16

An Adaptive Energy Discretization of the Neutron Transport Equation Based on a Wavelet Galerkin Method

D. Fournier and R. Le Tellier
CEA, DEN, DER/SPRC/LEPh, Cadarache, F-13108 Saint-Paul-lez-Durance France

1. Introduction

The time-independent neutron transport equation derived from the Boltzmann equation with a linear collision kernel models the neutron population in the six dimensional space defined by $\vec{r} \in \mathcal{D}$ the space variable, $\vec{\Omega} \in \mathcal{S}_2$ the direction of motion variable and $E \in \mathcal{B} =]E_{G+1}, E_1[$ the energy variable. It represents the balance between the neutrons entering the hypervolume $d^3r d^2\Omega dE$ about $\left(\vec{r}, \vec{\Omega}, E\right)$ by fission or scattering and those leaving by streaming or any kind of interactions. The unknown is the so-called neutron flux $\phi(\vec{r}, \vec{\Omega}, E) = v(E)n(\vec{r}, \vec{\Omega}, E)$ with $n(\vec{r}, \vec{\Omega}, E)$ the neutron density and $v(E)$ the neutron velocity. The problem is defined in terms of the neutron interaction properties of the different materials *i.e.* the cross sections.
The solution of this equation in a deterministic way proceeds by the successive discretization of the three variables: energy, angle and space. The treatment of the energy variable invariably consists in a multigroup discretization which considers the cross sections and the flux to be constant within a group (*i.e.* a cell of the 1D energy mesh). A pre-homogenization of the cross sections is performed at the library processing level using a spatially independent weighting flux (*e.g.* $1/E$ spectrum in the epithermal range).
With a broad group structure ($\approx$ 100 to 2000 energy groups), this prior homogenization is unsufficient to take into account the case-specific, spatially-dependent, self-shielding effect *i.e.* the flux local depression in the vicinity of resonances that largely affects the neutron balance. As a consequence, a neutron transport calculation has to incorporate a so-called self-shielding model to correct the group cross sections of resonant isotopes. This homogenization stage of a neutron transport calculation is known to be a main source of errors for deterministic methods; as a consequence, an important work has been carried out to improve it. An optimized energy mesh structure (Mosca et al., 2011) in addition to an advanced self-shielding model (Hébert, 2007) is incorporated in state-of-the-art transport codes.
A different treatment for the energy variable based on a finite element approach is the basis of the present work. Such an avenue was proposed in the past by (Allen, 1986) but seldom used in practice. Indeed, finite element methods are commonly based on polynomial function bases which are not appropriate for non-smooth behavior.
Recently, two independent works by (Le Tellier et al., 2009) and (Yang et al., 2010) have proposed wavelet-Galerkin methods to overcome this issue. In this chapter, after a review

of these two approaches, we will focus on the development in this framework of adaptive algorithms with a control (at least, partial) of the discretization error. Such algorithms have been partially presented in a previous conference presentation by (Fournier & Le Tellier, 2009) but this book chapter gives a more in-depth presentation and updated numerical results for algorithms that may be of interest for other applications of wavelet-based finite elements. Such algorithms are analyzed in a limited framework (the fine structure flux equation for a single isotope diluted in a mixture of non-resonant isotopes in an infinite homogeneous medium) but the relevant issues regarding their extension in the general case are discussed.

2. Wavelet-Galerkin based energy discretization of the neutron transport equation

2.1 Generalized weak multigroup neutron transport equation

As in (Allen, 1986), the transport equation is discretized starting from the Sobolev spaces

$$W_2^1(\mathcal{D}\times\mathcal{S}_2) = \{\phi \in L_2(\mathcal{D}\times\mathcal{S}_2), \text{ all the weak derivatives of } \phi \in L_2(\mathcal{D}\times\mathcal{S}_2)\}, \quad (1)$$

$$W_2^{1\dagger}(\mathcal{A}) = \left\{\phi \in L_2(\mathcal{A}), \phi \in W_2^1(\mathcal{D}\times\mathcal{S}_2)\right\}. \quad (2)$$

with $\mathcal{A} = \mathcal{D}\times\mathcal{S}_2\times\mathcal{B}$. In particular, the energy variable is discretized as follows. An energy mesh consisting of G groups such that $E_1 > E_2 \cdots > E_{G+1}$ ($I_g =]E_{g+1}, E_g[$) is selected and the finite-dimension space $W^h(\mathcal{A}) \subset W_2^{1\dagger}(\mathcal{A})$ considered is

$$W^h(\mathcal{A}) = \left\{\phi \in W_2^{1\dagger}(\mathcal{A}), \phi(\vec{r},\vec{\Omega},E) = \sum_{g=1}^{G} \Pi_{I_g}(E)\underline{f}^{g\mathrm{T}}(E)\underline{\phi}^g(\vec{r},\vec{\Omega}) \text{ with } \underline{\phi}^g \in \left(W_2^1(\mathcal{D}\times\mathcal{S}_2)\right)^{N_g}\right\}, \quad (3)$$

where Π_{I_g} is the characteristic function of group g, $\underline{f}^g \in \left(L_2(I_g)\right)^{N_g}$ is an orthonormal set of wavelet functions on I_g and the group flux unknowns are the flux wavelet modes *i.e.* $\underline{\phi}^g(\vec{r},\vec{\Omega}) = \int_{I_g} dE\underline{f}^g(E)\phi(\vec{r},\vec{\Omega},E)$.

Within this framework, a Ritz-Galerkin procedure casts the transport equation (written with isotropic scattering and an external source $Q(\vec{r},\vec{\Omega},E)$) in a generalized (weak) multigroup form: $\forall g \in [1,G]$,

$$\left(\vec{\Omega}\cdot\vec{\nabla} + \underline{\underline{\Sigma_t}}^g(\vec{r})\right)\underline{\phi}^g(\vec{r},\vec{\Omega}) = \frac{1}{4\pi}\sum_{g'=1}^{G}\left(\underline{\underline{\Sigma_s}}^{g\leftarrow g'}(\vec{r})\right)\underline{\Phi}^{g'}(\vec{r}) + \underline{Q}^g(\vec{r},\vec{\Omega}), \quad (4)$$

where $\underline{\Phi}^g(\vec{r}) = \int_{\mathcal{S}_2} d^2\Omega\underline{\phi}^g(\vec{r},\vec{\Omega})$ and the source vector is $\underline{Q}^g(\vec{r},\vec{\Omega}) = \int_{I_g} dE\underline{f}^g(E)Q(\vec{r},\vec{\Omega},E)$.

The matrices coupling the flux modes within a group are defined in terms of the total $\Sigma_t(\vec{r},E)$ and scattering transfer $\Sigma_s(\vec{r},E'\to E)$ cross sections as

$$\underline{\underline{\Sigma_t}}^g(\vec{r}) = \int_{I_g} dE\underline{f}^g(E)\Sigma_t(\vec{r},E)\underline{f}^{g\mathrm{T}}(E), \quad (5)$$

$$\underline{\underline{\Sigma_s}}^g(\vec{r}) = \int_{I_g} dE\underline{f}^g(E)\int_{I_{g'}} dE'\Sigma_s(\vec{r},E'\to E)\underline{f}^{g\mathrm{T}}(E'). \quad (6)$$

Note that Eq. 5 introduces a coupling between the different modes within a group on the left hand side of the transport equation *i.e.* a coupling of the angular flux projections $\underline{\phi}^g(\vec{r},\vec{\Omega})$ that is not present for the standard multigroup approach.

Two different approaches by (Le Tellier et al., 2009) and (Yang et al., 2010) based on compactly supported Daubechies wavelets (Daubechies, 1992) have been proposed so far to deal with this coupling:

1. in (Yang et al., 2010), a dilation order is fixed and the basis consists in the translates of the associated scaling function; in this case, $\underline{\underline{\Sigma_t}}^g(\vec{r})$ is a band matrix and the mode coupling is limited in such a way that a Richardson iterative scheme can be employed to resolve this coupling.
2. in (Le Tellier et al., 2009), both dilates and translates of the mother wavelet functions are retained in the basis according to a thresholding procedure applied to the discrete wavelet transform of either the total cross section Σ_t or an approximate flux. In this case, the basis selection can be optimized but the modes are tightly coupled; a procedure based on a change of basis through a matrix diagonalization have been proposed to explictly decouple the equations.

This second approach proceeds as follows. Let us consider that the nuclear data are known by their projections on a set of orthonormal functions $\left(g_k^g\right)_{k\in[1,N_g]}$ in each group *e.g.*

$$\Sigma_t(\vec{r},E) = \underline{\hat{\Sigma}_t}^{g\mathrm{T}}(\vec{r})\underline{g}^g(E). \tag{7}$$

At this stage, $\underline{g}^g$ is assumed to be spatially uniform. This condition is satisfied for example if the same set of functions is considered for all the isotopes of a given configuration.
Considering the isomorphism between the Hilbert space $F_g = \text{span}\left(g_1^g \ldots, g_{N_g}^g\right)$ and $\mathbb{R}^{N_g}$, we can construct an orthonormal basis $(f_n^g)_{n\in[1,N_g]}$ of F_g in such a way that the different functions f_n^g are Σ_t-orthogonal. Indeed,

$$\underline{\underline{\tilde{\Sigma}_t}}^g(\vec{r}) = \int_{I_g} dE\underline{g}^g(E)\Sigma_t(\vec{r},E)\underline{g}^{g\mathrm{T}}(E) \tag{8}$$

is unitary similar to a diagonal matrix (see (Le Tellier et al., 2009) for more details) *i.e.*

$$\underline{\underline{\tilde{\Sigma}_t}}^g(\vec{r}) = \underline{\underline{C}}^g(\vec{r})\underline{\underline{\Sigma_t}}^g(\vec{r})\underline{\underline{C}}^{g\mathrm{T}}(\vec{r}), \tag{9}$$

with

- $\underline{\underline{C}}^g(\vec{r})$ = a unitary matrix containing the eigenvectors of $\underline{\underline{\tilde{\Sigma}_t}}^g(\vec{r})$,
- $\underline{\underline{\Sigma_t}}^g(\vec{r})$ = a diagonal matrix containing its eigenvalues.

Thus, $\underline{f}^g(\vec{r},E) = \underline{\underline{C}}^{g\mathrm{T}}(\vec{r})\underline{g}^g(E)$.
The problem at this stage is that the diagonalization of this operator depends on the spatial position through $\Sigma_t(\vec{r},E)$ *i.e.* $\underline{f}^g(\vec{r},E)$ depends on $\vec{r}$ and in the general case the discretized streaming operator is no longer diagonal. However, in most of the practical cases, the total cross section is defined as a step function with respect to the space variable *i.e.* a set of uniform media is defined and used to represent the spatial distribution of the nuclear data. Let us consider that the spatial domain $\mathcal{D}$ is split into a set of non-overlapping uniform medium domains *i.e.* $\mathcal{D} = \bigcup_i \mathcal{D}_i$. The total cross section (along with the other nuclear data) is represented as

$$\Sigma_t(\vec{r},E)=\sum_g \Pi_{I_g}(E)\sum_i \Pi_{\mathcal{D}_i}(\vec{r})\underline{\hat{\Sigma}_{ti}}^{g\mathrm{T}}(\vec{r})\underline{g}^g(E), \tag{10}$$

where $\Pi_{\mathcal{D}_i}$ is the characteristic function of $\mathcal{D}_i$ and the flux is expanded as

$$\phi(\vec{r},\vec{\Omega},E)=\sum_g \Pi_{I_g}(E)\sum_i \Pi_{\mathcal{D}_i}(\vec{r})\underline{f}_i^{g\mathrm{T}}(E)\underline{\phi}_i^g(\vec{r},\vec{\Omega}). \tag{11}$$

For a given i, $\underline{f}_i^g$ is uniform on $\mathcal{D}_i$ and is obtained by diagonalizing $\underline{\underline{\tilde{\Sigma}}}_{ti}^{\ g}$ as previously described.
For $\vec{r}$ belonging to a uniform medium domain $\mathcal{D}_i$, Eq. 4 can be written without any complication of the streaming term. In fact, this formulation of the transport equation is similar to the standard multigroup form. In this case, the mode coupling only appears for the conditions at the interface Γ_{ij} between two uniform medium domains $\mathcal{D}_i$ and $\mathcal{D}_j$ along $\vec{\Omega}$. The continuity of $\phi(\vec{r},\vec{\Omega},E)$ at $\vec{r}\in\Gamma_{ij}$ implies directly the continuity of $\phi^g(\vec{r},\vec{\Omega})$ in the standard multigroup case:

$$\phi_j^g(\vec{r},\vec{\Omega})=\phi_i^g(\vec{r},\vec{\Omega}), \tag{12}$$

while, in our case, it translates into

$$\underline{\phi}_j^g(\vec{r},\vec{\Omega})=\underline{\underline{C}}_j^{g\mathrm{T}}\underline{\underline{C}}_i^g\underline{\phi}_i^g(\vec{r},\vec{\Omega}). \tag{13}$$

When crossing an interface between two uniform media domain, a change of basis with respect to the energy expansion has to be performed in order to maintain a diagonal group transport operator over the whole domain.

2.2 Case of study

The numerical study of the proposed algorithms will be limited to the fine structure flux equation for a single isotope diluted in a mixture of non-resonant isotopes in an infinite homogeneous medium in such a way that only the energy variable has to be discretized. The total cross section is written as $\Sigma_t^+ + N^*\sigma_t^*(E)$ considering that Σ_t^+ is constant; $*$ refers to the resonant isotope. Considering $\underline{f}^g(E)$, the σ_t^{*g}-orthogonal and orthonormal basis of F_g, the weak form of the fine structure flux equation is written as

$$\left(\underline{\underline{\sigma_t}}^{*g}+\sigma_d\right)\underline{\phi}^g=\sum_{g'=1}^{G}\underline{\underline{\sigma_s}}^{*g\leftarrow g'}\underline{\phi}^{g'}+\sigma_d\int_{I_g}dE\underline{f}^g(E), \tag{14}$$

In matrix-vector form, this linear system is summarized as

$$H\Phi=S\Phi+Q. \tag{15}$$

We will also consider that the source-flux coupling in Eq. 15 is solved by a simple Richardson iterative scheme under the form

$$H\Phi^{n+1}=S\Phi^n+Q. \tag{16}$$

H^{-1} will be denoted A in the remainder.

2.3 Wavelet-based elements

Let θ be some function in $L_2(\mathbb{R})$. We consider the translates and dilates of θ denoted $\theta_{j,k}$ such that $\theta_{j,k}(x) = 2^{j/2}\theta(2^j x - k)(j \in \mathbb{Z}, k \in \mathbb{Z})$ and $V_j = \text{span}\left\{\theta_{j,k}, k \in \mathbb{Z}\right\}$ the generated linear spaces. θ is called the father wavelet or scaling function and is constructed in such a way that $\left\{V_j, j \in \mathbb{Z}\right\}$ is a multiresolution analysis (MRA) *i.e.*

- $\left\{\theta_{0,k}, k \in \mathbb{Z}\right\}$ is an orthonormal system in $L_2(\mathbb{R})$,
- $V_j \subset V_{j+1}, \forall j \in \mathbb{Z}$,
- $\bigcup_{j\geq 0} V_j$ is dense in $L_2(\mathbb{R})$.

Moreover, for convenience, we consider that θ is normalized in such a way that $\int dx\theta(x) = 1$. In this case, defining W_j by $V_{j+1} = W_j \oplus V_j (j \in \mathbb{Z})$, we obtain $L_2(\mathbb{R}) = V_0 \oplus \bigoplus_{j=0}^{\infty} W_j$. The next step is to find a function $\gamma \in W_0$ ($\gamma_{j,k}$ is defined in a same way as $\theta_{j,k}$) such that $\left\{\gamma_{0,k}, k \in \mathbb{Z}\right\}$ is an orthonormal basis of W_0. The existence of such a function is guaranteed but it is not unique; in any case, it verifies $\int dx\gamma(x) = 0$. This function is called the mother wavelet. Consequently, $\left\{\gamma_{j,k}, k \in \mathbb{Z}\right\}$ is an orthonormal basis of W_j. Note that the mother wavelet is always orthogonal to the father wavelet.

Within such a framework, any function $\phi \in L_2(\mathbb{R})$ has a unique representation in terms of an L_2-convergent series: (see (Hardle et al., 1997))

$$\phi(x) = \sum_k \alpha_{0,k}\theta_{0,k}(x) + \sum_{j=0}^{\infty}\sum_k \beta_{j,k}\gamma_{j,k}(x), \tag{17}$$

where $\alpha_{j,k}$ and $\beta_{j,k}$ correspond to the orthogonal projection of ϕ on $\theta_{j,k}$ and $\gamma_{j,k}$ respectively. In the present work, we consider for the basis functions $\underline{g}^g$ in each group I_g a subset of $\left(\left(\theta_{0,k}\right)_k, \left(\gamma_{j,k}\right)_{j,k}\right)$ obtained by the sampling, discrete wavelet transform and thresholding of $\sigma_t^{*g}(E)$ or an approximate flux restricted to I_g. This is to be distinguished from the work of (Yang et al., 2010) where the basis is composed of the scaling functions for a given dilation order j *i.e.* $\underline{g}^g = \left(\theta_{j,k}\right)_k$.

In the following, we restrict ourselves to compactly supported wavelets introduced by (Daubechies, 1992) constructed starting from a function $m_0(\xi) = \frac{1}{\sqrt{2}}\sum_k h_k e^{-ik\xi}$ where h_k are real-valued coefficients such that only a finite number M (the support length) of h_k are non-zero. In this context, the MRA obeys

$$\theta_{j-1,l} = \sum_k h_{k-2l}\theta_{j,k}, \tag{18}$$

$$\gamma_{j-1,l} = \sum_k g_{k-2l}\gamma_{j,k}, \tag{19}$$

and the decomposition of a sampled $N-$length signal is obtained efficiently by the discrete wavelet transform (DWT) based on the cascade algorithm proposed in (Mallat, 1989).

Following such a wavelet decomposition, the thresholding consists in replacing Eq. 17 by

$$\phi(x) = \sum_k \alpha_{0,k}\theta_{0,k}(x) + \sum_{j=0}^{J}\sum_k \tilde{\beta}_{j,k}\gamma_{j,k}(x), \tag{20}$$

where $(\tilde{\beta}_{j,k})_{j,k}$ is obtained from $(\beta_{j,k})_{j,k}$ and $\#(\tilde{\beta}_{j,k})_{j,k} \ll \#(\beta_{j,k})_{j,k}$
A natural criterion is to discard coefficients lower than a given cut-off ε *i.e.*

$$\tilde{\beta}_{j,k} = \begin{cases} 0 & \text{if} \quad |\beta_{j,k}| \leq \varepsilon \max_{j,k}(\beta_{j,k}), \\ \beta_{j,k} & \text{otherwise.} \end{cases} \tag{21}$$

This method is called hard thresholding. We refer the interested reader to (Le Tellier et al., 2009) for a comparison of different wavelet filters and thresholding strategies in this context.

3. Adaptivity

In the context of Eq. 16, adaptive algorithms aim at improving the operators discretization during the iterative process by dynamically selecting the basis functions and consequently, optimizing the computational cost and control (at least partially) the error on the final solution. The proposed algorithms aim at reducing the computational cost defined as the sum of the supports size at each iteration:

$$\text{cost} = \sum_{i=1}^{nbIter} \left(\#\Lambda_i^A + \#\Lambda_i^S \right), \tag{22}$$

where Λ_i^S (resp. Λ_i^A) represents the support of operator S (resp. A) at iteration i. Actually, the computational cost required to solve Eq. 16 is directly linked to the size of the operators manipulated: Λ_i^S for the construction of matrix S_i and Λ_i^A the order of the system used for iterations. It justifies the use of Eq. 22 as a measure of the algorithm computational cost.
Our work differs from the approach in (Cohen, 2003) where the goal was to minimize the final support. Here, the purpose is to find a balance between the number of iterations and the support size. In the following, two different algorithms are presented and tested. Both are based on a decomposition of the error in terms of the Richardson iterations residual ($\delta\epsilon^{res}$) and the errors due to the discretization of A and S operators (denoted $\delta\epsilon^A$ and $\delta\epsilon^S$ respectively):

$$\frac{\left\|\Phi^{n+1} - \Phi\right\|}{\left\|\Phi^{n+1}\right\|} \leq \frac{1}{1 - \|AS\|} \left(\delta\epsilon^A + \delta\epsilon^S + \delta\epsilon^{res} \right) = NB. \tag{23}$$

Sections 3.2 and 3.3 explicit this bound for both algorithms. The first version, inspired from (Cohen, 2003), uses two levels of iterations: one in order to increase the support and one to converge the residual. The single-loop algorithm is proposed as a simplification of the first one and a way to correlate the errors on the operators and the residual is detailed.

3.1 Numerical cases of study

As $\|AS\|$ plays an important role in both algorithms presented in Sections 3.2 and 3.3, tests are performed on different isotopes and energy ranges (the energy mesh used for this study contains 172 groups) as presented in Table 1.

Isotope	$\|AS\|$	Energy range (eV)	Energy groups
^{238}U	0.26	6.16 - 7.52	88
^{56}Fe	0.10	1018 - 1230	56
^{16}O	0.01	273.2E3 - 498.9E3	26-29

Table 1. Numerical cases of study for the two adaptive algorithms

3.2 Two-loop algorithm

In this algorithm, an outer iteration loop (index j) is added. At a given iteration j, the following system is solved:

$$\Phi_{j+1}^{n+1} = A_{j+1}\left(S_{j+1}\Phi_{j+1}^{n} + Q\right), \tag{24}$$

with A_{j+1} (resp. S_{j+1}) representing matrix A (resp. S) restricted to Λ_{j+1}^{A} (resp. Λ_{j+1}^{S}) support. The error is given by:

$$\begin{aligned}\Phi_{j+1}^{n+1} - \Phi &= A_{j+1}\left(S_{j+1}\Phi_{j+1}^{n} + Q\right) - A(S\Phi + Q)\\ &= \left(A_{j+1} - A\right)\left(S_{j+1}\Phi_{j+1}^{n} + Q\right) + A\left(S_{j+1} - S\right)\Phi_{j+1}^{n} + AS\left(\Phi_{j+1}^{n} - \Phi\right).\end{aligned} \tag{25}$$

It follows that the relative error can be expressed by Eq. 23 with

$$\delta\epsilon^{S} = \|A\|\frac{\left\|\left(S_{j+1} - S\right)\Phi_{j+1}^{n}\right\|}{\left\|\Phi_{j+1}^{n+1}\right\|}, \tag{26}$$

$$\delta\epsilon^{A} = \frac{\left\|\left(A_{j+1} - A\right)\left(S_{j+1}\Phi_{j+1}^{n} + Q\right)\right\|}{\left\|\Phi_{j+1}^{n+1}\right\|}, \tag{27}$$

$$\delta\epsilon^{res} = \|AS\|\frac{\left\|\Phi_{j+1}^{n+1} - \Phi_{j+1}^{n}\right\|}{\left\|\Phi_{j+1}^{n+1}\right\|}. \tag{28}$$

A main issue is the choice of the matrices S_{j+1} and A_{j+1} or, in other words, the selection of the wavelet supports. The idea in the remainder is to monitor the errors related to the operator discretizations using the numerical residual in order to obtain a relation of the type:

$$\frac{\left\|\Phi_{j+1}^{n+1} - \Phi\right\|}{\left\|\Phi_{j+1}^{n+1}\right\|} \leq K\frac{\left\|\Phi_{j+1}^{n+1} - \Phi_{j+1}^{n}\right\|}{\left\|\Phi_{j+1}^{n+1}\right\|}, \tag{29}$$

where K is a given constant. The error on the flux is thus controlled by the residual at each iteration.

The error on $\delta\epsilon^{S}$ (resp. $\delta\epsilon^{A}$) can be practically controlled by a thresholding on the product $S\Phi^{n}$ (resp. $A\left(S_{j+1}\Phi_{j+1}^{n} + Q\right)$) ensuring:

$$\left\|(S_{j+1} - S)\Phi_{j+1}^{n}\right\| \leq \epsilon'_{j+1}\left\|\Phi_{j+1}^{n}\right\|, \tag{30}$$

$$\left\|\left(A_{j+1} - A\right)\left(S_{j+1}\Phi_{j+1}^{n} + Q\right)\right\| \leq \epsilon_{j+1}\left\|\Phi_{j+1}^{n}\right\|. \tag{31}$$

Remaining coefficients give the new supports Λ^S_{j+1} and Λ^A_{j+1} such that $\#\Lambda^S_{j+1} \ll \#\Lambda^S$ and $\#\Lambda^A_{j+1} \ll \#\Lambda^A$ where Λ^S and Λ^A are the support of S and A operators approximated by a large number of coefficients. The localization property of wavelets ensure that these two supports slowly increase when ϵ'_{j+1} and ϵ_{j+1} decrease (see (Cohen, 2003) for more details).
By applying the procedures exposed above to A and S, Eq. 23 becomes:

$$\frac{\left\|\Phi^{n+1}_{j+1}-\Phi\right\|}{\left\|\Phi^{n+1}_{j+1}\right\|} \leq \frac{1}{1-\|AS\|}\left(\epsilon_{j+1}\frac{\left\|\Phi^{n}_{j+1}\right\|}{\left\|\Phi^{n+1}_{j+1}\right\|}+\epsilon'_{j+1}\|A\|\frac{\left\|\Phi^{n}_{j+1}\right\|}{\left\|\Phi^{n+1}_{j+1}\right\|}+\|AS\|\frac{\left\|\Phi^{n+1}_{j+1}-\Phi^{n}_{j+1}\right\|}{\left\|\Phi^{n+1}_{j+1}\right\|}\right). \tag{32}$$

Note however that the thresholding procedure described for operator A in Eq. 31 cannot be applied in the general context of the spatially-dependent transport equation (Eq. 4). A possibility is to use the same support for operators A and S. Such a solution has been tested in (Fournier & Le Tellier, 2009). Even if the convergence is deteriorated compared to the solution with two different supports for S and A, results are interesting and show that the adaptive algorithms proposed in this book chapter are extensible to the general problem.
As proposed in (Cohen, 2003), a geometrical decreasing sequence $\left(\epsilon_j\right)$ is fixed and iterations on n are performed until the residual becomes inferior to the value imposed by this sequence. To link ϵ_j and ϵ'_j, we ensure that the first two terms defined in Eq. 32 decay at the same rate by imposing:

$$\epsilon'_{j+1} = \frac{\epsilon_{j+1}}{\|A\|}. \tag{33}$$

At a given iteration j, Richardson iterations are carried out in order to ensure:

$$\frac{\left\|\Phi_{j+1}-\Phi_j\right\|}{\left\|\Phi_{j+1}\right\|} \leq \frac{\epsilon_{j+1}}{\|AS\|}. \tag{34}$$

Combining Eqs. 33 and 34 with the bound of Eq. 32 guarantees the convergence of the error:

$$\frac{\left\|\Phi_{j+1}-\Phi\right\|}{\left\|\Phi_{j+1}\right\|} \lesssim \frac{3\epsilon_{j+1}}{1-\|AS\|}. \tag{35}$$

The devised algorithm is written in pseudocode in Algorithm 1.
The choice of $\left(\epsilon_j\right)$ is arbitrary and some numerical tests have been performed with different values. A possible choice is

$$\epsilon = \frac{\epsilon_{j+1}}{\epsilon_j} = \left\|A_j S_j\right\|,$$

the rate of convergence of Richardson method.
Indeed, two asymptotic behaviours can be observed depending on the ϵ value with respect to $\rho = \|AS\|$ as presented in Figure 1 for ^{16}O where $\rho = 0.01$:

- $\epsilon \gg \rho$ (case $\epsilon = 1/2$ in Figure 1): Richardson iterative scheme converges rapidly (and even in one iteration in the presented case) and the error decreases linearly at the same rate than the sequence (ϵ_j) but it needs many outer iterations. In our example, the slope of the straight line is equal to $0.3 = \log(1/2) = \log(\epsilon)$.

Result: Solve $H\Phi = S\Phi + Q$ thanks to an adaptive procedure
Input : Matrix H, S and vector Q calculated on an "infinite" support
given accuracy tol
Output: Φ: flux solution of $\Phi = H^{-1}(S\Phi + Q)$ at accuracy tol
Data: Λ: support, ϵ: accuracy

$\Phi_0 = 0$, $\Lambda_0 = \emptyset$, $\epsilon_0 = 1$;
$j = 0$;
$err = 1$;
while $\epsilon_j \geq tol\ \frac{1-\|AS\|}{3}$ **do**
  $j \leftarrow j+1$;
  $\epsilon_j \leftarrow \epsilon\epsilon_{j-1}$;
  $\Phi_j^0 \leftarrow \Phi_{j-1}$;
  $n \leftarrow 1$;
  while $err \geq \frac{\epsilon_j}{\|AS\|}$ **do**
    $tmp = S\Phi_{j-1}^{n-1}$;
    $prod = \texttt{Thresholding}(tmp, \epsilon_j)$;
    % *remove smallest coefficients of* tmp, *guarantee* $\|tmp - prod\| \leq \epsilon_j \left\|\Phi_{j-1}^{n-1}\right\|$
    $R_j^n = prod + Q$;
    $\Lambda^S{}_j^n = \texttt{Support}(R_j^n)$;
    $\Phi_j^n = H^{-1}R_j^n$;
    $\Phi_j^n = \texttt{Thresholding}(\Phi_j^n, \frac{\epsilon_j}{\|H^{-1}\|})$;
    $\Lambda^A{}_j^n = \texttt{Support}(\Phi_j^n)$;
    $err \leftarrow \frac{\left\|\Phi_j^n - \Phi_j^{n-1}\right\|}{\left\|\Phi_j^n\right\|}$;
    $n \leftarrow n+1$;
  end
  $\Phi_{j+1} = \Phi_j^n$;
end

Algorithm 1: two-loop adaptive algorithm

- $\epsilon \ll \|AS\|$: the number of coefficients kept increases rapidly and several Richardson iterations are necessary to converge at a given support.

$\epsilon = \rho$ seems a good compromise between increasing too slowly the support causing useless iterations and keeping too many coefficients which implies the resolution of a uselessly large linear system.
Figure 2 presents the L^2-error as a function of the cost for ^{238}U and ^{16}O. Showing this two cases is interesting because they exhibit a different spectral radius ($\|AS\| = 0.26$ for ^{238}U and 0.01 for ^{16}O). As ϵ decreases, the cost decreases to a minimum value ($\epsilon = \frac{1}{8}$ for ^{238}U), and then increases again. As ϵ decreases, less iterations are performed which improves the cost; below a given value too large systems are solved and the cost increases (these are the two

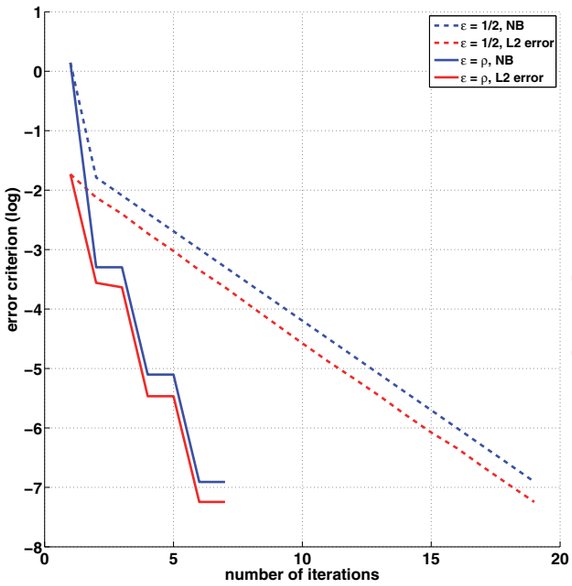

Fig. 1. L^2-error and numerical bound for different ϵ values on groups 26 to 29 of ^{16}O

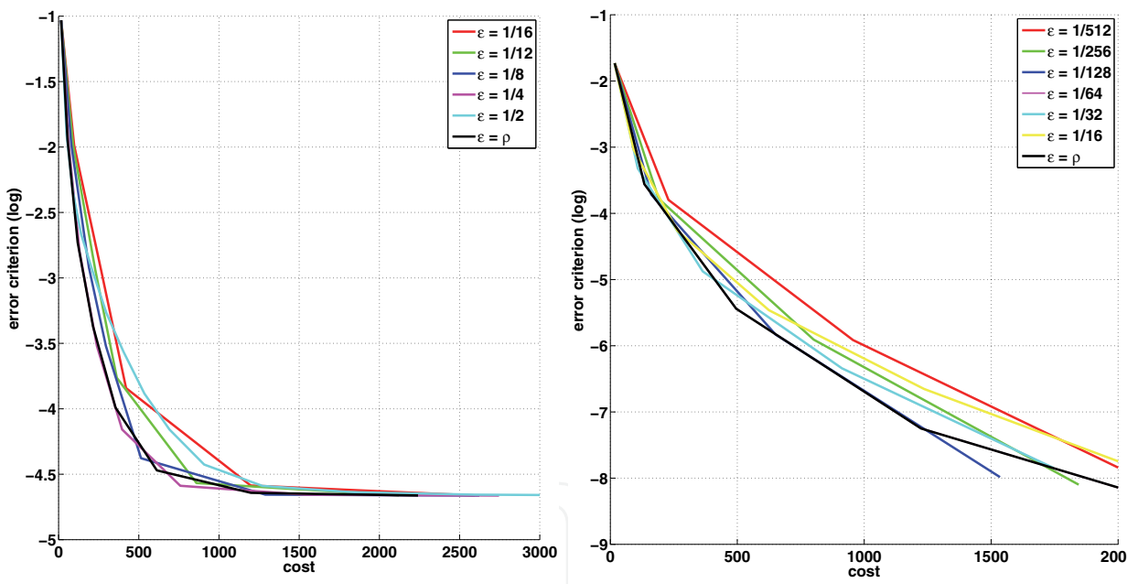

Fig. 2. Relative error versus cost for different ϵ values for group 88 of ^{238}U (left) and for groups 26 to 29 of ^{16}O (right)

behaviours illustrated in Figure 1). Obtaining the value of this minimum is not possible in the general case but let us mention that the use of the spectral radius $\left\|A_j S_j\right\|$ ensures a reasonable cost. Figure 3 further illustrates the evolution of the cost as a function of the parameter ϵ for ^{16}O and confirms the choice of $\left\|A_j S_j\right\|$ to minimize the cost.

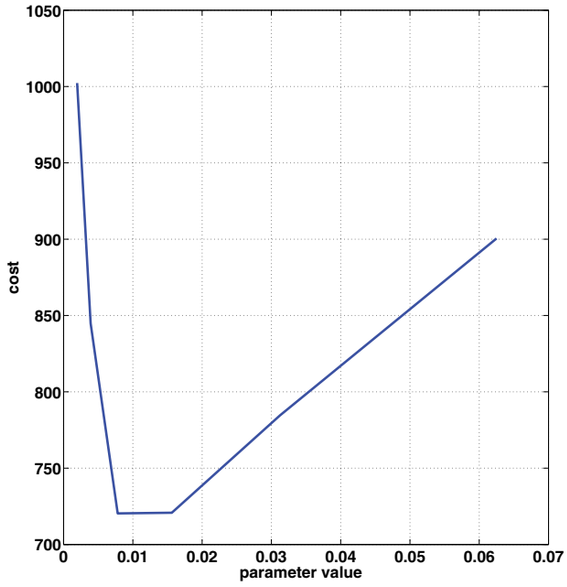

Fig. 3. Cost versus ϵ for a given accuracy of 10^{-6} for groups 26 to 29 of ^{16}O

3.3 Single-loop algorithm

The previous algorithm was directly inspired from Cohen (2003) and uses two levels of iterations which complicate the source iterations. Besides, the choice of the series (ϵ_j) is not obvious even if a geometrical sequence with a common ratio equal to $\left\| A_j S_j \right\|$ gives good results. As a simplification of this algorithm, a one-loop version is proposed, *i.e.* the iterative system is written as:

$$\Phi^{n+1} = A^{n+1}\left(S^{n+1}\Phi^n + Q\right). \tag{36}$$

A single loop means that the residual is no longer directly controlled and a strategy to handle this point has to be devised. At a given iteration, the residual is given by:

$$\begin{aligned}\Phi^{n+1} - \Phi^n &= A^{n+1}\left(S^{n+1}\Phi^n + Q\right) - A^n\left(S^n\Phi^{n-1} + Q\right)\\ &= \left(A^{n+1} - A^n\right)\left(S^{n+1}\Phi^n + Q\right) + A^n\left(S^{n+1} - S^n\right)\Phi^n + A^n S^n\left(\Phi^n - \Phi^{n-1}\right).\end{aligned} \tag{37}$$

And the same relationship as the one for the two-loop algorithm holds for the actual error:

$$(I - AS)\left(\Phi^{n+1} - \Phi\right) = \left(A^{n+1} - A\right)\left(S^{n+1}\Phi^n + Q\right) + A\left(S^{n+1} - S\right)\Phi^n - AS\left(\Phi^{n+1} - \Phi^n\right). \tag{38}$$

Substituting $\left(\Phi^{n+1} - \Phi^n\right)$ as given by Eq. 37 in Eq. 38 leads to an error bound given by Eq. 23

with

$$\delta\epsilon^S = \frac{\left\| A\left(\left(S^{n+1}-S\right) - SA^n\left(S^{n+1}-S^n\right)\right)\Phi^n\right\|}{\left\|\Phi^{n+1}\right\|}, \tag{39}$$

$$\delta\epsilon^A = \frac{\left\|\left(\left(A^{n+1}-A\right) - AS\left(A^{n+1}-A^n\right)\right)\left(S^{n+1}\Phi^n + Q\right)\right\|}{\left\|\Phi^{n+1}\right\|}, \tag{40}$$

$$\delta\epsilon^{res} = \|ASA^nS^n\| \frac{\left\|\Phi^n - \Phi^{n-1}\right\|}{\left\|\Phi^{n+1}\right\|}. \tag{41}$$

Such a bound for the operator-related error $\delta\epsilon^A$ (resp. $\delta\epsilon^S$) is interesting because it takes into account both $\|A_{n+1} - A\|$ (resp. $\|S_{n+1} - S\|$), the distance between the current operator and the complete one, and $\|A_{n+1} - A_n\|$ (resp. $\|S_{n+1} - S_n\|$), the distance between two successive operators. The direct control of the numerical residual with Richardson iterations in the previous algorithm is now "replaced" by the introduction of the distance between two successive operators in the error bounds on A and S. As the first term decreases with n until 0, the second one increases until $\|A - A_n\|$ (resp. $\|S - S_n\|$). Depending on the value of $\|AS\|$, $\left(\left\|A^{n+1} - A\right\| + \|AS\|\left\|A^{n+1} - A^n\right\|\right)$ can be strictly decreasing or presents a minimum or a maximum (Figure 4).

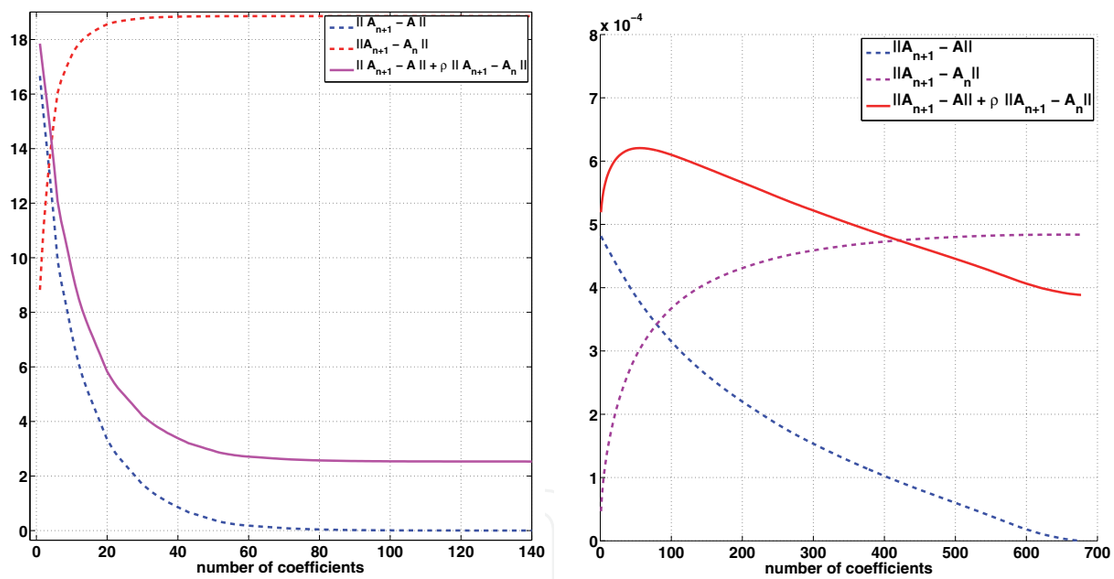

Fig. 4. Comparison of error terms defined in Eq. 40 for group 88 of ^{238}U with $\|AS\| = 0.26$ (left) and with $\|AS\|$ artificially increased to 0.8 (right)

Even if the general behaviour is not known, the initial and final bounds are given by:

$$\delta\epsilon^S_{(S_{n+1}=S)} = \delta\epsilon^S_{fin} = \|ASA^n\left(S - S^n\right)\Phi^n\|, \tag{42}$$

$$\delta\epsilon^A_{(A_{n+1}=A)} = \delta\epsilon^A_{fin} = \left\|AS\left(A - A^n\right)\left(S^{n+1}\Phi^n + Q\right)\right\|, \tag{43}$$

$$\delta\epsilon^S_{(S_{n+1}=S_n)} = \delta\epsilon^S_{ini} = \|A\left(S^n - S\right)\Phi^n\|, \tag{44}$$

$$\delta\epsilon^A_{(A_{n+1}=A_n)} = \delta\epsilon^A_{ini} = \left\|\left(A - A^n\right)\left(S^{n+1}\Phi^n + Q\right)\right\|. \tag{45}$$

As $\|AS\| < 1$ (ensuring the convergence of Richardson iterations), it guarantees that $\delta\epsilon^S_{fin} < \delta\epsilon^S_{ini}$ and $\delta\epsilon^A_{fin} < \delta\epsilon^A_{ini}$. These error bounds are at the basis of our algorithm. Three different cases are considered:

- $\delta\epsilon^{res} \in [\delta\epsilon^S_{fin}, \delta\epsilon^S_{ini}]$. It is possible to decrease the error due to operator S discretization to the numerical residual so S^{n+1} is chosen to ensure $\delta\epsilon^S \approx \delta\epsilon^{res}$.
- $\delta\epsilon^{res} < \delta\epsilon^S_{fin}$. Numerical residual is too small to be reached directly. Error on operator S is reduced to
$$\delta\epsilon^S = \alpha\delta\epsilon^S_{ini} + (1-\alpha)\delta\epsilon^S_{fin}, \tag{46}$$
with α fixed in $[0,1]$.
- $\delta\epsilon^{res} > \delta\epsilon^S_{ini}$. The numerical residual is not yet enough converged so the support of operator S is not modified, $S^{n+1} = S^n$.

The same approach is used to treat $\delta\epsilon^A$.
Figure 5 presents the behaviour of the three error terms and the numerical bound defined by Eq. 23. When $\|AS\|$ is low (Figure 5 (left)), Richardson iterations converge rapidly and do not slow the convergence of other terms. When $\|AS\|$ tends to 1 (Figure 5 (right)), more Richardson iterations are needed in order to converge the numerical residual and the operators support grows slowly and stepwise. It explains the decay by step observed for the operator discretization errors in Figure 5 (right).

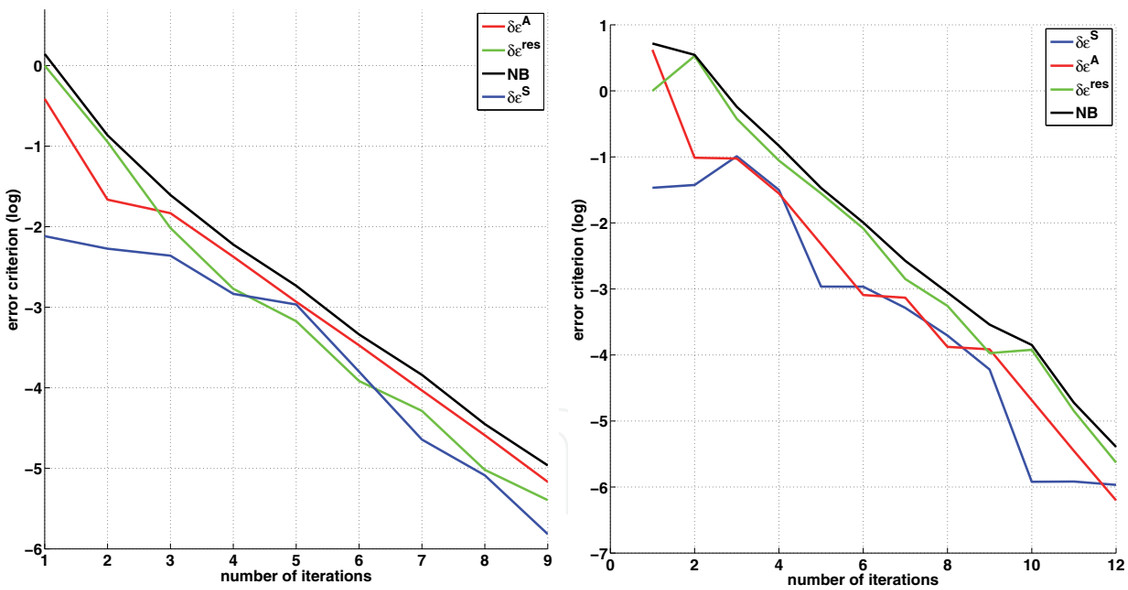

Fig. 5. Comparison of error terms on group 88 of ^{238}U with $\|AS\| = 0.26$ (left) and with $\|AS\|$ artificially increased to 0.8 (right)

The only remaining parameter is α. A numerical study is performed to give us some information about the optimal value.
Figure 6 shows that the choice of this parameter is important regarding the cost of the algorithm. If not enough coefficients are kept at each iteration, the rate of convergence is low which causes an important cost. On the opposite, if a large number is kept, large systems

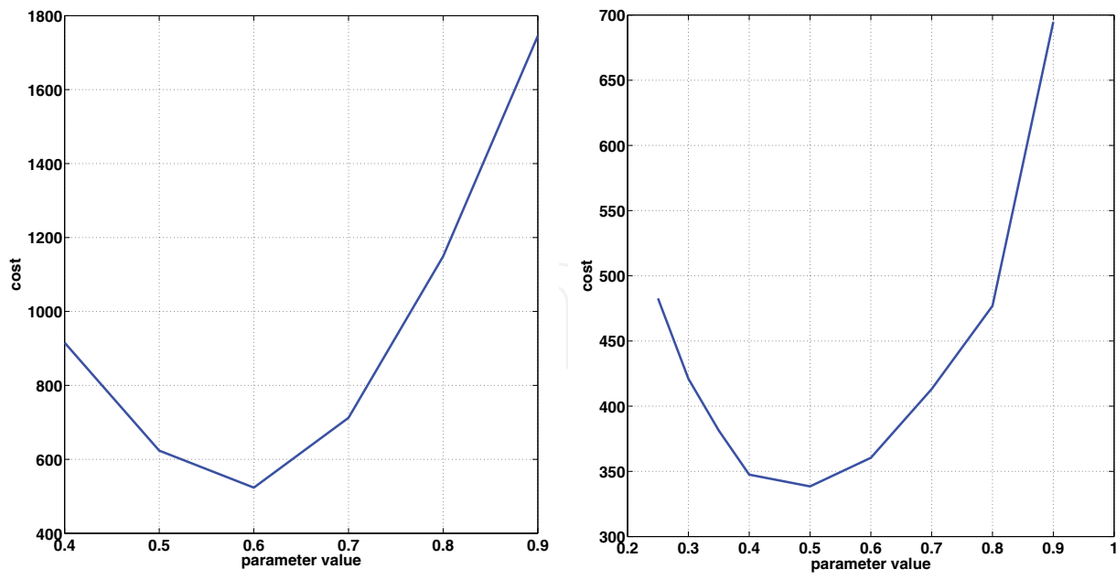

Fig. 6. Cost of the algorithm depending on α for a given accuracy $\epsilon = 10^{-5}$ on group 56 of ^{56}Fe (left) and $\epsilon = 10^{-4}$ on group 88 of ^{238}U (right)

have to be solved. An interesting compromise seems to keep coefficients in order to reduce the error by about half.

3.4 Comparison of the two algorithms

A comparison of the two algorithms and of the non-adaptive strategy is done in this section. All tests are performed by doing hard thresholding on an approximated flux and using symmlets of the 6^{th} order. All strategies are compared as a regard of the number of kept coefficients but also the cost defined by Eq. 22. To make non-adaptive and adaptive strategies comparable, non-adaptive Richardson iterations are stopped when $\delta\epsilon^{res}$ is of the same order as $\delta\epsilon^{S} + \delta\epsilon^{A}$ in such a way that the cost of the non-adaptive algorithm is nearly optimal.
Figure 7 (resp. Figure 8) presents results obtained on ^{238}U (resp. ^{56}Fe).
Figures 7 and 8 present coherent results and clearly highlight the interest of the two adaptive algorithms. The use of the spectral radius in the two-loop algorithm and the construction of our single-loop strategy make the convergence nearly independent of the case of study. Moreover, let us recall that the non-adaptive algorithm used in this study exhibits a nearly optimal cost and requires the control of the different error terms ($\delta\epsilon^{S}$, $\delta\epsilon^{A}$ and $\delta\epsilon^{res}$) as explained at the beginning of this section.
While both adaptive algorithms exhibit similar performances, the single-loop algorithm presents some advantages. First, the treatment of source iterations is easier with only one level of iteration. Then, the choice of the decreasing series (ϵ_j) is problem-dependent and more difficult to compute compared to the choice $\alpha = 0.5$ in Eq. 46 for the one-loop algorithm.

4. Conclusion

Considering a wavelet-based Galerkin discretization for treating the energy variable in the neutron transport equation, this chapter has proposed two adaptive algorithms for the Richardson iterative scheme that is commonly used to solve the source-flux coupling. While

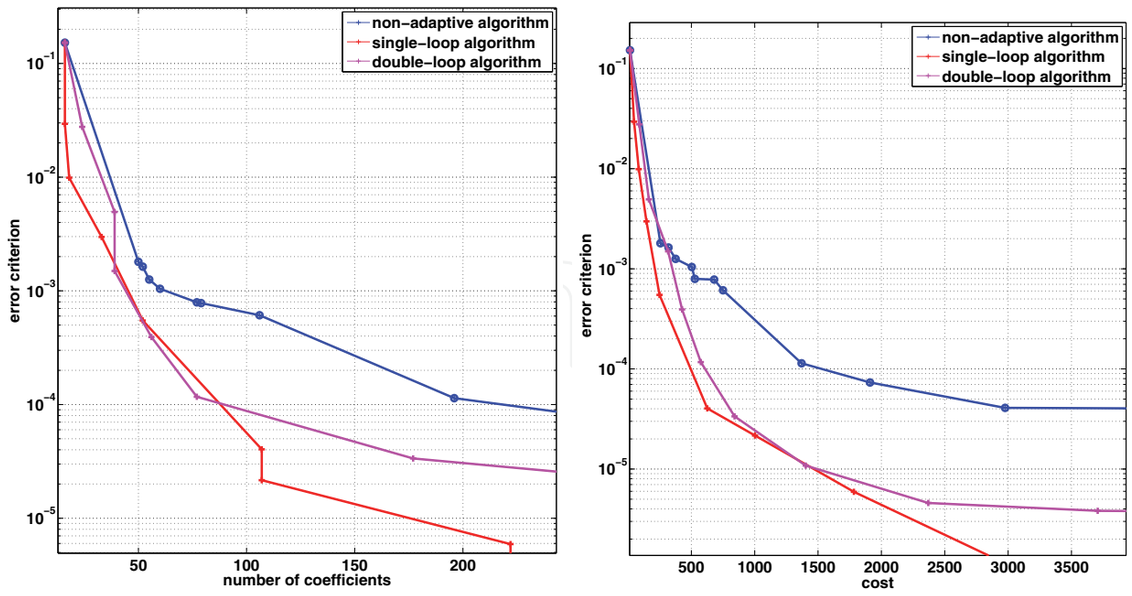

Fig. 7. Algorithms comparison in terms of the convergence (left) and the cost (right) for group 88 of ^{238}U

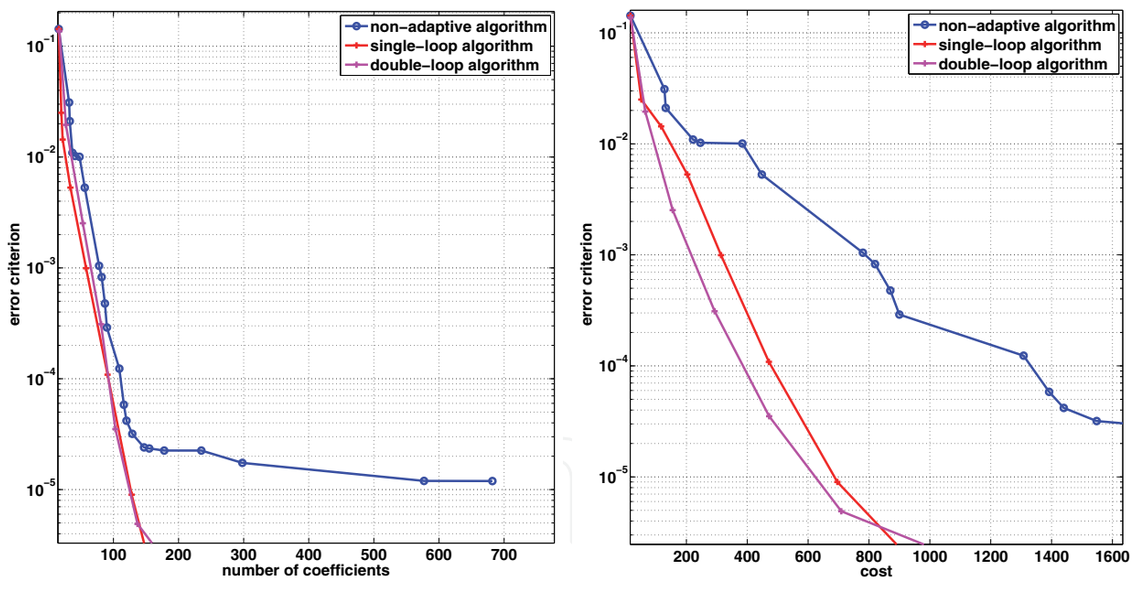

Fig. 8. Algorithms comparison in terms of the convergence (left) and the cost (right) for group 56 of ^{56}Fe

the first algorithm based on two nested loops is a modification of an algorithm previously proposed in the literature, the second one has been devised as a simplification that retains the same convergence properties. Both approaches are based on a formal decomposition of the error into three terms: two of them are related to the operators discretization while the third one is the Richardson residual. The algorithms then consist in a strategy to monitor and relate these three terms in such a way that error can be controlled by the Richardson iterations

residual. As a benefit of these algorithms, the accuracy of the final solution is known and the cost to obtain it has been decreased by adapting the size of the system during iterations. The performances of these algorithms have been demonstrated in the restricted framework of the fine structure flux equation in an homogeneous infinite medium. In the context of neutron transport calculations, the modifications necessary for spatially-dependent cases have been mentioned.

5. References

Allen, E. J. (1986). A finite element approach for treating the energy variable in the numerical solution of the neutron transport equation, *Transport Theory and Statistical Physics* **15**(4): 449–478.

Cohen, A. (2003). *Numerical Analysis of Wavelet Methods*, Vol. 32 of *Studies in Mathematics and its Application*, North Holland.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM.

Fournier, D. & Le Tellier, R. (2009). Adaptive algorithms for a self-shielding treatment using a wavelet-based Galerkin method, *Proc. of Int. Conf. on Mathematics, Computational Methods & Reactor Physics M&C 2009*, ANS, Saratoga Springs, USA.

Hardle, W., Kerkyacharian, G., Picard, D. & Tsybakov, A. (1997). Wavelets, approximation and statistical applications, Seminar Paris-Berlin.

Hébert, A. (2007). A review of legacy and advanced self-shielding models for lattice calculations, *Nuclear Science and Engineering* **155**(2): 310–320.

Le Tellier, R., Fournier, D. & Ruggieri, J. M. (2009). A wavelet-based finite element method for the self-shielding issue in neutron transport, *Nuclear Science and Engineering* **163**(1): 34–55.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**: 674–693.

Mosca, P., Mounier, C., Sanchez, R. & Arnaud, G. (2011). An adaptive energy mesh constructor for multigroup library generation for transport codes, *Nuclear Science and Engineering* **167**(1): 40–60.

Yang, W., Wu, H., Zheng, Y. & Cao, L. (2010). Application of wavelets scaling function expansion method in resonance self-shielding calculation, *Annals of Nuclear Energy* **37**(5): 653–663.

Discrete Wavelet Transforms - Algorithms and Applications

Edited by Prof. Hannu Olkkonen

ISBN 978-953-307-482-5
Hard cover, 296 pages
Publisher InTech
Published online 29, August, 2011
Published in print edition August, 2011

The discrete wavelet transform (DWT) algorithms have a firm position in processing of signals in several areas of research and industry. As DWT provides both octave-scale frequency and spatial timing of the analyzed signal, it is constantly used to solve and treat more and more advanced problems. The present book: Discrete Wavelet Transforms: Algorithms and Applications reviews the recent progress in discrete wavelet transform algorithms and applications. The book covers a wide range of methods (e.g. lifting, shift invariance, multi-scale analysis) for constructing DWTs. The book chapters are organized into four major parts. Part I describes the progress in hardware implementations of the DWT algorithms. Applications include multitone modulation for ADSL and equalization techniques, a scalable architecture for FPGA-implementation, lifting based algorithm for VLSI implementation, comparison between DWT and FFT based OFDM and modified SPIHT codec. Part II addresses image processing algorithms such as multiresolution approach for edge detection, low bit rate image compression, low complexity implementation of CQF wavelets and compression of multi-component images. Part III focuses watermaking DWT algorithms. Finally, Part IV describes shift invariant DWTs, DC lossless property, DWT based analysis and estimation of colored noise and an application of the wavelet Galerkin method. The chapters of the present book consist of both tutorial and highly advanced material. Therefore, the book is intended to be a reference text for graduate students and researchers to obtain state-of-the-art knowledge on specific applications.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元

Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen

IntechOpen