We now have all the necessary tools to prove necessary conditions for convergence of MGRiT and Parareal.

THEOREM 26 (Necessary conditions). *Let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, with coarsening factor $k$, and $N_c$ coarse-grid time points. Assume that $(\Psi - \Phi^k)$ is invertible and that $\Phi$ satisfies an F-TAP$_1$ with respect to $\Psi$, with minimum constant $\varphi_{F,1}$. Then,*

$$\left\|I - A_\Delta B_\Delta^{-1}\right\| \geq \frac{\varphi_{F,1}}{1 + O(1/\sqrt{N_c})}.$$

*If we further assume that $\Phi$ and $\Psi$ commute, that is, $\Phi\Psi = \Psi\Phi$, and that $\Phi$ satisfies an F-TAP$_p$ with respect to $\Psi$, with minimum constant $\varphi_{F,p}$, then*

$$\left\|(I - A_\Delta B_\Delta^{-1})^p\right\| \geq \frac{\varphi_{F,p}}{1 + O(1/\sqrt{N_c})}.$$

*Similarly, assume that $(\Psi - \Phi^k)$ is invertible and that $\Phi$ satisfies an FCF-TAP$_1$ with respect to $\Psi$, with constant minimum $\varphi_{FCF,1}$. Then*

$$\left\|(I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right\| \geq \frac{\varphi_{FCF,1}}{1 + O(1/\sqrt{N_c})}.$$

*If we further assume that $\Phi$ and $\Psi$ commute, that is, $\Phi\Psi = \Psi\Phi$, and that $\Phi$ satisfies an FCF-TAP$_p$ with respect to $\Psi$, with minimum constant $\varphi_{FCF,p}$, then*

$$\left\| \left( (I - A_\Delta B_\Delta^{-1}) A_{cf} A_{ff}^{-1} A_{fc} \right)^p \right\| \geq \frac{\varphi_{FCF,p}}{1 + O(1/\sqrt{N_c})}.$$

*Proof.* To bound $(I - A_\Delta B_\Delta^{-1})^p$ and $((I - A_\Delta B_\Delta^{-1}) A_{cf} A_{ff}^{-1} A_{fc})^p$ in norm, we note that the $\ell^2$-norm of an operator is given by its largest singular value, which is equal to one over the smallest nonzero singular value of the operator's inverse or pseudoinverse.

Now, notice that these operators exactly take the form of $A_0$ (38) and $A_1$ (39). For F-relaxation, $f = \Psi$, $g = (\Psi - \Phi^k)$, and $h = I$, and for FCF relaxation, $f = \Psi$, $g = (\Psi - \Phi^k)$, and $h = \Phi^k$. Lemma 21 gives an exact pseudoinverse for powers of such operators, and Lemma 25 proves that the minimum nonzero singular value of this pseudoinverse is bounded above by the minimum singular value of the Toeplitz operator $\widehat{\mathcal{T}}_p$ (47), with $a = h^{-1} f g^{-1}$ and $b = h^{-1} g^{-1}$ (45). Equivalently, we can consider the minimum nonzero eigenvalue of the corresponding normal equations. Appealing to Lemma 23, the Fourier generating functions for the nonzero Toeplitz block in these operators are given by

(52)
$$F_F(x,p) = \left( e^{\mathrm{i}x} \Psi (\Psi - \Phi^k)^{-1} - (\Psi - \Phi^k)^{-1} \right)^p \left( e^{\mathrm{i}x} \Psi (\Psi - \Phi^k)^{-1} - (\Psi - \Phi^k)^{-1} \right)^{*^p},$$

(53)
$$F_{FCF}(x,p) = \left( e^{\mathrm{i}x} \Phi^{-k} \Psi (\Psi - \Phi^k)^{-1} - \Phi^{-k} (\Psi - \Phi^k) \right)^p \left( e^{\mathrm{i}x} \Phi^{-k} \Psi (\Psi - \Phi^k)^{-1} - \Phi^{-k} (\Psi - \Phi^k)^{-1} \right)^{*^p},$$

where $F_F(x,p)$ will lead to a bound on $(I - A_\Delta B_\Delta^{-1})^p$ (F-relaxation) and $F_{FCF}(x,p)$ will lead to a bound on $((I - A_\Delta B_\Delta^{-1}) A_{cf} A_{ff}^{-1} A_{fc})^p$ (FCF-relaxation).

By Theorem 17 we seek the infimum over $x$ of the minimum nonzero eigenvalue of $F_F(x,p)$ and $F_{FCF}(x,p)$. Let $\lambda_k(A)$ and $\sigma_k(A)$ denote the $k$th eigenvalue and singular value of some operator $A$ and consider the case of $p = 1$ for $F_F(x,1)$:

$$\begin{aligned}
\min_{\substack{x \in [0,2\pi], \\ k}} \lambda_k(F_F(x,1)) &= \min_{\substack{x \in [0,2\pi], \\ k}} \sigma_k \left( (e^{\mathrm{i}x} \Psi - I)(\Psi - \Phi^k)^{-1} \right)^2 \\
&= \min_{\substack{x \in [0,2\pi], \\ \mathbf{v} \neq \mathbf{0}}} \frac{\left\| (e^{\mathrm{i}x} \Psi - I)(\Psi - \Phi^k)^{-1} \mathbf{v} \right\|^2}{\|\mathbf{v}\|^2} \\
&= \min_{\substack{x \in [0,2\pi], \\ \mathbf{v} \neq \mathbf{0}}} \frac{\left\| (e^{\mathrm{i}x} \Psi - I) \mathbf{v} \right\|^2}{\|(\Psi - \Phi^k)\mathbf{v}\|^2}.
\end{aligned}$$

Appealing to Theorem 17,

$$\begin{aligned}
\left\| I - A_\Delta B_\Delta^{-1} \right\| &\geq \frac{1}{\sqrt{\min_{\substack{x \in [0,2\pi], \\ \mathbf{v} \neq \mathbf{0}}} \frac{\left\| (e^{\mathrm{i}x} \Psi - I)\mathbf{v} \right\|^2}{\|(\Psi - \Phi^k)\mathbf{v}\|^2} + O(1/N_c)}} \geq \frac{1}{\min_{\substack{x \in [0,2\pi], \\ \mathbf{v} \neq \mathbf{0}}} \frac{\left\| (e^{\mathrm{i}x} \Psi - I)\mathbf{v} \right\|}{\|(\Psi - \Phi^k)\mathbf{v}\|} + O(1/\sqrt{N_c})} \\
&= \max_{\mathbf{v} \neq \mathbf{0}} \frac{\|(\Psi - \Phi^k)\mathbf{v}\|}{\min_{x \in [0,2\pi]} \|(I - e^{\mathrm{i}x} \Psi)\mathbf{v}\| + O(1/\sqrt{N_c})}. \qquad (54)
\end{aligned}$$

By assumption of an F-TAP$_1$ with constant $\varphi_{F,1}$[8]

$$\|(\Psi-\Phi^k)\mathbf{v}\| \leq \varphi_{F,1}\left[\min_{x\in[0,2\pi]}\left\|(I-e^{\mathrm{i}x}\Psi)\mathbf{v}\right\|\right] = \frac{\varphi_{F,1}}{1+O(1/\sqrt{N_c})}\left[\min_{x\in[0,2\pi]}\left\|(I-e^{\mathrm{i}x}\Psi)\mathbf{v}\right\|+O(1/\sqrt{N_c})\right]. \tag{55}$$

Assuming that $\varphi_{F,1}$ is a tight bound, there exists some $\mathbf{v}$ such that (55) holds with equality. Then, plugging (55) into (54),

$$\left\|I-A_\Delta B_\Delta^{-1}\right\| \geq \frac{\varphi_{F,1}}{1+O(1/\sqrt{N_c})}.$$

A similar derivation based on the assumption of an FCF-TAP$_1$ with constant $\varphi_{FCF,1}$ follows to bound

$$\left\|(I-A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right\| \geq \frac{\varphi_{FCF,1}}{1+O(1/\sqrt{N_c})}.$$

Finally, if $\Phi$ and $\Psi$ commute, then, for example,

$$\left(e^{\mathrm{i}x}\Psi(\Psi-\Phi^k)^{-1}-(\Psi-\Phi^k)^{-1}\right)^p = (e^{\mathrm{i}x}\Psi-I)^p(\Psi-\Phi^k)^{-p}.$$

Under the assumption of an F-TAP$_p$ and FCF-TAP$_p$ with constants $\varphi_{F,p}$ and $\varphi_{FCF,P}$, respectively, for $p\geq 1$, an analogous derivation as used for $p=1$ yields the bounds

$$\left\|(I-A_\Delta B_\Delta^{-1})^p\right\| \geq \frac{\varphi_{F,p}}{1+O(1/\sqrt{N_c})},$$
$$\left\|\left((I-A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right)^p\right\| \geq \frac{\varphi_{FCF,p}}{1+O(1/\sqrt{N_c})},$$

and similarly for the case of FCF-relaxation. ∎

Coupling Theorems 19 and 26 with the operator form of residual propagation for $p$ iterations (36) and the equivalence of $\|\mathcal{R}\| = \|\mathcal{E}\|_{A^*A}$ (18) completes the proof of Theorem 2. Recall from (33) and (34) that if $\Phi$ and $\Psi$ commute, then $I-B_\Delta^{-1}A_\Delta = I-A_\Delta B_\Delta^{-1}$, which proves Corollary 3.

**4.2. Error-propagation and $I-B_\Delta^{-1}A_\Delta$.** This section provides proofs of Theorem 10. The framework developed in previous sections allows for a streamlined presentation. First, sufficient conditions for convergence of error in the $\ell^2$-norm (based on $B_\Delta^{-1}$ as a left approximate inverse) are presented.

THEOREM 27 (Sufficient conditions ($\ell^2$-error)). *Let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, with coarsening factor $k$, and $N_c$ coarse-grid time points. Assume that $\Phi$ satisfies an F-ITAP with respect to $\Psi$, with constant $\tilde{\varphi}_F$. Then,*

$$\left\|I-B_\Delta^{-1}A_\Delta\right\| \leq \tilde{\varphi}_F\left(1+\|\Psi^{N_c}\|\right).$$

*Similarly, assume that $\Phi$ satisfies an FCF-ITAP with respect to $\Psi$, with constant $\tilde{\varphi}_{FCF}$. Then,*

$$\left\|(I-B_\Delta^{-1}A_\Delta)A_{cf}A_{ff}^{-1}A_{fc}\right\| \leq \tilde{\varphi}_{FCF}\left(1+\|\Psi^{N_c}\|\right).$$

[8]Note that in (55), the leading constant in the $O(1/\sqrt{N_c})$ terms changes; however, the change in constant is marginal for any moderate $N_c \gg O(1)$.

*Proof.* The proof is analogous to that of Theorem 19. Notice from (34) that $I - B_\Delta^{-1}A_\Delta$ is a block-Toeplitz matrix with generating coefficients $\alpha_i = \Psi^{-(1+i)}(\Psi - \Phi^k)$ for $i = -1, \dots, -N_c$ and $\alpha_i = \mathbf{0}$ for $i \geq 0$. Following Theorem 19, the generating function is given by

$$F_F(x) = e^{\mathrm{i}x}(I - e^{\mathrm{i}Nx}\Psi^N)(I - e^{\mathrm{i}x}\Psi)^{-1}(\Psi - \Phi^k).$$

Recall that, under the assumption of an F-ITAP, $\tilde{\varphi}_F\|\mathbf{v}\| \leq [\max_{x\in[0,2\pi]} \|(I - e^{\mathrm{i}x}\Psi)^{-1}(\Psi - \Phi^k)\mathbf{v}\|]$ for all $\mathbf{v}$, with equality for some $\mathbf{v}$. Theorem 18 then yields

$$\begin{aligned}
\|I - B_\Delta^{-1}A_\Delta\| &\leq \max_{x\in[0,2\pi]} \sigma_{max}(F_F(x)) \\
&= \max_{\substack{x\in[0,2\pi],\\ \mathbf{v}\neq\mathbf{0}}} \frac{\left\|(I - e^{\mathrm{i}Nx}\Psi^N)(I - e^{\mathrm{i}x}\Psi)^{-1}(\Psi - \Phi^k)\mathbf{v}\right\|}{\|\mathbf{v}\|} \\
&\leq \max_{\substack{x\in[0,2\pi],\\ \mathbf{v}\neq\mathbf{0}}} \frac{\left(1 + \|\Psi^{N_c}\|\right)\left\|(I - e^{\mathrm{i}x}\Psi)^{-1}(\Psi - \Phi^k)\mathbf{v}\right\|}{\|\mathbf{v}\|} \\
&= \tilde{\varphi}_F\left(1 + \|\Psi^{N_c}\|\right).
\end{aligned}$$

A similar proof follows for the case of FCF-relaxation, where the generator function, $F_{FCF}$, has coefficients $\alpha_i = \Psi^{-(1+i)}(\Psi - \Phi^k)\Phi^k$ for $i = -1, \dots, -N_c$ and $\alpha_i = \mathbf{0}$ for $i \geq 0$. By assumption of an FCF-ITAP with constant $\tilde{\varphi}_{FCF}$, the result follows. ∎

Now, we present a result similar to Theorem 26 which provides necessary conditions for convergence of error in the $\ell^2$-norm.

THEOREM 28 (Necessary conditions ($\ell^2$-error)). *Let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, with coarsening factor $k$, and $N_c$ coarse-grid time points. Assume that $(\Psi - \Phi^k)$ is invertible and that $\Phi$ satisfies an F-ITAP with respect to $\Psi$, with constant $\tilde{\varphi}_F$. Then,*

$$\left\|I - B_\Delta^{-1}A_\Delta\right\| \geq \frac{\tilde{\varphi}_F}{1 + O(1/\sqrt{N_c})}.$$

*Similarly, assume that $(\Psi - \Phi^k)$ is invertible and that $\Phi$ satisfies an FCF-ITAP with respect to $\Psi$, with constant $\tilde{\varphi}_{FCF}$. Then*

$$\left\|(I - B_\Delta^{-1}A_\Delta)A_{cf}A_{ff}^{-1}A_{fc}\right\| \geq \frac{\tilde{\varphi}_{FCF}}{1 + O(1/\sqrt{N_c})}.$$

*Proof.* This proof is analogous to that of Theorem 26, this time with $g = I$, $f = \Psi$, and $h = \Psi - \Phi^k$ for F-relaxation, and $h = (\Psi - \Phi^k)\Phi^k$ for FCF-relaxation. Similar to (52) and (53), the Fourier generating functions of interest are then given by

$$\begin{aligned}
F_F(x) &= \left(e^{\mathrm{i}x}(\Psi - \Phi^k)^{-1}\Psi - (\Psi - \Phi^k)^{-1}\right)\left(e^{\mathrm{i}x}(\Psi - \Phi^k)^{-1}\Psi - (\Psi - \Phi^k)^{-1}\right)^*, \\
F_{FCF}(x) &= \left(e^{\mathrm{i}x}\Psi(\Psi - \Phi^k)^{-1}\Phi^{-k} - \Phi^{-k}(\Psi - \Phi^k)\right)\left(e^{\mathrm{i}x}\Psi(\Psi - \Phi^k)^{-1}\Phi^{-k} - \Phi^{-k}(\Psi - \Phi^k)^{-1}\right)^*.
\end{aligned}$$

Following the algebraic steps in the proof of Theorem 26 completes the proof. ∎

**5. The diagonalizable case.** So far, results have been derived in terms of the time-stepping operators $\Phi$ and $\Psi$. In this section, we assume that $\Phi$ and $\Psi$ commute and are diagonalizable. In general this corresponds to the spatial operator being diagonalizable, which holds for many parabolic-type problems, among others. The purposes of this section are as follows:

1. Derive exact bounds on convergence for diagonalizable operators.
2. Show that theory developed in this paper is, in some sense, a strengthening and generalization of previous results in [5].

If $\Phi$ and $\Psi$ commute and are diagonalizable, this means that they are diagonalizable with the same eigenvectors, or simultaneously diagonalizable. Under the assumption of simultaneous diagonalizability, certain bounds can be derived on functions of $\Phi$ and $\Psi$. Let $\Phi = U\Lambda U^{-1}$, where $\Lambda_{ii} = \lambda_i$ for $i = 1, \dots, n$ is a diagonal matrix consisting of the eigenvalues of $\Phi$, and columns of $U$ are the corresponding eigenvectors. Similarly, let $\Psi = U\Xi U^{-1}$ for diagonal matrix $\Xi$, where $\Xi_{ii} = \mu_i$ are the eigenvalues of $\Psi$. Note that in this section subscript $i$ corresponds to eigenvalue index, as opposed to iteration number as used previously. Now let $\mathcal{A}$ be some matrix operator where each entry is a rational function of $\Phi$ and $\Psi$,

$$\mathcal{A}(\Phi,\Psi) = \begin{bmatrix} a_{00}(\Phi,\Psi) & a_{01}(\Phi,\Psi) & \dots \\ a_{10}(\Phi,\Psi) & a_{11}(\Phi,\Psi) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} U & & \\ & U & \\ & & \ddots \end{bmatrix} \begin{bmatrix} a_{00}(\Lambda,\Xi) & a_{01}(\Lambda,\Xi) & \dots \\ a_{10}(\Lambda,\Xi) & a_{11}(\Lambda,\Xi) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} U^{-1} & & \\ & U^{-1} & \\ & & \ddots \end{bmatrix}. \tag{56}$$

Denote $\mathbf{U}$ as the block diagonal matrix of eigenvectors, $U$, in (56), define $\mathcal{P}$ as the orthogonal permutation matrix such that $\mathcal{P}\mathcal{A}(\Lambda,\Xi)\mathcal{P}^T$ is block diagonal, with blocks given by $\mathcal{A}(\lambda_i,\mu_i)$, and let $\widetilde{U} = \mathbf{U}\mathcal{P}^T$. Then,

$$\begin{aligned} \|\mathcal{A}(\Phi,\Psi)\|_{(\widetilde{U}\widetilde{U}^*)^{-1}} &= \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\left\|\widetilde{U}^{-1}\mathbf{U}\mathcal{P}^T\mathcal{P}\mathcal{A}(\Lambda,\Xi)\mathcal{P}^T\mathcal{P}\mathbf{U}^{-1}\mathbf{x}\right\|}{\|\widetilde{U}^{-1}\mathbf{x}\|^2} \\ &= \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\left\|\mathcal{P}\mathcal{A}(\Lambda,\Xi)\mathcal{P}^T\mathbf{x}\right\|}{\|\mathbf{x}\|^2} = \sup_i \|\mathcal{A}(\lambda_i,\mu_i)\|. \end{aligned} \tag{57}$$

Thus, the $(\widetilde{U}\widetilde{U}^*)^{-1}$-norm of $\mathcal{A}(\Phi,\Psi)$ can be computed by maximizing the norm of $\mathcal{A}$ over eigenvalue indices of $\Phi$ and $\Psi$. In the case that $\Phi$ and $\Psi$ are normal matrices, $U$ is unitary and the $(\widetilde{U}\widetilde{U}^*)^{-1}$-norm reduces to the standard Euclidean 2-norm. More generally, we have the relation

$$\frac{1}{\kappa(U)}\left(\sup_i \|\mathcal{A}(\lambda_i,\mu_i)^k\|\right) \leq \|\mathcal{A}(\Phi,\Psi)^k\| \leq \kappa(U)\left(\sup_i \|\mathcal{A}(\lambda_i,\mu_i)^k\|\right), \tag{58}$$

where $\kappa(U)$ denotes the matrix condition number of $U$.[9]

Here, we are interested in $\mathcal{A}$ corresponding to the error- and residual-propagation operators of MGRiT, $\mathcal{R}$ and $\mathcal{E}$ (see (20), (22), (24), and (26)). For notation, let, for example, $\mathcal{R}_F(\lambda_i,\mu_i)$ denote residual propagation for F-relaxation (20) operating

[9] A similar modified norm also occurs in the case of integrating in time with a mass matrix [5].

on eigenvalues $\lambda_i$ and $\mu_i$ as opposed to operators $\Phi$ and $\Psi$. Convergence of MGRiT requires $\|\mathcal{R}^p\|, \|\mathcal{E}^p\| \to 0$ with iteration $p$; to that end, bounding $\sup_i \|\mathcal{R}(\lambda_i, \mu_i)^p\| < 1$ for all $i$ provides necessary and sufficient conditions for $\|\mathcal{R}(\Phi, \Psi)^p\| \to 0$ with $p$, and similarly for $\mathcal{E}(\lambda_i, \mu_i)$.

**5.1. Necessary conditions.** First, let us extend the necessary conditions for $p$ iterations (Theorems 12, 13, and 26) to the diagonalizable case. For notation, let, for example, $\left[I - A_\Delta B_\Delta^{-1}\right]_i$ denote the $N_x \times N_x$ matrix of $I - A_\Delta B_\Delta^{-1}$ evaluated at the $i$th eigenmode of $\Phi$ and $\Psi$, where $\Phi$ and $\Psi$ are $N_x \times N_x$. Then, by assumption of simultaneous diagonalizability and the fact that $I - A_\Delta B_\Delta^{-1} = I - B_\Delta^{-1} A_\Delta$ when $\Phi$ and $\Psi$ commute,

(59)
$$\begin{aligned}
\|(I - B_\Delta^{-1} A_\Delta)^p\|_{(UU^*)^{-1}} &= \|(I - A_\Delta B_\Delta^{-1})^p\|_{(UU^*)^{-1}} \\
&= \sup_i \left\| \left[I - A_\Delta B_\Delta^{-1}\right]_i^p \right\|, \\
\left\| \left( (I - B_\Delta^{-1} A_\Delta) A_{cf} A_{ff}^{-1} A_{fc} \right)^p \right\|_{(UU^*)^{-1}} &= \left\| \left( (I - A_\Delta B_\Delta^{-1}) A_{cf} A_{ff}^{-1} A_{fc} \right)^p \right\|_{(UU^*)^{-1}} \\
&= \sup_i \left\| \left[ (I - A_\Delta B_\Delta^{-1}) A_{cf} A_{ff}^{-1} A_{fc} \right]_i^p \right\|.
\end{aligned}$$

Now, we can follow the derivation in section 4. For a pseudoinverse (Lemmas 20 and 21), we have $f = \mu_i$, $g = \mu_i - \lambda_i^k$, and, for F-relaxation, $h = 1$, while for FCF-relaxation, $h = \lambda_i^k$. Because $f$, $g$, and $h$ now commute, we can remove a leading factor of $g^{-1}h^{-1}$, and we are interested in the smallest nonzero singular value of (Lemma 21)

(60)
$$(\mathcal{A}_0^p)^\dagger = \frac{1}{\mu_i(\mu_i - \lambda_i^k)} \begin{bmatrix} I & \\ & \mathbf{0}_{p\times p} \end{bmatrix} \mathcal{T}_0^p \begin{bmatrix} \mathbf{0}_{p\times p} & \\ & I \end{bmatrix}, \quad \text{where} \quad \mathcal{T}_0 = \begin{bmatrix} -\mu_i & 1 & & \\ & -\mu_i & \ddots & \\ & & \ddots & 1 \\ & & & -\mu_i \end{bmatrix}.$$

Following the further derivations in section 4, the minimum nonzero singular value of $(\mathcal{A}_0^p)^\dagger$ is bounded above by the minimum eigenvalue of $\widehat{\mathcal{T}}_p \widehat{\mathcal{T}}_p^*$, where $a = \mu_i$ and $b = 1$ (see (45), (47), and Lemmas 22, 23, and 25). The Fourier generating function for $\widehat{\mathcal{T}}_p \widehat{\mathcal{T}}_p^*$ from Lemma 23 (48) now takes the form

$$\begin{aligned}
F_p(x) &= \left(1 + |\mu_i|^2 - (\mu_i e^{-\mathrm{i}x} + \overline{\mu}_i e^{\mathrm{i}x})\right)^p \\
&= \Big(1 + \operatorname{Re}(\mu_i)^2 + \operatorname{Im}(\mu_i)^2 - 2\operatorname{Re}(\mu_i)\cos(x) - 2\operatorname{Im}(\mu_i)\sin(x)\Big)^p.
\end{aligned}$$

To obtain the minimum of $F_p(x)$, note that

$$\begin{aligned}
F_p'(x) &= p\Big(2\operatorname{Re}(\mu_i)\sin(x) - 2\operatorname{Im}(\mu_i)\cos(x)\Big)\left(1 + |\mu_i|^2 - (\mu_i e^{-\mathrm{i}x} + \overline{\mu}_i e^{\mathrm{i}x})\right)^{p-1} \\
&= p\Big(2\operatorname{Re}(\mu_i)\sin(x) - 2\operatorname{Im}(\mu_i)\cos(x)\Big)\Big((\sin(x) - \operatorname{Im}(\mu_i))^2 + (\cos(x) - \operatorname{Re}(\mu_i))^2\Big).
\end{aligned} \tag{61}$$

The first term in (61) has real roots given by $n\pi + \arctan(\operatorname{Im}(\mu_i)/\operatorname{Re}(\mu_i))$ for $n \in \mathbb{Z}$. Any other roots of $F_p'(x)$ must satisfy

$$(\sin(x) - \operatorname{Im}(\mu_i))^2 + (\cos(x) - \operatorname{Re}(\mu_i))^2 = 0,$$

or, equivalently, $\sin(x) = \mathrm{Im}(\mu_i)$ and $\cos(x) = \mathrm{Re}(\mu_i)$. Suppose this holds. Solving for $\tilde{x} = \arcsin(\mathrm{Im}(\mu_i))$, we must also have $\cos(\tilde{x}) - \mathrm{Re}(\mu_i) = 0$, which implies $\mathrm{Re}(\mu_i) = \sqrt{1 - \mathrm{Im}(\mu_i)^2}$. However, then $|\mu_i| = 1$, which violates the assumption that $|\mu_i| < 1$. It follows that the only real roots of $F_p'(x)$ are given by $\hat{x} := n\pi + \arctan(\mathrm{Im}(\mu_i)/\,\mathrm{Re}(\mu_i))$ for $n \in \mathbb{Z}$.

Given that $F_p(x)$ is continuous and $2\pi$-periodic, the infimum of $F_p(x)$ over $x \in \mathbb{R}$ is attained at one of the roots of $F_p'(x)$. It is easily confirmed that the infimum is achieved for even $n$, $\hat{x}_{\min} := 2\hat{n}\pi + \arctan(\mathrm{Im}(\mu_i)/\,\mathrm{Re}(\mu_i))$, where $\hat{n} \in \mathbb{Z}$, which yields

$$\begin{aligned}
\inf_{x\in\mathbb{R}} F_p(x) &= \Big(1 + \mathrm{Re}(\mu_i)^2 + \mathrm{Im}(\mu_i)^2 - 2\,\mathrm{Re}(\mu_i)\cos(\hat{x}) - 2\,\mathrm{Im}(\mu_i)\sin(\hat{x})\Big)^p \\
&= \left(1 + \mathrm{Re}(\mu_i)^2 + \mathrm{Im}(\mu_i)^2 - 2\frac{\mathrm{Re}(\mu_i)}{\sqrt{1 + \frac{\mathrm{Im}(\mu_i)}{\mathrm{Re}(\mu_i)}}} - 2\frac{\mathrm{Im}(\mu_i)^2/\,\mathrm{Re}(\mu_i)}{\sqrt{1 + \frac{\mathrm{Im}(\mu_i)}{\mathrm{Re}(\mu_i)}}}\right)^p \\
&= \left(1 + \mathrm{Re}(\mu_i)^2 + \mathrm{Im}(\mu_i)^2 - 2\frac{\mathrm{Im}(\mu_i)^2 + \mathrm{Re}(\mu_i)^2}{\sqrt{\mathrm{Im}(\mu_i)^2 + \mathrm{Re}(\mu_i)^2}}\right)^p \\
&= \Big(1 + |\mu_i|^2 - 2|\mu_i|\Big)^p \\
&= (1 - |\mu_i|)^{2p}.
\end{aligned}$$

Noting that $F_p(x) - (1-|\mu_i|)^{2p}$ has a zero of order two at $\hat{x}_{\min}$, it follows from Theorem 17 that $\lambda_{\min}(\widehat{\mathcal{T}}_p\widehat{\mathcal{T}}_p^*) = (1 - |\mu_i|)^{2p} + O(1/N_c^2)$. Note the faster convergence in $N_c$ of eigenvalues to the infimum over $F(x)$ in the diagonalizable case, $O(1/N_c^2)$, compared with the general case in section 4, where the first-order root in $x$ led to convergence $O(1/N_c)$.

This discussion is summarized in the following theorem on necessary conditions for convergence.

THEOREM 29 (Necessary conditions: The diagonalizable case). *Let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, with coarsening factor $k$, and $N_c$ coarse-grid time points. Assume that $\Phi$ and $\Psi$ commute and are diagonalizable, with eigenvectors as columns of $U$. Then, for number of iterations $p \geq 1$*

$$\begin{aligned}
\|(I - A_\Delta B_\Delta^{-1})^p\|_{(UU^*)^{-1}} &\geq \sup_i \frac{|\mu_i - \lambda_i^k|^p}{\sqrt{(1-|\mu_i|)^{2p} + O(1/N_c^2)}} > \sup_i \frac{|\mu_i - \lambda_i^k|^p}{(1-|\mu_i|)^p + O(1/N_c)}, \\
\left\|\left((I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right)^p\right\|_{(UU^*)^{-1}} &\geq \sup_i \frac{|\lambda_i^{kp}||\mu_i - \lambda_i^k|^p}{\sqrt{(1-|\mu_i|)^{2p} + O(1/N_c^2)}} > \sup_i \frac{|\lambda_i^{kp}||\mu_i - \lambda_i^k|^p}{(1-|\mu_i|)^p + O(1/N_c)}.
\end{aligned}$$

*Proof.* The proof follows from (59), Theorem 17, the minimum derived for $F_p(x)$, and the fact that $x + y > \sqrt{x^2 + y^2}$ for $x, y > 0$. ∎

**5.2. Sufficient conditions.** Now consider sufficient conditions for convergence under the assumption that $\Phi$ and $\Psi$ commute and are diagonalizable. To do so, we consider the minimum nonzero singular value of $(I - A_\Delta B_\Delta^{-1})^\dagger$. As in section 5.1 and (60), we can pull out leading constants, form the normal equations with the remaining block, and reduce the problem to finding the minimum nonzero singular value of the

following symmetric positive semidefinite matrix:

$$\begin{bmatrix} 0 & 0 & 0 & \dots & \\ 0 & 1+|\mu_i|^2 & -\overline{\mu}_i & & \\ 0 & -\mu_i & \ddots & \ddots & \\ \vdots & & \ddots & 1+|\mu_i|^2 & -\overline{\mu}_i \\ & & & -\mu_i & 1 \end{bmatrix}. \tag{62}$$

The nonzero block is a single-entry perturbation to a symmetric tridiagonal Toeplitz matrix, for which we can place tight bounds on the minimum nonzero eigenvalue (see Appendix A, Lemma 36, and (74)). Using the bounds derived in (74) leads to the following theorem.

THEOREM 30 (Tight bounds: The diagonalizable case). *Let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, with coarsening factor $k$ and $N_c$ coarse-grid time points. Assume that $\Phi$ and $\Psi$ commute and are diagonalizable, with eigenvectors as columns of $U$. Then,*

(63)
$$\sup_i \frac{|\mu_i-\lambda_i^k|}{\sqrt{(1-|\mu_i|)^2+\frac{\pi^2|\mu_i|}{N_c^2}}} \le \|I-A_\Delta B_\Delta^{-1}\|_{(UU^*)^{-1}} \le \sup_i \frac{|\mu_i-\lambda_i^k|}{\sqrt{(1-|\mu_i|)^2+\frac{\pi^2|\mu_i|}{6N_c^2}}},$$
$$\sup_i \frac{|\lambda_i^k||\mu_i-\lambda_i^k|}{\sqrt{(1-|\mu_i|)^2+\frac{\pi^2|\mu_i|}{N_c^2}}} \le \left\|(I-A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right\|_{(UU^*)^{-1}} \le \sup_i \frac{|\lambda_i^k||\mu_i-\lambda_i^k|}{\sqrt{(1-|\mu_i|)^2+\frac{\pi^2|\mu_i|}{6N_c^2}}}.$$

*Furthermore, for $p \ge 1$,*

$$\begin{aligned} \left\|(I-A_\Delta B_\Delta^{-1})^p\right\|_{(UU^*)^{-1}} &= \sup_i \frac{|\mu_i-\lambda_i^k|^p}{\sqrt{(1-|\mu_i|)^{2p}+O(1/N_c^2)}}, \\ \left\|\left((I-A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right)^p\right\|_{(UU^*)^{-1}} &= \sup_i \frac{|\lambda_i^k|^p|\mu_i-\lambda_i^k|^p}{\sqrt{(1-|\mu_i|)^{2p}+O(1/N_c^2)}}. \end{aligned} \tag{64}$$

*Proof.* The single-iteration bounds follow immediately from (59) and Lemma 36 (74). Applying the submultiplicative norm property to (63) yields an upper bound on $p$ iterations, and Theorem 29 yields lower bounds, each of which take the form, for example, with F-relaxation,

$$\sup_i \frac{|\mu_i-\lambda_i^k|^p}{\sqrt{(1-|\mu_i|)^{2p}+O(1/N_c^2)}}.$$

This completes the proof. ∎

**5.3. Relation to the TAP.** Returning to the general theoretical framework, suppose that $\Phi$ and $\Psi$ commute and are diagonalizable, $\Phi = U\Lambda U^{-1}$, and $\Psi = U\Xi U^{-1}$. Further suppose that $\Phi$ satisfies an F-TAP$_p$ with respect to $\Psi$, with constant $\varphi_{F,p}$, *in the $(UU^*)^{-1}$-norm*. This is equivalent to saying that there exists a constant $\varphi_{F,p}$ such that for all $\mathbf{v}$,

$$\begin{aligned} &\|(\Psi-\Phi^k)^p\mathbf{v}\|_{(UU^*)^{-1}} \le \varphi_{F,p}\left[\min_{x\in[0,2\pi]}\|(I-e^{\mathrm{i}x}\Psi)^p\mathbf{v}\|_{(UU^*)^{-1}}\right] \\ \iff\quad &\|(\Xi-\Lambda^k)^pU^{-1}\mathbf{v}\| \le \varphi_{F,p}\left[\min_{x\in[0,2\pi]}\|(I-e^{\mathrm{i}x}\Xi)^pU^{-1}\mathbf{v}\|\right]. \end{aligned} \tag{65}$$

Now note that if $\Phi$ and $\Psi$ are diagonalizable, the eigenvectors form a basis, and any vector $\mathbf{v}$ can be written as a linear combination of eigenvectors of $\Phi, \Psi$, where $\mathbf{v} = \sum_{\ell=1}^{N_x} \alpha_\ell \mathbf{u}_\ell$. Then, because $U^{-1}\mathbf{u}_i = \mathbf{e}_i$, where $\mathbf{e}_i$ is the $i$th canonical basis vector, (65) reduces to

$$\sum_{\ell=0}^{N_x-1} \alpha_\ell |\mu_\ell - \lambda_\ell^k|^p \leq \varphi_{F,1} \sum_{\ell=0}^{N_x-1} (1-|\mu_\ell|)^p.$$

Note that this is only satisfied for all $\mathbf{v}$ if, for every eigenvalue index $i$,

$$|\mu_i - \lambda_i^k|^p \leq \varphi_{F,p}(1-|\mu_i|)^p \qquad \Longleftrightarrow \qquad |\mu_i - \lambda_i^k| \leq \varphi_{F,p}(1-|\mu_i|).$$

Indeed, this is exactly the F-TEAP introduced in section 2.4. A similar property holds for FCF-relaxation, which is summarized in the following proposition.

PROPOSITION 31 (Equivalent approximation properties). *The F-TEAP is the same as the F-TAP$_p$ in the $(UU^*)^{-1}$-norm, for arbitrary $p$, and the FCF-TEAP is the same as the FCF-TAP$_p$ in the $(UU^*)^{-1}$-norm, for arbitrary $p$. If $\Phi$ and $\Psi$ are normal, the two types of approximation property are identical.*

*Proof.* The proof follows from the above discussion. ∎

We are now ready to present the final result.

THEOREM 32 (Tight bounds: Multiple iterations). *Let $\Phi$ denote the fine-grid time-stepping operator and $\Psi$ denote the coarse-grid time-stepping operator, with coarsening factor $k$ and $N_c$ coarse-grid time points. Assume that $\Phi$ and $\Psi$ commute and are diagonalizable, with eigenvectors given as columns of $U$. Suppose that $\Phi$ satisfies an F-TEAP with respect to $\Psi$, with constant $\varphi_F$. Then, for $p \geq 1$,*

$$\left\|(I - A_\Delta B_\Delta^{-1})^p\right\|_{(UU^*)^{-1}}^2 = \varphi_F^{2p} - O(1/N_c^2).$$

*Similarly, suppose that $\Phi$ satisfies an FCF-TEAP with respect to $\Psi$, with constant $\varphi_{FCF}$. Then, for $p \geq 1$,*

$$\left\|\left((I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right)^p\right\|_{(UU^*)^{-1}}^2 = \varphi_{FCF}^{2p} - O(1/N_c^2).$$

*Proof.* By assumption of the T-FEAP and Theorem 30, to order $O(1/N_c^2)$,

$$\left\|(I - A_\Delta B_\Delta^{-1})^p\right\|_{(UU^*)^{-1}}^2 = \sup_i \frac{\left(|\mu_i - \lambda_i^k|^p\right)^2}{(1-|\mu_i|)^{2p} + O(1/N_c^2)} = \frac{\varphi_F^{2p}(1-|\mu_{i_{max}}|)^{2p}}{(1-|\mu_{i_{max}}|)^{2p} + O(1/N_c^2)}.$$

Note that equality holds in the second relation because there exists some $i = i_{max}$ such that $\varphi_F$ is tight. A simple Taylor or Laurent series argument about $N_c = \infty$ confirms that for $N_c > 1/(1-|\mu_i|)^p$,

$$\frac{1}{(1-|\mu_i|)^{2p} + O(1/N_c^2)} = \frac{1}{(1-|\mu_i|)^{2p}} - \frac{1}{N_c^2(1-|\mu_i|)^{4p}} + \frac{1}{N_c^4(1-|\mu_i|)^{6p}} - \cdots.$$

This yields $\left\|(I - A_\Delta B_\Delta^{-1})^p\right\|_{(UU^*)^{-1}}^2 = \varphi_F^{2p} - O(1/N_c^2)$. An analogous proof confirms the result for FCF-relaxation. ∎

Interestingly, despite having $1-|\mu_i|$ in the denominator, it is typically *not* eigenvalues $|\mu_i| \approx 1$ for which the maximum $\varphi_F$ is obtained [5]. To that end, the $O(1/N_c^2)$ in Theorem 32 will generally be fairly small, except for potentially in the case of very large $p$. How tight the bounds are clearly depends on the size of $p$ and $N_c$; in practice, however, Theorem 32 indicates that the upper bound on convergence in the $(UU^*)^{-1}$-norm will generally not improve in later iterations, that is, $\|\mathcal{E}^p\|_{(UU^*)^{-1}} \approx \|\mathcal{E}\|^p_{(UU^*)^{-1}}$.

These results also lead to a corollary, which proves that, in some cases, the bounds derived in [5] are asymptotically exact in $N_c$ in a single-iteration sense.

COROLLARY 33 (Sharp matrix inequalities). *The matrix norm inequality* $\|M\|_2^2 \leq \|M\|_1\|M\|_\infty$ [2, Fact 11.9.27] *is asymptotically exact, for* $\left[I - A_\Delta B_\Delta^{-1}\right]_i$ *and* $[(I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}]_i$*; that is,*

$$\lim_{N_c\to\infty} \left\| \left[I - A_\Delta B_\Delta^{-1}\right]_i \right\|_2^2 = \left\| \left[I - A_\Delta B_\Delta^{-1}\right]_i \right\|_1 \left\| \left[I - A_\Delta B_\Delta^{-1}\right]_i \right\|_\infty, \tag{66}$$

*and likewise for* $\|[(I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}]_i\|_2^2$.

*Moreover,*

$$\begin{aligned}
\lim_{N_c\to\infty} \left\|(I - A_\Delta B_\Delta^{-1})^p\right\|_{(UU^*)^{-1}} &= \lim_{N_c\to\infty} \left\|I - A_\Delta B_\Delta^{-1}\right\|^p_{(UU^*)^{-1}} \\
&= \left(\sup_i \frac{|\mu_i - \lambda_i^k|}{1-|\mu_i|}\right)^p, \\
\lim_{N_c\to\infty} \left\|\left((I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right)^p\right\|_{(UU^*)^{-1}} &= \lim_{N_c\to\infty} \left\|(I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right\|^p_{(UU^*)^{-1}} \\
&= \left(\sup_i \frac{|\lambda_i^k||\mu_i - \lambda_i^k|}{1-|\mu_i|}\right)^p.
\end{aligned}$$

*Proof.* The inequality in (66) was used in [5] to establish bounds

$$\begin{aligned}
\left\|\left[I - A_\Delta B_\Delta^{-1}\right]_i\right\| &\leq \frac{|\mu_i - \lambda_i^k|(1-|\mu_i|^{N_c-1})}{1-|\mu_i|}, \\
\left\|\left[(I - A_\Delta B_\Delta^{-1})A_{cf}A_{ff}^{-1}A_{fc}\right]_i\right\| &\leq \frac{|\lambda_i^k||\mu_i - \lambda_i^k|(1-|\mu_i|^{N_c-2})}{1-|\mu_i|}.
\end{aligned}$$

It is easily verified that as $N_c \to \infty$, these bounds asymptote as in (63). The second result follows from a limiting argument applied to Theorem 32. ∎

## 6. Time-dependent operators.

**6.1. The general case.** The previous section focused on the specific case of commuting, diagonalizable time-stepping operators. This section moves on to the more general setting of (almost) arbitrary, linear time-stepping operators. In particular, we drop the assumption that $\Phi$ and $\Psi$ are fixed for all time-points, allowing for $\Phi$ and $\Psi$ to be time-dependent operators. Much of the theory so far has, on some level, been based on Toeplitz matrix theory. Allowing for time-dependent operators leads to non-Toeplitz block matrices, and such theory does not apply. Indeed, without some further assumptions or knowledge of $\Phi$ and $\Psi$, in general results cannot be extended to the time-dependent setting. However, this section shows that the pseudoinverse of $I - A_\Delta B_\Delta^{-1}$ derived in Lemma 20 can indeed be extended to the time-dependent setting. Although bounds for its minimum singular value are not clear, the resulting

bidiagonal matrix is still more amenable to analysis than the dense lower-triangular matrix of $I - A_\Delta B_\Delta^{-1}$ (33).

Note that, in the time-dependent setting, $I - B_\Delta^{-1} A_\Delta \neq I - A_\Delta B_\Delta^{-1}$, even if $\Phi$ and $\Psi$ commute. To that end, this section considers $I - A_\Delta B_\Delta^{-1}$, corresponding to residual propagation in the $\ell^2$-norm and error propagation in the $A^*A$-norm. However, similar results can be derived for error in the $\ell^2$-norm based on analogous derivations applied to $I - B_\Delta^{-1} A_\Delta$.

For preliminary notation, assume that we are considering $N$ time-points and a coarse grid of $N_c$ time-points. Then the linear system corresponding to time integration (1) takes the generalized form of

$$A\mathbf{u} = \begin{bmatrix} I & & & & \\ -\Phi_1 & I & & & \\ & -\Phi_2 & I & & \\ & & \ddots & \ddots & \\ & & & -\Phi_{N-1} & I \end{bmatrix} \begin{bmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_{N-1} \end{bmatrix} = \mathbf{f}, \tag{67}$$

where $\Phi_i$ denotes $\Phi$ evaluated at time-point $t_i$. As in the time-independent case, there is a closed form for inverses of the form of (67), which will prove useful for further derivations:

(68)

$$\begin{bmatrix} I & & & & \\ -\Phi_1 & I & & & \\ & -\Phi_2 & I & & \\ & & \ddots & \ddots & \\ & & & -\Phi_{N-1} & I \end{bmatrix}^{-1} = \begin{bmatrix} I & & & & & & \\ \Phi_1 & I & & & & & \\ \Phi_2\Phi_1 & \Phi_2 & I & & & & \\ \Phi_3\Phi_2\Phi_1 & \Phi_3\Phi_2 & \Phi_3 & I & & & \\ \vdots & \vdots & & \ddots & \ddots & & \\ \Phi_{N-1}...\Phi_1 & \Phi_{N-1}...\Phi_2 & ... & ... & \Phi_{N-1} & I \end{bmatrix}.$$

Excusing the slight abuse of notation, define $\Phi_i^j := \Phi_i \Phi_{i-1} \dots \Phi_j$. Then, using the inverse in (68) and analogous matrix derivations as in section 3.4 leads to a Schur complement coarse grid given by

$$A_\Delta = \begin{bmatrix} I & & & & \\ -\Phi_k^1 & I & & & \\ & -\Phi_{2k}^{k+1} & I & & \\ & & \ddots & \ddots & \\ & & & -\Phi_{(N_c-1)k}^{(N_c-2)k+1} & I \end{bmatrix}.$$

Notice that, despite the more complicated notation, the Schur complement coarse-grid operator in the time-dependent case does exactly what it does in the time-independent case (31): it takes exactly $k$ steps on the fine grid, in this case using the appropriate sequence of time-dependent operators. Let $\Psi_i$ denote the non-Galerkin approximation to $\Phi_{ik}^{(i-1)k+1}$. Then, the operator we are primarily interested in for error and residual propagation, $I - A_\Delta B_\Delta^{-1}$, is given by

(69)

$$= \begin{bmatrix} \mathbf{0} & & & & & \\ \Phi_k^1 - \Psi_1 & \mathbf{0} & & & & \\ (\Phi_{2k}^{k+1} - \Psi_2)\Psi_1 & (\Phi_{2k}^{k+1} - \Psi_2) & \mathbf{0} & & & \\ (\Phi_{3k}^{2k+1} - \Psi_3)\Psi_2^1 & (\Phi_{3k}^{2k+1} - \Psi_3)\Psi_2 & \Phi_{3k}^{2k+1} - \Psi_3 & \mathbf{0} & & \\ \vdots & \vdots & & \ddots & \ddots & \\ \left(\Phi_{(N_c-1)k}^{(N_c-2)k+1} - \Psi_{N_c-1}\right)\Psi_{N_c-2}^1 & & \cdots & & \Phi_{(N_c-1)k}^{(N_c-2)k+1} - \Psi_{N_c-1} & \mathbf{0} \end{bmatrix}.$$

Moreover, the pseudoinverses derived in Lemma 20 can be extended to the time-dependent case.

LEMMA 34 (Time-dependent pseudoinverse). *Let* $\{\Phi_i\}_{i=1}^{N-1}$ *and* $\{\Psi_i\}_{i=1}^{N_c-1}$ *denote two sets of operators and, for notation, define* $\Phi_i^j := \Phi_i\Phi_{i-1}\dots\Phi_j$. *Assume that* $(\Phi_{ik}^{(i-1)k+1} - \Psi_i)$ *is invertible for* $i = 1, \dots, N_c - 1$ *and define* $I - A_\Delta B_\Delta^{-1}$ *as in* (69). *Then,*

(70)

$$(I - A_\Delta B_\Delta^{-1})^\dagger = \begin{bmatrix} \mathbf{0} & (\Phi_k^1 - \Psi_1)^{-1} & & & \\ \mathbf{0} & -\Psi_1(\Phi_k^1 - \Psi_1)^{-1} & (\Phi_{2k}^{k+1} - \Psi_2)^{-1} & & \\ & \ddots & \ddots & & \\ & & & -\Psi_{N_c-2}\left(\Phi_{(N_c-2)k}^{(N_c-3)k+1} - \Psi_{N_c-2}\right)^{-1} & \left(\Phi_{(N_c-1)k}^{(N_c-2)k+1} - \Psi_{N_c-1}\right)^{-1} \\ & & & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

*Proof.* Following from the proof of Lemma 20, it is easy to confirm from (69) and (70) that $(I - A_\Delta B_\Delta^{-1})^\dagger$ satisfies the four properties of a pseudoinverse. ∎

As in the time-dependent case, we seek the minimum nonzero singular value of $(I - A_\Delta B_\Delta^{-1})^\dagger$, which is equivalent to the maximum singular value (and $\ell^2$-norm) of $I - A_\Delta B_\Delta^{-1}$. This can be expressed as a minimization over a linear combination of operators as follows:

$$\begin{aligned} \sigma_{\min}\left((I - A_\Delta B_\Delta^{-1})^\dagger\right)^2 &= \min_{\mathbf{v}\notin\ker\left((I - A_\Delta B_\Delta^{-1})^\dagger\right)} \frac{\left\|(I - A_\Delta B_\Delta^{-1})^\dagger\mathbf{v}\right\|^2}{\|\mathbf{v}\|^2} \\ &= \min_{\mathbf{v}_i, i=1,\dots,(N_c-1)} \frac{\|(\Phi_k^1 - \Psi_1)^{-1}\mathbf{v}_1\|^2 + \sum_{i=1}^{N_c-2}\left\|(\Phi_{(i+1)k}^{ik+1} - \Psi_{i+1})^{-1}\mathbf{v}_{i+1} - \Psi_i(\Phi_{ik}^{(i-1)k+1} - \Psi_i)^{-1}\mathbf{v}_i\right\|^2}{\sum_{i=1}^{N_c-1}\|\mathbf{v}_i\|^2} \\ &= \min_{\mathbf{v}_i, i=1,\dots,(N_c-1)} \frac{\|\mathbf{v}_1\|^2 + \sum_{i=1}^{N_c-2}\|\mathbf{v}_{i+1} - \Psi_i\mathbf{v}_i\|^2}{\sum_{i=1}^{N_c-1}\left\|(\Phi_{ik}^{(i-1)k+1} - \Psi_i)\mathbf{v}_i\right\|^2}. \end{aligned} \tag{71}$$

Further analysis likely requires some knowledge of $\{\Phi_i\}$ and $\{\Psi_i\}$. In particular, now we are letting $\{\Phi_i\}$ and $\{\Psi_i\}$ be completely arbitrary operators. In practice, there is typically some continuity in how operators change between time steps; that is, $\Phi_i$ and $\Phi_{i+1}$ are similar in some sense.

**6.2. The diagonalizable case.** Finally, suppose that $\Phi$ and $\Psi$ are time-dependent but simultaneously diagonalizable for all times, $t_i$. In particular this allows for time-dependent reaction terms in the spatial operator, $\mathcal{L}$, and for variable time-step size, as occurs in, for example, adaptive time stepping. To that end,

$$\|I - A_\Delta B_\Delta^{-1}\|_{(UU^*)^{-1}} = \max_i \left\|\left[I - A_\Delta B_\Delta^{-1}\right]_i\right\| = \frac{1}{\min_i \left\|\left[(I - A_\Delta B_\Delta^{-1})^\dagger\right]_i\right\|}$$

$$= \frac{1}{\min_i \sigma_{max}\left(\left[(I - A_\Delta B_\Delta^{-1})^\dagger\right]_i\right)},$$

where, recall, $\left[I - A_\Delta B_\Delta^{-1}\right]_i$ denotes (70) evaluated at eigenvalues of $\Phi$ and $\Psi$ as opposed to the actual operators. As previously, $\sigma_{max}\left(\left[(I - A_\Delta B_\Delta^{-1})^\dagger\right]_i\right)$ is given by the square root of the minimum nonzero eigenvalue of the corresponding normal residual equations (that is, $MM^*$ as opposed to $M^*M$). Eliminating the final zero-row and zero-column (corresponding to the zero eigenvalue), this is equivalent to the minimum eigenvalue of the tridiagonal matrix

(72)
$$\begin{bmatrix} \frac{1}{|\lambda_k^1-\mu_1|^2} & \frac{-\overline{\mu}_1}{|\lambda_k^1-\mu_1|^2} & & & \\ \frac{-\mu_1}{|\lambda_k^1-\mu_1|^2} & \frac{|\mu_1|^2}{|\lambda_k^1-\mu_1|^2}+\frac{1}{|\lambda_{2k}^{k+1}-\mu_2|^2} & \frac{-\overline{\mu}_2}{|\lambda_{2k}^{k+1}-\mu_2|^2} & & \\ & \frac{-\mu_2}{|\lambda_{2k}^{k+1}-\mu_2|^2} & \frac{|\mu_2|^2}{|\lambda_{2k}^{k+1}-\mu_2|^2}+\frac{1}{|\lambda_{3k}^{2k+1}-\mu_3|^2} & \frac{-\overline{\mu}_3}{|\lambda_{3k}^{2k+1}-\mu_3|^2} & \\ & & \ddots & \ddots & \ddots \end{bmatrix}.$$

*Note the change in notation*—here, for example, $\lambda_k^1$ denotes the product of the $i$th eigenvalue of $\Phi$ at times $t_1, \ldots, t_k$, that is, $\lambda_k^1 = \lambda_i(t_k)\lambda_i(t_{k-1})\ldots\lambda_i(t_1)$, and $\mu_1$ denotes the $i$th eigenvalue of $\Psi$ evaluated at time $t_1$. The $i$ is dropped from eigenvalues to limit subscript/superscript notation.

This leads to the final result on convergence in the time-dependent case.

THEOREM 35. *Let $\{\lambda_j\}^{(i)}$ and $\{\mu_j\}^{(i)}$ denote the sets of the ith eigenvalue of $\Phi$ and $\Psi$, respectively, evaluated at time indices, $j = 1, \ldots, N_c - 1$. Let $\hat{\sigma}_{\min}$ denote the minimum nonzero eigenvalue of* (72)*. Then,*

$$\|I - A_\Delta B_\Delta^{-1}\|_{(UU^*)^{-1}} = \frac{1}{\sqrt{\hat{\sigma}_{\min}}}.$$

*Now, assume that for each eigenvalue index $i$, a TEAP-like approximation property holds, where, for all $j = 1, \ldots, N_c - 1$,*

$$|\lambda_{jk}^{(j-1)k+1} - \mu_j|^2 \leq \widehat{\varphi}_j^{(i)}|1 - \mu_j|.$$

*Then,*

$$\|I - A_\Delta B_\Delta^{-1}\|_{(UU^*)^{-1}}^2 \leq \max_i \max\left\{\widehat{\varphi}_1^{(i)}, \max_{j=1,\ldots,N_c-2} \frac{\widehat{\varphi}_j^{(i)}\widehat{\varphi}_{j+1}^{(i)}}{\widehat{\varphi}_j^{(i)} + \mu_j\widehat{\varphi}_{j+1}^{(i)}}\right\}.$$

*Moreover, sufficient conditions for convergence are that for all $i$ and for all $j$,*

$$\frac{\widehat{\varphi}_j^{(i)}\widehat{\varphi}_{j+1}^{(i)}}{\widehat{\varphi}_j^{(i)} + \mu_j\widehat{\varphi}_{j+1}^{(i)}} < 1,$$

*and, in addition, that for all $i$, $\widehat{\varphi}_1^{(i)} < 1$.*

*Proof.* The proof follows by applying the Gershgorin circles theorem to (72) to bound the minimum eigenvalue from below, and using this to bound the maximum singular value of $\left[I - A_\Delta B_\Delta^{-1}\right]_i$ from above. For example, forming the Gershgorin disc for row one of (72) yields a lower bound:

$$\frac{1 - |\mu_2|}{|\lambda_{2k}^{k+1} - \mu_2|^2} + \frac{|\mu_1|(1 - |\mu_1|)}{|\lambda_k^1 - \mu_1|^2} \geq \frac{1}{\widehat{\varphi}_1^{(i)}} + \frac{|\mu_1|}{\widehat{\varphi}_2^{(i)}}.$$

Repeating for all rows and inverting yields the result. ∎

Note from section 5 that in the diagonalizable case, Gershgorin is indeed asymptotically tight on the nonboundary rows. The Gershgorin disc for row $i > 1$ of (62) bounds the minimum eigenvalue below by $1 + |\mu_i|^2 - 2|\mu_i| = (1 - |\mu_i|)^2$, which is exactly the minimum eigenvalue of (62) to $O(1/N_c^2)$. In practice, it is likely that the bound in Theorem 35 is not tight, in particular the boundary term $\widehat{\varphi}_1^{(i)}$, and that convergence will actually resemble $\widehat{\varphi}_j^{(i)}\widehat{\varphi}_{j+1}^{(i)}$.

Fortunately, although we cannot derive a closed form for eigenvalues of (72), a simple estimate of the convex hull allows us to compute exact convergence bounds on two-level Parareal and MGRiT by solving a tridiagonal eigenvalue problem, which is a computationally tractable task. This result allows for rigorous, problem-specific analysis in some time-dependent cases, such as adaptive time stepping.

**7. Conclusion.** This paper derives necessary and sufficient conditions for the convergence of Parareal and MGRiT, assuming that time-stepping operators $\Phi$ and $\Psi$ are linear and not time-dependent, and sufficient conditions for a subset of the general linear (time-dependent) case. This is accomplished by introducing a temporal approximation property (TAP), which gives a measure of how accurately $\Phi^k$ approximates the action of $\Psi$, for any vector $\mathbf{v}$. How accurately the TAP is satisfied then defines the $\ell^2$-norm and $A^*A$-norm of error reduction over successive iterates. With further assumptions on the diagonalizability of $\Phi$ and $\Psi$, these results are strengthened to give tight bounds on an arbitrary number of iterations.

For space-time PDEs with a symmetric positive semidefinite (or symmetric negative semidefinite) spatial component, the real eigenvalues of $\Phi$ and $\Psi$ can be explicitly computed as a function of time-step size, $\delta t$, and eigenvalues of the spatial operator. With a simple estimate of the minimum and maximum eigenvalues of the spatial operator, exact bounds on the convergence of Parareal and MGRiT can be easily computed by evaluating the TEAP over the range of eigenvalues of $\Phi$ and $\Psi$. In the general case, for example, that arises in hyperbolic PDEs, the eigenvectors no longer form an orthogonal basis and the TAP does not reduce to just considering eigenvalues. However, for most time-stepping schemes applied to some operator $\mathcal{L}$, it is straightforward to expand $\Phi$ and $\Psi$ in terms of $\mathcal{L}$. This permits a robust method to derive the expected convergence of Parareal and MGRiT applied to a problem of the form $\mathbf{u}_t = \mathcal{L}\mathbf{u} + \mathbf{g}$, for arbitrary $\mathcal{L}$.

Further research regarding the optimal $\Psi$ with respect to $\Phi$, the difficulties in solving hyperbolic problems, and the more general time-dependent and nonlinear cases are ongoing work.

## Appendix A.

LEMMA 36 (Minimum eigenvalue of tridiagonal Toeplitz perturbation). *Define the $n \times n$ tridiagonal matrix*

$$\mathcal{D}_i = \begin{bmatrix} 1+|\mu_i|^2 & -\overline{\mu}_i & & \\ -\mu_i & \ddots & \ddots & \\ & \ddots & 1+|\mu_i|^2 & -\overline{\mu}_i \\ & & -\mu_i & 1 \end{bmatrix}. \tag{73}$$

*The minimum eigenvalue of $\mathcal{D}_i$, denoted $\lambda_n$, is bounded by*

$$\begin{aligned}(1-|\mu_i|)^2+\frac{\pi^2|\mu_i|}{6n^2} &\leq 1+|\mu_i|^2+2|\mu_i|\cos\left(\frac{n\pi}{n+1/2}\right)\\ &\leq \lambda_n\\ &\leq 1+|\mu_i|^2+2|\mu_i|\cos\left(\frac{n\pi}{n+1}\right)\\ &\leq (1-|\mu_i|)^2+\frac{\pi^2|\mu_i|}{n^2}.\end{aligned} \tag{74}$$

*Proof.* Let $\widehat{\mathcal{D}}_i$ denote the self-adjoint, tridiagonal, Toeplitz matrix for which $\mathcal{D}_i$ is a rank-one perturbation. In the scalar setting, there is a closed form for eigenvalues of a tridiagonal Toeplitz matrix of size $n\times n$, given by

$$\lambda\left(\widehat{\mathcal{D}}_i\right)=\left\{1+|\mu_i|^2+2|\mu_i|\cos\left(\frac{\ell\pi}{n+1}\right)\ \mid\ \ell=1,\dots,n\right\}. \tag{75}$$

Returning to (73), consider the rank-one perturbation in $\mathcal{D}_i$. Following from [44], the spectrum of a tridiagonal Toeplitz matrix and single-entry perturbations is derived by building and solving a three-term recursion relation. One of the general results in [44, eq. (7)] states that all eigenvalues of the matrices in (73) take the form

$$\lambda=1+|\mu_i|^2+2|\mu_i|\cos(\theta) \tag{76}$$

for a given $\theta\neq m\pi$, $m\in\mathbb{Z}$. In the case of a Toeplitz matrix, the necessary conditions on $\theta$ are $\sin((n+1)\theta)=0$, which is satisfied for $\widehat{\theta}_\ell=\frac{\ell\pi}{n+1}$, $\ell=1,\dots,n$, yielding the result in (75). For the perturbation in $\mathcal{D}_i$, necessary conditions on $\theta$ are that [44, eq. (6)]

$$T(\theta):=\sin((n+1)\theta)+|\mu_i|\sin(n\theta)=0. \tag{77}$$

Unfortunately, (77) does not have a closed-form solution as found in the Toeplitz case and several other perturbations with closed-form spectra, introduced in [44, 4]. However, each eigenvalue of $\mathcal{D}_i$ can be shown to be a small perturbation to eigenvalues of $\widehat{\mathcal{D}}_i$. Denote $\{\widehat{\lambda}_\ell\}_{\ell=1}^n$ as the eigenvalues of $\widehat{\mathcal{D}}_i$, with corresponding $\theta$-values $\{\widehat{\theta}_\ell\}_{\ell=1}^n$, and $\{\lambda_\ell\}_{\ell=1}^n$ the eigenvalues of $\mathcal{D}_i$, with corresponding $\theta$-values $\{\theta_\ell\}_{\ell=1}^n$.

Consider $\widehat{\theta}_\ell=\frac{\ell\pi}{n+1}$, $\ell=1,\dots,n$, which yields all $n$ eigenvalues of $\widehat{\mathcal{D}}_i$ in the context of necessary conditions for $\mathcal{D}_i$ (77):

$$T\left(\widehat{\theta}_\ell\right)=|\mu_i|\sin\left(\frac{n}{n+1}\ell\pi\right)\mapsto\begin{cases}<0 & 2|\ell,\\ >0 & 2\not|\ell.\end{cases} \tag{78}$$

Now, define $\widetilde{\theta}_\ell=\frac{\ell\pi}{n+\frac{1}{2}}$ for $\ell=1,\dots,n$. Then, under the assumption that $|\mu_i|<1$,

$$\begin{aligned}T(\widetilde{\theta}_\ell)&=\sin\left(\frac{n+1}{n+\frac{1}{2}}\ell\pi\right)+|\mu_i|\sin\left(\frac{n}{n+\frac{1}{2}}\ell\pi\right)\\ &=-\sin\left(\frac{n}{n+\frac{1}{2}}\ell\pi\right)+|\mu_i|\sin\left(\frac{n}{n+\frac{1}{2}}\ell\pi\right)\\ &=-(1-|\mu_i|)\sin\left(\frac{n}{n+\frac{1}{2}}\ell\pi\right)\end{aligned}$$

$$\mapsto \begin{cases} > 0 & 2|\ell, \\ < 0 & 2 \not| \ell. \end{cases} \tag{79}$$

From (78), (79), and the continuity of $T(\theta)$, it follows that there exists $\theta_\ell \in (\frac{\ell\pi}{n+1}, \frac{\ell\pi}{n+\frac{1}{2}})$, $\theta_\ell \neq m\pi, m \in \mathbb{Z}$, such that $T(\theta_\ell) = 0$, for $\ell = 1, \ldots, n$. Following from (76), eigenvalues of $\mathcal{D}_i$ take the form

$$\lambda_\ell = 1 + |\mu_i|^2 + 2|\mu_i| \cos(\theta_\ell).$$

The smallest nonzero eigenvalue of $\mathcal{D}_i$ is given by $\lambda_n = 1 + |\mu_i|^2 + 2|\mu_i| \cos(\theta_n)$, where $\cos(\theta_\ell) \to -1$ as $\ell \to \infty$. Given $\frac{n\pi}{n+1} \leq \theta_n \leq \frac{n\pi}{n+1/2}$, $\lambda_n$ can then be bounded by

$$1 + |\mu_i|^2 + 2|\mu_i| \cos\left(\frac{n\pi}{n+1/2}\right) \leq \lambda_n \leq 1 + |\mu_i|^2 + 2|\mu_i| \cos\left(\frac{n\pi}{n+1}\right). \tag{80}$$

With a little extra work, we can show that $\lambda_n = 1 + |\mu_i|^2 + O(1/n^2)$, which leads to necessary and sufficient conditions for convergence.

Consider the term $\cos(\frac{n\pi}{n+\frac{1}{2}})$. As $n \to \infty$, $\cos(\frac{n\pi}{n+1}) \to^+ -1$, that is, from above. Consider a series expansion of $f(n) = \cos(\frac{n\pi}{n+\frac{1}{2}})$ about $n = \infty$. To accomplish this, apply the change of variable $n = \frac{1}{w}$ to get $f(w) = \cos(\frac{\frac{1}{w}\pi}{\frac{1}{w}+\frac{1}{2}})$, and expand about $w = 0$. Formally, this would be expanded as a Laurent series, but recognizing that $w^k f(w) = w^k \cos(\frac{\frac{1}{w}\pi}{\frac{1}{w}+\frac{1}{2}}) = w^k \cos(\frac{\pi}{1+\frac{w}{2}})$ is holomorphic about $w = 0$ for $k = 0, 1, \ldots,$ negative coefficients in the Laurent series are zero by the Cauchy Integral Theorem. Our expansion reduces to a Taylor expansion for $\cos(\frac{\pi}{1+\frac{w}{2}})$ about $w = 0$,

$$\cos\left(\frac{\pi}{1+\frac{w}{2}}\right) = -1 + \frac{\pi^2 w^2}{8} - \frac{\pi^3 w^3}{8} + O(w^4).$$

In fact, truncating the Taylor expansion at $k = 1$ leads to a remainder term

$$\begin{aligned} \cos\left(\frac{\pi}{1+\frac{w}{2}}\right) &= -1 + \int_0^w (w-t) \frac{\partial}{\partial_{tt}} \left[\cos\left(\frac{\pi}{1+\frac{t}{2}}\right)\right] dt \\ &= -1 + \int_0^w (w-t) \frac{-4\pi \left[\pi \cos\left(\frac{\pi}{1+\frac{t}{2}}\right) + (2+t)\sin\left(\frac{\pi}{1+\frac{t}{2}}\right)\right]}{(2+t)^4} dt. \end{aligned} \tag{81}$$

Note that

$$0 < \frac{\pi^2}{6} < \frac{-4\pi \left[\pi \cos\left(\frac{\pi}{1+\frac{t}{2}}\right) + (2+t)\sin\left(\frac{\pi}{1+\frac{t}{2}}\right)\right]}{(2+t)^4} \leq \frac{\pi^2}{4}$$

for all $t \in [0, 1/10]$ (this range is not tight; it is just sufficient for our purposes). By positivity of the two terms being integrated in the remainder (81), the remainder can be bounded above and below. For $w \in [0, 1/10]$, substituting for $n \geq 10$ yields

$$-1 + \frac{\pi^2}{12n^2} \leq \cos\left(\frac{n\pi}{n+\frac{1}{2}}\right) \leq -1 + \frac{\pi^2}{8n^2}. \tag{82}$$

A similar expansion on $\cos(\frac{n\pi}{n+1})$ yields a truncated Taylor expansion and remainder given by

$$\cos\left(\frac{\pi}{w+1}\right) = -1 + \int_0^w (w-t) \frac{-\pi \left[\pi \cos\left(\frac{\pi}{t+1}\right) + 2(1+t)\sin\left(\frac{\pi}{t+1}\right)\right]}{(1+t)^4} dt.$$

Here, note that

$$0 < \frac{\pi^2}{2} < \frac{-\pi\left[\pi\cos\left(\frac{\pi}{t+1}\right) + 2(1+t)\sin\left(\frac{\pi}{t+1}\right)\right]}{(1+t)^4} \leq \pi^2 \tag{83}$$

for $t \in [0, 1/10]$. Integrating the product of two positive functions, the bounds in (83) can be pulled out, and, for $w \in [0, 1/10]$, substituting for $n \geq 10$ yields

$$-1 + \frac{\pi^2}{4n^2} \leq \cos\left(\frac{n\pi}{n+1}\right) \leq -1 + \frac{\pi^2}{2n^2}.$$

Returning to the upper and lower bounds on $\lambda_n$ in (80) completes the proof. ∎

**Acknowledgments.** The author would like to thank Andreas Hessenthaler for bringing up convergence of MGRiT as a research topic and providing useful numerical comparisons, Professor Tom Manteuffel for his helpful comments and discussions, and Professor Stefano Serra-Capizzano for generously sharing his insight and expertise on block-Toeplitz matrix theory.

## REFERENCES

[1] G. Bal, *On the convergence and the stability of the parareal algorithm to solve partial differential equations*, in Domain Decomposition Methods in Science and Engineering, Springer, Berlin, 2005, pp. 425–432.

[2] D. S. Bernstein, *Scalar, Vector, and Matrix Mathematics: Theory, Facts, and Formulas*, revised and expanded edition, Princeton University Press, Princeton, NJ, 2018.

[3] W. L. Briggs, V. E. Henson, and S. F. McCormick, *A Multigrid Tutorial*, 2nd ed., SIAM, Philadelphia, 2000, https://doi.org/10.1137/1.9780898719505.

[4] S. S. Cheng and W. C. Yueh, *Explicit eigenvalues and inverses of tridiagonal Toeplitz matrices with four perturbed corners*, ANZIAM J., 49 (2008), pp. 361–387.

[5] V. A. Dobrev, Tz. V. Kolev, N. A. Petersson, and J. B. Schroder, *Two-level convergence theory for multigrid reduction in time (MGRIT)*, SIAM J. Sci. Comput., 39 (2017), pp. S501–S527, https://doi.org/10.1137/16M1074096.

[6] M. Emmett and M. L. Minion, *Toward an efficient parallel in time method for partial differential equations*, Commun. Appl. Math. Comput. Sci., 7 (2012), pp. 105–132.

[7] R. Falgout, S. Friedhoff, Tz. V. Kolev, S. MacLachlan, J. Schroder, and S. Vandewalle, *Multigrid methods with space–time concurrency*, Comput. Vis. Sci., 18 (2017), pp. 123–143.

[8] R. D. Falgout, S. Friedhoff, Tz. V. Kolev, S. P. MacLachlan, and J. B. Schroder, *Parallel time integration with multigrid*, SIAM J. Sci. Comput., 36 (2014), pp. C635–C661, https://doi.org/10.1137/130944230.

[9] R. D. Falgout, M. Lecouvez, and C. S. Woodward, *A Parallel-in-Time Algorithm for Variable Step Multistep Methods*, manuscript, 2017.

[10] R. D. Falgout and P. S. Vassilevski, *On generalizing the algebraic multigrid framework*, SIAM J. Numer. Anal., 42 (2004), pp. 1669–1693, https://doi.org/10.1137/S0036142903429742.

[11] S. Friedhoff and S. P. MacLachlan, *A Generalized predictive analysis tool for multigrid methods*, Numer. Linear Algebra Appl., 22 (2015), pp. 618–647.

[12] H. Gahvari, V. A. Dobrev, R. D. Falgout, Tz. V. Kolev, J. B. Schroder, M. Schulz, and U. M. Yang, *A performance model for allocating the parallelism in a multigrid-in-time solver*, in Proceedings of the IEEE International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS), IEEE, Washington, DC, 2016, pp. 22–31.

[13] M. J. Gander and E. Hairer, *Nonlinear convergence analysis for the parareal algorithm*, in Domain Decomposition Methods in Science and Engineering XVII, Springer, Berlin, Heidelberg, 2008, pp. 45–56.

[14] M. J. Gander, F. Kwok, and H. Zhang, *Multigrid interpretations of the parareal algorithm leading to an overlapping variant and MGRiT*, Comput. Vis. Sci., 19 (2018), pp. 59–74.

[15] M. J. Gander and S. Vandewalle, *Analysis of the parareal time-parallel time-integration method*, SIAM J. Sci. Comput., 29 (2007), pp. 556–578, https://doi.org/10.1137/05064607X.

[16] M. J. Gander and S. Vandewalle, *On the superlinear and linear convergence of the parareal algorithm*, in Domain Decomposition Methods in Science and Engineering XVI, Springer, Berlin, Heidelberg, 2007, pp. 291–298.

[17] A. Goddard and A. Wathen, *A note on parallel preconditioning for all-at-once evolutionary PDEs*, Electron. Trans. Numer. Anal., submitted; preprint, https://arxiv.org/abs/1810.00615, 2018.

[18] A. Hessenthaler, D. Nordsletten, O. Röhrle, J. B. Schroder, and R. D. Falgout, *Convergence of the multigrid reduction in time algorithm for the linear elasticity equations*, Numer. Linear Algebra Appl., 25 (2018), e2155, https://doi.org/10.1002/nla.2155.

[19] A. Hessenthaler, B. S. Southworth, D. Nordsletten, O. Röhrle, R. D. Falgout, and J. B. Schroder, *Multilevel convergence analysis of multigrid reduction in time*, SIAM J. Sci. Comput., in review.

[20] G. Horton and S. Vandewalle, *A space-time multigrid method for parabolic partial differential equations*, SIAM J. Sci. Comput., 16 (1994), pp. 848–864, https://doi.org/10.1137/0916050.

[21] A. Howse, H. De Sterck, S. MacLachlan, R. Falgout, and J. Schroder, *Multigrid Reduction in Time with Adaptive Spatial Coarsening for the Linear Advection Equation*, Tech. report, Lawrence Livermore National Laboratory (LLNL), Livermore, CA, 2017.

[22] R. J. LeVeque, *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*, SIAM, Philadelphia, 2007, https://doi.org/10.1137/1.9780898717839.

[23] J.-L. Lions, Y. Maday, and G. Turinici, *Resolution d'EDP par un schéma en temps "pararéel,"* C. R. Acad. Sci. Paris Sér. I Math., 332 (2001), pp. 661–668.

[24] S. P. MacLachlan, T. A. Manteuffel, and S. F. McCormick, *Adaptive reduction-based AMG*, Numer. Linear Algebra Appl., 13 (2006), pp. 599–620.

[25] T. A. Manteuffel, S. Münzenmaier, J. W. Ruge, and B. S. Southworth, *Nonsymmetric reduction-based algebraic multigrid*, SIAM J. Sci. Comput., in review.

[26] T. A. Manteuffel, J. W. Ruge, and B. S. Southworth, *Nonsymmetric algebraic multigrid based on local approximate ideal restriction (ℓAIR)*, SIAM J. Sci. Comput., 40 (2018), pp. A4105–A4130, https://doi.org/10.1137/17M1144350.

[27] E. McDonald, J. Pestana, and A. Wathen, *Preconditioning and iterative solution of all-at-once systems for evolutionary partial differential equations*, SIAM J. Sci. Comput., 40 (2018), pp. A1012–A1033, https://doi.org/10.1137/16M1062016.

[28] N. Mendelsohn, *Some properties of approximate inverses of matrices*, Trans. Roy. Soc. Canada Sect. III (3), 50 (1956), pp. 53–59.

[29] M. Miranda and P. Tilli, *Asymptotic spectra of Hermitian block Toeplitz matrices and preconditioning results*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 867–881, https://doi.org/10.1137/S0895479896313036.

[30] J. M. Reynolds-Barredo, D. E. Newman, R. Sanchez, D. Samaddar, L. A. Berry, and W. R. Elwasif, *Mechanisms for the convergence of time-parallelized, parareal turbulent plasma simulations*, J. Comput. Phys., 231 (2012), pp. 7851–7867.

[31] M. Ries, U. Trottenberg, and G. Winter, *A note on MGR methods*, Linear Algebra Appl., 49 (1983), pp. 1–26.

[32] D. Ruprecht, *Convergence of parareal with spatial coarsening*, Proc. Appl. Math. Mech., 14 (2014), pp. 1031–1034.

[33] D. Ruprecht, *Wave propagation characteristics of parareal*, Comput. Vis. Sci., 19 (2018), pp. 1–17.

[34] S. Serra, *Asymptotic results on the spectra of block Toeplitz preconditioned matrices*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 31–44, https://doi.org/10.1137/S0895479896310160.

[35] S. Serra, *On the extreme eigenvalues of Hermitian (block) Toeplitz matrices*, Linear Algebra Appl., 270 (1998), pp. 109–129.

[36] S. Serra, *Spectral and computational analysis of block Toeplitz matrices having nonnegative definite matrix-valued generating functions*, BIT, 39 (1999), pp. 152–175.

[37] S. Serra and P. Tilli, *Extreme singular values and eigenvalues of non-Hermitian block Toeplitz matrices*, J. Comput. Appl. Math., 108 (1999), pp. 113–130.

[38] P. Tilli, *Singular values and eigenvalues of non-Hermitian block Toeplitz matrices*, Linear Algebra Appl., 272 (1998), pp. 59–89.

[39] S.-L. Wu, *Convergence analysis of some second-order parareal algorithms*, IMA J. Numer. Anal., 35 (2015), pp. 1315–1341.

[40] S.-L. Wu and T. Zhou, *Convergence analysis for three Parareal solvers*, SIAM J. Sci. Comput., 37 (2015), pp. A970–A992, https://doi.org/10.1137/140970756.
[41] W. Wyss, *Quasi-multiplication and *-algebras*, Rep. Math. Phys., 18 (1980), pp. 87–95.
[42] W. Wyss, *Two Non-Commutative Binomial Theorems*, preprint, https://arxiv.org/abs/1707.03861, 2017.
[43] X. Xu and C.-S. Zhang, *On the ideal interpolation operator in algebraic multigrid methods*, SIAM J. Numer. Anal., 56 (2018), pp. 1693–1710, https://doi.org/10.1137/17M1162779.
[44] W. C. Yueh, *Eigenvalues of several tridiagonal matrices*, Appl. Math. E-Notes, 5 (2005), pp. 210–230.
[45] H. Zha, *Computing the generalized singular values/vectors of large sparse or structured matrix pairs*, Numer. Math., 72 (1996), pp. 391–417.