

# COMPARISON OF CHEBYSHEV AND ANDERSON ACCELERATIONS FOR THE NEUTRON TRANSPORT EQUATION

**Ansar Calloo<sup>1</sup>, Romain Le Tellier<sup>2</sup> and David Couyras<sup>3</sup>**

<sup>1</sup>DEN - Service d'études des réacteurs et de mathématiques appliquées (SERMA)  
 CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

<sup>2</sup>DEN - DTN/STMA/LMAG  
 CEA Cadarache, F-13108 Saint-Paul-lez-Durance, France

<sup>3</sup> EDF R&D/PERICLES  
 EDF Lab Paris-Saclay 7, boulevard Gaspard Monge, F-91120 Palaiseau, France

ansar.calloo@cea.fr, romain.le-tellier@cea.fr, david.couyras@edf.fr

## ABSTRACT

This work focuses on the  $k$ -eigenvalue problem of the neutron transport equation. The variables of interest are the largest eigenvalue ( $k_{eff}$ ) and the corresponding eigenmode is called the fundamental mode. Mathematically, this problem is usually solved using the power iteration method. However, the convergence of this algorithm can be very slow, especially if the dominance ratio is high as is the case in some reactor physics applications. Thus, the power iteration method has to be accelerated in some ways to improve its convergence. One such acceleration is the Chebyshev acceleration method which has been widely applied to legacy codes. In recent years, nonlinear methods have been applied to solve the  $k$ -eigenvalue problem. Nevertheless, they are often compared to the unaccelerated power iteration. Hence, the goal of this paper is to apply the Anderson acceleration to the power iteration, and compare its performance to the Chebyshev acceleration.

KEYWORDS: Power iteration,  $k$ -eigenvalue, Chebyshev acceleration, Anderson acceleration

## 1. INTRODUCTION

This work investigates the solution of the multigroup neutron transport equation with a discrete ordinates ( $S_n$ ) method. More specifically, it focuses on the  $k$ -eigenvalue problem of the equation. The multigroup discrete-ordinate transport equation[1] is written in operator form as:

$$\mathcal{H}\psi = \mathcal{F}\psi \text{ with } \begin{cases} \mathcal{H}\psi = \boldsymbol{\Omega}_n \cdot \nabla \psi_n^g(\mathbf{r}) + \Sigma_t^g \psi_n^g(\mathbf{r}) - \frac{1}{2\pi} \sum_{g'=1}^G \sum_{n'=1}^N \omega_{n'} \Sigma_s^{g' \rightarrow g}(\mathbf{r}, \boldsymbol{\Omega}_{n'} \cdot \boldsymbol{\Omega}_n) \psi_{n'}^{g'}(\mathbf{r}) \\ \mathcal{F}\psi = \frac{\chi^g}{4\pi} \sum_{g'=1}^G \sum_{n'=1}^N \omega_{n'} \nu \Sigma_f^{g'} \psi_{n'}^{g'}(\mathbf{r}) \end{cases} \quad (1)$$

Therefore, the critical reactor core is obtained by establishing an equilibrium between fission sources and the removal term, without external sources, and is equivalent to solving for the largest eigenvalue  $k_{eff}$  such that  $(k_{eff}, \psi)$  is the solution to  $\mathcal{H}^{-1}\mathcal{F}\psi = k_{eff}\psi$ .

In this case, the variables of interest are the largest eigenvalue ( $k_{eff}$ ) and the corresponding eigenmode is called the fundamental mode. Mathematically, this problem is usually solved using the power iteration method [2]. However, the convergence of this algorithm can be very slow and, the power iteration method has to be accelerated in some ways to improve its convergence. One such acceleration is the Chebyshev acceleration method [3] which has been applied to legacy codes. In recent years, nonlinear methods have been applied to solve the  $k$ -eigenvalue problem. Nevertheless, these are often compared to the unaccelerated power iteration method. Hence, the goal of this paper is to apply the Anderson acceleration method to the power iteration method, and to compare its performance to the Chebyshev acceleration method.

## 2. MATHEMATICAL BACKGROUND

### 2.1. Power iteration

The power iteration method consists in iterating on the fission sources to solve the eigenvalue problem and is as described in Algorithm 1. The approximate eigenvector  $\psi^{(n)}$  can be decomposed on the basis of eigenvectors of matrix  $\mathcal{H}^{-1}\mathcal{F}$  as  $\psi^{(n)} = \lambda_1^n \alpha_1 \psi_1 + \mathcal{O}\left(\left(\frac{\lambda_2}{\lambda_1}\right)^n\right)$ , where  $\lambda_i$  is the  $i^{th}$  (in order of decreasing magnitude) eigenvalue,  $\psi_i$  the associated eigenvector,  $\alpha_i$  a constant. For thermal reactors, the dominance ratio,  $\kappa = \frac{\lambda_2}{\lambda_1}$ , is usually above 0.95, and the power iteration method converges slowly. Thus, over the years, several acceleration methods have been conceived among which are rebalance techniques, Chebyshev acceleration and more recent nonlinear methods [4].

---

#### Algorithm 1: Power iteration

---

```

1 while  $\frac{|k_{eff}^{(n+1)} - k_{eff}^{(n)}|}{k_{eff}^{(n)}} > \epsilon_{k_{eff}}$  or  $\frac{\|\mathcal{F}\psi^{(n+1)} - \mathcal{F}\psi^{(n)}\|}{\|\mathcal{F}\psi^{(n)}\|} > \epsilon_{\psi}$  do
2     Apply source iteration for new flux  $\mathcal{H}\psi^{(n+1)} = \frac{1}{k_{eff}^{(n)}} \mathcal{F}\psi^{(n)}$ 
3     Compute the new eigenvalue  $k_{eff}^{(n+1)} = k_{eff}^{(n)} \frac{\langle \mathcal{F}\psi^{(n+1)}, \mathcal{F}\psi^{(n+1)} \rangle}{\langle \mathcal{F}\psi^{(n+1)}, \mathcal{F}\psi^{(n)} \rangle}$ 
4 end
    
```

---

### 2.2. Chebyshev acceleration

This consists in modifying the initial problem into an equivalent one with a smaller dominance ratio to improve the convergence of the power iteration algorithm. The Chebyshev acceleration consists in applying power iteration method on a polynomial of  $\mathcal{H}^{-1}\mathcal{F}$  such that the dominance ratio is as small as possible without changing the eigenvectors of the latter.

Thus, the starting point is to expand  $\psi^{(n)}$  on the basis of eigenvectors of  $A = \frac{\mathcal{H}^{-1}\mathcal{F}}{k_{eff}}$  as a polynomial  $P_l$ . Given that  $P_l(A)\psi = \sum_{i=1}^m \alpha_i P_l(A)\psi_i$  and that  $A^l\psi = \sum_{i=1}^m \alpha_i \lambda_i^l \psi_i$ :

$$P_l(A)\psi = \alpha_1 P_l(\lambda_1)\psi_1 + P_l(\lambda_1)(\alpha_2 \frac{P_l(\lambda_2)}{P_l(\lambda_1)}\psi_2 + \dots + \alpha_m \frac{P_l(\lambda_m)}{P_l(\lambda_1)}\psi_m) \quad (2)$$

and by minimising the fractional part of the former expression, the polynomial  $P_l$  obtained may be applied to accelerate the power iteration [3]. One such polynomial is obtained from the Chebyshev polynomial defined as  $P_n(x) = \frac{C_n(\frac{2x}{\kappa} - 1)}{C_n(\frac{2}{\kappa} - 1)}$  where  $C_n(x) = \cos(n \arccos(x))$  if  $|x| < 1$  or  $C_n(x) = \cosh(n \operatorname{arccosh}(x))$  otherwise. Following the recurrence property of the Chebyshev polynomials, the acceleration step is hence expressed as:

$$\begin{cases} \gamma = \frac{2}{2 - \kappa}, \sigma = \frac{\kappa}{2 - \kappa}, \rho^{(1)} = 1, \rho^{(2)} = \frac{2}{2 - \sigma^2 \rho^{(1)}} \text{ and } \rho^{(n+1)} = \frac{2}{2 - \sigma^2 \rho^{(n)}} \\ \mathcal{F}\psi^{(n+1)} = \frac{\rho^{(n+1)}\gamma}{\lambda_1} \mathcal{F}\psi^{(n+1)} - \rho^{(n+1)}(1 - \gamma)\mathcal{F}\psi^{(n)} + (1 - \rho^{(n+1)})\mathcal{F}\psi^{(n-1)} \end{cases} \quad (3)$$

The main drawback of this method is that it requires *a priori* knowledge of  $\lambda_1$  and the dominance ratio  $\kappa$  - its efficiency depends on an appropriate estimate for these. Algorithm 2 illustrates an implementation of the Chebyshev acceleration in APOLLO3<sup>®</sup>/MINARET[5].

---

**Algorithm 2:** Chebyshev acceleration for iteration ( $n$ )

---

- 1 Compute an estimation of the dominance ratio  $\kappa^{(n+1)} = \frac{\epsilon^{(n)}}{\epsilon^{(n-1)}}$  at iteration  $n$
  - 2 Accelerate source vector
  - 3 **if**  $n == n_{start}$  **then**
  - 4     **if**  $\kappa^{(n+1)} < 1$  &  $\|\kappa^{(n+1)} - \kappa^{(n)}\| < \epsilon$  **then**
  - 5          $\alpha = \frac{2}{\kappa^{(n+1)}}, \rho = \frac{1}{\alpha - 1}, r_0 = \alpha \cdot \rho, r_1 = -\rho$
  - 6          $\mathcal{F}\psi^{(n+1)} = r_0 \cdot \mathcal{F}\psi^{(n+1)} - r_1 \cdot \mathcal{F}\psi^{(n-1)}$
  - 7     **end**
  - 8 **end**
  - 9 **else if**  $n > n_{start}$  **then**
  - 10      $\rho = \frac{4}{4(\alpha - 1) - \rho}, r_0 = \frac{\rho}{\rho - \alpha}, r_0 = \alpha \cdot \rho, r_1 = -\rho, r_2 = 1 - r_0 - r_1$
  - 11      $\mathcal{F}\psi^{(n+1)} = r_0 \cdot \mathcal{F}\psi^{(n+1)} + r_1 \cdot \mathcal{F}\psi^{(n)} + r_2 \cdot \mathcal{F}\psi^{(n-1)}$
  - 12 **end**
- 

### 2.3. Anderson acceleration

In this research effort, the Anderson acceleration [6] is applied to the power iteration algorithm. The power iteration works with two iterates  $\psi^{(n+1)}$  and  $\psi^{(n)}$  - it corresponds to a Krylov subspace of dimension 1. The method can be improved by using information from previous iterates. Chebyshev acceleration or Successive Over-Relaxation update the current iteration using a linear combination of previous iterates. In this sense, Anderson acceleration (Anderson mixing in quantum chemistry[7]) attempts to benefit from past iterations. It has been applied in[8] to accelerate

neutron-thermal hydraulics problems. Anderson acceleration is an extrapolation method which has been successfully applied recently to neutron transport by [4,9].

Anderson belongs to the class of Newton-Krylov methods except that the subspace projection is orthogonalised using only the number of vectors prescribed by the history length of the method. Unlike Newton methods, it does not require the forming of the Jacobian matrix (similar to Jacobian-Free Newton-Krylov). No formal proof of convergence exists in literature but [7] showed that under some assumptions, convergence is observed in practice. The power iterations described in Algorithm 1 (lines 2 and 3) can be recast as fixed point iterations under the following form:

$$\mathbf{u}^{(n+1)} = \frac{\langle \mathbf{u}^{(n)}, \mathcal{FH}^{-1}\mathbf{u}^{(n)} \rangle}{\langle \mathcal{FH}^{-1}\mathbf{u}^{(n)}, \mathcal{FH}^{-1}\mathbf{u}^{(n)} \rangle} \mathcal{FH}^{-1}\mathbf{u}^{(n)} = f(\mathbf{u}^{(n)}) \quad (4)$$

associated with the non-linear problem  $u = f(u)$  where  $u$  is the ratio of the fission source to the dominant eigenvalue  $\frac{\mathcal{F}\psi}{k_{eff}}$ . Anderson acceleration can be applied to accelerate such sequences as described in Algorithm 3. Each Anderson iteration requires the solution to a constrained least-

---

**Algorithm 3:** Anderson acceleration

---

```

1 Given  $\mathbf{u}^{(0)}$  and a history length of  $m \geq 1$ 
2  $\mathbf{u}^{(1)} = f(\mathbf{u}^{(0)})$ 
3 while  $\frac{\|\mathbf{u}^{(n+1)} - f(\mathbf{u}^{(n)})\|}{\|f(\mathbf{u}^{(n)})\|} > \epsilon_{\mathbf{u}}$  do
4    $m_n = \min(m, n)$ 
5   Compute current residual
6    $\mathcal{G}^{(n)} = (\mathbf{g}_{n-m_n}, \dots, \mathbf{g}_n)$  with  $\mathbf{g}_n = f(\mathbf{u}^{(n)}) - \mathbf{u}^{(n)}$ 
7   Determine  $\alpha^{(n)} = (\alpha_0^{(n)}, \dots, \alpha_{m_n}^{(n)})$  which solves the following minimisation problem
8    $\min_{\alpha} \|\mathcal{G}^{(n)}\alpha\|$  such that  $\sum_{i=0}^{m_n} \alpha_i = 1$ 
9   Compute new estimate for unknown vector with a damping factor  $\beta_n$ 
10   $\mathbf{u}^{(n+1)} = (1 - \beta_n) \sum_{i=0}^{m_n} \alpha_i^{(n)} \mathbf{u}^{(n-m_n+i)} + \beta_n \sum_{i=0}^{m_n} \alpha_i^{(n)} f(\mathbf{u}^{(n-m_n+i)})$ 
11 end
    
```

---

squares problem that is solved in our implementation in a naive way using the normal equations. On the difference of the Chebyshev algorithm, Anderson acceleration offers only two free parameters: the damping coefficient  $\beta$  and the history length  $m_n$ . Besides, on the computational side, it should be noted that increasing the order of the method leads to higher memory considerations to store the “longer” history of vectors.

### 3. NUMERICAL RESULTS

The Chebyshev and Anderson acceleration methods were implemented in a mock-up in-house  $S_n$  code and several benchmarks have been considered. First, 1D cases with reflector/active core/reflector were studied to obtain a general trend of the acceleration methods. Then, 2D Cartesian benchmarks have been set up with varying core sizes up to sizes with  $10 \times 10$  assemblies of sidelength 20 cm (with larger cores, the power iteration method converges more and more slowly) and one reflector assembly of similar dimensions. All calculations were carried out with  $S_2$  level-symmetric quadratures [1]. Two-group cross sections with upscattering have been employed as

given in Table 1. In forthcoming sections, the results illustrated are for the 2D Cartesian geometry

**Table 1: Benchmark cross sections. Fuel and reflector share the same cross sections, except the reflector has a nil  $\nu\Sigma_f$ .**

Cross section	Group 1	Group 2
$\Sigma_t$ (cm <sup>-1</sup> )	0.53	1.30
$\nu\Sigma_f$ (cm <sup>-1</sup> )	0.0071	0.14
$\Sigma_{s0}$ (cm <sup>-1</sup> )	$\Sigma_{s0}^{1\rightarrow 1} = 0.500$ and $\Sigma_{s0}^{1\rightarrow 2} = 0.016$	$\Sigma_{s0}^{2\rightarrow 1} = 0.002$ and $\Sigma_{s0}^{2\rightarrow 2} = 1.200$

with  $10 \times 10$  assemblies as it is the most challenging case considered. The reference solution is a well-converged calculation, with a tolerance of  $10^{-13}$  with free iterations to avoid any effect of the acceleration method. The plots given next correspond to the error on the fission source compared to the reference solution given against the total number of inner iterations (transport sweeps). These figures are for a maximum of 5 spatial sweep iteration (Richardson), and 5 multigroup iterations (Gauss-Seidel type). The maximum number of power iterations was arbitrarily set to 500. In each plot, the total number of transport sweeps for unaccelerated (or free) iterations are given as  $N_{total}^{free}$  (if  $N_{total}^{free} = 5 \times 5 \times 500 = 12500$ , the unaccelerated calculation is deemed unconverged.)

### 3.1. Anderson acceleration results

Here, some general trends for the Anderson acceleration are highlighted.

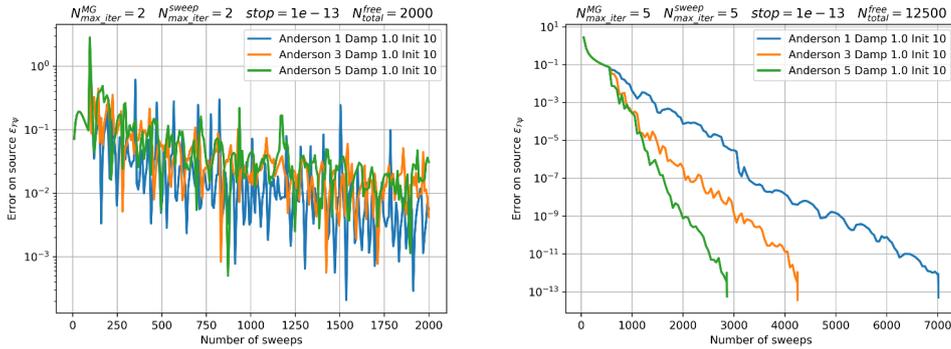
#### 3.1.1. Impact of the convergence of the inner iterations

The transport equation is solved using a series of nested iterations (Gauss-Seidel, Richardson). As shown in Fig. 1 (left figure), if insufficient transport sweeps are carried out, the calculation does not converge at all, as the initial solution vectors are inaccurate. Allowing for more transport sweeps (Fig. 1 right figure) mitigates this problem.

Besides, other works like [10] concluded that for Newton-like methods, the accuracy of the inner iterations must be increased as the outer iteration approaches the solution. For nested iterations with outer Krylov such as Anderson acceleration, the initial iterates must be computed with good accuracy, which can be relaxed later on in the process [11]. Even when affected, the Chebyshev acceleration is less sensitive to the inner loop convergence compared to the Anderson acceleration.

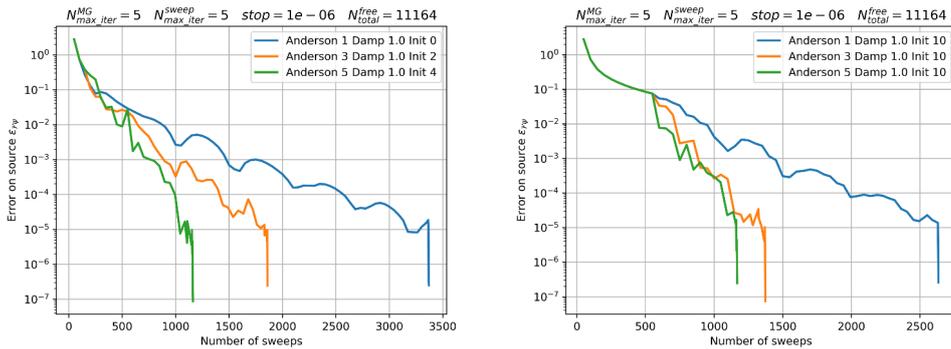
#### 3.1.2. Impact of the history length

Starting the acceleration method as soon as possible (when sufficient iterations have been carried out to extrapolate the source vector using past iterations for a given order) is not always a right strategy. From Fig. 2, it can be observed that imposing sufficient free iterations beforehand leads to better eigenvectors to start building the solution space, and thus increases the rate of convergence. Besides, it has been observed that for less strict stopping criteria, the impact was more significant



**Figure 1: Convergence of the Anderson algorithm for inner loops with different number of maximum iterations but with same stopping criteria of  $10^{-13}$**

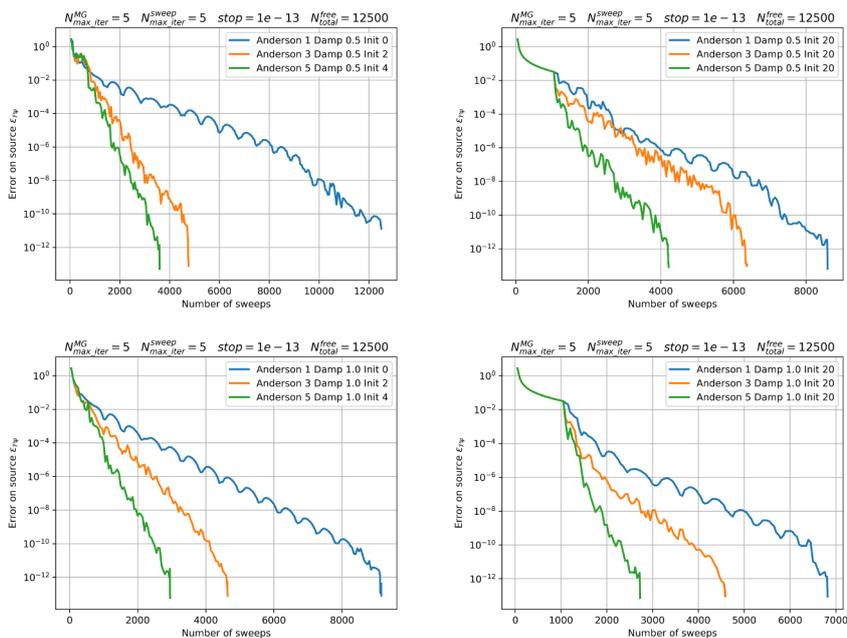
as the accuracy is greatly improved. In addition, as the order increases, the effect of the history length is less appreciable.



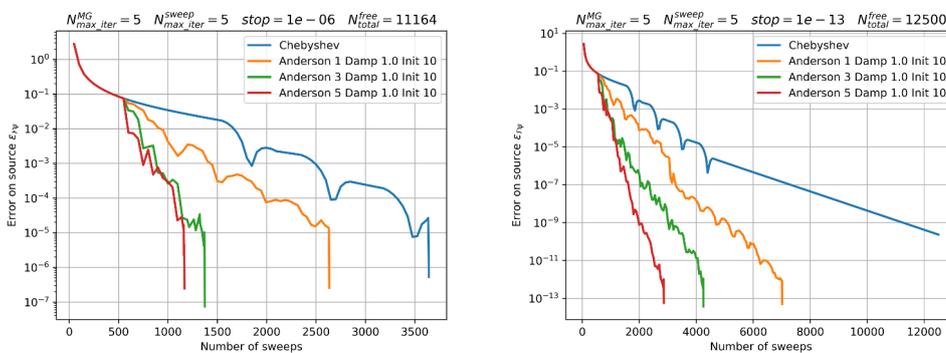
**Figure 2: Effect of triggering the Anderson acceleration as soon the history length is enough to compute the correct Anderson order (left) or by imposing a fixed number of free iterations - here 10 - before starting the acceleration (right).**

### 3.1.3. Impact of the damping factor

Fig. 3 illustrates the effect of the damping coefficient  $\beta_n$  on the rate of convergence. Two sets of results are provided:  $\beta_n$  of 0.5 and 1, and for minimal history lengths and with a fixed number of initial free iterations. In any case, a damping coefficient of 1 is optimal with respect to the rate of convergence. Furthermore, it appears that as the order increases the impact of the damping coefficient is less significant ( $\sim 1000$  sweeps difference for order 5, whereas  $> 5000$  for order 1).



**Figure 3: Effect of the damping ( $\beta_n$ ) on the Anderson acceleration. Two sets of results are given: top row for  $\beta_n = 0.5$  with vector sizes of “as soon as possible” (left) and 20 (right), and below for  $\beta_n = 1$ .**



**Figure 4: Anderson vs. Chebyshev with stopping criteria of  $10^{-6}$ (left) and  $10^{-13}$ (right).**

### 3.2. Anderson vs. Chebyshev

In any case as shown in Fig. 4, even for less stringent stopping criteria, Anderson 3 - which employs the same number of previous iterations for extrapolation as the Chebyshev acceleration - is more efficient than Chebyshev acceleration. It converges with almost three times less iterations than required by Chebyshev acceleration. With strict tolerances, the Chebyshev algorithm fails to

accelerate the source to the required criteria.

#### 4. CONCLUSIONS

The Anderson acceleration method has been applied successfully to the power iteration as a good candidate to substitute the Chebyshev acceleration. The Anderson acceleration relies on a strong mathematical background with two parameters left to user: the history length and the damping factor. Triggering the acceleration too early is not desirable, sufficient free iterations are required to obtain better initial histories and thus more precise initial vectors for the subspace. As for the damping parameter, it is observed that although it can be varied, a value of unity (1) is optimal for the application of Anderson acceleration to the power iteration method for the neutron transport equation. Anderson acceleration is also sensitive to the convergence of inner iterations and further work must be carried out to investigate the optimisation of the computational cost by introducing relaxation techniques for the convergence of nested loops.

#### REFERENCES

- [1] A. HEBERT. *Applied Reactor Physics*. Presses Internationales Polytechnique (2009).
- [2] E. E. LEWIS, W. F. MILLER Jr. *Computational Methods for Neutron Transport*. John Wiley & Sons Inc. (1984).
- [3] D. FERGUSON, K. DERSTINE. “Optimized Iteration Strategies and Data Management Considerations for Fast Reactor Finite Difference Diffusion Theory Codes.” *Nuclear Science & Engineering*, **volume 64**, pp. 593–604 (1977).
- [4] E. FICHTL, J. WARSA, M. CALEF. “Nonlinear acceleration of  $S_n$  transport calculations.” In *Proc. of Int. Conf. on Mathematics and Computational Methods Applied to Nuclear Science and Engineering M&C 2011*. Rio de Janeiro, RJ, Brazil (2011).
- [5] D. SCHNEIDER, F. DOLCI, F. GABRIEL, and J.-M. P. et. al. “APOLLO3<sup>®</sup>: CEA/DEN deterministic multipurpose platform for core physics analysis.” In *Proc. of Int. Mtg. on the Physics of Fuel Cycles and Advanced Nuclear Systems PHYSOR 2016*. Sun Valley, USA (2016).
- [6] D. ANDERSON. “Iterative procedures for nonlinear integral equation.” *Journal of Association for Computing Machinery*, **volume 4**, pp. 547–560 (1965).
- [7] H. WALKER, P. NI. “Anderson acceleration for fixed-point iteration.” *SIAM Journal of Numerical Analysis*, **volume 49**, pp. 1715–1735 (2011).
- [8] A. TOTH, C. KELLEY, S. SLATTERY, S. HAMILTON, K. CLARNO, R. PAWLOWSKI. “Analysis of Anderson acceleration on a simplified neutronics-thermal hydraulics system.” In *Proceedings of M&C 2015*. Nashville, TN, USA (2015).
- [9] M. CALEF, E. FICHTL, J. WARSA, M. BERNDT, N. CARLSON. “Nonlinear Krylov acceleration applied to a discrete ordinates formulation of the  $k$ -eigenvalue problem.” *Journal of Computational Physics*, **volume 238**, pp. 188–209 (2013).
- [10] S. EISENSTAT, H. WALKER. “Globally convergent inexact Newton methods.” *SIAM Journal Optimization*, **volume 4**, pp. 393–422 (1994).
- [11] A. BOURAS, V. FRAYSSE. “A relaxation strategy for the Arnoldi method in eigen problems.” Technical Report TR/PA/00/16, CERFACS (2000).