Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 227 (2007) 244–263
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Linear multifrequency-grey acceleration recast for preconditioned Krylov iterations

Jim E. Morel [a,*], T.-Y. Brian Yang [b], James S. Warsa [c]

[a] Texas A&M University, College Station, TX 77843, United States
[b] Lawrence Livermore National Laboratory, Livermore, CA 94550, United States
[c] Los Alamos National Laboratory, Los Alamos, NM 87545, United States

Received 19 April 2007; received in revised form 20 July 2007; accepted 23 July 2007
Available online 17 August 2007

Abstract

The linear multifrequency-grey acceleration (LMFGA) technique is used to accelerate the iterative convergence of multigroup thermal radiation diffusion calculations in high energy density simulations. Although it is effective and efficient in one-dimensional calculations, the LMFGA method has recently been observed to significantly degrade under certain conditions in multidimensional calculations with large discontinuities in material properties. To address this deficiency, we recast the LMFGA method in terms of a preconditioned system that is solved with a Krylov method (LMFGK). Results are presented demonstrating that the new LMFGK method always requires fewer iterations than the original LMFGA method. The reduction in iteration count increases with both the size of the time step and the inhomogeneity of the problem. However, for reasons later explained, the LMFGK method can cost more per iteration than the LMFGA method, resulting in lower but comparable efficiency in problems with small time steps and weak inhomogeneities. In problems with large time steps and strong inhomogeneities, the LMFGK method is significantly more efficient than the LMFGA method.

Keywords: Thermal radiation diffusion; Preconditioned Krylov methods; Numerical solution methods

1. Introduction

The linear multifrequency-grey acceleration (LMFGA) technique has proven to be very effective for improving the iterative convergence of multigroup thermal radiation diffusion calculations in high energy density simulations [1]. Although it appears to be unconditionally effective in one-dimensional problems, the method has recently been observed by one of the authors (B. Yang) and others at Lawrence Livermore National Laboratory to significantly degrade in multidimensional problems with large discontinuities in material properties. A similar deficiency has been observed for the diffusion-synthetic acceleration (DSA) scheme associated with the transport equation [2]. Using the DSA method as a preconditioner for Krylov iterations

* Corresponding author. Tel.: +1 979 845 6072; fax: +1 979 845 6075.
E-mail address: morel@tamu.edu (J.E. Morel).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2007.07.033

addresses this deficiency, restoring the effectiveness and efficiency of the method [3]. The purpose of this paper is to recast the LMFGA method for preconditioned Krylov iterations, i.e., to recast it as a preconditioned system solved with a standard Krylov method. This is done with the intent of obtaining an iterative solution technique for solving the multigroup thermal radiation diffusion equations that is effective and efficient under all conditions. We refer to our solution method as the linear multifrequency-grey Krylov (LMFGK) technique. Our computational testing indicates that the LMFGK method always requires fewer iterations than the original LMFGA method. The reduction in iteration count increases with both the size of the time step and the inhomogeneity of the problem. However, for reasons later explained, the LMFGK method can cost more per iteration than the LMFGK method, resulting in lower but comparable efficiency in problems with small time steps and weak inhomogeneities. In problems with large time steps and strong inhomogeneities, the LMFGK method is significantly more efficient than the LMFGA method.

The remainder of this paper is organized as follows. In Section 2 we give a brief description of the use of Krylov methods. In Section 3 we describe the equations of thermal radiation diffusion. In Section 4 we present an overview of the LMFGA technique. In Section 5 we show how to recast the LMFGA technique as a preconditioned Krylov method for solving the multigroup thermal radiation diffusion equations (LMFGK). In Section 6 we show a simple relationship between the LMFGA and LMFGK techniques. In Section 7 we describe our discrete equations. Computational results are presented in Section 8 comparing the performance of the LMFGK method with that of the original LMFGA method. Finally, in Section 9, we give conclusions and recommendations for future work.

2. Krylov methods

It is sufficient for our purposes to briefly describe how a Krylov iteration is used in an implementation. A theoretical description of Krylov methods is given in Ref. [4]. Suppose we wish to solve a linear system of the form

$$\mathbf{M}\vec{x} = \vec{b}, \tag{1}$$

for the solution $\vec{x}$ with a Krylov iterative method software package. The user supplies the right hand side and an initial guess. At each Krylov iteration, the package supplies a vector $\vec{z}$ and the user returns a vector $\vec{w} = \mathbf{M}\vec{z}$ to the package. This is referred to as the action of the operator $\mathbf{M}$ on $\vec{z}$ because the operator need not be explicitly constructed to generate $\vec{w}$ (which is the case in our application). Most Krylov methods require the action of $\mathbf{M}$ only once per iteration, CG and GMRES for example, but some may require the action more than once, BiCGStab for example.

The convergence rate for Krylov iterative methods can be difficult to predict in general. Nonetheless, some qualitative criteria for the spectrum of an operator (matrix) $\mathbf{M}$ that can lead to rapid iterative convergence are that the eigenvalues are bounded away from zero, as tightly clustered as possible, and preferably close to unity. If an operator does not have these characteristics, a good preconditioner will have the effect of clustering the eigenvalues and moving them away from zero and towards unity. A precise description of how these very loose criteria result in good convergence rates for GMRES, in particular, can be found in Ref. [5].

3. The thermal radiation multigroup diffusion equations

The equations of multigroup thermal radiation diffusion can be written as follows:

$$\frac{1}{c}\frac{\partial \phi_g}{\partial t} - \vec{\nabla}\cdot D_g \vec{\nabla}\phi_g + \sigma_{a,g}\phi_g = \sigma_{a,g} 4\pi B_g(T), \quad g = 1, G, \tag{2}$$

and

$$C_v \frac{\partial T}{\partial t} = \sum_{g=1}^{G} \sigma_{a,g}[\phi_g - 4\pi B_g(T)], \tag{3}$$

where $\phi_g(\vec{r})$ is the angularly integrated radiation intensity for group g, $T(\vec{r})$ is the material temperature, c is the speed of light, $D_g(\vec{r}, T)$ is the diffusion coefficient for group g, $\sigma_{a,g}(\vec{r}, T)$ is the macroscopic absorption cross-

section for group g, $C_v(\vec{r},T)$ is the material heat capacity, and $B_g(T)$ is the Planck function integrated over group g:

$$B_g(T) = \int_{\Delta_g} \frac{2E^3}{h^3c^2}\left[\exp\left(\frac{E}{kT}\right)-1\right]^{-1} \mathrm{d}E, \tag{4}$$

where h is Planck's constant and k is Boltzmann's constant. The standard definition for the diffusion coefficient is

$$D_g = \frac{1}{3\sigma_{t,g}}, \tag{5}$$

where $\sigma_{t,g}$ denotes the total macroscopic cross-section (absorption plus Thompson scattering). Linearizing Eqs. (2) and (3) about an arbitrary temperature, T^*, and discretizing them in time over the interval $\left[t^{n-\frac{1}{2}}, t^{n+\frac{1}{2}}\right]$ using the backward Euler method, we obtain the following multigroup diffusion equations:

$$-\vec{\nabla}\cdot D_g^*\vec{\nabla}\phi_g + \sigma_{\tau,g}^*\phi_g - v\chi_g\sum_{k=1}^{G}\sigma_{a,k}^*\phi_k = \xi_g, \quad g = 1,G, \tag{6}$$

where:

$$\sigma_{\tau,g}^* = \sigma_{a,g}^* + \tau, \tag{7a}$$

$$\tau = \frac{1}{c\Delta t^n}, \tag{7b}$$

$$v = \frac{\sum_{g=1}^{G}\sigma_{a,g}^* 4\pi\frac{\partial B_g^*}{\partial T}}{\frac{C_v^*}{\Delta t^n}+\sum_{g=1}^{G}\sigma_{a,g}^* 4\pi\frac{\partial B_g^*}{\partial T}}, \tag{7c}$$

$$\chi_g = \frac{\sigma_{a,g}^* 4\pi\frac{\partial B_g^*}{\partial T}}{\sum_{k=1}^{G}\sigma_{a,k}^* 4\pi\frac{\partial B_k^*}{\partial T}}, \tag{7d}$$

$$\xi_g = 4\pi\sigma_{a,g}^* B_g^* + \tau\phi_g^{n-1/2} - v\chi_g\left[\sum_{k=1}^{G}\sigma_{a,k}^* 4\pi B_k^* + \frac{C_v^*}{\Delta t^n}(T^{n-1/2}-T^*)\right], \tag{7e}$$

with the material temperature given by

$$T = T^* + \frac{\sum_{g=1}^{G}\sigma_{a,k}^*(\phi_g - 4\pi B_g^*) + \frac{C_v^*}{\Delta t^n}(T^{n-1/2}-T^*)}{\frac{C_v^*}{\Delta t^n}+\sum_{g=1}^{G}\sigma_{a,g}^* 4\pi\frac{\partial B_g^*}{\partial T}}. \tag{8}$$

Note that a superscript "*" denotes evaluation at T^*, and that the time index $n+\frac{1}{2}$ has been suppressed in Eqs. (6)–(8). The temperatures can be locally obtained once the multigroup diffusion equations are solved. Thus we focus on solving Eq. (6). We generally lag the nonlinearities in the material properties by setting $T^* = T^n$. However, the nonlinearities can be consistently converged if desired by putting a non-linear iteration loop around the linearized solution of these equations.

4. Linear multifrequency-grey acceleration

The traditional method for solving the multigroup diffusion equation is source iteration. Denoting the iteration index by ℓ, this iteration can be represented as follows:

$$-\vec{\nabla}\cdot D_g\vec{\nabla}\phi_g^{\ell+1} + \sigma_{\tau,g}^*\phi_g^{\ell+1} = v\chi_g f^{\ell} + \xi_g, \tag{9}$$

where

$$f^{\ell} = \sum_{k=1}^{G}\sigma_{a,k}^*\phi_k^{\ell}. \tag{10}$$

Lacking knowledge of a better initial guess, f^0 is generally computed using the intensities at the beginning of the time step. Note that if the absorption rate f is known, the process converges in one iteration. Thus the convergence rate of this iteration process is determined by the rate at which errors in the absorption rate are attenuated. When $v \approx 1$, this iteration process can converge arbitrarily slowly. This is known as the strong material–radiation coupling limit and is physically characterized by large absorption and small heat capacity. One can always make this iteration process rapidly convergent by taking a sufficiently small time step, but this generally requires time steps much smaller than the characteristic time scale of the problem. Hence it is not an effective strategy. The LMFGA technique accelerates the convergence of this outer iteration process. In particular, it can be represented as follows:

$$-\vec{\nabla}\cdot D_g\vec{\nabla}\phi_g^{\ell+\frac{1}{2}} + \sigma^*_{\tau,g}\phi_g^{\ell+\frac{1}{2}} = v\chi_g f^{\ell} + \xi_g, \tag{11a}$$

$$-\vec{\nabla}\cdot\langle D\rangle\vec{\nabla}\delta\Phi + [\langle\sigma_a\rangle(1-v)+\tau]\delta\Phi = v\left(f^{\ell+\frac{1}{2}} - f^{\ell}\right), \tag{11b}$$

$$f^{\ell+1} = f^{\ell+\frac{1}{2}} + \langle\sigma_a\rangle\delta\Phi. \tag{11c}$$

where

$$\langle D\rangle = \sum_{g=1}^{G}\frac{\varsigma_g}{3\sigma^*_{a,g}}, \tag{12}$$

$$\langle\sigma_a\rangle = \sum_{g=1}^{G}\sigma^*_{a,g}\varsigma_g, \tag{13}$$

$$\varsigma_g = \frac{\frac{\chi_g}{\sigma^*_{\tau,g}}}{\sum_{k=1}^{G}\frac{\chi_k}{\sigma^*_{\tau,k}}}. \tag{14}$$

The principle behind this method is fully explained in Ref. [1]. A brief description follows. A one-dimensional Fourier analysis for a homogeneous infinite-medium can be performed that is based upon a decomposition of the errors in f into Fourier modes, each having a spatial dependence of the form $\exp(\mathrm{j}\lambda x)$, where $\mathrm{j} = \sqrt{-1}$, λ is any real number, and x is the spatial variable. Note that small values of λ correspond to slowly varying errors while large values of λ correspond to rapidly varying errors. The analysis indicates that the outer source iteration at step $\ell+\frac{1}{2}$, defined by Eq. (11a), strongly attenuates error modes of f with large values of $|\lambda|$ (rapidly varying errors) and weakly attenuates those with small values of $|\lambda|$ (slowly varying errors). An exact multigroup diffusion equation can be written for the additive errors in the multigroup intensities at iteration step $\ell+\frac{1}{2}$, and the exact additive error in the absorption rate can be directly calculated from these intensity errors. However, this exact multigroup error equation is no easier to solve than the original equation, so solving it is not a viable strategy. Instead one substitutes a low-rank approximation for the exact equation that is reasonably easy to solve with the intent that it be accurate for modes with $\lambda \approx 0$. The Fourier analysis shows that the errors in the *intensity* take on a specific energy shape in the limit as $\lambda \to 0$. By assuming that the solution to the exact error equation is the product of this energy shape modulated by a purely space-dependent function, one can derive a grey equation for the error in the energy-integrated intensity that is exact in the limit as $\lambda \to 0$. This grey diffusion equation for the additive error in the energy-integrated intensity is given in Eq. (11b), with the normalized energy shape function given by Eq. (14). Once the error in the integrated intensity has been estimated from a solution of the approximate grey equation, the corresponding error in the absorption rate is calculated from the error in the energy-integrated intensity and added to the absorption rate iterate at step $\ell+\frac{1}{2}$, resulting in the "accelerated" absorption rate at step $\ell+1$. This entire process is performed in Eq. (11c). Thus the grey diffusion equation enables the errors in the absorption rate to be perfectly attenuated in the limit as $\lambda \to 0$. Nonetheless, one must be concerned about the effect of the grey diffusion approximation on the rapidly varying error components. The grey diffusion equation is of low rank and thus cannot be accurate for all error components. If the grey diffusion equation were to sufficiently overestimate the rapidly varying errors, it would cause the accelerated iteration process to diverge. Fortunately, the grey diffusion approximation grossly *underestimates* the rapidly varying errors and thus does no harm. The overall result for the idealized homo-

geneous infinite-medium problem is a rapidly convergent iteration process that almost perfectly attenuates both very slowly and very rapidly varying errors, and strongly attenuates errors with an intermediate variation.

Computational experience indicates that in real one-dimensional calculations, this method is unconditionally effective. However, as previously noted, it has recently been found that this method can significantly degrade in multidimensional problems with large discontinuities in material properties. The multifrequency-grey acceleration technique can be interpreted as a two-grid method with the grey-diffusion operator playing the role of the ''coarse-grid'' operator. It is well known within the computational mathematics community that multigrid methods are far more robust when recast as preconditioned Krylov methods. Indeed, as previously noted, the DSA scheme for the transport equation (a two-grid diffusion-based method) becomes ineffective in multidimensional problems with large discontinuities in material properties, but when recast as a preconditioned Krylov method has been has been observed to remain effective under all conditions [3]. As previously noted, this is our motivation for recasting the LMFGA method as a preconditioned Krylov method.

5. A preconditioned Krylov method

In this section we derive a preconditioned Krylov method for the multigroup radiative diffusion equations. We first derive an equation for the absorption rate f. The action of the operator associated with this equation requires the same set of independent one group diffusion solutions required by a source iteration. We then derive a diffusion-based preconditioner that is closely related to the diffusion operator used in the LMFGA method to estimate the intensity errors at step $\ell + \frac{1}{2}$.

5.1. The absorption rate equation

We choose not to directly solve Eq. (6) with a Krylov method. Rather, we solve an equation that has the absorption rate f as its unknown. Once the absorption rate has been calculated, the multigroup intensities can be obtained by solving an independent set of one group diffusion equations:

$$-\vec{\nabla}\cdot D_g\vec{\nabla}\phi_g + \sigma^*_{\tau,g}\phi_g = v\chi_g f + \xi_g, \quad g = 1, G. \tag{15}$$

We assume that an efficient method exists for solving such a system. Obviously, this assumption is also required for the LMFGA method because the basic unaccelerated iteration scheme requires the solution of an equivalent system as illustrated by Eq. (11a).

We now proceed with the derivation of an equation for the absorption rate. Let us first rewrite Eq. (6) in operator form as follows:

$$\mathbf{A}_g\phi_g = v\chi_g f + \xi_g, \quad g = 1, G, \tag{16}$$

where

$$\mathbf{A}_g \equiv -\vec{\nabla}\cdot D_g\vec{\nabla} + \sigma^*_{\tau,g}. \tag{16a}$$

Solving Eq. (16) for ϕ_g, we get

$$\phi_g = \mathbf{A}_g^{-1}[v\chi_g f + \xi_g], \quad g = 1, G. \tag{17}$$

Multiplying Eq. (17) by $\sigma^*_{a,g}$ and summing over all groups, we obtain:

$$f = \sum_{g=1}^{G} \sigma^*_{a,g}\mathbf{A}_g^{-1}[v\chi_g f + \xi_g]. \tag{18}$$

Moving all terms containing f to the left side of Eq. (18), we obtain the desired equation for the absorption rate:

$$\mathbf{B}f = \sum_{g=1}^{G} \sigma^*_{a,g}\mathbf{A}_g^{-1}\xi_g, \tag{19}$$

where

$$\mathbf{B} = \left[\mathbf{I} - \sum_{g=1}^{G} \sigma_{a,g}^{*} \mathbf{A}_g^{-1} \nu \chi_g\right], \tag{19a}$$

and $\mathbf{I}$ denotes the identity operator. Eq. (19) is solved via a Krylov method rather than Eq. (6) because this strategy can require far less memory. Assuming one spatial unknown per group per cell, Eq. (6) has a solution vector dimension equal to the number of spatial cells times the number of groups; but Eq. (19) has a solution vector dimension equal to the number of spatial cells. Furthermore, Eq. (19) has a much more compact spectrum than Eq. (6). In particular, under the assumption of an infinite homogeneous medium, it can be shown that the eigenvalues of the operator associated with Eq. (18) are real, positive, and lie in the open interval (0,1). For instance, let us assume an infinite homogeneous medium and a one-dimensional Fourier dependence for f:

$$f(x) = f_0 \exp \mathrm{j}\lambda x. \tag{20}$$

Applying the operator $\mathbf{B}$ to f, we obtain

$$\mathbf{B}f = \left[\mathbf{I} - \sum_{g=1}^{G} \frac{\sigma_{a,g}^{*} \nu \chi_g}{D_g \lambda^2 + \sigma_{a,g}^{*} + \tau}\right] f = \omega(\lambda) f, \tag{21}$$

where

$$\omega(\lambda) = \sum_{g=1}^{G} \chi_g \frac{D_g \lambda^2 + \sigma_{a,g}^{*}(1-\nu) + \tau}{D_g \lambda^2 + \sigma_{a,g}^{*} + \tau}. \tag{22}$$

Note that f is an eigenfunction of $\mathbf{B}$ with eigenvalue $\omega(\lambda)$. Assuming a nonzero and bounded time step, it can be seen from Eq. (7c) that ν has a greatest lower bound of zero and a least upper bound of unity. Under the same assumptions it can be seen from Eq. (7d) that the sum over g of χ_g is unity. Considering this information, it is not difficult to see that $\omega(\lambda)$ has a greatest lower bound of zero and a least upper bound of unity. Small eigenvalues correspond to low frequency modes, while eigenvalues near unity correspond to high frequency modes. This spectrum is much more compact than that associated with Eq. (6), which not only has real and positive eigenvalues near zero, but also has real and positive eigenvalues of arbitrarily large magnitude. Even if we assume a symmetric positive definite (SPD) discretization for each one group diffusion operator associated with $\mathbf{B}$, it is clear that $\mathbf{B}$ is not self-adjoint with space-dependent material properties. It is desirable to use the conjugate gradient Krylov method [4] whenever possible. However, this method is limited to symmetric positive definite systems. Thus we cannot use the CG method to solve Eq. (19) for the absorption rate. However, this is not necessarily a significant disadvantage, because the CG method can still be used to perform the group-dependent diffusion solutions required to form the action of $\mathbf{B}$, and to calculate the angularly integrated intensities using Eq. (15) after the absorption rate has been calculated. Note from Eqs. (9) and (19) that the steps required to form the action of the operator associated with Eq. (19) are identical to that required to perform a source iteration.

5.2. The preconditioner for multigroup diffusion

We next derive the preconditioner for the operator $\mathbf{B}$. The first step in the derivation is to decompose the solution to Eq. (6) into two components, $\phi_g^{(0)}$ and $\phi_g^{(1)}$, where

$$\phi_g = \phi_g^{(0)} + \phi_g^{(1)}, \quad g = 1, G. \tag{23}$$

The equation satisfied by $\phi_g^{(0)}$ is

$$\mathbf{A}_g \phi_g^{(0)} = \xi_g, \quad g = 1, G, \tag{24}$$

and the equation satisfied by $\phi_g^{(1)}$ is

$$\mathbf{A}_g\phi_g^{(1)} - v\chi_g \sum_{k=1}^{G} \sigma_{a,k}^* \phi_k^{(1)} = v\chi_g \sum_{k=1}^{G} \sigma_{a,k}^* \phi_k^{(0)}, \quad g = 1, G. \tag{25}$$

Solving Eq. (24) for $\phi_g^{(0)}$, we obtain

$$\phi_g^{(0)} = \mathbf{A}_g^{-1}\xi_g, \quad g = 1, G. \tag{26}$$

Substituting from Eq. (26) into Eq. (25), we get

$$\mathbf{A}_g\phi_g^{(1)} - v\chi_g \sum_{k=1}^{G} \sigma_{a,k}^* \phi_k^{(1)} = v\chi_g \sum_{k=1}^{G} \sigma_{a,k}^* \mathbf{A}_k^{-1}\xi_k, \quad g = 1, G. \tag{27}$$

We next define a grey approximation to Eq. (27) that is accurate when ϕ_g is slowly varying in space. Applying $\mathbf{A}_g^{-1}$ to Eq. (27) from the left, we obtain

$$\phi_g^{(1)} = \mathbf{A}_g^{-1} v\chi_g \sum_{k=1}^{G} \sigma_{a,k}^* [\phi_k^{(1)} + \mathbf{A}_k^{-1}\xi_k], \quad g = 1, G. \tag{28}$$

It is easily seen from Eq. (28) that $\phi_g^{(1)}$ has a group-dependent shape proportional to $\mathbf{A}_g^{-1}\chi_g$. Recalling Eqs. (16a) and (20), we find that

$$\mathbf{A}_g^{-1}\chi_g \equiv \frac{\chi_g}{D_g\lambda^2 + \sigma_{\tau,g}^*}. \tag{29}$$

Thus in the limit as $\lambda \to 0$, the normalized shape of $\phi_g^{(1)}$ is given by

$$\varsigma_g = \frac{\frac{\chi_g}{\sigma_{\tau,g}^*}}{\sum_{k=1}^{G} \frac{\chi_k}{\sigma_{\tau,k}^*}}. \tag{30}$$

Note from Eq. (14) that this shape is identical to that of the weakly attenuated intensity errors associated with the source iteration process. We next assume this shape for $\phi_g^{(1)}$. In particular, we assume that

$$\phi_g^{(1)} = \varsigma_g \Phi^{(1)}, \tag{31}$$

where $\Phi^{(1)}$ is a space-dependent modulation function corresponding to the energy-integrated intensity. Substituting from Eq. (31) into Eq. (27), and summing over all groups, we obtain the following grey drift-diffusion approximation:

$$-\vec{\nabla}\cdot\langle D\rangle\vec{\nabla}\Phi^{(1)} - \vec{\nabla}\cdot\langle\vec{D}\rangle\Phi^{(1)} + [\langle\sigma_a\rangle(1-v) + \tau]\Phi^{(1)} = v\sum_{k=1}^{G} \sigma_{a,k}^* \mathbf{A}_k^{-1}\xi_k, \tag{32}$$

where

$$\langle D\rangle = \sum_{g=1}^{G} D_g\varsigma_g, \tag{32a}$$

$$\langle\vec{D}\rangle = \sum_{g=1}^{G} D_g\vec{\nabla}\varsigma_g, \tag{32b}$$

and

$$\langle\sigma_a\rangle = \sum_{g=1}^{G} \sigma_{a,g}^*\varsigma_g, \tag{32c}$$

For reasons that are explained later, we drop the drift term from Eq. (32) to obtain

$$\mathbf{H}\Phi^{(1)} = v\sum_{k=1}^{G} \sigma_{a,k}^* \mathbf{A}_k^{-1}\xi_k, \tag{33}$$

where

$$\mathbf{H} \equiv -\vec{\nabla} \cdot \langle D \rangle \vec{\nabla} + [\langle \sigma_a \rangle (1 - v) + \tau]. \tag{33a}$$

Note from Eqs. (11b) and (33a) that this grey diffusion operator is identical to that associated with the LMFGA method. In early 1D versions of the multifrequency-grey method, the drift term was retained because the resulting drift-diffusion equation could be easily solved using direct techniques. However, when the method was applied in multidimensions, the drift term was dropped because doing so did not seem to significantly affect the convergence rate in 1D calculations and iterative solution techniques for the diffusion equation were significantly less costly than those for the drift-diffusion equation. This remains true, though to perhaps a lesser extent. Hence we follow what is now standard practice and drop the drift term. Solving Eq. (33) for $\Phi^{(1)}$, we obtain

$$\Phi^{(1)} = \mathbf{H}^{-1} v \sum_{k=1}^{G} \sigma_{a,k}^{*} \mathbf{A}_k^{-1} \xi_k. \tag{34}$$

Recognizing that the grey absorption rate is given by $\langle \sigma_a \rangle \Phi$, we manipulate Eq. (34) to yield the first component of the absorption rate:

$$f^{(1)} = \langle \sigma_a \rangle \mathbf{H}^{-1} v \sum_{k=1}^{G} \sigma_{a,k}^{*} \mathbf{A}_k^{-1} \xi_k. \tag{35}$$

Following Eq. (24), the zero'th component of the absorption rate is rigorously given as follows:

$$f^{(0)} = \sum_{g=1}^{G} \sigma_{a,g}^{*} \mathbf{A}_g^{-1} \xi_g. \tag{36}$$

Adding Eqs. (35) and (36), we obtain the total absorption rate with the zero'th component computed exactly, and the first component computed approximately via the grey approximation:

$$f = (\mathbf{I} + \langle \sigma_a \rangle \mathbf{H}^{-1} v) \sum_{g=1}^{G} \sigma_{a,g}^{*} \mathbf{A}_g^{-1} \xi_g. \tag{37}$$

Let us next solve Eq. (19) for the absorption rate:

$$f = \mathbf{B}^{-1} \sum_{g=1}^{G} \sigma_{a,g}^{*} \mathbf{A}_g^{-1} \xi_g. \tag{38}$$

Comparing Eqs. (37) and (38), it is clear that

$$(\mathbf{I} + \langle \sigma_a \rangle \mathbf{H}^{-1} v) \approx \mathbf{B}^{-1}. \tag{39}$$

Thus our preconditioner is

$$\mathbf{C} \equiv (\mathbf{I} + \langle \sigma_a \rangle \mathbf{H}^{-1} v), \tag{40}$$

and our preconditioned equation is

$$\mathbf{C}\mathbf{B} f = \mathbf{C} \sum_{g=1}^{G} \sigma_{a,g}^{*} \mathbf{A}_g^{-1} \xi_g. \tag{41}$$

Assuming an infinite homogeneous medium and the Fourier spatial dependence assumed in Eq. (20), it is easily shown that

$$\mathbf{C}\mathbf{B} = \mathbf{I} + \mathrm{O}(\lambda^2). \tag{42}$$

Remembering that small values of λ correspond to the eigenfunctions of $\mathbf{B}$ with the smallest eigenvalues, it follows that $\mathbf{C}$ will move the smallest eigenvalues away from zero to essentially unity. We are concerned with the eigenfunctions of $\mathbf{B}$ having the largest eigenvalues since the preconditioner cannot be exactly equal to the

inverse of **B** for all of its eigenfunctions. In analogy with the behavior of the grey diffusion approximation in the LMFGA method, the inverse grey diffusion operator in **C** grossly underestimates the eigenvalues of **B** that are large. They are in fact so underestimated that the identity term in **C** dominates the diffusion term with the result that the preconditioner is effectively just the identity. Thus the overall effect of the preconditioner is ideal: it moves the smallest eigenvalues essentially to unity and leaves the largest eigenvalues alone. While this is guaranteed only for an infinite homogeneous medium, it nonetheless suggests that **C** will be a very effective preconditioner for the general case. The properties of the grey diffusion operator that cause degradation in the classic LMFGA method can possibly generate some eigenvalues in the preconditioned operator that are greater than the maximum value of unity associated with **B**. However, we conjecture that this effect will have a small adverse effect on the convergence of **CB** relative to **B**. The movement of the smallest eigenvalues away from zero will likely have much greater impact than any increase in the largest eigenvalues, resulting in a preconditioner that is very effective overall. We later provide computational results that give insight into the spectrum of the preconditioned operator.

Even if we assume an SPD diffusion discretization, **C** is not SPD with spatially dependent material properties. As previously noted, **B** is also not SPD with spatially dependent material properties. Thus, unless there is some symmetrization process of which we are currently not aware, the preconditioned system **CB** will not be SPD, which implies that Eq. (41) cannot be solved using the CG method. Nonetheless, all of the diffusion solutions associated with forming the action of **CB** can be performed using the CG method.

Since the action of the operator **B** requires the inversion of an independent set of one-group diffusion equations, a nested Krylov strategy is required to solve Eq. (41). By this we mean that each Krylov iteration for Eq. (41) requires the solution of an independent set of one-group diffusion equations, each of which should ideally be solved via a preconditioned CG method. To achieve the highest efficiency for nested calculations, while still converging to the correct solution, the tolerance for the "inner" CG iterations could be relaxed as the "outer" Krylov iteration proceeds according to the strategy presented in Refs. [6,7].

6. Relating the LMFGA and LMFGK methods

In this section we demonstrate a simple relationship between the LMFGA and LMFGK methods. We begin by re-expressing the LMFGA method in terms of operators previously defined for the LMFGK method. In particular, we first solve Eq. (11a) for $\phi_g^{\ell+\frac{1}{2}}$, and then use that expression to solve for $f^{\ell+\frac{1}{2}}$:

$$f^{\ell+\frac{1}{2}} = (\mathbf{I}-\mathbf{B})f^{\ell} + \sum_{g=1}^{G} \sigma_{a,g}^{*}\mathbf{A}_g^{-1}\xi_g, \tag{43}$$

where $\mathbf{A}_g$ and **B** are defined by Eqs. (16a) and (19a), respectively. Next we substitute from Eq. (43) into Eq. (11b) to obtain

$$\delta\Phi = \mathbf{H}^{-1}v\left(\sum_{g=1}^{G} \sigma_{a,g}^{*}\mathbf{A}_g^{-1}\xi_g - \mathbf{B}f^{\ell}\right). \tag{44}$$

Finally, we substitute from Eqs. (43) and (44) into Eq. (11c), to obtain a single expression that defines a complete LMFGA iteration for the absorption rate:

$$f^{\ell+1} = (\mathbf{I}-\mathbf{CB})f^{\ell} + \mathbf{C}\sum_{g=1}^{G} \sigma_{a,g}^{*}\mathbf{A}_g^{-1}\xi_g, \tag{45}$$

where **C** is defined by Eq. (40). Comparing Eq. (45) with Eq. (41), one finds that the LMFGA method is simply Richardson iteration [4] applied to the preconditioned equation that is solved in the LMFGK method. Richardson iteration applied to a general linear system of the form given in Eq. (1) is

$$\vec{x}^{\ell+1} = (\mathbf{I}-\mathbf{M})\vec{x}^{\ell} + \vec{b}. \tag{46}$$

Thus, as one would expect, the two methods are closely and simply related. They each represent a method for solving the preconditioned equation given by Eq. (41). In the case of the LMFGA method, this equation is

solved via Richardson iteration; and in the case of the LFMGK method, this equation is solved via a standard Krylov method.

7. The discretized equations

To test our LMFK method, we use Palmer's node-centered polygonal-mesh discretization of the diffusion operator in two-dimensional r–z geometry [8] in conjunction with a discontinuous representation for the material temperatures. The radiation intensities are located at nodes, but the temperatures are located at "corners". A corner is uniquely associated with both a node and a cell that subtends that node. Every corner coincides with a node, but there is a separate corner for each cell that subtends the node. Corners are illustrated for a polygonal mesh in Fig. 1. Thus the material temperatures are also located at the nodes, but there is a separate node temperature for each cell that subtends the node. Consequently, material temperature is said to be discontinuous at each node. Since there is only one set of multigroup radiation intensities per node, the radiation intensity treatment is said to be continuous. A continuous temperature treatment would simply have one temperature at each node. Since material properties are cell-centered, the discontinuous temperature treatment yields separate temperatures in each material, which is highly desirable in radiation-hydrodynamics calculations. The hybrid node-centered discretization that we use with continuous radiation intensities and discontinuous temperatures is simpler and easier to solve than a fully cell-centered discretization, but it is nonetheless compatible with standard cell-centered and staggered-mesh hydrodynamics discretizations. Although we do not consider radiation-hydrodynamics calculations in this study, such calculations are a major application for multigroup thermal radiation diffusion algorithms. Thus it is appropriate to test our LMFGK method using a multigroup thermal radiation diffusion discretization that is practical for radiation-hydrodynamics calculations.

Before describing our radiation/temperature discretization for general polygonal meshes, we first consider a simple 1D slab-geometry mesh. The center, node, and corner indexing for such a mesh is shown in Fig. 2. Let h_i denote the width of cell i. The effective width associated with corner $i, i+\frac{1}{2}$ is

$$h_{i,i+\frac{1}{2}} = \frac{h_i}{2}, \tag{47}$$

and the effective width associated with node $i+\frac{1}{2}$ is the sum of the associated corner widths

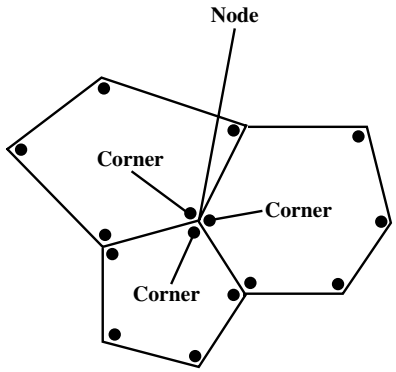

Fig. 1. Illustration of corners and nodes. Each corner coincides with a node but is uniquely associated with cell that subtends that node.

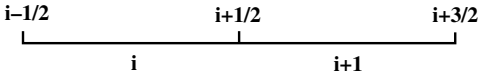

Fig. 2. Center, node, and corner indexing in 1D slab geometry. The cell centers carry integer indices, the nodes carry half-integer indices, and corners carry both cell and node indices.

$$h_{i+\frac{1}{2}} = h_{i+1,i+\frac{1}{2}} + h_{i,i+\frac{1}{2}} = \frac{h_{i+1}}{2} + \frac{h_i}{2}. \tag{48}$$

Many of our discrete equations contain node-centered quantities obtained by volume-weighted averaging of corner-centered quantities. The normalized weight associated with corner $i, i+\frac{1}{2}$ is

$$w_{i,i+\frac{1}{2}} = \frac{h_{i,i+\frac{1}{2}}}{h_{i+\frac{1}{2}}} = \frac{h_i}{h_{i+1} + h_i}. \tag{49}$$

It is often convenient for our purposes to collectively apply an index to a group of terms enclosed within parentheses, brackets, etc. This implies that all the enclosed quantities individually carry that index unless otherwise indicated.

On the mesh interior, the discrete analog of Eq. (16) for the intensity in group g at node $i+\frac{1}{2}$ can be expressed as follows:

$$(\mathbf{A}_g \phi_g)_{i+\frac{1}{2}} = \left[(v\chi_g f + \xi_g)_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + (v\chi_g f + \xi_g)_{i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}} \right] h_{i+\frac{1}{2}}, \tag{50}$$

where

$$(\mathbf{A}_g \phi_g)_{i+\frac{1}{2}} = -\frac{D_{g,i+1}}{h_{i+1}} \left(\phi_{g,i+\frac{3}{2}} - \phi_{g,i+\frac{1}{2}} \right) + \frac{D_{g,i,}}{h_i} \left(\phi_{g,i+\frac{1}{2}} - \phi_{g,i-\frac{1}{2}} \right) + \sigma^*_{\tau,g,i+\frac{1}{2}} \phi_{g,i+\frac{1}{2}} h_{i+\frac{1}{2}}, \tag{50a}$$

$$f_{i,i+\frac{1}{2}} = \sum_{k=1}^{G} \sigma^*_{a,k,i+1,i+\frac{1}{2}} \phi_{g,i+\frac{1}{2}}, \tag{50b}$$

$$\sigma^*_{\tau,g,i+\frac{1}{2}} = \sigma^*_{\tau,g,i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + \sigma^*_{\tau,g,i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}. \tag{50c}$$

The diffusion coefficient $D_{g,i,}$, which appears in Eq. (50a), is cell-centered and evaluated for the material in cell i at the cell-centered temperature,

$$T^*_i = \frac{\left[\left(T^*_{i,i+\frac{1}{2}} \right)^4 + \left(T^*_{i,i-\frac{1}{2}} \right)^4 \right]^{\frac{1}{4}}}{2}. \tag{50d}$$

The cross-section, $\sigma^*_{a,g,i,i+\frac{1}{2}}$, which appears in Eq. (50b), is corner-centered and evaluated for the material in cell i at the corner temperature, $T^*_{i,i+\frac{1}{2}}$. The cross-section, $\sigma^*_{\tau,g,i,i+\frac{1}{2}}$, which appears in Eq. (50c), is constructed from $\sigma^*_{a,g,i,i+\frac{1}{2}}$ in accordance with Eqs. (7a) and (7b). The quantities, $v_{i,i+\frac{1}{2}}$ and $\chi_{g,i,i+\frac{1}{2}}$, which appear in Eq. (50), are corner-centered and constructed from other corner-centered quantities in accordance with Eqs. (7c) and (7d). The quantity, $\xi_{g,i,i+\frac{1}{2}}$, which appears in Eq. (50), is corner-centered and constructed from other corner-centered quantities and one node-centered quantity in accordance with Eq. (7e). This node-centered quantity corresponds to $\tau\phi^{n-\frac{1}{2}}_{g,i+\frac{1}{2}}$.

The diffusion operator appearing in Eq. (50) and defined by Eq. (50a) is node-centered, but the node-centered cross-section appearing in that operator and defined in Eq. (50c) is a volume-weighted average of corner-centered cross-sections. Similarly, the right side of Eq. (50) is node-centered but formed from a volume-weighted average of corner-centered quantities.

The discrete analog of Eq. (8) for the temperature at corner $i, i+\frac{1}{2}$ is given by

$$T_{i,i+\frac{1}{2}} = T^*_{i,i+\frac{1}{2}} + \frac{\sum_{g=1}^{G} \sigma^*_{a,g,i,i+\frac{1}{2}} \left(\phi_{g,i+\frac{1}{2}} - 4\pi B^*_{g,i,i+\frac{1}{2}} \right) + \frac{C^*_{v,i,i+\frac{1}{2}}}{\Delta t^n} \left(T^{n-\frac{1}{2}}_{i,i+\frac{1}{2}} - T^*_{i,i+\frac{1}{2}} \right)}{\frac{C^*_{v,i,i+\frac{1}{2}}}{\Delta t^n} + \sum_{g=1}^{G} \sigma^*_{a,g,i,i+\frac{1}{2}} 4\pi \frac{\partial B^*}{\partial T}_{g,i,i+\frac{1}{2}}}, \tag{51}$$

where $C^*_{v,i,i+\frac{1}{2}}$ denotes the heat capacity for the material in cell i evaluated at the corner-centered temperature, $T_{i,i+\frac{1}{2}}$, $B^*_{g,i,i+\frac{1}{2}}$ denotes the Planck function evaluated at the corner-centered temperature, $T_{i,i+\frac{1}{2}}$, and $\frac{\partial B^*}{\partial T}_{g,i,i+\frac{1}{2}}$ denotes the temperature-derivative of the Planck function evaluated at the corner-centered temperature, $T_{i,i+\frac{1}{2}}$.

Symbolically inverting the discrete diffusion operator in Eq. (50), we obtain a discrete analog of Eq. (17):

$$\phi_{g,i+\frac{1}{2}} = \mathbf{A}_g^{-1}\left[(v\chi_g f + \xi_g)_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + (v\chi_g f + \xi_g)_{i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}\right] h_{i+\frac{1}{2}}. \tag{52}$$

Multiplying Eq. (52) by $\sigma^*_{a,g,i,i+\frac{1}{2}}$ and summing over all groups, we obtain the discrete analog of Eq. (18):

$$f_{i,i+\frac{1}{2}} = \sum_{g=1}^{G} \sigma^*_{a,g,i,i+\frac{1}{2}} \mathbf{A}_g^{-1}\left[(v\chi_g f + \xi_g)_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + (v\chi_g f + \xi_g)_{i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}\right] h_{i+\frac{1}{2}}, \tag{53}$$

Note that this is an equation for the absorption rate at corner $i, i+\frac{1}{2}$. Thus the absorption-rate equation is corner-centered even though the radiation intensity equation is node-centered. This occurs because the temperatures are corner-centered rather than node-centered.

Assuming a discrete decomposition analogous to that given in Eqs. (23) and (24), the discrete analog of Eq. (28) is

$$\phi^{(1)}_{g,i+\frac{1}{2}} = \mathbf{A}_g^{-1}\left\{\left[v\chi_g\left(f^{(1)} + \sum_{k=1}^{G} \sigma^*_{a,k}\mathbf{A}_k^{-1}\xi_k\right)\right]_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + \left[v\chi_g\left(f^{(1)} + \sum_{k=1}^{G} \sigma^*_{a,k}\mathbf{A}_k^{-1}\xi_k\right)\right]_{i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}\right\} h_{i+\frac{1}{2}}, \tag{54}$$

where

$$f^{(1)} = \sum_{k=1}^{G} \sigma^*_{a,k}\phi^{(1)}_k. \tag{54a}$$

A difficulty arises when trying to obtain an energy shape function for the grey diffusion equation from Eq. (54) because the solution to Eq. (54) does not assume a unique shape when the solution is spatially constant. In particular, setting the spatial derivative terms to zero in $\mathbf{A}_g^{-1}$, Eq. (54) becomes

$$\phi^{(1)}_{g,i+\frac{1}{2}} = \frac{1}{\sigma^*_{\tau,g,i+\frac{1}{2}}}\left\{\left[v\chi_g\left(f^{(1)} + \sum_{k=1}^{G} \frac{\sigma^*_{a,k}\xi_k}{\sigma^*_{\tau,k,i+\frac{1}{2}}}\right)\right]_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + \left[v\chi_g\left(f^{(1)} + \sum_{k=1}^{G} \frac{\sigma^*_{a,k}\xi_k}{\sigma^*_{\tau,k,i+\frac{1}{2}}}\right)\right]_{i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}\right\}. \tag{55}$$

There is a separate contribution to $\phi^{(1)}_{g,i+\frac{1}{2}}$ in Eq. (55) from each corner associated with node $i+\frac{1}{2}$, and each of these contributions has a unique shape. Thus the overall shape of $\phi^{(1)}_{g,i+\frac{1}{2}}$ depends upon the relative magnitudes of these contributions. The solution would have a unique shape if the temperatures were node-centered. Even though a unique shape is not assumed by the solution, a grey equation nonetheless can be obtained given node-centered shape functions. Our procedure for obtaining shape functions is not unique. We simply made a reasonable choice and found that it worked well.

To obtain node-centered shape functions, we first define unnormalized corner-centered shape functions. In particular, the unnormalized shape function at corner $i, i+\frac{1}{2}$ is defined as follows:

$$\alpha_{g,i,i+\frac{1}{2}} = \frac{v_{i,i+\frac{1}{2}}\chi_{g,i,i+\frac{1}{2}}}{\sigma^*_{\tau,g,i+\frac{1}{2}}}. \tag{56}$$

The unnormalized shape function at node $i+\frac{1}{2}$ is obtained via a volume-weighted average of the unnormalized corner shape functions:

$$\alpha_{g,i+\frac{1}{2}} = \alpha_{g,i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + \alpha_{g,i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}. \tag{57}$$

Thus the normalized shape function at node $i+\frac{1}{2}$ is given by

$$\varsigma_{i+\frac{1}{2}} = \frac{\alpha_{g,i+\frac{1}{2}}}{\sum_{k=1}^{G} \alpha_{k,i+\frac{1}{2}}}. \tag{58}$$

The generation of the discrete grey diffusion equation begins by multiplying Eq. (54) from the left with $\mathbf{A}_g$:

$$(\mathbf{A}_g\phi^{(1)})_{g,i+\frac{1}{2}} - \left[(v\chi_g f^{(1)})_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + (v\chi_g f^{(1)})_{i,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}}\right] h_{i+\frac{1}{2}}$$
$$= \left[\left(v\chi_g \sum_{k=1}^{G} \sigma^*_{a,k}\mathbf{A}_k^{-1}\xi_k\right)_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + \left(v\chi_g \sum_{k=1}^{G} \sigma^*_{a,k}\mathbf{A}_k^{-1}\xi_k\right)_{i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}\right] h_{i+\frac{1}{2}}. \tag{59}$$

We directly use the node-centered shape functions for the absorptive components of $\mathbf{A}$ and for the second term in brackets on the left side of Eq. (59). The diffusion components of $\mathbf{A}$ require some care to avoid an effective drift term. More specifically, each difference of intensities associated with a single cell-centered diffusion coefficient must be averaged with a single shape function. These cell-centered shape functions are obtained simply by linearly averaging the two node-centered shape functions associated with each cell:

$$\varsigma_{g,i} = \frac{1}{2}\left(\varsigma_{i-\frac{1}{2}} + \varsigma_{i+\frac{1}{2}}\right). \tag{60}$$

The following discrete grey diffusion equation is obtained from the averaging process:

$$(\mathbf{H}\Phi^{(1)})_{i+\frac{1}{2}} = \left[\left(v\sum_{k=1}^{G} \sigma^*_{a,k}\mathbf{A}_k^{-1}\xi_k\right)_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + \left(v\sum_{k=1}^{G} \sigma^*_{a,k}\mathbf{A}_k^{-1}\xi_k\right)_{i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}\right] h_{i+\frac{1}{2}}, \tag{61}$$

where

$$(\mathbf{H}\Phi^{(1)})_{i+\frac{1}{2}} = -\frac{\langle D_{i+1}\rangle}{h_{i+1}}\left(\Phi^{(1)}_{i+\frac{3}{2}} - \Phi^{(1)}_{i+\frac{1}{2}}\right) + \frac{\langle D\rangle_i}{h_i}\left(\Phi^{(1)}_{i+\frac{1}{2}} - \Phi^{(1)}_{i-\frac{1}{2}}\right) + \left\{[\langle\sigma_a\rangle(1-v)]_{i+\frac{1}{2}} + \tau\right\}\Phi^{(1)}_{i+\frac{1}{2}} h_{i+\frac{1}{2}}, \tag{61a}$$

$$\langle D\rangle_i = \sum_{g=1}^{G} D_{g,i}\varsigma_{g,i}, \tag{61b}$$

$$[\langle\sigma_a\rangle(1-v)]_{i+\frac{1}{2}} = [\langle\sigma_a\rangle(1-v)]_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + [\langle\sigma_a\rangle(1-v)]_{i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}, \tag{61c}$$

$$\langle\sigma_a\rangle_{i,i+\frac{1}{2}} = \sum_{g=1}^{G} \sigma^*_{a,g,i,i+\frac{1}{2}}\varsigma_{g,i+\frac{1}{2}}. \tag{61d}$$

Like the discrete multigroup diffusion operator defined by Eq. (50a), the discrete grey diffusion equation defined by Eq. (61a) is node-centered. The grey equation for the absorption rate is obtained by first symbolically inverting the operator on the left side of Eq. (61a), and then multiplying that equation by $\langle\sigma_a\rangle_{i,i+\frac{1}{2}}$:

$$f^{(1)}_{i,i+\frac{1}{2}} = \langle\sigma_a\rangle_{i,i+\frac{1}{2}}\mathbf{H}^{-1}\left[\left(v\sum_{k=1}^{G} \sigma^*_{a,k}\mathbf{A}_k^{-1}\xi_k\right)_{i+1,i+\frac{1}{2}} w_{i+1,i+\frac{1}{2}} + \left(v\sum_{k=1}^{G} \sigma^*_{a,k}\mathbf{A}_k^{-1}\xi_k\right)_{i,i+\frac{1}{2}} w_{i,i+\frac{1}{2}}\right] h_{i+\frac{1}{2}}. \tag{62}$$

Note that the grey equation for the absorption rate given in Eq. (62), like the multigroup equation for the absorption rate given in Eq. (53), is corner-centered.

The procedure for generating the polygonal-mesh equations is almost completely analogous to that used to generate the slab-geometry equations. Only a few aspects of the procedure require explicit discussion. As previously noted, the polygonal diffusion discretization is node-centered [8]. As in the slab-geometry case, all terms relating to absorption and emission are treated using a one-point or diagonal approximation. Let us consider a node on the mesh interior, and denote it as "node a". The diffusion stencil couples the intensity at node a to each intensity located at a node that shares a polygon with node a. This is illustrated in Fig. 3. Each polygon is decomposed into triangles as shown in Fig. 4. Each triangle is associated with three intensities: two nodal intensities and an intensity at the center of the polygon. The center intensity is not an independent unknown but rather is an average of the nodal intensities associated with the polygon. Each triangle is also associated with a gradient computed from the three triangular intensities. While there is only one material in each polygon, there is a separate diffusion coefficient for each triangle. This diffusion coefficient is analogous to the cell-centered coefficient in slab geometry, but is not really "centered" in the triangle. It is

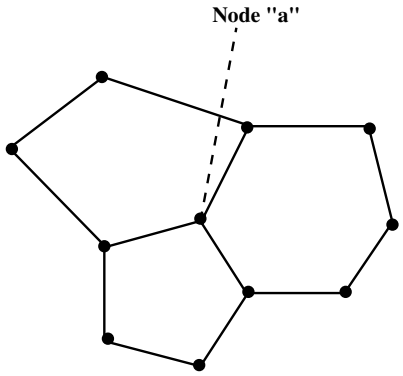

Fig. 3. The polygonal diffusion stencil. All of the nodes associated with the diffusion stencil for node "a" are shown. Every node that shares a polygon with node "a" is in the stencil.

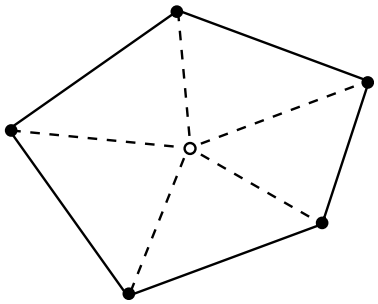

Fig. 4. The triangular decomposition associated with each polygon.

evaluated in analogy with Eq. (50d) using the two corner temperatures associated with each triangle. Thus it can be thought of as being located at the center of the outer edge of the triangle. The shape functions associated with these two nodes are also averaged in analogy with Eq. (60) to obtain a unique shape function for each triangle, and this shape function is used to generate a grey diffusion coefficient. While there are only two corners per node in slab geometry, there can be an arbitrary number of corners per node on a polygonal mesh. Corner-centered quantities are averaged onto the nodes using the same volume-weighted technique used in slabs. Each corner is uniquely associated with two triangles, and each corner carries half of those triangular volumes. Polygonal corner volumes are illustrated in Fig. 5. This concludes our description of the discrete polygonal-mesh equations.

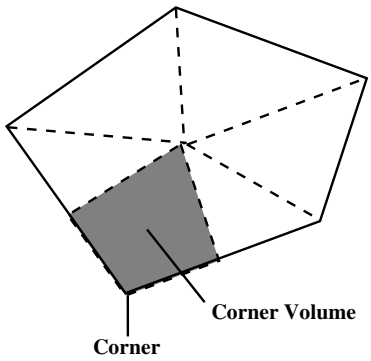

Fig. 5. The volume associated with a polygonal corner.

8. Calculations and results

Two sets of calculations in two-dimensional $r-z$ geometry were performed to compare the LMFGA and LMFGK methods. As elucidated in Section 6, both the LMFGA and LMFGK methods solve the same preconditioned equation given in Eq. (41). Nonetheless, considerable effort was required to define the specific steps used to execute each method so as to achieve as fair a comparison as possible. Thus we next give full descriptions of the LMFGA and LMFGK algorithms that were used for the test calculations.

To achieve a fair comparison, the same stopping criterion was used for both methods:

$$\frac{\|R^{\ell}\|_2}{\|\mathbf{C}\sum_{g=1}^{G}\sigma_{a,g}^{*}\mathbf{A}_g^{-1}\xi_g\|_2} < \epsilon. \tag{63}$$

Here, $\|\cdots\|_2$ stands for the L2-norm, ϵ is the convergence tolerance, and R^{ℓ} is the residual vector evaluated with f^{ℓ}, the solution at the ℓth iteration:

$$R^{\ell} = \mathbf{C}\sum_{g=1}^{G}\sigma_{a,g}^{*}\mathbf{A}_g^{-1}\xi_g - \mathbf{CB}f^{\ell}. \tag{64}$$

Note that the left side of Eq. (63) is the norm of the residual divided by the norm of the source vector.

It is evident from Eq. (45) that the residual vector R^{ℓ} for the LMFGA method is equal to $f^{\ell+1}-f^{\ell}$; therefore, the LMFGA method can be expressed as follows:

Algorithm 1 (*LMFGA algorithm*).

(1) Compute the right-hand-side of Eq. (41).
(2) Set the initial guess as the solution of the zeroth iteration f^0.
(3) For $\ell = 0, 1, \ldots$ Do:
- Compute $f^{\ell+1}$ by solving Eqs. (11a)–(11c).
- Set $R^{\ell} = f^{\ell+1} - f^{\ell}$.
- If the stopping criterion Eq. (63) is satisfied go to 5.

(4) EndDo
(5) Accept $\phi_g^{\ell+\frac{1}{2}} = \mathbf{A}_g^{-1}[\nu\chi_g f^{\ell} + \xi_g]$ as the angularly integrated radiation intensity for group g ($g = 1,G$).

Since $\phi_g^{\ell+\frac{1}{2}}$ has been computed during the computation of $f^{\ell+1}$, step 5 does not require any further computation. Computing the right-hand-side of Eq. (41) and the computation of $f^{\ell+1}$ in each iteration both require the inverses of the operators $\mathbf{A}_g$ and $\mathbf{H}$, defined in Eqs. (16a) and (33a), respectively; therefore, if the iteration stops at the kth iteration, the algorithm above requires solving $G+1$ diffusion equations $k+2$ times.

Without going into the details of the GMRES algorithm, the steps of the LMFGK method are as follows:

Algorithm 2 (*LMFGK algorithm*).

(1) Compute the right-hand-side of Eq. (41).
(2) Compute the initial residual vector R^0 using the initial guess f^0.
(3) If the stopping criterion Eq. (63) is satisfied set ℓ to 0 and go to 7.
(4) Set $v_1 = R^0/\|R^0\|_2$ as the first vector of the Krylov space.
(5) For $\ell = 1, 2, \ldots$ Do :
- Compute $w_{\ell} = \mathbf{CB}\ \mathrm{v}_{\ell}$ to extend the Krylov space.
- Compute $\|R^{\ell}\|_2$ and f^{ℓ} according to the GMRES algorithm.
- If the stopping criterion Eq. (63) is satisfied go to 7.
- If the dimension of the Krylov space reaches the maximum size:
  Restart using f^{ℓ} as the new initial guess and set $v_{\ell+1} = R^{\ell}/\|R^{\ell}\|_2$ as the first vector of the new Krylov space.

- Else:
  Compute $v_{\ell+1}$ from w_ℓ according to the GMRES algorithm.
- Endif

(6) Enddo
(7) Compute $\phi_g^{\ell+1} = \mathbf{A}_g^{-1}[v\chi_g f^\ell + \xi_g]$, the angularly integrated radiation intensity for group g ($g = 1,G$).

In the LMFGK method, the inverses of the operators $\mathbf{A}_g$ and $\mathbf{H}$ are needed in the computations of w_ℓ for each iteration and the right-hand-side of Eq. (41). In addition, computation of $\phi_g^{\ell+1}$ in step (7) needs the inverse of the operator $\mathbf{A}_g$; therefore, if the iteration stops at the kth iteration, the algorithm above requires solving the diffusion equations for G photon groups $k+3$ times and the grey diffusion equation $k+2$ times.

The calculations used 10 photon groups with the following group boundaries (in units of keV): 0.0001, 0.000316, 0.001, 0.00316, 0.01, 0.0316, 0.1, 0.316, 1.0, 3.16, 10.0. Photons with energies less than 0.0001 keV or greater than 10.0 keV are ignored. The absorption cross-section for each group was the geometric-mean of the opacities at the group boundaries evaluated with the following power-law formula:

$$\sigma_{a,g}(T,E) = 10.0\ \mathrm{cm}^{-1}\left(\frac{\rho}{\mathrm{g/cm}^3}\right)\left(\frac{T}{\mathrm{keV}}\right)^{-1/2}\left(\frac{E}{\mathrm{keV}}\right)^{-3}, \tag{65}$$

where T and E are material temperature and photon energy in units of keV, and ρ is the material density in units of g/cm^3. The diffusion coefficient for each group was

$$D_g = \min\left(\frac{1}{3\sigma_{a,g}}, \frac{\phi_g}{|\vec{\nabla}\phi_g|}\right), \tag{66}$$

evaluated at the beginning of each time step. This is a flux-limited coefficient that becomes equal to the standard diffusion coefficient in optically thick diffusive regions, but becomes smaller in optically thin regions to prevent non-physically large radiation fluxes. The material had a constant specific heat of 0.05 jerks/keV/g. The material temperatures were initially set to a constant value of 0.005 keV, and the radiation intensities were initially set to a Planck distribution at the initial material temperaure.

The geometry of the test problems was a 7-cm long cylinder with a 1-cm radius. This geometry is illustrated in Fig. 6. The spatial zoning of the cylinder follows:

(1) There are three regions.
  (a) Region 1 is defined by $r \in [0.0, 0.5]$ and $z \in [0.0, 7.0]$.
  (b) Region 2 is defined by $r \in [0.5, 0.6]$ and $z \in [0.0, 7.0]$.
  (c) Region 3 is defined by $r \in [0.6, 1.0]$ and $z \in [0.0, 7.0]$.
(2) Regions 1 and 3 were tessellated with 0.1 cm by 0.1 cm square zones.
(3) Region 2 was tessellated with zones that were 0.1 cm wide in the z-direction and had varying widths in the r-direction. In particular, the r-widths of the zones were chosen such that the r-width of each successive zone was 1.47392 times larger than that of the zone beneath it. Region 2 was tessellated with 10 such geometrically expanding zones. The r-width of the first of these zones (the zone with a minimum radius of 0.5 cm), was 1×10^{-3} cm.

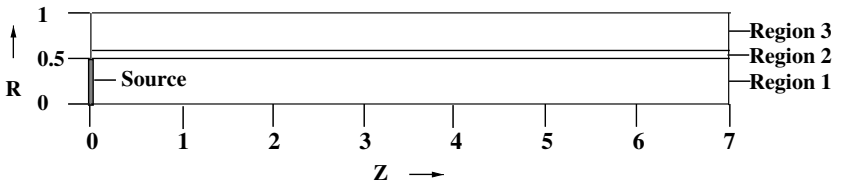

Fig. 6. The basic problem geometry for all calculations. Region 2 has a width of 0.1 cm.

The boundary conditions for the radiation follow:

(1) The following Marshak source boundary condition was applied on the left boundary subface defined by $z = 0.0$ cm and $r \leqslant 0.5$ cm:

$$D_g \vec{\nabla}\phi_g \cdot \hat{n} = \frac{4\pi B_g(T_s) - \phi_g}{2}, \tag{67}$$

where $\hat{n}$ is the outward-directed unit vector normal to the boundary, B_g is the integrated Planck function defined in Eq. (4), and T_s is a time-dependent source temperature. This source temperature had an initial value of 0.05 keV, increased linearly to reach 0.5 keV in 2.0 sh, and remained at 0.5 keV thereafter.

(2) A vacuum boundary condition was applied on the right boundary face defined by $z = 7.0$ cm. A vacuum boundary condition is a Marshak boundary condition defined in Eq. (67) with zero source temperature T_s.

(3) A reflecting boundary condition, $\vec{\nabla}\phi_g \cdot \hat{n} = 0$, was applied on the rest of the boundaries.

The first set of calculations corresponds to a problem with strong discontinuities in material properties. The material density of Region 1 was 0.01 g/cm^3 and that of Regions 2 and 3 was 2.0 g/cm^3. Each calculation started at $t = 0.0$ sh and ran to a final time of 20 sh with a fixed time-step size. The time steps were varied between calculations. Table 1 compares the LMFGA and LMFGK methods for three choices of time-step sizes Δt (0.1 sh, 0.01 sh, and 0.001 sh). Three numbers are compared:

- The wall-clock run time per time step, averaged over all the time steps. The calculations ran on a dedicated computing processor; therefore, the wall-clock time is very close to the CPU time.
- Number of iterations for solving the preconditioned equation Eq. (41) per time step, averaged over all the time steps, therein after referred to as number of "outer" iterations.
- Sum of the numbers of iterations taken by the linear solver for all the diffusion equations within each time step, averaged over all the time steps, therein after referred to as number of "inner" iterations.

The tolerance ϵ in the outer iteration stopping criterion, Eq. (63), was set to 10^{-5} for this set of calculations, and the maximum dimension of the Krylov space was set to 10. The diffusion equations associated with the inner iterations were solved using the GMRES algorithm with an algebraic multigrid (AMG) preconditioner. The GMRES algoithm was used rather than the conjugate-gradient algorithm because the polyhedral-mesh diffusion discretization is non-symmetric. The iterations for each diffusion equation were terminated when the L_2 norm of the residual divided by the L_2 norm of the source vector was less than 10^{-10}.

It can be seen from Table 1 that LMFGK takes fewer "outer" iterations than LMFGA. This is expected because GMRES should take no more iterations than Richardson iteration in solving the same set of linear equations. As the time-step size Δt increases, the advantage of LMFGK becomes more significant. Note, however, that the ratio of the number of "inner" iterations to the number of "outer" iterations is significantly smaller for LMFGA than for LMFGK. Consequently, the advantage of LMFGK in terms of the number of "outer" iterations does not translate proportionally to shorter run-times. The reason for this behavior lies in the difference in the right-hand-side of the diffusion equations. For LMFGA, the diffusion equation to be solved for each group is Eq. (11a). In solving Eq. (11a), the solution of the previous iteration $\phi_g^{l-(1/2)}$ is used as an initial guess for the current solution $\phi_g^{l+(1/2)}$. As the iteration proceeds, the initial guess becomes an

Table 1
Comparisons for the variable-density calculations

Δt (sh)	Method	Time (s)	No. of "Outer" iterations	No. of "Inner" iterations
0.1	LMFGK	1.553	7.98	1424.42
	LMFGA	2.992	25.73	2929.14
0.01	LMFGK	0.848	3.55	668.40
	LMFGA	0.986	8.07	785.00
0.001	LMFGK	0.450	1.26	260.32
	LMFGA	0.417	1.56	190.12

increasingly better approximation to the solution $\phi_g^{l+(1/2)}$. In contrast, for the LFMGK method the operator **B** defined in Eq. (19a) is applied to the vector v_ℓ each time the GMRES method extends its Krylov vector space (see Algorithm 2). Since v_ℓ is orthogonal to the Krylov vector space of iteration $(\ell - 1)$, there is no appropriate initial guess for solving the diffusion equation associated with the computation of $\mathbf{A}_g^{-1} v\chi_g v_\ell$. The significance of this fact was not evident to us before performing this computational study. For the variable-density problem with $\Delta t = 0.1$ sh the LMFGK method takes a factor of 3.2 fewer iterations than the LMFGA method, but is only a factor of 1.9 faster than the LMFGA method. For $\Delta t = 0.01$ sh, the LMFGK method takes a factor of 2.3 fewer iterations than the LMFGA method, but is only a factor of 1.16 faster than the LMFGA method. For $\Delta t = 0.001$ sh, the LMFGK method takes a factor of 1.24 fewer iterations than the LMFGA method, but is actually a factor of 1.08 *slower* than the LMFGA method.

It is possible for LMFGK to achieve higher efficiency by relaxing the tolerance for the "inner" iterations while the "outer" Krylov iteration proceeds according to the strategy presented in Refs. [6,7]. In the interest of fairness, one would like to apply a similar strategy to the LMFGA method, but it is not clear that such a strategy is valid for a traditional iteration technique since it was specifically developed for Krylov methods. Thus we chose not to apply this strategy in this initial study, and used an inner iteration tolerance for both methods that was quite small relative to the outer iteration tolerance. Inner versus outer iteration tolerances and dynamic modification of these tolerances are clearly topics for future research.

The second set of calculations corresponds to a homogeneous problem. The material density was 2.0 g/cm^3 in all regions. The tolerance ϵ in the outer iteration stopping criterion, Eq. (63), was set to 10^{-4} for this set of calculations. The remaining problem parameters were identical to those in the first set of calculations. The comparisons are given in Table 2 for three time-step sizes, Δt (0.2 sh, 0.1 sh, and 0.01 sh).

As one would expect, the advantage of the LMFGK method relative to the LMFGA method is not as significant for the homogeneous problem since the LMFGA method does not suffer the degradation associated with highly inhomogeneous problems. Although the LMFGK method always requires fewer outer iterations than the LMFGA method, the need for more inner iterations per outer iteration has a significant impact upon the efficiency of the LMFGK method. For the uniform-density problem with $\Delta t = 0.2$ sh the LMFGK method takes a factor of 1.8 fewer iterations than the LMFGA method, but is only a factor of 1.07 faster than the LMFGA method. For $\Delta t = 0.1$ sh, the LMFGK method takes a factor of 1.6 fewer iterations than the LMFGA method, but has the same run time as the LMFGA method. For $\Delta t = 0.01$ sh, the LMFGK method takes about a factor of 1.12 fewer iterations than the LMFGA method, but is actually a factor of 1.1 *slower* than the LMFGA method.

Even though only calculation-averaged iteration counts were given in Tables 1 and 2, it is important to recognize that iteration counts can vary significantly during a calculation. For instance, the outer iteration counts for both methods are plotted in Fig. 7 as a function of the time-step number for Problem 1 with a time step of 0.1 sh. It can be seen from Fig. 7 that the iteration counts for each method not only vary over the course of the calculation, but also vary relative to each other. For instance, the LMFGA method requires about five times more iterations than the LMFGK method near the beginning of the calculation, but both methods rapidly converge near the end. The latter effect is simply due to the fact that the solution is approaching steady-state at the end of the calculation.

The degradation exhibited by the LMFGA method in strongly heterogeneous problems is highly problem-dependent. Divergence of the LMFGA method has been observed in calculations at LLNL by one of the authors (B. Yang) and others, but when divergence occurs, it generally occurs over only a few critical time

Table 2
Comparisons for the uniform-density calculations

Δt (sh)	Method	Time (s)	No. of "Outer" iterations	No. of "inner" iterations
0.2	LMFGK	0.601	3.41	455.71
	LMFGA	0.646	6.14	472.31
0.1	LMFGK	0.561	2.83	403.83
	LMFGA	0.561	4.57	376.99
0.01	LMFGK	0.425	1.228	249.49
	LMFGA	0.388	1.375	180.97

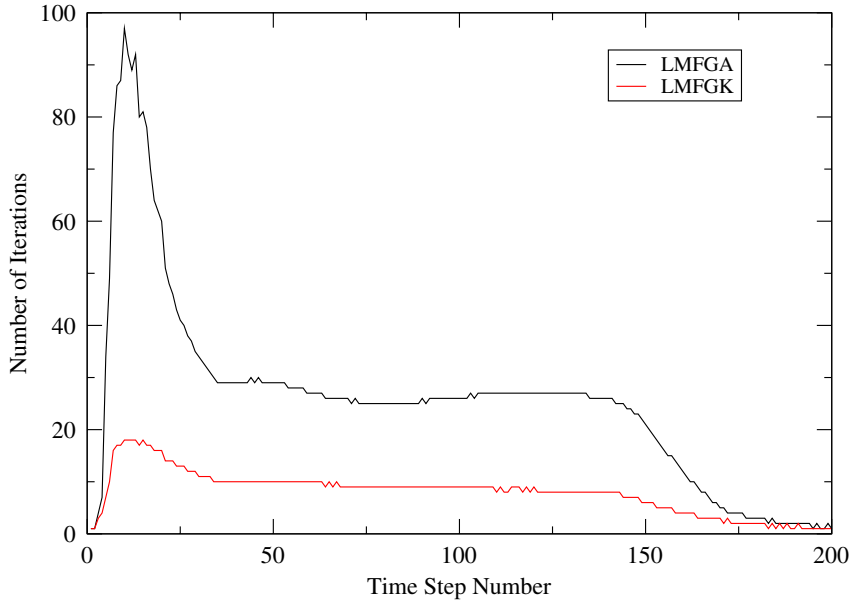

Fig. 7. Outer iteration counts as a function of time-step number for Problem 1 with the 0.1 sh time step.

steps. Heuristic measures have been defined that seem to allow such calculations to continue to completion, but the lack of robustness is clearly undesirable. We fully expect the LMFGK method to be far more robust than the LMFGA method, but robustness cannot be reliably gauged with a few test problems. Indeed, years of computational experience are required.

9. Conclusions and recommendations for future work

Our results indicate that the LMFGK method is indeed less sensitive to material inhomogeneities than the LMFGA method. They further indicate that the LMFGK method can be expected to have comparable efficiency at worst and significantly better efficiency at best relative to the LMFGA method. Thus we conclude that the LMFGK method is a superior alternative to the LMFGA method. Nonetheless, the increased number of inner iterations required per outer iteration with the LMFGK method clearly diminishes the impact of its rapid outer iteration convergence rate. The strategy of using a convergence tolerance for the inner iterations that changes as the outer iterations proceed clearly has the potential to significantly reduce the cost of the inner iterations. In principle, this type of strategy should be applicable to both the LMFGK and LMGFA methods. However, it is not clear that these methods will be equally amenable to this type of strategy. For instance, the strategy described in Refs. [6,7] was specifically developed for Krylov methods. Thus the potential exists to significantly alter the relative efficiencies of the LMFGK and LMFGA methods. We intend to address this issue in the near future.

A variant of the LMFGA method exists for multigroup thermal radiation transport calculations [9]. We are presently investigating a corresponding LMFGK method. Interestingly, there are two very different options for formulating such a method. Each has potential advantages and disadvantages relative to the other. We are investigating both of them and intend to report our results in the near future.

Acknowledgements

The research described here was performed in part under the auspices of the US Department of Energy at the Lawrence Livermore National Laboratory under contract No. W-7405-ENG-48. This work was supported in part by Los Alamos National Laboratory, operated for the US Department of Energy by Los Alamos National Security, LLC, under contract No. DE-AC52-06NA25396.

References

[1] J.E. Morel, E.W. Larsen, M.K. Matzen, A synthetic acceleration scheme for radiative diffusion calculations, J. Quant. Spectrosc. Radiat. Transfer 34 (1985) 243.
[2] Y. Azmy, Unconditionally stable and robust adjacent-cell diffusive preconditioning of weighted-difference particle transport methods is impossible, J. Comp. Phys. 182 (2002) 213.
[3] J.S. Warsa, T.A. Wareing, J.E. Morel, Krylov iterative methods and the degraded effectiveness of diffusion synthetic acceleration for multidimensional S_N calculations in problems with material discontinuities, Nucl. Sci. Eng. 147 (2004) 218.
[4] Y. SAAD, Iterative Methods for Sparse Linear Systems, PWS Publishing Company, Boston, 1996.
[5] S.L. Campbell, I.C.F. Ipsen, C.T. Kelley, C.D. Meyer, GMRES and the minimal polynomial, BIT Numer. Math. 36 (1996) 664.
[6] A. Bouras, V. Frayssé, Inexact matrix-vector products in Krylov methods for solving linear systems: a relaxation strategy, SIAM J. Matrix Anal. Appl. 26 (2005) 660.
[7] V. Simoncini, D. Szyld, Theory of inexact Krylov subspace methods and applications to scientific computing, SIAM J. Sci. Comput. 25 (2003) 454.
[8] T. Palmer, Discretization the diffusion equation on unstructured polygonal meshes in two dimensions, Ann. Nucl. Energ. 28 (2001) 1851.
[9] J.E. Morel, T.A. Wareing, K. Smith, A linear-discontinuous spatial differencing scheme for S_n radiative transfer calculations, J. Comput. Phys. 128 (1996) 445.