Initial Development of Depletion Capability in the GPU-Based Monte Carlo Code PRAGMA

Kyung Min Kim, Namjae Choi, Han Gyu Lee, Han Gyu Joo*
Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea
**Corresponding author: joohan@snu.ac.kr*

1. Introduction

Although the Monte Carlo (MC) method has the most attractive advantage of simulating neutrons in continuous phase space, it also has the most critical drawback that it requires excessive computational burden associated with using significant amount of particles to retain statistically reliable results, which prevents the MC method from being widely used in reactor design. The drawback had been resolved by relying on supercomputers employing a massive number of CPU cores, such as the Shift code [1] of the Oak Ridge National Laboratory (ORNL) or the OpenMC code [2] of MIT.

However, it is not practical in terms of real application since such computing systems are not easily accessible in most institutions. The advent of GPU and its spread in nuclear reactor physics field, nevertheless, enabled more researchers to achieve high computing performance in the MC simulation through its sophisticated parallelism, even with smaller computing systems. For the first time, WARP [3] developed at U.C. Berkeley demonstrated the potential of powerful GPU computing performance in continuous energy MC calculations, and the OpenMC and the Shift also started to exploit GPUs as in [4] and [5].

Motivated by this trend, PRAGMA [6], a GPU-based continuous energy MC code, has been developed at Seoul National University. It is capable of carrying out high-speed simulations with massive number of particles within cabinet-scale computing clusters, which can in turn produce high-fidelity results in practical time.

As a subsequent stage, we planned to realize whole-core cycle depletion calculations using PRAGMA. The rationales in the development of the PRAGMA depletion capability are introduced in this paper. The limitations imposed by the hardware memory size prohibit adopting conventional methodologies used in burnup calculations. Therefore, a couple of alternative methods are suggested and accuracy examinations on the suggested methods are presented.

2. Backgrounds and Limitations

2.1 Overview of PRAGMA Burnup Procedure

The PRAGMA burnup analysis module is composed of neutronics and depletion calculations in the same manner as the deterministic codes. During the neutronics calculations, one-group microscopic reaction rates, which gives the information of the production and the loss of nuclides, are tallied by the MC simulation. Then, they are plugged into the Bateman equation which is given as follows.

$$\frac{dN_i}{dt} = -(\lambda_i + \overline{\sigma_i \phi})N_i + \sum_j (\lambda_{ji} + \overline{\sigma_{ji}\phi})N_j \qquad (1)$$

where

N_i	Number density of nuclide *i*
λ_i	Decay constant of nuclide *i*
$\overline{\sigma_i \phi}$	Microscopic absorption rate of nuclide *i*
λ_{ji}	Decay yield of nuclide *i* produced by nuclide *j*
$\overline{\sigma_{ji}\phi}$	Microscopic reaction rate of nuclide *j* leading to creation of nuclide *i*

The Bateman equation above can be written as a linear system expressed as follows.

$$\frac{d\vec{N}(t)}{dt} = \mathbf{A}\vec{N}(t). \qquad (2)$$

After temporal discretization, the number density at time $t_0 + \Delta t$ can be written in terms of the matrix exponential.

$$\vec{N} = \exp(\mathbf{A}\Delta t)\vec{N}_0 \quad \text{where } \vec{N}_0 \equiv \vec{N}(t_0) \text{ and } \vec{N} \equiv \vec{N}(t_0 + \Delta t). \qquad (3)$$

This can be efficiently solved by the Chebyshev Rational Approximation Method (CRAM) [7], where the matrix exponential is converted to several sets of linear systems as follows.

$$\vec{N} = \alpha_0 \vec{N}_0 + 2\,\mathrm{Re}\left(\sum_j \alpha_j \left(\mathbf{A}\Delta t - \theta_j \mathbf{I} \right)^{-1} \vec{N}_0 \right) \qquad (4)$$

2.2 Limitations on the Reaction Rate Tally

There are two conventional approaches to tally the one-group microscopic reaction rates. One is so-called online tally used in McCARD [8], in which the reaction rates are accumulated in every neutron track for every nuclide of interest during the MC simulation. The other one, employed in the Shift code [9], is to accumulate flux spectra with very fine energy structure for all depletion regions, and produce the reaction rates after the MC simulation.

The former, however, has a critical shortcoming, in that the memory access burden is heavily imposed, especially on GPUs. Since the cross-section lookup takes most of the computing time on GPUs due to its random-

access nature, calculating the reaction rates for every neutron track severely degrades the performance.

In addition, both approaches require too large memory such that the practical-sized computing clusters cannot handle. To give an estimate of the memory requirement for an APR1400 core depletion problem, approximately 205 GB is required in the former and 3.5 TB for the latter, respectively. However, no more than 100 GB is allowed in our case and thus our objective is to devise a method using less than 100 GB.

3. Multilevel Spectral Collapse

Due to the physical limitation of the memory size, an alternative approach is suggested, which is referred to as multilevel spectral collapse (MSC). In this procedure, two types of flux spectra are tallied: finer one with tens of thousands energy groups for background geometries and coarser one with hundreds of energy groups for each unit depletion region. By using the fine flux spectra, the multigroup cross sections are produced. Then, using the cross sections and the coarse flux spectra, one-group reaction rates are generated, as illustrated in **Figure 1**.

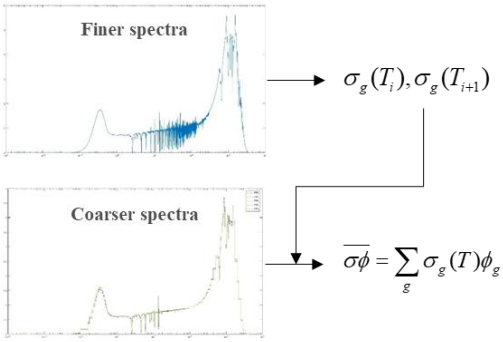

Figure 1. Procedure for multilevel spectral collapse.

3.1 Energy Group Structure

For the background geometries having fine spectra, since certain neutron reactions show sensitive behaviors on the spectral energy grids as illustrated in **Figure 2**, 56,000 equilethargic energy points are assigned.

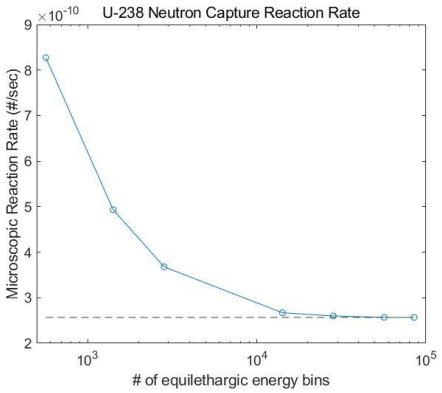

Figure 2. Reaction rate sensitivity on energy grids.

For the coarse flux spectra which are tailed for each depletion region, only a few hundred groups are allowed due to the memory limitation. For example, an APR1400 full-core problem has more than 11 million depletion regions, and the corresponding memory requirements are presented in **Table 1**.

Table 1. Memory requirements depending on the number of groups.

# of Energy Groups	Memory (GB)
100	8
500	41
1,000	82

Therefore, the number of energy groups is set to 500, but not purely equilethargic. The structure is extended from the standard SCALE [10] 252-group structure, in which additional energy bins are added in thermal energy region to capture the spatial self-shielding effect which is especially significant in burnable absorbers.

3.2 Adopting Background Geometry

There are two options in configuring the background geometry: assembly segment and fuel pin as illustrated in **Figure 3**. In the first option, an axial segment of assembly is set as the background geometry so that all the depletion regions in the segment share the same fine spectrum. The second option is to set a single fuel pin as the background geometry, and depletion regions in the three-dimensional fuel pin share the same fine spectrum. Both of the options are feasible in terms of the memory requirement for the core depletion as shown in **Table 2**.

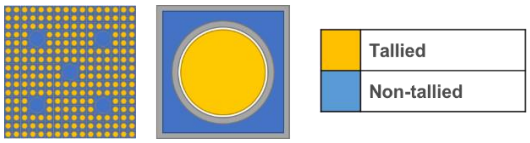

Figure 3. Assembly segment (left) and fuel pin (right).

Table 2. Memory requirements for the options.

Background Geometry	Memory (GB)
Assembly Segment	4
Fuel Pin	18

4. Accuracy Examinations

4.1 Validity of Employing the Fine Grid Spectra

Before investigating the MSC scheme, it should be preceded to confirm the validity of generating one-group reaction rates from the fine flux spectra. **Figure 4** shows the eigenvalue behaviors with burnup for three different types of fuel pins. The eigenvalues of PRAGMA agree with those of McCARD for all fuel pins within 30 pcm.

In addition, **Figure 5** shows the relative difference in group-wise fission reaction rates (McCARD generates three-group reaction rates) of fissiles between PRAGMA and McCARD for the 1.72% problem at BOC, where the maximum difference is only 0.1%. These results confirm the one-group reaction rates obtained from the fine flux spectra is equivalent to the directly tallied reaction rates.

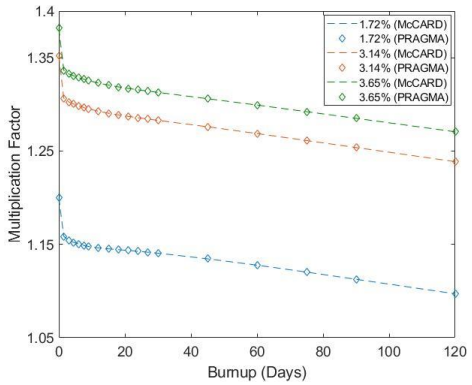

Figure 4. Comparison of single pin burnup problems

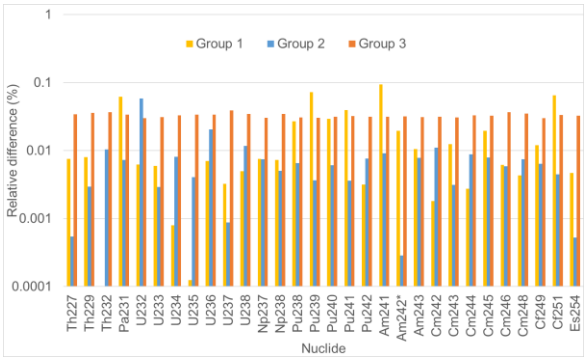

Figure 5. Relative difference in fission reaction rates

4.2 Application of MSC in 2D Assembly Problems

Five APR1400 2D assembly problems were examined: A0, B0 and C0 which do not include gadolinium, and B3 and C3 which have the largest amount of Gd contents. The references are using the fine flux spectra for each depletion region. **Figure 6** shows the eigenvalues of non-Gd assemblies with different geometry collapse schemes. The errors of the two algorithms from the reference are within 10 pcm. **Figure 7**, however, shows relatively high difference in the assembly-level MSC from the reference. This is contributed mostly by the fact that the fine flux spectrum, which is averaged over the assembly, is biased toward typical UO_2 spectra as illustrated as **Figure 8**. The biased spectrum generates the group-wise cross sections which are not properly weighted for the Gd fuel pins. Furthermore, this also makes it difficult to capture the spatial self-shielding effect in the burnable absorbers. On the other hand, it can be observed that the pin-level MSC is more accurate than the assembly-level MSC, since the error from the bias is eliminated. Hence, even though the assembly-level MSC uses much less memory, it is not flexible as it is inadequate for the problems involving high radial heterogeneity.

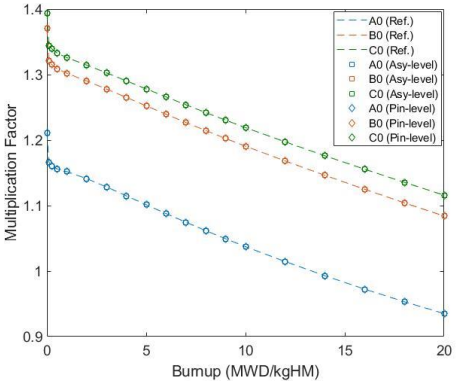

Figure 6. 2D non-Gd assembly results

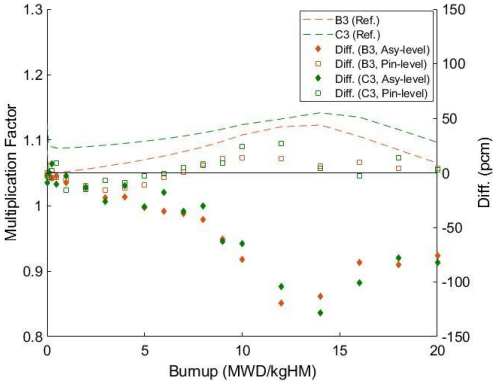

Figure 7. 2D Gd-bearing assembly results

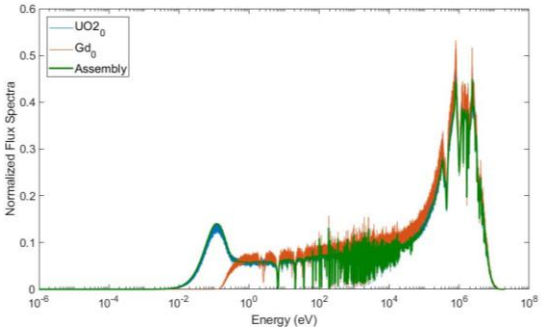

Figure 8. Spectra of Gd, UO_2 pins and assembly

4.2 Application of MSC in 3D Assembly Problems

Since the pin-level MSC takes average of the spectra in the axial direction where burnup variation exists, its 3D extension should be examined thoroughly. The test problems are the corresponding 3D assemblies with spacer grid modelled. Due to axial instability originated from xenon oscillation, xenon equilibrium is applied.

Since the differences in non-Gd assemblies at each burnup point are within 10 pcm as illustrated in **Figure 9**, it is revealed that the degree of axial heterogeneity by

the burnup variation can be neglected in generating one-group reaction rates. In addition, **Figure 10** shows that the spatial self-shielding effect of the burnable absorbers can be captured with the pin-level MSC. Furthermore, **Figure 11** presents the maximum axial power difference at each burnup in the C3 assembly for two cases: one between the pin-level MSC case and the reference case and the other between two independent reference cases. It confirms that the degree of error in the pin-level MSC is merely the uncertainty level.

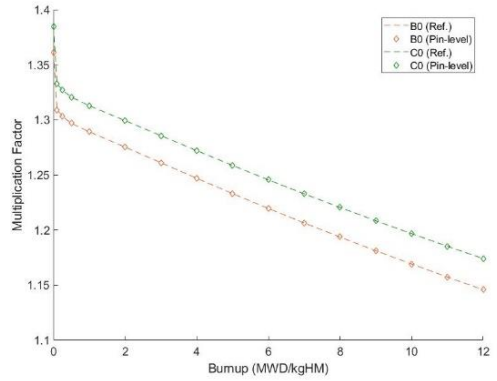

Figure 9. 3D non-Gd assembly results

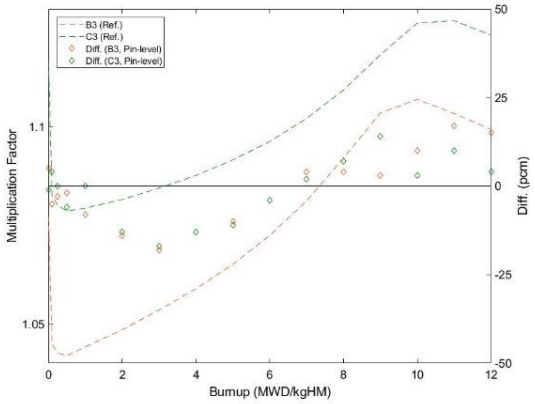

Figure 10. 3D Gd-bearing assembly results

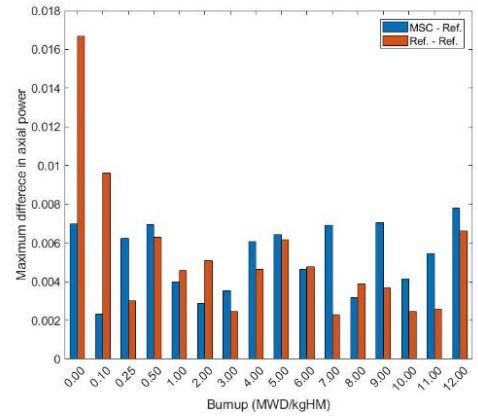

Figure 11. Axial power difference in 3D C3 assembly

5. Conclusion and Future Work

The initial development for realizing whole-core cycle depletion of PRAGMA is completed. Although it was impossible to adopt the conventional methods in reaction rate tally due to the limited memory of GPUs, sufficient accuracy was achieved by the multilevel spectral collapse scheme. For the background, the pin-level collapse showed a decent performance in accuracy for both 2D and 3D, non-Gd and Gd assemblies.

Meanwhile, there is a persistent instability in 3D problems despite xenon equilibrium. It is triggered by the MC uncertainty which causes axial power oscillation in later burnup stages. This issue will be resolved by full predictor-corrector scheme and sub-step methods.

As the future work, online reaction rate tally for some nuclides will be introduced to further enhance the accuracy, especially for problems involving burnable absorbers, and the shadow effect caused by the control rod insertion will be examined. In addition, porting of the whole depletion procedure to GPUs is being considered to reduce the computing burden of generating one-group reaction rates and solving the Bateman equation. Finally, several ways to reduce the memory consumption will be devised to retain sufficient memory margin in the whole-core calculations.

ACKNOWLEDGEMENTS

This work was supported by KOREA HYDRO & NUCLEAR POWER CO., LTD (No. 2018-Tech-09)

REFERENCES

[1] T. Pandya *et al*., "Implementation, Capabilities, and Benchmarking of Shift, a Massively Parallel Monte Carlo Radiation Transport Code," Journal of Computational Physics, 308, pp. 273-272 (2016).
[2] P. Romano and B. Forget, "The OpenMC Monte Carlo Particle Transport Code," Annals of Nuclear Energy, 51, pp. 274-281 (2013).
[3] R. Bergmann, "The Development of WARP – A Framework for Continuous Energy Monte Carlo Transport in General 3D Geometries on GPUs," Ph.D. Dissertation, University of California, Berkeley (2014).
[4] J. L. Salmon and S. M. Smith, "Exploiting Hardware-Accelerated Ray Tracing for Monte Carlo Particle Transport with OpenMC," IEEE/ACM Performance Modeling, Benchmarking and Simulation of High Performance Computer System, Denver, CO, USA, Nov. 18 (2019).
[5] S. P. Hamilton and T. M. Evans, "Continuous Monte Carlo neutron transport on GPUs in the Shift Code," Annals of Nuclear Energy, 128, pp. 236-247 (2019).
[6] N. Choi and H. G. Joo, "Domain Decomposition for GPU-Based Continuous Energy Monte Carlo Power Reactor Calculation," Nuclear Engineering and Technology, 52(11), pp. 2667-2677 (2020).
[7] M. Pusa, "Higher-Order Chebyshev Rational Approximation Method and Application to Burnup Equations," Nuclear Science and Engineering, 182, pp. 297-318 (2016).
[8] H. J. Shim *et al.*, "McCARD: Monte Carlo Code for Advanced Reactor Design and Analysis," Nuclear Engineering and Technology, 44(2), pp. 161-176 (2012).
[9] G. G. Davidson *et al.*, "Nuclide Depletion Capabilities in the Shift Monte Carlo Code," Annals of Nuclear Energy, 114, pp. 259-276 (2018).
[10] B. T. Rearden and M. A. Jessee, "SCALE Code System 6.2.1," ORNL/TM-2016/352, Oak Ridge National Laboratory, USA (2016)