Correlations in Monte Carlo Eigenvalue Simulations: Uncertainty Quantification, Prediction and Reduction

by

Jilang Miao

B.S, University of Science and Technology of China (2013)

Submitted to the Department of Nuclear Science and Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Nuclear Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2018

Author .....
Department of Nuclear Science and Engineering
Aug 10, 2018

Certified by .....
Benoit Forget, Ph.D.
Associate Professor of Nuclear Science and Engineering
Thesis Supervisor

Certified by .....
Kord S. Smith, Ph.D.
KEPCO Professor of the Practice of Nuclear Science and Engineering
Thesis Supervisor

Accepted by .....
Ju Li, Ph.D.
Battelle Energy Alliance Professor of Nuclear Science and Engineering
Professor of Materials Science and Engineering
Chairman, Department Committee on Graduate Students

Correlations in Monte Carlo Eigenvalue Simulations: Uncertainty Quantification, Prediction and Reduction

by

Jilang Miao

Submitted to the Department of Nuclear Science and Engineering
on Aug 10, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Nuclear Science and Engineering

Abstract

Monte Carlo methods have mostly been used as a benchmark tool for other transport and diffusion methods in nuclear reactor analysis. One important feature of Monte Carlo calculations is the report of the variance of the estimators as a measure of uncertainty. In the current production codes, the assumption of independence of neutron generations in Monte Carlo eigenvalue simulations leads to the oversimplified estimate of the uncertainty of tallies. The correlation of tallies between neutron generations can make reported uncertainty underestimated by a factor of 8 in assembly size tallies in a typical LWR.

This work analyzes the variance/uncertainty convergence rate in Monte Carlo eigenvalue simulations and develops different methods to properly report the variance. To correct the underestimated variance as a post-processing step, a simple correction factor can be calculated from the correlation coefficients estimated from a sufficient number of active generations and fitted to decreasing exponentials. If the variance convergence rate is needed before or during the simulation to optimize the run strategy (number of generations and neutrons per generation), a discrete model can be constructed from the inactive generations that can predict the correlation behavior of the original problem. Since it is not efficient to perform variance correction to all tallies on all problems, a simple correlation indicator is also developed to quickly determine the potential impact of correlations on a given tally in a given problem. This can help decide if more complicated correction analysis or the use of independent simulations should be used to calculate the true variance.

Run strategy to reduce correlations is also investigated by introducing the notion of delayed neutrons. A predictive model for the new source update scheme was developed to help identify optimal delayed neutron parameters before implementing in OpenMC. Optimal run strategies in terms of delayed bank size, frequency of delayed bank sampling and true simulation costs are proposed.

Thesis Supervisor: Benoit Forget, Ph.D.

Title: Associate Professor of Nuclear Science and Engineering

Thesis Supervisor: Kord S. Smith, Ph.D.
Title: KEPCO Professor of the Practice of Nuclear Science and Engineering

Acknowledgments

The author would like to express his sincere gratitude to thesis advisors Professors Benoit Forget and Kord Smith, who not only suggested this problem and provided him with the constant guidance and ever-ready help so kindly and generously, but also gave him much needed encouragement throughout the course of this work.

This research was supported in part by the Consortium for Advanced Simulation of Light Water Reactors (CASL), an Energy Innovation Hub for Modeling and Simulation of Nuclear Reactors under U.S. Department of Energy Contract No. DE-AC05-00OR22725.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Numerical modeling of neutron physics in nuclear reactors has long played an important role in the nuclear industry to ensure the safety and reliability of the operating fleet of nuclear reactors and to evaluate the technical capabilities of advanced designs. Monte Carlo methods have long been considered a reference for neutron transport simulations since they make very few approximations in simulating the random walk of neutrons in a system. Monte Carlo simulations permit an accurate representation of the geometry and energy dependence of reaction cross sections and thus provide high-fidelity core spectral calculations.

Previous work has observed correlation effects between neutrons in systems with fission, particularly when performing eigenvalue simulations based on the power iteration [3] [22] [17] [42]. In the presence of neutron correlation, Monte Carlo simulation will typically underestimate variance and thus report inaccurate confidence intervals. Correlation will also reduce the tally convergence rate thus requiring many additional generations of neutrons in order to obtain statistically accurate results.

1.1 Background and Literature Review

The general procedure to solve a Monte Carlo eigenvalue problem is described here to introduce the unfamiliar reader to this concept. Individual neutron histories are tracked through an heterogeneous system and information is gathered as they cross

boundaries and collide with nuclides. At the level of the neutron history, a neutron is tracked from birth (source site) to death (absorption or leakage) by sampling path lengths and collision events based on neutron and material properties. In the eigenvalue problem, a number of such neutron histories are grouped into a Fission Source Generation (FSG). In a FSG, when a neutron is terminated by absorption and a fission event is sampled from the cross section data, the fission source site is accumulated into a fission bank for the next FSG. After all neutrons of a generation have been tracked, normalization of the ensuing fission bank is performed to avoid neutron number explosion or extinction. After normalization, a new FSG starts from the normalized fission bank. Once a sufficient number of FSGs (i.e. inactive generations) are performed to achieve a stationary distribution (i.e. distribution follows the fundamental mode), additional FSGs (i.e. active generations) are performed to score tallies of interest. Tallies from all active generations form a sample from which estimate parameters such as k_{eff}, flux, reaction rates and their associated variance can be estimated The procedure is schematically plotted in Fig 1-1.

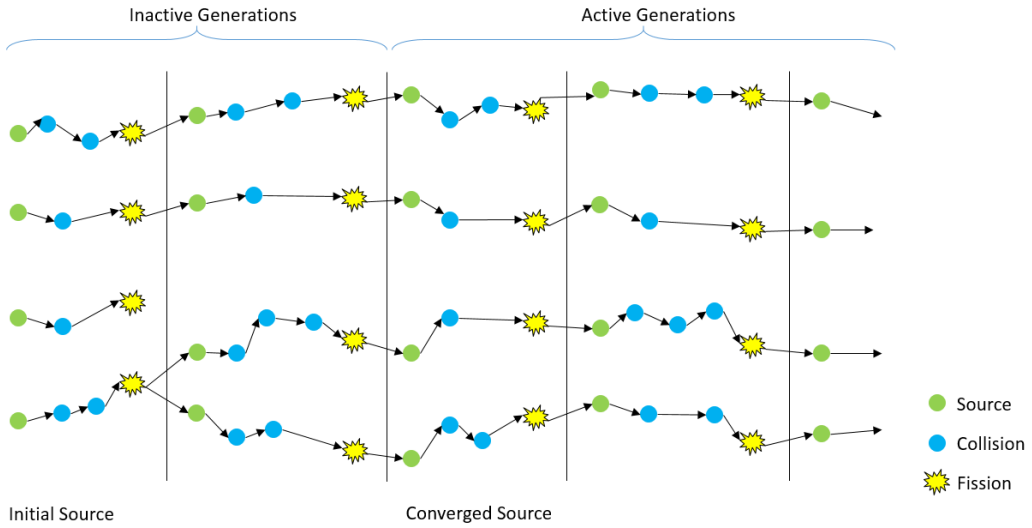


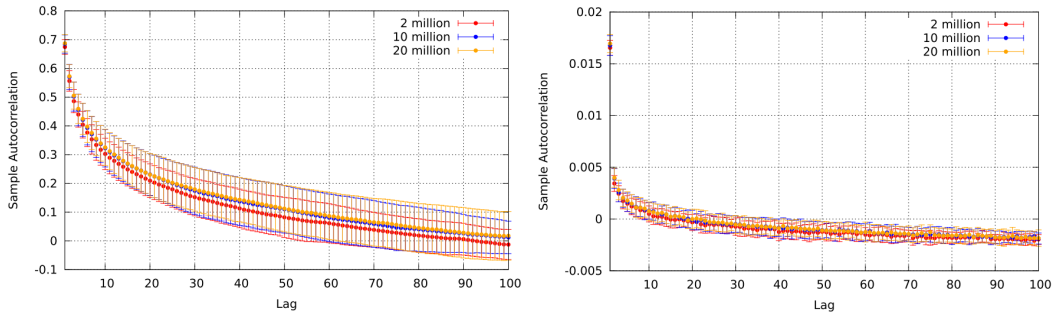
Figure 1-1: Monte Carlo Eigenvalue Problem Calculation Procedure

If the FGS's are independent, the variance of the commonly used estimators decrease inversely proportional to the number of generations (the $1/N$ rate). Further if the estimators are unbiased (bias small enough to be neglected when using a large gen-

eration size [3]), the expected square error decreases with the same $1/N$ rate. Previous work has observed correlation effects between neutrons in systems with fission, particularly when performing eigenvalue simulations based on the power iteration [22] [21].

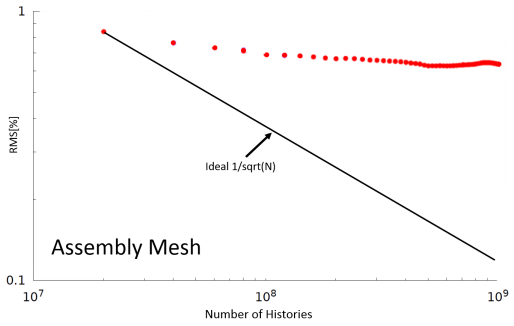
Fig 1-2(c) shows the impact of generation-to-generation correlation on the convergence rate of a tally on an assembly size mesh for the 2D BEAVRS benchmark [22] [23]. In the presence of correlation, additional generations of neutrons will only slightly improve the sample mean and may not be worth the additional effort. In Monte Carlo eigenvalue simulations, quantities of interests are usually estimated as an average over many generations once a stationary fission source is obtained. The sample mean is reported accompanied by an estimate of the variance indicating the level of statistical uncertainty of the simulation and allowing for the definition of confidence intervals. If generation-to-generation correlation is neglected, the variance of the estimator will be evaluated as the variance of the sample of generation tallies divided by N, the generation number. This underestimates the variance of the estimator and will report inaccurate confidence intervals.

Neglecting this correlation effect results in an underestimate of uncertainty reported by Monte Carlo calculations, the magnitude of which depends on the dominance ratio of the problem and the size of the phase space being tallied [46]. The simple approach to avoid underestimation of the variance is to perform multiple independent simulations with different initial random seeds. This will lead to good variance estimates but requires lots of additional work since each independent simulation needs an independent fission source and will not improve convergence rates [22]. Another approach is to tally neutrons from batches (one batch contains more than one fission source generations)[13]. This is based on the observation that generation-to-generation correlations decrease as a function of generation lag, which leads to lower correlation between tallies from simulation batches than those from generations. The many-generations-in-one-batch method provides a better estimate of the variance with a single simulation at the cost of many more generations. The superhistory method [3] tracks each neutron through more than one generations (instead of just one) before any fissions are banked for the next batch. This is similar to the



(a) Autocorrelation coefficients assembly-size tally

(b) Autocorrelation coefficients pin-size tally



(c) RMS Convergence assembly-size tally

Figure 1-2: Plots of autocorrelation coefficients and convergence rate of assembly-size tally on 2D BEAVRS. Generation-to-generation correlation makes statistical error deviate from ideal convergence rate. (From "Monte Carlo and Thermal Hydraulic Coupling using Low-Order Nonlinear Diffusion Acceleration" by B. Herman, 2014 [22])

many-generation-in-one-batch method except that tallies are only performed on the last generation in each batch. Wielandt's method, originally developed for deterministic nuclear reactor calculations [7] [37] was applied to Monte Carlo eigenvalue calculation [6] by tacking some neutron for more than one generations in one batch. Wielandt shift method reduces correlation by effectively solving a different problem with lower eigenvalues. This method requires an estimate of the upper bound of the eigenvalue. Significant reduction requires a compact estimate of the upper bound.

The generation-to-generation correlation of tally X between generation n and $n+k$ can be directly characterized by the covariance matrix of a tally from different

generations.

$$\rho(n, n+k) = \frac{\mathrm{Cov}[X(n), X(n+k)]}{\sqrt{\mathrm{Var}[X(n)]\,\mathrm{Var}[X(n+k)]}} \tag{1.1.1}$$

When the generation tally values can be viewed as stationary time series, the covariance between generations only depends on generation distance or lag [47] [22].

$$\rho(n, n+k) = \rho(n; k) = \rho(k) \tag{1.1.2}$$

In recent years, more effort has been dedicated at evaluating the auto-correlation coefficients (ACC) observed in the Monte Carlo eigenvalue simulations. Recent work by Herman [21] [22] extensively demonstrated the impact of generation-to-generation correlations on a realistic 2D full core PWR benchmark. This work also numerically showed the dependency on the mesh size as well as the insensitivity to the number of neutrons per generation (Figure 1-2(a) and Figure 1-2(b)).

Previous work has also proposed methods to predict the underestimation ratio between the correlated and uncorrelated estimates. The investigated methods can be classified into two categories. The first performs data fitting on simulation outputs to capture the correlation. The second directly computes covariance between the Monte Carlo generations based on corresponding approximations of the Monte Carlo simulation [44]. Demaret et al [14] fitted AR (auto-regressive) and MA (Moving Average) models to the results of Monte Carlo eigenvalue calculations and used the AR and MA models to give variance estimator of k_{eff}. Yamamoto et al [49] [50] expanded the fission source distribution with diffusion equation modes, performed numerical simulation of the AR process of the expansion coefficients and used the correlation of the AR process to predict the Monte Carlo eigenvalue simulation. Approximating the original neutron transport problem with a diffusion problem lead to a non-negligible lost in accuracy. Inspired by Brissenden and Garlick's work [3], Sutton [42] applied the discretized phase space (DPS) approach to predict the underestimation ratio but the method cannot predict the ratio when one neutron generates offspring in different

phase space regions or generates a random number of offspring. Ueki [45] developed variance estimator with orthonormally weighted standardized time series (OWSTS). The estimator is based on the convergence of step-wise interpolation of standardized tallies (SIST) to a Brownian bridge. SIST weighted by a trigonometric family of weighting functions gives a new statistical estimator for the variance. Asymptotic behavior of the expectation of the new estimator leads to a variance estimator that is not affected by the correlation effect thus converging at a 1/N rate. Numerical results showed that the variance can be accurately estimated after approximately $5000 \sim 10000$ active generations for problems with non-negligible auto-correlation up to 100 generation lags (as observed in a typical LWR). A convergence diagnosis is also necessary to determine when the asymptotic behavior is reached. However, the convergence diagnosis cannot be implemented on-the-fly.

In summary, among the current methods to reduce the impact of correlations, the methods that decorrelate batches become very costly due to an increase in the number of generations. The methods that correct the variance underestimation using post-processing techniques are very memory intensive since tallies of all active generations need to be stored. Finally, the methods that predict the correlation coefficients and thus correct the variance with approximate models do not capture a sufficient level of accuracy.

Previous work also proposes quantitative explanations of the generation-to-generation of tallies. Dumonteil et al [18] attributed the generation-to-generation correlation to spatial correlation of the fission process and ensuing asymmetry between neutron creation and annihilation in a stochastic branching process. Houchmandzadeh [24] and Zoia [51] investigated the issue of neutral particle clustering, the phenomenon where particles without interaction can concentrate locally. In simple homogeneous problems, spatial covariance between particle number density at different positions proves existence of clustering theoretically and thus explains the existence of generation-to-generation correlation of tallies contributed by these clustered particles.

1.2 Thesis Objectives

The subject matter of this thesis is organized along three main themes:

- Correlation Analysis – Quantify correlation coefficients in different problems, analyze the effect of correlations on variance estimation, analyze the dependency of correlation on problems and simulation parameters. The goal is to understand the nature of the correlations and determine appropriate variance estimators if the correlations are known.

- Correlation prediction – Develop models to predict correlation coefficients as a function of tally position, region and simulation parameters accurately without storing all tallies. The goal is to develop a simple deterministic model that can predict correlations prior to performing the bulk of the simulation thus providing an accurate estimate of the variance correction needed and an estimate as to when the desired level of accuracy will be reached.

- Run Strategy in the presence of correlations – Allocate computation resources in optimal ways to reach better results given cost requirement. The goal is to develop a simple diagnosis indicating the need for corrective action, and once identified, provide an optimal run strategy to achieve the target accuracy faster.

1.3 Thesis Outline

Chapter 2 discusses the correlation coefficients of tallies in Monte Carlo simulations and their impact on the underestimation of variances. Behavior of correlation coefficients and variance underestimation ratios are analyzed in different situations. A post-processing method is proposed by assuming the correlation coefficients $\rho(n;k)$ do not depend on generation n and decrease exponentially as a function of the generation lag k. This method estimates correlation coefficients from all tallies and fits the estimated correlation coefficients to exponentials. The fitted correlation coefficients can accurately predict variance of accumulated tallies at any generation n. The fitted

correlation coefficients also explain the particular features of the variance convergence rates observed numerically.

Chapter 3 focuses on predicting correlation coefficients before starting the active generations of the simulation. This provides not only correction to the underestimated variance but also provides an estimate on when a target accuracy will be reached. A Markov Chain Monte Carlo model is developed to show that the dependence of neutron source bank between consecutive generations contributes a significant fraction of the correlation of tallies between generations but also misses an important part caused by the multiplying effect of fission. The method of multitype branching process (MBP) is developed to capture both the source bank dependence and the correlation due to multiplication. The MBP model is exact in predicting correlations for neutrons transported in a discrete phase space and provides acceptable accuracy in continuous problems by constructing an approximate discrete problem from tallies. Since the MBP model is a discrete approximation, it can predict correlations and variances on tally meshes coarser than the discretized model mesh accurately.

Chapter 4 provides a prediction model of the correlation level at any mesh size without the requirement of high accuracy discrete tallies in order to quickly determine if variance correction is needed. This chapter generalizes the MBP model to a continuous space and solves correlation coefficients (and thus variance correction factors) exactly in a homogeneous problems. The real heterogeneous problem can then be homogenized while imposing preservation of the neutron migration area. The correlation coefficients and asymptotic variance underestimation ratio of the homogenized problem can give a rough estimate of those in the real problem.

Chapter 5 investigates how to modify the current Monte Carlo eigenvalue simulations to reduce the impact of correlations and improve the variance convergence rate in the active generations. For a given active generation size, more neutrons are simulated in the inactive generations to enable a split of the source bank into two parts, neutrons in one part are transported in every active generation (i.e. prompt bank) and contribute to tallies, neutrons in the other bank (i.e. delayed bank) are carried forward until included in the prompt at later generations. The delayed neutron

method provides better variance behavior at the cost of more work in the inactive generations. A predictive model for the delayed scheme is also developed to help identify an optimal run strategy and provide an estimate of the variance correction.

Chapter 2

Estimating Correlation and Correcting Variance

2.1 Problems for Illustration

Since performing meaningful analysis on a large problem is inherently difficult, this section defines two simple problems to illustrate the impact of correlations and provide references for the predictive methods developed in this work.

The first problem tracks neutrons in discrete phase space regions and has an analytic solution that can easily be used to evaluate spatial and temporal moments, and correlation coefficients of tallies.

The problem is also characterized by simple parameters which facilitates the interpretation of the observed correlation induced behavior. This test problem will be referred to as $TD1$(discrete test problem) in the following chapters.

The second problem tracks neutrons in a simple homogeneous reflective cube with constant (one group) cross section. This problem has an exact solution for flux and reaction rates. The geometry and material properties of the cube are selected to be representative of and exhibit similar behavior to the full core PWR by matching the migration area of neutrons of a typical LWR.

This test problem will be referred to as $TC1$(continuous test problem) in the following chapters.

2.1.1 Neutrons in Discrete Phase Space

In the first problem, neutrons are being transferred among M phase space regions. The process is specified by a Markov Chain transfer matrix P, with $P_{i,j}$ being the probability that a neutron at region i will be in region j in the next generation and a random variable ξ_i representing new neutrons per fission at region i with $p_{i,a}$ being the probability that a fission at region i produces a new neutrons.

In OpenMC [39], like in many other Monte Carlo codes, the fission process is simulated by first evaluating the expected number of neutrons per fission

$$\nu \equiv \mathbb{E}[\xi \,|\xi \neq 0] \tag{2.1.1}$$

where ξ denotes the random variable of new neutrons per absorption and adding the condition $\xi \neq 0$ makes the conditional expectation equal to the expected number of new neutrons per fission.

Then sample $\lceil \nu \rceil$ (the closest integer that is larger than or equal to ν) and $\lfloor \nu \rfloor$ (the closest integer that is smaller than or equal to ν) with appropriate probability to satisfy the constraint on expectation of ξ (Eq 2.1.1).

That is, after a neutron is determined to induce a fission event and the expected number of new neutrons per fission is determined to be ν, the number of new neutrons is sampled to be $\lceil \nu \rceil$ with probability $p'_{\lceil \nu \rceil}$ and $\lfloor \nu \rfloor$ with probability $p'_{\lfloor \nu \rfloor}$ such that

$$\lfloor \nu \rfloor p'_{\lfloor \nu \rfloor} + \lceil \nu \rceil p'_{\lceil \nu \rceil} = \nu \tag{2.1.2}$$

In the simple test problem, it is not necessary to simulate the neutrons as above by first determining whether the neutron induces a capture or fission event then determining the number of new neutrons per fission. The test problem specifies the probability of number of new neutrons per absorption directly. Then the distribution

of ξ_i can be solved from the below equations

$$\begin{aligned} \sum_a p_{i,a} &= 1 \\ \sum_a a\ p_{i,a} &= \mathbb{E}\xi \\ \sum_a a\ \frac{p_{i,a}}{1-p_{i,0}} &= \nu_i \end{aligned} \tag{2.1.3}$$

If a further assumption is made to make the system critical, that is $\mathbb{E}\xi = 1$. The probabilities are determined as

$$\begin{aligned} p_{i,0} &= 1 - \frac{1}{\nu_i} \\ p_{i,\lfloor\nu\rfloor} &= \frac{\lceil\nu\rceil - \nu}{\nu} \\ p_{i,\lceil\nu\rceil} &= \frac{\nu - \lfloor\nu\rfloor}{\nu} \\ p_{i,a} &= 0 \quad (otherwise) \end{aligned} \tag{2.1.4}$$

where the material is assumed to be fissile and the equations above will only be applied to test problems with fissile material.

The numerical values of the matrix P are selected such that

$$\begin{aligned} P_{i,i} &= 0.5 \\ P_{i,i-1} = P_{i,i+1} &= 0.25 \quad (i > 1, i < M) \\ P_{1,2} = P_{M,M-1} &= 0.5 \\ P_{i,j} &= 0 \quad (otherwise) \end{aligned} \tag{2.1.5}$$

where the M phase space regions are indexed from 1 to M.

It can be shown that the Markov matrix specified according to Eq 2.1.5 has the normalized equilibrium distribution π with

Table 2.1: Parameters of demonstration problem

Geometry		ν	Macro Cross-Section				k_{eff}
Boundary	Width(cm)		$\Sigma_s(cm^{-1})$	$\Sigma_c(cm^{-1})$	$\Sigma_f(cm^{-1})$	$\Sigma_t(cm^{-1})$	
Reflective	400	2.45	0.270	0.018	0.012	0.300	1

$$\begin{aligned} \pi_i &= \frac{1}{M-1} \quad (i = 2, \cdots, M-1) \\ \pi_i &= \frac{1}{2(M-1)} \quad (i = 1, M) \end{aligned} \tag{2.1.6}$$

For any given neutron at region i, the destination region j is sampled according to $P_{i,j}$. Then for the neutron absorbed at region j, the number of new neutrons a is sampled according to $p_{j,a}$. The a neutrons in region j will then be transferred to new locations as described above.

2.1.2 Neutrons in a Homogeneous Cube

This problem is designed to mimic the correlation behavior of real PWR neutron eigenvalue simulation. Analyzing correlation coefficients on a full core realistic problem is a very costly endeavor. Herman et al [22] were able to compute such coefficients on the 2D BEAVRS benchmark using extensive computational time making any substantive analysis impractical. In order to accelerate the process, a simple benchmark was developed that preserves the correlation effects, reduces run time and has a simple analytical reference solution of neutron distribution. Parameters of the homogenized cubic reactor are given in Table 2.1. The simple benchmark was chosen as a $400cm$ reflective cube since it has dimensionality similar to that of a full core PWR. Additionally, the one group cross sections were selected such that the system is critical and preserves the migration length of neutrons.

In this problem, neutrons are not transferred among discrete states but instead within a continuous phase space in a realistic way. The transfer probability density

for a collision from position x to x' is defined as

$$P(x, x') = \Sigma_t e^{-\Sigma_t |x - x'|} \tag{2.1.7}$$

The probability mass function of new neutrons after an absorption event at any position is

$$\begin{aligned} p_0 &= \frac{\Sigma_c}{\Sigma_a} \\ p_1 &= 0 \\ p_2 &= \frac{\Sigma_f}{\Sigma_a} 0.55 \\ p_3 &= \frac{\Sigma_f}{\Sigma_a} 0.45 \\ \nu &= 0.55 \times 2 + 0.45 \times 3 \end{aligned} \tag{2.1.8}$$

where p_0 is the probability of neutron capture, p_2 and p_3 represent a probability for producing a discrete number of neutrons from fission.

The stochastic behavior of fission events in the simple benchmark can be evaluated analytically. For a tally region m occupying η_m fraction of volume of the whole system, the expected number of fission events Z_m induced by s independent neutrons is

$$\langle Z_m \rangle = s \frac{\Sigma_f}{\Sigma_t - \Sigma_s} \eta_m \tag{2.1.9}$$

where Σ_t, Σ_f, Σ_s are the total cross section, fission cross section, scattering cross section of the cube respectively. And the variance of Z_m is

$$\mathrm{Var}[Z_m] = s \frac{\Sigma_f \eta_m}{\Sigma_a} \left(1 - \frac{\Sigma_f \eta_m}{\Sigma_a} \right) \tag{2.1.10}$$

Derivation of the expectation and variance of Z_m is shown in Appendix A1.1.

To analyze this benchmark a simple Monte Carlo code was developed on a GPU to accelerate the analysis. The number of neutrons at each generation is equal to the

number of threads (number of blocks times number of threads per block) launched.

Each thread has a local collision tally in each spatial bin and a reduction algorithm is performed after each generation to obtain the global tally. When the generation size (number of neutrons per generation) exceeds the number of threads on the GPU, kernels are launched sequentially. Generation sizes are selected to the power of 2 for more efficient use of the GPU hardware. The fission source distribution is then obtained by multiplying the collision source distribution by the constant factor $\frac{\Sigma_f}{\Sigma_t}$.

2.2 Background

2.2.1 Variance of correlated sample average

In Monte Carlo eigenvalue simulations, quantities of interests are usually estimated as an average over many generations once a stationary fission source is obtained. The sample mean is reported accompanied by an estimate of the variance indicating the level of statistical uncertainty of the simulation and allowing for the definition of confidence intervals.

2.2.1.1 Correlation coefficients in variance underestimation

For tally X, by definition, the variance of the sample mean is given by

$$\begin{aligned} \operatorname{Var}\left[\bar{X}(N)\right] &= \operatorname{Var}\left[\frac{\sum_{n=1}^{N} X(n)}{N}\right] = \frac{1}{N^2}\operatorname{Var}\left[\sum_{n=1}^{N} X(n)\right] \\ &= \frac{1}{N^2}\mathbb{E}\left[\left(\sum_{n=1}^{N} X(n) - \mathbb{E}\left[\sum_{n=1}^{N} X(n)\right]\right)^2\right] \end{aligned} \tag{2.2.1}$$

where $X(n)$ is the result obtained from generation n and N is the total number of active generations.

The commutability of expectation and sum decomposes the variance into two

groups of terms

$$\begin{aligned}\mathrm{Var}\left[\bar{X}(N)\right] &= \frac{1}{N^2}\left(\sum_n \mathbb{E}\left[(X(n)-\mathbb{E}[X(n)])^2\right] + \sum_{n\neq j}\mathbb{E}\left[(X(n)-\mathbb{E}[X(n)])(X(j)-\mathbb{E}[X(j)])\right]\right) \\ &= \frac{1}{N^2}\left(N\sigma^2 + 2\sum_{n<j}\mathrm{Cov}\left[X(n),X(j)\right]\right)\end{aligned} \tag{2.2.2}$$

where $\{X(n)\}_{n=1}^{N}$ is assumed to have identical distribution and σ^2 is the variance of $X(n)$ for any n.

$$\sigma^2 = \mathrm{Var}[X(n)] \quad \forall n \tag{2.2.3}$$

If further assumptions are made that the generations are uncorrelated, as often assumed in Monte Carlo simulations, the covariance terms would vanish and the variance of the sample mean would be

$$\mathrm{Var}[\bar{X}(N)] = \frac{\sigma^2}{N} \tag{2.2.4}$$

In the presence of correlation, since the generation tally values can be modeled as stationary time series, the covariance between two batches only depends on batch distance or lag, k, [22] [47]

$$\begin{aligned}\mathrm{Var}\left[\bar{X}(N)\right] &= \frac{1}{N^2}\left(N\sigma^2 + 2\sum_{k=1}^{N-1}\sum_{i=1}^{N-k}\mathrm{Cov}[X(i),X(i+k)]\right) \\ &= \frac{1}{N^2}\left(N\sigma^2 + 2\sum_{k=1}^{N-1}(N-k)\,\mathrm{Cov}[X(i),X(i+k)]\right)\end{aligned} \tag{2.2.5}$$

where the first equality restructures the summation over indices i, j into $i, j = i + k$ with k being the generation lag, the second equality recognizes the $N - k$ identical covariance terms $\mathrm{Cov}[X(i), X(i + k)]$ based on the stationarity assumption.

Taking into account the covariance between generations, the variance of the sample mean of observable X can be evaluated with Eq 2.2.6, where σ^2 is the variance of

X, N is the number of active generations, and $\rho(k)$ is the autocorrelation coefficient between X's with generation lag k.

$$\mathrm{Var}[\bar{X}] = \frac{\sigma^2}{N}\left(1 + 2\sum_{k=1}^{N-1}\left(1 - \frac{k}{N}\right)\rho(k)\right) \tag{2.2.6}$$

$$\rho(k) = \rho(n;k) = \rho(n, n+k) = \frac{\mathrm{Cov}[X(n), X(n+k)]}{\sqrt{\mathrm{Var}[X(n)]\,\mathrm{Var}[X(n+k)]}} \tag{2.2.7}$$

It is worthwhile to note that $\rho(k)$ should have been more naturally written as $\mathrm{Cov}[X(i), X(i+k)]/\sigma^2$ following the derivation above. σ^2 is expanded back into $\sqrt{\mathrm{Var}[X(i)]\,\mathrm{Var}[X(i+k)]}$ with the stationarity assumption in order to make $\rho(k)$ match the definition of the Pearson correlation coefficient [9].

As can be seen from Eq 2.2.6, if the variance of the sample formed by the generation results $\{X(1), \cdots, X(N)\}$ is divided by the total number of active generations N to estimate the variance of their average, the variance of the estimator is underestimated by a factor $r(N)$ defined as

$$r(N) \equiv \frac{\mathrm{Var}[\bar{X}(N)]}{\mathrm{Var}[X]} = 1 + 2\sum_{k=1}^{N-1}\left(1 - \frac{k}{N}\right)\rho(k) \tag{2.2.8}$$

where $r(N) < N$ unless $\rho(k) = 1$ for all k.

2.2.1.2 Correlation coefficients in predicting RMS

When estimating observable X (with expectation $\langle X \rangle$) by estimator $\hat{X}$, it can be shown that the expected square error (ESE) equals the variance of the variable plus the bias of the estimator.

$$ESE = \mathbb{E}[(\hat{X} - \langle X \rangle)^2] = \mathrm{Var}[\hat{X}] + (\mathbb{E}[\hat{X}] - \langle X \rangle)^2 \tag{2.2.9}$$

If the estimator is unbiased, the expected square error becomes

$$ESE = \mathrm{Var}[\hat{X}] \tag{2.2.10}$$

Eq 2.2.10 shows the equivalence between square error and variance of estimator. Therefore, the error and convergence rate of the estimator (the sample mean) can be predicted by its ESE or variance. The ESE can be used as a predictor for the RMS as follows.

First, from the above unbiasedness assumption, for any tally in region m, the expectation of square error is equal to the variance

$$ESE_m = \mathrm{Var}[\hat{X}_m] \tag{2.2.11}$$

One common metric of interest is the relative square error (RSE),

$$RSE_m = \frac{(\hat{X}_m - \langle X_m \rangle)^2}{\langle X_m \rangle^2} \tag{2.2.12}$$

whose expectation is related to the variance as before

$$E[RSE_m] = \frac{ESE_m}{\langle X_m \rangle^2} = \frac{\mathrm{Var}[\hat{X}_m]}{\langle X_m \rangle^2} \tag{2.2.13}$$

This work focuses on predicting variance, and the predicted variance is used as an approximation of RSE.

$$RSE_{m,predict} = \frac{\mathrm{Var}[\hat{X}_m]}{\langle X_m \rangle^2} \tag{2.2.14}$$

Finally, a global metric of error is constructed as spatially averaging the RSE_m for each region.

$$RMS = \sqrt{\frac{\sum_{m=1}^{M} RSE_m}{M}} \tag{2.2.15}$$

where the subscript m denotes index of a tally region, M is the total number of tally regions; RMS stands for Root Mean Square error, since it is the square root of average of the relative square error over all tally regions. RMS synthesizes the relative square error of all tally regions into a one number metric.

Similarly, prediction of the comprising RSE_m's are spatially averaged as global

prediction

$$RMS_{predict} = \sqrt{\frac{\sum_{m=1}^{M} RSE_{m,predict}}{M}} = \sqrt{\mathbb{E}[RMS^2]} \tag{2.2.16}$$

If the generation average $\bar{X}_m(N)$ is selected to be the estimator $\hat{X}_m = \bar{X}_m(N)$, the variance of the estimator can be expanded using the definition in Eq 2.2.6. Thus, the predicted RMS can be written as:

$$\begin{aligned} RMS_{predict} &= \sqrt{\frac{1}{M}\sum_{m=1}^{M}\frac{\mathrm{Var}[\bar{X}_m(N)]}{\langle X_m\rangle^2}} \\ &= \sqrt{\frac{1}{MN}\sum_{m=1}^{M}\frac{\sigma_m^2}{\langle X_m\rangle^2}\left(1+2\sum_{k=1}^{N-1}\left(1-\frac{k}{N}\right)\rho_m(k)\right)} \end{aligned} \tag{2.2.17}$$

Switching the order of summation over tally regions and summation over active generations, Eq 2.2.17 can be expressed in a neat way hiding tally region dependence, which eases later analysis. By defining an average variance for all tally regions as

$$\bar{\sigma^2} = \frac{1}{M}\sum_{m=1}^{M}\frac{\sigma_m^2}{\langle X_m\rangle^2} \tag{2.2.18}$$

and an average variance-weighted ACC,

$$\bar{\rho}(k) = \frac{\sum_{m=1}^{M}\frac{\sigma_m^2}{\langle X_m\rangle^2}\rho_m(k)}{\sum_{m=1}^{M}\frac{\sigma_m^2}{\langle X_m\rangle^2}} \tag{2.2.19}$$

the predicted $RMS_{predict}$ can be cast in the same form as Eq 2.2.6

$$RMS_{predict} = \sqrt{\frac{\bar{\sigma^2}}{N}\left(1+2\sum_{k=1}^{N-1}(1-\frac{k}{N})\bar{\rho}(k)\right)} \tag{2.2.20}$$

With knowledge of the ACC's or a suitable approximation, this formula permits pre-

dicting the convergence rate of the tallies as a function of the number of generations, N, and the number of neutrons per generation, via the $\bar{\sigma}^2$ term.

2.2.2 Bias of Sample Variance Estimated from Generation Tallies

Eq 2.2.20 shows that the variance, $\bar{\sigma}^2$, divided by the sample size, N, underestimates the variance of the mean. On the other hand, if the variance is unknown, it must be estimated from samples that are correlated. This section provides the correction to the variance estimator needed to take into account sample correlation.

For a sample of observables of random variable X, $\{X(1), \cdots, X(N)\}$, sample variance $\widehat{s^2}$ is a common estimate of variance of X [34].

$$\begin{aligned}
\widehat{s^2}(N) &= \frac{1}{N}\sum_{n=1}^{N}(X(n) - \bar{X}(N))^2 \\
&= \frac{\sum_{n=1}^{N} X(n)^2}{N} - \left(\frac{\sum_{n=1}^{N} X(n)}{N}\right)^2 \\
&= \frac{1}{N}\sum_{n=1}^{N} X(n)^2 - \frac{1}{N^2}\sum_{n,n'=1}^{N} X(n)X(n') \\
&= \frac{1}{N}\sum_{n=1}^{N} X(n)^2 - \frac{1}{N^2}\left(\sum_{n=1}^{N} X(n)^2 + \sum_{n\neq n'} X(n)X(n')\right)
\end{aligned} \tag{2.2.21}$$

The derivation below finds the expectation of the sample variance under different assumptions and adjusts it to be an unbiased estimator.

Expectation of $\widehat{s^2}$ in Eq 2.2.21 can be first expanded to

$$\mathbb{E}\widehat{s^2}(N) = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[X(n)^2\right] - \frac{1}{N^2}\left(\sum_{n=1}^{N}\mathbb{E}\left[X(n)^2\right] + \sum_{n\neq n'}\mathbb{E}\left[X(n)X(n')\right]\right) \tag{2.2.22}$$

The factor after $\frac{1}{N}$ in Eq 2.2.22 contains N terms. The factor after $\frac{1}{N^2}$ in Eq 2.2.22 contains N^2 terms, where the first summation ranges from 1 to N and the second

summation contains $N(N-1)$ pairs.

Assume $\mathbb{E}[X(n)]$ does not change over generation n and denote it with $\mathbb{E}[X]$. Subtracting $\mathbb{E}[X]^2$ from both terms on RHS of Eq 2.2.22 leads to Eq 2.2.23.

$$\begin{aligned}\mathbb{E}\widehat{s^2}(N) = &\frac{1}{N}\left\{\sum_{n=1}^{N}\mathbb{E}\left[X(n)^2\right] - N\mathbb{E}[X]^2\right\} \\ &- \frac{1}{N^2}\left(\sum_{n=1}^{N}\mathbb{E}\left[X(n)^2\right] + \sum_{n\neq n'}\mathbb{E}\left[X(n)X(n')\right] - N^2\mathbb{E}[X]^2\right)\end{aligned} \tag{2.2.23}$$

Then merge the N number of $\mathbb{E}[X]^2$ terms back to the summation indexed from $n=1$ to N. Among the N^2 number of $\mathbb{E}[X]^2$ terms, merge N of them are to the summation indexed by n, merge the remaining $N(N-1)$ to the summation indexed by n, n'. Expectation of $\widehat{s^2}(N)$ is now written as

$$\begin{aligned}\mathbb{E}\widehat{s^2}(N) = &\frac{1}{N}\sum_{n=1}^{N}\left\{\mathbb{E}\left[X(n)^2\right] - \mathbb{E}[X]^2\right\} \\ &- \frac{1}{N^2}\left\{\sum_{n=1}^{N}\left(\mathbb{E}\left[X(n)^2\right] - \mathbb{E}[X]^2\right) + \sum_{n\neq n'}\left(\mathbb{E}\left[X(n)X(n')\right] - \mathbb{E}[X]^2\right)\right\}\end{aligned} \tag{2.2.24}$$

Since $\mathbb{E}X = \mathbb{E}X(n)$ for any n, recover $\mathbb{E}X$ as $\mathbb{E}X(n)$ or $\mathbb{E}X(n')$ in Eq 2.2.24 whenever needed to match the summation index. The expectations with matched generation indices n and n' in Eq 2.2.25 are simplified to covariances.

$$\begin{aligned}\mathbb{E}\widehat{s^2}(N) = &\frac{1}{N}\sum_{n=1}^{N}\left\{\mathbb{E}\left[X(n)^2\right] - \mathbb{E}[X(n)]^2\right\} \\ &- \frac{1}{N^2}\left\{\sum_{n=1}^{N}\left(\mathbb{E}\left[X(n)^2\right] - \mathbb{E}[X(n)]^2\right) + \sum_{n\neq n'}\left(\mathbb{E}\left[X(n)X(n')\right] - \mathbb{E}[X(n)]\mathbb{E}[X(n')]\right)\right\}\end{aligned} \tag{2.2.25}$$

$$= \frac{1}{N}\sum_{n=1}^{N}\mathrm{Var}[X(n)] - \frac{1}{N^2}\left(\sum_{n=1}^{N}\mathrm{Var}[X(n)] + \sum_{n\neq n'}\mathrm{Cov}[X(n), X(n')]\right) \tag{2.2.26}$$

When simplifying Eq 2.2.25 to Eq 2.2.26, the following equations are used.

$$\begin{aligned}\mathrm{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ \mathrm{Cov}[X, X'] &= \mathbb{E}[XX'] - \mathbb{E}X\mathbb{E}X'\end{aligned} \tag{2.2.27}$$

The above derivation holds for any processes $\{X(n)\}_{n=1}^N$ with constant $\mathbb{E}X(n)$. In the three sections below, Eq 2.2.22 is applied to relate the $\widehat{s^2}$ to useful variance estimators under different assumptions: independent identical $X(n)$, correlated but identical $X(n)$ and correlated $X(n)$ with only identical $\mathbb{E}X(n)$.

2.2.2.1 Independent and identical observables

If we assume that $\{X(n)\}_{n=1}^N$ belong to an identical distribution and are uncorrelated, all the $\mathrm{Var}[X(n)]$ terms are identical and equal to $\mathrm{Var}[X]$. This leads to the covariance term vanishing which simplifies Eq 2.2.26 to

$$\mathbb{E}\widehat{s^2}(N) = \frac{N-1}{N}\mathrm{Var}[X] \tag{2.2.28}$$

This reveals the classic unbiased variance estimator for uncorrelated samples [9]:

$$\widehat{\sigma_0^2}(N) = \widehat{s^2}(N)\frac{N}{N-1} \tag{2.2.29}$$

2.2.2.2 Correlated but identical observables

If the observables are correlated but still with identical $\mathrm{Var}[X(n)]$ terms, Eq 2.2.26 is simplified to

$$\begin{aligned}\mathbb{E}\widehat{s^2}(N) &= \mathrm{Var}[X]\left(\frac{1}{N}N - \frac{1}{N^2}\left(N + \sum_{n\neq n'}\frac{\mathrm{Cov}[X(n), X(n')]}{\mathrm{Var}[X]}\right)\right) \\ &= \frac{\mathrm{Var}[X]}{N}\left(N - \left(1 + \frac{1}{N}\sum_{n\neq n'}\frac{\mathrm{Cov}[X(n), X(n')]}{\sqrt{\mathrm{Var}[X(n)]\,\mathrm{Var}[X(n')]}}\right)\right) \\ &= \frac{\mathrm{Var}[X]}{N}\left(N - \left(1 + \sum_{k=1}^{N-1}\left(1 - \frac{k}{N}\right)\rho(k)\right)\right)\end{aligned} \tag{2.2.30}$$

where the last equality follows the same assumption that generation-to-generation correlation coefficient depends only on generation lag as in Eq 2.2.5. The definition of variance underestimation ratio $r(N)$ (Eq 2.2.8) is recognized in Eq 2.2.30. After re-organization, this leads to the unbiased variance estimator for correlated sample with identical distribution

$$\sigma_1^2(N) = \widehat{s^2}(N)\frac{N}{N - r(N)} \tag{2.2.31}$$

where $\rho(k)$ (and therefore $r(N)$) are assumed to be known parameters rather a statistic calculated from the sample $\{X(n)\}_{n=1}^{N}$.

Comparing $\widehat{\sigma_1^2}$ (Eq 2.2.31) and $\widehat{\sigma_0^2}$(Eq 2.2.29) shows that the bias of the estimator $\widehat{\sigma_0^2}$ is related with the variance underestimation ratio r through

$$b(N) \equiv \frac{\widehat{\sigma_0^2}(N)}{\widehat{\sigma_1^2}(N)} = \frac{N - r(N)}{N - 1}. \tag{2.2.32}$$

Note that in $\widehat{s^2}$ (Eq 2.2.21), the comprising N terms in the sum cannot vary independently due to the constraint that $\sum_{n=1}^{N}\left(X(n) - \bar{X}(N)\right) = 0$, which leads to the reduced degree of freedom $N-1$ in the estimator $\widehat{\sigma_0^2}$ (denominator in Eq 2.2.29). For the correlated sample, the degree of freedom in $\widehat{\sigma_1^2}$ (Eq 2.2.31) is further reduced to $N - r(N)$.

Therefore, finding $r(N)$ is only part of the problem in estimating $\mathrm{Var}[\bar{X}]$. The ratio $r(N)$ corrects the convergence rate from $1/N$ to $r(N)/N$, but the estimator of the leading term $\mathrm{Var}[X]$ is biased due to the correlation. The factor $b(N)$ is referred to as the bias of the estimator of $\mathrm{Var}[X]$.

2.2.2.3 Correlated observables with identical expectation

In this case, since $\mathrm{Var}[X(n)]$ is no longer constant with n, it is more convenient to estimate its average of N generations. The expectation of the sample variance $\widehat{s^2}$ is thus a useful quantity to investigate. Define the average variance up to generation N

(of sample size N) as

$$\bar{\sigma}^2(N) = \frac{1}{N}\sum_{n=1}^{N} \mathrm{Var}[X(n)] \tag{2.2.33}$$

And define the normalized covariance between generation n and n' as

$$\bar{\rho}(n, n', N) = \frac{\mathrm{Cov}[X(n), X(n')]}{\bar{\sigma}^2(N)} \tag{2.2.34}$$

The above definitions simplify Eq 2.2.26 to

$$\mathbb{E}\widehat{s^2}(N) = \bar{\sigma}^2(N)\left(1 - \frac{1}{N}\left(1 + \frac{1}{N}\sum_{n\neq n'} \bar{\rho}(n, n', N)\right)\right) \tag{2.2.35}$$

Similar to $r(N)$, define the excess degree of freedom $R(N)$ as

$$R(N) \equiv 1 + \frac{1}{N}\sum_{n\neq n'} \bar{\rho}(n, n', N) \tag{2.2.36}$$

Then Eq 2.2.35 reveals an unbiased estimator of the average sample variance up to generation N.

$$\widehat{\bar{\sigma}^2} = \widehat{s^2}\frac{N}{N - R(N)} \tag{2.2.37}$$

Note that since $\mathrm{Var}[X(n)]$ is not stationary, Eq 2.2.5 cannot be used to evaluate $\mathrm{Var}[\bar{X}(N)]$. Eq 2.2.38 should be used instead.

$$\mathrm{Var}\left[\bar{X}(N)\right] = \frac{1}{N^2}\left(\sum_{n=1}^{N} \mathrm{Var}[X(n)] + 2\sum_{k=1}^{N-1}\sum_{n=1}^{N-k} \mathrm{Cov}[X(n), X(n+k)]\right) \tag{2.2.38}$$

$$= \widehat{\bar{\sigma}^2}(N)\frac{R(N)}{N} \tag{2.2.39}$$

2.2.3 Reference Calculations

Before entering the details of the correlation analysis and prediction methods, this section discusses how reference solutions (and correlation coefficients) can be obtained from independent simulations. These reference values will then be used to assess the accuracy of the predictive methods.

Accuracy of the predictive models will be assessed using the three following quantities:

- Correlation coefficients of tally in region I as function of generation lag k: $\rho_I(k)$
- Bias of the estimator of sample variance as function of active generations: $b_I(N)$
- Underestimation ratio of the variance of the estimator (tally averaged over active generations) as function of active generations: $r_I(N)$ and $R_I(N)$.

Suppose D independent simulations are performed and each simulation contains N' active generations.

2.2.3.1 Correlation coefficients

The ACC, $\rho(k)$, can be estimated by first calculating $\rho(n,k)$ (Eq 2.2.7) as the Pearson correlation coefficient of the two samples, $\{X_l^{(1)}(n),\cdots,X_l^{(D)}(n)\}$ and $\{X_l^{(1)}(n+k),\cdots,X_l^{(D)}(n+k)\}$ and then taking the average of all n's using the fact that $\rho_{n,k}$ is stationary. $X_l^{(a)}$ uses the same notation as Section 3.3.1 with a superscript (a) to identify the independent simulations from $a=1$ to $a=D$.

Eq 2.2.40 is equivalent to the definition in Eq 2.2.7 by simply replacing the expectation operator in Eq 2.2.7 with average over the independent simulations:

$$\hat{\rho}_I(n;k) = \frac{s\sum_{a=1}^{D} X_I^{(a)}(n)X_I^{(a)}(n+k) - \sum_{a=1}^{D} X_I^{(a)}(n)\sum_{a=1}^{D} X_I^{(a)}(n+k)}{\sqrt{s\sum_{a=1}^{D} X_I^{(a)}(n)^2 - \left(\sum_{a=1}^{D} X_I^{(a)}(n)\right)^2}\sqrt{s\sum_{a=1}^{D} X_I^{(a)}(n+k)^2 - \left(\sum_{a=1}^{D} X_I^{(a)}(n+k)\right)^2}} \tag{2.2.40}$$

where I denotes the location, D denotes the number of simulations, n denotes the index of the N' active generations in each of the D simulations.

When the correlation coefficients are stationary, $\rho(n;k) = \rho(n';k)\ ,\forall n,n',\ \hat{\rho}(k)$

can be calculated as average over the $\hat{\rho}(n;k)$'s for $n = 1, \cdots, N'$.

$$\hat{\rho}_I(k) = \frac{1}{N'-k} \sum_{n=1}^{N'-k} \hat{\rho}_I(n;k) \tag{2.2.41}$$

2.2.3.2 Bias of variance estimator

The reference of $b_I(N)$ can be calculated by first having a good estimate of the variance of X_I, then dividing it by the biased estimator using tally $X_I(n)$ with n ranging from 1 to N. The good estimate of the variance of X_I is obtained by first estimating $\mathrm{Var}[X_I(n)]$ over the sample $\{X_I^{(a)}(n)\}_{a=1}^{D}$ from D independent simulations using the classic unbiased estimator $\widehat{\sigma^2_{0,X_I(n)}}(D)$ (Eq 2.2.29) then taking the average over all active generations to reduce statistic noise assuming stationarity.

$$\begin{aligned}
\widehat{\mathrm{Var}[X_I]} &= \frac{1}{N'} \sum_{n=1}^{N'} \widehat{\sigma^2_{0,X_I(n)}}(D) \\
&= \frac{1}{N'} \sum_{n=1}^{N'} \frac{\sum_{a=1}^{D} \left(X_I^{(a)}(n)\right)^2 - \left(\sum_{a=1}^{D} X_I^{(a)}(n)\right)^2 / D}{D-1}
\end{aligned} \tag{2.2.42}$$

Then the traditional "unbiased" estimator which turns out biased in the situation of correlation is also calculated with the classic unbiased estimator $\widehat{\sigma^2_{0,X_I}}(N)$ (Eq 2.2.29) but over the sample $\{X_I^{(a)}(n)\}_{n=1}^{N'}$ for each simulation (a) then averaged over the D simulations to reduce statistic noise.

$$\widehat{\sigma^2_{0,X_I}}(N) = \frac{1}{D} \sum_{a=1}^{D} \frac{1}{N-1} \sum_{n=1}^{N} \left(X_I^{(a)}(n) - \bar{X}_I^{(a)}(n)\right)^2 \tag{2.2.43}$$

The ratio between the above two variance estimators is used as the numerical reference for $b_I(N)$

$$\hat{b}_I(N) = \frac{\widehat{Var[X_I]}}{\widehat{\sigma^2_{0,X_I}}(N)} \tag{2.2.44}$$

2.2.3.3 Variance underestimation ratio

The reference of $r_I(N)$ can be obtained by first calculating the variance of $\overline{X}_I(N)$ and then dividing it by a good reference of $\mathrm{Var}[X_I]$.

$\mathrm{Var}[\overline{X}_I(N)]$ can be estimated directly from the sample formed by tallies at any N over the D independent simulations $\{\overline{X}_I(N)^{(a)}\}_{a=1}^{D}$ using the classic unbiased estimator $\widehat{\sigma^2_{0,\overline{X}_I(N)}}(D)$.

$$\widehat{\mathrm{Var}[\overline{X_I}(N)]} = \widehat{\sigma^2_{0,\overline{X}_I(N)}}(D) = \frac{\sum_{a=1}^{D}\left(\overline{X_I^{(a)}}(N)\right)^2 - \left(\sum_{a=1}^{D}\overline{X_I^{(a)}}(N)\right)^2/D}{D-1} \tag{2.2.45}$$

Since the D simulations are independent, $\mathrm{Var}\left[\overline{X_I}(N)\right]$ estimated by Eq 2.2.45 is unbiased.

$\widehat{Var[\overline{X_I}(N)]}$ should then be divided by an estimate of the leading term $\mathrm{Var}[X_I]$ to give a reference of $\frac{r(N)}{N}$. The denominator of Eq 2.2.46 could have naturally been $\widehat{\mathrm{Var}[\overline{X_I}(1)]}$ using the same formula in Eq 2.2.45 by setting $N=1$. Assuming stationarity and to reduce statistic noise, the average over all $\widehat{\mathrm{Var}[X_I(n)]}$ $(n=1,\cdots,N')$ is used, which is equivalent to $\widehat{\mathrm{Var}[X_I]}$ estimated in Eq 2.2.42. That is, $\widehat{\frac{r_I(N)}{N}}$ is calculated as

$$\widehat{\frac{r_I(N)}{N}} = \frac{\widehat{Var\left[\overline{X_I}(N)\right]}}{\widehat{\mathrm{Var}[X_I]}} \tag{2.2.46}$$

It is worthwhile to note that Eq 2.2.46 can be modified slightly to be the reference of $\frac{R_I(N)}{N}$. Recall from section 2.2.2.3 that when the variance of $X(n)$ is not stationary, $R(N)$ is defined as $\mathrm{Var}[\bar{X}(N)]/\bar{\sigma}(N)$, where $\bar{\sigma}(N)$ is the average over the first N number of $\mathrm{Var}[X(n)]$'s. Also, the $\widehat{\mathrm{Var}[X_I]}$ defined in Eq 2.2.42 is the average over all the N' estimators (references) of $\mathrm{Var}[X(n)]$'s. Therefore, replacing $\widehat{\mathrm{Var}[X_I]}$ (or more accurately $\widehat{\mathrm{Var}[X_I]}(N')$) with $\widehat{\mathrm{Var}[X_I]}(N)$ changes the reference of $r(N)/N$ to reference of $R(N)/N$.

$$\widehat{\frac{R_I(N)}{N}} = \frac{\widehat{Var\left[\overline{X_I}(N)\right]}}{\widehat{\mathrm{Var}[X_I]}(N)} \tag{2.2.47}$$

2.3 Correlation analysis

2.3.1 Estimating autocorrelation coefficients

From the analysis in section 2.2, as long as autocorrelation coefficients $\rho(k)$ are known, the bias in tally variance at each generation and averaged over multiple generations can be corrected using Eq 2.2.6 and Eq 2.2.31 respectively.

Though the covariance and variance terms in the calculation of $\rho(k)$ (Eq 2.2.7) are not known, they can be estimated under the assumption that the correlation between generations only depends on generation lag. In order to calculate $\hat{\rho}(k)$, $\{X(i)\}_{i=1}^{N-k}$ and $\{X(i+k)\}_{i=1}^{N-k}$ are treated as two separate data sets and $\hat{\rho}(k)$ is calculated as the normalized covariance between the two sets. The estimate is given in Eq 2.3.1 [22].

$$\widehat{\rho}(k) = \frac{(N-k)\sum_{i=1}^{N-k} X(i)X(i+k) - \sum_{i=1}^{N-k} X(i)\sum_{i=1}^{N-k} X(i+k)}{\sqrt{(N-k)\sum_{i=1}^{N-k} X(i)^2 - \left(\sum_{i=1}^{N-k} X(i)\right)^2}\sqrt{(N-k)\sum_{i=1}^{N-k} X(i+k)^2 - \left(\sum_{i=1}^{N-k} X(i+k)\right)^2}} \tag{2.3.1}$$

Eq 2.3.1 only gives reasonable estimate with large enough sample size $N-k$. Since Eq 2.3.1 is used to evaluate $\hat{\rho}(k)$ for $k \ll N$, it is safe to discard the correction to variance estimators due to correlation between the $X(i)'s$.

To numerically verify the assumption that the autocorrelation coefficients are function of generation lag only, one instance of test problem $TD1$ was constructed to transport neutrons in discrete phase space regions as described in section 2.1.1. In $TD1$, neutrons are transferred among $M = 17$ discrete states according to Eq 2.1.5. And the probability mass function of new neutrons out of each absorption is set to be homogeneous with values $\mathbb{P}(\xi = 0) = p_0 = 0.6$, $\mathbb{P}(\xi = 2) = p_2 = 0.2$ and $\mathbb{P}(\xi = 3) = p_3 = 0.2$. The probabilities are specified according to Eq 2.1.4 such that $\nu = 2.5$. Since $\mathbb{E}\xi = 1$, $TD1$ is a critical system.

An initial source of 32000 neutrons were started in the system and transferred among the $M = 17$ states independently following Eq 2.1.5 (Transport step). After the destination state is determined, the number of new neutrons is sampled according

to the probabilities p_0,p_2 and p_3 (Branching step). The total number of new neutrons does not necessarily add up to 32000 due to statistical fluctuations. Denote the total number of new neutrons as n. In order to make the simulation comparable to typical neutron transport codes where normalization is performed between each generation, 32000 neutrons are sampled from the new neutrons at the destination states by repeating or discarding the last $|n - 32000|$ neutrons (Normalization step) [39]. These 32000 neutrons start a new generation and the process is repeated for 1050 generations. The neutron source distribution from the last 1000 generations are stored for analysis. The above simulation was repeated 2000 times using a different initial random seed.

The reference autocorrelation coefficients for the tally at each state can be calculated according to Eq 2.2.7 by using the independent simulations to compute the covariance terms.

$\rho(n, n + k)$ for the central region calculated for different starting generation n are plotted in Figure 2-1, which clearly indicates that the ACC's only depend on generation lag. It is worthwhile to note that the stationarity of ACC's is valid for the simulation process described above due to the Markov property of the simulation. If more complicated generation-to-generation dependence were introduced into the system through source acceleration such as CMFD [31] [30] [27], generation dependent ACC's may appear [32].

The ACC estimator formula is first applied to the homogeneous cube. A tally mesh of $16 \times 16 \times 16$ was selected since it is representative of an assembly size tally in a PWR. Eq 2.3.1 estimates $\hat{\rho}(k)$ for each region m, and Eq 2.2.19 incorporates ACC for all tally regions into one global $\bar{\rho}(k)$. Homogeneity of the cube simplifies Eq 2.2.19 into a simple average of ACC over all tally regions. Fig 1-2(a) and Fig 2-2(a) illustrate the similarities in autocorrelation coefficients between the 2D BEAVRS with assembly size tallies and the simple cube with $25cm(= 400cm/16)$ size tallies. The simple benchmark also illustrates an important feature of the problem, that of lower autocorrelation coefficients with smaller tally regions, as seen in the BEAVRS benchmark from figures 1-2(a) and 1-2(b) and illustrated in the benchmark by figures 2-2(a)

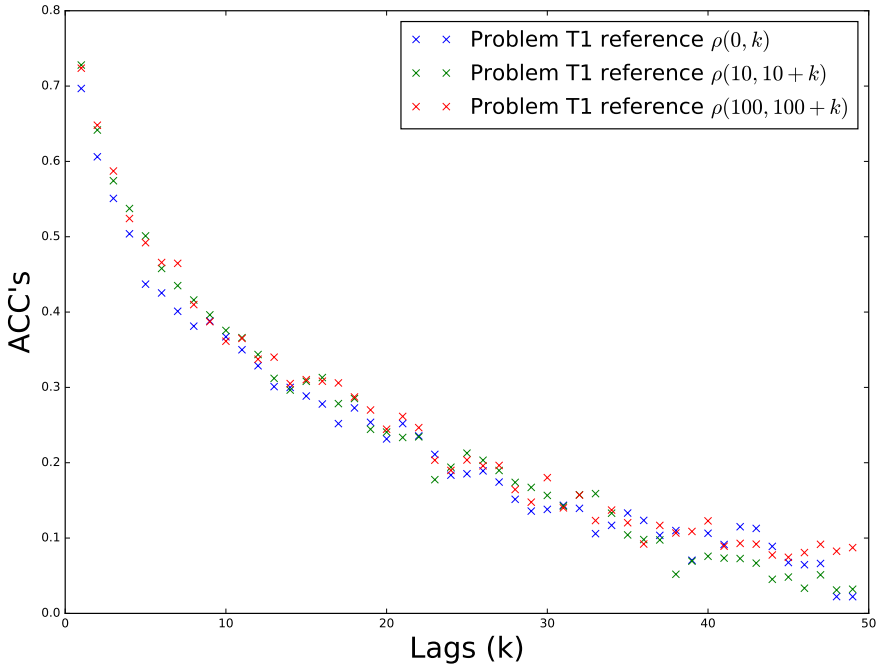


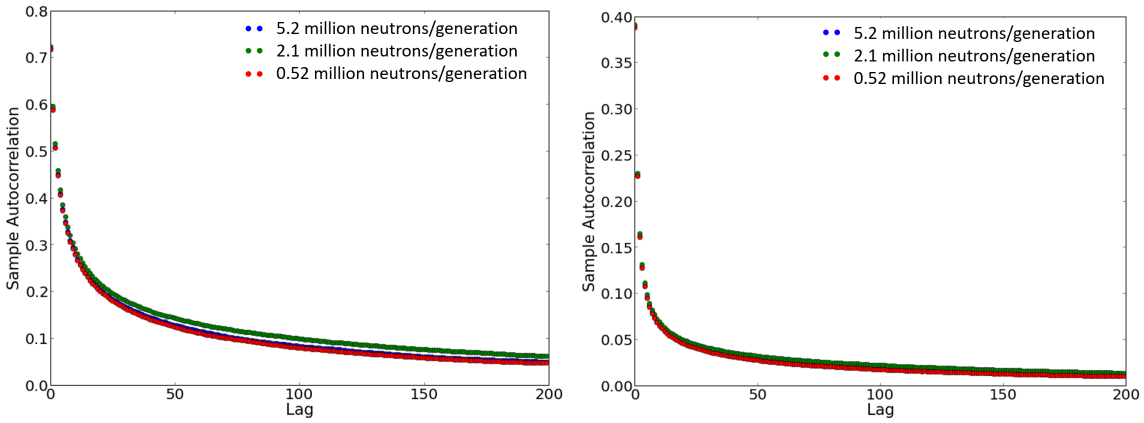
Figure 2-1: Autocorrelation coefficients of test problem TD1

and 2-2(b). Additionally, Figures 1-2(c) and 2-2(c) present similar deviations from the ideal convergence rate which is directly caused by the presence of correlations.

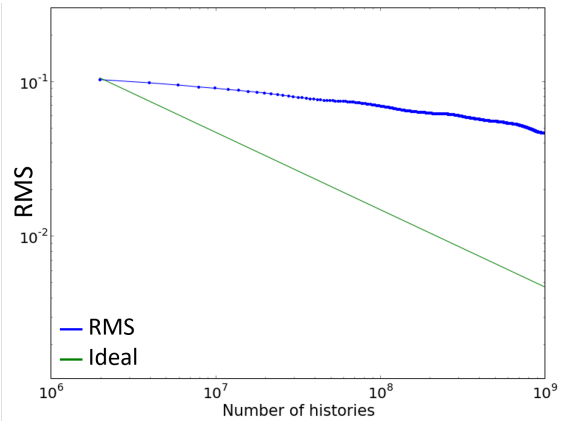
2.3.2 Fitting autocorrelation coefficients

The ACC estimator only gives $\rho(k)$ for $k \ll N$ for a given tally composed of N generations. However, the variance underestimation ratio $r(n)$ up to $n = N$ is required to properly correct the variance, thus needing knowledge of $\rho(n)$ up to $n = N$. One heuristic method is to fit the estimated ACC to some simple model and then extrapolate the model up to N.

From the autocorrelation coefficients observed in section 2.3.1, it is reasonable to assume that the autocorrelation coefficients decay exponentially as a function of the generation lag [49]. This hypothesis can be further validated by observing that the tallies in a region are both Gaussian and Markovian (depends only on previous generation) and that such Gauss-Markov processes present exponential autocorrelation



(a) Autocorrelation coefficients $16 \times 16 \times 16$ tally (b) Autocorrelation coefficients $32 \times 32 \times 32$ tally



(c) RMS Convergence $16 \times 16 \times 16$ tally

Figure 2-2: Simple benchmark autocorrelation coefficients and assembly-size tally convergence rate

coefficients [19].

The exponential decay assumption of ACC's is expressed as

$$\bar{\rho}(k) = \sum_{j=1}^{J} c_j q_j^k \tag{2.3.2}$$

where J is the number of decay modes used to fit the ACC and $0 < q_j < 1$ since ACC decays as lag increases.

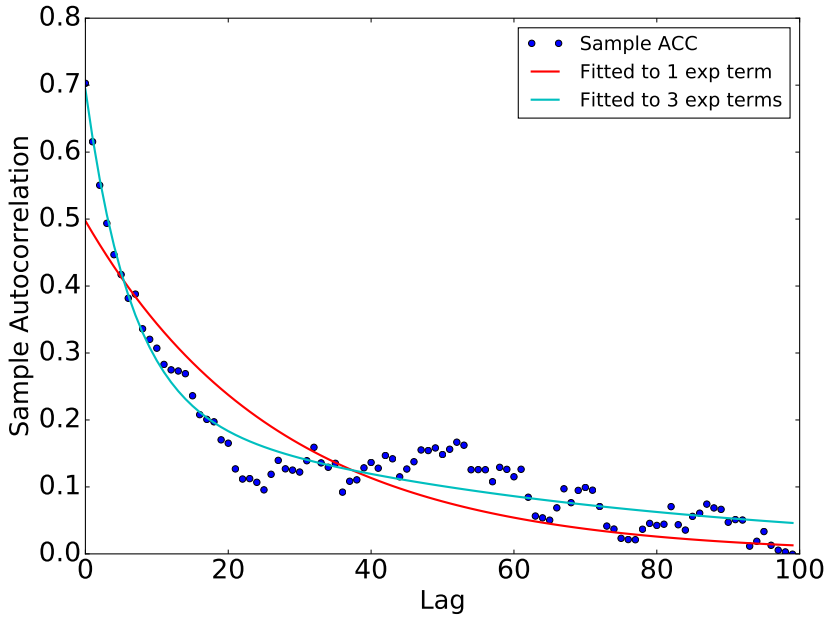
The sampled autocorrelation coefficients for the central region in test problem $TD1$ are fitted to Eq 2.3.2 with $J = 1$ and $J = 3$. The fitted autocorrelation coefficients are plotted along with the sampled autocorrelation coefficient in Fig 2-3(a). The agreement between the fitted curves and original data points is clearly observed. More indirect verifications of the exponential fit via $r(N)$ calculated from ACC are discussed in next section.

It is worthwhile to note that the sample autocorrelation in Fig 2-3(a) is different from the reference ACC shown in Fig 2-1. The reference ACC's in Figure 2-1 were calculated from the multiple independent simulations, while the sample ACC's in Figure 2-3(a) were calculated using the estimator in Eq 2.3.1. Fitting to the reference for demonstration purposes can help support the approximation, but the real test comes when using a single simulation to predict the convergence profile.

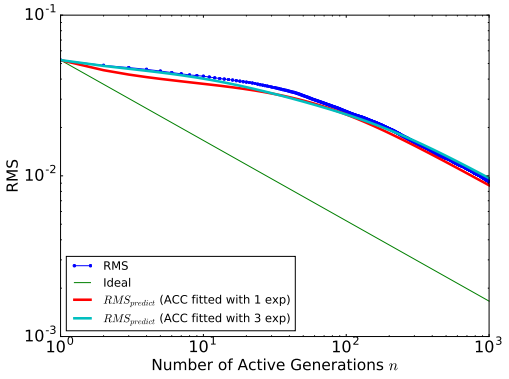
As for the homogeneous cube problem, $TC1$, it is very costly to generate and store all tally regions per generations for many independent simulations. Instead, the (spatially averaged) sampled autocorrelation coefficients of the homogeneous cube (as shown in Fig 2-2(a)) are plotted along with the (spatially averaged) exponential fit of autocorrelation coefficients in Fig 2-4(a). Two fitting schemes $J = 1$ and $J = 3$ are plotted. With statistical noise significantly reduced from spatial averaging, the $J = 3$ exponential fitting matches the original data nearly perfectly.

2.3.3 Convergence rate in the presence of correlation

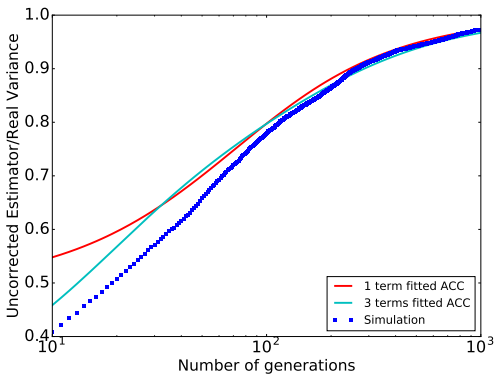
This section analyzes the convergence rate of the variance $\mathrm{Var}[\bar{X}]$ (thus the expectation of the squared error ESE for an unbiased $\bar{X}$) with generation-to-generation



(a) Sample ACC and fitted ACC in the cube

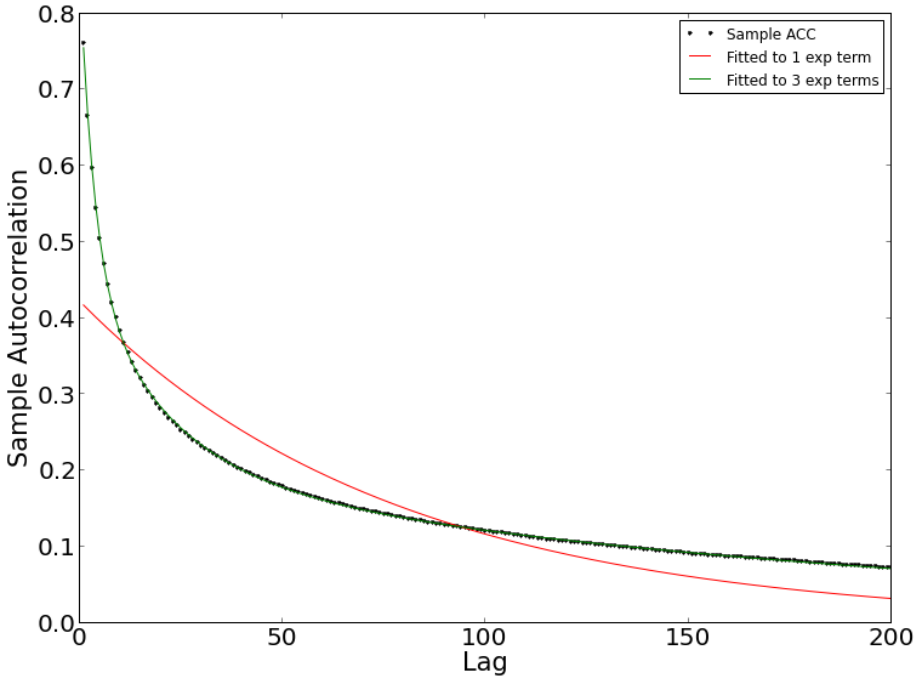


(b) RMS and $RMS_{predict}$ from fitted ACC

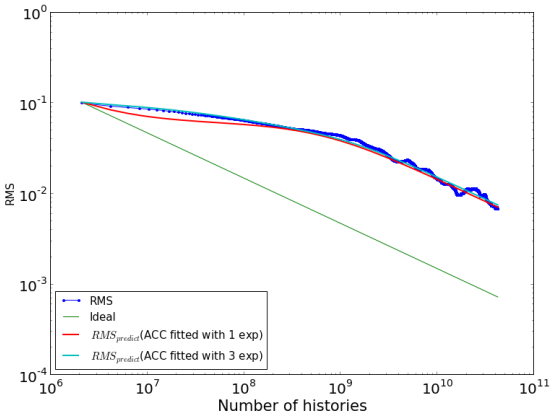


(c) Sample variance underestimate ratio and its prediction from fitted ACC

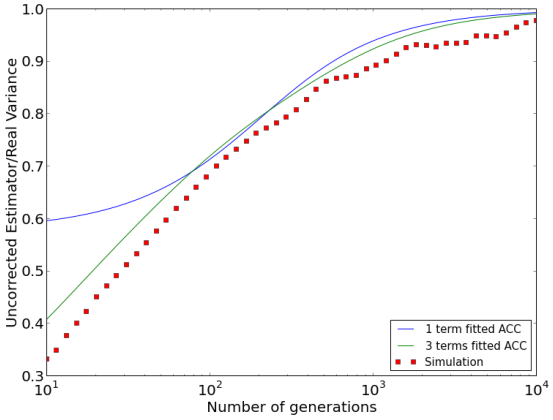
Figure 2-3: Verification of exponential ACC fitting in central region of test problem $TD1$



(a) Sample ACC and fitted ACC in the cube



(b) RMS and $RMS_{predict}$ from fitted ACC



(c) Sample variance underestimate ratio and its prediction from fitted ACC

Figure 2-4: Verification of exponential ACC fitting of test problem TC1

correlation coefficients modelled with exponential fits as shown in Eq 2.3.2. Without loss of generality, the following analysis is based on the simple fit with $J = 1$, but all conclusions still hold for any $J > 1$. Using the simple exponential fit,

$$\bar{\rho}(k) = \rho_0 q^k \tag{2.3.3}$$

the sum Eq 2.2.8 can be evaluated as

$$r(N) = 1 + 2\rho_0 \left(\frac{q}{1-q} - \frac{q^2 - q^N}{(1-q)^2 N} - \frac{q + q^N}{(1-q)N} \right) \tag{2.3.4}$$

Knowing $r(N)$, $RMS^2_{predict}$(Eq 2.2.20) can be evaluated as a function of the number of generations N, spatially averaged variance σ^2 (Eq 2.2.18), and the two ACC fitting parameters ρ_0, q,

$$RMS^2_{predict}(N, \sigma^2, \rho_0, q) = \frac{\sigma^2}{N} \left[1 + 2\rho_0 \left(\frac{q}{1-q} - \frac{q^2 - q^N}{(1-q)^2 N} - \frac{q + q^N}{(1-q)N} \right) \right] \tag{2.3.5}$$

As N approaches infinity, the first two terms in the bracket of $RMS^2_{predict}$ dominate and the overall asymptotic behavior follows a $1/N$ trend. This limit also indicates that the asymptotic ratio of variance underestimation with uncorrelated variance (Eq 2.2.4) is equal to $1 + 2\rho_0 \frac{q}{1-q}$. The convergence rate of RMS can also be obtained from the asymptotic slope of $RMS^2_{predict}$ versus N in the log-log scale.

$$\begin{aligned} &\lim_{N \to \infty} \frac{\partial log\left(RMS^2_{predict}\right)}{\partial log(N)} \\ &= \frac{N\left((q-1)^2 - 2q\rho_0(q-1)\right) + O(q^N) + O(Nq^N) + O(1)}{N\left(2q\rho_0(q-1) - (q-1)^2\right) + O(q^N) + O(1)} \\ &\to -1 \end{aligned} \tag{2.3.6}$$

This indicates that the asymptotic convergence rate of $RMS^2_{predict}$ follows $1/N$, thus RMS follows $1/\sqrt{N}$ even in the presence of generation-to-generation correlation.

In addition to the comparison between sampled ACC and fitted ACC, the exponential decay assumption can be verified with the factor $r(N)$ (Eq 2.2.8). First,

$r(N)$ from fitted autocorrelation coefficient is substituted into $RMS_{predict}$ (Eq 2.3.5), illustrated in Fig 2-4(b). Second, $r(N)$ from fitted ACC predicts the sample variance bias factor $\frac{N-r}{N-1}$ (Eq 2.2.32), given in Fig 2-4(c). It can be seen from Fig 2-4 that although 1 term exponential fitting does not predict the ACC as accurately as a 3 term fitting, it is sufficiently well to predict the asymptotic behavior of RMS and sample variance bias. This can be attributed to the $1 - \frac{k}{n}$ factor beside $\rho(k)$ in the definition of $r(N)$ (Eq 2.2.8). The autocorrelation coefficients of larger generation lag contribute less to the variance underestimation ratio. Fig 2-4(b) also numerically verifies the unbiasedness of the sample mean since $RMS_{predict}$ stems from the variance of the estimator while RMS stems from the square error of the estimator. Their agreement demonstrates unbiasedness as given in Eq 2.2.31.

2.3.4 Independence of autocorrelation coefficients on generation size

If q and ρ_0 in Eq 2.3.5 are function of the generation size s, it would be possible to find an optimal value of work $N \cdot s$ (assuming a known stationary fission source distribution and a minimal value of s to avoid undersampling issues) to reach a target accuracy.

This optimization problem simplifies immensely if we assume that $q(s)$ and $\rho_0(s)$ are constant as a function of generation size, as evidenced by Fig 2-2(a). The ACC's estimated with Eq 2.3.1 for 3 different generation sizes almost overlap with one another. Two additional observations can be made supporting the assumption of constant $q(s)$ and $\rho_0(s)$ as a function of generation size. The observations will be presented in terms of $RMS_{predict}$ calculated from fitted ACC's and RMS from the simulation. Define the total amount of work $N_t = N \cdot s$.

The first observation is that the RMS vs N_t curves of different generation sizes merge when N_t approaches infinity (Fig 2-5). Without loss of generality, we investigate $RMS^2_{predict}$ of two simulations with generations size s_1 and s_2 respectively. The

dependence of $RMS^2_{predict}$ on s can be written explicitly

$$\begin{aligned} RMS^2_{predict}\left(N,\sigma^2,\rho_0,q\right) &= RMS^2_{predict}\left(N_t/s,\sigma(s)^2,\rho_0(s),q(s)\right) \\ &= \frac{\sigma^2(s)}{N_t/s}\left[1+2\rho_0(s)\left(\frac{q(s)}{1-q(s)}-\frac{q(s)^2-q(s)^{N_t/s}}{(1-q(s))^2N_t/s}-\frac{q(s)+q(s)^{N_t/s}}{(1-q(s))N_t/s}\right)\right] \end{aligned} \tag{2.3.7}$$

Regardless of the generation size, s, the $RMS^2_{predict}$ is unchanged as N_t approaches infinity, as illustrated in Fig 2-5 and shown mathematically below

$$\lim_{N_t\to\infty}\frac{RMS^2_{predict}(N_t/s_1,\sigma^2(s_1),\rho_0(s_1),q(s_1))}{RMS^2_{predict}(N_t/s_2,\sigma^2(s_2),\rho_0(s_2),q(s_2))}=1,\ \forall s_1,s_2 \tag{2.3.8}$$

Eq 2.3.8 simplifies to Eq 2.3.9 after substituting Eq 2.3.7 and taking the limit

$$\frac{s_1(1-q(s_2)\sigma^2(s_1)(1-q(s_1)+2q(s_1)\rho_0(s_1)}{s_2(1-q(s_1)\sigma^2(s_2)(1-q(s_2)+2q(s_2)\rho_0(s_2)}=1,\ \forall s_1,s_2 \tag{2.3.9}$$

where $\sigma^2(s_1)$ is the variance of the sample obtained from N_t/s_1 generations. At each generation, X_i is calculated from the tally average over the s_1 neutrons. Thus, if the variance of the tally for one neutron is c and the neutrons are independent, $\sigma^2(s_1)$ becomes c/s_1. In reality, the s_1 neutrons are not necessarily independent because neutrons generated from the same fission site are correlated. However, it is still reasonable to assume $\sigma^2(s_1)\propto\frac{1}{s_1}$

$$\sigma(s_1)^2=\frac{c}{s_1},\sigma(s_1)^2=\frac{c}{s_2},\ \forall s_1,s_2\Rightarrow\sigma^2(s_1)s_1=\sigma^2(s_2)s_2=c \tag{2.3.10}$$

From this, Eq 2.3.9 simplifies to equation Eq 2.3.11

$$\frac{\rho_0(s_1)q(s_1)}{1-q(s_1)}=\frac{\rho_0(s_2)q(s_2)}{1-q(s_2)},\ \forall s_1,s_2 \tag{2.3.11}$$

If $\rho_0(s)$ and $q(s)$ are constant as a a function of s, Eq 2.3.8 and Eq 2.3.11 hold, and the $RMS^2_{predict}$ curves will merge at large N_t as observed in Fig 2-5.

The second observation is that $RMS\times\sqrt{s}$ as a function of the number of gener-

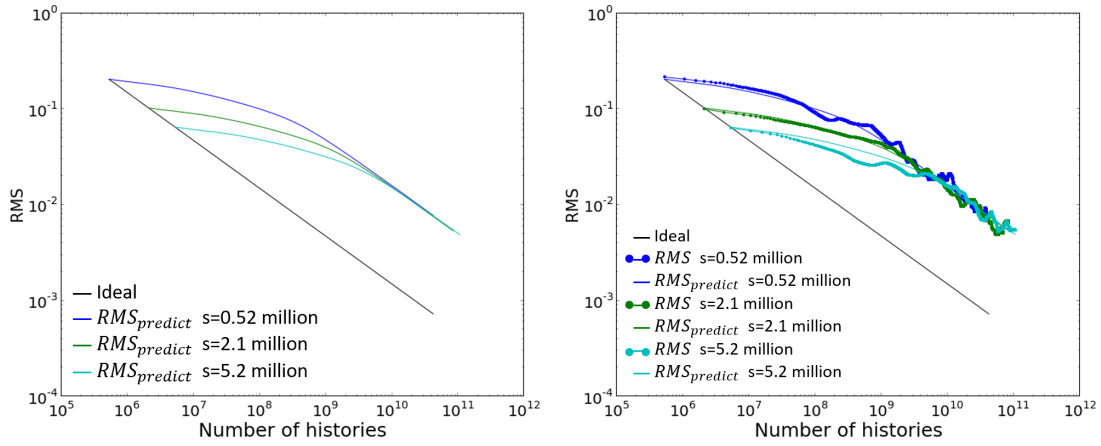


Figure 2-5: $RMS - N_t$ curves merge; left: prediction with fitted ACC; right: with RMS data from MC simulation

ations, N, for different generation size overlap with each other (Fig 2-6). Similar to the above analysis, this dependence can be written explicitly

$$\begin{aligned} RMS^2_{predict}(N,\sigma^2,\rho_0,q) &= RMS^2_{predict}(N,\sigma(s)^2,\rho_0(s),q(s)) \\ &= \frac{\sigma^2(s)}{N}\left[1+2\rho_0(s)\left(\frac{q(s)}{1-q(s)}-\frac{q(s)^2-q(s)^N}{(1-q(s))^2N}-\frac{q(s)+q(s)^N}{(1-q(s))N}\right)\right] \end{aligned} \tag{2.3.12}$$

If $\rho_0(s)$ and $q(s)$ are constant as a function of generation size, $RMS^2_{predict}$ depends on s only through the σ^2 term. Substituting Eq 2.3.10 into Eq 5.1.1,

$$\frac{RMS^2_{predict}(N,\sigma^s(s_1),\rho_0,q)s_1}{RMS^2_{predict}(N,\sigma^s(s_2),\rho_0,q)s_2}=1 \tag{2.3.13}$$

Therefore, the $RMS\times\sqrt{s}$ vs N curves will overlap regardless of generation size as observed in Fig 2-6.

In this section, it was numerically verified that ρ_0 and q are constant as a function of generation size by directly investigating ACC's and observing the convergence behavior of RMS. This can also be supported by Yamamoto, et al. [50] and Sutton [42]'s work on predicting underestimation of variance, where the underestimate ratio is insensitive to generation size. The underestimation ratio in this paper is represented as $1+2\rho_0\frac{q}{1-q}$ according to Eq 2.3.5 and its insensitivity to generation size is

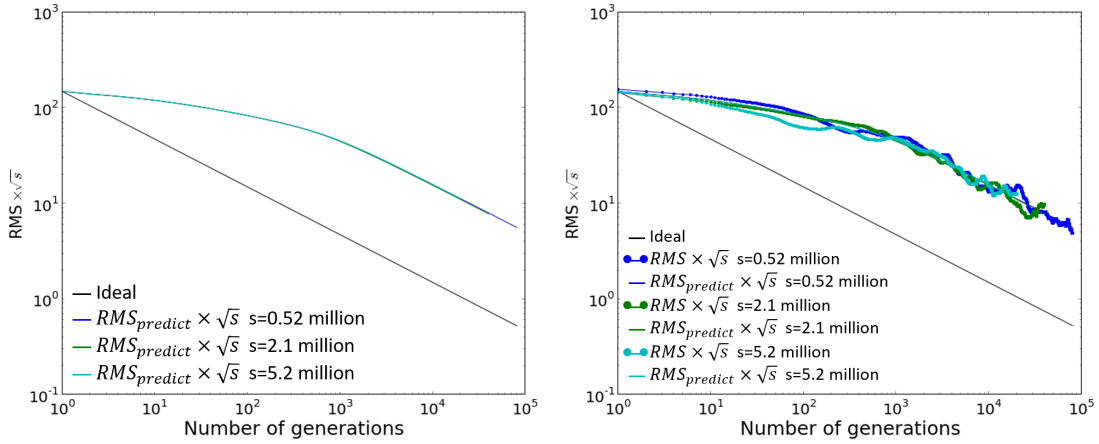


Figure 2-6: RMS×$\sqrt{s}-N$ curves overlap; left: prediction with fitted ACC; right: with RMS data from MC simulation

consistent with the insensitivity of ACC's.

2.4 Interval Estimation

This section gives a brief review of interval estimate of uncorrelated samples and then discusses the extension to correlated samples. The basic idea of interval estimate is to construct a pivot quantity Q, which depends on the statistic T calculated from a sample and the parameter θ to be estimated but whose distribution does not depend on θ as expressed in the following equation

$$\mathbb{P}(t_1 < Q(T,\theta) < t_2) = P \tag{2.4.1}$$

which can be converted to

$$\begin{aligned} \mathbb{P}(f_1(T,t_1) <\theta < f_2(T,t_2) = P \\ or\ \mathbb{P}(f_1(T,t_2) <\theta < f_2(T,t_1) = P \end{aligned} \tag{2.4.2}$$

since the distribution of Q does not depend on θ, neither do t_1 and t_2. Therefore $(f_1(T,t_1), f_2(T,t_2))$ or $(f_1(T,t_2), f_2(T,t_1))$ is the confidence interval of θ with probability P.

Suppose in generation i, the s histories provide tallies of quantity x, $\{x_{i,1}, \cdots, x_{i,s}\}$

from which a generation mean is defined

$$X_i = \frac{1}{s}\sum_{j=1}^{s} x_{i,j} \tag{2.4.3}$$

Additionally, for N generations with sample $\{X_1, X_2, \cdots, X_N\}$, the overall simulation mean is given by

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{2.4.4}$$

which is an estimate of the expectation of the true mean μ. If the neutrons within a generation are assumed independent of each other, thus tallied quantities $\{x_1, \cdots, x_s\}$ are independent. According to the central limit theorem, the average over these neutrons (X_i) converges to a normal distribution,

$$X_i \overset{d}{\Rightarrow} N(\mu, \sigma^2) = N(\mu, \frac{c}{s}) \tag{2.4.5}$$

where $\overset{d}{\Rightarrow}$ denotes convergence in distribution, σ^2 is the variance of X, and c is the variance of x.

2.4.1 Uncorrelated Generations

This subsection reviews the confidence intervals that would be obtained if generation correlation is neglected. If the sample $\{X_1, \cdots, X_N\}$ is uncorrelated, an interval estimate can be obtained by constructing a Student's t distribution pivot. For $i.i.d$ random variables $X_i \sim N(\mu, \sigma^2)$, if μ is estimated with Eq 2.4.4 and σ^2 with $\widehat{\sigma^2}$ (Eq 2.2.29), the pivot can be written as

$$Q_{UG1} = \frac{\bar{X} - \mu}{\sqrt{\widehat{\sigma^2}/N}} \sim t_{N-1} \tag{2.4.6}$$

where t_{N-1} does not dependent on μ and σ^2 [28].

Substituting Q_{UG1}, $\theta = \mu$, $t_1 = -t_{N-1,\alpha/2}$, $t_2 = t_{N-1,\alpha/2}$ and $P = 1-\alpha$ in Eq 2.4.1

provides the uncorrelated generation statistics confidence interval ($UGCI$) of μ as

$$UGCI_g = \bar{X} \pm t_{N-1,\alpha/2} \frac{\sqrt{\widehat{\sigma^2}}}{\sqrt{N}} \tag{2.4.7}$$

where $t_{N-1,\alpha/2}$ is the $\alpha/2$ quantile of student's t distribution of degrees of freedom $N-1$ and $\widehat{\sigma^2}$ is the estimator of σ^2 (Eq 2.2.29). $UGCI_g$ only requires a large generation size to approximate X_i by a normal distribution but applies to any generation number $N > 2$. When the generation number, N, approaches infinity, the limit of $UGCI$ is

$$UGCI_g = \bar{X} \pm Z_{\alpha/2} \frac{\sqrt{\widehat{\sigma^2}}}{\sqrt{N}} \tag{2.4.8}$$

This can be obtained by taking the limit of $t_{N-1,\alpha/2}$ to $Z_{\alpha/2}$.

Alternatively, σ^2 can be calculated from $\hat{c}$ (history based variance estimate of c), which does not require generation statistics over the sample $\{X_1, \cdots, X_N\}$. Therefore σ^2 can be assumed to be known and a normal pivot can be constructed

$$Q_{UG2} = \frac{\bar{X} - \mu}{\sqrt{\sigma^2}/\sqrt{N}} \overset{d}{\Rightarrow} N(0,1) \tag{2.4.9}$$

where

$$\begin{aligned} \sigma^2 &= \frac{\hat{c}}{s} \\ \hat{c} &= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{s-1} \sum_{j=1}^{s} (x_{i,j} - X_i)^2 \equiv \widehat{\sigma^2}_h \end{aligned} \tag{2.4.10}$$

Each generation gives an estimate of c, while $\hat{c}$ is the average.

Substituting Q_{UG2}, $\theta = \mu$, $t_1 = -Z_{\alpha/2}$, $t_2 = Z_{\alpha/2}$ and $P = 1 - \alpha$ in Eq 2.4.1 gives the uncorrelated generation confidence interval of μ as

$$UGCI_h = \bar{X} \pm Z_{\alpha/2} \frac{\sqrt{\hat{c}}}{\sqrt{Ns}} \tag{2.4.11}$$

2.4.2 Correlated Generations

Since the generation results are correlated with each other, it is impossible to construct the traditional student's t-distribution pivot to perform interval estimate. If the sample $\{X_1, \cdots, X_N\}$ is viewed as a Markov Chain, as N approaches infinity, the central limit theorem of Markov Chain holds [26][4] and a normal distribution pivot can be constructed for interval estimates. More generally, Ibragimov [25] proved a variant of the central limit theorem for stationary sequence with weak dependence ($\rho_k \to 0$ as $k \to \infty$):

$$Q_{CG} = \frac{\bar{X} - \mu}{\sigma\left(\sum_{i=1}^{N} X_i\right)/N} \overset{d}{\Rightarrow} N(0,1) \tag{2.4.12}$$

Following the procedure in section 2.2.1, it can be shown

$$\sigma\left(\sum_{i=1}^{N} X_i\right) = \sqrt{\sigma^2 N r} \tag{2.4.13}$$

where r is defined in Eq 2.2.8. Similar to the analysis of uncorrelated generations, there are two methods to obtain confidence interval from Eq 2.4.12: the first estimates $\sigma\left(\sum_{i=1}^{N} X_i\right)$ from generation statistics and the second assumes $\sigma\left(\sum_{i=1}^{N} X_i\right)$ is known from history statistics.

If σ^2 is estimated from generation statistics substituted into the original pivot in Eq 2.4.12. Slutsky's theorem [10] determines the distribution of pivot Q_{CG1}:

$$\sqrt{\widehat{\sigma^2} N r} \overset{P}{\Rightarrow} \sigma\left(\sum_{i=1}^{N} X_i\right) \tag{2.4.14}$$

$$Q_{CG1} = \frac{\bar{X} - \mu}{\sqrt{\widehat{\sigma^2} r}/\sqrt{N}} \overset{d}{\Rightarrow} N(0,1) \tag{2.4.15}$$

Substituting Q_{CG1}, $\theta = \mu$, $t_1 = -Z_{\alpha/2}$, $t_2 = Z_{\alpha/2}$ and $P = 1 - \alpha$ in Eq 2.4.1 provides

the correlated generation confidence interval ($CGCI$)of μ as

$$CGCI_g = \bar{X} \pm Z_{\alpha/2} \frac{\sqrt{\widehat{\sigma^2}}}{\sqrt{N}} \sqrt{r} \tag{2.4.16}$$

where $\widehat{\sigma^2}$ is the estimator of σ^2 from generation-based statistics

$$\widehat{\sigma^2} = \frac{1}{N-r} \sum_{i=1}^{N} (X_i - \bar{X})^2 \equiv \widehat{\sigma^2}_g \tag{2.4.17}$$

Alternatively, the denominator of Q_{CG} in Eq 2.4.12 can be approximated from history statistics. Thus the pivot can be written as

$$Q_{CG2} = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/Nr}} \overset{d}{\Rightarrow} N(0,1) \tag{2.4.18}$$

Substituting Q_{CG2}, $\theta = \mu$, $t_1 = -Z_{\alpha/2}$, $t_2 = Z_{\alpha/2}$ and $P = 1 - \alpha$ in Eq 2.4.1 provides the correlated generation confidence interval of μ as

$$CGCI_h = \bar{X} \pm Z_{\alpha/2} \frac{\sqrt{\sigma^2}}{\sqrt{N}} \sqrt{r} \tag{2.4.19}$$

where σ^2 can be approximated as given in Eq 2.4.10. Since each generation gives an unbiased estimate of c, although generations are correlated, their average is unbiased.

Table 2.2 summarizes confidence interval options from history statistics and generation statistics using the same total number of simulated neutrons. Generation statistics are based on tallies from N generations with size s while history statistics are based on tallies from all generation with $N \cdot s$ neutrons. $\bar{X}$ will be the same no matter whether history statistics or generation statistics are used and the confidence intervals are symmetric about the common center, therefore, only the size of confidence intervals are compared.

Table 2.2 also shows that the confidence intervals are larger by a factor of $\sqrt{r}$ in the presence of generation-to-generation correlation. The bias of the variance estimator as noted in Eq 2.2.32 can be neglected since $r \ll N$. Since the confidence intervals are derived from pivots where $N \to \infty$, the asymptotic value of r can be used to

Table 2.2: Comparison of Confidence Intervals

	Uncorrelated generations	Correlated generations
generation statistics (unknown variance)	$\frac{2t_{N-1,\alpha/2}}{\sqrt{N}}\sqrt{\widehat{\sigma^2}}$ ($UGCI_g$,Eq 2.4.8)	$\frac{2Z_{\alpha/2}}{\sqrt{N}}\sqrt{\widehat{\sigma^2}}\sqrt{r}$ ($CGCI_g$,Eq 2.4.16)
history statistics (known variance)	$\frac{2Z_{\alpha/2}}{\sqrt{N}}\sqrt{\frac{\hat{c}}{s}}$ ($UGCI_h$,Eq 2.4.11)	$\frac{2Z_{\alpha/2}}{\sqrt{N}}\sqrt{\frac{\hat{c}}{s}}\sqrt{r}$($CGCI_h$,Eq 2.4.19)

calculate confidence intervals. Using the simple benchmark with $16 \times 16 \times 16$ mesh, $\sqrt{r}$ approaches 8 asymptotically. The factor r is defined in Eq 2.2.8 and is a function of the number of generations N. Additionally, it can be observed in Table 2.2 that history statistics and generation statistics will give the same confidence interval as long as both $\widehat{\sigma^2}$ and $\hat{c}/s$ estimate the variance of X (σ^2) accurately.

2.4.3 Numerical Example

2.4.3.1 Method 1: Asymptotic Confidence Interval with large N

The simplest way to obtain confidence intervals is to run a large number of generations and assume N to approach infinity such that $CGCI_g$ (Eq 2.4.16) or $CGCI_h$ (Eq 2.4.19) can be used directly. The problem is that there is no criterion to determine whether N is large enough to validate the central limit theorem for correlated generation tallies or the asymptotic normal distribution of pivot Q_{CG} (Eq 2.4.12).

2.4.3.2 Method 2: Asymptotic Confidence Interval without explicit r dependence

Using a batching algorithm enables evaluating the confidence interval without the need for r [11] [20]. This method has been implemented in OpenMC [39], MCNP [43] and MC21[13]. The total number of active generations N is divided into $N_B = N/B$ batches, where a batch consists of B consecutive fission generations. The batch k

tally value is obtained as

$$\mathcal{X}_k = \frac{1}{B} \sum_{i=(k-1)M+1}^{kM} X_i \tag{2.4.20}$$

where X_i is the generation i tally value. Statistics over the sample $\{\mathcal{X}_1, \cdots, \mathcal{X}_{N_B}\}$ are performed to construct a pivot and obtain a confidence interval. According to Kelly et al, by making B sufficiently large, the correlation between batches can be reduced to a negligible level. Therefore with a pivot similar to Q_{UG1} (Eq 2.4.6), confidence interval in the form of $UGCI_g$ (Eq 2.4.8) can be obtained by replacing N with N_B, and $\widehat{\sigma^2}$ with $\widehat{\sigma^2_{\mathcal{X}}}$. This is denoted as $UBCI$ (Uncorrelated Batches Confidence Interval) and is evaluated as

$$UBCI = \bar{\mathcal{X}} \pm \frac{t_{N_B-1,\alpha/2}}{\sqrt{N_B}} \sqrt{\widehat{\sigma^2_{\mathcal{X}}}} \tag{2.4.21}$$

$$\widehat{\sigma^2_{\mathcal{X}}} = \frac{1}{(N_B - 1)} \sum_{k=1}^{N_B} (\mathcal{X}_k - \bar{\mathcal{X}})^2 \tag{2.4.22}$$

$UBCI$ is consistent with $CGCI_g$ or $CGCI_h$. By definition of $\mathcal{X}$ (Eq 2.4.20)

$$\sigma^2_{\mathcal{X}} = \frac{\sigma^2}{B} r(B) \tag{2.4.23}$$

where the dependence of r on the number of generations is recalled. Therefore $\widehat{\sigma^2_{\mathcal{X}}}$ along with $\sqrt{N_B}$ in $UGCI$ is consistent with $\sqrt{\frac{\widehat{\sigma^2} r}{N}}$ in $CGCI_g$ and $\sqrt{\frac{\hat{c}r}{sN}}$ in $CGCI_h$. They are all estimators of the variance of $\bar{X}$.

The main problem of this method is that there is no criterion to determine whether B is sufficiently large to validate the independence between batches. Theory developed so far in this paper can be applied to provide a criterion with respect to the correlation between batches. Similarly to the analysis in Sec 2.2.1, covariance between batch J and K can be evaluated as

$$\mathrm{Cov}[\mathcal{X}_J, \mathcal{X}_K] = \frac{\sigma^2}{B^2} \sum_{l=(\Delta-1)B+1}^{(\Delta+1)B-1} (\Delta B - |\Delta B - l|)\, \rho_l \tag{2.4.24}$$

where $\Delta = |J - K| \geq 1$ is the batch lag, ρ_l is the correlation coefficient between

generations with lag l as defined in Eq 2.2.7. The correlation coefficient between batches with lag Δ can be further obtained as

$$\varrho_\Delta = \frac{\mathrm{Cov}[\mathcal{X}_J, \mathcal{X}_K]}{\sigma^2_{\mathcal{X}}} = \frac{1}{Br(B)} \sum_{l=(\Delta-1)B+1}^{(\Delta+1)B-1} (\Delta B - |\Delta B - l|)\, \rho_l \tag{2.4.25}$$

With the verified assumption of exponential decay of ρ_l, batch correlation coefficient ϱ_Δ can be evaluated with the exponential decay parameters as introduced in Sec 2.3.2. With the one-term exponential form of ρ_k, ϱ_1 can be obtained as

$$\varrho_1 = \frac{q\rho_0 \left(1-q^B\right)^2}{B(1-q)\left(2\rho_0 q - q + 1\right) - 2q\rho_0\left(1-q^B\right)} \tag{2.4.26}$$

Finally, given the exponential decay parameter of generation correlation coefficients ρ_k, the "sufficiently" large B can be found through $\varrho_1 \sim 0$.

2.4.3.3 Method 3: Confidence Interval for finite N

In addition to the central limit theorem, distribution of pivots constructed earlier in this section can also be obtained with more assumptions. If the asymptotic distribution of X(Eq 2.4.5) is substituted into the pivot Q_{CG2}, the distribution of Q_{CG2} can be obtained for any N. It can be shown that for N random variables $X_i \sim N(\mu, \sigma^2)$, their average is still normal.

Denote the N Gaussian random variables as an N dimension random vector $\vec{X} \sim (N(\vec{\mu}, \overline{\overline{\Sigma}})$, where $\overline{\overline{\Sigma}}$ is the covariance matrix with $\Sigma_{i,j} = \mathrm{Cov}[X_i, X_j]$. The characteristic function of the Gaussian random vector is

$$\phi_{\vec{X}}\left(\vec{t}\right) = E\left[e^{-i\vec{t}^T\vec{X}}\right] = e^{-i\vec{t}^T\vec{\mu} + \frac{1}{2}\vec{t}^T\overline{\overline{\Sigma}}\vec{t}} \tag{2.4.27}$$

The average of the N X's can be denoted as

$$\bar{X} = \vec{a}^T\vec{X} \tag{2.4.28}$$

$$\vec{a}^T = \frac{1}{N}(1, \cdots, 1) \tag{2.4.29}$$

Then the characteristic function of $\bar{X}$ can be found as

$$\phi_{\bar{X}}(s) = E\left[e^{-is\vec{a}^T\vec{X}}\right] = \phi_{\vec{X}}(s\vec{a}) = e^{-is\vec{a}^T\vec{\mu}+\frac{1}{2}s^2\vec{a}^T\overline{\overline{\Sigma}}\vec{a}} \tag{2.4.30}$$

Therefore $\bar{X}$ is a Gaussian random variable with mean $\vec{a}^T\vec{\mu} = \mu$ and variance $\sigma^2(\bar{X}) = \vec{a}^T\overline{\overline{\Sigma}}\vec{a}$. This leads to a Gaussian pivot for finite N, Q_N,

$$Q_N = \frac{\bar{X} - \mu}{\sqrt{\sigma^2 r(N)}/\sqrt{N}} \sim N(0,1) \tag{2.4.31}$$

where σ^2 and $r(N)$ are assumed to be known and σ^2 can be evaluated with Eq 2.4.10. In comparison with Q_{CG2}, the distribution of Q_N does not require large N and the variance underestimation ratio need not take the asymptotic value. Substituting Q_N, $\theta = \mu$, $t_1 = -Z_{\alpha/2}$, $t_2 = Z_{\alpha/2}$ and $P = 1 - \alpha$ in Eq 2.4.1 provides the correlated generation confidence interval of μ as

$$CGCI_N = \bar{X} \pm Z_{\alpha/2}\frac{\sqrt{\hat{c}}}{\sqrt{Ns}}\sqrt{r(N)} \tag{2.4.32}$$

2.4.3.4 Results - Homogeneous Cube (TC1)

This section compares variance estimators of $\bar{X}$ (Eq 2.4.4), as a proxy to confidence intervals. The variance estimators calculated from history statistics (Eq 2.4.10), generation statistics (Eq 2.4.17) and batch statistics (Eq 2.4.22) are summarized below.

$$\widehat{\sigma^2_{\bar{X}h}} = \frac{\widehat{\sigma^2}_h}{sN} r(N) \tag{2.4.33}$$

$$\widehat{\sigma^2_{\bar{X}g}} = \frac{\widehat{\sigma^2}_g}{N} r(N) \tag{2.4.34}$$

$$\widehat{\sigma^2_{\bar{X}b}} = \frac{\widehat{\sigma^2}_{\mathcal{X}}}{N_B} \tag{2.4.35}$$

The reference variance value can be calculated as

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{N} = \frac{c}{N\ s} \tag{2.4.36}$$

where σ^2 can be obtained analytically and r can be approximated as described in sec 2.3.2. The assumption that the tallies obtained by neutrons in one generation are independent are not true in traditional source update methods due to the possibility of re-using source sites in a given generation leading to history-to-history correlations. This correlation invalidates the relation between history variance c and generation variance σ^2 described by Eq 2.3.10 but can be avoided by removing this source site multiplicity. In the simple cube problem, sites can be re-sampled from a uniform distribution (exact solution), but more generally sites could be replaced by source sites from a previous generation.

Under the assumption of independent history tallies, the analytical variance of the fission tally is

$$c = \frac{P_f}{M^2(1-P_s)^2}(P_c + (1-P_s)(M-1)) \tag{2.4.37}$$

where M is the number of meshes in the cube, and the fission probability P_f, scattering probability P_s and capture probability P_c can be calculated from the cross section parameters in Table 2.1. Eq 2.4.37 is derived in Appendix A1.1.

Knowing c and approximating $r(N)$, Eq 2.4.36 provides a reference value of the true variance.

Sixty independent simulations were performed using $10,000$ generations of $983,040$ neutrons per generation on problem $TC1$. Analog fission rates were tallied on $4\times4\times4$ mesh and results compared on a single tally region. The factor $r(N)$ is calculated following the procedure described in Sec 2.3.2 and assumed to be known. The variance using the three methods for the 60 simulations are presented in Fig 2-7. The batch statistics results were produced assuming batch groupings of $B = 200$. The first observation in Fig 2-7 is that using uncorrelated statistics leads to underestimating the variance by a factor of 2. This figure also demonstrates consistency between the three proposed methods for correlated statistics. The variance estimates from history and generation statistics are closely distributed around the theoretical value, however, batch statistics presents much greater dispersion. Appendix A1.2 demon-

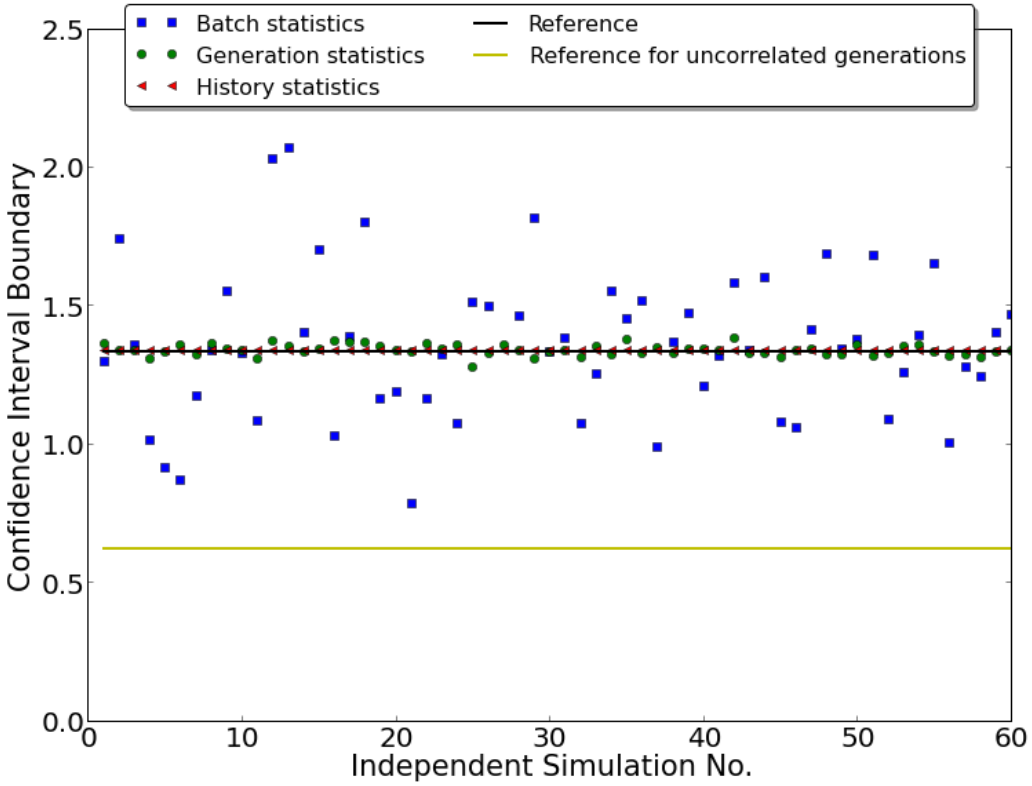


Figure 2-7: Comparison of CI evaluated from different methods
Batch statistics: Estimate $\widehat{\sigma_{\mathcal{X}}^2}$ *in* $UBCI$ *(Eq 2.4.21) with Eq 2.4.22.*
Generation statistics: Estimate $\widehat{\sigma^2}$ *in* $CGCI_h$ *(Eq 2.4.16) with Eq 2.4.17.*
History statistics: Estimate $\widehat{c^2}$ *in* $CGCI_N$ *(Eq 2.4.32) with Eq 2.4.10.*

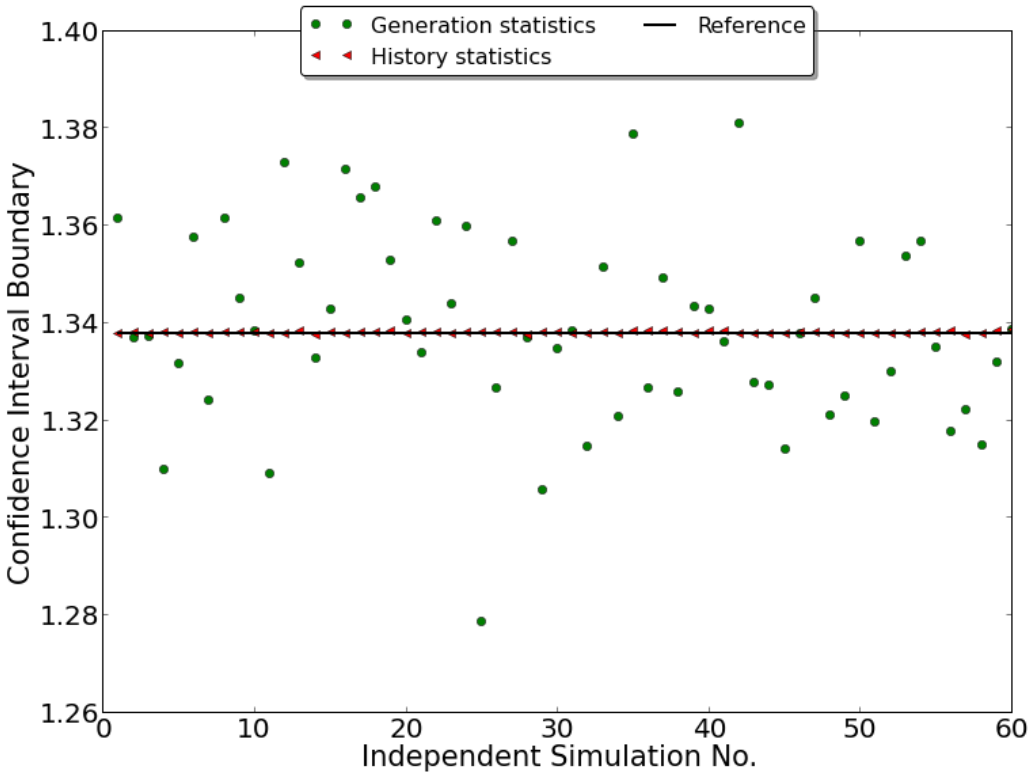
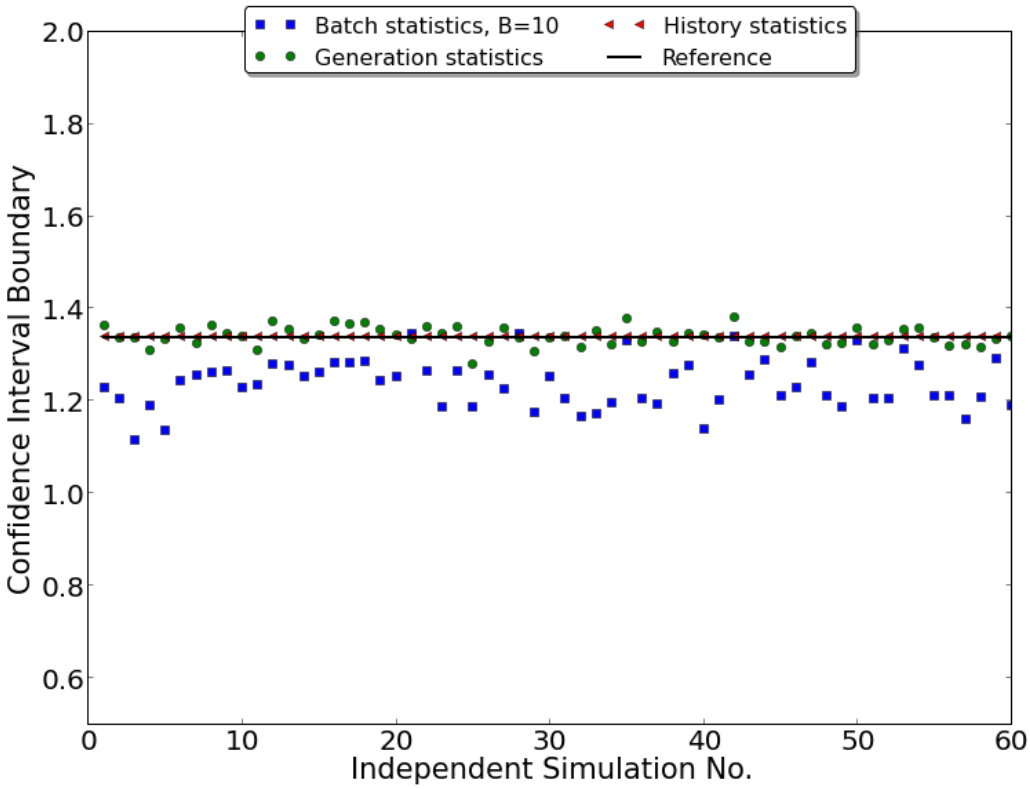


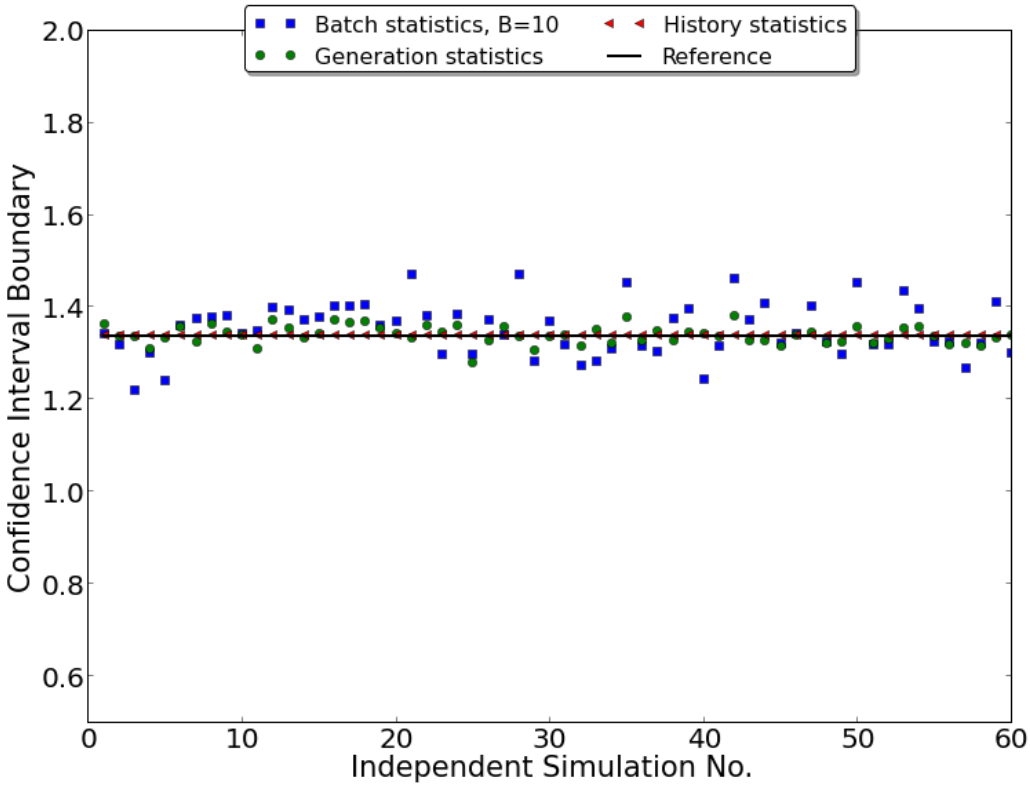
Figure 2-8: Comparison of CI: history and generation statistics

strates that history statistics will provide an estimation with a smaller uncertainty than generation-based estimates and that generation-based estimates provide an estimation with a smaller uncertainty than batch-based variance estimates. This can be observed in Fig 2-7 and Fig 2-8. For this particular problem, $B = 200$ might be an overkill. The large B leads to smaller number of batches, N_B, and further increases the uncertainty of the estimator $\widehat{\sigma^2_{\mathcal{X}}}$. If B is instead set to 10, the dispersion of $\widehat{\sigma^2_{\mathcal{X}}}$ becomes, as expected, much narrower but the average value deviates from the reference value (Fig 2-9(a)). The deviation can be viewed as an underestimation of the batch variance and attributed to correlation between batches. Reducing the batch size B from 200 to 10 increased the batch correlation coefficient of lag 1 (Eq 2.4.26) from 0.0022 to 0.048. Therefore, calculating the batch correlation, the variance can be corrected by a factor $\frac{r(N)}{r(B)}$. This corrected estimator matches the reference value as shown in Fig 2-9(b)).

From these results it can be seen that history statistics (once history-to-history



(a) $B = 10$, $\widehat{\sigma^2_{\mathcal{X}}}$ uncorrected



(b) $B = 10$, $\widehat{\sigma^2_{\mathcal{X}}}$ corrected by $\frac{r(N)}{r(B)}$

Figure 2-9: Comparison of CI: batch statistics of different batch sizes

correlation is eliminated) can provide a suitable and accurate estimate of the variance. Thus, in support of the optimization conclusion to be discussed in section 5.1, few generations with large particle counts is a possibility to provide the lowest possible tally errors while still providing an accurate measure of the confidence intervals. One caveat to this approach is that it requires an approximation of the underestimation factor. The batching scheme avoids this need, but the run strategy favors multiple generations which was shown to be sub-optimal with respect to the convergence rate. Additionally, history statistics require performing accumulation of tallies after each history which can become computationally prohibitive when many tallies are needed.

Chapter 3

Predicting Correlations

Suppose generation n yields tally $X_I(n)$ for tally region I, the simulation typically reports the average $\bar{X}_I(N) = \sum_{n=1}^{N} X_I(N)/N$ and $\sigma^2_{X_I}/N$ as an approximation of $\sigma^2_{\bar{X}_I(N)}$. Due to the correlation between generations in the power iteration process, $\sigma_{X_I}/\sqrt{N}$ underestimates $\sigma_{\bar{X}_I(N)}$. The $\mathrm{Cov}[X_I(i), X_I(j)]$, where i,j are the active generation indexes, is required to correctly evaluate the uncertainty.

These correlations across generations result from the fission site update process where the source of generation $n+1$ come from the fission sites created during generation n. The correlation of any tallied quantity can be calculated from the correlation of the fission source distribution [41].

Recall from chapter 2 that the auto-correlation coefficient of tally in region I between generation n and $n+k$ is defined as

$$\rho_I(n,k) = \frac{\mathrm{Cov}[X_I(n), X_I(n+k)]}{\sqrt{Var[X_I(n)]Var[X_I(n+k)]}} \tag{3.0.1}$$

The correlation estimation method discussed in section 2.3.1 requires tallies from many active generations. This chapter develops methods that extract sufficient information about the system from the first few active generations. This information can be used to build approximate models of the system, where correlation behaviors can be predicted. This method avoids the need of simulating many generations and enables forecasting the correlation behavior and statistical convergence in real-time.

Methods in this chapter are developed to predict the *auto-correlation coefficients* (ACC) of the fission source distribution. All methods to be developed in this section are based on discretizing the neutron phase space. A neutron can be in different phase space regions and during a generation, it can be transferred to another region using a Markov chain transfer matrix. Once transferred to the new region, a neutron can induce a fission event and produce new neutrons. These new neutrons are then transferred through the system independently and can cause new fission events.

Section 3.1 discusses two important differences between predicted correlation results on the discretized phase space regions and results on tally regions of interest. Section 3.1.1 describes how to condense results on finer discrete phase space regions to coarser tally regions. Section 3.1.2 addresses a way to extract accurate correlation results of realistic problems from the analytic model that does not perform neutron number normalization. Section 3.2 applies the Markov Chain Monte Carlo method to predict correlation coefficients by neglecting the fission events. The observed discrepancy between predicted and reference correlation coefficients (and consequentially variance underestimation ratio) implies the generation-to-generation neutron source dependence only captures part of the tally correlation and highlights the impact of source site multiplicity on correlation. Section 3.3 develops the Multitype Branching Processes (MBP) model that can address both generation-to-generation source dependence and fission site multiplicity. The MBP model can predict the correlation behavior of neutron count tallies exactly in the situation where neutron number normalization is not enforced. In order to predict correlation behavior in the scenario with neutron number normalization, the MBP model is developed to predict correlation behavior of neutron count tallies normalized by number of neutrons per generation with covariance expansion. The lowest order expansion gives accurate correlation coefficients of normalized neutron count tallies in the simple test problem $TD1$. Finally, the MBP model is applied to BEAVRS benchmark and shows acceptable accuracy.

3.1 Predictions and references in tally regions

This section discusses the relation between quantities of discretized phase space regions and simulation tally in coarser regions. These relations are the foundation for building a dicretized predictive model.

3.1.1 Tally score condensation

By discretizing the system where neutrons transport and branch, various quantities, especially neutron source count moments can be calculated for the discretized phase space regions. In realistic simulations, auto-correlation coefficients might be needed on discretized regions of different size (e.g. pin vs assembly size tallies in a reaction simulation). To avoid recomputing multiple moments on varying mesh sizes, a condensation process is proposed allowing to evaluate ACCs on coarser meshes from the finely discretized solution.

Suppose tally region I contains phase space cells $I_1, \cdots, I_i, \cdots$, the number of neutrons in region I (denoted by $Z_I(n)$) is the sum of neutrons in the phase space cells. $Z_I(n)$ is defined in Eq 3.1.1.

$$Z_I(n) \equiv \sum_{I_i \in I} Z_{I_i}(n) \tag{3.1.1}$$

From the discussion in chapter 2, the covariance of tallies in the same region I across different generations are needed to predict correlation coefficients and variance underestimation ratio. The covariance terms of tally region I can be directly calculated from finer mesh moments as

$$Var[Z_I(n)] = \sum_{I_i, I_j \in I} \mathrm{Cov}[Z_{I_i}(n), Z_{I_j}(n)] \tag{3.1.2}$$

$$\mathrm{Cov}[Z_I(n), Z_I(n+k)] = \sum_{I_i, I_j \in I} \mathrm{Cov}[Z_{I_i}(n), Z_{I_j}(n+k)] \tag{3.1.3}$$

Explicit forms of all the terms on the right hand side of equations in Eq 3.1.3 will be derived in this chapter.

3.1.2 Source normalization

The other discrepancy between the prediction from discretized phase space region and real simulation is the source particle normalization process at the end of each generation.

For a simulation with s neutrons per generation, s neutron sources are started at the next generation regardless of how many fission sites (denoted by s') were generated. If $s' > s$, $s' - s$ neutrons are discarded randomly. If $s' < s$, $s - s'$ neutrons are sampled randomly from the s' neutrons to fill the neutron source bank of size s. The details on discarded or sampled neutrons differs in different codes. This normalization while essential creates a new type of correlation between each individual particle and all others in a generation. This type of correlation is difficult and will not be treated while building models to predict correlation behaviors for the transport and branching processes of neutrons.

Therefore, the models developed below technically only simulate the neutron transport and branching processes without particle number normalization. However, it will be shown numerically that although the correlation behavior of Z's (raw neutron counts) are different between the predictive models and the real neutron transport processes with particle number normalization, the correlation behavior of X's (neutron counts normalized by neutron particle number) are similar between the predictive models and the real processes.

In previous sections $X_l(n)$ denotes any tally in region l at generation n. From now on, $Z_l(n)$ denotes the number of source neutrons in region l at generation n. And $X_l(n)$ is $Z_l(n)$ normalized by the total number of source neutrons at generation n.

$$X_l(n) = \frac{Z_l(n)}{\sum_{l=1}^{M} Z_l(n)} \tag{3.1.4}$$

To demonstrate that the correlation from normalization is small, test problem TD1 was simulated with $s = 32000$ neutrons per generation. In one case (denoted as $TD1_n$), neutron number normalization is applied at the end of each generation. In

the other case (denoted as $TD1_f$), the neutron population fluctuates freely along with the number of generated fission sites. Both cases are simulated for 300 active generations, and 1000 independent simulations are performed.

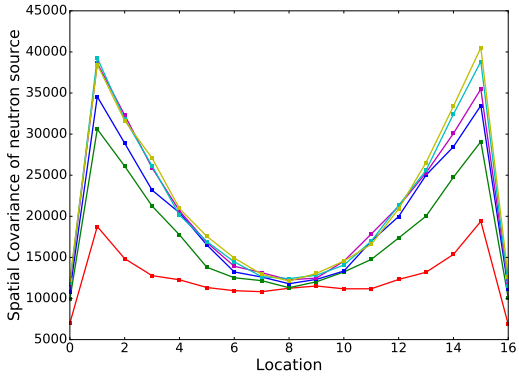
The variances of both raw tally Z and normalized tally X are plotted for both cases $TD1_n$ and $TD1_f$. In Figure 3-1, $\text{Var}[X_I(n)]$ and $\text{Var}[Z_I(n)]$ are plotted for all $M = 17$ regions at generations $n \in \{25, 75, 125, 175, 225, 275\}$. Figure 3-1(a) plots $\{\text{Var}[Z_I]\}_{I=1}^{17}$ from $TD1_n$, and Figure 3-1(b) plots $\{\text{Var}[Z_I]\}_{I=1}^{17}$ from $TD1_f$. It can be seen from these two figures that the variance grows unbounded without neutron number normalization ($TD1_f$) and the variance seems to be clamped by the normalization step ($TD1_n$). On the contrary, the variances of normalized tallies $\{\text{Var}[X_I]\}_{I=1}^{17}$ have similar behaviors for both $TD1_n$ and $TD1_f$ (Figure 3-1(c),Figure 3-1(d)).

The correlation coefficients of both raw count tally Z and normalized tally X in the center region of $TD1$ are plotted for both cases $TD1_n$ and $TD1_f$ in Figure 3-2. Similarly to the behavior of spatial moments, the correlation coefficients of Z tallies differ considerably when neutron number normalization is applied. However, the correlation coefficients of X tallies are indistinguishable between $TD1_n$ and $TD1_f$. This observation allows us to neglect neutron number normalization if correlation coefficients of X tallies are investigated.

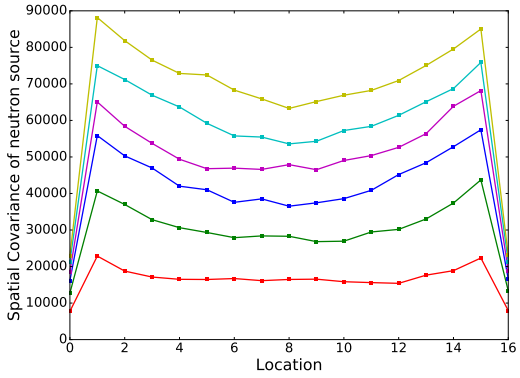
3.2 Markov Chain Monte Carlo

Within a generation, s neutrons from the source bank are simulated and generate s new fission sites. The above process can be approximated by independently simulating s neutrons according to a Markov Chain [4] transfer matrix which can then be used to predict the correlation between the different generations. In this special case, fission is the only reaction and each fission event produces only 1 new neutron. This problem essentially becomes a pure scattering problem and removes within-generation site multiplicity.

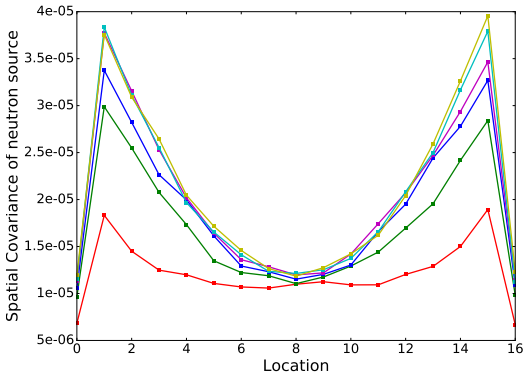
Denote the number of neutrons in region l at generation n as $Z_l(n)$ and the



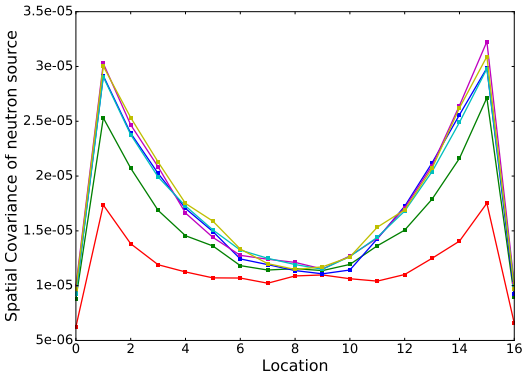
(a) tally Z, with particle number normalization



(b) tally Z, without particle number normalization

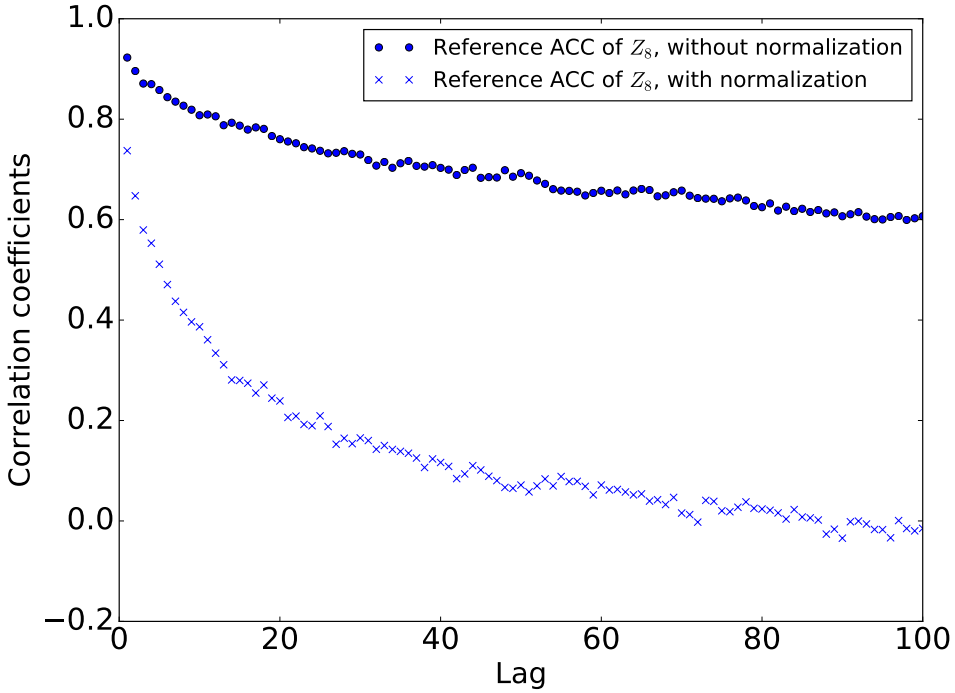


(c) tally X, with particle number normalization

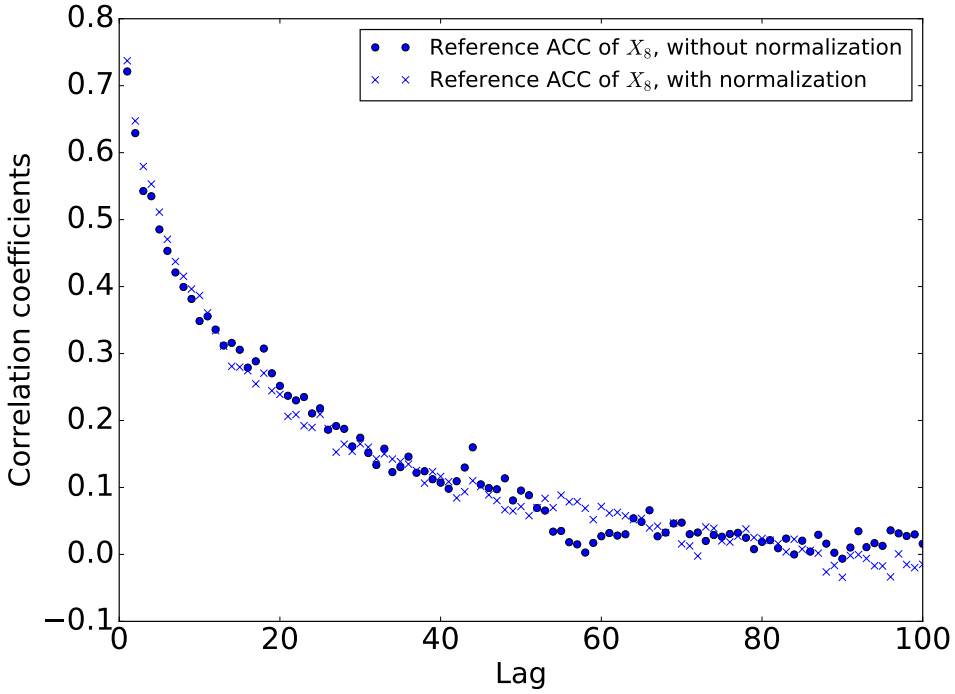


(d) tally X, without particle number normalization

Figure 3-1: Variance of tallies as function of location at different generations.



(a) raw count tally at center region, $\rho_{Z_8}(k)$



(b) normalized tally at center region, $\rho_{X_8}(k)$

Figure 3-2: Correlation coefficients of tallies as function of location at different generations.

corresponding tally $X_l(n) = \frac{Z_l(n)}{\sum_{l'=1}^{M} Z_{l'}}$.

Since the s neutrons contribute to the tallies independently, the correlation behavior contributed by these s neutrons is expected to be a simple transformation of the correlation behavior contributed by one neutron. The tallies induced by an individual neutron are investigated first.

3.2.1 Independence of the neutrons in a generation

This subsection demonstrates that the correlation coefficient between tally $X_I(n)$ and $X_I(n+d)$ (Eq 2.2.7) contributed by the s neutrons in the generation is equivalent to the correlation observed from a generation with a single neutron.

First, if the number of neutrons per generation does not change from generation to generation, it is sufficient to analyze the correlation coefficients of $Z_I(n)$ and $Z_I(n+d)$ (instead of $X_I(n)$ and $X_I(n+d)$).

$$\begin{aligned}\rho_I(d) &= \frac{\mathrm{Cov}[X_I(n), X_I(n+d)]}{\sqrt{\mathrm{Var}[X_I(n)]\,\mathrm{Var}[X_I(n+d)]}} = \frac{\mathrm{Cov}[Z_I(n), Z_I(n+d)]}{\sqrt{\mathrm{Var}[Z_I(n)]\,\mathrm{Var}[Z_I(n+d)]}} \\ &= \frac{\mathbb{E}Z_I(n)Z_I(n+d) - \mathbb{E}Z_I(n)\mathbb{E}Z_I(n+d)}{\sqrt{(\mathbb{E}Z_I(n)^2 - (\mathbb{E}Z_I(n))^2)\,(\mathbb{E}Z_I(n+d)^2 - (\mathbb{E}Z_I(n+d))^2)}}\end{aligned} \tag{3.2.1}$$

Use $z_l(n)$ to denote the number of neutrons in region l at generation n due to one neutron. Index the contribution from s neutrons as $\{z_l^{(1)}, \cdots, z_l^{(\alpha)}, \cdots, z_l^{(s)}\}$. Then

$$Z_l(n) = \sum_{\alpha=1}^{s} z_l^{(\alpha)}(n) \tag{3.2.2}$$

which gives

$$\mathbb{E}Z_l(n) = \mathbb{E}\sum_{\alpha=1}^{s} z_l^{(\alpha)}(n) = \sum_{\alpha=1}^{s} \mathbb{E}z_l^{(\alpha)}(n) = s\mathbb{E}z_l(n) \tag{3.2.3}$$

where the last equality recognizes the s summands are equal.

Then using Eq 3.2.2 to evaluate the expectation of $Z_l(n)Z_{l'}(n+d)$ needed for $Z_I(n)Z_I(n+d)$ in Eq 3.2.1.

$$
\begin{aligned}
\mathbb{E}Z_l(n)Z_{l'}(n+d) &= \mathbb{E}\sum_{\alpha=1}^{s} z_l^{(\alpha)}(n)\sum_{\beta=1}^{s} z_{l'}^{(\beta)}(n+d) \\
&= \sum_{\alpha,\beta=1}^{s} \mathbb{E}z_l^{(\alpha)}(n)z_{l'}^{(\beta)}(n+d) \\
&= \sum_{\alpha}^{s} \mathbb{E}z_l^{(\alpha)}(n)z_{l'}^{(\alpha)}(n+d) + \sum_{\alpha\neq\beta}^{s} \mathbb{E}z_l^{(\alpha)}(n)z_{l'}^{(\beta)}(n+d) \qquad (3.2.4) \\
&= \sum_{\alpha}^{s} \mathbb{E}z_l^{(\alpha)}(n)z_{l'}^{(\alpha)}(n+d) + \sum_{\alpha}^{s} \mathbb{E}z_l^{(\alpha)}(n)\sum_{\beta\neq\alpha}^{s} \mathbb{E}z_{l'}^{(\beta)}(n+d) \qquad (3.2.5) \\
&= s\mathbb{E}z_l(n)z_{l'}(n+d) + s(s-1)\mathbb{E}z_l(n)\mathbb{E}z_{l'}(n+d) \qquad (3.2.6)
\end{aligned}
$$

where Eq 3.2.4 separates the $< \alpha, \beta >$ pairs into two groups, one with $\alpha = \beta$, the other with $\alpha \neq \beta$; the second part of Eq 3.2.5 comes from the fact that neutrons α β are assumed independent; the last equality recognizes the identical summands within each summation.

Similarly, Eq 3.2.2 evaluates the expectation of $Z_l(n)Z_{l'}(n)$ needed for $Z_I(n)^2$ in Eq 3.2.1 as

$$
\begin{aligned}
\mathbb{E}Z_l(n)Z_{l'}(n) &= \mathbb{E}\sum_{\alpha=1}^{s} z_l^{(\alpha)}(n)\sum_{\beta=1}^{s} z_{l'}^{(\beta)}(n) \\
&= \sum_{\alpha=1}^{s}\sum_{\beta=1}^{s} \mathbb{E}z_l^{(\alpha)}(n)z_{l'}^{(\beta)}(n) \\
&= \sum_{\alpha=1}^{s} \mathbb{E}z_l^{(\alpha)}(n)z_{l'}^{(\alpha)}(n) + \sum_{\alpha\neq\beta}^{s} \mathbb{E}z_l^{(\alpha)}(n)z_{l'}^{(\beta)}(n) \qquad (3.2.7) \\
&= \sum_{\alpha=1}^{s} \mathbb{E}z_l^{(\alpha)}(n)z_{l'}^{(\alpha)}(n) + \sum_{\alpha=1}^{s} \mathbb{E}z_l^{(\alpha)}(n)\sum_{\beta\neq\alpha}^{s} \mathbb{E}z_{l'}^{(\beta)}(n) \\
&= s\mathbb{E}z_l(n)z_{l'}(n) + s(s-1)\mathbb{E}z_l(n)\mathbb{E}z_{l'}(n)
\end{aligned}
$$

Eq 3.2.6, Eq 3.2.7,Eq 3.2.3 enable expressing moments of tallies from all s neutrons ($Z_l(n)$'s) in terms of moments of tallies from one single neutron ($z_l(n)$'s), which will

be derived in the next section.

Eq 3.2.6 and Eq 3.2.3 expresses the covariance between neutron count in region I at generations n and $n+d$ into Eq 3.2.8.

$$\begin{aligned}
&\mathbb{E}Z_I(n)Z_I(n+d) - \mathbb{E}Z_I(n)\mathbb{E}Z_I(n+d) \\
&\quad = \sum_{l,l' \in I} \{\mathbb{E}Z_l(n)Z_{l'}(n+d) - \mathbb{E}Z_l(n)\mathbb{E}Z_{l'}(n+d)\} \\
&\quad = \sum_{l,l' \in I} \left\{s\mathbb{E}z_l(n)z_{l'}(n+d) + s(s-1)\mathbb{E}z_l(n)\mathbb{E}z_{l'}(n+d) - s^2\mathbb{E}z_l(n)\mathbb{E}z_{l'}(n+d)\right\} \\
&\quad = \sum_{l,l' \in I} \{s\mathbb{E}z_l(n)z_{l'}(n+d) - s\mathbb{E}z_l(n)\mathbb{E}z_{l'}(n+d)\} \\
&\quad = s\mathbb{E}z_I(n)z_I(n+d) - s\mathbb{E}z_I(n)\mathbb{E}z_I(n+d) \qquad (3.2.8)
\end{aligned}$$

Eq 3.2.7 and Eq 3.2.3 expresses the variance between neutron count in region l at generation n into Eq 3.2.9.

$$\begin{aligned}
&\mathbb{E}Z_I(n)^2 - (\mathbb{E}Z_I(n))^2 \\
&\qquad = \sum_{l,l' \in I} \mathbb{E}Z_l(n)Z_{l'}(n) - (\mathbb{E}Z_I(n))^2 \\
&\qquad = \sum_{l,l' \in I} \{s\mathbb{E}z_l(n)z_{l'}(n) + s(s-1)\mathbb{E}z_l(n)\mathbb{E}z_{l'}(n)\} - s^2(\mathbb{E}z_I(n))^2 \\
&\qquad = s\mathbb{E}z_I(n)^2 - s(\mathbb{E}z_I(n))^2 \qquad (3.2.9)
\end{aligned}$$

Simply replacing n with $n+d$, the variance between neutron count in region l at generation $n+d$ is

$$\mathbb{E}Z_I(n+d)^2 - (\mathbb{E}Z_I(n+d))^2 = s\mathbb{E}z_I(n+d)^2 - s(\mathbb{E}z_I(n+d))^2 \qquad (3.2.10)$$

Eq 3.2.8,Eq 3.2.9,Eq 3.2.10 simplify Eq 3.2.1 to

$$\rho_I(d) = \frac{\mathbb{E}z_I(n)z_I(n+d) - \mathbb{E}z_I(n)\mathbb{E}z_I(n+d)}{\sqrt{\left(\mathbb{E}z_I(n)^2 - (\mathbb{E}z_I(n))^2\right)\left(\mathbb{E}z_I(n+d)^2 - (\mathbb{E}z_I(n+d))^2\right)}} \qquad (3.2.11)$$

3.2.2 Correlation of individual neutrons across generations

The above section shows that the correlation behavior of Z tallies contributed by s tallies is identical to the correlation behavior of z tallies contributed by a single neutron. Now we characterize the transport of a single neutron by a markov transfer matrix P such that $P_{i,j} = \mathbb{P}(\text{neutron moves from region } i \text{ to } j)$ after one generation.

For the first term in the numerator of Eq 3.2.11, the expectation can be expressed from the probability of every configuration of the neutron at generation n and $n+d$ and the corresponding $z_l(n)z_l(n+d)$ under the configurations. If the neutron is in region i at generation n and at region j at generation $n+d$, we will denote the configuration of the neutron at the two generations as $< i, j >$. When the neutron is in region i at generation n, the count of neutron at region l is 1 if and only if $l = i$, that is $z_l(n) = \delta_{l,i}$, $z_l(n+d) = \delta_{l,j}$.

First, write the expectation of $z_l(n)z_l(n+d)$ by summing the corresponding value and probability over all possible $< i, j >$ configurations.

$$\mathbb{E}z_l(n)z_{l'}(n+d) = \sum_{i,j=1}^{M} \mathbb{P}\left(z_l(n) = \delta_{l,i}; z_{l'}(n+d) = \delta_{l',j}\right) \delta_{l,i}\delta_{l',j} \tag{3.2.12}$$

Then decompose the probability of the configuration $< i, j >$ into the product of probability of distribution at generation i and the conditional probability.

$$\mathbb{E}z_l(n)z_{l'}(n+d) = \sum_{i,j=1}^{M} \mathbb{P}\left(z_{l'}(n+d) = \delta_{l',j} \,|z_l(n) = \delta_{l,i}\right) \mathbb{P}\left(z_l(n) = \delta_{l,i}\right) \delta_{l,i}\delta_{l',j} \tag{3.2.13}$$

Then recognize $\mathbb{P}(z_{l'}(n+d) = \delta_{l',j}|z_l(n) = \delta_{l,i})$ as the matrix element $P^d_{i,j}$, where P is the Markov transfer matrix introduced above and P^d denotes the d^{th} power of the matrix. Due to the Markov property, the probability does not depend on n but depends on d only.

$$\mathbb{E}z_l(n)z_{l'}(n+d) = \sum_{i,j=1}^{M} P^d_{i,j}\delta_{l,i}\mathbb{P}\left(z_i(n) = \delta_{l,i}\right) \delta_{l',j} \tag{3.2.14}$$

Then assume that the neutron distribution reaches equilibrium before generation n. Denote the neutron distribution $\pi^{(eq)}$ such that $\pi_i^{(eq)}$ the probability that a neutron is in region i after the distribution reaches equilibrium. Mathematically, $\pi^{(eq)}$ as the equilibrium distribution of the Markov process satisfies

$$\pi^{(eq)} P = \pi^{(eq)} \tag{3.2.15}$$

This assumption leads to $\mathbb{P}(z_l(n) = \delta_{l,i}) = \pi_i^{(eq)}$ ans simplifies Eq 3.2.14 to Eq 3.2.16

$$\mathbb{E} z_l(n) z_{l'}(n+d) = P_{ll'}^d \pi_l^{(eq)} \tag{3.2.16}$$

The assumption that the neutron distribution is in equilibrium also gives

$$\mathbb{E} z_l(n) = \mathbb{E} z_l(n+d) = \pi_l^{(eq)} \tag{3.2.17}$$

Because $z_l(n) \in \{0, 1\}$,

$$\mathbb{E} z_l(n)^2 = \mathbb{E} z_l(n) = \pi_l^{(eq)} \tag{3.2.18}$$

Because a neutron cannot be at l and l' simultaneously,

$$\mathbb{E} z_l(n) z_{l'}(n) = \delta_{l,l'} \pi_l^{(eq)} \tag{3.2.19}$$

The moments of $z_l(n)$ evaluated above (Eq 3.2.16, Eq 3.2.17, Eq 3.2.19) are sufficient to calculate correlation coefficients.

$$\begin{aligned} \rho_l(d) &= \frac{P_{l,l}^d \pi_l^{(eq)} - (\pi_l^{(eq)})^2}{\pi_l^{(eq)} - (\pi_l^{(eq)})^2} \\ &= \frac{P_{l,l}^d - \pi_l^{(eq)}}{1 - \pi_l^{(eq)}} \end{aligned} \tag{3.2.20}$$

If an eigenvalue decomposition is performed on matrix P, $P_{l,l}^d$ must be of the form $\sum_{i=1}^{M} c_i \lambda_i^d$, where λ_i is the i^{th} eigenvalue of P^T, the transpose of P. This observation validates the assumption made in section 2.3.1 that the autocorrelation coefficients

as a function of generation lags can be approximated by a sum of exponentials. The observation also reveals that the asymptotic behavior of the ACC is dominated by λ_2/λ_1, the dominance ratio of the system.

And for a finite region I containing some discrete phase space regions.

$$\begin{aligned}\rho_I(d) &= \frac{\sum_{l,l'\in I}\left\{P^d_{l,l'}\pi^{(eq)}_l - \pi^{(eq)}_l\pi^{(eq)}_{l'}\right\}}{\sum_{l,l'\in I}\left\{\pi^{(eq)}_l\delta_{l,l'} - \pi^{(eq)}_l\pi^{(eq)}_{l'}\right\}} \\ &= \frac{\sum_{l,l'\in I}\left\{P^d_{l,l'}\pi^{(eq)}_l\right\} - \pi^{(eq)}_I\pi^{(eq)}_I}{\pi^{(eq)}_I - \pi^{(eq)}_I\pi^{(eq)}_I} \\ &= \frac{P^d_{I,I} - \pi^{(eq)}_I}{1-\pi^{(eq)}_I}\end{aligned} \tag{3.2.21}$$

where we recognize

$$P^d_{I,I} = \frac{\sum_{l\in I}\pi^{(eq)}_l \sum_{l'\in I} P^d_{l,l'}}{\sum_{l\in I}\pi^{(eq)}_l} \tag{3.2.22}$$

as the probability that a neutron in region I is still found in the same region after d generations.

Under the approximation that the neutrons per generation are independent, according to Eq 3.2.20, the correlation coefficients can be predicted from the Markov Chain transfer matrix describing the independent neutrons. In a real problem with continuous phase space, the transfer matrix P can be approximated by tallying fission-to-absorption matrix as $P_{i,j}$ (probability that a neutron source from tally region i is absorbed in region j at the end of one generation). In next section, the prediction method derived above is tested on problems with a given Markov transfer matrix.

3.2.3 Comparison to references

In this section, the test problem $TD1$ (described in section 2.1.1) and one variation of it, $TD2$ are used to illustrate the autocorrelation coefficients prediction method

derived above.

In both problems, neutrons are transferred among $M = 17$ discrete states with transfer probabilities being specified by Eq 2.1.5. $TD2$ and $TD1$ only differ in the distribution of number of new neutrons per absorption. In $TD2$, exactly one neutron is created per absorption to preserve balance with $k_{eff} = 1$. That is, $\mathbb{P}(\xi = 1) = p_1 = 1$. In comparison, the distribution of ξ in $TD1$ has possible values 2 and 3 in Eq 2.1.8. $TD2$ is an exact representation of the Markov chain process described previously, while $TD1$ captures the more realistic branching of the fission process.

The difference between $TD2$ and $TD1$ can be first viewed from a basic quantity, the spatial variance $\mathrm{Var}[Z_l(n)]$. In $TD2$, the neutrons are independently distributed in the system according to the distribution π in Eq 2.1.6. Therefore,

$$z_l(n) \sim Bernoulli(\pi_l) \tag{3.2.23}$$

which gives

$$\mathrm{Var}[Z_l(n)] = s\,\mathrm{Var}[z_l(n)] = s\pi_l(1-\pi_l) \tag{3.2.24}$$

Since $X_l(n) = Z_l(n)/s$,

$$\mathrm{Var}[X_l(n)] = \pi_l(1-\pi_l)/s \tag{3.2.25}$$

The reference of $\mathrm{Var}[X_l(120)]$ of $TD2$ and $TD1$ are plotted in Figure 3-3 along with the prediction from Eq 3.2.25. Eq 3.2.25 and reference $\mathrm{Var}[X_l(120)]$ of $TD2$ match perfectly. However, the reference $\mathrm{Var}[X_l(120)]$ of $TD1$ are roughly one magnitude higher than that of $TD2$ and are plotted along a different y-axis. The significantly different variance behavior in Figure 3-3 implies that the assumption of independent neutrons within one generation oversimplifies real problems.

Given the transfer matrix P in Eq 2.1.5, we can apply the prediction method expressed by Eq 3.2.20. In Figure 3-4(a), the predicted correlation coefficients of neutron source tally in the central region $\rho_{Z_8}(d)$ from Eq 3.2.20 are plotted along with the numerical references from independent simulations of $TD1$ and $TD2$. The prediction of autocorrelation coefficients using Eq 3.2.20 match the reference of $TD1$

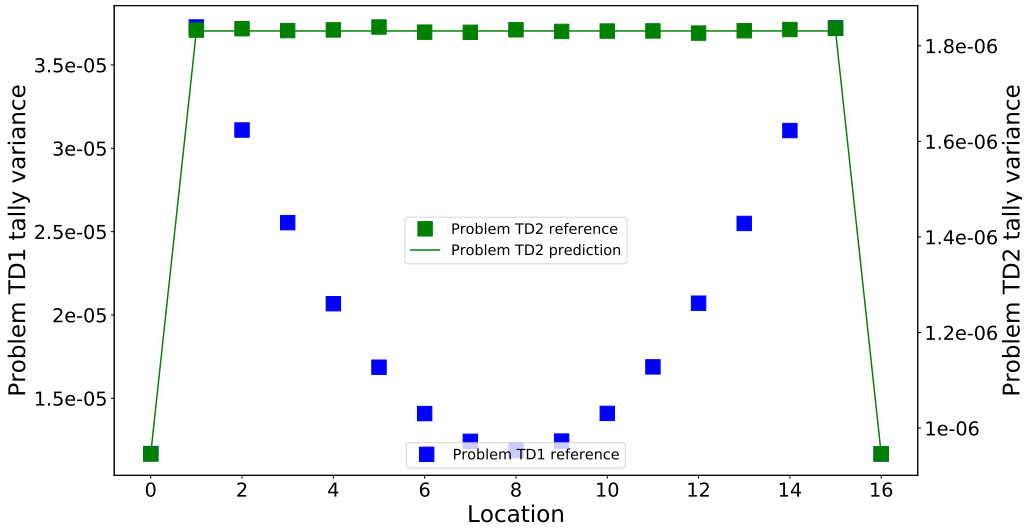
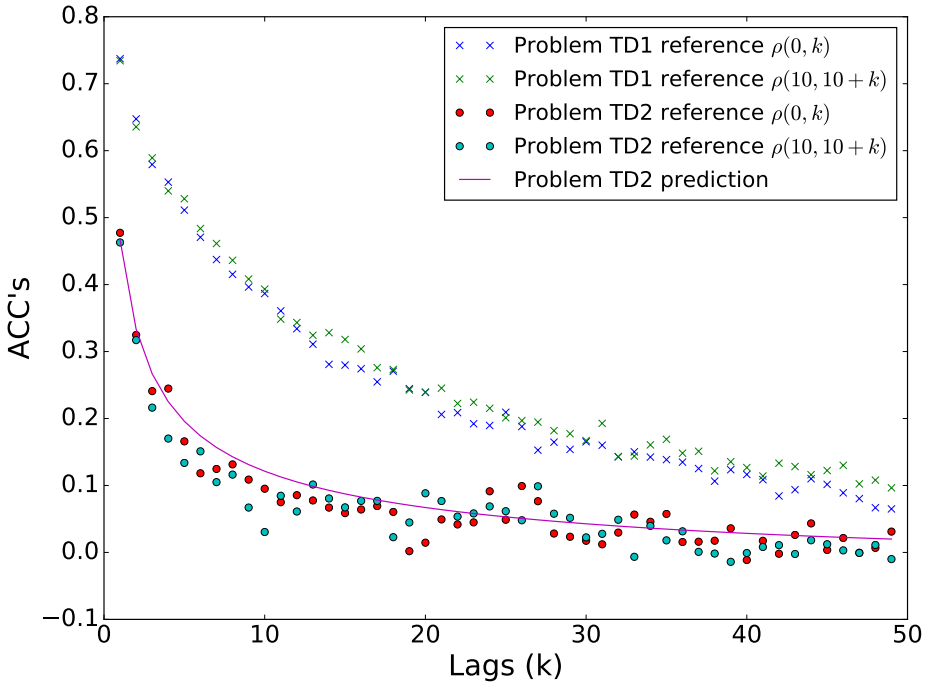


Figure 3-3: Variance of fission source tally of problem TD1 and TD2

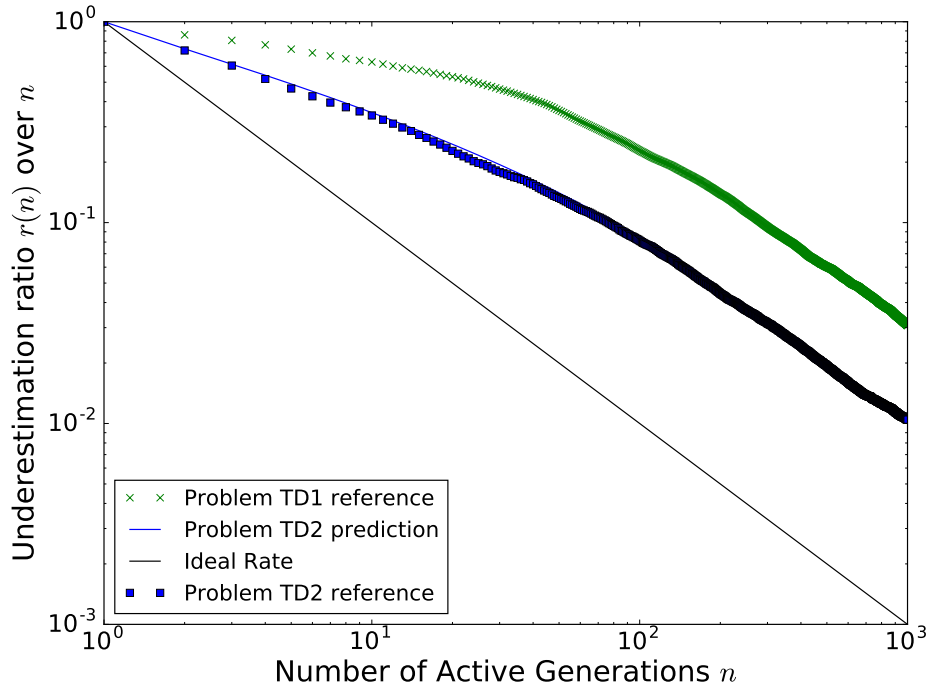
perfectly. This is because $TD2$ is fully characterized by the Markov Chain model. The deviation between predicted correlation coefficients and reference values for $TD1$ is not negligible.

The derivations in section 3.2.2 actually predict the correlation of fission source tally from one individual neutron and the result is then generalized to fission source tallies contributed by all neutrons in one generation under the assumption that all neutrons are independent. The fission event (the multiplicity effect discussed in section 2.4) was not taken into account in this model. Therefore, in addition to the generation-to-generation source dependence, the neutron multiplicity induced by fission events and the consequential correlation among neutrons within one generation contributes to a significant part of the generation-to-generation correlation of neutron tallies.

In Figure 3-4(b), the predicted variance underestimation ratio divided by active generation number $r_{Z_8}(n)/n$ calculated from correlation coefficients of neutron source tally in the central region $\rho_{Z_8}(d)$ from Eq 3.2.20 are plotted along with the numerical references from independent simulations of $TD1$ and $TD2$. For $TD1$, the prediction and reference values match well as expected from result of correlation coefficients. For



(a) correlation coefficients



(b) Variance Convergence Rate (variance underestimation ratio over active generation number)

Figure 3-4: Correlation Prediction from Markov Chain Method and reference of TD2

$TD2$, The deviation between predicted correlation coefficients $\rho_{Z_8}(d)$ and reference values also leads to the deviation of predicted variance underestimation ratio $r_{Z_8}(n)$.

Eq 3.2.20 approximates the autocorrelation coefficients of the fission source using information from the Markov transfer probability P only. This is a good estimation for a critical system with low variance on the number of new neutrons per absorption. The Markov method captures the correlation of the neutron source distribution at one generation with the previous generation which captures a large part of the autocorrelation of the fission source tally. However, fission/branching processes also play an important role that must be captured to improve the agreement which will be discussed in the next section.

3.3 Method of Multitype Branching Processes

3.3.1 Theory of Multitype Branching Processes

A branching process describes a population of individuals where each individual produces offsprings independently with identical distributions. Multitype branching process extends the model to a population of finite types of individuals, such as spatial position, energy and angle. An MBP model can approximate the Monte Carlo power iteration if the neutron phase space is discretized where each cell is treated as a unique type.

This subsection first defines the moment generating function (MGF) related to the MBP model. Then the MGF is used to extract the serial-spatial moments which can then be related to the ACCs.

3.3.1.1 Moment Generating Functions

(Factorial) Moment Generating Functions are used to systematically derive moments of the number of neutrons in different discretized phase space regions and at different generations. In probability theory, the factorial moment is a mathematical quantity defined as the expectation of the falling factorial of a random variable. Factorial

moments are useful for studying non-negative integer-valued random variables, and arise in the use of probability-generating functions to derive the moments of discrete random variables [2] [12].

For a random variable X, the moment generating function (MGF) of X of argument s is defined as

$$f_X(s) = \mathbb{E}\left[s^X\right] \tag{3.3.1}$$

Some important properties of MGF are used frequently in this work:

$$f_X(s)\big|_{s=1} = \mathbb{E}\left[s^X\right]\big|_{s=1} = 1 \tag{3.3.2}$$

$$\frac{\partial}{\partial s} f_X(s)\bigg|_{s=1} = \mathbb{E}\left[X s^{X-1}\right]\big|_{s=1} = \mathbb{E}X \tag{3.3.3}$$

$$\frac{\partial^2}{\partial s^2} f_X(s)\bigg|_{s=1} = \mathbb{E}\left[X(X-1) s^{X-2}\right]\big|_{s=1} = \mathbb{E}[X(X-1)] \tag{3.3.4}$$

$$\frac{\partial^3}{\partial s^3} f_X(s)\bigg|_{s=1} = \mathbb{E}\left[X(X-1)(X-2) s^{X-3}\right]\big|_{s=1} = \mathbb{E}[X(X-1)(X-2)] \tag{3.3.5}$$

Eq 3.3.2 shows a moment generating function evaluated at argument $s = 1$ which equals to 1. Eq 3.3.3, Eq 3.3.4 and Eq 3.3.5 relate the (factorial) moments of X to the derivatives of the MGF of X evaluated at $s = 1$. It is worthwhile to note that the n^{th} order derivative of MGF evaluated at $s = 1$ is not necessarily equal to the n^{th} order moment $\mathbb{E}[X^n]$ but instead the power of X grows in a way similar to that of a factorial of natural numbers. The moments in Eq 3.3.3, Eq 3.3.4 and Eq 3.3.5 are actually named factorial moments. And the MGF in Eq 3.3.1 is actually a factorial moment generating function.

The model developed from multitype branching process discretizes the neutron phase space over all independent variables into M discrete regions and denotes the system state at generation n with a vector $\vec{Z}(n)$ (Eq 3.3.6).

The l^{th} component of the vector corresponds to the number of neutrons belonging to the discrete phase space l at generation n. A neutron in region l is defined to be

of type l in the terminology of multitype branching processes.

$$\vec{Z}(n) = (Z_1(n), \cdots, Z_l(n), \cdots, Z_m(n)) \tag{3.3.6}$$

The total number of neutrons of all phase space regions is the sum of all components of $\vec{Z}(n)$ and is denoted as $Z(n)$ (Eq 3.3.7)

$$Z(n) = \sum_{i=1}^{M} Z_i(n). \tag{3.3.7}$$

The state vector at generation n is related to generation $n-1$ through

$$\vec{Z}(n) = \sum_{i=1}^{M} \sum_{j=1}^{Z_i(n-1)} \vec{Y}_{ij}, \tag{3.3.8}$$

where $\overrightarrow{Y_{ij}}$ is the state vector generated by the j^{th} neutron of type i at generation $n-1$.

The moment generating function of $\vec{Z}(n)$ is defined as

$$F_n(\vec{r}_0, \vec{s}) \equiv \mathbb{E}\left[\prod_{i=1}^{M} s_i^{Z_i(n)} \middle| \vec{Z}(0) = \vec{r}_0\right] = \sum_{\vec{r}} \mathbb{P}\left(\vec{Z}(n) = \vec{r} \middle| \vec{Z}(0) = \vec{r}_0\right) \prod_{i=1}^{M} s_i^{r_i}, \tag{3.3.9}$$

where $\vec{r}_0$ denotes the initial configuration of neutrons in the discretized phase space and $\vec{s}$ is the argument of the moment generating function.

If $\vec{r}_0$ is a point source of type i ($\vec{r_0} \equiv \vec{e}_i$,$(r_0)_j = \delta^i_j$), we denote $F_n\left(\vec{r}_0, \vec{s}\right)$ as $F_n(i, \vec{s})$. The random vector $\vec{Z}(n)$ initiated by $\vec{Z}(0)$ is the sum of the random vectors initiated by the $Z(0)$ neutrons represented by the vector $\vec{Z}(0)(= \vec{r_0})$

$$\vec{Z}(n)\Big|_{\vec{Z}(0)=\vec{r_0}} = \sum_{i=1}^{M} \sum_{j=1}^{r_0^i} \vec{Z}(n)\Big|_{\vec{Z}(0)=\vec{e}_i} \tag{3.3.10}$$

Eq 3.3.10 expresses $\vec{Z}(n)$ as a sum of contributions from all neutrons in the 0^{th} generation. Since the $Z(0)$ random vectors are independent, the moment generating function of the state vector $\vec{Z}(n)\Big|_{\vec{Z}(0)=\vec{r_0}}$ is written as the product of the moment gen-

erating function of the state vectors of the $Z(0)$ components $\vec{Z}(n)\Big|_{\vec{Z}(0)=\vec{e}_i}$. Therefore, $F_n(\vec{r}_0, \vec{s})$ and $F_n(i, \vec{s})$ are related by

$$F_n(\vec{r}_0, \vec{s}) = \prod_{i=1}^{M} F_n(i, \vec{s})^{r_0^i}. \tag{3.3.11}$$

$F_n(i, \vec{s})$ can be evaluated from $F_1(i, \vec{s})$ recursively according to its definition in Eq 3.3.9,

$$F_n(i, \vec{s}) = \mathbb{E}\left[\prod_{l=1}^{M} s_l^{Z_l(n)} \,\middle|\, \vec{Z}(0) = \vec{e}_i\right] \tag{3.3.12}$$

$$= \mathbb{E}\left[\mathbb{E}\left[\prod_{l=1}^{M} s_l^{Z_l(n)} \,\middle|\, \vec{Z}(n-1)\right] \,\middle|\, \vec{Z}(0) = \vec{e}_i\right]. \tag{3.3.13}$$

where Eq 3.3.13 is based on the Markov property of process $\vec{Z}(n)$.

Substituting the relation between $\vec{Z}(n)$ and $\vec{Z}(n-1)$ (Eq 3.3.8) into the conditional expectation yields

$$\begin{aligned}
\mathbb{E}\left[\prod_{l=1}^{M} s_l^{Z_l(n)} \,\middle|\, \vec{Z}(n-1)\right] =& \mathbb{E}\left[\prod_{l=1}^{M}\prod_{k=1}^{M}\prod_{j=1}^{Z_k(n-1)} s_l^{(\overrightarrow{Y_{k,j}})_l} \,\middle|\, \vec{Z}(n-1)\right] \\
=& \mathbb{E}\left[\prod_{k=1}^{M}\prod_{j=1}^{Z_k(n-1)}\prod_{l=1}^{M} s_l^{(\overrightarrow{Y_{k,j}})_l} \,\middle|\, \vec{Z}(n-1)\right] \\
=& \prod_{k=1}^{M}\prod_{j=1}^{Z_k(n-1)} \mathbb{E}\left[\prod_{l=1}^{M} s_l^{(\overrightarrow{Y_{k,j}})_l}\right] \\
=& \prod_{k=1}^{M}\left(\mathbb{E}\left[\prod_{l=1}^{M} s_l^{(\overrightarrow{Y_{k,j}})_l}\right]\right)^{Z_k(n-1)}
\end{aligned} \tag{3.3.14}$$

Because the neutrons at generation $n-1$ reproduce independently, the product over k and j on the second line of Eq 3.3.14 can be pulled out of the expectation operator. By the definition of $\overrightarrow{Y_{k,j}}$,

$$\mathbb{E}\left[\prod_{l=1}^{M} s_l^{(\overrightarrow{Y_{k,j}})_l}\right] = F_1(k,\vec{s}). \tag{3.3.15}$$

Combining the conditional expectation and Eq 3.3.13 gives

$$F_n(i,\vec{s}) = \mathbb{E}\left[\prod_{k=1}^{M} (F_1(k,\vec{s}))^{Z_k(n-1)} \,\middle|\, \vec{Z}(0) = \vec{e}_i\right] \tag{3.3.16}$$

which is the moment generating function of $\vec{Z}(n-1)$ with $\vec{f}(1,\vec{s})$ as the argument, where the vector $\vec{f}(1,\vec{s})$ is defined as a vector with $F_1(k,\vec{s})$ as its k^{th} component:

$$\vec{f}(1,\vec{s}) = (F_1(1,\vec{s}),\cdots,F_1(m,\vec{s})) \tag{3.3.17}$$

Therefore the recursive relation between $F_n(i,\vec{s})$ and $F_{n-1}(i,\vec{s})$ reads

$$F_n(i,\vec{s}) = F_{n-1}(i,\vec{f}(1,\vec{s})) \tag{3.3.18}$$

This equivalence can be demonstrated by induction on the above equation which gives

$$\begin{aligned} F_n(i,\vec{s}) &= F_{n-1}\left(i,\vec{f}(1,\vec{s})\right) \\ &= F_{n-2}\left(i,\vec{f}\left(1,\vec{f}(1,\vec{s})\right)\right) \\ &\equiv F_{n-2}\left(i,\vec{f}(2,\vec{s})\right) \\ &= \cdots \\ &= F_1(i,\vec{f}(n-1,\vec{s})) \end{aligned} \tag{3.3.19}$$

where $\vec{f}(n,\vec{s})$ is naturally defined as:

$$\vec{f}(n,\vec{s}) = \vec{f}(n-1,\vec{f}(1,\vec{s})),\ n \geq 2 \tag{3.3.20}$$

Combining the definitions of $\vec{f}(n,\vec{s})$(Eq 3.3.17, Eq 3.3.20) and the last recursive

evaluation of $F_n(i, \vec{s})$ (Eq 3.3.19) yields

$$F_n(i, \vec{s}) = (\vec{f}(n, \vec{s}))_i \tag{3.3.21}$$

which means $F_n(i, \vec{s})$ is the i^{th} element of the vector function $\vec{f}(n, \vec{s})$.

For convenience, denote the l^{th} component of $\vec{f}(1, \vec{s})$ (also $F_1(l, \vec{s})$) as $f_l(\vec{s})$

$$f_l(\vec{s}) \equiv \mathbb{E}\left[\prod_{i=1}^{M} s_i^{Z_i(1)} \,|Z(0)_i = \delta_{i,l}\right] = \sum_{\vec{r}} \mathbb{P}\left(\vec{Z}(1) = \vec{r} \,\middle|\, \vec{Z}(0) = \vec{r}_0\right) \prod_{i=1}^{M} s_i^{r_i}, \tag{3.3.22}$$

3.3.1.2 Spatial Moments

The spatial moments of $Z_l(n)$ defined in Eq 3.3.23 can be evaluated by taking derivatives of $F_n(\vec{r}_0, \vec{s})$.

$$\begin{aligned} \mu_l(n) &\equiv \mathbb{E}[Z_l(n)] \\ C_{l,j}(n) &\equiv \mathbb{E}[Z_l(n) Z_j(n)] \\ T_{l,j,k}(n) &\equiv \mathbb{E}[Z_l(n) Z_j(n) Z_k(n)] \end{aligned} \tag{3.3.23}$$

This would involve commuting the derivatives into the recursive definition of $\vec{f}(n, \vec{s})$ (Eq 3.3.20) which can be quite lengthy.

A simpler way is to evaluate the expectations at generation $n+1$ conditional on $\vec{Z}(n)$. Then take the expectation with respect to the distribution of $\vec{Z}(n)$. Generally, for any function f of $\vec{Z}(n)$

$$\mathbb{E}[f(\vec{Z}(n+1))] = \mathbb{E}[\mathbb{E}[f(\vec{Z}(n+1))|\vec{Z}(n)]] \tag{3.3.24}$$

$$= \mathbb{E}[g_{n+1,n}(\vec{Z}(n))] \tag{3.3.25}$$

where Eq 3.3.24 inserts the conditional expectation and Eq 3.3.25 evaluates the expectation with respect to the distribution of $\vec{Z}(n)$. $g_{n+1,n}(\vec{Z}(n))$ is written to emphasize the dependence of $\vec{Z}(n)$ only.

From Eq 3.3.25, an equation that relates the moments at $n+1$ with moments

at n will show up. In most instances of $f(\vec{Z}(n)$ derived below, $g_{n+1,n}(\vec{Z}(n))$ includes dependence of $f(\vec{Z}(n)$. Therefore, $\mathbb{E}[f(\vec{Z}(n+1))]$ contains dependence of $\mathbb{E}[f(\vec{Z}(n))]$. Thus the moments at generation $n+1$ can be evaluated from the moments at generation n via conditional expectation. The recursive evaluation of moments is termed evolution. For simplicity in derivations to follow, $\mathbb{E}[X|Y]$ is denoted as $\mathbb{E}[X]|_Y$.

Similarly to the development of Eq 3.3.16, the moment generating function at generation $n+1$ can be written as

$$\begin{aligned} F_{n+1}\left(\vec{r_0}, \vec{s}\right)\Big|_{\vec{Z}(n)} &= F_1(\vec{Z}(n), \vec{s}) \\ &= \prod_{k=1}^{M} F_1(k, \vec{s})^{Z_k(n)} \\ &= \prod_{k=1}^{M} f_k(\vec{s})^{Z_k(n)}. \end{aligned} \tag{3.3.26}$$

From the moment generating function in Eq 3.3.26, $\mu_i(n+1)$ is first evaluated as a conditional expectation on $\vec{Z}(n)$ and then expressed as a function of $\mu_i(n)$.

$$\begin{aligned} \mu_i(n+1)|_{\vec{Z}(n)} &= \frac{\partial}{\partial s_i} \prod_{k=1}^{M} f_k(\vec{s})^{Z_k(n)}\Bigg|_{\vec{s}=\vec{1}} \\ &= \sum_{l=1}^{M} f_l(\vec{s})^{Z_l(n)-1} Z_l(n) \frac{\partial}{\partial s_i} f_l(\vec{s}) \prod_{k\neq l} f_k(\vec{s})^{Z_k(n)}\Bigg|_{\vec{s}=\vec{1}} \\ &= \sum_{l=1}^{M} Z_l(n) \frac{\partial}{\partial s_i} f_l(\vec{s})\Big|_{\vec{s}=\vec{1}} \qquad (3.3.27) \\ &= \sum_{l=1}^{M} Z_l(n) M_i^l \qquad (3.3.28) \end{aligned}$$

The identity $F_1(l, \vec{s})|_{\vec{s}=\vec{1}} = 1$ (Eq 3.3.2) is used in Eq 3.3.27. In Eq 3.3.28 M_i^l is used to denote $\frac{\partial}{\partial s_i} f_l(\vec{s})\Big|_{\vec{s}=\vec{1}}$. By definition of the moment generating function, M_i^l is the expected number of neutrons found of type i after one generation given a source

neutron of type l. It will be referred as the first spatial moment response:

$$M_i^l \equiv \frac{\partial}{\partial s_i} f_l(\vec{s})\Big|_{\vec{s}=\vec{1}} \tag{3.3.29}$$

$$= \mathbb{E}[Z_i(1)|\vec{r}_0 = \vec{e}_l] \tag{3.3.30}$$

By adding a superscript to the notation $\mu_i(n)$ to describe the type of source neutron, M is also recognized as

$$M_i^l \equiv \mu_i^l(1) \tag{3.3.31}$$

Due to the quite frequent usage of the first order moment response, the new notation M_i^l will be used.

The above derivation and many to come require the derivatives of the moment generating function $F_{n+1}\left(\vec{r_0}, \vec{s}\right)\Big|_{\vec{Z}(n)}$. The needed derivatives are evaluated in Appendix A1.4. For example, Eq 3.3.28 can be obtained from Eq A1.4.2 by neglecting the $\vec{\sigma}$ term. The following derivations will rely on the material in Appendix A1.4.

Evaluating the expected value of Eq 3.3.28 with respect to the distribution of $\vec{Z}(n)$ removes the dependence on $\vec{Z}(n)$ and yields

$$\mu_i(n+1) = \mu_l(n) M_i^l. \tag{3.3.32}$$

where the Einstein tensor notation is used (the sum is taken over all values of the index whenever the same symbol appears as a subscript and superscript in the same term) and will be used throughout this work.

A similar procedure will provide the second order factorial moments, $\text{Ⅽ}_{l,j}(n)$, which can eventually be related to $C_{l,j}(n)$ (Eq 3.3.23) through Eq 3.3.33.

$$\text{Ⅽ}_{l,j}(n) = \frac{\partial}{\partial s_l}\frac{\partial}{\partial s_j} F_n(\vec{r}_0, \vec{s}) = \mathbb{E}[Z_l(n)Z_j(n) - \delta_{l,j}Z_l(n)] \tag{3.3.33}$$

The random variable within the expectation of $\text{Ⅽ}_{l,j}(n)$, $Z_l(n)Z_j(n) - \delta_{l,j}Z_l(n)$ can interpreted as the number of pair of neutrons with one neutron in region l and one

neutron in region j at generation n. When l and j are not the same region, the number of pairs is just the product of the number of neutrons in each region. Otherwise, for $Z_l(n)$ neutrons in region l, there are $Z_l(n)(Z_l(n)-1)$ pairs formed by these neutrons.

$\text{Ç}_{i,j}(n+1)$ is first evaluated conditional on $\vec{Z}(n)$ and then expressed as a function of $\text{Ç}_{i,j}(n)$,

$$\text{Ç}_{j,i}(n+1)|_{\vec{Z}(n)} = \frac{\partial}{\partial s_j}\frac{\partial}{\partial s_i}\prod_{k=1}^{M} f_k(\vec{s})^{Z_k(n)}\bigg|_{\vec{s}=\vec{1}} \tag{3.3.34}$$

$$= \sum_{l=1}^{M}\sum_{h=1}^{M} Z_l(n)Z_h(n)M_i^l M_j^h + Z_l(n)(\bar{\text{V}}_{j,i}^l - M_j^l M_i^l). \tag{3.3.35}$$

where Eq 3.3.34 follows definition and Eq 3.3.35 follows Eq A1.4.10. A new response moment of the second order is defined in Eq 3.3.35. $\bar{\text{V}}_{a,b}^l$ is defined as

$$\begin{aligned}\bar{\text{V}}_{a,b}^l &\equiv \frac{\partial^2}{\partial s_a \partial s_b} f_l(\vec{s})\bigg|_{\vec{s}=\vec{1}} \\ &= \mathbb{E}\left[Z_a(1)Z_b(1) - Z_a(1)\delta_{a,b} \middle| Z_i(0) = \delta_{i,l}\right]\end{aligned} \tag{3.3.36}$$

Similarly, the new notation $\bar{\text{V}}_{a,b}^l$ is created even though it can be recognized as $\text{Ç}_{j,i}$ with an explicitly specified type of source neutron.

$$\bar{\text{V}}_{j,i}^l \equiv \text{Ç}_{j,i}^l(1) \tag{3.3.37}$$

Taking the expected value of $\text{Ç}_{j,i}(n+1)|_{\vec{Z}(n)}$ (Eq 3.3.35) removes the dependence of $\vec{Z}(n)$ and yields

$$\text{Ç}_{j,i}(n+1) = M_j^h M_i^l C_{h,l}(n) + \mu_l(n)\left(\bar{\text{V}}_{j,i}^l - M_j^l M_i^l\right) \tag{3.3.38}$$

Substituting the relation between $\text{Ç}_{j,i}$ and $C_{j,i}$ (Eq 3.3.33) simplifies Eq 3.3.38 to

$$C_{j,i}(n+1) = M_j^h M_i^l C_{h,l}(n) + \mu_l(n) V_{j,i}^l \tag{3.3.39}$$

where

$$V_{i,j}^l \equiv \bar{\mathbb{V}}_{i,j}^l + M_i^l \delta_{i,j} - M_i^l M_j^l \tag{3.3.40}$$

$V_{i,j}^l$ is constructed from $\bar{\mathbb{V}}_{i,j}^l$ by first adding the diagonal terms $M_i^l \delta_{i,j}$ due to the difference between factorial and ordinary moments and then subtracting the first order moments $M_i^l M_j^l$ to produce central moments. And $V_{i,j}^l$ becomes

$$V_{i,j}^l = \mathbb{E}[(Z_i(1) - \mu_i(1))(Z_j(1) - \mu_j(1))|\vec{r}_0 = \vec{e}_l], \tag{3.3.41}$$

with $V_{i,j}^l$ being the covariance between the number of new neutrons of type i and type j after one generation given a neutron born of type l. It will be referred to as the second spatial (central) moment response.

3.3.1.3 Numerical results

This section verifies the recursive formula of first and second order moments (Eq 3.3.32 and Eq 3.3.39) derived so far The test problem $TD1$ is used to calculate numerical references and perform MBP predictions on. The response moments M and V required can be calculated from the transfer matrix (Eq 2.1.5). The details will be discussed in section 3.3.4.1.

The spatial moments derived above are now compared to reference calculations on test problem $TD1$. In Figure 3-5, the solid curves of different colors correspond to the prediction at different generations. The squares correspond to reference values calculated from independent simulations. The prediction curves and reference value squares share the same color for results at the same generation. For convenience, the color code legends are given in Figure 3-5(a) and 3-5(b) only.

Figure 3-5(a) plots the expectation of neutron counts as a function of tally region location for generations $n \in \{25, 75, 125, 175, 225, 275\}$. The expectation in neutron count distribution has reached the eigenmode of the test problem $TD1$ (the equilibrium distribution of Eq 2.1.5 scaled by number of neutrons per generation). And the prediction of expectation of neutron counts matches the reference well. The results over different generations overlap at the eigenmode, and almost only results at

generation 275 are visible.

Figure 3-5(c) plots the variance of neutron tally in all regions along with the corresponding expectations

$$\mathrm{Var}[Z_i(n)] = C_{i,i}(n) - \mu_i(n)\mu_i(n) \tag{3.3.42}$$

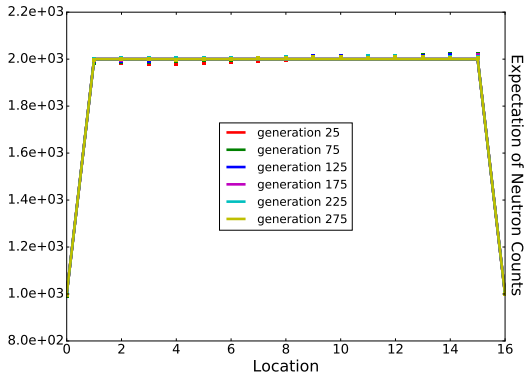
as function of tally region at generations $n \in \{25, 75, 125, 175, 225, 275\}$. The variance of the counts keeps growing as the simulation continues. The predicted variance also matches the reference calculated from $TD1_f$ simulations.

It is worthwhile to mention that the simulations performed on $TD1_f$ follows the same process as the MBP model where no neutron number normalization is performed which explains the good agreement between prediction and reference. The reference values of $\mu_i(n)$ and $\mathrm{Var}[Z_i(n)]$ calculated from $TD1_n$ are also plotted alongside the prediction from the MBP model in Figure 3-5(b) and Figure 3-5(d). The MBP model does not accurately predict the variance on $TD1_n$ since it differs considerably from $TD1_f$. As discussed in section 3.1.2, this issue will be resolved by converting the predictions on Z tallies to X tallies.

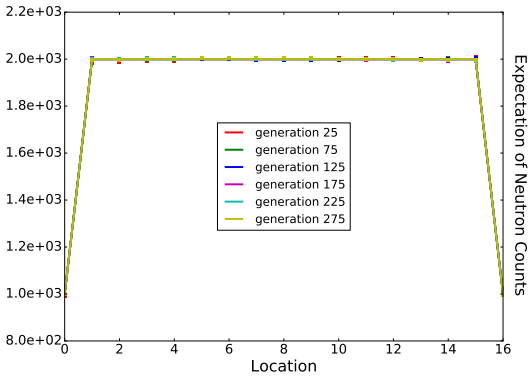
The focus for the moment is on verifying the predictive capability of moments using the MBP model. Figure 3-5(e) and Figure 3-5(f) show the covariance

$$\mathrm{Cov}[Z_i(n), Z_j(n)] = C_{i,j}(n) - \mu_i(n)\mu_j(n) \tag{3.3.43}$$

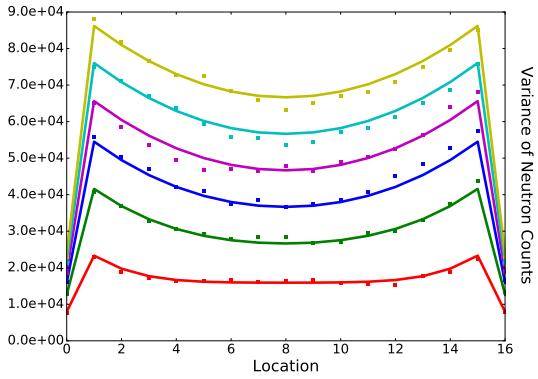
as a function of the tally region j at generations $n \in \{25, 75, 125, 175, 225, 275\}$ for i equal to 8 and 1 respectively. Both figures demonstrate the consistency between the numerical reference and the MBP prediction. It also can be seen that the covariance between region i and region j is highest when $j = i$ and decreases as function of the distance between i and j.



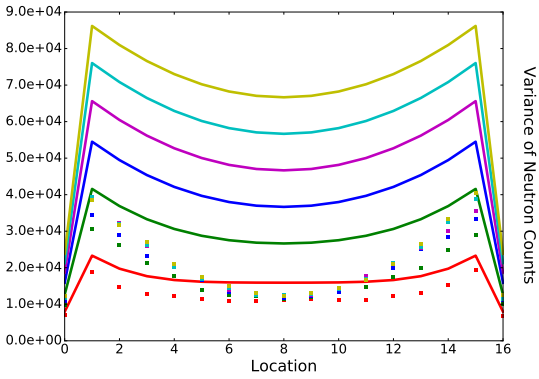
(a) $\mu_i(n)$ and reference from $TD1_f$, simulation without neutron number normalization



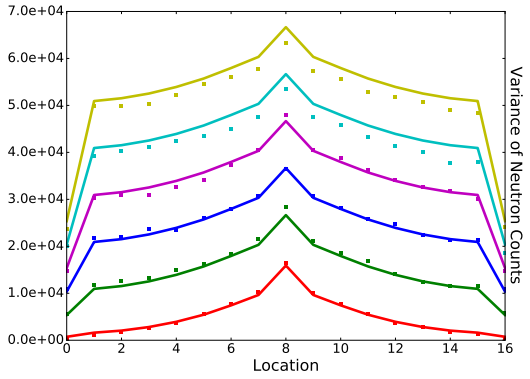
(b) $\mu_i(n)$ and reference from $TD1_n$, simulation with neutron number normalization



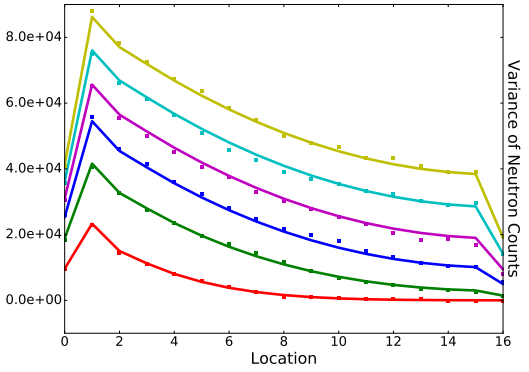
(c) Var $Z_i(n)$ and reference from simulation on $TD1_f$



(d) Var $Z_i(n)$ and reference from simulation on $TD1_n$



(e) $\mathrm{Cov}(Z_8(n), Z_i(n))$ and reference from simulation on $TD1_f$



(f) $\mathrm{Cov}(Z_1(n), Z_i(n))$ and reference from simulation on $TD1_f$

Figure 3-5: Spatial moments of test problem $TD1$. The solid curves in both figures correspond to $C_{i,i}(n) - \mu_i(n)\mu_i(n)$ at generations $n \in \{25, 75, 125, 175, 225, 275\}$ calculated according to Eq 3.3.32 and Eq 3.3.39.

3.3.1.4 Serial Moments

To calculate the generation-to-generation correlation, the spatial moments in the form of $\mathbb{E}[Z_i(n)]$, $\mathbb{E}[Z_i(n)Z_j(n)]$, $\mathbb{E}[Z_i(n)Z_j(n)Z_k(n)]$, are not sufficient, the serial-spatial moments in the form of $\mathbb{E}[Z_i(n)Z_j(n+k)]$, are also required. Therefore, the joint moment generating function of $\vec{Z}(n)$ and $\vec{Z}(n+k)$ is needed (where k is the generation lag):

$$\begin{aligned} F_{n,n+k}\left(\vec{r_0},\vec{s},\vec{t}\right) =& \mathbb{E}\left[\vec{Z}(n)\vec{Z}(n+k)\middle|\vec{Z}(0)=\vec{r_0}\right] \\ =& \sum_{\vec{r},\vec{q}}\mathbb{P}\left(\vec{Z}(n)=\vec{r},\vec{Z}(n+k)=\vec{q}\,\middle|\,\vec{Z}(0)=\vec{r_0}\right)\prod_{i=1}^{M}s_i^{r_i}t_i^{q_i} \end{aligned} \tag{3.3.44}$$

The relation between the moment generating function ($F_n(\vec{r_0},\vec{s})$) of $\vec{Z}(n)$ and the joint moment generating function ($F_{n,n+k}(\vec{r_0},\vec{s},\vec{t})$) of $\vec{Z}(n)$ and $\vec{Z}(n+k)$ can be found:

$$\begin{aligned} F_{n,n+k}\left(\vec{r_0},\vec{s},\vec{t}\right) &= \sum_{\vec{r},\vec{q}}\mathbb{P}\left(\vec{Z}(n)=\vec{r},\vec{Z}(n+k)=\vec{q}\,\middle|\,\vec{Z}(0)=\vec{r_0}\right)\prod_{i=1}^{M}s_i^{r_i}t_i^{q_i} \\ &= \sum_{\vec{r},\vec{q}}\mathbb{P}\left(\vec{Z}(n+k)=\vec{q}\,\middle|\,\vec{Z}(n)=\vec{r}\right)\mathbb{P}\left(\vec{Z}(n)=\vec{r}\,\middle|\,\vec{Z}(0)=\vec{r_0}\right)\prod_{i=1}^{M}s_i^{r_i}t_i^{q_i} \\ &= \sum_{\vec{r}}\mathbb{P}\left(\vec{Z}(n)=\vec{r}\,\middle|\,\vec{Z}(0)=\vec{r_0}\right)\left[\sum_{\vec{q}}\mathbb{P}\left(\vec{Z}(n+k)=\vec{q}\,\middle|\,\vec{Z}(n)=\vec{r}\right)\prod_{j=1}^{M}t_j^{q_j}\right]\prod_{i=1}^{M}s_i^{r_i} \\ &= \sum_{\vec{r}}\mathbb{P}\left(\vec{Z}(n)=\vec{r}\,\middle|\,\vec{Z}(0)=\vec{r_0}\right)\left[F_k(\vec{r},\vec{t}\,)\right]\prod_{i=1}^{M}s_i^{r_i} \\ &= \sum_{\vec{r}}\mathbb{P}\left(\vec{Z}(n)=\vec{r}\,\middle|\,\vec{Z}(0)=\vec{r_0}\right)\left[\prod_{j=1}^{M}F_k(j,\vec{t})^{r_j}\right]\prod_{i=1}^{M}s_i^{r_i} \\ &= \sum_{\vec{r}}\mathbb{P}\left(\vec{Z}(n)=\vec{r}\,\middle|\,\vec{Z}(0)=\vec{r_0}\right)\left[\prod_{j=1}^{M}F_k(j,\vec{t})^{r_j}s_j^{r_j}\right] \end{aligned} \tag{3.3.45}$$

It shows that the joint moment generating function $F_{n,n+k}(\vec{r_0},\vec{s},\vec{t})$ satisfies the

functional equation

$$\begin{aligned} F_{n,n+k}\left(\vec{r}_0, \vec{s}, \vec{t}\right) &= F_n(\vec{r}_0, \vec{u}(k)) \\ u_i(k) &= s_i F_k(i, \vec{t}) \equiv s_i f_i^{(k)}(\vec{t}) \end{aligned} \tag{3.3.46}$$

From the joint moment generating function, moments in the form of $\mathbb{E}[Z_I(n)Z_I(n')]$ can be evaluated and contribute directly to the evaluation of correlation coefficients $\rho_I(n, n'-n)$ as defined in Eq 3.0.1. A generalized version of the moment $\mathbb{E}[Z_l(n)Z_j(n+k)]$ will also be required and is derived here.

$$\mathbb{E}Z_l(n)Z_j(n+k) = \frac{\partial}{\partial s_l}\frac{\partial}{\partial t_j}F_n(\vec{r_0}, \vec{u}(k))\Big|_{\vec{s}=\vec{1},\vec{t}=\vec{1}} \tag{3.3.47}$$

where the derivative can be evaluated explicitly using the chain-rule and repeatedly replacing in Eq 3.3.46.

The second order derivative of the joint moment generating function is derived and appendix (Eq A1.3.6) and stated below

$$\begin{aligned} &\frac{\partial^2 F_n(\vec{Z}(0), \vec{u}(k))}{\partial s_l \partial t_j} \\ &= \frac{\partial F_k(h, \vec{t})}{\partial t_j}\frac{\partial^2 F_n\left(\vec{Z}(0); \vec{u}(k)\right)}{\partial u_l(k)\partial u_h(k)}F_k(l, \vec{t})s_h + \frac{\partial F_k(l, \vec{t})}{\partial t_j}\frac{\partial F_n\left(\vec{Z}(0); \vec{u}(k)\right)}{\partial u_l(k)} \end{aligned} \tag{3.3.48}$$

Evaluating at $\vec{s} = \vec{t} = \vec{1}$ (and consequentially $\vec{u}(k) = \vec{1}$), the above equation leads to

$$\mathbb{E}Z_l(n)Z_j(n+k) = C_{l,h}(n)\mu_j^h(k) + \mu_j^l(k)\mu_l(n) = C_{l,h}(n)\mu_j^h(k), \tag{3.3.49}$$

where $\mu_j^l(n)$ is an extension of the definition of $M_j^l (= \mu_j^l(1))$.

$$\mu_j^l(n) \equiv \mathbb{E}\left[Z_j(n) \,|\vec{r_0} = \vec{e}_l\right] \tag{3.3.50}$$

Then insert the relation between ordinary and factorial second order spatial mo-

ments $\mathcal{C}_{i,j}(n)$ and $C_{i,j}(n)$ of Eq 3.3.33 in Eq 3.3.49

$$\mathbb{E}Z_l(n)Z_j(n+k) = C_{l,h}(n)\mu_j^h(k), \tag{3.3.51}$$

Eq 3.3.51 reveals a large component of the origin of the auto-correlation. Given the number of neutrons in one cell at one generation, the number of neutrons in this cell at future generations is correlated in two parts: 1) neutrons from a given cell remain or re-enter a given cell at a later generation, and 2) neutrons initially in another cell enters a given cell at a later generation. The first part is determined by the diagonal elements of the spatial covariance matrix. Contributions from neighboring cells is determined by the off-diagonal elements of the spatial covariance matrix.

The explicitly written moments enables evaluating correlation coefficients of tally Z (if replacing all X's with Z's in Eq 3.0.1).

$$\rho_{I,Z}(n,k) = \frac{Cov[Z_I(n), Z_I(n+k)]}{\sqrt{Var[Z_I(n)]\,\mathrm{Var}[Z_I(n+k)]}} \tag{3.3.52}$$

To illustrate, the predicted correlation coefficients for tally Z_8 (center region of problem $TD1_f$) are plotted along with the reference values calculated from independent simulations in Figure 3-6.

Since problem $TD1_f$ can be completely characterized by an MBP model, good agreement is expected and shown with the independent simulations. However, the corresponding reference correlation coefficients for $TD1_n$, also plotted in Figure 3-6, indicate that the normalization correlation is quite impactful.

3.3.2 Approximating the ACCs

Section 3.3.1 derived various moments to evaluate the correlation coefficients for Z tallies of an MBP model. Due to the neutron number normalization steps in realistic Monte Carlo simulations, the correlation behavior of Z tallies in a realistic simulation is quite different then the proposed MBP model. Prior discussion in section 3.1.2 implied that the normalization correlation should be negligible for X tallies. At

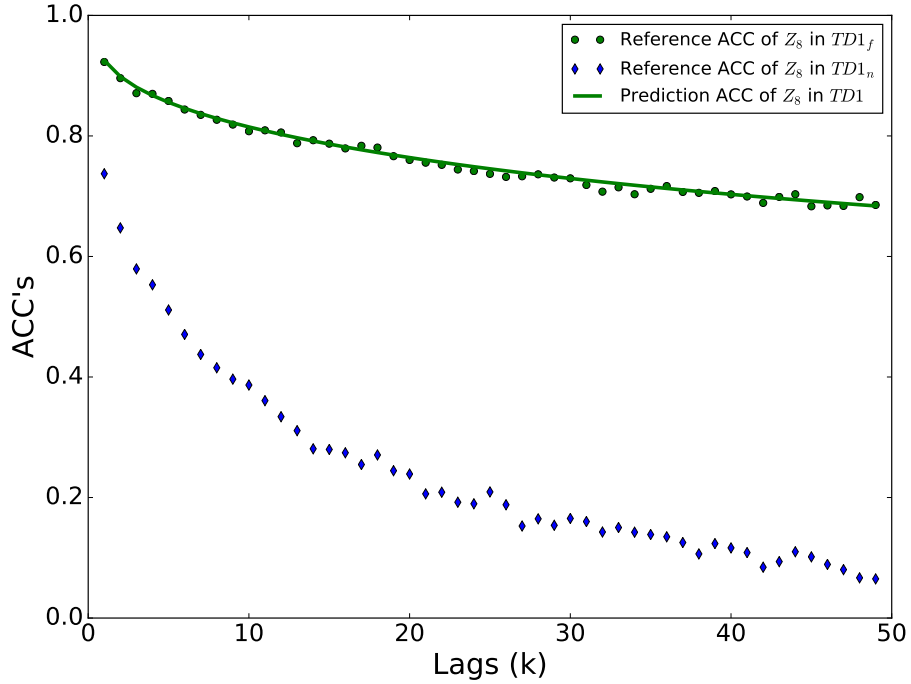


Figure 3-6: Correlation coefficients of problem TD1 for tally Z_8. The green solid curve is the predictive model. The green dots are the reference values from independent simulations of problem $TD1_f$. The blue crosses are the corresponding reference correlation coefficients of problem $TD1_n$, where neutron normalization at each generation is enforced.

the same time, in realistic simulations, an X tally is merely the corresponding Z tally scaled by a constant, the number of neutrons simulated per generation. This section will adjust the MBP model from Z tallies to X tallies to better represent the Monte Carlo power iteration with normalization and explicitly write the correlation coefficients of the X tallies in term of the spatial and serial moments derived in section 3.3.1.

3.3.2.1 Expansion of fission source distribution

Before specifying the serial-spatial moments to be evaluated from the moment generating functions $F_n(\vec{r_0}, \vec{s})$ and $F_{n,n+k}(\vec{r_0}, \vec{s}, \vec{t})$, correlation coefficients of X tallies must be explicitly expressed in the form of such moments.

Taking derivatives of moment generating functions in Eq 3.3.9 and Eq 3.3.44 gives an expectation of products of $Z_l(n)$ such as $\mathbb{E}[Z_l(n)]$, $\mathbb{E}[Z_l(n)Z_j(n)]$ and $\mathbb{E}[Z_l(n)Z_l(n+k)]$. However, to evaluate the variance and covariance terms in Eq 3.0.1, expectations in the form of $\mathbb{E}[X_l(n)]$ and $\mathbb{E}[X_l(n)X_l(n+k)]$ are required.

To perform the transformation, the definition of $X_l(n)$ (Eq 3.1.4) is viewed as a function of $\vec{Z}(n)$ and expanded around $\mathbb{E}[\vec{Z}(n)]$ ($\equiv \vec{\mu}(n)$):

$$\begin{aligned} X_l(n) &\equiv g_l(\vec{Z}(n)) = g_l(\vec{\mu}(n)) \\ &+ \sum_{i=1}^{M} \left.\frac{\partial g_l(\vec{Z}(n))}{\partial Z_i(n)}\right|_{\vec{Z}(n)=\vec{\mu}(n)} (Z_i(n) - \mu_i(n)) \\ &+ \sum_{i,j=1}^{M} \left.\frac{\partial^2 g_l(\vec{Z}(n))}{\partial Z_i(n)\partial Z_j(n)}\right|_{\vec{Z}(n)=\vec{\mu}(n)} (Z_i(n) - \mu_i(n))(Z_j(n) - \mu_j(n)) \\ &+ \cdots \end{aligned} \tag{3.3.53}$$

where the derivatives of $g_l(\vec{Z}(n))$ can be calculated explicitly from Eq 3.1.4 as

$$\left.\frac{\partial g_l(\vec{Z}(n))}{\partial Z_i(n)}\right|_{\vec{Z}(n)=\vec{\mu}(n)} \equiv g_l'_i(n) = \frac{\delta_{i,l}}{\mu(n)} - \frac{\mu_l(n)}{\mu(n)^2} \tag{3.3.54}$$

$$\left.\frac{\partial^2 g_l(\vec{Z}(n))}{\partial Z_i(n)\partial Z_j(n)}\right|_{\vec{Z}(n)=\vec{\mu}(n)} \equiv g_{l\,i,j}''(n) = -\frac{\delta_{i,l}}{\mu(n)^2} - \frac{\delta_{j,l}}{\mu(n)^2} + 2\frac{\mu_l(n)}{\mu(n)^3} \tag{3.3.55}$$

where

$$\mu(n) \equiv \sum_{i=1}^{M} \mu_i(n). \tag{3.3.56}$$

The covariance terms required to evaluate $\rho_{X_l}(n,k)$ (Eq 3.0.1) are of the form $\mathbb{E}\left[(X_l - \mathbb{E}X_l)(X_l - \mathbb{E}X_l)\right]$. Therefore X_l and $\mathbb{E}X_l$ must preserve terms of the same order following an expansion. $g_l(\vec{Z}(n))$ (corresponding to X_l) and $\mathbb{E}g_l(\vec{Z}(n))$ (corresponding to $\mathbb{E}X_l$) will be expressed explicitly below to obtain consistent expansion orders. For simplicity, the dependence on generation n and neutron type l is suppressed and the first and second order derivatives of $g(\vec{Z}(n))$ evaluated at $\vec{\mu}(n)$ are denoted as g_i' and $g_{i,j}''$ respectively.

Eq 3.3.54 and Eq 3.3.55 show that g_i' is of order $\frac{1}{\mu(n)}$, and $g_{i,j}''$ is of order $\frac{1}{\mu(n)^2}$. And g_i' and $g_{i,j}''$ are coefficients of $Z_i - \mu_i \equiv \Delta Z_i$ and $(Z_i - \mu_i)(Z_j - \mu_j) \equiv \Delta Z_i \Delta Z_j$ respectively. Therefore, as long as the system has a large number ($\mu(n)$) of expected neutrons, the expansion in Eq 3.3.53 is justified.

Eq 3.3.53 can be conceptually written as Eq 3.3.57:

$$g(Z) = g(\mu) + \frac{\Delta Z}{\mu} + \left(\frac{\Delta Z}{\mu}\right)^2 + o\left(\epsilon^2\right) \tag{3.3.57}$$

where

$$\epsilon \equiv \frac{\Delta Z}{\mu} \tag{3.3.58}$$

It is worthwhile to note that if a function of $\vec{Z}(n)$ is defined as

$$h\left(\vec{Z}(n)\right) = \frac{\sum_{i=1}^{M} Z_i(n)}{\sum_{i=1}^{M} Z_i(n-1)} \tag{3.3.59}$$

the MBP methodology can also be applied to analyze correlation and variance underestimate of k_{eff} eigenvalue tallies.

3.3.2.2 Variance terms

From the expansion of fission source distribution

$$g(\vec{Z}) = g(\vec{\mu}) + g'_i(Z^i - \mu^i) + \frac{1}{2}g''_{i,j}(Z^i - \mu^i)(Z^j - \mu^j) + o\left(\epsilon^2\right) \tag{3.3.60}$$

its expectation is given by

$$\mathbb{E}g(\vec{Z}) = g(\vec{\mu}) + \frac{1}{2}g''_{i,j}Cov[Z^i, Z^j] + o\left(\mathbb{E}\epsilon^2\right) \tag{3.3.61}$$

From Eq 3.3.60 and Eq 3.3.61, an expression for $\mathrm{Var}[g(\vec{Z})]$ can be obtained

$$\begin{aligned}\mathrm{Var}[g(\vec{Z})] &= \mathbb{E}\left(g'_i(Z^i - \mu^i) + \frac{1}{2}g''_{i,j}(Z^i - \mu^i)(Z^j - \mu^j) - \frac{1}{2}g''_{i,j}Cov[Z^i, Z^j] + o(\epsilon^2)\right)^2 \\ &= g'_i g'_{i'}\mathbb{E}(Z^i - \mu^i)(Z^{i'} - \mu^{i'}) \\ &+ g'_{i'}g''_{i,j}\mathbb{E}(Z^{i'} - \mu^{i'})(Z^i - \mu^i)(Z^j - \mu^j) + o(\mathbb{E}\epsilon^3).\end{aligned} \tag{3.3.62}$$

Since the lowest order of $g(\vec{Z}) - \mathbb{E}g(\vec{Z})$ is $\mathcal{O}(\epsilon)$, a second order expansion of $g(\vec{Z})$ is sufficient to provide the third order expansion of $\mathrm{Var}[g(\vec{Z})]$. Otherwise, third order expansion is required to comprise all third order terms in $\mathrm{Var}[g(\vec{Z})]$ with the leading constant term in $g(\vec{Z})$. The third order terms of the form $g'_i g''_{i',j'}(Z^i - \mu^i)Cov[Z^{i'}, Z^{j'}]$

vanish because $\mathbb{E}Z^i = \mu^i$.

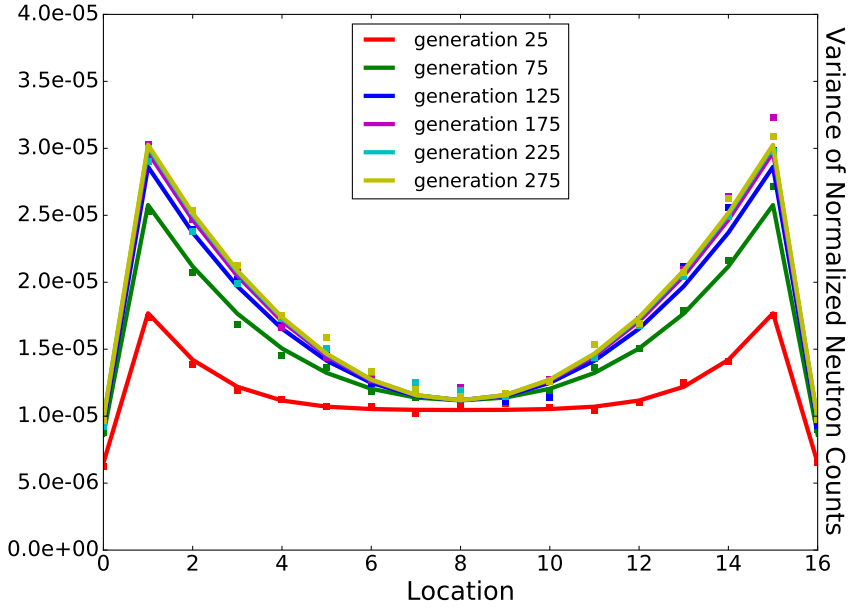
Inserting the expansion coefficients (Eq 3.3.54, Eq 3.3.55) and spatial moments, (Eq 3.3.23) and restoring the neutron type and generation dependence, the variance can be explicitly expressed as

$$\begin{aligned}\mathrm{Var}[g_l(\vec{Z}(n))] = &\left(\frac{\delta_l^i}{\mu(n)} - \frac{\mu_l(n)}{\mu(n)^2}\right)(C_{i,j}(n) - \mu_i(n)\mu_j(n))\left(\frac{\delta_l^j}{\mu(n)} - \frac{\mu_l(n)}{\mu(n)^2}\right)\\ &+ \left(\frac{\delta_l^h}{\mu(n)} - \frac{\mu_l(n)}{\mu(n)^2}\right)\left(-\frac{\delta_l^i}{\mu(n)^2} - \frac{\delta_l^j}{\mu(n)^2} + 2\frac{\mu_l(n)}{\mu(n)^3}\right) \times\\ &\{T_{h,i,j}(n) - \mu_i(n)C_{j,h}(n) - \mu_j(n)C_{h,i}(n) - \mu_h(n)C_{i,j}(n) + 2\mu_i(n)\mu_j(n)\mu_h(n)\} + o\left(\mathbb{E}\epsilon^2\right)\end{aligned} \tag{3.3.63}$$

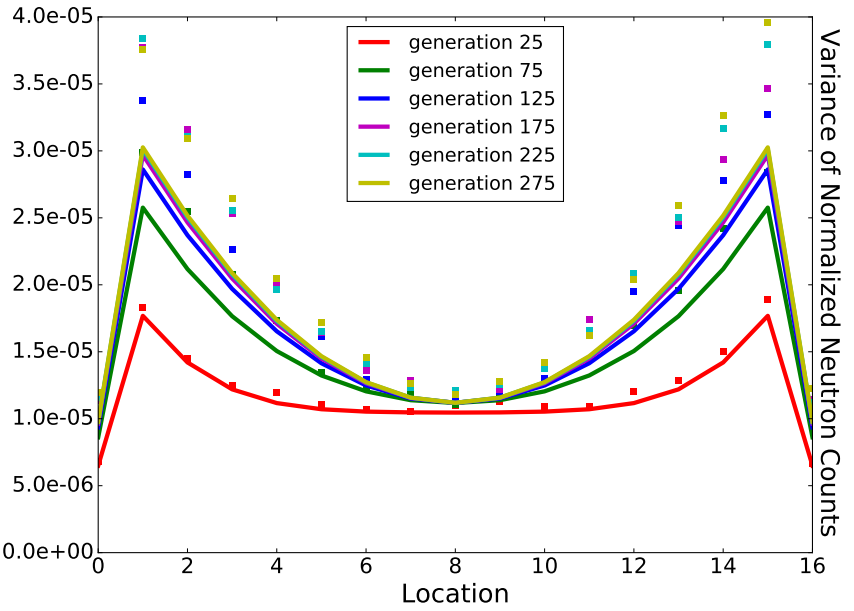
With the formula for $\mu_i(n)$ (Eq 3.3.32) and $C_{i,j}(n)$ (Eq 3.3.39), approximations of $\mathrm{Var}[X_i(n)]$ in Eq 3.3.63 with the leading term ($\mathcal{O}\left(\mathbb{E}\epsilon^2\right)$) can be calculated. For the test problem $TD1_f$, predicted $\mathrm{Var}[X_I(n)]$ are plotted for all $M = 17$ regions at generations $n \in \{25, 75, 125, 175, 225, 275\}$ in Figure 3-7. On the left figure, predicted $\mathrm{Var}[X_I(n)]$ is plotted along with the reference which indicates good agreement. The predicted values are computed as the expectation of the references truncated at $o\left(\mathbb{E}\epsilon^2\right)$ as shown in Eq 3.3.63. The right figure shows the predicted $\mathrm{Var}[X_I(n)]$ and the $TD1_n$ reference. The predicted values are not mathematically the expectation of the $TD1_n$ reference since the MBP model does not include the neutron number normalization present in $TD1_n$. However, $X_i(n)$ in both the MBP model and $TD1_n$ represent the empirical neutron distribution of the $TD1$ problem and they are expected to have similar mean and variance.

3.3.2.3 Covariance terms

The expansion of $g(\vec{Z}(n))$(Eq 3.3.60) and $\mathbb{E}g(\vec{Z}(n))$ (Eq 3.3.61) can also be used to evaluate the covariance terms. Since the dependence on generation is suppressed, g_1 and g_2 are used to denote $g(\vec{Z}(n))$ and $g(\vec{Z}(n+k))$. μ_1, μ_2, Z_1 and Z_2 are introduced similarly.



(a) Var $X_i(n)$ and reference from $TD1_f$



(b) Var $X_i(n)$ and reference from $TD1_n$

Figure 3-7: Variances of normalized tally for all regions of problem $TD1$ The solid curves correspond to $\mathrm{Var}[X_i(n)]$ *predicted by Eq 3.3.63 up to order* $\mathcal{O}\left(\mathbb{E}\epsilon^2\right)$*. The dots in Figure 3-7(a) correspond to references calculated from simulations without neutron number normalization. The dots in Figure 3-7(b) correspond to references calculated from simulations with neutron number normalization.*

$$
\begin{aligned}
&Cov[g(\vec{Z}(n), g(\vec{Z}(n+k))] \\
&= \mathbb{E}\left[\left(g'_{1i}(Z_1^i - \mu_1^i) + \frac{1}{2}g''_{1\,i,j}(Z_1^i - \mu_1^i)(Z_1^j - \mu_1^j) - \frac{1}{2}g''_{i,j}Cov[Z_1^i, Z_1^j] + o(\epsilon^2)\right) \times \right. \\
&\quad \left.\left(g'_{2i'}(Z_2^{i'} - \mu_2^{i'}) + \frac{1}{2}g''_{2\,i',j'}(Z_2^{i'} - \mu_2^{i'})(Z_2^{j'} - \mu_2^{j'}) - \frac{1}{2}g''_{i',j'}Cov[Z_2^{i'}, Z_2^{j'}] + o(\epsilon^2)\right)\right] \\
&= g'_{1i}g'_{2i'}\mathbb{E}(Z_1^i - \mu_1^i)(Z_2^{i'} - \mu_2^{i'}) \\
&+ \frac{1}{2}g'_{1i}g''_{2\,i',j'}\mathbb{E}(Z_1^i - \mu_1^i)(Z_2^{i'} - \mu_2^{i'})(Z_2^{j'} - \mu_2^{j'}) \\
&+ \frac{1}{2}g'_{2i'}g''_{1\,i,j}\mathbb{E}(Z_2^{i'} - \mu_2^{i'})(Z_1^i - \mu_1^i)(Z_1^j - \mu_1^j) + o(\mathbb{E}\epsilon^3)
\end{aligned}
\tag{3.3.64}
$$

Inserting the expansion coefficients (Eq 3.3.54, Eq 3.3.55) and reintroducing the neutron type and generation dependence, the covariance can be explicitly written as

$$
\begin{aligned}
&Cov[g_{l1}(\vec{Z}(n)), g_{l2}(\vec{Z}(n+k))] \\
&= \left(\frac{\delta_{l1}^i}{\mu(n)} - \frac{\mu_{l1}(n)}{\mu(n)^2}\right)(\mathbb{E}Z_i(n)Z_j(n+k) - \mu_i(n)\mu_j(n+k))\left(\frac{\delta_{l2}^j}{\mu(n+k)} - \frac{\mu_{l2}(n+k)}{\mu(n+k)^2}\right) \\
&+ \frac{1}{2}\left(\frac{\delta_{l1}^h}{\mu(n)} - \frac{\mu_{l1}(n)}{\mu(n)^2}\right)\left(-\frac{\delta_{l2}^i}{\mu(n+k)^2} - \frac{\delta_{l2}^j}{\mu(n+k)^2} + 2\frac{\mu_{l2}(n+k)}{\mu(n+k)^3}\right) \times \\
&\quad \mathbb{E}(Z_h(n) - \mu_h(n))(Z_i(n+k) - \mu_i(n+k))(Z_j(n+k) - \mu_j(n+k)) \\
&+ \frac{1}{2}\left(\frac{\delta_{l_2}^h}{\mu(n+k)} - \frac{\mu_{l2}(n+k)}{\mu(n+k)^2}\right)\left(-\frac{\delta_{l1}^i}{\mu(n)^2} - \frac{\delta_{l1}^j}{\mu(n)^2} + 2\frac{\mu_{l1}(n)}{\mu(n)^3}\right) \times \\
&\quad \mathbb{E}(Z_h(n+k) - \mu_h(n+k))(Z_i(n) - \mu_i(n))(Z_j(n) - \mu_j(n)) + o(\mathbb{E}\epsilon^3)
\end{aligned}
\tag{3.3.65}
$$

Eq 3.3.65 indicates the moments that need be evaluated. The first term is of lowest expansion order ($\mathcal{O}(\epsilon^2)$). It requires the value of serial-spatial moments $\mathbb{E}Z_i(n)Z_j(n+k)$. The second and third terms are of the higher expansion order ($\mathcal{O}(\epsilon^3)$). They require the value of serial-spatial moments $\mathbb{E}Z_h(n)Z_i(n+k)Z_j(n+k)$ and $\mathbb{E}Z_h(n+k)Z_i(n)Z_j(n)$.

So far, the explicitly written moments are $\mathbb{E}Z_i(n)$ (Eq 3.3.32), $\mathbb{E}[Z_i(n)Z_j(n)]$

(Eq 3.3.39) and $\mathbb{E}[Z_i(n)Z_j(n+k)]$ (Eq 3.3.51), which enable evaluating correlation coefficients in Eq 3.3.65 up to order $\mathcal{O}(\mathbb{E}\epsilon^2)$.

The predicted correlation coefficients for X tallies in two spatial locations are plotted in Figure 3-8 along with the reference values calculated from independent simulations. Both references from $TD1_f$ and $TD1_n$ are plotted since similar correlation behaviors were observed.

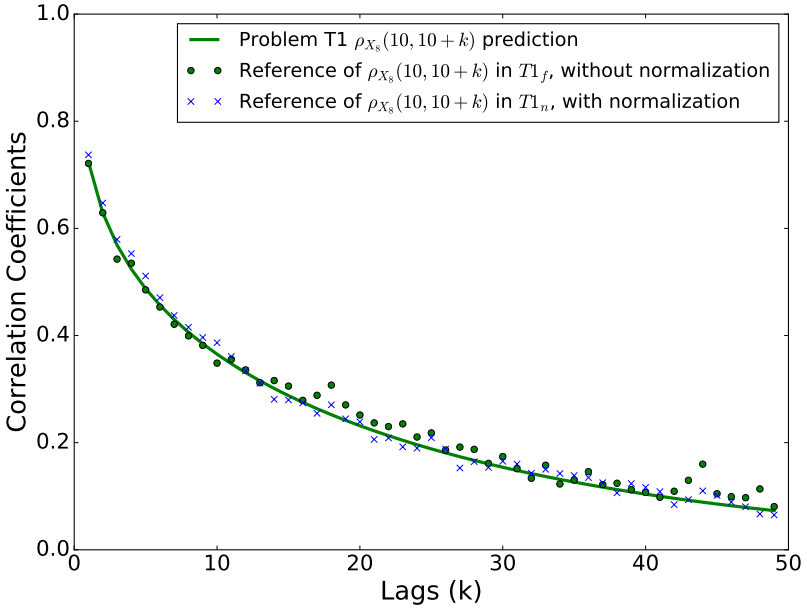
The central region $I = 8$ was selected because the reference values have been used for illustration in previous figures. Region $I = 1$ was chosen because of the large discrepancy illustrated in Figure 3-7(b). Despite the relatively poor variance prediction of $\mathrm{Var}[X_i(n)]$, Figure 3-8(b) shows acceptable accuracy in the correlation prediction. Since region $I = 1$ is closer to the reflective boundary, there is a higher probability that neutrons observed in one generation re-enters the same region in future generations. This contributes to the higher correlation coefficients in region $I = 1$ when comparing to the central region.

The predicted correlation coefficients match the reference from $TD1_f$ because the prediction is performed on the same physical problem, which merely implies that the covariance expansions derived above are correct. The agreement between the prediction curves and reference from $TD1_n$ is more useful and suggests the capability of the MBP model in predicting correlation coefficients and thus correcting variance underestimation without the need for many independent simulations.

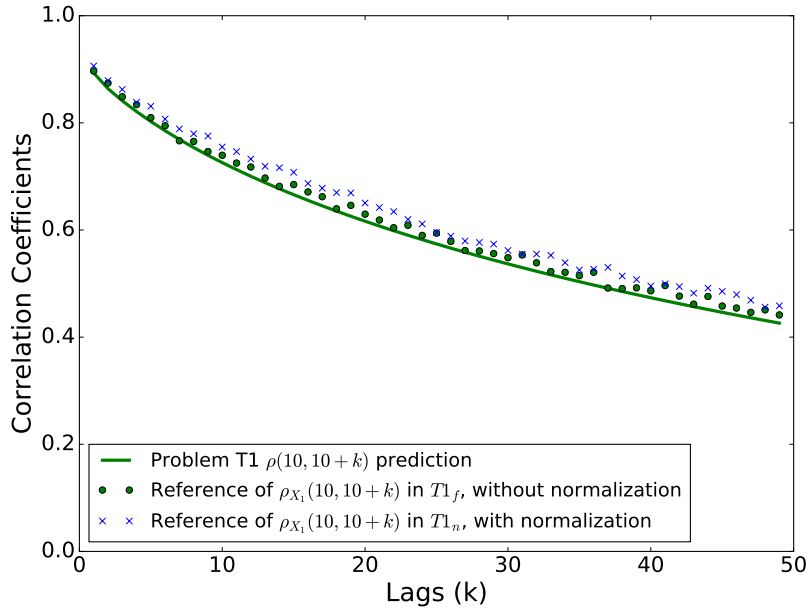
The above results show that second order moments are sufficient to predict accurate correlation coefficients. In situations where ϵ (Eq 3.3.58) is large, for example, predicting correlation in a region that occupies a large fraction of the whole system, higher order expansions might be needed. $\mathcal{O}(\mathbb{E}\epsilon^3)$ terms in Eq 3.3.63 and Eq 3.3.65 require the third order moments.

Following a similar but more lengthy derivation (section A1.3.2) to the second order derivations, the third order spatial moments are given as

$$\begin{aligned} T_{k,j,i}(n+1) = {} & T_{g,h,l}(n)M_k^g M_j^h M_i^l + V_{j,i}^l M_k^\alpha C_{l,\alpha}(n) \\ & + V_{k,i}^l M_j^\alpha C_{l,\alpha}(n) + V_{j,k}^l M_i^\alpha C_{l,\alpha}(n) + W_{i,j,k}^l \mu_l(n) \end{aligned} \tag{3.3.66}$$



(a)



(b)

Figure 3-8: Correlation coefficients of X tallies in problem TD1. The green solid curves correspond to predicted correlation coefficients of $\rho_{X,I}(k)$ for $I = 8$ (Figure 3-8(a)) and $I = 1$ (Figure 3-8(b)). The green dots correspond to reference correlation coefficients of X_8 (Figure 3-8(a)) and X_1 from simulations on $TD1_f$ (Figure 3-8(b)). The blue crosses correspond to reference correlation coefficients of X_8 (Figure 3-8(a)) and X_1 from simulations on $TD1_n$ (Figure 3-8(b)).

where $W^l_{i,j,k}$ is the third order (central) spatial moment response to a neutron source of type l:

$$W^l_{i,j,k} = \mathbb{E}\left[(Z_i(1)-\mu_i(1))(Z_j(1)-\mu_j(1))(Z_k(1)-\mu_k(1))|\,\vec{r}_0 = \vec{e}_l\right]. \tag{3.3.67}$$

And the third order serial-spatial moments are found as

$$\mathbb{E}Z_i(n)Z_j(n)Z_l(n+k) = \mathbb{T}_{h,i,j}\mu^h_l(k) + \mathbb{C}_{i,j}(n)(\mu^j_l(k)+\mu^i_l(k)) \tag{3.3.68}$$

$$\begin{aligned}\mathbb{E}Z_i(n+k)Z_j(n+k)Z_l(n) = {}& \mathbb{T}_{ghl}(n)\mu^g_i(k)\mu^h_j(k) + \mathbb{C}_{l,h}(n)\mathbb{C}^h_{i,j}(k) \\ & + \mathbb{C}_{l,h}(n)(\mu^h_j(k)\mu^l_i(k)+\mu^h_i(k)\mu^l_j(k)) + \mathbb{C}^l_{i,j}(k)\mu_l(n).\end{aligned} \tag{3.3.69}$$

Similarly, $\mathbb{C}^h_{i,j}(n)$ is defined like $\mathbb{C}_{i,j}(n)$ with the exception that $\mathbb{C}^h_{i,j}(n)$ is a conditional expectation on $\vec{r}_0 = \vec{e}_h$. It should be noted that there is an exception to Einstein notation in the above equation. The left hand side depends on index l and l is not dummy (summed over) on the right hand side.

3.3.2.4 Auto-Correlation $\rho_{n,k}$ of coarse tally regions

In the above derivation and verification process, the tally regions are the same as the discrete phase space regions. In realistic applications, one tally region could contain more than one phase space region. The relation between the (co)variance terms needed to evaluate correlation coefficients for a tally region I and the comprising phase space regions $I_i \in I$ was discussed in section 3.1.1 Eq 3.1.3. All the terms on the right hand side of Eq 3.1.3 have already been explicitly written in Eq 3.3.65. Though Eq 3.3.65 only evaluates moments in the form of $Cov[Z_l(n), Z_l(n+k)]$ and Eq 3.1.3 requires moments in the form of $Cov[Z_i(n), Z_j(n+k)]$, it is straightforward to substitute subscript l of terms at generation n with I_i and substitute subscript l at generation $n+1$ with I_j.

Moreover, the summation in Eq 3.1.3 and the kronecker deltas in Eq 3.3.65 can

simplify the covariance of a tally in region I

$$
\begin{aligned}
&Cov[X_I(n), X_I(n+k)] \\
&= \frac{\mu_I(n)}{\mu(n)}\mathbb{E}\left[\left(\frac{Z(n)}{\mu(n)} - \frac{Z_I(n)}{\mu_I(n)}\right)\left(\frac{Z(n+k)}{\mu(n+k)} - \frac{Z_I(n+k)}{\mu_I(n+k)}\right)\right]\frac{\mu_I(n+k)}{\mu(n+k)} \\
&- \frac{\mu_I(n)}{\mu(n)}\mathbb{E}\left[\left(\frac{Z(n+k)}{\mu(n+k)} - \frac{Z_I(n+k)}{\mu_I(n+k)}\right)\left(\frac{Z(n)}{\mu(n)} + \frac{Z(n+k)}{\mu(n+k)} - 2\right)\left(\frac{Z(n)}{\mu(n)} - \frac{Z_I(n)}{\mu_I(n)}\right)\right]\frac{\mu_I(n+k)}{\mu(n+k)} \\
&+ o(\mathbb{E}\epsilon^3)
\end{aligned}
\tag{3.3.70}
$$

the derivation is shown in Eq A1.3.26 and Eq A1.3.31 (see Appendix A1.3.3). Further simplification to Eq 3.3.70 was not sought in order to keep the $\mathcal{O}(\mathbb{E}\epsilon^2)$ and the $\mathcal{O}(\mathbb{E}\epsilon^3)$ terms separate.

Eq 3.3.70 also shows that the correlation of normalized tallies (X tallies) within one specific region I depend on spatial correlations over the whole space instead of spatial correlations within the region I only. The complexity that all behavior of an individual neutron depends on all other neutrons due to the neutron number normalization discussed before has been treated by the complexity of tallies.

3.3.3 Procedure to Calculate $\rho_{n,k}$

In summary, knowing the spatial moments, serial-spatial moments, evolution equations and expansions of $X_l(n)$, generation-to-generation correlation coefficients, $\rho_{n,k}(X_l)$, can be calculated following these steps:

1. Calculate the spatial moment responses M_i^l, $V_{i,j}^l$ and $W_{i,j,k}^l$ given in Eq 3.3.29,Eq 3.3.41 and Eq 3.3.67

2. Evolve the spatial moments $\mu_i(n)$, $C_{i,j}(n)$ and $T_{i,j,k}(n)$ according to Eq 3.3.32,Eq 3.3.39 and Eq 3.3.66

3. Combine the spatial moments of different generations to evaluate the serial-spatial moments according to Eq 3.3.51,Eq 3.3.68 and Eq 3.3.69

4. Condense the covariances of discretized phase space cells into tally regions of interest with Eq 3.1.3

5. Substitute the moments into the expansion of $Var[Z_I(n)]$ and $Cov[Z_I(n), Z_I(n+k)]$ to evaluate $\rho_{n,k}(Z_I)$

The above procedure yields $\rho_{n,k}$ after evolving the MBP for n generations. It should be noted that the above procedure should be performed prior to performing a fully converged simulation. The moment responses M_j^l, $V_{i,j}^l$ with all phase space cells being finely discretized into flat source regions can be tallied from a fixed source calculation of uniformly sampled neutron sources. Since such a fine mesh can be costly, the moment responses should be tallied from a stationary source on a mesh slightly finer than the tallied quantities.

Generation-to-generation correlation of a tally region is captured by the generation-to-generation and spatial correlation among the comprising phase space cells. Prediction should be accurate if the migration distance of neutrons between generations is on the same order or larger than the MBP phase space cells for a given tally region.

Following the above procedure, evolving the constructed MBP model for k generations gives the ACC of generation lag k and can thus provide a correction for the variance estimator after k active generations. In order to predict termination criteria for tallies, very long MBP calculations are required to evaluate ACC's up to an arbitrary generation lag k. Alternatively, the method of fitting ACC to a sum of exponentially decaying terms can be applied on the ACC's found from the MBP model instead of the noisy estimated value as in [35]. With data fitting, the MBP model can be evolved for a fixed number of generations and the fit can be used to predict ACC and variance underestimation to any arbitrary number of generations.

Results from MBP also verify the data fitting method where ACC's are fitted to a sum of exponentially decaying terms. If ACC's predicted from MBP are written explicitly, a quadratic form of the response matrices appears, whose dominant eigenvalues are equivalent to the exponential terms used in data fitting.

3.3.4 Application

3.3.4.1 Evaluate spatial moment responses

With the procedure given in Sec 3.3.3, the spatial moment responses are needed to initialize the calculation of ACC's. Based on their definition, spatial moment responses are just spatial moments except that they specify the starting phase space region and are one generation away from the source. Therefore, these spatial moment responses can be tallied from a single generation of a Monte Carlo simulation. Due to computational cost, it is unreasonable to tally directly higher order spatial moments. For example, the moment response $V_{i,j}^l$ of second order spatial moments is a third order tensor that depends on three phase space regions and would require many neutrons per generation to accurately resolve.

We propose instead to tally the fission-to-absorption probability for each region pair and moments of new neutrons (from fission) for each region and use these to construct the spatial moment responses. The feasibility of constructing the spatial moment responses is limited to cases where only spatial (excluding energy, angle) correlation effects are of concern.

Assuming that the probability of absorption and the number of neutrons born from fission are independent, we can denote the probability that a neutron born at phase space region l is absorbed at phase space region i as P_i^l. Defining the random variable ν_i as the number of new neutrons out of absorption at phase space region i, then by definition of M_i^l:

$$M_i^l = P_i^l \mathbb{E}\nu_i \tag{3.3.71}$$

As for the second order spatial moment responses, Eq 3.3.41 can be rearranged as

$$V_{i,j}^l = \mathbb{E}(Z_i(1)Z_j(1)|_{\vec{r_0}=\vec{e}_l}) - M_i^l M_j^l \tag{3.3.72}$$

Since energy and angle are not discretized, new neutrons out of absorption cannot appear at different phase space regions,

$$Z_i(1)Z_j(1) = \nu_i^2 \delta_{i,j} \tag{3.3.73}$$

Therefore the second order spatial moment responses $V_{i,j}^l$ can be constructed from fission-to-absorption probabilities and moments of new neutrons from fission as

$$V_{i,j}^l = P_i^l \mathbb{E}\nu_i^2 \delta_{i,j} - P_i^l P_j^l \mathbb{E}\nu_i \mathbb{E}\nu_j \tag{3.3.74}$$

Although $V_{i,j}^l$ is a third order tensor, it can be constructed from a matrix P_i^l and a vector $\mathbb{E}\nu_i$. The matrix P_i^l and vectors $\mathbb{E}\nu_i$ and $\mathbb{E}\nu_i^2$ can be either computed from the unconverged or converged fission source. If unconverged fission source is used, a fine mesh is needed such that the flat source approximation yields a suitable approximation of the Monte Carlo eigenvalue problem. Under this circumstance, the distribution of fission source does not alter P_i^l, $\mathbb{E}\nu_i$ and $\mathbb{E}\nu_i^2$. If starting from a converged fission source, P_i^l, $\mathbb{E}\nu_i$ and $\mathbb{E}\nu_i^2$ can be tallied directly on a mesh as coarse as the tally regions of interest and used to construct an MBP model.

The above construction of spatial moment responses M_i^l and $V_{i,j}^l$ are based on the reasonable assumption that the occurrence of branching defines the concept of generation. However, in Monte Carlo simulations, other branching processes like (n, xn) are treated within a generation [39].

To describe such processes exactly, a more detailed structure of transfer is required which can be shown to require 3 additional matrices and 2 additional vectors. The detailed derivation is given in A1.3.4, but results were shown unnecessary for capturing the bulk correlation behavior in nuclear systems and were thus omitted.

3.3.4.2 Approximating source normalization

As discussed in section 3.1.2 and through the development of the correlation prediction capability of the MBP model, the neutron number normalization step cannot be incorporated into the MBP model easily but the effect can be compensated by focusing on the normalized tallies.

In OpenMC [39], the number of new neutrons for each fission event is normal-

ized by the estimated k_{eff} to maintain the neutron population near constant with subsequent normalization steps to enforce a constant fission bank. For a non-critical problem, the neutron number normalization step is equivalently forcing criticality. This effect can be captured by making the constructed MBP critical. With the tallied fission-to-absorption matrix P and the expected number of new neutrons for each cell $\mathbb{E}\nu$, a fission matrix can be constructed according to Eq 3.3.71 and then scaled by the estimated eigenvalue of the fission matrix to force the largest eigenvalue to be 1.

3.3.5 Results and Analysis

This section applies the MBP model to predict correlation on a realistic problem, the 2D BEAVRS benchmark [23]. Three quantities can be solved for the MBP model and thus used as a prediction for the real problem. The quantities are correlation coefficients $\rho(k)$ (Eq 3.0.1), the variance underestimation ratio $r(N)$ (Eq 2.2.8) and the bias of sample variance $b(N)$ (Eq 2.2.32).

The predictive capability of the MBP model was compared on assembly size tallies using the 2D BEAVRS benchmark at Hot Zero Power conditions. The 2D model is represented by a 10 cm axial slice at mid-height with reflective top and bottom boundary conditions. To serve as a reference for the variance and ACCs, 450 independent simulations were performed. Each simulation contained 250 inactive generations and 800 active generations. Each generation tracked 4×10^6 neutrons. These independent simulations provide a true measure of variance at each generation.

To limit the amount of tally data while still exploring the spatial dependency of the predicted quantities, three representative assemblies were selected as shown in Figure 3-9. Among the three fuel assemblies selected, one is at the edge (Assembly 1), one is at the center (Assembly 3) and the last one is in the middle (Assembly 2). $U235$ enrichment of Assembly 1, 2, 3 are 3.1%, 1.6% and 1.6% respectively.

As can be seen from Figure 3-9, the fuel assemblies are in a 15×15 mesh. For example, Assembly 1 is at row 1, column 5, Assembly 2 is at row 5 column 5 and Assembly 3 is at row 8 column 8. The MBP model was constructed over a region on

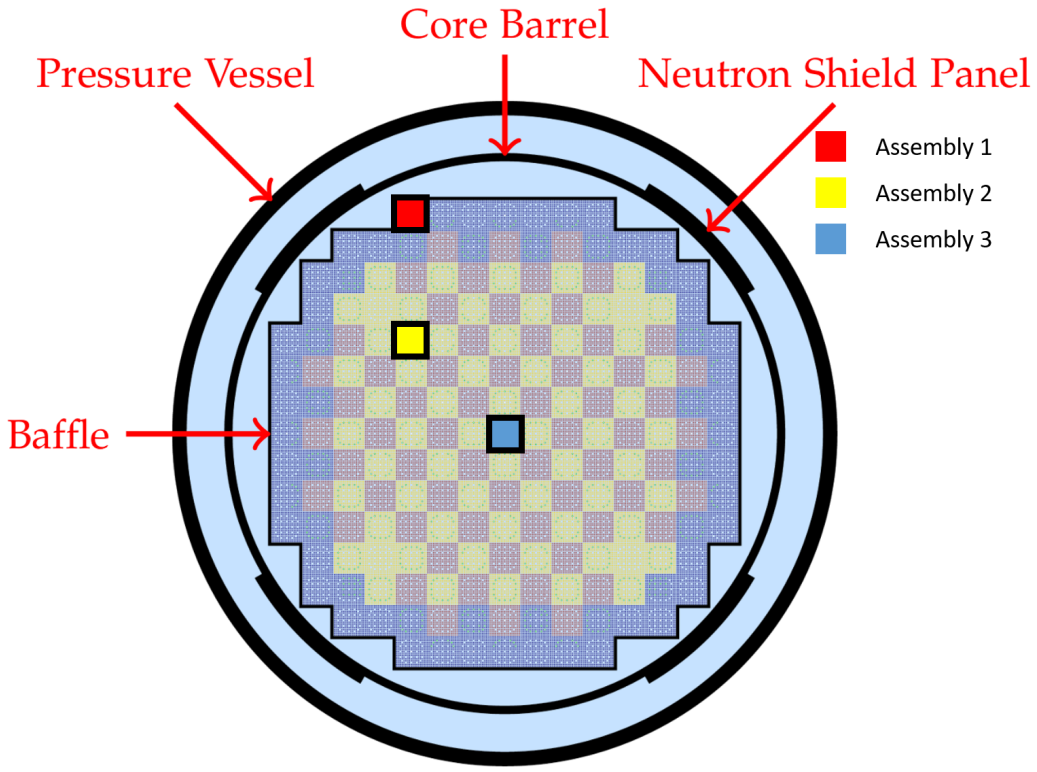


Figure 3-9: Layout of BEAVRS Benchmark and selected assemblies

17×17 assembly-sized meshes that extends one row into the reflector region. Two discretizations were used, one that is assembly size (17×17) and one quarter assembly size 34×34.

The MBP model's transfer matrix P_i^l, the moments of number of new neutrons per absorption $\mathbb{E}\nu_i$, $\mathbb{E}\nu_i^2$ on the above discretization meshes were computed using the first 8 active generations.

3.3.5.1 Comparing to reference

3.3.5.1.1 *ACC* **Prediction** The predicted and reference *ACC*'s for the three assemblies as a function of the generation lag are shown in Figure 3-10. The exponentials fitted to the predicted *ACC*'s are also given. In Figure 3-10, the black solid curves correspond to predictions from the MBP model. The blue diamonds are the reference *ACC*'s calculated from the 450 independent simulations. The green dashed curves correspond to exponential fittings of correlation coefficients to those predicted by the MBP model. The fitted *ACC* curves capture the behavior of predicted values.

The left figures correspond to the quarter-assembly size mesh, while the right figures are for the assembly mesh. In the assembly mesh case, Eq 3.3.63 and Eq 3.3.65 are used to predict the correlation coefficients directly on the selected assemblies. In the quarter-assembly mesh case, the same equations are used in addition with the condensation formula, Eq 3.1.3.

From Figure 3-10, predictions in the right column (assembly mesh MBP model) underestimate correlation coefficients compared to the prediction in left column (quarter-assembly mesh MBP model). This can be intuitively explained by the observation in Eq 3.3.51 and Eq 3.1.3. Correlation coefficient of region I tally between generation n and $n+k$ are due to the spatial correlation of between region I tally and the whole system. The spatial correlation can be characterized by the distance a neutron travels in one generation. If the discrete MBP model were built such that the discretized phase space regions are separated by a distance smaller or equal to the characteristic length, the correlation from the tally region of interest to the whole system can be captured. Otherwise, if the discrete MBP model were built on a coarser mesh where a neutron stays in the same discrete phase space region, the spatial correlation would be truncated by the local tally region and thus underestimate the correlation coefficients. In a typical water moderated reactor, the average distance a neutron travels is around $5cm \sim 6cm$ estimated from the migration area [29]. The assembly mesh in a PWR is $21.5cm \times 21.5cm$ while the quarter-assembly mesh is $10.25cm \times 10.25cm$. The results in Figure 3-10 show empirically that the MBP model with quarter-assembly mesh predicts correlation coefficients with acceptable accuracy.

The correlation level decreases from the boundary assembly to the center assembly. This is consistent with the observation of the test problem $TD1$ in section 3.3.2.1, Figure 3-8 because water near the boundary reflects (scatters) neutrons.

The small deviation between the predicted and reference values in Figure 3-10 also implies that the estimation of the moments M_j^l and $V_{i,j}^l$ from the first 8 active generations is accurate enough to capture the spatial correlations in the system. The larger deviations observed with the fitted curves is caused by data fitting instabilities [1]. Fitting to a sum of exponentials is an ill-posed problem where small changes in

data can lead to large discrepancies in the fit.

In the following comparisons, only the predictions from the quarter-assembly MBP model will be used.

3.3.5.1.2 Variance convergence rate prediction

The ACC's can then be used to correct the variance correction rate ($r(N)/N$) from underestimation ratio ($r(N)$). The real convergence rates from independent simulations according to Eq 2.2.45 are shown in Figure 3-11 with the convergence rates predicted from the MBP model and from the fitted ACC of the MBP prediction. As mentioned in Section 2.2.3, $\widehat{Var[\overline{X}_I(N)]}$ is renormalized to give $r(N)/N$ (Eq 2.2.46). In the peripheral assembly with the highest correlation level among the three, real variance convergence rate deviates from $1/N$ by a factor of 7 ($r(\infty) \sim 7$). In the central assembly with lowest correlation level, this factor is around 4 ($r(\infty) \sim 4$).

3.3.5.1.3 Variance bias prediction

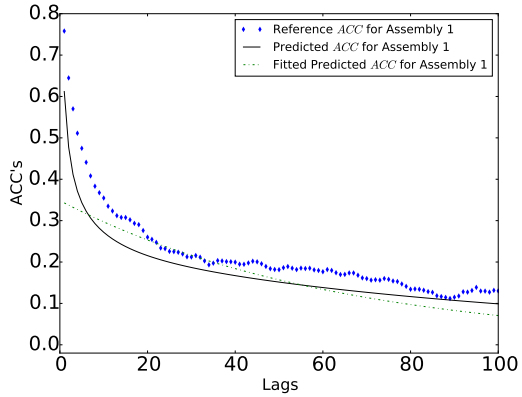
The impact of the correlation coefficients on the bias of the estimator of $Var[X_I(N)]$ is shown in Fig 3-12. In the selected assembly tally regions, the variance for each generation is underestimated by a factor of ~ 2 ($b(10) \sim 0.5$) if 10 active generations are simulated for statistics. In comparison with the deviation from $1/N$, which remains constant asymptotically, the bias of leading variance becomes negligible as more active generations are simulated.

Figure 3-13 plots the RMS (Eq 2.2.14) calculated from correcting variance by incorporating both bias ratio ($b(N)$) and underestimation of $1/N$ ($r(N)$). In addition to the reference values and RMS corrected by predicted correlations, we also show what the RMS would look like if the traditional variance estimator was used.

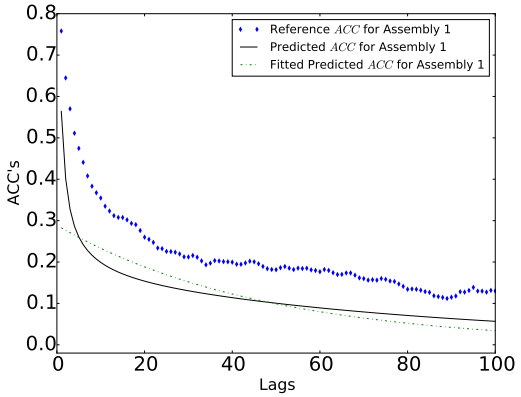
3.3.5.2 Cost of the prediction method

Theoretical predictions above are based on a MBP model on a quarter assembly mesh (resulting from discretizing the BEAVRS 2D benchmark into $M = 34 \times 34 = 1156$ cells) and the second order expansion of fission source distributions (Sec 3.3.2.1).

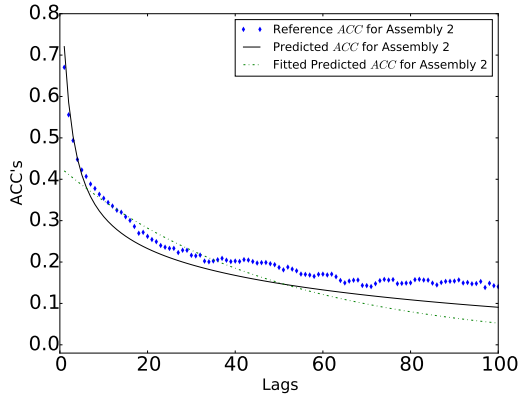
To perform the second order calculation, moment responses M_j^l and $V_{i,j}^l$ and mo-



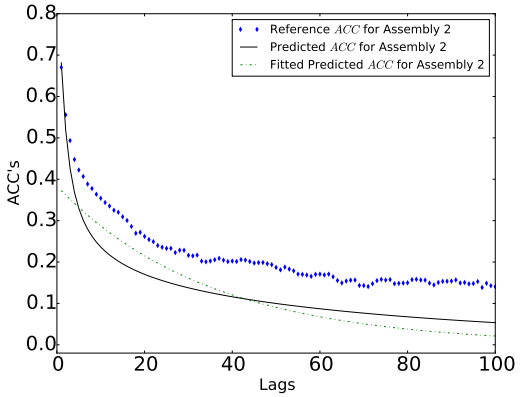
(a) Prediction on Assembly 1 from quarter-assembly mesh



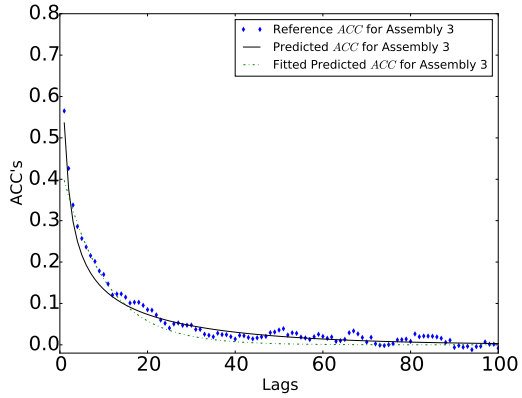
(b) Prediction on Assembly 1 from assembly mesh



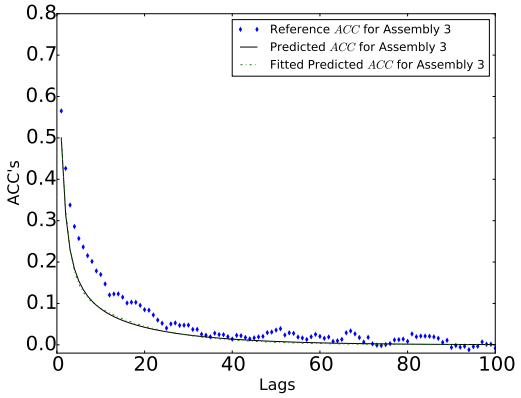
(c) Prediction on Assembly 2 from quarter-assembly mesh



(d) Prediction on Assembly 2 from assembly mesh

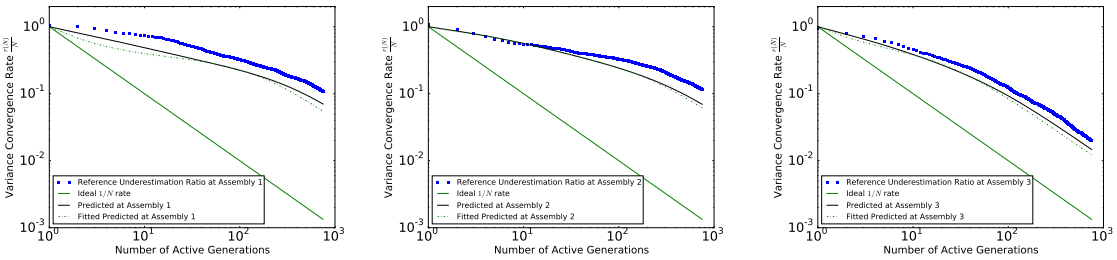


(e) Prediction on Assembly 3 from quarter-assembly mesh



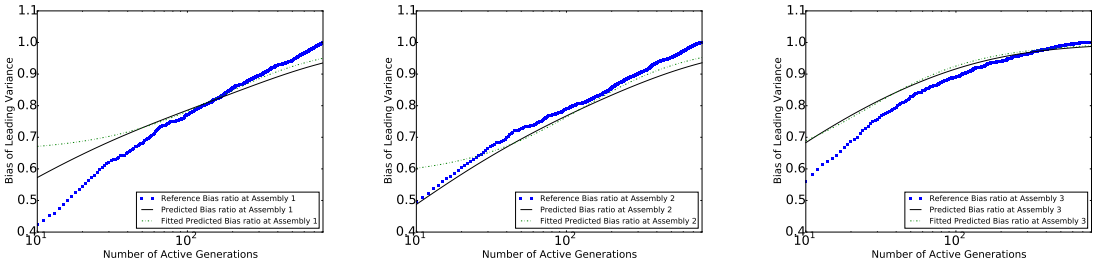
(f) Prediction on Assembly 3 from assembly mesh

Figure 3-10: Predicted and Reference ACC for the three selected assemblies. The blue diamonds are references calculated from 450 *independent simulations according to Eq 2.2.41. The black solid curves correspond to the prediction out of the MBP model built from first* 8 *active generations. The green dashed curves fit the predicted ACC to three-term exponential* ($J = 3$ *in Eq 2.3.2).*



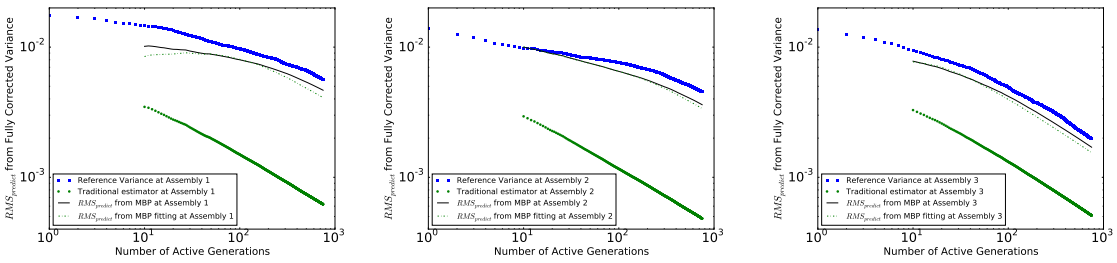
(a) Assembly 1 (b) Assembly 2 (c) Assembly 3

Figure 3-11: Predicted and Reference variance convergence rate $r(N)/N$ at three selected assemblies. The blue diamonds are references calculated from 450 *independent simulations according to Eq 2.2.46. The black solid curves correspond to the prediction out of the quarter-assembly MBP model built from first* 8 *active generations. The green dashed curves correspond to the calculation from fitting the predicted ACC $\rho_I(d)$ up to $d = 100$ to sum of exponentials.*



(a) Assembly 1 (b) Assembly 2 (c) Assembly 3

Figure 3-12: Predicted and Reference bias of variance at three selected assemblies. The blue diamonds are references calculated from 450 *independent simulations according to Eq 2.2.44. The black solid curves correspond to the prediction out of the quarter-assembly MBP model built from first* 8 *active generations. The green dashed curves correspond to the calculation from fitting the predicted ACC $\rho_I(d)$ upto $d = 100$ to sum of exponentials.*



(a) Assembly 1 (b) Assembly 2 (c) Assembly 3

Figure 3-13: Predicted RMS (calculated from variance fully corrected by $r(N)$ and $b(N)$) and Reference RMS at three selected assemblies. The blue diamonds are references calculated from 450 independent simulations according to Eq 2.2.14. The black solid curves correspond to the prediction out of MBP model built from first 8 active generations. The green dashed curves correspond to the calculation from fitting the predicted ACC $\rho_I(d)$ up to $d = 100$ to sum of exponentials. The green dots correspond to the RMS prediction if traditional variance estimator were used.

ments $\mu_l(n)$, $\mu_l^i(k)$ and $C_{i,j}(n)$ are needed. M_j^l and $V_{i,j}^l$ can be constructed from the fission to absorption matrix (P_i^l) and first ($\mathbb{E}\nu_i$) and second ($\mathbb{E}\nu_i^2$) order moments of ν_i, where ν_i is the number of neutrons born from absorption in phase space region i. The quantities P_i^l, $\mathbb{E}\nu_i$ and $\mathbb{E}\nu_i^2$ are tallied during the first few active generations. These correspond to one $M \times M$ matrix and two length M vectors, respectively. M_j^l and $V_{i,j}^l$ are used to find the stationary distribution of $\mu_l(n)$ and spatial covariance $C_{i,j}(n)$.

Then M_j^l is used to evolve the first order point response moment $\mu_j^h(k)$ with $\mu_j^h(0) = \delta_j^h$. According to Eq 3.1.3 and Eq 3.3.65, $\mu_l(n)$, $C_{i,j}(n)$ and $\mu_j^h(k)$ are sufficient to evaluate $\rho_{n,k}(X_I)$.

To provide a better variance estimate, $r(N)$ rather than $\rho(k)$ is required for each tally region. According to Eq 2.2.8, two accumulated summation of ACC's for each tally region should be maintained: $\sum_{k=1}^{N-1} \rho(k)$ and $\sum_{k=1}^{N-1} k\rho(k)$, thus needing to store two vectors of length M. In total, the minimal memory requirement of a second order MBP calculation is four M length vectors ($\mathbb{E}\nu_i$,$\mathbb{E}\nu_i^2$,$\sum_{k=1}^{N-1} k\rho_k(X_i)$,$\sum_{k=1}^{N-1} \rho_k(X_i)$)) and three $M \times M$ matrices (M_j^l,$C_{i,j}(n)$,$\mu_j^h(k)$). For the BEAVRS 2D benchmark discretized to $34 \times 34 = 1156$ cells with all matrix element represented as double precision numbers, the memory requirements are approximately $30MB$.

However, once the ACCs for each tally region have been calculated for enough MBP generations (~ 100) to perform exponential fitting, only $2 \sim 3$ coefficients per tally are needed to provide a better variance estimate.

Chapter 4

Correlation Diagnosis

The previous chapter explored ways of quantifying the correlation level in order to provide better variance estimates. However, in many problems, correlation levels are quite low thus not needing any correction to the typical variance estimator. This chapter aims to provide a simple indicator as to the potential impact of correlations on the problem at hand. To accomplish this, the derivation in Chapter 3 will be generalized to a continuous space and applied to an homogenized version of the problem.

Section 4.1 simplifies the derivation of the MBP model in section 3.3 by assuming criticality of the system and stationarity of the neutron source distribution. Section 4.2 generalizes the MBP model to a continuous phase space and homogeneous problems with reflective boundary conditions are used in section 4.3 and section 4.4 to verify this model with the analytic solutions. This will demonstrate that the asymptotic variance underestimation ratio calculated from the correlation coefficients can be used as a numerical indicator of the correlation in the system. Section 4.3 investigates a $1D$ homogeneous problem with explicitly defined response moments to calculate correlation coefficients while section 4.4 generalizes the exact results of the $1D$ problem to $3D$ using the diffusion approximation. Section 4.5 applies the diffusion approximation of the continuous MBP model to predict correlation coefficients and asymptotic variance underestimation ratio of a realistic problem, the BEAVRS benchmark.

4.1 Simplification of covariance formula

Starting from the recursive formula for the spatial moments $C_{j,i}(n)$ (Eq 3.3.39). and shifting the focus to central moments

$$Cov_{j,i}(n) = C_{j,i}(n) - \mu_j(n)\mu_i(n) \tag{4.1.1}$$

where $C_{j,i}(n)$ is the expectation of product of number of neutrons in region j and i at generation n, and $\mu_i(n)$ is the expected number of neutrons in region i at generation n (Eq 3.3.23).

The recursive formula for $C_{j,i}(n)$ in Eq 3.3.39 can be transformed to $Cov_{j,i}(n)$ as

$$\begin{aligned} Cov_{j,i}(n+1) &= M_j^h M_i^l C_{h,l}(n) + \mu_l(n)V_{j,i}^l - \mu_j(n+1)\mu_i(n+1) \\ &= M_j^h M_i^l C_{h,l}(n) + \mu_l(n)V_{j,i}^l - \mu_h(n)M_j^h\mu_l(n)M_i^l \\ &= M_j^h M_i^l \left(C_{h,l}(n) - \mu_h(n)\mu_l(n)\right) + \mu_l(n)V_{j,i}^l \\ &= M_j^h M_i^l Cov_{h,l}(n) + \mu_l(n)V_{j,i}^l \\ &= M_j^h M_i^l Cov_{h,l}(n) + \not{V}_{j,i}(n) \end{aligned} \tag{4.1.2}$$

where the first equality uses the evolution of $\mu(n)$ in Eq 3.3.32, and the last equality defines $\not{V}_{j,i}(n) \equiv \mu_l(n)V_{j,i}^l$. For convenience, Eq 4.1.2 is rewritten in matrix form

$$Cov(n+1) = M^T Cov(n) M + \not{V}(n). \tag{4.1.3}$$

Suppose the initial source $\mu_l(0)$ is identical to the fundamental eigenmode of the fission matrix M, that is

$$\mu_l(0)M_i^l = \mu_i(0) \tag{4.1.4}$$

Then the processes $\mu_l(n)$ as function of n for each l are stationary

$$\mu_i(n+1) = \mu_l(n)M_i^l \quad \forall n \tag{4.1.5}$$

And the matrix $\mu_l(n)V^l_{j,i}$ is stationary and defined as

$$\not{V}_{j,i} \equiv \mu_l(n)V^l_{j,i} \tag{4.1.6}$$

For the fixed initial source, the spatial moment $C(0)$ is deterministic and given by

$$C_{i,j}(0) = \mu_i(0)\mu_j(0) \tag{4.1.7}$$

therefore the $Cov_{i,j}(0) = 0$. Inserting in the recursive formula for $Cov(n)$ leads to

$$Cov(n) = \sum_{i=n}^{1} \left(M^T\right)^{i-1} \not{V} M^{i-1} \tag{4.1.8}$$

If we further assume that there is no branching processes within each generation and that the MBP model is critical, Eq 3.3.71 and Eq 3.3.74 simplify $V^l_{i,j}$ to

$$V^l_{i,j} = \sigma_i^2 M_i^l \delta_{i,j} - M_i^l M_j^l \tag{4.1.9}$$

where $\sigma_i^2 \equiv \mathbb{E}\nu_i^2$. Multiplying $V^l_{i,j}$ by eigendistribution μ_l yields

$$\not{V}_{i,j} = \sigma_i^2 \mu_i \delta_{i,j} - \mu_l M_i^l M_j^l \tag{4.1.10}$$

Inserting the explicit form of $\not{V}$ (Eq 4.1.10) in the equation for $Cov(n)$ (Eq 4.1.8) gives

$$Cov(n)_{h,k} = \sum_{g=n}^{1} \left\{ \left(\left(M^T\right)^{g-1} \right)^i_h \sigma_i^2 \mu_i \delta_{i,j} \left(M^{g-1}\right)^j_k - \left(\left(M^T\right)^{g-1} \right)^i_h \mu_l M_i^l M_j^l \left(M^{g-1}\right)^j_k \right\} \tag{4.1.11}$$

4.2 Continuous space

In a continuous space, the summations over the phase space region indices are replaced with integrals over the whole space composed of an infinite number of infinitesimal

intervals. We can thus define $\mu(x)$ such that $\mu(r)dr$ is the expected number of neutrons in region dr around r which corresponds to the fundamental eigenmode.

It is worthwhile to note that in this section, dr (and similarly dr' and dr'') denotes any infinitesimal volume in the neutron phase space. Define $M^{(1)}(r, r')$ such that $M(r, r')dr'$ is the expected number of neutrons in region dr' around r' induced by one source neutron at position r from the previous generation.

For a critical system, $M(r, r')dr'$ is also equal to the probability that a source neutron from position x is absorbed in dr' around r' after one generation.

$\mu(r)$ is the solution (fundamental eigenmode) of the integral equation corresponding to eigenvalue 1 (since our system is assumed critical).

$$\mu(r) \equiv \int dr' \mu(r') M^{(1)}(r', r) \tag{4.2.1}$$

where the integral over the whole phase space is omitted. Taking the of matrix M can be written as

$$M^{(g)}(r, r') \equiv \int dr'' M^{(g-1)}(r, r'') M^{(1)}(r'', r') \tag{4.2.2}$$

where $M^{(g)}(r, r')dr'$ represents the probability that a source neutron from position r is absorbed in dr' around r' after g generations.

Define $V(r, r', r'')$ such that $V(r, r', r'')dr'dr''$ is the covariance of the number of neutrons at region dr' and dr'' induced by one source neutron from position r from the last generation.

From Eq 4.1.9, it can be explicitly written as

$$V(r, r', r'') = \sigma(r')^2 M^{(1)}(r, r')\delta(r', r'') - M^{(1)}(r, r')M^{(1)}(r, r'') \tag{4.2.3}$$

$\not{V}(r', r'')$ is then naturally defined as

$$\begin{aligned} \not{V}(r', r'') &= \int dr \mu(r) V(r, r', r'') \\ &= \sigma(r')^2 \mu(r')\delta(r', r'') - \int dr \mu(r) M^{(1)}(r, r')M^{(1)}(r, r'') \end{aligned} \tag{4.2.4}$$

Define $\mathrm{Cov}^{(n)}(r,r')$ such that $\mathrm{Cov}^{(n)}(r,r')drdr'$ is the covariance of the number of neutrons at region dr around r and dr around r' at generation n. Similarly, define $\mathrm{Cov}^{(n,n+d)}(r,r')$ such that $\mathrm{Cov}^{(n)}(r,r')drdr'$ is the covariance of the number of neutrons in region dr around r at generation n and in region dr' around r' at generation $n+d$. Also, Eq 3.3.51 that relates the spatial covariance within one generation and spatial covariance across different generations can be generalized to continuous space as

$$\mathrm{Cov}^{(n,n+d)}(r,r') = \int dr''\,\mathrm{Cov}^{(n)}(r,r'')M^{(d)}(r'',r') \tag{4.2.5}$$

The condensation equations for the neutron count in region I and covariance of the neutron count in region I_1 at generation n and region I_2 at generation $n+d$ are given by

$$\mu_I(n) = \int_I dr\mu(r) \tag{4.2.6}$$

$$\mathrm{Cov}[Z_{I_1}(n), Z_{I_2}(n+d)] = \int_{I_1} dr \int_{I_2} dr'\,\mathrm{Cov}^{(n,n+d)}(r,r') \tag{4.2.7}$$

The spatial covariance of a tally in a finite region, I, at two different generations corresponds to the condensation equation (Eq 3.3.70). It suffices to only keep the $\mathcal{O}(\mathbb{E}\epsilon^2)$ terms in Eq 3.3.70.

$$\begin{aligned}
&Cov[X_I(n), X_I(n+d)] \\
&= \frac{\mu_I(n)}{\mu(n)}\mathbb{E}\left[\left(\frac{Z(n)}{\mu(n)} - \frac{Z_I(n)}{\mu_I(n)}\right)\left(\frac{Z(n+d)}{\mu(n+d)} - \frac{Z_I(n+d)}{\mu_I(n+d)}\right)\right]\frac{\mu_I(n+d)}{\mu(n+d)} + o(\mathbb{E}\epsilon^2) \\
&= \frac{\mu_I(n)}{\mu(n)}\left[\frac{\mathrm{Cov}[Z(n), Z(n+d)]}{\mu(n)\mu(n+d)} - \frac{\mathrm{Cov}[Z(n), Z_I(n+d)]}{\mu(n)\mu_I(n+d)}\right. \\
&\qquad \left. - \frac{\mathrm{Cov}[Z_I(n), Z(n+d)]}{\mu_I(n)\mu(n+d)} + \frac{\mathrm{Cov}[Z_I(n), Z_I(n+d)]}{\mu_I(n)\mu_I(n+d)}\right]\frac{\mu_I(n+d)}{\mu(n+d)} + o(\mathbb{E}\epsilon^2)
\end{aligned} \tag{4.2.8}$$

where Eq 3.3.70 is rewritten in terms of covariances for convenience.

If $\mu(r)$ is scaled by a factor s, the volume ratios outside the bracket in Eq 4.2.8 do

not change. The covariance terms within the bracket will also be scaled by s, while the μ terms in the denominator will be scaled by $1/s^2$. Therefore, the covariance of X tallies will be scaled by $1/s$.

4.3 Analysis of 1D homogeneous model

The above model can be applied to a continuous homogeneous problem. Let's suppose the we have a $1D$ system of dimension $[-L, L]$ with reflective boundaries with a macroscopic absorption cross section c.

And in this section, the infinitesimal volume dx is understood as an interval on the real line.

4.3.1 Transition Kernel and Eigenmode

First, we evaluate $M^{(1)}(x, x')$, the probability density for a neutron starting at x being absorbed around x'. $M^{(1)}(x, x')$ can be decomposed into all the components of a neutron reaching x' from x through all possible paths.

The probability that a neutron from x is absorbed in region $[x', x' + dx']$ without any reflection at $x = -L$ or $x = L$ is proportional to $e^{-c|x-x'|}dx'$. The corresponding probability for a neutron being absorbed after reflection once at $x = L$ is proportional to $e^{-c|2L-x-x'|}dx'$ because the source neutron can be equivalently viewed as starting from $2L - x$, the mirror image of x about the right boundary $x = L$. Similarly, the probability for a neutron being absorbed after reflection once at $x = -L$ is proportional to $e^{-c|-2L-x-x'|}dx'$ due to the image source at $-2L - x$.

It can be shown by induction that if a neutron reaches x' from x after reflecting at $x = -L$ for n_- times (number of reflections on left surface) and reflecting at $x = L$ for n_+ times (number of reflections on right surface), the coordinate of the equivalent image x_{img} satisfies

$$x_{img} = \begin{cases} (2(n_- + n_+)L - x)\,(-1)^{n_- + n_+ + 1} \equiv x_{img}(n_- + n_+, +) & if\ first\ reflected\ at\ L \\ (2(n_- + n_+)L + x)\,(-1)^{n_- + n_+} \equiv x_{img}(n_- + n_+, -) & if\ first\ reflected\ at\ -L \end{cases} \tag{4.3.1}$$

x_{img} only depends on the initial direction and number of reflections.

The probability density $M^{(1)}(x, x')$ can be calculated from all possible $x_{img}(n, +)$'s and $x_{img}(n, -)$'s.

$$M^{(1)}(x, x') \propto e^{-c|x - x'|} + \sum_{n=1} e^{-c|x_{img}(n,+) - x'|} + \sum_{n=1} e^{-c|x_{img}(n,-) - x'|} \tag{4.3.2}$$

Then we insert the explicit form of x_{img} (Eq 4.3.1) into Eq 4.3.2.

$$M^{(1)}(x, x') \propto e^{-c|x - x'|} + \sum_{n=1} e^{-c|(2nL - x)(-1)^{n+1} - x'|} + \sum_{n=1} e^{-c|(2nL + x)(-1)^{n} - x'|} \tag{4.3.3}$$

Then separating the summation over index n into odd ($n = 2k-1$) and even ($n = 2k$) integers simplifies Eq 4.3.3 into Eq 4.3.4.

$$\begin{aligned} M^{(1)}(x, x') \propto e^{-c|x - x'|} &+ \sum_{k=1} e^{-c|(4kL - x)(-1) - x'|} + \sum_{k=1} e^{-c|(4kL - 2L - x) - x'|} \\ &+ \sum_{k=1} e^{-c|(4kL + x) - x'|} + \sum_{k=1} e^{-c|(4kL - 2L + x)(-1) - x'|} \end{aligned} \tag{4.3.4}$$

Realizing that $x - x' \in [-L, L]$ and $x + x' \in [-2L, 2L]$, the absolute value notation can be removed in the last four terms. Eq 4.3.4 is then simplified to Eq 4.3.5.

$$\begin{aligned} M^{(1)}(x, x') \propto e^{-c|x - x'|} &+ \sum_{k=1} e^{-c(4kL - x + x')} + \sum_{k=1} e^{-c(4kL - 2L - x - x')} \\ &+ \sum_{k=1} e^{-c(4kL + x - x')} + \sum_{k=1} e^{-c(4kL - 2L + x + x')} \end{aligned} \tag{4.3.5}$$

Finally, we evaluate the geometric series indexed by k and reach Eq 4.3.6.

$$\begin{aligned} M^{(1)}(x,x') &\propto e^{-c|x-x'|} + \left(e^{c(x-x')} + e^{c(2L+x+x')} + e^{-c(x-x')} + e^{c(2L-(x+x'))}\right) \sum_{k=1} e^{-4kcL} \\ &= e^{-c|x-x'|} + \left(2cosh[c(x-x')] + e^{2cL} 2sinh[c(x+x')]\right) \frac{e^{-4cL}}{1-e^{-4cL}} \\ &= e^{-c|x-x'|} + \frac{e^{-2cL}cosh[c(x-x')] + sinh[c(x+x')]}{sinh[2cL]} \end{aligned} \tag{4.3.6}$$

Integrating Eq 4.3.6 over $x' \in [-L, L]$ is equal to $\frac{2}{c}$, which allows to define the normalization constant and write

$$M^{(1)}(x,x') = \frac{c}{2}\left\{e^{-c|x-x'|} + \frac{e^{-2cL}cosh[c(x-x')] + sinh[c(x+x')]}{sinh[2cL]}\right\} \tag{4.3.7}$$

since the probability that a neutron starting from x ends between $[-L, L]$ is 1. It is worthwhile to note that in an infinite system,

$$M^{(1)}_{(0)}(x,x') = \frac{c}{2} e^{-c|x-x'|} \tag{4.3.8}$$

The identical normalization factor in Eq 4.3.7 and Eq 4.3.8 verifies that the contribution from reflection at boundaries has been treated correctly.

In order to predict the correlation (and the covariance) of a specific region, the power of the kernel as defined in Eq 4.2.2 is needed. A simple closed form for the power of matrix $M^{(1)}(x,x')$ evaluated in Eq 4.3.7 does not exist. Alternatively, $M^{(1)}(x,x')$ can be expanded using an orthonormal basis in region $[-L, L]$. The power of the kernel is then converted to the power of the corresponding matrix elements. For convenience, the bra-ket notation is used. For the basis set $\{|\alpha\rangle\}$, the matrix element is calculated as

$$\begin{aligned} m_{\alpha,\beta} &\equiv \langle\alpha|M^{(1)}|\beta\rangle \\ &= \langle\alpha|x\rangle \ \langle x|M^{(1)}|x'\rangle \ \langle x'|\beta\rangle \\ &= \int_{-L}^{L} dx \int_{-L}^{L} dx' \alpha(x) M^{(1)}(x,x') \beta(x') \end{aligned} \tag{4.3.9}$$

and the kernel $M^{(1)}(x,x')$ can be reconstructed from the matrix elements as

$$\begin{aligned} M^{(1)}(x,x') &= \langle x|\, M^{(1)}\, |x'\rangle \\ &= \langle x|\alpha\rangle\ \langle\alpha|M^{(1)}|\beta\rangle\ \langle\beta|x'\rangle \\ &= \sum_{\alpha,\beta} \alpha(x) m_{\alpha,\beta}\beta(x') \end{aligned} \tag{4.3.10}$$

The power of the kernel $M^{(1)}(x,x')$ (Eq 4.2.2) can be represented in matrix form as

$$\begin{aligned} M^{(g)}(x,x') &= \langle x|M^{(g-1)}|x''\rangle\ \langle x''|M^{(1)}|x'\rangle = \langle x|M^{(g-1)}M^{(1)}|x'\rangle \\ &= \langle x|\alpha\rangle\ \langle\alpha|M^{(g-1)}|\gamma\rangle\ \langle\gamma|M^{(1)}|\beta\rangle\ \langle\beta|x'\rangle \\ &= M^{(g-1)}_{\alpha,\gamma} m_{\gamma,\beta}\ \langle x|\alpha\rangle\ \langle\beta|x'\rangle \\ &= \sum_{\alpha,\beta} (M^{(g-1)}m)_{\alpha,\beta}\alpha(x)\beta(x') \end{aligned} \tag{4.3.11}$$

For the simple homogeneous problem in 1D, the eigenfunctions of the kernel can be calculated and normalized to form an orthonormal basis such that the corresponding matrix is diagonal.

Because the kernel $M^{(1)}(x,x')$ is symmetric, the normalization condition directly shows that $\mu(x) = 1$ is an eigenfunction corresponding to an eigenvalue of 1.

$$\int_{-L}^{L} M^{(1)}(x,x')\mu(x) = 1\mu(x') \tag{4.3.12}$$

Additionally, reflective boundaries restrict the eigenfunctions to those with 0 derivative at $-L$ and L. The eigenfunctions are $sin(\frac{2k-1}{2}\frac{\pi x}{L})$ and $cos(k\frac{pix}{L})$ for $k \in \mathbb{Z}^+$. The corresponding eigenvalues can be found by

$$\int_{-L}^{L} M^{(1)}(x,x') sin\left(\frac{2k-1}{2}\frac{\pi x}{L}\right) = \frac{1}{1+\left(\frac{2k-1}{2}\frac{\pi}{cL}\right)^2} sin\left(\frac{2k-1}{2}\frac{\pi x'}{L}\right) \tag{4.3.13}$$

$$\int_{-L}^{L} M^{(1)}(x,x') cos\left(k\frac{\pi x}{L}\right) = \frac{1}{1+\left(\frac{k\pi}{cL}\right)^2} cos\left(k\frac{\pi x}{L}\right) \tag{4.3.14}$$

We index the eigenfunction $\mu(x) = 1$ as $|0\rangle$, the sin functions as $|1\rangle, |3\rangle, \cdots$ and the

cos functions as $|2\rangle, |4\rangle, \cdots$. With appropriate normalization, the eigenfunctions of $M^{(1)}(x, x')$ are

$$\begin{aligned} \langle x|0\rangle &= \frac{1}{\sqrt{2L}} \\ \langle x|2k-1\rangle &= \frac{1}{\sqrt{L}} sin\left(\frac{2k-1}{2}\frac{\pi x}{L}\right) \\ \langle x|2k\rangle &= \frac{1}{\sqrt{L}} cos\left(\frac{2k}{2}\frac{\pi x}{L}\right) \end{aligned} \tag{4.3.15}$$

From Eq 4.3.13 and Eq 4.3.14, the corresponding eigenvalues are

$$\lambda_k = \frac{1}{1 + \left(\frac{k\pi}{2cL}\right)^2} \tag{4.3.16}$$

It can be shown that the basis in Eq 4.3.15 is orthonormal.

$$\langle \alpha|\beta\rangle = \delta_{\alpha,\beta} \tag{4.3.17}$$

Then the matrix elements are defined as

$$m_{k,j} = \langle k|M^{(1)}|j\rangle = \lambda_j \langle k|j\rangle = \lambda_k \delta_{k,j} \tag{4.3.18}$$

According to Eq 4.3.11,

$$M^{(g)}(x, x') = \sum_k \lambda_k^g k(x) k(x') \tag{4.3.19}$$

Using the eigenfunctions in Eq 4.3.15 and eigenvalues in Eq 4.3.16, $M^{(g)}(x, x')$ can

be explicitly written as

$$M^{(g)}(x,x') = \frac{1}{2L} + \frac{1}{L}\sum_{k=1}^{\infty}\left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^g cos\left(\frac{k\pi x}{L}\right) cos\left(\frac{k\pi x'}{L}\right) + \frac{1}{L}\sum_{k=1}^{\infty}\left(\frac{1}{1+\left(\frac{(2k-1)\pi}{2cL}\right)^2}\right)^g sin\left(\frac{(2k-1)\pi x}{2L}\right) sin\left(\frac{(2k-1)\pi x'}{2L}\right) \tag{4.3.20}$$

$M^{(1)}(x,x')$ is symmetric and all eigenvalues are positive.

4.3.2 Spatial covariance

The explicit form of the kernel $M^{(1)}(x,x')$ for the homogeneous problem allows us to perform an analysis of the spatial correlations. With

$$\mu(x) = \frac{1}{2L} \tag{4.3.21}$$

as the eigenfunction corresponding to the largest eigenvalue 1 and the transition kernel being symmetric, $\not{V}(r',r'')$ (Eq 4.1.10) is simplified to

$$\not{V}(x',x'') = \frac{1}{2L}\left(\sigma^2\delta(x',x'') - M^{(2)}(x',x'')\right) \tag{4.3.22}$$

The spatial covariance in generation n, $\mathrm{Cov}(n)_{i,j}$ (Eq 4.1.11) becomes

$$Cov^{(n)}(x,x') = \frac{1}{2L}\sum_{g=n}^{1}\left\{\sigma^2 M^{(2(g-1))}(x,x') - M^{(2g)}(x,x')\right\} \tag{4.3.23}$$

where $M^{(0)}$ is defined as

$$M^{(0)}(x,x') \equiv \delta(x,x') \tag{4.3.24}$$

Then, relating the spatial covariances within one generation to the spatial covariances across different generations (Eq 4.2.5), we obtain

$$Cov^{(n,n+d)}(x,x') = \frac{1}{2L}\sum_{g=n}^{1}\left\{\sigma^2 M^{(2(g-1)+d)}(x,x') - M^{(2g+d)}(x,x')\right\} \tag{4.3.25}$$

From Eq 4.3.23 and Eq 4.3.25 along with the explicit expressions of $\mu(x)$ (Eq 4.3.21) and $M^{(g)}(x,x')$ (Eq 4.3.20) enables explicit evaluation of spatial correlation $\mathrm{Cov}[X_I(n), X_I(n+d)]$ according to Eq 4.2.8. This will require the integral over region $I \times I$ of $M^{(g)}(x,x')$.

Without loss of generality, let's define

$$I = [-a, a] \tag{4.3.26}$$

to investigate the correlation of the central tally region of width $2a$. Also, let's define

$$W = [-L, L] \tag{4.3.27}$$

The first order moments in Eq 4.2.8 are simply

$$\begin{aligned} \mu(n) &= \mu(n+d) = 1 \\ \mu_I(n) &= \mu_I(n+d) = \frac{a}{L} \end{aligned} \tag{4.3.28}$$

The integral of the matrix $M^{(g)}(x,x')$ is calculated as

$$\begin{aligned} &\int_I dx \int_I dx' M^{(g)}(x,x') = \frac{1}{2L}\int_I dx \int_I dx' \\ &+ \frac{1}{L}\sum_{k=1}^{\infty}\left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^g \int_I dx cos\left(\frac{k\pi x}{L}\right)\int_I dx' cos\left(\frac{k\pi x'}{L}\right) \\ &+ \frac{1}{L}\sum_{k=1}^{\infty}\left(\frac{1}{1+\left(\frac{(2k-1)\pi}{2cL}\right)^2}\right)^g \int_I dx sin\left(\frac{(2k-1)\pi x}{2L}\right)\int_I dx' sin\left(\frac{(2k-1)\pi x'}{2L}\right) \end{aligned} \tag{4.3.29}$$

Integration of the odd function $sin\left(\frac{(2k-1)\pi x}{2L}\right)$ over the region $I = [-a, a]$ symmetric

about the origin vanishes. Only the integration of the *cosine* functions remain in Eq 4.3.30. And they are evaluated explicitly in Eq 4.3.31.

$$\int_I dx \int_I dx' M^{(g)}(x,x') = \frac{2a^2}{L} + \frac{1}{L}\sum_{k=1}^{\infty}\left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^g \int_I dx cos\left(\frac{k\pi x}{L}\right)\int_I dx' cos\left(\frac{k\pi x'}{L}\right) \tag{4.3.30}$$

$$= \frac{2a^2}{L} + \frac{4L}{\pi^2}\sum_{k=1}^{\infty}\left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^g \frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2} \tag{4.3.31}$$

We can then evaluate the $\mathrm{Cov}[Z_I(n), Z_I(n+d)]$ as below. First, we integrate the covariance density in Eq 4.3.25 over region I.

$$\mathrm{Cov}[Z_I(n), Z_I(n+d)] =$$
$$\frac{1}{2L}\sum_{g=n}^{1}\left\{\sigma^2\int_I dx\int_I dy' M^{(2(g-1)+d)}(x,x') - \int_I dx\int_I dx' M^{(2g+d)}(x,x')\right\} \tag{4.3.32}$$

Then, we insert the explicit expression of $M^{(2(g-1)+d)}$ and $M^{(2g+d)}$,

$$\mathrm{Cov}[Z_I(n), Z_I(n+d)] =$$
$$= \frac{1}{2L}\sum_{g=n}^{1}\left\{\frac{2a^2}{L}(\sigma^2-1) + \frac{4L}{\pi^2}\sum_{k=1}^{\infty}\left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^{2(g-1)+d}\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}\left(\sigma^2 - \left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^2\right)\right\} \tag{4.3.33}$$

Next, we change the order of the summation over basis index k and the summation over generation index g in Eq 4.3.33 and evaluate the geometric series with index g.

$$\mathrm{Cov}[Z_I(n), Z_I(n+d)] = \frac{a^2}{L^2}(\sigma^2-1)n +$$
$$+ \frac{2}{\pi^2}\sum_{k=1}^{\infty}\frac{\sigma^2 - \left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^2}{1-\left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^2}\left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^d \frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}\left(1-\left(\frac{1}{1+\left(\frac{k\pi}{cL}\right)^2}\right)^{2n}\right) \tag{4.3.34}$$

Finally, we recognize λ_{2k} in Eq 4.3.34 and with the assumption that stationarity is reached after sufficiently many generations n, $\lambda_{2k}^{2n} \to 0$.

$$\mathrm{Cov}[Z_I(n), Z_I(n+d)] = \frac{a^2}{L^2}(\sigma^2 - 1)n + \frac{2}{\pi^2}\sum_{k=1}^{\infty}\frac{\sigma^2 - \lambda_{2k}^2}{1-\lambda_{2k}^2}\lambda_{2k}^d\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2} \tag{4.3.35}$$

Next, we can evaluate $\mathrm{Cov}[Z(n), Z_I(n+d)]$. The $\langle x|k\rangle$ part of $M^{(g)}(x,x')$ for all components will be integrated over the whole region W and the $\langle x'|k\rangle$ part will be integrated over the region of interest I. It can be seen that

$$\int_W dx\,\langle x|k\rangle = \begin{cases}\sqrt{2L} & k=0\\ 0 & otherwise\end{cases} \tag{4.3.36}$$

$$\int_I dx\,\langle x|0\rangle = \frac{2a}{\sqrt{2L}} \tag{4.3.37}$$

Therefore, only the 0^{th} term in the expansion of $M^{(g)}(x,x')$ contribute to the integral.

$$\begin{aligned}\mathrm{Cov}[Z(n), Z_I(n+d)] &= \frac{1}{2L}\left\{\sigma^2\sqrt{2L}\frac{2a}{\sqrt{2L}} - \sqrt{2L}\frac{2a}{\sqrt{2L}}\right\}\\ &= \frac{a}{L}(\sigma^2-1)n\end{aligned} \tag{4.3.38}$$

By symmetry,

$$\mathrm{Cov}[Z_I(n), Z(n+d)] = \frac{a}{L}(\sigma^2-1)n \tag{4.3.39}$$

Similarly, only the 0^{th} term in the expansion of $M^{(g)}(x,x')$ contribute to the integral of $\mathrm{Cov}[Z(n), Z(n+d)]$.

$$\begin{aligned}\mathrm{Cov}[Z(n), Z(n+d)] &= \frac{1}{2L}\left\{\sigma^2\sqrt{2L}\sqrt{2L} - \sqrt{2L}\sqrt{2L}\right\}\\ &= (\sigma^2-1)n\end{aligned} \tag{4.3.40}$$

All the terms required to calculate $\mathrm{Cov}[X_I(n), X_I(n+d)]$ in Eq 4.2.8 are known (Eq 4.3.28, Eq 4.3.35, Eq 4.3.38, Eq 4.3.39 and Eq 4.3.40). The covariance between

normalized neutron tally in region I accross generation lag d is thus

$$
\begin{aligned}
&Cov[X_I(n), X_I(n+d)] \\
&= \left(\frac{a}{L}\right)^2 \left[\frac{\left(\frac{a}{L}\right)^2(\sigma^2-1)n}{\frac{a}{L}\frac{a}{L}} - \frac{\frac{a}{L}(\sigma^2-1)n}{1\frac{a}{L}} - \frac{\frac{a}{L}(\sigma^2-1)n}{\frac{a}{L}1} + \frac{(\sigma^2-1)n}{11}\right] \\
&+ \frac{2}{\pi^2}\sum_{k=1}^{\infty}\frac{\sigma^2-\lambda_{2k}^2}{1-\lambda_{2k}^2}\lambda_{2k}^d\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2} \qquad (4.3.41) \\
&= \frac{2}{\pi^2}\sum_{k=1}^{\infty}\frac{\sigma^2-\lambda_{2k}^2}{1-\lambda_{2k}^2}\lambda_{2k}^d\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2} \qquad (4.3.42)
\end{aligned}
$$

where Eq 4.3.41 combines the 0^{th} term in $\mathrm{Cov}[Z_I(n), Z_I(n+d)]$ (Eq 4.3.35) with the other covariance terms and cancels them out. Note that though the comprising covariances of Z tallies are dependent on the generation number n, the approximated covariances of X tallies are stationary. The covariances of X tallies are function of generation lag d only (besides the system and tally region parameters) and the covariances decay exponentially as function of d.

4.3.3 Correlation coefficients

Similar calculations are performed for $\mathrm{Var}[X_I(n)]$ which will allow us to obtain $\rho_I(d)$ defined as $\mathrm{Cov}[X_I(n), X_I(n+d)]/\mathrm{Var}[X_I(n)]$.

$\mathrm{Var}[X_I(n)]$ is evaluated via $\mathrm{Cov}[X_I(n), X_I(n+d)]$ with $d = 0$. Thus $\mathrm{Cov}[Z_I(n), Z_I(n)]$, $\mathrm{Cov}[Z_I(n), Z(n)]$ and $\mathrm{Cov}[Z(n), Z(n)]$ are needed. Different from the covariance terms evaluated in last section, $d = 0$ leads to the appearance of $M^{(0)}(x, x') = \delta(x, x')$ in the summation indexed by generation number g. The following integrals of $\delta(x, x')$

are useful.

$$\int_I dx \int_I dx' \delta(x, x') = \int_I dx = 2a \tag{4.3.43}$$

$$\int_I dx \int_W dx' \delta(x, x') = \int_I dx = 2a \tag{4.3.44}$$

$$\int_W dx \int_I dx' \delta(x, x') = \int_W dx \mathbb{1}_{x \in I} = 2a \tag{4.3.45}$$

$$\int_W dx \int_W dx' \delta(x, x') = 2L \tag{4.3.46}$$

While integrating $\text{Cov}^{(n)}(x, x')$ (Eq 4.3.23) into $\text{Cov}[Z_I(n), Z_I(n)]$, $g = 1$ and $g \geq 2$ should be treated separately.

$$\begin{aligned}
&\text{Cov}[Z_I(n), Z_I(n)] = \\
&\frac{1}{2L}\left[\sigma^2 2a - \left(\left(\frac{2a}{\sqrt{2L}}\right)^2 + \frac{4L}{\pi^2}\sum_{k=1}\lambda_{2k}^2 \frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}\right)\right] \\
&+ \frac{1}{2L}\sum_{g=n}^{2}\left\{\frac{2a^2}{L}(\sigma^2 - 1) + \frac{4L}{\pi^2}\sum_{k=1}^{\infty}\lambda_{2k}^{2(g-1)}\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}\left(\sigma^2 - \lambda_{2k}^2\right)\right\}
\end{aligned} \tag{4.3.47}$$

Then Eq 4.3.48 evaluates the geometric series indexed by g in Eq 4.3.47 in a similar way to Eq 4.3.35.

$$\begin{aligned}
\text{Cov}[Z_I(n), Z_I(n)] = \frac{a}{L}\sigma^2 - \left(\frac{a}{L}\right)^2 + \left(\frac{a}{L}\right)^2(\sigma^2 - 1)(n-1) \\
- \frac{2}{\pi^2}\sum_{k=1}\lambda_{2k}^2\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2} + \frac{2}{\pi^2}\sum_{k=1}\frac{\sigma^2 - \lambda_{2k}^2}{1 - \lambda_{2k}^2}\lambda_{2k}^2\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}
\end{aligned} \tag{4.3.48}$$

Eq 4.3.49 merges the two summations over index k into one.

$$\begin{aligned}
\text{Cov}[Z_I(n), Z_I(n)] = \frac{a}{L}\sigma^2 - \left(\frac{a}{L}\right)^2\sigma^2 + \left(\frac{a}{L}\right)^2(\sigma^2 - 1)n \\
+ \frac{2}{\pi^2}\sum_{k=1}\frac{\sigma^2 - 1}{1 - \lambda_{2k}^2}\lambda_{2k}^2\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}
\end{aligned} \tag{4.3.49}$$

Next, we evaluate $\text{Cov}[Z(n), Z_I(n)]$, which is similar to the previous expression

for $\mathrm{Cov}[Z(n), Z_I(n+d)]$.

$$\mathrm{Cov}[Z(n), Z_I(n)] = \frac{1}{2L}\left\{\sigma^2\left(2a + \sum_{g=n}^{2}\sqrt{2L}\frac{2a}{\sqrt{2L}}\right) - \sum_{g=n}^{1}\sqrt{2L}\frac{2a}{\sqrt{2L}}\right\}$$
$$= \frac{a}{L}(\sigma^2 - 1)n \tag{4.3.50}$$

By symmetry,

$$\mathrm{Cov}[Z_I(n), Z(n)] = \frac{a}{L}(\sigma^2 - 1)n \tag{4.3.51}$$

Similarly,

$$\mathrm{Cov}[Z(n), Z(n)] = \frac{1}{2L}\left\{\sigma^2\left(2L + \sum_{g=n}^{2}\sqrt{2L}\frac{2L}{\sqrt{2L}}\right) - \sum_{g=n}^{1}\sqrt{2L}\frac{2L}{\sqrt{2L}}\right\}$$
$$= (\sigma^2 - 1)n \tag{4.3.52}$$

Then, we combine the covariance of neutron counts Z to the variance of relative counts X according to Eq 4.2.8.

$$Var[X_I(n)] = \mathrm{Cov}[X_I(n), X_I(n)]$$
$$= \left(\frac{a}{L}\right)^2\left[\frac{\left(\frac{a}{L}\right)^2(\sigma^2-1)n + \sigma^2\frac{a}{L}\left(1-\frac{a}{L}\right)}{\frac{a}{L}\frac{a}{L}} - \frac{\frac{a}{L}(\sigma^2-1)n}{1\frac{a}{L}} - \frac{\frac{a}{L}(\sigma^2-1)n}{\frac{a}{L}1} + \frac{(\sigma^2-1)n}{11}\right]$$
$$+ \frac{2}{\pi^2}\sum_{k=1}^{\infty}\frac{\sigma^2-1}{1-\lambda_{2k}^2}\lambda_{2k}^d\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2} \tag{4.3.53}$$
$$= \sigma^2\frac{a}{L}\left(1-\frac{a}{L}\right) + \frac{2}{\pi^2}\sum_{k=1}^{\infty}\frac{\sigma^2-1}{1-\lambda_{2k}^2}\lambda_{2k}^d\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2} \tag{4.3.54}$$

where Eq 4.3.53 combines the 0^{th} term of $\mathrm{Cov}[Z_I(n), Z_I(n)]$ (Eq 4.3.49) with the other variance terms and cancels them out. Note that the covariances of Z tallies are dependent on the generation number n and that the approximated covariances of X tallies are stationary.

With the covariances that are functions of only the generation lag and variances

that are stationary, the correlation coefficients of a normalized tally in region I are function of the generation lag d only.

$$\begin{aligned}\rho_I(d) &= \frac{\mathrm{Cov}[X_I(n), X_I(n+d)]}{\mathrm{Cov}[X_I(n), X_I(n)]} \\ &= \frac{\sum_{k=1}^{\infty} \frac{\sigma^2-\lambda_{2k}^2}{1-\lambda_{2k}^2} \lambda_{2k}^d \frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}}{\frac{\pi^2}{2}\sigma^2 \frac{a}{L}\left(1-\frac{a}{L}\right) + \sum_{k=1}^{\infty} \frac{\sigma^2-1}{1-\lambda_{2k}^2} \lambda_{2k}^d \frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}}\end{aligned} \tag{4.3.55}$$

A test problem is created to numerically verify the above derivations. Denote the problem as $TC1-1D$.

The material properties of $TC1-1D$ are identical to the homogeneous cube described in section 2.1.2,Table 2.1.

In the simple problem, only absorption events need to be sampled. $M^{(1)}(x, x')$ (Eq 4.3.7) can be used to sample the first absorption site x' for a neutron started at x. It is not straightforward to find the CDF of $M^{(1)}(x, x')$ and directly sample from it. However, $M^{(0)}(x, x')$, the probability density for the first absorption site in an infinite system can easily be sampled to find the first absorption site $x'_{(0)}$. $x'_{(0)}$ is thus not necessarily within $[-L, L]$, but an equivalent image within the interval can be found. It can be shown by induction that $x'_{(0)}$ and x' are related through

$$x' = (-1)^n sign(x'_{(0)}) \left(|x'_{(0)}| - 2nL\right) \tag{4.3.56}$$

where n satisfies

$$n < \frac{|x'_{(0)}|}{2L} + \frac{1}{2} \leq n+1 \tag{4.3.57}$$

After the first absorption site $x'_{(0)}$ is sampled according to $M^{(0)}(x, x')$ and transformed to x'. New neutrons are sampled at x'. The distribution of new neutrons per absorption follows Eq 2.1.8.

The simulation was performed with $100,000$ neutrons per generation for 1000 active generations. Additionally, 200 independent simulations were performed to generate reference statistics. Source neutrons are counted in region $I = [-a, a]$ for

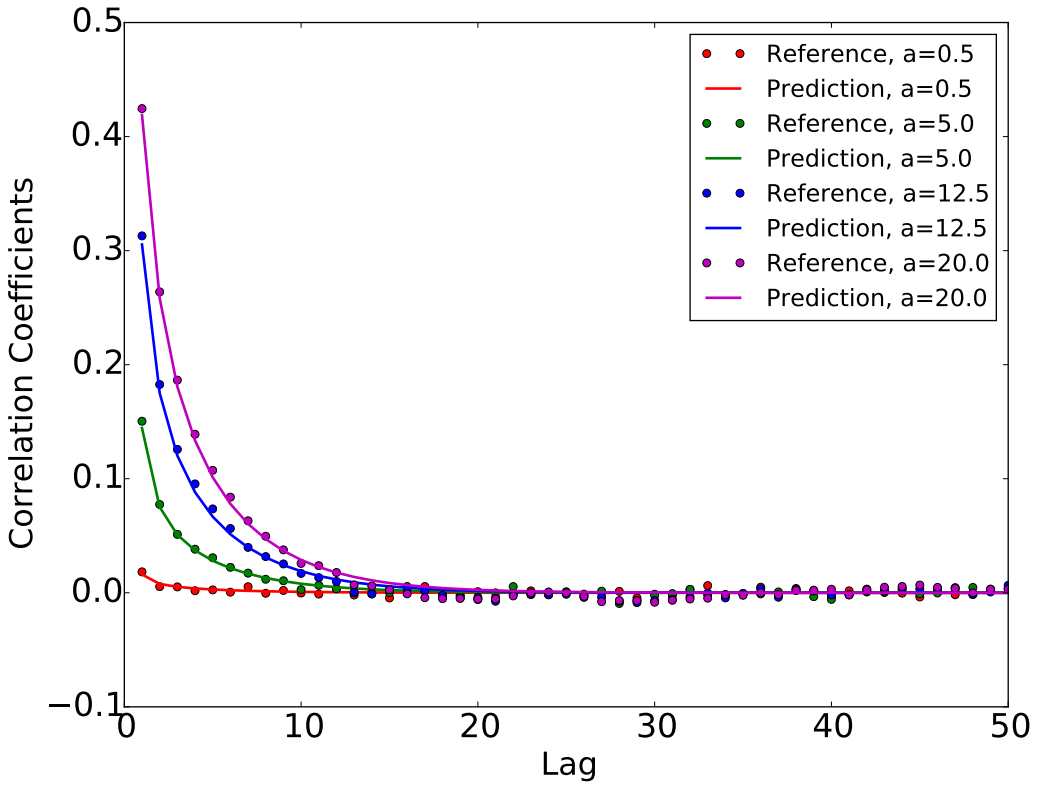


Figure 4-1: Correlation coefficients for central tally region with different sizes in $TC1-1D$. The dots correspond to correlation coefficients estimated from independent simulations. The solid line curves correspond to correlation coefficients predicted from the continuous model.

different values of a and the correlation coefficients estimated from the simulation are plotted along with the prediction from $\rho_I(d)$ in Eq 4.3.55 in Figure 4-1. Due to the stationarity of correlation coefficients discussed and numerically verified in section 2.3.1, $\rho(n, n+d)$ calculated above is then averaged over all available n's to reduce statistical noise. As expected, the correlation coefficients are higher in larger tally regions.

4.3.4 Variance underestimation ratio and asymptotic behavior

The covariances of normalized X tallies can now be used to predict the variance of tallies averaged over the active generations, $\mathrm{Var}\left[\bar{X}(N)\right]$, according to Eq 2.2.5. Equivalently, the correlation coefficients can be converted to the variance estimation

ratio $r(N)$ as defined in Eq 2.2.8.

$$r_I(N) = 1 + 2\sum_{d=1}^{N-1}\rho_I(d) - \frac{2}{N}\sum_{d=1}^{N-1} d\ \rho_I(d) \tag{4.3.58}$$

$$= 1 + 2\frac{\sum_{k=1}^{\infty}\frac{2}{\pi^2}\left(\sum_{d=1}^{N-1}\lambda_{2k}^d - \frac{1}{N}\sum_{d=1}^{N-1} d\ \lambda_{2k}^d\right)\frac{\sigma^2-\lambda_{2k}^2}{1-\lambda_{2k}^2}\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}}{\mathrm{Var}[X_I]} \tag{4.3.59}$$

$$= 1 + 2\frac{\sum_{k=1}^{\infty}\frac{2}{\pi^2}\left(\frac{\lambda_{2k}(1-\lambda_{2k}^{N-1})}{1-\lambda_{2k}} - \frac{1}{N}\frac{\lambda_{2k}(1-\lambda_{2k}^{N-1})-(N-1)\lambda_{2k}^N(1-\lambda_{2k})}{(1-\lambda_{2k})^2}\right)\frac{\sigma^2-\lambda_{2k}^2}{1-\lambda_{2k}^2}\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}}{\mathrm{Var}[X_I]} \tag{4.3.60}$$

where Eq 4.3.58 separates the summation in Eq 2.2.8 over generation lag d into two parts that depends on d in different ways. Eq 4.3.59 inserts the explicit form of $\rho_I(d)$, switches the order of summation over expansion series and over generation lag, and separates the summands that depend on the generation lag. Eq 4.3.60 evaluates the summation over the generation lag explicitly. $\mathrm{Var}[X_I]$ was written in Eq 4.3.54 as $\mathrm{Var}[X_I(n)]$ and its generation independence is emphasized in Eq 4.3.60.

In Figure 4-2, numerical reference of the variance underestimation ratio from the independent simulations for the problem $TC1-1D$ are plotted with the derivation of $r_I(N)$ above.

The variance underestimation ratio in (Eq 4.3.60) also contains the real variance of the normalized tallies averaged over the active generations.

$$\mathrm{Var}[\bar{X}_I(N)] = \mathrm{Var}[X_I]r_I(N) \tag{4.3.61}$$

The expectation of relative square error (as defined in Eq 2.2.13) can be calculated as

$$\mathbb{E}RSE_I(N) = \frac{\mathrm{Var}[\bar{X}_I(N)]}{s\mu_I^2} = \frac{\mathrm{Var}[X_I]}{s}r_I(N)\left(\frac{L}{a}\right)^2 \tag{4.3.62}$$

where s is the number of neutrons per generation. The normalized $\mu(x)$ chosen above (Eq 4.3.21) corresponds to 1 neutron per generation. $\mathrm{Var}[X_I]$ is scaled by $1/s$ to match the results from simulating $100,000$ neutrons per generation in $TC1-1D$.

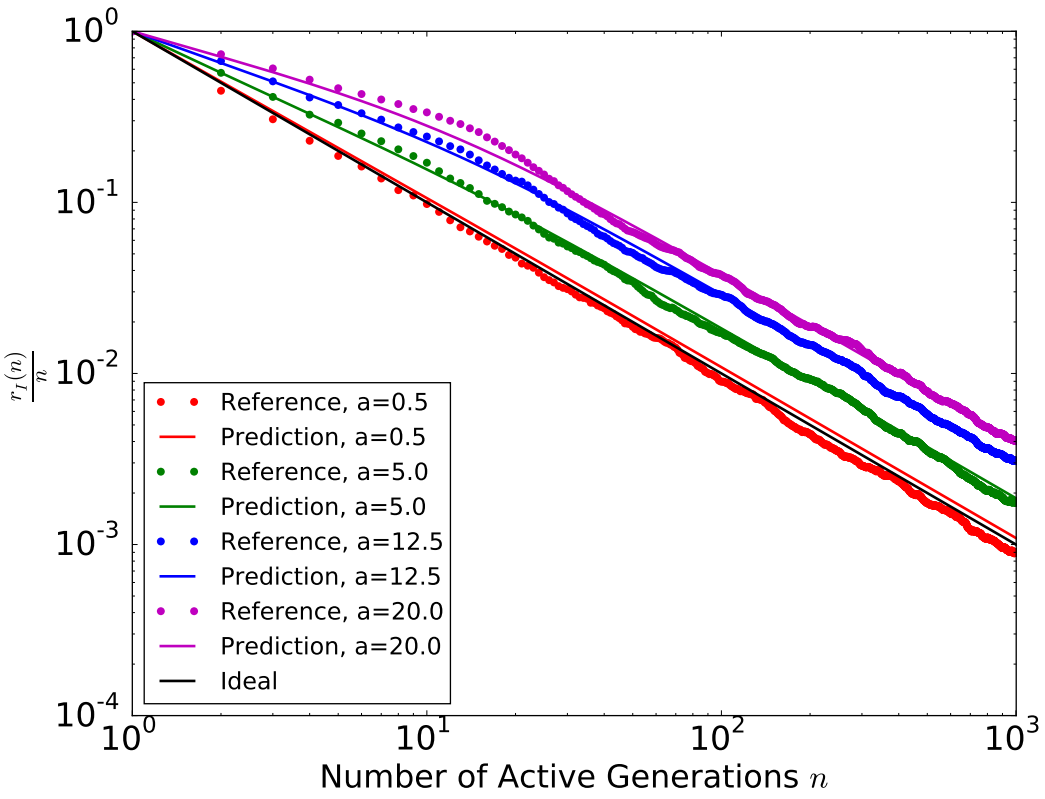


Figure 4-2: Variance convergence rate calculated from underestimation ratio for central tally region with different sizes in $TC1-1D$. The dots correspond to $\frac{r_I(n)}{n}$ estimated from independent simulations. The solid line curves correspond to predictions from the continuous model.

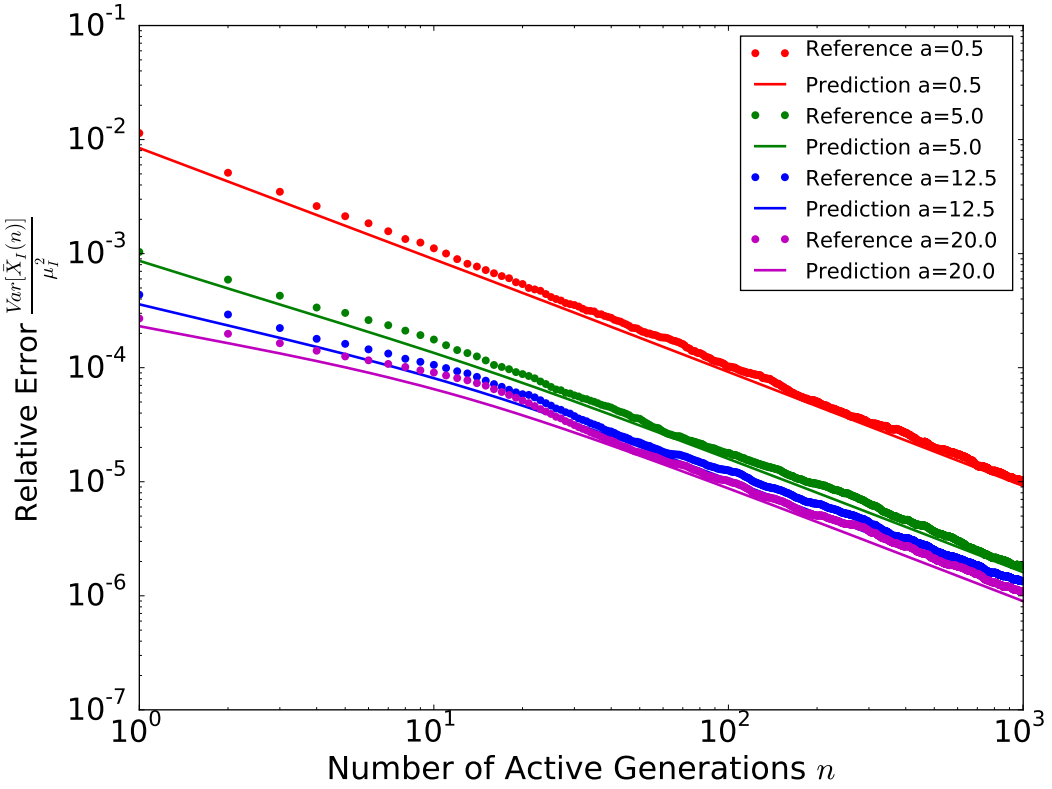


Figure 4-3: Relative square error for central tally region with different sizes in $TC1-1D$. The dots correspond to RSE_I estimated from independent simulations. The solid line curves correspond to predictions from the continuous model.

Under the assumption that the estimators are unbiased, $\mathbb{E}RSE_I$ is equal to the expectation of relative square error RSE_I (as defined in Eq 2.2.12). The $\mathbb{E}RSE_I$ derived above in Eq 4.3.62 is plotted with the reference relative square error obtained from many independent simulations in Figure 4-3.

Although smaller regions have lower correlation coefficients and thus their variance decreases at a rate closer to $1/N$, the relatively lower tally counts in the region contributes to a higher uncertainty magnitude.

As previously stated, the homogeneous continuous model will be used to simply determine if a large variance underestimation is expected for a given problem and associated tally size. The asymptotic variance underestimation ratio $r_I(\infty)$ is a good indicator of correlation level. In comparison with correlation coefficients which are characterized by the generation lag, $r_I(N)$ can be characterized by one scalar num-

ber, $r_I(\infty)$, that directly impacts variance. It can be seen from previous variance underestimation curves that in $log - log$ scale, the distance between the $\frac{r(n)}{n}$ vs n curves and ideal rate $\frac{1}{n}$ vs n curve approaches a constant as N approaches infinity which can be stated by Eq 4.3.63.

$$\begin{aligned} &\log\left(\frac{r(n)}{n}\right) - \log\left(\frac{1}{n}\right) = const \\ &\Leftrightarrow r(n) = e^{const} \end{aligned} \tag{4.3.63}$$

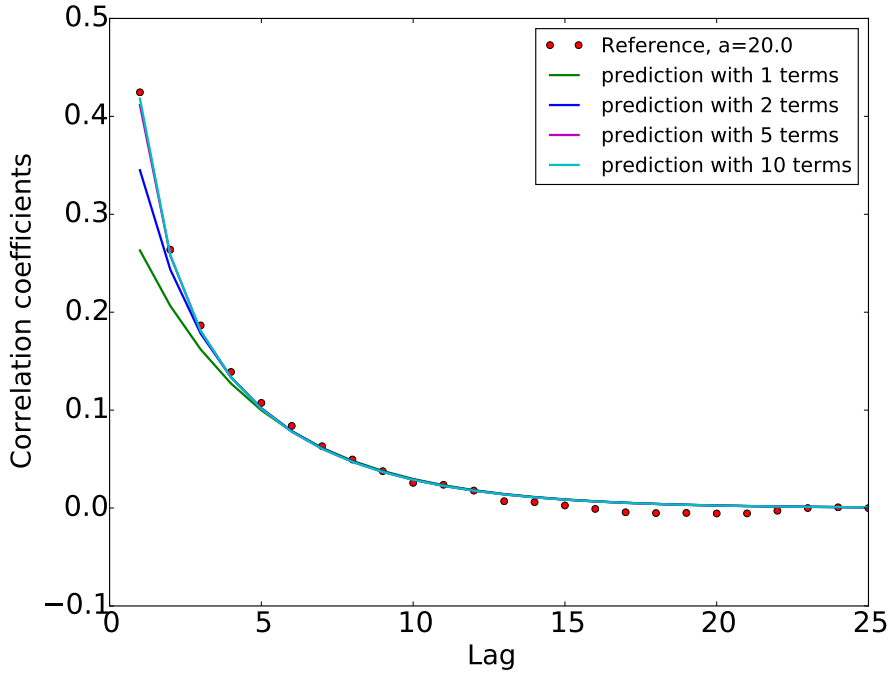
In addition, since the variance underestimation ratio $r_I(N)$ is determined by the cumulative behavior of correlation coefficients, fewer terms are required accurately predict the asymptotic behavior. Correlation coefficients and variance underestimation ratio for a tally of width $a = 20.0cm$ in $TC1-1D$ are plotted in Figure 4-4 along with the respective prediction with the expansion series truncated to different orders. Large deviations are observed for $\rho_I(d)$ values with a short generation lag, but the asymptotic value of r_I is fairly constant across all approximations.

The asymptotic variance underestimation ratio can be calculated by taking the limit of $N \to \infty$ in $r_I(N)$ (Eq 4.3.60).

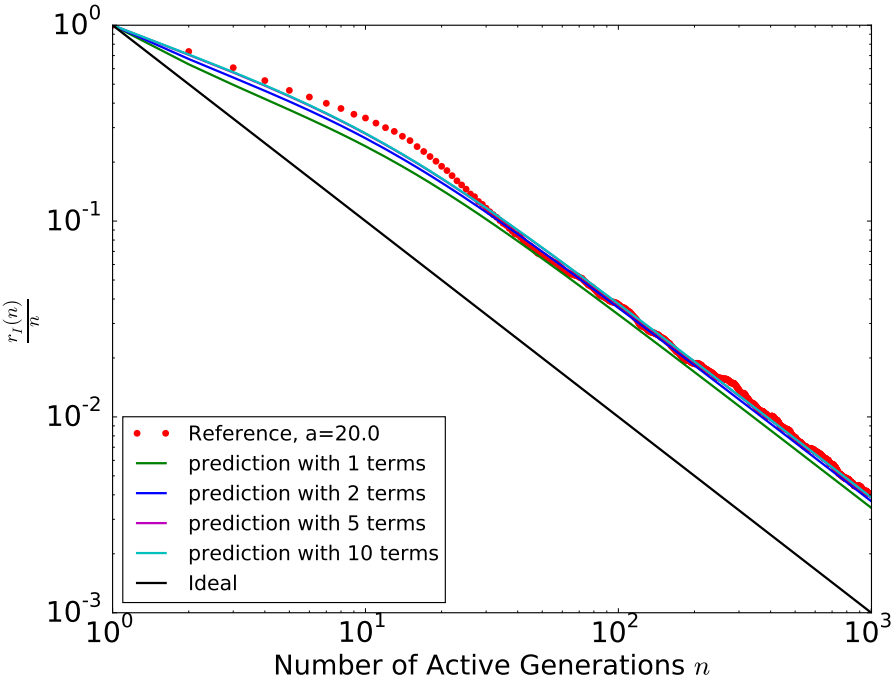
$$\begin{aligned} r_I &\equiv \lim_{N\to\infty} r_I(N) \\ &= 1 + 2\frac{\sum_{k=1}^{\infty}\left(\frac{\lambda_{2k}}{1-\lambda_{2k}}\right)\frac{\sigma^2-\lambda_{2k}^2}{1-\lambda_{2k}^2}\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}}{\frac{\pi^2}{2}\sigma^2\frac{a}{L}\left(1-\frac{a}{L}\right) + \sum_{k=1}^{\infty}\frac{\sigma^2-1}{1-\lambda_{2k}^2}\lambda_{2k}^d\frac{sin^2\left(\frac{k\pi a}{L}\right)}{k^2}} \end{aligned} \tag{4.3.64}$$

The asymptotic r_I values are plotted as a function of tally size in Figure 4-5. The numerical reference is approximated by $r_I(1000)$.

Consistency between the predicted values and references (Figure 4-1, Figure 4-2 and Figure 4-5) indicate that the continuous model generalized from the discrete model of Multitype Branching Processes can be applied to provide an approximate assessment of the correlation in a Monte Carlo simulation with generation-to-generation dependence and branching processes.



(a) Correlation coefficients



(b) Variance convergence rate

Figure 4-4: $\rho_I(d)$ and $\frac{r(n)}{n}$ and prediction series truncation. $a = 20.0cm$ is fixed. The dots correspond to estimation from independent simulations for $TC1-1D$. And the solid line curves correspond to prediction with different truncations applied to tally in the central region.

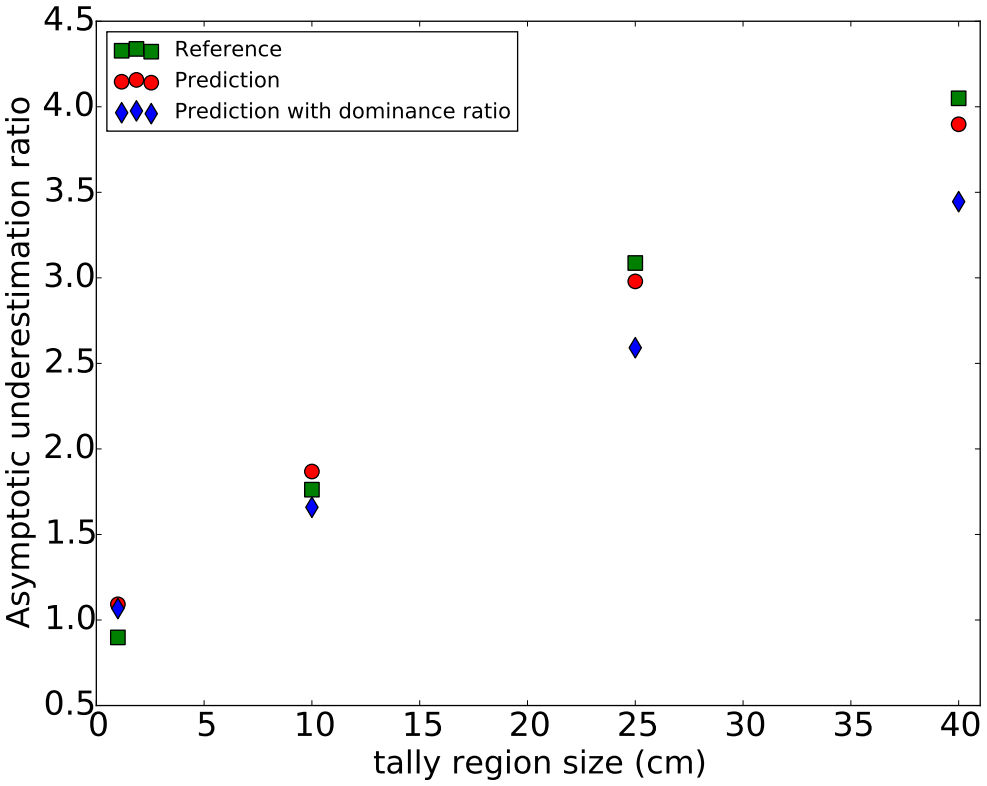


Figure 4-5: Asymptotic variance underestimation ratio for central tally region with different sizes in $TC1-1D$. The squares correspond to r_I estimated from independent simulations. The dots correspond to predictions from the continuous model with sufficient expansion terms. The diamonds correspond to predictions from the continuous model with the first expansion term.

However, the values of the correlation coefficients and asymptotic variance underestimation ratios are much lower than those observed in a typical nuclear reactor (Figure 1-2,Figure 2-2). From a discrete point of view, as the dimension of the system increases, one phase space cell has exponentially more neighboring phase space regions to be correlated with, and thus contributes to much higher correlation level in the tally region containing these phase space cells. Negligible correlation level for tally region with width $2a$ in a $1D$ system with width $2L$ does not necessarily imply negligible correlations for a tally region with volume $(2a)^3$ in a $3D$ system with volume $(2L)^3$. The analysis in the $1D$ system must be extended to realistic $3D$ systems.

4.4 Analysis of 3D homogeneous model

Let's now generalize the homogeneous $1D$ problem into $3D$. Suppose the system is in the volume $[-L_x, L_x] \times [-L_y, L_y] \times [-L_z, L_z]$ with constant macroscopic cross sections and reflective boundaries. Denote any position in the system as $\vec{r} = (x, y, z)$.

4.4.1 Transition Kernel and Diffusion approximation

For the $3D$ system, it is not easy to find the exact form of $M^{(1)}(\vec{r}, \vec{r}')$ as was doen for $M^{(1)}(x, x')$ in Eq 4.3.7. Instead, this section builds a relation between the integral kernel $M^{(1)}(\vec{r}, \vec{r}')$ and the diffusion operator in a $1D$ system and then generalize the relation as an approximation in $3D$.

Recall that the one group diffusion approximation for the neutron transport equation in a homogeneous system reads as [38]

$$-D\nabla^2\phi(x) + \Sigma_a\phi(x) = \frac{\nu\Sigma_f}{k_{eff}}\phi(x) \tag{4.4.1}$$

Eq 4.4.1 can be converted to an eigenvalue form of the Laplace operator

$$\nabla^2\phi(x) = \frac{1 - \frac{k_\infty}{k_{eff}}}{M^2}\phi(x) \tag{4.4.2}$$

where k_∞ and the migration area M^2 are defined as

$$k_\infty \equiv \frac{\nu\Sigma_f}{\Sigma_a} \tag{4.4.3}$$

$$M^2 \equiv \frac{D}{\Sigma_a} \tag{4.4.4}$$

where the definition of ν is in Eq 2.1.1.

If we assume that the problem is critical, which implies that $\mathbb{E}\xi = 1$, where ξ denotes the number of new neutrons per absorption. The criticality condition is equivalent to

$$1 = \nu\mathbb{P}(\xi \neq 0) = \nu\frac{\Sigma_f}{\Sigma_a} \tag{4.4.5}$$

The eigenvalue problem of the Laplace operator is thus simplified to

$$\nabla^2\phi(x) = \frac{1 - \frac{1}{k_{eff}}}{M^2}\phi(x) \tag{4.4.6}$$

The boundary conditions determine the eigenvalues of the Laplacian indexed by integer k, as seen below

$$\nabla^2\phi(x) = l_k\phi(x) \tag{4.4.7}$$

Eq 4.4.7 and Eq 4.4.6 can be used to express the relation between the eigenvalues of the transport equation, k_{eff} and the eigenvalues of the Laplacian equation, l_k, as

$$k_{keff,k} = \frac{1}{1 - l_k M^2} \tag{4.4.8}$$

In a $1D$ problem in $[-L, L]$ with reflective boundary, the following functions $\left\{sin(\frac{2k-1}{2}\frac{\pi x}{L})\right\}_{k\in\mathbb{Z}^+}$ and $\left\{cos(k\frac{\pi x}{L})\right\}_{k\in\mathbb{Z}^+}$ satisfy the reflective condition with 0 derivative at $x = -L$ and L. The eigenvalues of the Laplacian are indicated in Eq 4.4.9

$$\begin{aligned}
\nabla^2 sin(\frac{2k-1}{2}\frac{\pi x}{L}) &= -\left(\frac{2k-1}{2}\frac{\pi}{L}\right)^2 sin(\frac{2k-1}{2}\frac{\pi x}{L}) \\
\nabla^2 cos(k\frac{\pi x}{L}) &= -\left(\frac{k\pi}{L}\right)^2 cos(k\frac{\pi x}{L}) \\
\nabla^2 1 &= 0 \times 1
\end{aligned} \tag{4.4.9}$$

as

$$l_k = -\left(\frac{\pi k}{2L}\right)^2 \tag{4.4.10}$$

if the eigenfunctions $\left\{sin(\frac{2k-1}{2}\frac{\pi x}{L})\right\}_{k\in\mathbb{Z}^+}$ and $\left\{cos(k\frac{\pi x}{L})\right\}_{k\in\mathbb{Z}^+}$ are indexed the same way as in Eq 4.3.16.

Then, Eq 4.4.10 and Eq 4.4.8 solve the eigenvalue of the neutron transport equation (Eq 4.4.1) as

$$k_{eff,k} = \frac{1}{1+\left(\frac{\pi k M}{2L}\right)^2} \tag{4.4.11}$$

Comparing the eigenvalues of $M^{(1)}(x,x')$, λ_k in Eq 4.3.16 and the eigenvalues of the transport (diffusion) equation, $k_{eff,k}$ in Eq 4.4.11 and their corresponding eigenfunctions, it can be shown that the integral kernel $M^{(1)}(x,x')$ and the diffusion equation have the same eigensystem as long as

$$\begin{aligned} M^2 &= \frac{D}{\Sigma_a} = \frac{1}{\Sigma_a^2} \\ \Leftrightarrow D &= \frac{1}{\Sigma_a} \end{aligned} \tag{4.4.12}$$

Note that the relationship between diffusion coefficient D and macroscopic cross sections in Eq 4.4.12 only holds for this simple $1D$ problem.

The observation above leads to the assumption that the continuous transfer kernel $M^{(1)}(x,x')$ and the diffusion equation share the same eigensystem in the $3D$ problem as well. With the method of separation of variables,

$$\langle x,y,z|k_x,k_y,k_z\rangle = \langle x|k_x\rangle \langle y|k_y\rangle \langle z|k_z\rangle \tag{4.4.13}$$

the eigenvalues and eigenvectors are indexed by a vector $\vec{k}$ with three components for each dimension,

$$\vec{k} = (k_x,k_y,k_z) \tag{4.4.14}$$

And the eigenvalue $l_{\vec{k}}$ for the eigenfunction $\langle x,y,z|k_x,k_y,k_z\rangle$ is the sum of eigenvalues

for each dimension.

$$\begin{aligned}
&\nabla^2 \langle x,y,z|k_x,k_y,k_z\rangle = \nabla^2 \langle x|k_x\rangle \langle y|k_y\rangle \langle z|k_z\rangle \\
&= (\nabla_x^2+\nabla_y^2+\nabla_z^2) \langle x|k_x\rangle \langle y|k_y\rangle \langle z|k_z\rangle \\
&= (\nabla_x^2 \langle x|k_x\rangle) \langle y|k_y\rangle \langle z|k_z\rangle + \langle x|k_x\rangle (\nabla_y^2 \langle y|k_y\rangle) \langle z|k_z\rangle + \langle x|k_x\rangle \langle y|k_y\rangle (\nabla_z^2 \langle z|k_z\rangle) \\
&= (l_{k_x} \langle x|k_x\rangle) \langle y|k_y\rangle \langle z|k_z\rangle + \langle x|k_x\rangle (l_{k_y} \langle y|k_y\rangle) \langle z|k_z\rangle + \langle x|k_x\rangle \langle y|k_y\rangle (l_{k_z} \langle z|k_z\rangle) \\
&= (l_{k_x}+l_{k_y}+l_{k_z}) \langle x,y,z|k_x,k_y,k_z\rangle \qquad (4.4.15) \\
&\Rightarrow l_{\vec{k}} = \left(\frac{\pi k_x}{2L}\right)^2 + \left(\frac{\pi k_y}{2L}\right)^2 + \left(\frac{\pi k_z}{2L}\right)^2 \qquad (4.4.16)
\end{aligned}$$

The relation between k_{eff} and $l_{\vec{k}}$ (Eq 4.4.11) holds for any dimension.

$$\begin{aligned}
\lambda_{\vec{k}} &\equiv k_{keff,\vec{k}} \\
&= \frac{1}{1+\left(\frac{\pi k_x M}{2L_x}\right)^2 + \left(\frac{\pi k_y M}{2L_y}\right)^2 + \left(\frac{\pi k_z M}{2L_z}\right)^2} \qquad (4.4.17)
\end{aligned}$$

From now on, $\lambda_{\vec{k}}$ instead of k_{eff} are used for consistency in notation with the previous sections.

The expansion of the transfer kernel $M^{(1)}(x,x')$ can thus be approximated with the diffusion equation. What remains undetermined is the migration area M^2, which can often be approximated using classical models [15] or the more rigorous cumulative migration method [33].

If we assume that the migration area can be approximated and knowing the eigensystem from diffusion theory, $M^{(g)}(\vec{r},\vec{r}')$ can be conceptually written as

$$\begin{aligned}
&M^{(g)}(\vec{r},\vec{r}') = \frac{1}{8L_xL_yL_z} + \frac{1}{L_xL_yL_z}\times \\
&\sum_{\vec{k}\neq\vec{0}} \left\{ \left(\frac{1}{1+\left(\frac{k_xM\pi}{2L_x}\right)^2 + \left(\frac{k_yM\pi}{2L_y}\right)^2 + \left(\frac{k_zM\pi}{2L_z}\right)^2} \right)^g \langle x|k_x\rangle \langle y|k_y\rangle \langle z|k_z\rangle \langle k_x|x'\rangle \langle k_y|y'\rangle \langle k_z|z'\rangle \right\}
\end{aligned} \qquad (4.4.18)$$

For convenience $\sum_{\vec{k}\neq\vec{0}}$ is denoted as $\sum_{\vec{k}}^{'}$ from now on. With the approximated expansion series of the transfer kernel $M^{(1)}(\vec{r},\vec{r}')$, the derivations of spatial covariances, correlation coefficients and variance underestimation ratios are straightforward and in parallel with the development in section 4.3.

4.4.2 Spatial covariance

In the homogeneous problem with reflective boundaries, the eigenfunction of the largest eigenvalue 1 is

$$\mu(\vec{r}) = \frac{1}{8L_xL_yL_z} \tag{4.4.19}$$

With the stationary $\mu(\vec{r})$ being constant in space and the transition kernel $M^{(1)}(\vec{r},\vec{r}')$ being symmetric in $\vec{r}$ and $\vec{r}'$, $\not{V}(\vec{r},\vec{r}')$ (Eq 4.1.10) is simplified to

$$\not{V}(\vec{r},\vec{r}') = \frac{1}{8L_xL_yL_z}\left(\sigma^2\delta^{(3)}(\vec{r},\vec{r}') - M^{(2)}(\vec{r},\vec{r}')\right) \tag{4.4.20}$$

The continuous space version of the spatial correlation within generation n, $\mathrm{Cov}(n)_{i,j}$ (Eq 4.1.11) becomes

$$Cov^{(n)}(\vec{r},\vec{r}') = \frac{1}{8L_xL_yL_z}\sum_{g=n}^{1}\left\{\sigma^2 M^{(2(g-1))}(\vec{r},\vec{r}') - M^{(2g)}(\vec{r},\vec{r}')\right\} \tag{4.4.21}$$

where $M^{(0)}(\vec{r},\vec{r}')$ is defined as

$$M^{(0)}(\vec{r},\vec{r}') \equiv \delta^{(3)}(\vec{r},\vec{r}') \tag{4.4.22}$$

Similarly, the continuous space version of the serial-spatial moments becomes

$$Cov^{(n,n+d)}(\vec{r},\vec{r}') = \frac{1}{8L_xL_yL_z}\sum_{g=n}^{1}\left\{\sigma^2 M^{(2(g-1)+d)}(\vec{r},\vec{r}') - M^{(2g+d)}(\vec{r},\vec{r}')\right\} \tag{4.4.23}$$

Next, we choose a tally region I on which the correlation coefficients are to be predicted through $Cov^{(n)}(\vec{r},\vec{r}')$ and $Cov^{(n,n+d)}(\vec{r},\vec{r}')$. In the $1D$ problem, the sym-

metric tally region I was chosen as $[-a, a]$. Symmetry makes the contribution from all eigenfunctions that are odd function of x to $\mathrm{Cov}[X_I(n), X_I(n+d)]$ and $\mathrm{Var}[X_I(n)]$ vanish, which simplifies calculation.

In the $3D$ problem, to make the prediction more challenging and expose the impact of the dominance ratio, whose contribution vanishes due to symmetry of the central region, the tally region is moved off-center, I is chosen to be a box with volume $8a_x a_y a_z$ but with one corner at $(0, 0, 0)$.

$$I \equiv I_x \times I_y \times I_z = [0, 2a_x] \times [0, 2a_y] \times [0, 2a_z] \tag{4.4.24}$$

In the $3D$ problem, the whole space is denoted as W for convenience.

$$W \equiv W_x \times W_y \times W_z = [-L_x, L_x] \times [-L_y, L_y] \times [-L_z, L_z] \tag{4.4.25}$$

The first order moments in Eq 4.2.8 are simply

$$\begin{aligned} \mu(n) &= \mu(n+d) = 1 \\ \mu_I(n) &= \mu_I(n+d) = \frac{a_x}{L_x}\frac{a_y}{L_y}\frac{a_z}{L_z} \end{aligned} \tag{4.4.26}$$

The following integrals will be useful and hold for any dimension.

$$\int_{I_x} dx \, \langle x|k\rangle = \frac{2a_x}{\sqrt{2L_x}} \quad , k = 0 \tag{4.4.27}$$

$$\int_{I_x} dx \, \langle x|k\rangle = \frac{1}{\sqrt{L_x}} \cos\left(\frac{\pi k_x x}{2L_x}\right) = \frac{2\sqrt{L_x}}{k_x \pi} \sin\left(\frac{\pi k_x a_x}{L_x}\right) \quad , k \ even \tag{4.4.28}$$

$$\begin{aligned} \int_{I_x} dx \, \langle x|k\rangle &= \frac{1}{\sqrt{L_x}} \sin\left(\frac{\pi k_x x}{2L_x}\right) = \frac{2\sqrt{L_x}}{k_x \pi} \left(1 - \cos\left(\frac{\pi k_x a_x}{L_x}\right)\right) \\ &= \frac{4\sqrt{L_x}}{k_x \pi} \sin^2\left(\frac{\pi k_x a_x}{2L_x}\right) \qquad , k \ odd \end{aligned} \tag{4.4.29}$$

$$\int_{W_x} dx \, \langle x|k\rangle = \begin{cases} \sqrt{2L_x} & k = 0 \\ 0 & otherwise \end{cases} \tag{4.4.30}$$

Then we perform the integrals over region $I \times I$ of $M^{(g)}(\vec{r}, \vec{r}')$.

$$\int_I d^3\vec{r}\int_I d^3\vec{r}' M^{(g)}(\vec{r},\vec{r}') = \frac{1}{8L_xL_yL_z}\int_I d^3\vec{r}\int_I d^3\vec{r}'$$
$$+\sum_{\vec{k}}{}' \left(\lambda_{\vec{k}}\right)^g \int_I d^3\vec{r}\,\langle x,y,z|k_x,k_y,k_z\rangle \int_I d^3\vec{r}'\,\langle k_x,k_y,k_z|x',y',z'\rangle$$
$$= \frac{2a_x^2}{L_x}\frac{2a_y^2}{L_y}\frac{2a_z^2}{L_z} + \sum_{\vec{k}}{}' \left(\lambda_{\vec{k}}\right)^g \times$$
$$\int_{I_x} dx\,\langle x|k_x\rangle \int_{I_y} dy\,\langle y|k_y\rangle \int_{I_z} dz\,\langle z|k_z\rangle \int_{I_x} dx'\,\langle k_x|x'\rangle \int_{I_y} dy'\,\langle k_y|y'\rangle \int_{I_z} dz'\,\langle k_z|z'\rangle \tag{4.4.31}$$
$$= \frac{2a_x^2}{L_x}\frac{2a_y^2}{L_y}\frac{2a_z^2}{L_z} + \sum_{\vec{k}}{}' \lambda_{\vec{k}}^g f(\vec{k}) \tag{4.4.32}$$

where Eq 4.4.31 separates the integral over the tally region into a product of integrals along each dimension. Eq 4.4.32 evaluates the component of $\vec{k}=\vec{0}$ since the explicit form of the leading term is useful in derivations below. The final form of the remaining integrals depends on k of the corresponding dimension and are written in Eq 4.4.27, Eq 4.4.28, Eq 4.4.29 and Eq 4.4.30.

We then evaluate $\mathrm{Cov}[Z_I(n), Z_I(n+d)]$ below, first substitute the explicit form of kernel M into Eq 4.1.8 and Eq 4.1.10

$$\mathrm{Cov}[Z_I(n), Z_I(n+d)] =$$
$$\frac{1}{8L_xL_yL_z}\sum_{g=n}^{1}\left\{\sigma^2\int_I d^3\vec{r}\int_I d^3\vec{r}' M^{(2(g-1)+d)}(\vec{r},\vec{r}') - \int_I d^3\vec{r}\int_I d^3\vec{r}' M^{(2g+d)}(\vec{r},\vec{r}')\right\} \tag{4.4.33}$$

Then substitute the integral of the matrix elements of the transfer kernel, Eq 4.4.33

becomes Eq 4.4.34

$$\mathrm{Cov}[Z_I(n), Z_I(n+d)] = \frac{1}{8L_xL_yL_z}\left\{\sigma^2\sum_{g=n}^{1}\left(\frac{8a_x^2a_y^2a_z^2}{L_xL_yL_z} + \sum_{\vec{k}}{}' \lambda_{\vec{k}}^{2(g-1)+d} f(\vec{k})\right) - \sum_{g=n}^{1}\left(\frac{8a_x^2a_y^2a_z^2}{L_xL_yL_z} + \sum_{\vec{k}}{}' \lambda_{\vec{k}}^{2g+d} f(\vec{k})\right)\right\} \quad (4.4.34)$$

Then switch the order of the summation over the eigenfunction expansion and the summation over the generation number g and separate the leading $\vec{k} = \vec{0}$ term from the remaining $\sum_{\vec{k}}{}'$ summation terms, Eq 4.4.34 is simplified into Eq 4.4.35.

$$\begin{aligned}&\mathrm{Cov}[Z_I(n), Z_I(n+d)]\\ &= \frac{1}{8L_xL_yL_z}\left\{\frac{8a_x^2a_y^2a_z^2}{L_xL_yL_z}(\sigma^2-1)n + \sum_{\vec{k}}{}'\left(\sum_{g=n}^{1}\lambda_{\vec{k}}^{2(g-1)}\right)(\sigma^2-\lambda_{\vec{k}}^2)\lambda_{\vec{k}}^d f(\vec{k})\right\}\end{aligned} \quad (4.4.35)$$

Then evaluate the geometric series of index g (Eq 4.4.36), and take the limit $n \to \infty$ assuming that stationarity is reached after sufficiently many generations n and recognizes $\lambda_{\vec{k}}^{2n} \to 0$ (Eq 4.4.37).

$$\begin{aligned}&\mathrm{Cov}[Z_I(n), Z_I(n+d)]\\ &= \frac{1}{8L_xL_yL_z}\left\{\frac{8a_x^2a_y^2a_z^2}{L_xL_yL_z}(\sigma^2-1)n + \sum_{\vec{k}}{}'\frac{1-\lambda_{\vec{k}}^{2n}}{1-\lambda_{\vec{k}}^2}(\sigma^2-\lambda_{\vec{k}}^2)\lambda_{\vec{k}}^d f(\vec{k})\right\} \quad (4.4.36)\\ &= \frac{1}{8L_xL_yL_z}\left\{\frac{8a_x^2a_y^2a_z^2}{L_xL_yL_z}(\sigma^2-1)n + \sum_{\vec{k}}{}'\frac{\sigma^2-\lambda_{\vec{k}}^2}{1-\lambda_{\vec{k}}^2}\lambda_{\vec{k}}^d f(\vec{k})\right\} \quad (4.4.37)\end{aligned}$$

Next, we can evaluate $\mathrm{Cov}[Z(n), Z_I(n+d)]$. The $\langle x, y, z|k_x, k_y, k_z\rangle$ part of $M^{(g)}(\vec{r}, \vec{r}')$ for all components will be integrated over the whole region W and the $\langle k_x, k_y, k_z|x', y', z'\rangle$ part will be integrated over the region of interest I. According to Eq 4.4.30, only the

leading $\lambda_{\vec{k}} = \lambda_{\vec{0}}$ term contributes non-zero value to the integral $\int_I d^3\vec{r} \int_I d^3\vec{r}'$.

$$\begin{aligned}
&\mathrm{Cov}[Z(n), Z_I(n+d)] \\
&= \frac{1}{8L_xL_yL_z}\left\{\sigma^2\left(\prod_{i=x,y,z}\frac{2L_i}{\sqrt{2L_i}}\frac{2a_i}{\sqrt{2L_i}} + 0\right) - \left(\prod_{i=x,y,z}\frac{2L_i}{\sqrt{2L_i}}\frac{2a_i}{\sqrt{2L_i}} + 0\right)\right\} \quad (4.4.38)\\
&= \frac{a_xa_ya_z}{L_xL_yL_z}(\sigma^2 - 1)n \quad (4.4.39)
\end{aligned}$$

By symmetry,

$$\mathrm{Cov}[Z_I(n), Z(n+d)] = \frac{a_xa_ya_z}{L_xL_yL_z}(\sigma^2 - 1)n \quad (4.4.40)$$

Similarly, only the $\vec{k} = \vec{0}$ term in the expansion of $M^{(g)}(\vec{r}, \vec{r}')$ contribute to the integral of $\mathrm{Cov}[Z(n), Z(n+d)]$.

$$\begin{aligned}
\mathrm{Cov}[Z(n), Z(n+d)] = \frac{1}{8L_xL_yL_z}&\left\{\sigma^2\left(\prod_{i=x,y,z}\frac{2L_i}{\sqrt{2L_i}}\frac{2L_i}{\sqrt{2L_i}} + 0\right) - \right.\\
&\left.\left(\prod_{i=x,y,z}\frac{2L_i}{\sqrt{2L_i}}\frac{2L_i}{\sqrt{2L_i}} + 0\right)\right\} \quad (4.4.41)\\
= (\sigma^2 - 1)n& \quad (4.4.42)
\end{aligned}$$

where Eq 4.4.41 is the same as Eq 4.4.38 except that a_i's are replaced with corresponding L_i's.

All the terms required to calculate $\mathrm{Cov}[X_I(n), X_I(n+d)]$ in Eq 4.2.8 are known (Eq 4.4.26, Eq 4.4.37, Eq 4.4.39, Eq 4.4.40 and Eq 4.4.42). The covariance between

the normalized neutron tally in region I across the generation lag d is

$$
\begin{aligned}
&Cov[X_I(n), X_I(n+d)] \\
&= \left(\frac{a_x a_y a_z}{L_x L_y L_z}\right)^2 \left[\frac{\left(\frac{a_x a_y a_z}{L_x L_y L_z}\right)^2 (\sigma^2-1)n}{\frac{a_x a_y a_z}{L_x L_y L_z}\frac{a_x a_y a_z}{L_x L_y L_z}} - 2\frac{\frac{a_x a_y a_z}{L_x L_y L_z}(\sigma^2-1)n}{1\frac{a_x a_y a_z}{L_x L_y L_z}} + \frac{(\sigma^2-1)n}{11}\right] \\
&\quad + \frac{1}{8L_x L_y L_z}{\sum_{\vec{k}}}' \frac{\sigma^2-\lambda_{\vec{k}}^2}{1-\lambda_{\vec{k}}^2}\lambda_{\vec{k}}^d f(\vec{k}) \qquad (4.4.43) \\
&= \frac{1}{8L_x L_y L_z}{\sum_{\vec{k}}}' \frac{\sigma^2-\lambda_{\vec{k}}^2}{1-\lambda_{\vec{k}}^2}\lambda_{\vec{k}}^d f(\vec{k}) \qquad (4.4.44)
\end{aligned}
$$

where Eq 4.4.43 combines the 0^{th} term in $\text{Cov}[Z_I(n), Z_I(n+d)]$ (Eq 4.4.37) with the other covariance terms and cancels them out. Note that Eq 4.4.44 is a complete analog to Eq 4.3.42. The covariances of X tallies decay exponentially as a function of d.

4.4.3 Correlation coefficients

Similar calculations are performed for $\text{Var}[X_I(n)]$ to obtain $\rho_I(d)$ as $\text{Cov}[X_I(n), X_I(n+d)]/\text{Var}[X_I(n)]$. $\text{Var}[X_I(n)]$ is evaluated via $\text{Cov}[X_I(n), X_I(n+d)]$ with $d=0$. Thus $\text{Cov}[Z_I(n), Z_I(n)]$, $\text{Cov}[Z_I(n), Z(n)]$ and $\text{Cov}[Z(n), Z(n)]$ are needed. One difference from the covariance terms evaluated in the previous section, $d=0$ leads to the appearance of $M^{(0)}(\vec{r}, \vec{r}') = \delta(\vec{r}, \vec{r}')$ in the summation indexed by the generation number g. The following integrals of $\delta^{(3)}(\vec{r}, \vec{r}')$ are will be useful in upcoming evaluations.

$$
\begin{aligned}
\int_I d^3\vec{r}\int_I d^3\vec{r}'\,\delta(\vec{r},\vec{r}') &= \int_I d^3\vec{r} = 8a_x a_y a_z \qquad (4.4.45)\\
\int_I d^3\vec{r}\int_W d^3\vec{r}'\,\delta(\vec{r},\vec{r}') &= \int_I d^3\vec{r} = 8a_x a_y a_z \qquad (4.4.46)\\
\int_W d^3\vec{r}\int_I d^3\vec{r}'\,\delta(\vec{r},\vec{r}') &= \int_W d^3\vec{r}\,\mathbb{1}_{\vec{r}\in I} = 8a_x a_y a_z \qquad (4.4.47)\\
\int_W d^3\vec{r}\int_W d^3\vec{r}'\,\delta(\vec{r},\vec{r}') &= 8L_x L_y L_z \qquad (4.4.48)
\end{aligned}
$$

While integrating $\mathrm{Cov}^{(n)}(x,y)$ (Eq 4.3.23), $g=1$ and $g\geq 2$ should be treated separately.

$$\begin{aligned}
&\mathrm{Cov}[Z_I(n),Z_I(n)] = \\
&\frac{1}{8L_xL_yL_z}\left\{\sigma^2 8a_xa_ya_z+\sigma^2\sum_{g=n}^{2}\left[\left(\frac{2a_x}{\sqrt{2L_x}}\frac{2a_y}{\sqrt{2L_y}}\frac{2a_z}{\sqrt{2L_z}}\right)^2+{\sum_{\vec{k}}}'\lambda_{\vec{k}}^{2(g-1)}f(\vec{k})\right]\right. \\
&\qquad\left.-\sum_{g=n}^{1}\left[\left(\frac{2a_x}{\sqrt{2L_x}}\frac{2a_y}{\sqrt{2L_y}}\frac{2a_z}{\sqrt{2L_z}}\right)^2+{\sum_{\vec{k}}}'\lambda_{\vec{k}}^{2a}f(\vec{k})\right]\right\} \qquad (4.4.49)\\
&=\frac{1}{8L_xL_yL_z}\left[\sigma^2 8a_xa_ya_z-\left(\frac{8a_x^2a_y^2a_z^2}{L_xL_yL_z}+{\sum_{\vec{k}}}'\lambda_{\vec{k}}^2f(\vec{k})\right)\right] \\
&+\frac{1}{8L_xL_yL_z}\sum_{g=n}^{2}\left\{(\sigma^2-1)\frac{8a_x^2a_y^2a_z^2}{L_xL_yL_z}+{\sum_{\vec{k}}}'\lambda_{\vec{k}}^{2(g-1)}f(\vec{k})\left(\sigma^2-\lambda_{\vec{k}}^2\right)\right\} \qquad (4.4.50)\\
&=\frac{a_xa_ya_z}{L_xL_yL_z}\sigma^2-\left(\frac{a_xa_ya_z}{L_xL_yL_z}\right)^2+\left(\frac{a_xa_ya_z}{L_xL_yL_z}\right)^2(\sigma^2-1)(n-1) \\
&\qquad+\frac{1}{8L_xL_yL_z}\left[-{\sum_{\vec{k}}}'\lambda_{\vec{k}}^2f(\vec{k})+{\sum_{\vec{k}}}'\frac{\sigma^2-\lambda_{\vec{k}}^2}{1-\lambda_{\vec{k}}^2}\lambda_{\vec{k}}^2f(\vec{k})\right] \qquad (4.4.51)\\
&=\frac{a_xa_ya_z}{L_xL_yL_z}\sigma^2-\left(\frac{a_xa_ya_z}{L_xL_yL_z}\right)^2\sigma^2+\left(\frac{a_xa_ya_z}{L_xL_yL_z}\right)^2(\sigma^2-1)n \\
&+\frac{1}{8L_xL_yL_z}{\sum_{\vec{k}}}'\frac{\sigma^2-1}{1-\lambda_{\vec{k}}^2}\lambda_{\vec{k}}^2f(\vec{k}) \qquad (4.4.52)
\end{aligned}$$

where Eq 4.4.49 separates the summation over index g into $g=1$ and $g\geq 2$ for the $\sigma^2\,\mathrm{Cov}^{(2(g-1))}[\vec{r},\vec{r}']$ part. Eq 4.4.50 combines the $g=1$ and $g\geq 2$ terms. Eq 4.4.51 switches the order of the summation over $\vec{k}$ and the summation over g, and evaluates the geometric series of index g with $n\to\infty$. Eq 4.4.52 recognizes common factors in the two summations over index $\vec{k}$.

Next, we evaluate $\mathrm{Cov}[Z(n),Z_I(n)]$ which shows similarities to $\mathrm{Cov}[Z(n),Z_I(n+$

d)].

$$\begin{aligned}\mathrm{Cov}[Z(n), Z_I(n)] &= \frac{1}{8L_xL_yL_z}\left\{\sigma^2\left(8a_xa_ya_z + \sum_{g=n}^{2}\left(\prod_{i=x,y,z}\frac{2L_i}{\sqrt{2L_i}}\frac{2a_i}{\sqrt{2L_i}} + 0\right)\right)\right. \\ &\left. - \sum_{g=n}^{1}\left(\prod_{i=x,y,z}\frac{2L_i}{\sqrt{2L_i}}\frac{2a_i}{\sqrt{2L_i}} + 0\right)\right\} \\ &= \frac{a_xa_ya_z}{L_xL_yL_z}(\sigma^2-1)n \end{aligned} \tag{4.4.53}$$

By symmetry,

$$\mathrm{Cov}[Z_I(n), Z(n)] = \frac{a_xa_ya_z}{L_xL_yL_z}(\sigma^2-1)n \tag{4.4.54}$$

Similarly,

$$\begin{aligned}\mathrm{Cov}[Z(n), Z(n)] &= \frac{1}{8L_xL_yL_z}\left\{\sigma^2\left(8L_xL_yL_z + \sum_{g=n}^{2}\left(\prod_{i=x,y,z}\frac{2L_i}{\sqrt{2L_i}}\frac{2L_i}{\sqrt{2L_i}} + 0\right)\right)\right. \\ &\left. - \sum_{g=n}^{1}\left(\prod_{i=x,y,z}\frac{2L_i}{\sqrt{2L_i}}\frac{2L_i}{\sqrt{2L_i}} + 0\right)\right\} \\ &= (\sigma^2-1)n \end{aligned} \tag{4.4.55}$$

Then we combine the covariance of the neutron counts Z to the variance of relative counts X according to Eq 4.2.8.

$$\begin{aligned} Var[X_I(n)] &= \mathrm{Cov}[X_I(n), X_I(n)] \\ &= \left(\frac{a_xa_ya_z}{L_xL_yL_z}\right)^2\left[\frac{\left(\frac{a_xa_ya_z}{L_xL_yL_z}\right)^2(\sigma^2-1)n + \sigma^2\frac{a_xa_ya_z}{L_xL_yL_z}\left(1-\frac{a_xa_ya_z}{L_xL_yL_z}\right)}{\frac{a_xa_ya_z}{L_xL_yL_z}\frac{a_xa_ya_z}{L_xL_yL_z}} - 2\frac{\frac{a_xa_ya_z}{L_xL_yL_z}(\sigma^2-1)n}{\frac{a_xa_ya_z}{L_xL_yL_z}1} + \frac{(\sigma^2-1)n}{11}\right] \\ &+ \frac{1}{8L_xL_yL_z}\sum_{\vec{k}}{}'\frac{\sigma^2-1}{1-\lambda_{\vec{k}}^2}\lambda_{\vec{k}}^2 f(\vec{k}) \end{aligned} \tag{4.4.56}$$

$$= \sigma^2\frac{a_xa_ya_z}{L_xL_yL_z}\left(1-\frac{a_xa_ya_z}{L_xL_yL_z}\right) + \frac{1}{8L_xL_yL_z}\sum_{\vec{k}}{}'\frac{\sigma^2-1}{1-\lambda_{\vec{k}}^2}\lambda_{\vec{k}}^2 f(\vec{k}) \tag{4.4.57}$$

where Eq 4.4.56 combines the 0^{th} term in $\mathrm{Cov}[Z_I(n), Z_I(n)]$ (Eq 4.4.52) with the other variance terms and cancels them out. Note that Eq 4.4.57 is a complete analog to Eq 4.3.54, with the volume ratio $\frac{a}{L}$ replaced by $\frac{a_x a_y a_z}{L_x L_y L_z}$ and the absence of n in the final form indicates that Eq 4.4.57 is stationary.

With the covariances that are a function of the generation lag only and variances that are stationary, the correlation coefficients of the normalized tally in region I are a function of generation lag d only.

$$\begin{aligned}\rho_I(d) &= \frac{\mathrm{Cov}[X_I(n), X_I(n+d)]}{\mathrm{Cov}[X_I(n), X_I(n)]} \\ &= \frac{\sum_{\vec{k}}' \frac{\sigma^2-\lambda_{\vec{k}}^2}{1-\lambda_{\vec{k}}^2} \lambda_{\vec{k}}^d f(\vec{k})}{8\sigma^2 a_x a_y a_z \left(1 - \frac{a_x a_y a_z}{L_x L_y L_z}\right) + \sum_{\vec{k}}' \frac{\sigma^2-1}{1-\lambda_{\vec{k}}^2} \lambda_{\vec{k}}^2 f(\vec{k})}\end{aligned} \tag{4.4.58}$$

Eq 4.4.58 also exhibits two important features of the correlation coefficients. Firstly, the dependence on n in the moments of the Z tallies disappeared after being transformed to X tallies and leads to the correlation coefficients being a function of the generation lag only. Secondly, the dependence on the eigenvalues and consequentially the asymptotic dependence on the dominance ratio corresponds to the correlation due to the source dependence during the source update in consecutive generations and the dependence on σ^2 corresponds to the correlation due to the branching processes. Thirdly, the correlation coefficients do not depend on generation size.

4.4.4 Variance underestimation ratio and asymptotic behavior

Covariances of normalized X tallies can now be used to predict the variance of a tally averaged over the active generations, $\mathrm{Var}\left[\bar{X}(N)\right]$, according to Eq 2.2.5. Equivalently, the correlation coefficients can be converted to a variance estimation ratio

$r(N)$ as defined in Eq 2.2.8.

$$r_I(N) = 1 + 2\sum_{d=1}^{N-1}\rho_I(d) - \frac{2}{N}\sum_{d=1}^{N-1} d\ \rho_I(d) \tag{4.4.59}$$

$$= 1 + 2\frac{\frac{1}{8L_xL_yL_z}\sum_{\vec{k}}{}'\left(\sum_{d=1}^{N-1}\lambda_{\vec{k}}^d - \frac{1}{N}\sum_{d=1}^{N-1} d\ \lambda_{\vec{k}}^d\right)\frac{\sigma^2-\lambda_{\vec{k}}^2}{1-\lambda_{\vec{k}}^2}f(\vec{k})}{\mathrm{Var}[X_I]} \tag{4.4.60}$$

$$= 1 + 2\frac{\frac{1}{8L_xL_yL_z}\sum_{\vec{k}}{}'\left(\frac{\lambda_{\vec{k}}(1-\lambda_{\vec{k}}^{N-1})}{1-\lambda_{\vec{k}}} - \frac{1}{N}\frac{\lambda_{\vec{k}}(1-\lambda_{\vec{k}}^{N-1})-(N-1)\lambda_{\vec{k}}^N(1-\lambda_{\vec{k}})}{(1-\lambda_{\vec{k}})^2}\right)\frac{\sigma^2-\lambda_{\vec{k}}^2}{1-\lambda_{\vec{k}}^2}f(\vec{k})}{\mathrm{Var}[X_I]} \tag{4.4.61}$$

$$= 1 + 2\frac{\sum_{\vec{k}}{}'\left(\frac{\lambda_{\vec{k}}(1-\lambda_{\vec{k}}^{N-1})}{1-\lambda_{\vec{k}}} - \frac{1}{N}\frac{\lambda_{\vec{k}}(1-\lambda_{\vec{k}}^{N-1})-(N-1)\lambda_{\vec{k}}^N(1-\lambda_{\vec{k}})}{(1-\lambda_{\vec{k}})^2}\right)\frac{\sigma^2-\lambda_{\vec{k}}^2}{1-\lambda_{\vec{k}}^2}f(\vec{k})}{\sigma^2 v_I\left(1-\frac{v_I}{v}\right) + \sum_{\vec{k}}{}'\frac{\sigma^2-1}{1-\lambda_{\vec{k}}^2}\lambda_{\vec{k}}^2 f(\vec{k})} \tag{4.4.62}$$

where Eq 4.4.59 separates the summation in Eq 2.2.8 over generation lag d into two parts that depends on d in different ways. Eq 4.4.60 inserts the explicit form of $\rho_I(d)$ and changes the order of summation over the eigenfunction expansion and generation lag, and separates the summands that depend on the generation lag. Eq 4.4.61 evaluates the summation over the generation lag explicitly. $\mathrm{Var}[X_I]$ was written in Eq 4.4.57 as $\mathrm{Var}[X_I(n)]$ and its generation independence is emphasized in Eq 4.4.62. Eq 4.4.62 inserts the explicit form of $\mathrm{Var}[X_I]$ and recognize volume of tally region v_I and volume of the system v.

Finally, the asymptotic variance underestimation ratio can be obtained by taking the limit of $N \to \infty$ in $r_I(N)$ (Eq 4.4.62).

$$\begin{aligned} r_I &\equiv \lim_{N\to\infty} r_I(N) \\ &= 1 + 2\frac{\sum_{\vec{k}}{}'\frac{\lambda_{\vec{k}}}{1-\lambda_{\vec{k}}}\frac{\sigma^2-\lambda_{\vec{k}}^2}{1-\lambda_{\vec{k}}^2}f(\vec{k})}{\sigma^2 v_I\left(1-\frac{v_I}{v}\right) + \sum_{\vec{k}}{}'\frac{\sigma^2-1}{1-\lambda_{\vec{k}}^2}\lambda_{\vec{k}}^2 f(\vec{k})} \end{aligned} \tag{4.4.63}$$

The derivation of correlation coefficients $\rho_I(d)$, variance underestimation ratio

$r_I(N)$ and its asymptotic value r_I can be applied to many types of tallies by adjusting the values of a_x, a_y, a_z. For example, setting $a_z = L$ and $a_x = a_y = \delta$, the tally region I corresponds approximately to a nuclear fuel pin with cross section area δ^2. And setting $a_z = a_y = L$ and $a_x = \delta$, the tally region I corresponds to a slab of width δ. Note that, $f(\vec{k})$ depends on the tally region choice through the integrals in Eq 4.4.27,Eq 4.4.28 and Eq 4.4.29.

4.4.5 Numerical results

Simulations of the test problem $TC1$ (section 2.1.2) are used to verify the previous derivations numerically. $TC1$ corresponds to $L_x = L_y = L_z = 200cm$. The tally regions of interest are selected to be cubes of different sizes parameterized by $(2a)^3$. $1,000,000$ neutrons per generation are simulated in the $3D$ system for 1000 active generations, and 200 such independent simulations are performed to generate reference results.

For the predictive mode, M^2 is approximated as $1/(3\Sigma_t\Sigma_a)$ since the diffusion coefficient is calculated assuming isotropic scattering in the LAB reference system

$$D = \frac{1}{3\Sigma_t} \tag{4.4.64}$$

The summation over $\sum_{\vec{k}}^{'}$ in all of the eigenmode expansion equations such as Eq 4.4.58 and Eq 4.4.62 can be written explicitly by separating the fundamental eigenmode, as well as the even and odd eigenmodes.

The correlation coefficients estimated from the independent simulation tallies are plotted along with the prediction from $\rho_I(d)$ in Eq 4.4.58 in Figure 4-6.

The reference autocorrelation coefficients at each state are calculated according to Eq 2.2.7 by simply replacing the expectation operators with averages calculated over the independent simulations. Due to the stationarity of correlation coefficients discussed previously and numerically verified in section 2.3.1, all $\rho(n, n+d)$ of the same lag are combined to reduce statistical noise.

As expected, the correlation coefficients are higher in larger tally regions. Addi-

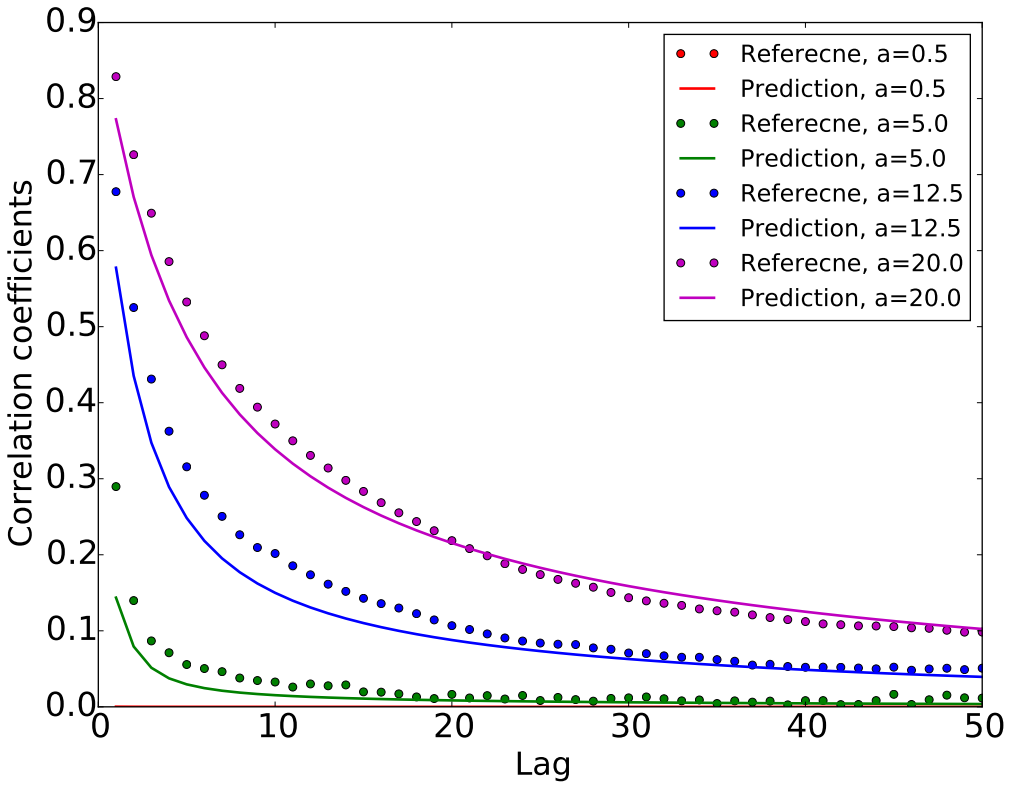


Figure 4-6: Correlation coefficients for central tally region with different sizes in $TC1$. The dots correspond to correlation coefficients estimated from independent simulations. The solid line curves correspond to correlation coefficients predicted from the continuous model.

tionally, it can be seen that when using the diffusion approximation, the predicted $\rho_I(d)$ follows the correct trend of that calculated from independent simulations but the prediction is not exact as previously observed with the $1D$ problem in section 4.3 (Figure 4-1).

The numerical reference of the variance underestimation ratio from the independent simulations and the predicted values are plotted in Figure 4-7. As expected, the underestimation ratio is larger as the tally size increases. When $a = 0.5cm$, the volume of the tally region is only 1.5625×10^{-8} of the total system volume. Even with 1000000 neutrons per generation simulated, for problem $TC1$, the expected number of neutron observed per generation in the tally region is 0.015625 or 1 neutron every 64 generations. This naturally implies that this small tally region will suffer from large statistical noise as observed in Figure 4-7.

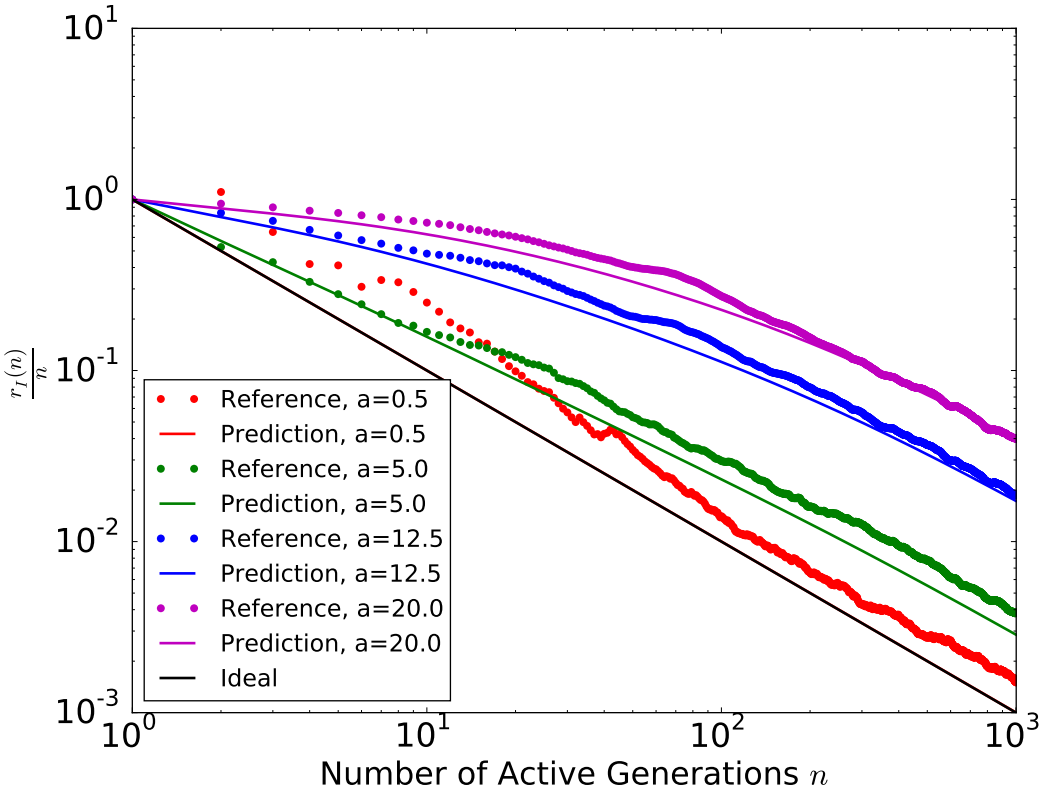


Figure 4-7: Variance convergence rate calculated from underestimation ratio for central tally region with different sizes in TC1. The dots correspond to $\frac{r_I(n)}{n}$ estimated from independent simulations. The solid line curves correspond to predictions from the continuous model.

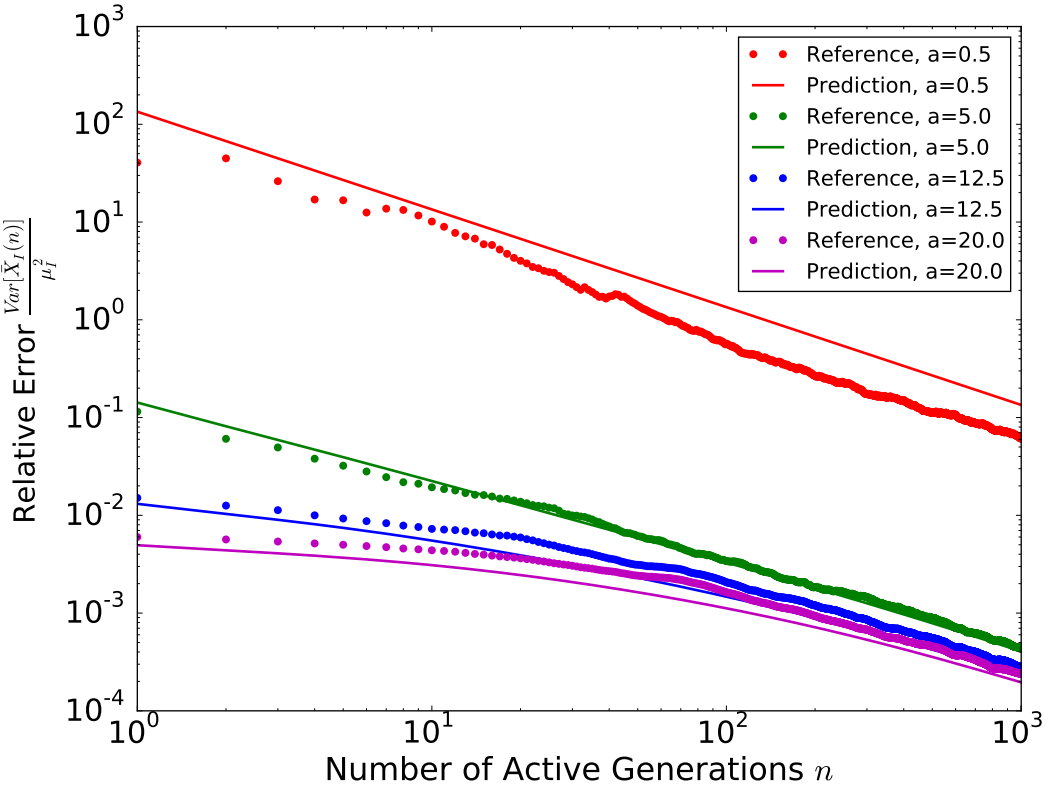


Figure 4-8: Relative square error for central tally region with different sizes in $TC1$. The dots correspond to RSE_I estimated from independent simulations. The solid line curves correspond to predictions from the continuous model.

Similarly to the discussion related to the $TC1-1D$ problem, the real variance of the normalized tally averaged over all active generations can also be extracted from the derivation of the variance underestimation ratio. The real variance can be used to predict the relative square error as expressed below

$$\mathbb{E}RSE_I(N) = \frac{\mathrm{Var}[\bar{X}_I(N)]}{s\mu_I^2} = \frac{\mathrm{Var}[X_I]}{s} r_I(N) \left(\frac{L}{a}\right)^6 \tag{4.4.65}$$

where s is the number of neutrons per generation.

The $\mathbb{E}RSE_I$ derived above in Eq 4.4.65 is plotted along with the relative square error obtained from independent simulations in Figure 4-8.

Although smaller regions have lower correlation coefficients and thus their variance decreases closer to $1/N$, the relatively lower tally counts in the region lead to larger

uncertainties.

Further, it can be seen from Figure 4-8 that the relative square error monotonically decreases as tally region size increases. By the law of large numbers, the relative square error approaches its prediction, real variance divided by squared reference, which can be interpreted as the inverse of squared signal to noise ratio (SNR). SNR is higher for larger tally regions which explains the lower relative square error observed.

Quantitatively, it can be shown for the homogeneous problem that the real variance divided by the squared reference value in tally region i cannot be lower than the larger tally region I comprised of smaller regions which includes i.

$$\begin{aligned}
\mathrm{Var}[\bar{X}_I(N)] &= \sum_{n,n'\leq N} \mathrm{Cov}[X_I(n), X_I(n')] \\
&= \sum_{n,n'\leq N} \sum_{i,j\in I} \mathrm{Cov}[X_i(n), X_j(n')] \\
&= \sum_{i\in I} \sum_{n,n'\leq N} \mathrm{Cov}[X_i(n), X_i(n')] + \sum_{n,n'\leq N} \sum_{i\neq j\in I} \mathrm{Cov}[X_i(n), X_j(n')] \\
&= \sum_{i\in I} \mathrm{Var}[\bar{X}_i(N)] + \sum_{i\neq j\in I} \sum_{n,n'\leq N} \mathrm{Cov}[X_i(n), X_j(n')]
\end{aligned} \tag{4.4.66}$$

For $i \neq j$, $\mathrm{Cov}[X_i(n), X_j(n')] < \mathrm{Cov}[X_i(n), X_i(n')]$ since a tally in region i is more correlated with itself than with other regions.

$$\begin{aligned}
\mathrm{Var}[\bar{X}_I(N)] &< \sum_{i\in I} \mathrm{Var}[\bar{X}_i(N)] + \sum_{i\neq j\in I} \sum_{n,n'\leq N} \mathrm{Cov}[X_i(n), X_i(n')] \\
&= \sum_{i\in I} \mathrm{Var}[\bar{X}_i(N)] + \sum_{i\neq j\in I} \mathrm{Var}[\bar{X}_i(N)]
\end{aligned} \tag{4.4.67}$$

For a homogeneous problem, $\mathrm{Var}[X_i(n)] = \mathrm{Var}[X_j(n)]$, then $\mathrm{Var}[\bar{X}_i(N)] = \mathrm{Var}[\bar{X}_j(N)]$. Therefore,

$$\begin{aligned}
\mathrm{Var}[\bar{X}_I(N)] &< I\,\mathrm{Var}[\bar{X}_i(N)] + I(I-1)\,\mathrm{Var}[\bar{X}_i(N)] \\
&= I^2\,\mathrm{Var}[\bar{X}_i(N)]
\end{aligned} \tag{4.4.68}$$

Since the system is reflective and homogeneous, $\mathbb{E}[X_i(n)] = \mathbb{E}[X_j(n)]$, then $\mathbb{E}[\bar{X}_i(N)] = \mathbb{E}[\bar{X}_j(N)]$ and $\mathbb{E}[\bar{X}_I(N)] = I\mathbb{E}[\bar{X}_i(N)]$. Finally,

$$\begin{aligned} \frac{\mathrm{Var}[\bar{X}_I(N)]}{\left(\mathbb{E}\bar{X}_I(N)\right)^2} &< \frac{I^2\,\mathrm{Var}[\bar{X}_i(N)]}{\left(\mathbb{E}\bar{X}_I(N)\right)^2} \\ &= \frac{\mathrm{Var}[\bar{X}_i(N)]}{\left(\mathbb{E}\bar{X}_i(N)\right)^2} \end{aligned} \tag{4.4.69}$$

which demonstrates that $RMS_I < RMS_i$.

The asymptotic value r_I for different tally sizes are plotted explicitly as a function of the tally size in Figure 4-9. The numerical reference values are approximated by using the value at $r_I(1000)$. Since the variance underestimation ratio measures the cumulative effect of correlation coefficients, the diffusion approximation can predict the asymptotic variance underestimation ratio with acceptable accuracy despite somewhat inaccurate correlation coefficients at each generation. Figure 4-9 also shows that using the lowest truncation of the kernel expansion (single higher mode) gives unsatisfactory predictions of the underestimation ratio. However, as long as the homogenization is performed appropriately, the evaluation of the eigenfunctions of the homogeneous diffusion problem is not very costly.

4.5 Application to the 2D BEAVRS benchmark

Section 4.3 solves the correlation issue exactly in a homogeneous $1D$ system, while section 4.4 shows that the correlations can be reasonably estimated using diffusion theory in a homogeneous $3D$ system. It was shown that a good indicator in determining situations where correlations can become an issue is the asymptotic variance underestimation ratio r_I.

This section will demonstrate how the r_I of a realistic problem can be used to identify tally sizes where the usual variance estimator will fail. The 2D BEAVRS benchmark was previously described in section 3.3.5 which corresponds to a mid-plane axial slice of the 3D BEAVRS benchmark. In the notation of this chapter, the

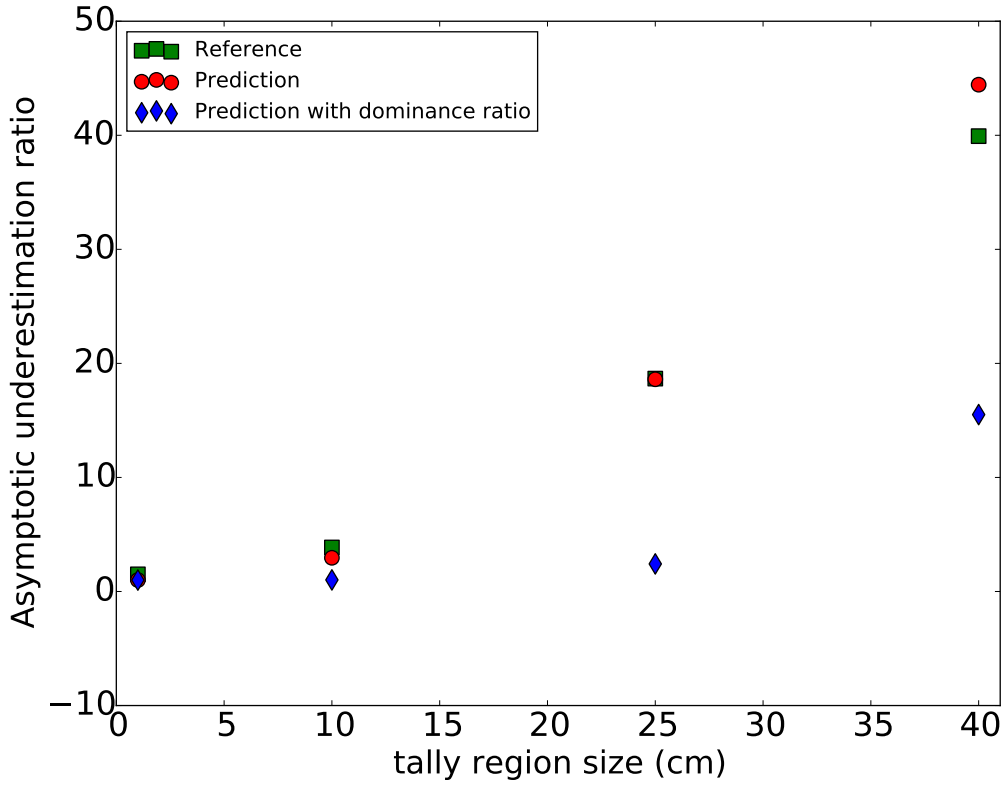


Figure 4-9: Asymptotic variance underestimation ratio for central tally region with different sizes in TC1. The squares correspond to r_I estimated from independent simulations. The dots correspond to predictions from the continuous model with sufficient expansion terms. The diamonds correspond to predictions from the continuous model with the first expansion term.

geometry parameters of the problem are

$$\begin{aligned} L_x &= L_y = 182.78cm \\ L_z &= 5.0cm \end{aligned} \tag{4.5.1}$$

The cumulative migration area method [33] was used to estimate the migration area of the core as

$$M^2 = 55.06cm^2 \tag{4.5.2}$$

The assemblies including the inter-assembly gaps are of size $21.5cm \times 21.5cm$ which in the notation of this chapter corresponds to $a_x = a_y = 10.75cm$ and $a_z = L_z = 5.0cm$. Correlations were investigated on four different tally sizes and references values were obtained from 450 independent simulations.

The reference and predicted correlation coefficients are plotted in Figure 4-10. As expected, the diffusion approximation does not match exactly the reference values, but does follow similar trends for all tally sizes.

Consequentially, a similar behavior can be observed for the variance underestimation ratio $r_I(n)$ as shown in Figure 4-11.

The asymptotic value r_I for different tally sizes $[0, 2a_x] \times [0, 2a_x] \times [-L_z, L_z]$ is illustrated in Figure 4-12. The reference values correspond to the value of r_I after 1000 generations. While not exact, especially for larger tallies, the diffusion approximation follows the appropriate trend and can provide a very simple way of deciding if independent simulations are needed or an alternative strategy to reduce or estimate the correlations should be used.

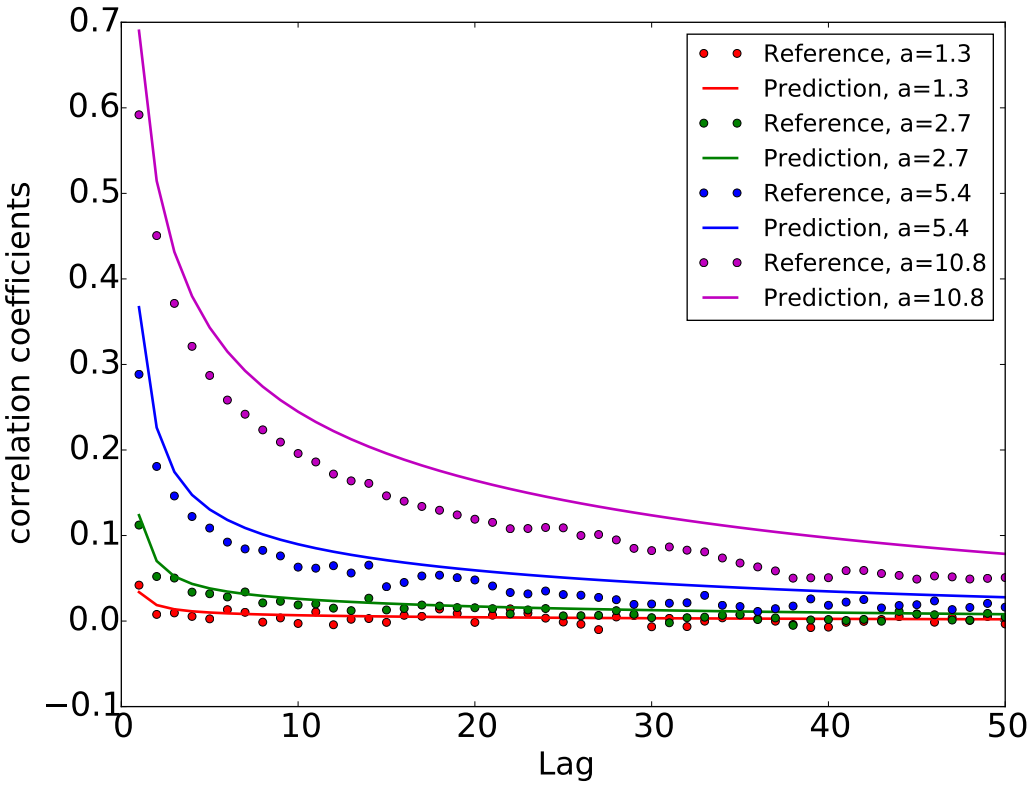


Figure 4-10: Correlation coefficients for central tally region with different sizes in 2D BEAVRS. The dots correspond to correlation coefficients estimated from independent simulations. The solid line curves correspond to correlation coefficients predicted from the continuous model.

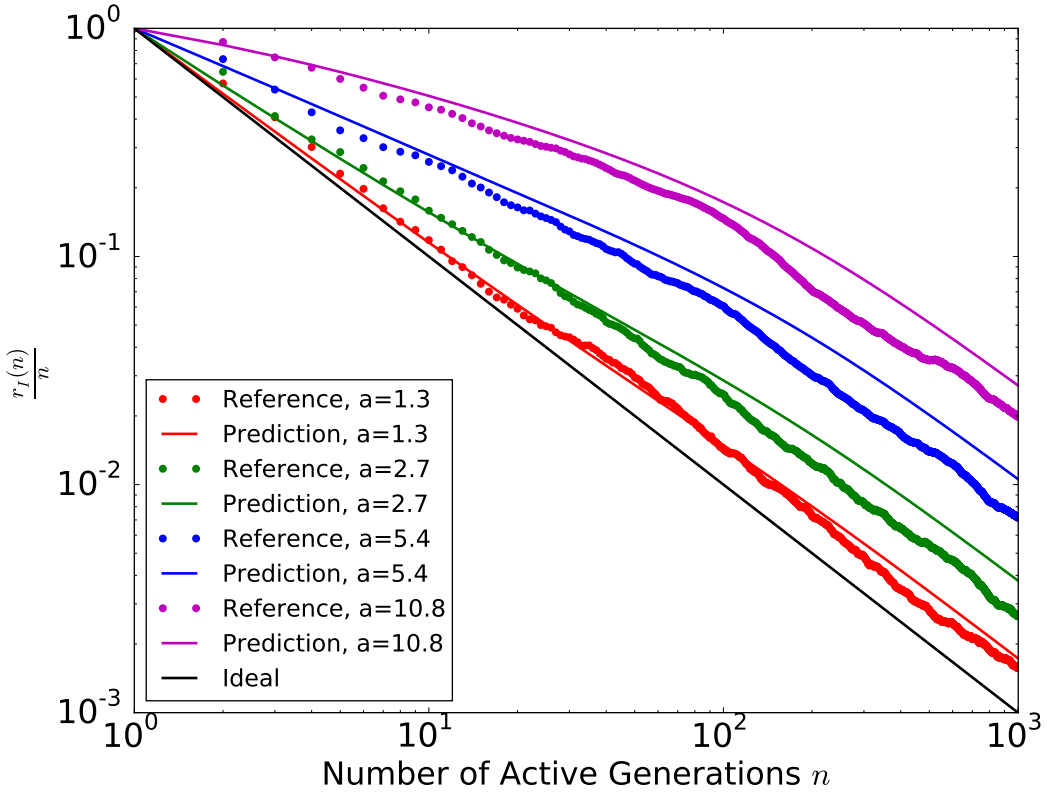


Figure 4-11: Variance convergence rate for central tally region with different sizes in 2D BEAVRS. The dots correspond to $\frac{r_I(n)}{n}$ estimated from independent simulations. The solid line curves correspond to predictions from the continuous model.

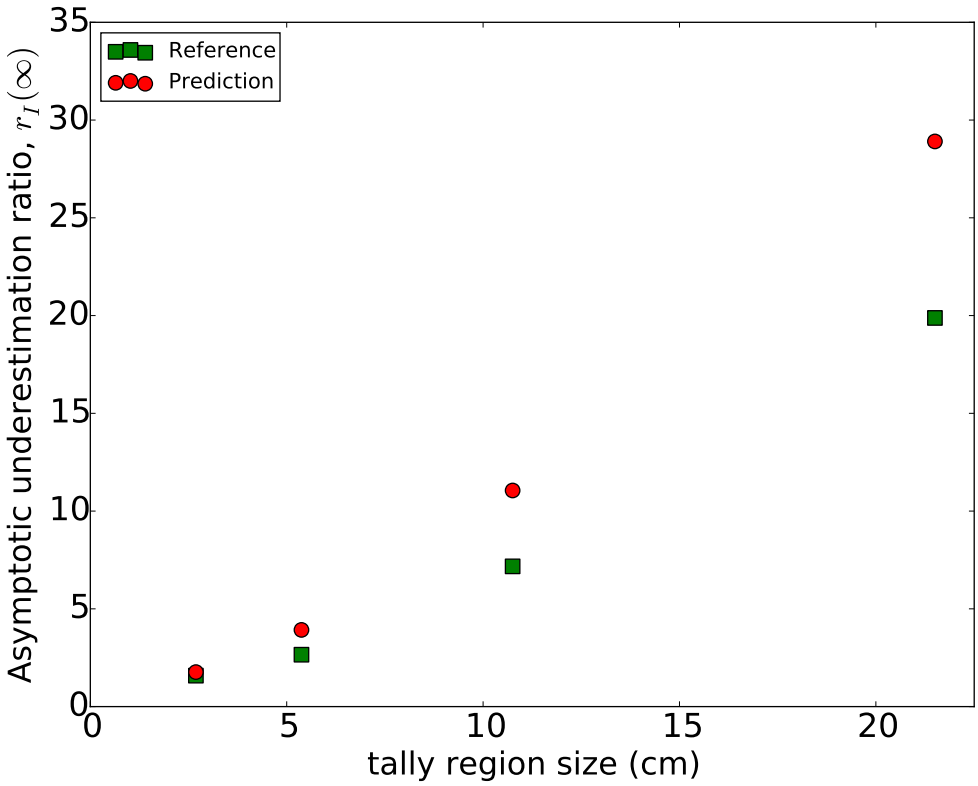


Figure 4-12: Asymptotic variance underestimation ratio for central tally region with different sizes in 2D BEAVRS. The squares correspond to r_I estimated from independent simulations. The dots correspond to predictions from the continuous model with sufficient expansion terms.

Chapter 5

Reducing the impact of correlations

When performing large Monte Carlo simulations, restrictions often come in the form of total runtime which is proportional to the total number of histories N_t. We denote the number of inactive generations as N_0, the relative cost of simulating one neutron in the inactive generations as β, (assuming the cost per neutron in the active generations is 1) and the number of neutrons per inactive and active generations as s' and s, respectively. Thus, the total cost after N active generations is

$$N_t = \beta N_0 s' + N s. \tag{5.0.1}$$

Alternatively, reactor simulations often seek answers to a target accuracy [40], thus leading to two possible types of optimization:

1. given a total computational amount N_t, find N and s to minimize the tally variance

2. given a target variance, find N and s to minimize N_t

Section 5.1 seeks to identify the optimal balance between generations and neutrons per generation in highly correlated systems. Section 5.2 proposes a modified source update algorithm in Monte Carlo eigenvalue simulations that introduces the concept of delayed neutrons in order to reduce the impact of generation-to-generation correlations on tally variance. The correlation prediction method based on the MBP

model developed in section 3.3 can be generalized to predict correlation behavior in the new delayed neutron simulation scheme. Section 5.3 identifies the optimal run strategy when using the proposed delayed neutron scheme.

5.1 Optimizing active generations

In this analysis, we assume that we are starting from a stationary source. Obviously, obtaining a stationary source for large correlated problems can be quite expensive, but it is assumed that this portion of the work can be greatly accelerated by other schemes currently in development [22] [31] [30] [27] [5] [8] [48] [16][36].

To find an optimal run strategy of either type 1 or type 2 requires an exact form of the variance or alternatively the error metric, $RMS^2_{predict}$, as a function of N and s. Results from both chapter 3 and chapter 4 show that the correlation coefficients can be expressed as a sum of decaying exponentials as a function of the generation lag.

The analytic form of $RMS^2_{predict}$ was derived in section 2.3 considering only the leading term of the correlation coefficients expressed by a sum of decaying exponentials. The conclusions derived in this section hold for correlation coefficients expressed with any number of decaying exponentials, however, for simplicity, only one exponential term will be considered. Starting from Eq 2.3.5 and re-writing it in terms of N and s

$$RMS^2_{predict}(N,s) = \frac{c}{N\ s}[1+2\rho_0(\frac{q}{1-q}-\frac{q^2-q^N}{(1-q)^2N}-\frac{q+q^N}{(1-q)N})] \equiv \frac{c}{N\ s}f(N) \quad (5.1.1)$$

where c is the constant first introduced in Eq 2.3.10.

The optimization can be performed in the N vs s plane but can be simplified in the $\log N$ vs $\log s$ plane since the fixed $N \times s$ constraint becomes a straight line with slope -1. Using parameters c,q and ρ_0 evaluated from the cube problem with a $16\times16\times16$ mesh, $RMS_{predict}$ is plotted in Fig 5-1. To perform an optimization of type 1, we need to find the point on the $\log N + \log s = \log N_t$ curve where $RMS_{predict}$

is minimal. To perform optimization of type 2, we need to find the point on the $RMS_{predict}(N,s) = RMS_{goal}$ curve where $N \times s$ is minimal.

Two equal-$RMS_{predict}$ curves are plotted in Fig 5-1(a) using thick black lines. One important feature of these curves in $\log N$ vs $\log s$ plane that affects optimization can be stated as such

$$\frac{d\log s}{d\log N} \in (-1, 0) \tag{5.1.2}$$

which is demonstrated below.

Proof. Since $\log s$ and $\log N$ are on the $equal - RMS^2_{predict}$ curve, the slope of the curve, $\frac{d\log s}{d\log N}$ can be related to the derivative of $RMS^2_{predict}$ with respect to $\log s$ and $\log N$ as shown in Eq 5.1.3.

$$\frac{d\log s}{d\log N} = -\frac{\frac{\partial RMS^2_{predict}(N,s)}{\partial \log N}}{\frac{\partial RMS^2_{predict}(N,s)}{\partial \log s}} \tag{5.1.3}$$

Converting the derivatives with respect to $\log N$ and $\log s$ to N and s, and knowing the dependence of $RMS^2_{predict}$ on s from Eq 5.1.1, Eq 5.1.3 is evaluated to yield Eq 5.1.4.

$$\frac{d\log s}{d\log N} = -\frac{N}{s}\frac{\frac{\partial RMS^2_{predict}(N,s)}{\partial N}}{\frac{\partial RMS^2_{predict}(N,s)}{\partial s}} = -\frac{N}{s}\frac{\frac{\partial RMS^2_{predict}(N,s)}{\partial N}}{-\frac{c}{Ns^2}f(N)} \tag{5.1.4}$$

Recognizing the definition of $f(N)$ as in Eq 5.1.1 simplifies Eq 5.1.4 to Eq 5.1.5.

$$\frac{d\log s}{d\log N} = N\frac{\frac{\partial RMS^2_{predict}(N,s)}{\partial N}}{RMS^2_{predict}(N,s)} \tag{5.1.5}$$

Since $RMS^2_{predict}(N,s) > 0$, the following equivalence holds

$$\begin{aligned} &\frac{d\log s}{d\log N} \in (-1, 0) \\ &\Leftrightarrow N\frac{\partial RMS^2_{predict}(N,s)}{\partial N} + RMS^2_{predict}(N,s) > 0 \end{aligned} \tag{5.1.6}$$

which can be related with $\frac{\partial RMS^2_{predict}(N,s)}{\partial s}$ for a fixed $N_t = N \times s$ as shown below.

Recognizing the constant N_t in $RMS^2_{predict}(N,s)$ reduces the derivative of $RMS^2_{predict}(N,s)$ to the derivative of $f(N)$

$$\frac{\partial RMS^2_{predict}(N_t,s)}{\partial s} = \frac{c}{N_t}\frac{df(N)}{ds} \tag{5.1.7}$$

The dependence of $f(N)$ on s is through $N = N_t/s$, the chain rule is used by taking the derivative with respect to N then taking the derivative of N with respect to s.

$$\frac{\partial RMS^2_{predict}(N_t,s)}{\partial s} = \frac{c}{N_t}\frac{df(N)}{dN}\frac{dN}{ds} = -\frac{c}{N_t}\frac{df(N)}{dN}\frac{N_t}{s^2} = -\frac{c}{s^2}\frac{df(N)}{dN} \tag{5.1.8}$$

Then recall the relation between $RMS^2_{predict}(N,S)$ and $f(N)$ in Eq 5.1.1, we have

$$\begin{aligned}\frac{\partial RMS^2_{predict}(N,s)}{\partial N} &= \frac{c}{s}\left(f(N)\frac{\partial}{\partial N}\frac{1}{N} + \frac{1}{N}\frac{\partial f(N)}{\partial N}\right)\\ &= \frac{c}{s}\left(-\frac{f(N)}{N^2} + \frac{1}{N}\frac{df(N)}{dN}\right)\end{aligned} \tag{5.1.9}$$

Inserting Eq 5.1.9 into Eq 5.1.8,

$$\begin{aligned}\frac{\partial RMS^2_{predict}(N_t,s)}{\partial s} &= -\frac{N}{s}\frac{\partial RMS^2_{predict}(N,s)}{\partial N} - \frac{c}{s^2}\frac{f(N)}{N}\\ &= -\left(N\frac{\partial RMS^2_{predict}(N,s)}{\partial N} + RMS^2_{predict}(N,s)\right)\frac{1}{s}\end{aligned} \tag{5.1.10}$$

According to the equivalence in Eq 5.1.6 and the relation in Eq 5.1.10, Eq 5.1.2 can be demonstrated by showing that $RMS^2_{predict}(N_t,s)$ decreases monotonically as a function of s for constant N_t.

The monotonicity of $RMS^2_{predict}(s)$ for a fixed N_t can be easily proven by first

writing $RMS^2_{predict}$ (Eq 2.3.5) as a function of s with parameter N_t in Eq 5.1.11

$$RMS^2_{predict}(N_t; s) = \frac{c}{N_t}\left[1 + 2\rho_0\left(\frac{q}{1-q} - \frac{q^2 - q^{N_t/s}}{(1-q)^2 N_t/s} - \frac{q + q^{N_t/s}}{(1-q)N_t/s}\right)\right] \quad (5.1.11)$$

and then taking the derivative with respect to s as in Eq 5.1.12.

$$\frac{dRMS^2_{predict}(N_t; s)}{ds} = -\frac{2q\rho_0}{N_t(1-q)^2 s}\left(s - q^{N_t/s}s + N_t q^{N_t/s} log(q)\right) < 0 \quad (5.1.12)$$

where the sign of the derivative is always negative. Therefore $RMS_{predict}$ decreases with generation size s. Note that if we use a sum of exponentials instead of a single term in the expression of correlation coefficients, Eq 5.1.12 still holds.

This can also be easily observed from Fig 2-5. For the same number of total histories, before reaching the asymptotic regime, a larger s always produces a lower RMS, thus indicating strict monotonous decrease of $RMS^2_{predict}(N_t; s)$ as a function of s.

Validity of Eq 5.1.2 has thus been demonstrated by the derivation of Eq 5.1.12 and Eq 5.1.10. □

Defining the tangent vector, τ, and the gradients of N_t and $RMS^2_{predict}$ can facilitate the analysis of both optimization types.

$$\boldsymbol{\nabla} RMS^2_{predict} = \left(\frac{\partial RMS^2_{predict}(N, s)}{\partial \log N}, \frac{\partial RMS^2_{predict}(N, s)}{\partial \log s}\right) \equiv (g_1, g_2) \quad (5.1.13)$$

$$\boldsymbol{\tau}(RMS^2_{predict}) = \left(1, -\frac{g_1}{g_2}\right) \quad (5.1.14)$$

$$\boldsymbol{\nabla} N_t = (1, 1) \quad (5.1.15)$$

$$\boldsymbol{\tau}(N_t) = (1, -1) \quad (5.1.16)$$

where $\boldsymbol{\nabla} RMS^2_{predict}$ indicates the direction where $RMS^2_{predict}$ increases faster on the $(\log(N), \log(s))$ plane; $\boldsymbol{\tau}(RMS^2_{predict})$ is the direction where $RMS^2_{predict}$ stays constant (because it is perpendicular to $\boldsymbol{\nabla} RMS^2_{predict}$) while $\log(N)$ increases (because the first component is positive); $\boldsymbol{\nabla} N_t$ indicates the direction where N_t increases faster on the

$(\log(N), \log(s))$ plane; $\boldsymbol{\tau}(N_t)$ is the direction where N_t remains constant (because it is perpendicular to $\boldsymbol{\nabla} N_t$) while $\log(N)$ increases.

5.1.1 Type 1

Optimization of type 1 is to find the lowest $RMS^2_{predict}$ on the straight line of constant cost N_t ($N \times s = 10^5$ in Fig 5-1(a)). From Fig 5-1(a), this clearly indicates that the lowest $RMS_{predict}$ predict is obtained with the fewest number of generations. Intuitively, this is because the $equal - RMS^2_{predict}$ curves are less steep than the constraint equal-N_t curve (Eq 5.1.2). When moving along the equal-N_t curve with increasing $\log(N)$, the equal-N_t curve intersects equal-$RMS^2_{predict}$ curves with higher $RMS^2_{predict}$ values. Mathematically, we have

$$\boldsymbol{\tau}(N_t) \cdot \boldsymbol{\nabla} RMS^2_{predict} = g_1 - g_2 \tag{5.1.17}$$

From the definition of g_1 and g_2 in Eq 5.1.13,

$$g_1 < 0; \qquad g_2 < 0 \tag{5.1.18}$$

From Eq 5.1.2, we have

$$-\frac{g1}{g2} > -1 \tag{5.1.19}$$

Combining Eq 5.1.17,Eq 5.1.18 and Eq 5.1.19 gives

$$\boldsymbol{\tau}(N_t) \cdot \boldsymbol{\nabla} RMS^2_{predict} > 0 \tag{5.1.20}$$

Eq 5.1.20 says that given a constant N_t, increasing $\log(N)$ (equivalently increasing N) also increases $RMS^2_{predict}$. That is, fewer N (generations) is preferred. It should be reminded that this assumes that a stationary source had previously been achieved.

5.1.2 Type 2

Optimization of type 2 is to find the lowest N_t for a desired $RMS_{predict}$ ($RMS^2_{predict} = 1.88 \times 10^{-6}$ in Fig 5-1(b)). From Fig 5-1(b), this clearly indicates that the lowest N_t cost is obtained with the fewest number of generations. Intuitively, this is because equal-N_t curves are steeper than the constraint $equal - RMS^2_{predict}$ curve (Eq 5.1.2). When moving along the equal-$RMS^2_{predict}$ curve with increasing $\log(N)$, the equal-$RMS^2_{predict}$ curve intersects equal-N_t curves with higher N_t. Mathematically, we have

$$\boldsymbol{\tau}(RMS^2_{predict}) \cdot \boldsymbol{\nabla} N_t = 1 - \frac{g_1}{g_2} \tag{5.1.21}$$

Combining Eq 5.1.21 and Eq 5.1.19 gives

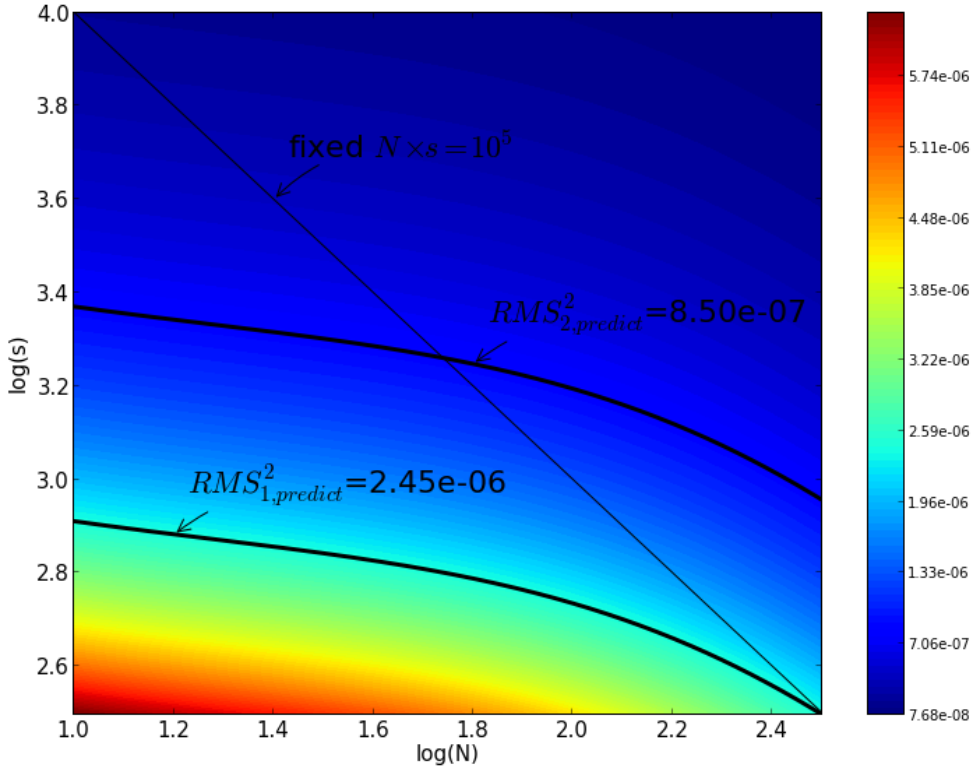
$$\boldsymbol{\tau}(RMS^2_{predict}) \cdot \boldsymbol{\nabla} N_t > 0 \tag{5.1.22}$$

Eq 5.1.20 says that given constant $RMS^2_{predict}$, increasing $\log(N)$ (equivalently increasing N) increases N_t. Thus, once again, fewer N is preferred.

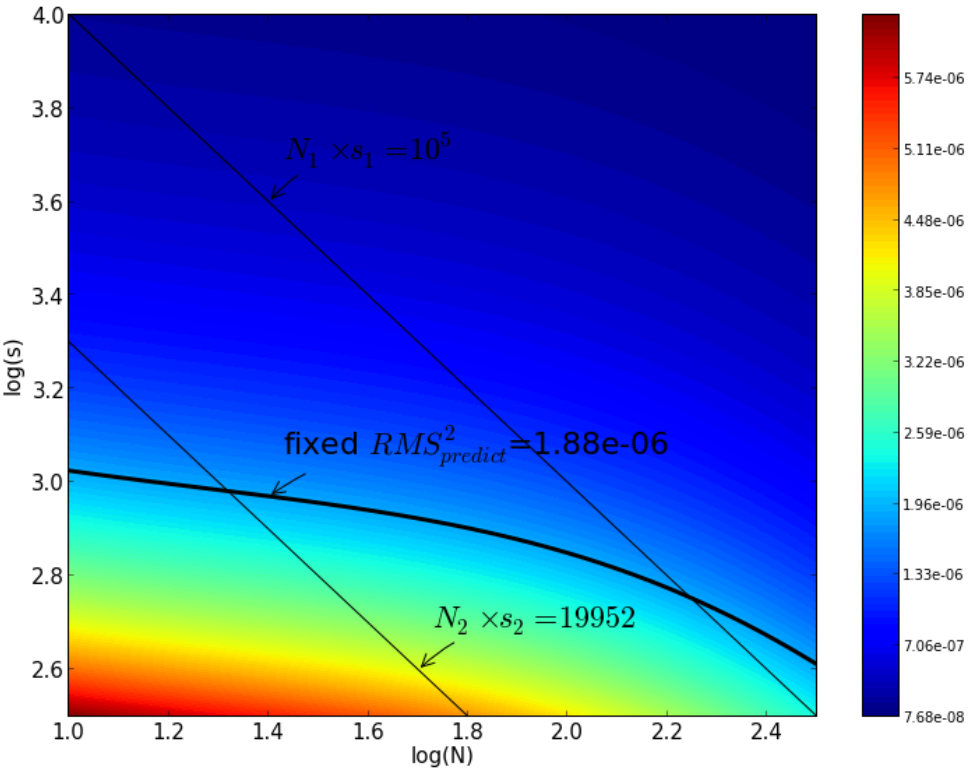
Everything we have done so far indicates a need for large generation size, s, and few generations which is not always practical. Large generation size will impact the inactive generations and can require an excessive amount of memory for the storage of the fission source. Additionally, common practice requires the definition of confidence intervals often done over generations, which was discussed in section 2.4.

5.2 Delayed Neutron Method

Section 5.2.1 proposes a new source update in Monte Carlo eigenvalue simulations, while section 5.2.2 builds a model to predict correlation behavior of the new simulation scheme.



(a) Optimization given N_t



(b) Optimization given $RMS_{predict}$

Figure 5-1: Two types of optimization

5.2.1 New source update scheme with delayed neutrons

In a typical Monte Carlo simulation with s neutrons per generation, the s neutrons are first transported according to the physics determined by the material properties of the system and end up producing s' new neutrons in the fission bank. s neutrons are then sampled from the s' bank to serve as the starting source for the next generation. The procedure is illustrated in Figure 5-2(a).

A new source iteration is proposed as follows. Suppose there are s neutrons per generation in the inactive generations. At the end of inactive generations, the source bank is split into two parts: the prompt source bank and the delayed source bank. αs neutrons are sampled into the prompt source bank and the remaining $(1-\alpha)s$ are placed in the delayed source bank. Then in each active generation, αs neutrons from the prompt source bank are simulated and generate s_f new neutrons in the fission bank. A fraction f of the fission bank, and a fraction p of the delayed source bank are sampled into the intermediate prompt source bank. There are thus $fs_f + p(1-\alpha)s$ source neutrons in the intermediate prompt source bank. And the remaining $1-f$ fraction of the fission bank and $1-p$ fraction of the delayed source bank are placed into the intermediate delayed source bank. There are $(1-f)s_f + (1-p)(1-\alpha)s$ source neutrons in the intermediate delayed source bank. The procedure is illustrated in Figure 5-2(b). For convenience, a neutron in the prompt source bank will be referred to as a prompt neutron. And a neutron in the delayed source bank will be referred to as a delayed neutron.

The steps of simulating neutrons in the prompt source bank and updating the source banks can be continued for any number of generations. In the traditional OpenMC scheme, the number of new neutrons for each fission event is normalized by the estimated k_{eff} to maintain the neutron population near constant. That is, $s_f \approx s$. In the delayed scheme, we require the intermediate prompt source bank to have the same number of neutrons as in the prompt source bank when $s_f = \alpha s$, that

is

$$\begin{aligned} \alpha s &= s_f f + (1-\alpha) sp \\ \Leftrightarrow \alpha &= \alpha f + (1-\alpha) p \end{aligned} \tag{5.2.1}$$

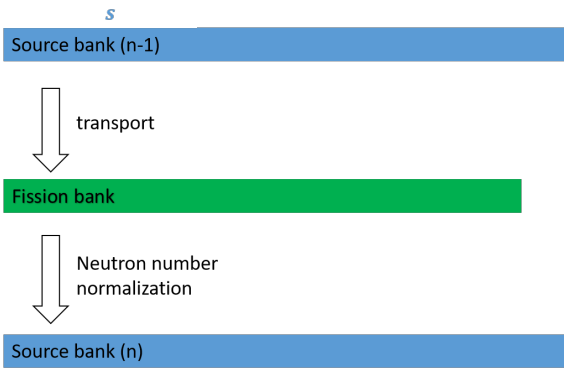
The total number of neutrons in both the prompt source bank and delayed source bank is kept at s. The normalization also extend to the delayed source banks written below

$$\begin{aligned} (1-\alpha) s &= s_f (1-f) + (1-\alpha) s (1-p) \\ \Leftrightarrow 1-\alpha &= \alpha(1-f) + (1-\alpha)(1-p) \end{aligned} \tag{5.2.2}$$

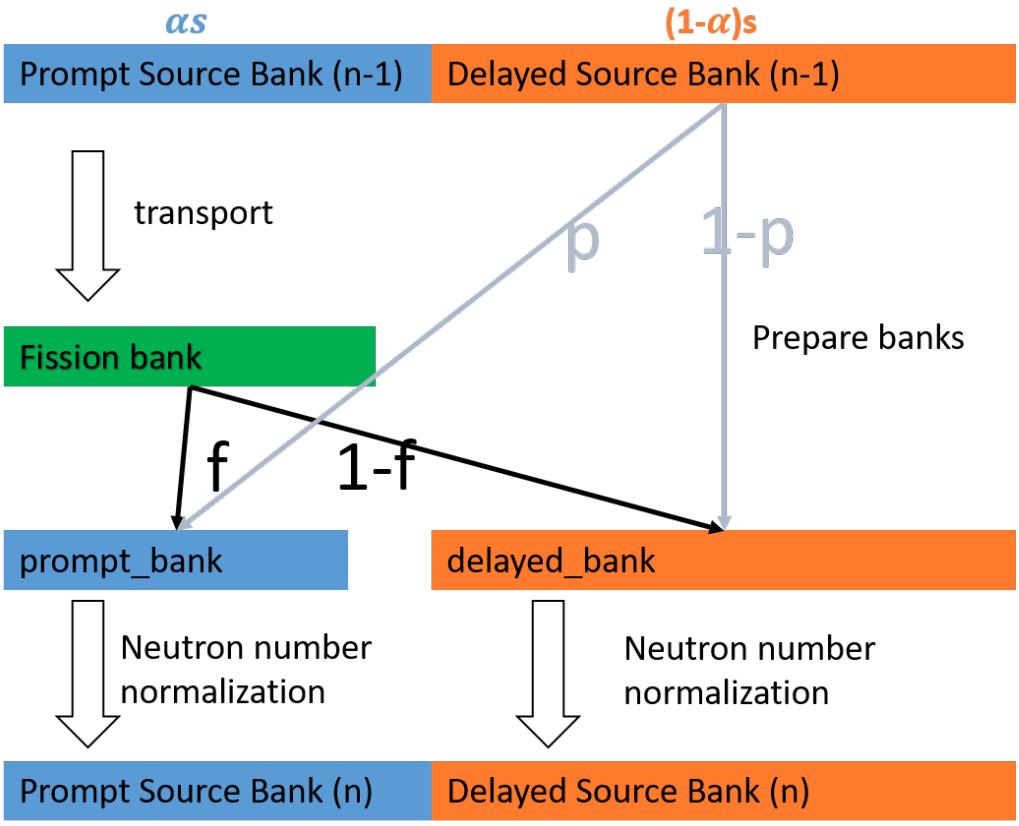
Any one of the two equations (Eq 5.2.1 and Eq 5.2.2) suffices to determine one of the three parameters given the other two. Since $\alpha, p, f \in (0,1)$, it can be shown that p must be chosen from $(0, \frac{\alpha}{1-\alpha})$ if $\alpha < 0.5$ and f must be chosen from $(2-\frac{1}{\alpha}, 1)$ if $\alpha > 0.5$.

5.2.2 Generalizing MBP model

Now, we will generalize the MBP model developed in Chapter 3 to describe the neutron transport process discussed above. We once again discretize the neutron phase space over all independent variables (space, angle, energy) into M discrete regions. We then characterize the system state at generation n with two vectors $\vec{Z}(n)$ and $\vec{U}(n)$ such that $Z_l(n)$ is the number of neutrons in the prompt source bank in region l at generation n and $U_l(n)$ is the number of neutrons in the delayed source bank in region l at generation n. There is an intermediate state between the states characterized by $\vec{Z}(n)$, $\vec{U}(n)$ and $\vec{Z}(n+1)$, $\vec{U}(n+1)$. The neutrons characterized by $\vec{Z}(n)$ are transported into the fission bank (Figure 5-2(b)) before being sampled into $\vec{Z}(n+1)$ and $\vec{U}(n+1)$. We denote the state of the fission bank at generation n with vector $\vec{Z}^*(n)$ such that $Z_l^*(n)$ is the number of neutrons in region l in the fission bank generated by the prompt source bank at generation n. The tallies are still contributed



(a) Traditional method



(b) Method of delayed neutrons

Figure 5-2: Source update schemes

by neutrons from the prompt source bank, the vector $\vec{Z}(n)$. This section finds the recursive formula for the moments of $\vec{Z}$ to evaluate correlation coefficients of tallies. The new model is named Delayed Multitype Branching Processes (DMBP).

We then seek the moment generating function of $\vec{Z}(n)$ and $\vec{U}(n)$ in the form of an expectation conditioned on $\vec{Z}(n-1)$ and $\vec{U}(n-1)$ such that the recursive formula can be derived to write the moments of $\vec{Z}(n)$ and $\vec{U}(n)$ in terms of moments of $\vec{Z}(n-1)$ and $\vec{U}(n-1)$.

The moment generating function of $\vec{Z}(n)$ and $\vec{U}(n)$ given $\vec{Z}(n-1)$ and $\vec{U}(n-1)$ is defined in Eq 5.2.3.

$$F_n(\vec{s},\vec{\sigma}|\vec{Z}(n-1),\vec{U}(n-1)) \equiv \mathbb{E}\left[\prod_i s_i^{Z_i(n)}\sigma_i^{U_i(n)}|\vec{Z}(n-1),\vec{U}(n-1)\right] \tag{5.2.3}$$

Then following the same derivation as in Eq 3.3.16, the moment generating function in Eq 5.2.3 can be written in terms of moment generating functions given a prompt neutron source $f_i(\vec{s},\vec{\sigma})$ and moment generating functions given a delayed neutron source $g_i(\vec{s},\vec{\sigma})$ as in Eq 5.2.4.

$$F_n(a,\alpha;\vec{s},\vec{\sigma}|\vec{Z}(n-1),\vec{U}(n-1)) = \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\prod_l \left(g_l(\vec{s},\vec{\sigma})\right)^{U_l(n-1)} \tag{5.2.4}$$

where the moment generation functions given one point source are defined in Eq 5.2.5 and Eq 5.2.6.

$$f_l(\vec{s},\vec{\sigma}) \equiv \mathbb{E}\left[\prod_i s_i^{Z_i(1)}\sigma_i^{U_i(1)}\middle| Z_i(0)=\delta_{i,l}, U_i(0)=0\right] \tag{5.2.5}$$

$$g_l(\vec{s},\vec{\sigma}) \equiv \mathbb{E}\left[\prod_i s_i^{Z_i(1)}\sigma_i^{U_i(1)}\middle| Z_i(0)=0, U_i(0)=\delta_{i,l}\right] \tag{5.2.6}$$

5.2.3 Evolution of DMBP moments

In order to predict the correlation coefficients of X tallies (normalized from prompt neutron counts Z as in Eq 3.1.4), the moments of $\vec{Z}$ in the form of $\mathbb{E}Z_l(n)$, $\mathbb{E}[Z_l(n)Z_j(n)]$ and $\mathbb{E}[Z_l(n)Z_j(n+k)]$ should be evaluated. Similar to the MBP model, the derivations

below first use the derivative of the MGF (Eq 5.2.4) to evaluate conditional expectations, then take the expectation of the conditional expectation to obtain recursive formula for the moments needed to evaluate the generation-to-generation correlation. The recursive evaluation of moments is termed evolution.

5.2.3.1 First Order Moments

This section derives the recursive formula for the expectation of the number of prompt neutrons ($Z_a(n)$) and the number of delayed neutrons ($U_a(n)$) as a function of region (or type) and generation.

The conditional expectation of the number of prompt neutrons in region a at generation n given a number of prompt neutrons in each region $\vec{Z}(n-1)$ and a number of delayed neutrons in each region $\vec{U}(n-1)$ is obtained by taking the derivative of the moment generation function in Eq 5.2.4 with respect to s_a and evaluating at $\vec{s}=\vec{\sigma}=\vec{1}$,

$$\begin{aligned}\mu_a(n)|_{\vec{Z}(n-1),\vec{U}(n-1)} &= \frac{\partial}{\partial s_a}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right]\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\ &= U_l(n-1)m_a^l + Z_l(n-1)M_a^l\end{aligned} \tag{5.2.7}$$

where the second equality is derived in Appendix A1.4 (Eq A1.4.26). Two response moments are defined in Eq 5.2.7, M_a^l is defined as

$$M_a^l \equiv \frac{\partial}{\partial s_a} f_l(\vec{s},\vec{\sigma})\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} = \mathbb{E}\left[Z_a(1)|Z_i(0)=\delta_{i,l}, U_i(0)=0\right] \tag{5.2.8}$$

and m_a^l is defined as

$$m_a^l \equiv \frac{\partial}{\partial s_a} g_l(s_l,\sigma_l)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} = \mathbb{E}\left[Z_a(1)|Z_i(0)=0, U_i(0)=\delta_{i,l}\right] \tag{5.2.9}$$

M_a^l is the expected number of neutrons observed at region a if a prompt neutron is started from region l. Similarly, m_a^l is the expected number of neutrons observed in region a if a delayed neutron is started in region l.

Similarly, taking the derivative of MGF in Eq 5.2.4 with respect to σ_a yields the conditional expectation of $U_a(n)$ given $\vec{Z}(n-1)$ and $\vec{U}(n-1)$

$$\begin{aligned}\rho_a(n)|_{\vec{Z}(n-1),\vec{U}(n-1)} &= \frac{\partial}{\partial\sigma_a}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right] \\ &= U_l(n-1)\psi_a^l + Z_l(n-1)\Psi_a^l\end{aligned} \tag{5.2.10}$$

where the second equality is derived in Appendix A1.4 (Eq A1.4.27). There are two newly defined response moments in Eq 5.2.10, Ψ_a^l is defined as

$$\Psi_a^l \equiv \frac{\partial}{\partial\sigma_a} f_l(\vec{s},\vec{\sigma})\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} = \mathbb{E}\left[U_a(1)|Z_i(0)=\delta_{i,l}, U_i(0)=0\right] \tag{5.2.11}$$

and ψ_a^l is defined as

$$\psi_a^l \equiv \frac{\partial}{\partial\sigma_a} g_l(s_l,\sigma_l)\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} = \mathbb{E}\left[U_a(1)|Z_i(0)=0, U_i(0)=\delta_{i,l}\right] \tag{5.2.12}$$

By definition of $f_l(\vec{s},\vec{\sigma})$, Ψ_a^l is the expected number of delayed neutrons observed in region a if a prompt neutron is started in region l. Similarly, ψ_a^l is the expected number of delayed neutrons observed in region a if a delayed neutron is started in region l.

Evaluating the above conditional expectations with respect to the distributions of Z_{n-1} and U_{n-1} leads to

$$\mu_a(n) = \rho_l(n-1)m_a^l + \mu_l(n-1)M_a^l \tag{5.2.13}$$

$$\rho_a(n) = \rho_l(n-1)\psi_a^l + \mu_l(n-1)\Psi_a^l \tag{5.2.14}$$

Recursively calling Eq 5.2.13 and Eq 5.2.14 gives the expectation of the number of prompt and delayed neutrons in reach region (of every type) in the system at any generation n.

The response moments defined in Eq 5.2.8,Eq 5.2.9,Eq 5.2.11 and Eq 5.2.12 follow the below notational convention. The moments of prompt neutrons M and m are

evaluated by taking derivatives with respect to argument $\vec{s}$ in the MGF (Eq 5.2.4). The moments of delayed neutrons Ψ and ψ are evaluated by taking derivatives with respect to argument $\vec{\sigma}$ in the MGF (Eq 5.2.4). It is clear that M and m belong to the same set and represent the generation of prompt neutrons, and that Ψ and ψ also belong to the same set representing the generation of delayed neutrons. In addition, M and Ψ are moments induced by a prompt neutron and are capitalized. m and ψ are moments induced by a delayed neutron and are lower case letters.

5.2.3.2 Second Order Moments

This section derives the recursive formula for the expectation of the product of the number of prompt neutrons ($Z_a(n)$) and the number of delayed neutrons ($U_a(n)$) as a function of region (or type) and generation. There are three categories of such products:

- product of the number of prompt neutrons

$$C_{a,b}(n) \equiv \mathbb{E}[Z_a(n)Z_b(n)] \tag{5.2.15}$$

- product of the number of delayed neutrons

$$\Gamma_{a,b}(n) \equiv \mathbb{E}[U_a(n)U_b(n)] \tag{5.2.16}$$

- product of the number of prompt and delayed neutrons

$$C\Gamma_{a,b}(n) \equiv \mathbb{E}[Z_a(n)U_b(n)] \tag{5.2.17}$$

Following the same procedure as in the last section, the recursive formulas are derived by first obtaining the conditional expectation from the derivatives of the MGF (Eq 5.2.4) and then writing the expectations at generation n in terms of the expectations at generation $n-1$.

5.2.3.2.1 $C_{a,b}(n)$

The conditional expectation of the product of the number of prompt neutrons in region a and the number of prompt neutrons in region b at generation n given the information on prompt and delayed neutrons in generation $n-1$ is obtained by taking the derivative of the MGF (Eq 5.2.4) with respect to s_a and s_b and evaluating at $\vec{s} = \vec{\sigma} = \vec{1}$,

$$\begin{aligned}
&\mathbb{E}\left[Z_a(n)Z_b(n) - Z_a(n)\delta_{a,b} | \, \vec{Z}(n-1), \vec{U}(n-1) \right] \\
&\equiv \text{Ⅽ}_{a,b}(n)|_{\vec{Z}(n-1),\vec{U}(n-1)} \\
&= \frac{\partial^2}{\partial s_a \partial s_b} \left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \right] \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= Z_l(n-1)(\bar{\text{Ⅴ}}^l_{a,b} - M^l_b M^l_a) + Z_l(n-1)M^l_a Z_j(n-1)M^j_b \\
&+ Z_l(n-1)M^l_a U_k(n-1)m^k_b + U_k(n-1)m^k_a Z_l(n-1)M^l_b \\
&+ U_k(n-1)(\bar{\text{ⅴ}}^k_{a,b} - m^k_b m^k_a) + U_k(n-1)m^k_a U_j(n-1)m^j_b
\end{aligned} \tag{5.2.18}$$

where the last equality is derived in Appendix A1.4 (Eq A1.4.28). Two new second order response moments are defined in Eq 5.2.18, $\bar{\text{Ⅴ}}^l_{a,b}$ is defined as

$$\begin{aligned}
\bar{\text{Ⅴ}}^l_{a,b} &\equiv \frac{\partial^2}{\partial s_a \partial s_b} f_l(\vec{s},\vec{\sigma}) \Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \mathbb{E}\left[Z_a(1)Z_b(1) - Z_a(1)\delta_{a,b} | \, Z_i(0) = \delta_{i,l}, U_i(0) = 0 \right]
\end{aligned} \tag{5.2.19}$$

and $\bar{\text{ⅴ}}^k_{a,b}$ is defined as

$$\begin{aligned}
\bar{\text{ⅴ}}^k_{a,b} &\equiv \frac{\partial^2}{\partial s_a \partial s_b} g_l(s_l,\sigma_l) \Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \mathbb{E}\left[Z_a(1)Z_b(1) - Z_a(1)\delta_{a,b} | \, Z_i(0) = 0, U_i(0) = \delta_{i,l} \right]
\end{aligned} \tag{5.2.20}$$

Taking second order derivatives of the MGF (Eq 5.2.4) does not necessarily yield the expectation of the product of prompt neutrons. When $a = b$, the derivative corresponds to factorial moments (discussion in section 3.3.1.1) and results in the $Z_a(n)\delta_{a,b}$ term in Eq 5.2.18. The difference between ordinary moments and factorial

moments is also exhibited by denoting the former with $C_{a,b}$ and the latter with $\text{Y}C_{a,b}$. The same convention also holds for the definition of $\text{Y}\bar{V}$(Eq 5.2.19) and $\text{Y}\bar{\nu}$(Eq 5.2.20).

Next, evaluating the expectation of $\vec{Z}(n-1)$ and $\vec{U}(n-1)$ on both sides of Eq 5.2.18

$$\begin{aligned}\text{Y}C_{a,b}(n) = {} & \mu_l(n-1)(\text{Y}\bar{V}^l_{a,b} - M^l_b M^l_a) + C_{l,j}(n-1)M^l_a M^j_b \\ & + M^l_a C\Gamma_{l,k}(n-1)m^k_b + m^k_a C\Gamma_{l,k}(n-1)M^l_b \\ & + \rho_k(n-1)(\text{Y}\bar{\nu}^k_{a,b} - m^k_b m^k_a) + \Gamma_{k,j}(n-1)m^k_a m^j_b\end{aligned} \tag{5.2.21}$$

Eq 5.2.21 contains both ordinary and factorial moments of the prompt neutron counts (C and $\text{Y}C$). Since the ordinary moments C will be used to predict correlation and thus the variance estimator ratio, $\text{Y}C$ should be removed from Eq 5.2.21. According to the definition of C (Eq 5.2.15) and $\text{Y}C$ (first equation in Eq 5.2.18), C and $\text{Y}C$ are related through:

$$\text{Y}C_{a,b}(n) = C_{a,b}(n) - \mu_a(n)\delta_{a,b} \tag{5.2.22}$$

For consistency in notation, the ordinary moments form of the response moments $\text{Y}\bar{V}$ and $\text{Y}\bar{\nu}$ are used. Response moments $\bar{V}$ (response of the prompt neutron pairs to a prompt neutron) and $\bar{\nu}$ (response of the prompt neutron pairs to a delayed neutron) are naturally defined as

$$\bar{V}^l_{a,b} \equiv \mathbb{E}\left[Z_a(1)Z_b(1)|\, Z_i(0) = \delta_{i,l}, U_i(0) = 0\right] \tag{5.2.23}$$

$$\bar{\nu}^l_{a,b} \equiv \mathbb{E}\left[Z_a(1)Z_b(1)|\, Z_i(0) = 0, U_i(0) = \delta_{i,l}\right] \tag{5.2.24}$$

$\bar{V}$ and $\text{Y}\bar{V}$ are related through:

$$\text{Y}\bar{V}^l_{a,b} = \bar{V}^l_{a,b} - \mathbb{E}\left[Z_a(1)\delta_{a,b}|\, Z_i(0) = \delta_{i,l}, U_i(0) = 0\right] = \bar{V}^l_{a,b} - M^l_a\delta_{a,b} \tag{5.2.25}$$

where the first equality compares the definition of $\bar{V}$(Eq 5.2.23) and $\text{Y}\bar{V}$ (Eq 5.2.19), the second equality realizes that the difference between $\bar{V}$ and $\text{Y}\bar{V}$ is the definition of M (first order moment of prompt neutron count in response to a prompt neutron) in Eq 5.2.8.

In complete analogy to $\bar{V}$ and $\bar{\bar{V}}$, $\bar{v}$ and $\bar{\bar{v}}$ are related through:

$$\bar{\bar{v}}^l_{a,b} = \bar{v}^l_{a,b} - \mathbb{E}\left[Z_a(1)\delta_{a,b}|\, Z_i(0) = 0, U_i(0) = \delta_{i,l}\right] = \bar{v}^l_{a,b} - m^l_a\delta_{a,b} \tag{5.2.26}$$

With the relation between ordinary and factorial moments for V (Eq 5.2.25) , v (Eq 5.2.26) and C (Eq 5.2.22), Eq 5.2.21 is transformed to

$$\begin{aligned} C_{a,b}(n) = {} & \mu_a(n)\delta_{a,b} + \\ & \mu_l(n-1)(\bar{V}^l_{a,b} - M^l_a\delta_{a,b} - M^l_b M^l_a) + C_{l,j}(n-1)M^l_a M^j_b \\ & + M^l_a C\Gamma_{l,k}(n-1)m^k_b + m^k_a C\Gamma_{l,k}(n-1)M^l_b \\ & + \rho_k(n-1)(\bar{v}^k_{a,b} - m^l_a\delta_{a,b} - m^k_b m^k_a) + \Gamma_{k,j}(n-1)m^k_a m^j_b \end{aligned} \tag{5.2.27}$$

Given the second order moments $C, C\Gamma$ and Γ at generation $n-1$, Eq 5.2.27 leads to the second order moment C at n with the exception that the expectation of the prompt neutron count μ is at generation n. Fortunately, $\mu(n)$ can be expressed in terms of the last generation moments $\mu(n-1)$ and $\rho(n-1)$ through Eq 5.2.13. This substitution also cancels the $M^l_a\delta_{a,b}$ and $m^l_a\delta_{a,b}$ terms and simplifies Eq 5.2.27 to

$$\begin{aligned} C_{a,b}(n) = {} & \mu_l(n-1)V^l_{a,b} + M^l_a C_{l,j}(n-1)M^j_b \\ & + M^l_a C\Gamma_{l,k}(n-1)m^k_b + m^k_a C\Gamma_{l,k}(n-1)M^l_b \\ & + \rho_k(n-1)v^k_{a,b} + m^k_a\Gamma_{k,j}(n-1)m^j_b \end{aligned} \tag{5.2.28}$$

where the central moments V and v are defined as

$$V^l_{a,b} \equiv \bar{V}^l_{a,b} - M^l_b M^l_a = Cov\left[Z_a(1), Z_b(1)|\, Z_i(0) = \delta_{i,l}, U_i(0) = 0\right] \tag{5.2.29}$$

$$v^l_{a,b} \equiv \bar{v}^l_{a,b} - m^l_b m^l_a = Cov\left[Z_a(1), Z_b(1)|\, Z_i(0) = 0, U_i(0) = \delta_{i,l}\right] \tag{5.2.30}$$

The new central moments are introduced since $\bar{V}^l_{a,b}$ always appear with $-M^l_a M^l_b$ and $\bar{v}^l_{a,b}$ always exists with $-m^l_a m^l_b$ and they can be interpreted as response covariances. $V^l_{a,b}$ is the covariance (2^{nd} order central moment) of the prompt neutron count in region a and region b in response to a prompt neutron started in region l. $v^l_{a,b}$ is

the covariance (2^{nd} order central moment) of the prompt neutron count in region a and region b in response to a delayed neutron started in region l.

As can be seen from Eq 5.2.28, the expectations of μ, ρ and C at generation $n-1$ are not sufficient to determine the expectation of C at generation n. The expectations of prompt and delayed neutron pairs $C\Gamma$ and delayed neutron pairs Γ are also needed, whose recursive formulas are derived in the next two sections.

5.2.3.2.2 $\Gamma_{a,b}(n)$

The conditional expectation of the product of delayed neutrons in region a and region b at generation n given information on prompt and delayed neutrons in generation $n-1$ is obtained by taking the derivative of the MGF (Eq 5.2.4) with respect to σ_a and σ_b and evaluating at $\vec{s}=\vec{\sigma}=\vec{1}$,

$$
\begin{aligned}
&\mathbb{E}\left[U_a(n)U_b(n)-U_a(n)\delta_{a,b}|\,\vec{Z}(n-1),\vec{U}(n-1)\right]\\
&\equiv \Upsilon_{a,b}(n)|_{\vec{Z}(n-1),\vec{U}(n-1)}\\
&=\frac{\partial^2}{\partial\sigma_a\partial\sigma_b}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right]\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}}\\
&=Z_l(n-1)(\bar{\lambda}^l_{a,b}-\Psi^l_b\Psi^l_a)+Z_l(n-1)\Psi^l_aZ_j(n-1)\Psi^j_b\\
&+Z_l(n-1)\Psi^l_aU_k(n-1)\psi^k_b+U_k(n-1)\psi^k_aZ_l(n-1)\Psi^l_b\\
&+U_k(n-1)(\bar{\lambda}^l_{a,b}-\psi^k_b\psi^k_a)+U_k(n-1)\psi^k_aU_j(n-1)\psi^j_b
\end{aligned}
\tag{5.2.31}
$$

where the last equality is derived in Appendix A1.4 (Eq A1.4.30). Two new second order response moments are defined in Eq 5.2.31, $\bar{\lambda}^l_{a,b}$ is defined as

$$
\begin{aligned}
\bar{\lambda}^l_{a,b}&\equiv\frac{\partial^2}{\partial\sigma_a\partial\sigma_b}f_l(\vec{s},\vec{\sigma})\Big|_{\vec{s}=\vec{\sigma}=\vec{1}}\\
&=\mathbb{E}\left[U_a(1)U_b(1)-U_a(1)\delta_{a,b}|\,Z_i(0)=\delta_{i,l},U_i(0)=0\right]
\end{aligned}
\tag{5.2.32}
$$

and $\bar{\not\lambda}^l_{a,b}$ is defined as

$$\begin{aligned}\bar{\not\lambda}^l_{a,b} &\equiv \frac{\partial^2}{\partial\sigma_a\partial\sigma_b} g_l(s_l,\sigma_l)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\ &= \mathbb{E}\left[U_a(1)U_b(1) - U_a(1)\delta_{a,b}\middle|\, Z_i(0)=0, U_i(0)=\delta_{i,l}\right]\end{aligned} \tag{5.2.33}$$

The extra $U_a\delta_{a,b}$ terms are present in Eq 5.2.31, Eq 5.2.32 and Eq 5.2.33 because the MGF (Eq 5.2.4) is in fact a factorial moment generating function (Paragraph 5.2.3.2.1).

We evaluate the expectation with respect to the distributions $\vec{Z}(n-1)$ and $\vec{U}(n-1)$ on both sides of Eq 5.2.31:

$$\begin{aligned}\not\Gamma_{a,b}(n) = {} & \mu_l(n-1)(\bar{\not\Lambda}^l_{a,b} - \Psi^l_b\Psi^l_a) + \Psi^l_a C_{l,j}(n-1)\Psi^j_b \\ & + \Psi^l_a C\Gamma_{l,k}(n-1)\psi^k_b + \psi^k_a C\Gamma_{l,k}(n-1)\Psi^l_b \\ & + \rho_k(n-1)(\bar{\not\lambda}^l_{a,b} - \psi^k_b\psi^k_a) + \psi^k_a\Gamma_{k,j}(n-1)\psi^j_b\end{aligned} \tag{5.2.34}$$

Γ and $\not\Gamma$ are related through:

$$\not\Gamma_{a,b}(n) = \Gamma_{a,b}(n) - \rho_a(n)\delta_{a,b} \tag{5.2.35}$$

Response moments $\bar{\Lambda}$ (response of delayed neutron pairs to a prompt neutron) and $\bar{\lambda}$ (response of delayed neutron pairs to a delayed neutron) are naturally defined as

$$\bar{\Lambda}^l_{a,b} \equiv \mathbb{E}\left[U_a(1)U_b(1)\middle|\, Z_i(0)=\delta_{i,l}, U_i(0)=0\right] \tag{5.2.36}$$

$$\bar{\lambda}^l_{a,b} \equiv \mathbb{E}\left[U_a(1)U_b(1)\middle|\, Z_i(0)=0, U_i(0)=\delta_{i,l}\right] \tag{5.2.37}$$

$\bar{\Lambda}$ and $\bar{\not\Lambda}$ are related through:

$$\bar{\not\Lambda}^l_{a,b} = \bar{\Lambda}^l_{a,b} - \mathbb{E}\left[U_a(1)\delta_{a,b}\middle|\, Z_i(0)=\delta_{i,l}, U_i(0)=0\right] = \bar{\Lambda}^l_{a,b} - \Psi^l_a\delta_{a,b} \tag{5.2.38}$$

where the first equality compares the definition of $\bar{\Lambda}$(Eq 5.2.36) and $\bar{\not\Lambda}$ (Eq 5.2.32),

the second equality realizes the difference between $\bar{\Lambda}$ and $\not{\bar{\Lambda}}$ as the definition of Ψ (first order moment of prompt neutron count in response to a prompt neutron) in Eq 5.2.11.

Analogous to $\bar{\Lambda}$ and $\not{\bar{\Lambda}}$, $\bar{\lambda}$ and $\not{\bar{\lambda}}$ are related through:

$$\not{\bar{\lambda}}^l_{a,b} = \bar{\lambda}^l_{a,b} - \mathbb{E}\left[U_a(1)\delta_{a,b} \middle| Z_i(0) = 0, U_i(0) = \delta_{i,l}\right] = \bar{\lambda}^l_{a,b} - m^l_a \delta_{a,b} \tag{5.2.39}$$

Now having a relation between ordinary and factorial moments for Λ (Eq 5.2.38) , λ (Eq 5.2.39) and Γ (Eq 5.2.35) Eq 5.2.34 are transformed to

$$\begin{aligned}\Gamma_{a,b}(n) = {} & \rho_a(n)\delta_{a,b} + \\ & \mu_l(n-1)(\bar{\Lambda}^l_{a,b} - \Psi^l_a\delta_{a,b} - \Psi^l_b\Psi^l_a) + \Psi^l_a C_{l,j}(n-1)\Psi^j_b \\ & + \Psi^l_a C\Gamma_{l,k}(n-1)\psi^k_b + \psi^k_a C\Gamma_{l,k}(n-1)\Psi^l_b \\ & + \rho_k(n-1)(\bar{\lambda}_{a,b} - \psi^l_a\delta_{a,b} - \psi^k_b\psi^k_a) + \psi^k_a\Gamma_{k,j}(n-1)\psi^j_b\end{aligned} \tag{5.2.40}$$

Given the second order moments $C, C\Gamma$ and Γ at generation $n-1$, Eq 5.2.40 leads to the second order moment Γ at n except that the expectation of the prompt neutron count ρ is at generation n. Fortunately, $\rho(n)$ can be expressed in terms of the last generation moments $\mu(n-1)$ and $\rho(n-1)$ through Eq 5.2.14. This substitution also cancels the $\Psi^l_a\delta_{a,b}$ and $\psi^l_a\delta_{a,b}$ terms and simplifies Eq 5.2.40 to

$$\begin{aligned}\Gamma_{a,b}(n) = {} & \mu_l(n-1)\Lambda^l_{a,b} + \Psi^l_a C_{l,j}(n-1)\Psi^j_b \\ & + \Psi^l_a C\Gamma_{l,k}(n-1)\psi^k_b + \psi^k_a C\Gamma_{l,k}(n-1)\Psi^l_b \\ & + \rho_k(n-1)\lambda^k_{a,b} + \psi^k_a\Gamma_{k,j}(n-1)\psi^j_b\end{aligned} \tag{5.2.41}$$

where the central moments Λ and λ are defined as

$$\Lambda^l_{a,b} \equiv \bar{\Lambda}^l_{a,b} - \Psi^l_b\Psi^l_a = Cov\left[U_a(1), U_b(1)\middle| Z_i(0) = \delta_{i,l}, U_i(0) = 0\right] \tag{5.2.42}$$

$$\lambda^l_{a,b} \equiv \bar{\lambda}^l_{a,b} - \psi^l_b\psi^l_a = Cov\left[U_a(1), U_b(1)\middle| Z_i(0) = 0, U_i(0) = \delta_{i,l}\right] \tag{5.2.43}$$

The new central moments are defined because $\bar{\Lambda}^l_{a,b}$ always appears with $-\Psi^l_a\Psi^l_b$ and

$\bar{\lambda}^l_{a,b}$ always exists with $-\psi^l_a\psi^l_b$ and they can be interpreted as response covariances. $\Lambda^l_{a,b}$ is the covariance (2^{nd} order central moment) of the delayed neutron count in region a and region b in response to a prompt neutron started in region l. $\lambda^l_{a,b}$ is the covariance (2^{nd} order central moment) of the delayed neutron count in region a and region b in response to a delayed neutron started in region l.

5.2.3.2.3 $C\Gamma_{a,b}(n)$

The conditional expectation of the product of the number of prompt neutrons in region a and the number of delayed neutrons in region b at generation n given information on prompt and delayed neutrons in generation $n-1$ is obtained by taking the derivative of the MGF (Eq 5.2.4) with respect to s_a and σ_b and evaluating at $\vec{s}=\vec{\sigma}=\vec{1}$,

$$\begin{aligned}
&\mathbb{E}\left[Z_a(n)U_b(n)|\,\vec{Z}(n-1),\vec{U}(n-1)\right]\\
&\equiv C\Gamma_{a,b}(n)|_{\vec{Z}(n-1),\vec{U}(n-1)}\\
&=\frac{\partial^2}{\partial s_a\partial\sigma_b}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right]\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}}\\
&=Z_l(n-1)(\bar{\Omega}^l_{a,b}-\Psi^l_bM^l_a)+Z_l(n-1)M^l_aZ_j(n-1)\Psi^j_b\\
&+Z_l(n-1)M^l_aU_k(n-1)\psi^k_b+U_k(n-1)m^k_aZ_l(n-1)\Psi^l_b\\
&+U_k(n-1)(\bar{\omega}^l_{a,b}-\psi^k_bm^k_a)+U_k(n-1)m^k_aU_j(n-1)\psi^j_b
\end{aligned}\tag{5.2.44}$$

where the last equality is derived in Appendix A1.4 (Eq A1.4.29). Two new second order response moments are defined in Eq 5.2.44, the response of the number of prompt and delayed neutron pairs from a prompt neutron $\bar{\Omega}^l_{a,b}$ is defined as

$$\begin{aligned}
\bar{\Omega}^l_{a,b}&\equiv\frac{\partial^2}{\partial s_a\partial\sigma_b}f_l(\vec{s},\vec{\sigma})\Big|_{\vec{s}=\vec{\sigma}=\vec{1}}\\
&=\mathbb{E}\left[Z_a(1)U_b(1)|\,Z_i(0)=\delta_{i,l},U_i(0)=0\right]
\end{aligned}\tag{5.2.45}$$

and the response of the number of prompt and delayed neutron pairs from a delayed neutron $\bar{\bar{\omega}}^l_{a,b}$ is defined as

$$\begin{aligned}\bar{\bar{\omega}}^l_{a,b} &\equiv \frac{\partial^2}{\partial s_a \partial \sigma_b} g_l(s_l, \sigma_l)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\ &= \mathbb{E}\left[Z_a(1) U_b(1) \middle| Z_i(0) = 0, U_i(0) = \delta_{i,l} \right]\end{aligned} \tag{5.2.46}$$

Since the two derivatives were not possible to be taken with respect to the same variable, the factorial moments in Eq 5.2.44 are identical to the corresponding ordinary moments.

Next, we evaluate the expectation of $\vec{Z}(n-1)$ and $\vec{U}(n-1)$ on both sides of Eq 5.2.44

$$\begin{aligned}C\Gamma_{a,b}(n) = \mu_l(n-1)(\bar{\bar{\Omega}}^l_{a,b} - \Psi^l_b M^l_a) + M^l_a C_{l,j}(n-1)\Psi^j_b \\ + M^l_a C\Gamma_{l,k}(n-1)\psi^k_b + m^k_a C\Gamma_{l,k}(n-1)\Psi^l_b \\ + \rho_k(n-1)(\bar{\bar{\omega}}^l_{a,b} - \psi^k_b m^k_a) + m^k_a \Gamma_{k,j}(n-1)\psi^j_b\end{aligned} \tag{5.2.47}$$

Since the ordinary moments and factorial moments are identical for $C\Gamma$, Ω and ω,

$$\bar{\bar{\Omega}}^l_{a,b} = \bar{\Omega}^l_{a,b} \tag{5.2.48}$$

$$\bar{\bar{\omega}}^l_{a,b} = \bar{\omega}^l_{a,b} \tag{5.2.49}$$

Eq 5.2.47 is already in the form of ordinary moments. Then we define the central moments directly and simplify Eq 5.2.47 to Eq 5.2.50.

$$\begin{aligned}C\Gamma_{a,b}(n) = \mu_l(n-1)\Omega^l_{a,b} + M^l_a C_{l,j}(n-1)\Psi^j_b \\ + M^l_a C\Gamma_{l,k}(n-1)\psi^k_b + m^k_a C\Gamma_{l,k}(n-1)\Psi^l_b \\ + \rho_k(n-1)\omega^k_{a,b} + m^k_a \Gamma_{k,j}(n-1)\psi^j_b\end{aligned} \tag{5.2.50}$$

where the central moments Ω and ω are defined as

$$\Omega^l_{a,b} \equiv \bar{\Omega}^l_{a,b} - \Psi^l_b M^l_a = \bar{\Omega}^l_{a,b} - \Psi^l_b M^l_a \tag{5.2.51}$$

$$= Cov\left[Z_a(1), U_b(1) \middle| Z_i(0) = \delta_{i,l}, U_i(0) = 0\right] \tag{5.2.52}$$

$$\omega^l_{a,b} \equiv \bar{\omega}^l_{a,b} - \psi^l_b m^l_a = \bar{\omega}^l_{a,b} - \psi^l_b m^l_a \tag{5.2.53}$$

$$= Cov\left[Z_a(1), U_b(1) \middle| Z_i(0) = 0, U_i(0) = \delta_{i,l}\right] \tag{5.2.54}$$

The new central moments are defined because $\bar{\Omega}^l_{a,b}$ always appears with $-M^l_a \Psi^l_b$ and $\bar{\omega}^l_{a,b}$ always exists with $-m^l_a \psi^l_b$ and they can be interpreted as response covariances. $\Omega^l_{a,b}$ is the covariance (2^{nd} order central moment) of the prompt neutron count in region a and the delayed neutron count in region b in response to a prompt neutron started in region l. $\omega^l_{a,b}$ is the covariance (2^{nd} order central moment) of the prompt neutron count in region a and the delayed neutron count in region b in response to a delayed neutron started in region l.

5.2.4 Temporal-Spatial moments

With the equations in the previous sections, moments in the form $\mathbb{E}Z_l(n)$, $\mathbb{E}[Z_l(n)Z_j(n)]$ can be evaluated recursively to any generation n. This section will derive moments in the form of $\mathbb{E}[Z_l(n)Z_j(n+k)]$ where k is the generation lag.

5.2.4.1 Joint moment generating function for neutron count at different generations

Define the joint (factorial) moment generating function of the prompt neutron count at generation n: $\vec{Z}(n)$, the delayed neutron count at generation n: $\vec{U}(n)$, the prompt neutron count at generation $n+k$: $\vec{Z}(n+k)$ and the delayed neutron count at generation $n+k$: $\vec{U}(n+k)$ with corresponding function arguments $\vec{s}$, $\vec{\sigma}$, $\vec{t}$, $\vec{\theta}$ as

$$\begin{aligned} &F_{n,n+k}\left(\vec{Z}(0), \vec{U}(0); \vec{s}, \vec{\sigma}, \vec{t}, \vec{\theta}\right) \\ &\equiv \mathbb{E}\left[\prod_i s_i^{Z_i(n)} \sigma_i^{U_i(n)} t_i^{Z_i(n+k)} \theta_i^{U_i(n+k)} \middle| \vec{Z}(0), \vec{U}(0)\right] \end{aligned} \tag{5.2.55}$$

The joint MGF can be reformulated in the form of the MGF at one generation (Eq 5.2.4) with different arguments, which simplifies the process of taking derivatives and evaluating them to obtain the desired moments. First, we evaluate the expectation in the definition of $F_{n,n+k}\left(\vec{Z}(0),\vec{U}(0);\vec{s},\vec{\sigma},\vec{t},\vec{\theta}\right)$ with condition on the neutron distributions at generation n:

$$\begin{aligned}
&F_{n,n+k}\left(\vec{Z}(0),\vec{U}(0);\vec{s},\vec{\sigma},\vec{t},\vec{\theta}\right)\\
&=\mathbb{E}\left[\mathbb{E}\left[\prod_i s_i^{Z_i(n)}\sigma_i^{U_i(n)}t_i^{Z_i(n+k)}\theta_i^{U_i(n+k)}\middle|\vec{Z}(n),\vec{U}(n)\right]\middle|\vec{Z}(0),\vec{U}(0)\right]\\
&=\mathbb{E}\left[\prod_j s_j^{Z_j(n)}\sigma_j^{U_j(n)}\mathbb{E}\left[\prod_i t_i^{Z_i(n+k)}\theta_i^{U_i(n+k)}\middle|\vec{Z}(n),\vec{U}(n)\right]\middle|\vec{Z}(0),\vec{U}(0)\right]
\end{aligned}\tag{5.2.56}$$

The conditional expectation on n^{th} generation neutrons $\mathbb{E}\left[\prod_i t_i^{Z_i(n+k)}\theta_i^{U_i(n+k)}\middle|\vec{Z}(n),\vec{U}(n)\right]$ takes into consideration all prompt and delayed neutrons in generation n.

The prompt and delayed neutron counts at generation $n+k$ are comprised of neutrons created at generation n and expressed as a sum of independent contributions. Therefore, the conditional expectation on all neutrons at generation n can be rearranged to a product of conditional expectations on individual neutrons at generation n. The product of expectations induced by neutrons from the same region can be grouped into the expectation to the power of neutrons in that region.

$$\begin{aligned}
&\mathbb{E}\left[\prod_i t_i^{Z_i(n+k)}\theta_i^{U_i(n+k)}\middle|\vec{Z}(n),\vec{U}(n)\right]\\
&=\prod_l\left(\mathbb{E}\left[\prod_i t_i^{Z_i(n+k)}\theta_i^{U_i(n+k)}\middle|\vec{Z}(n)=\vec{e}_l,\vec{U}(n)=\vec{0}\right]\right)^{Z_l(n)}\times\\
&\prod_k\left(\mathbb{E}\left[\prod_i t_i^{Z_i(n+k)}\theta_i^{U_i(n+k)}\middle|\vec{Z}(n)=\vec{0},\vec{U}(n)=\vec{e}_k\right]\right)^{U_k(n)}
\end{aligned}\tag{5.2.57}$$

The conditional expectation in Eq 5.2.57 depends on k but not on n. n can be set to be 0 without loss of generality, which makes the two conditional expectations match the moment generating function of prompt and delayed neutrons at generation

k in response to an individual prompt or delayed source neutron at generation 0 respectively (Eq 5.2.5, Eq 5.2.6). That is

$$\mathbb{E}\left[\prod_i t_i^{Z_i(n+k)}\theta_i^{U_i(n+k)}\middle|\vec{Z}(n),\vec{U}(n)\right]=\prod_l f_l^{(k)}(\vec{t},\vec{\theta})^{Z_l(n)}\prod_j g_j^{(k)}(\vec{t},\vec{\theta})^{U_j(n)} \tag{5.2.58}$$

where

$$f_l^{(k)}(\vec{s},\vec{\sigma})\equiv\mathbb{E}\left[\prod_i s_i^{Z_i(k)}\sigma_i^{U_i(k)}\middle|Z_i(0)=\delta_{i,l},U_i(0)=0\right] \tag{5.2.59}$$

$$g_l^{(k)}(\vec{s},\vec{\sigma})\equiv\mathbb{E}\left[\prod_i s_i^{Z_i(k)}\sigma_i^{U_i(k)}\middle|Z_i(0)=0,U_i(0)=\delta_{i,l}\right] \tag{5.2.60}$$

Substitute the simplified conditional expectation of $\vec{Z}(n)$ and $\vec{U}(n)$ (Eq 5.2.58) into the rearranged definition of $F_{n,n+k}\left(\vec{Z}(0),\vec{U}(0);\vec{s},\vec{\sigma},\vec{t},\vec{\theta}\right)$ (Eq 5.2.56).

$$\begin{aligned}
&F_{n,n+k}\left(\vec{Z}(0),\vec{U}(0);\vec{s},\vec{\sigma},\vec{t},\vec{\theta}\right)\\
&=\mathbb{E}\left[\prod_i s_i^{Z_i(n)}\sigma_i^{U_i(n)}\prod_l f_l^{(k)}(\vec{t},\vec{\theta})^{Z_l(n)}\prod_j g_j^{(k)}(\vec{t},\vec{\theta})^{U_j(n)}\middle|\vec{Z}(0),\vec{U}(0)\right]\\
&=\mathbb{E}\left[\prod_l s_l^{Z_l(n)}f_l^{(k)}(\vec{t},\vec{\theta})^{Z_l(n)}\prod_j \sigma_j^{U_j(n)}g_j^{(k)}(\vec{t},\vec{\theta})^{U_j(n)}\middle|\vec{Z}(0),\vec{U}(0)\right]
\end{aligned} \tag{5.2.61}$$

Eq 5.2.61 is recognized as the moment generating function of $\vec{Z}(n)$ and $\vec{U}(n)$ given initial prompt and delayed neutrons characterized by $\vec{Z}(0)$ and $\vec{U}(0)$ with new function arguments $\vec{u}(k)$ and $\vec{\xi}(k)$, whose vector elements are defined as

$$u_i(k)=s_i f_i^{(k)}(\vec{t},\vec{\theta}) \tag{5.2.62}$$

$$\xi_i(k)=\sigma_i g_i^{(k)}(\vec{t},\vec{\theta}) \tag{5.2.63}$$

In summary, the joint moment generating function $F_{n,n+k}\left(\vec{Z}(0),\vec{U}(0);\vec{s},\vec{\sigma},\vec{t},\vec{\theta}\right)$ satisfies the functional equation

$$F_{n,n+k}\left(\vec{Z}(0),\vec{U}(0);\vec{s},\vec{\sigma},\vec{t},\vec{\theta}\right)=F_n\left(\vec{Z}(0),\vec{U}(0);\vec{u}(k),\vec{\xi}(k)\right) \tag{5.2.64}$$

with $\vec{u}$ and $\vec{\xi}$ defined in Eq 5.2.62, Eq 5.2.63.

5.2.4.2 Expectation of prompt neutron pair at different generations

The derivative of the joint moment generating function derived previously (Eq 5.2.64) is related to the temporal-spatial moment $\mathbb{E}\left[Z_a(n)Z_b(n+k)\right]$, which is necessary to evaluate the generation-to-generation correlation between generation n and $n+k$.

By definition of $F_{n,n+k}\left(\vec{Z}(0),\vec{U}(0);\vec{s},\vec{\sigma},\vec{t},\vec{\theta}\right)$, the temporal-spatial moment $\mathbb{E}\left[Z_a(n)Z_b(n+k)\right]$ can be evaluated as:

$$\mathbb{E}\left[Z_a(n)Z_b(n+k)\right] = \frac{\partial^2}{\partial s_a \partial t_b} F_{n,n+k}\left(\vec{Z}(0),\vec{U}(0);\vec{s},\vec{\sigma},\vec{t},\vec{\theta}\right)\Big|_{\vec{s}=\vec{\sigma}=\vec{t}=\vec{\theta}=\vec{1}} \tag{5.2.65}$$

The functional equation of $F_{n,n+k}\left(\vec{Z}(0),\vec{U}(0);\vec{s},\vec{\sigma},\vec{t},\vec{\theta}\right)$ (Eq 5.2.64) enables calculating the temporal-spatial moments as

$$\mathbb{E}\left[Z_a(n)Z_b(n+k)\right] = \frac{\partial^2}{\partial s_a \partial t_b} F_n\left(\vec{Z}(0),\vec{U}(0);\vec{u}(k),\vec{\xi}(k)\right)\Big|_{\vec{s}=\vec{\sigma}=\vec{t}=\vec{\theta}=\vec{1}} \tag{5.2.66}$$

The derivatives in Eq 5.2.66 can be evaluated by expanding the derivative with respect to $\vec{s}$ into the derivative of the joint moment generating function with respect to the new arguments $\vec{u}$ and $\vec{\xi}$ and the derivative of $\vec{u}$ and $\vec{\xi}$ with respect to $\vec{s}$.

According to Eq 5.2.62 and Eq 5.2.63 the derivatives of $\vec{u}$ and $\vec{\xi}$ with respect to $\vec{s}$ and $\vec{t}$ are

$$\frac{\partial u_i(k)}{\partial s_a} = \delta_i^a f_i^{(k)}(\vec{t},\vec{\theta}) \tag{5.2.67}$$

$$\frac{\partial u_i(k)}{\partial t_a} = s_i \frac{\partial}{\partial t_a} f_i^{(k)}(\vec{t},\vec{\theta}) \tag{5.2.68}$$

$$\frac{\partial \xi_i(k)}{\partial s_a} = 0 \tag{5.2.69}$$

$$\frac{\partial \xi_i(k)}{\partial s_a} = \sigma_i \frac{\partial}{\partial t_a} f_i^{(k)}(\vec{t},\vec{\theta}) \tag{5.2.70}$$

The Jacobian in Eqs 5.2.67 - 5.2.70 reduces Eq 5.2.66 to

$$
\begin{aligned}
&\mathbb{E}\left[Z_a(n)Z_b(n+k)\right] \\
&= \sum_j \left(\frac{\partial^2}{\partial u_a \partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right) \right) f_a^{(k)}(\vec{t},\vec{\theta}) s_j \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t},\vec{\theta}) \\
&+ \left(\frac{\partial}{\partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right) \right) \delta_j^a \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t},\vec{\theta}) \\
&+ \left(\frac{\partial^2}{\partial u_a \partial \xi_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right) \right) f_a^{(k)}(\vec{t},\vec{\theta}) \sigma_j \frac{\partial}{\partial t_b} g_j^{(k)}(\vec{t},\vec{\theta}) \Big|_{\vec{s}=\vec{\sigma}=\vec{t}=\vec{\theta}=\vec{1}}
\end{aligned}
\tag{5.2.71}
$$

The derivation of Eq 5.2.71 by manipulating the functional equation of the joint MGF is included in Appendix A1.4, Eq A1.4.36.

We now express the temporal-spatial moment $\mathbb{E}\left[Z_a(n)Z_b(n+k)\right]$ in terms of other moments by evaluating the moment generating functions and their derivatives in Eq 5.2.71.

When the vectors $\vec{s} = \vec{\sigma} = \vec{t} = \vec{\theta} = \vec{1}$, the MGF's are evaluated to be 1.

$$
f_a^{(k)}(\vec{t},\vec{\theta})\Big|_{\vec{t}=\vec{\theta}=\vec{1}} = 1 \tag{5.2.72}
$$

The derivative of $f_j^{(k)}(\vec{t},\vec{\theta})$ with respect to t_b is the expectation of the prompt neutron count in region b after k generations initialized by one prompt neutron in region j.

$$
\frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t},\vec{\theta})\Big|_{\vec{t}=\vec{\theta}=\vec{1}} = \mu_b^{(j)}(k) \tag{5.2.73}
$$

where the superscript denotes the prompt neutron source region.

The derivative of $g_j^{(k)}(\vec{t},\vec{\theta})$ with respect to t_b is the expectation of the prompt neutron count in region b after k generations initialized by one delayed neutron in region j

$$
\frac{\partial}{\partial t_b} g_j^{(k)}(\vec{t},\vec{\theta})\Big|_{\vec{t}=\vec{\theta}=\vec{1}} = \mu_b^{\underline{(j)}}(k) \tag{5.2.74}
$$

where the underlined superscript denotes the delayed neutron source region.

The new arguments of the joint MGF also evaluate to 1.

$$u_a|_{\vec{s}=\vec{t}=\vec{\theta}=\vec{1}} = s_a f_a^{(k)}(\vec{t},\vec{\theta})\Big|_{\vec{s}=\vec{t}=\vec{\theta}=\vec{1}} = 1 \tag{5.2.75}$$

$$\xi_a|_{\vec{s}=\vec{t}=\vec{\theta}=\vec{1}} = \sigma_a g_a^{(k)}(\vec{t},\vec{\theta})\Big|_{\vec{s}=\vec{t}=\vec{\theta}=\vec{1}} = 1 \tag{5.2.76}$$

The derivative of $F_n(\vec{u},\vec{\xi})$ with respect to u_j is the expectation of the number of prompt neutrons in region j after n generations

$$\frac{\partial}{\partial u_j} F_n\left(\vec{u},\vec{\xi}\right)\Big|_{\vec{u}=\vec{\xi}=\vec{1}} = \mu_j(n) \tag{5.2.77}$$

The derivative of $F_n(\vec{u},\vec{\xi})$ with respect to u_j and u_a is the expectation of prompt neutron pairs in region a and region j after n generations

$$\left(\frac{\partial^2}{\partial u_a \partial u_j} F_n\left(\vec{u},\vec{\xi}\right)\right)\Big|_{\vec{u}=\vec{\xi}=\vec{1}} = \not{C}_{a,j}(n) \tag{5.2.78}$$

The derivative of $F_n(\vec{u},\vec{\xi})$ with respect to ξ_j and u_a is the expectation of prompt and delayed neutron pairs in region a and region j after n generations

$$\frac{\partial^2}{\partial u_a \partial \xi_j} F_n\left(\vec{u},\vec{\xi}\right)\Big|_{\vec{u}=\vec{\xi}=\vec{1}} = C\not{\Gamma}_{a,j}(n) \tag{5.2.79}$$

This simplifies Eq 5.2.71 to

$$\begin{aligned}&\mathbb{E}\left[Z_a(n)Z_b(n+k)\right]\\ &= \sum_j \not{C}_{a,j}(n)\mu_b^{(j)}(k) + \mu_j(n)\delta_a^j\mu_b^{(j)}(k) + C\not{\Gamma}_{a,j}(n)\mu_b^{(j)}(k)\end{aligned} \tag{5.2.80}$$

Taking into consideration the relation between the factorial and ordinary moments of prompt neutron pairs (Eq 5.2.18) leads to

$$\mathbb{E}\left[Z_a(n)Z_b(n+k)\right] = \sum_j C_{a,j}(n)\mu_b^{(j)}(k) + C\not{\Gamma}_{a,j}(n)\mu_b^{(j)}(k) \tag{5.2.81}$$

Eq 5.2.81 reveals that the expectation of prompt neutron pairs across different

generations is attributed to two parts. 1) transport process of prompt neutrons in the prompt neutron pair through k generations, and 2) transport process of delayed neutrons in the prompt-delayed neutron pair through k generations.

5.2.5 Surrogate to Monte Carlo simulation

The previous section derived the recursive formula for the expectations of the prompt neutron count in each region $\mu_a(n)$, the delayed neutron count in each region $\rho_a(n)$, the prompt neutron pairs in every two regions $C_{a,b}(n)$, the delayed neutron pairs in every two regions $\Gamma_{a,b}(n)$ and the prompt and delayed neutron pairs in every two regions $C\Gamma_{a,b}(n)$ in terms of the response moments M (prompt neutron count in response to a prompt neutron), m (prompt neutron count in response to a delayed neutron), Ψ (delayed neutron count in response to a prompt neutron), ψ (delayed neutron count in response to a delayed neutron), V (covariance of prompt neutron pairs in response to a prompt neutron), v (covariance of prompt neutron pairs in response to a delayed neutron), Λ (covariance of delayed neutron pairs in response to a prompt neutron), λ (covariance of delayed neutron pairs in response to a delayed neutron), Ω (covariance of prompt and delayed neutron pairs in response to a prompt neutron), ω (covariance of prompt and delayed neutron pairs in response to a delayed neutron). This section discusses the relation between these response moments and the Monte Carlo simulation the DMBP method is approximating.

The response moments denoted with upper case letters M,Ψ, V, Λ and Ω are responses to a prompt neutron. Therefore they are related to the spatial moments V and M derived in section 3.3.4.1 for the original Multitype Branching Process method. For ease, the corresponding moments V and M in section 3.3.4.1 are denoted as V_0 and M_0 in this section.

The response moments denoted with lower case letters m, ψ, v, λ and ω are responses to a delayed neutron. Therefore these moments can be characterized by free parameters that control the delayed simulation scheme.

5.2.5.1 Response to a delayed neutron

According the source update scheme with delayed neutrons described in section 5.2.1, a fraction p of neutrons in the delayed source bank will be sampled into the prompt source bank for the next generation. For mathematical ease in building the DMBP model, a slightly different scheme is modelled. Each neutron in the delayed source bank is sampled in the prompt source bank with probability p and probability $1-p$ into the delayed source bank. Under this sampling scheme, not an exact fraction p of neutrons will necessarily be sampled . Instead the fraction is a random variable with mean p and variance $\frac{p(1-p)}{(1-\alpha)s}$. We assume s is large enough to neglect the variance. In addition, similar to the MBP model, the DMBP model does not treat the neutron number normalization. The sizes of prompt source bank and delayed source bank in the DMBP model fluctuates freely but are kept stable by enforcing the system to be critical and maintaining enough neutrons per generation.

The sampling probability p could be generalized to be region dependent. Given a delayed neutron in region a, it could be sampled as a prompt neutron source for the next generation with probability p_a and as a delayed neutron source for the next generation with probability $1-p_a$. Note that the region dependent sampling probabilities are used to derive a general theory but are not tested in this work.

This section first evaluates the moment generating function of $\vec{Z}(1)$ and $\vec{U}(1)$ given one delayed neutron as defined in Eq 5.2.82.

$$g_l(\vec{s},\vec{\sigma}) = \mathbb{E}\left[\prod_j s_j^{Z_j(1)} \prod_k \sigma_k^{U_k(1)} \middle| Z_i(0)=0, U_i(0)=\delta_{i,l}\right] \tag{5.2.82}$$

Then we evaluate the response moments by taking the derivative of the moment generating function.

We use the random variable ξ_l to denote whether the delayed neutron in region l is sampled into the prompt source bank for the next generation. ξ_l is a Bernoulli random variable with parameter p_l. It is clear that $Z_a(1) = \xi_l \delta_{l,a}$ and $U_a(1) = (1-\xi_l)\delta_{l,a}$.

The probabilities then follow in Eq 5.2.83.

$$\begin{aligned}\mathbb{P}\left(Z_a(1)=1\,|Z_i(0)=0,U_i(0)=\delta_{i,l}\right)&=p_a\delta_{l,a}\\ \mathbb{P}\left(U_a(1)=1\,|Z_i(0)=0,U_i(0)=\delta_{i,l}\right)&=(1-p_a)\delta_{l,a}\end{aligned}\tag{5.2.83}$$

Therefore

$$\begin{aligned}g_l(\vec{s},\vec{\sigma})&=\mathbb{E}\left[\prod_j s_j^{Z_j(1)}\sigma_j^{U_j(1)}\,\middle|\,Z_i(0)=0,U_i(0)=\delta_{i,l}\right]\\ &=\sum_a s_a p_a\delta_{l,a}+\sum_a\sigma_a(1-p_a)\delta_{l,a}\\ &=s_l p_l+\sigma_l(1-p_l)\end{aligned}\tag{5.2.84}$$

With the MGF in Eq 5.2.84 and the definition of the moments in terms of $g_l(\vec{s},\vec{\sigma})$, the moments m,ψ,v,λ, ω can be written as follows.

5.2.5.1.1 First order response moments

First order derivatives of $g_l(\vec{s},\vec{\sigma})$ with respect to s_a (or σ_a) are non-zero only when $a=l$.

$$m_a^l\equiv\frac{\partial}{\partial s_a}g_l(\vec{s},\vec{\sigma})\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}}=p_a\delta_{l,a}\tag{5.2.85}$$

$$\psi_a^l\equiv\frac{\partial}{\partial \sigma_a}g_l(\vec{s},\vec{\sigma})\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}}=(1-p_a)\delta_{l,a}\tag{5.2.86}$$

In addition to the definition in terms of derivatives, more intuitive interpretation of moments m and ψ comes from the equivalent definition in the form of conditional expectations. Given a delayed neutron from region l, the induced neutron for the next generation can be either a prompt neutron in the same region l with probability p_l or a delayed neutron in the same region l with probability $1-p_l$. Consequentially the expectations are equal to the corresponding probability only if the region is the same from one generation to the next.

5.2.5.1.2 Second order response moments

Eq 5.2.84 also implies that all second order derivatives vanish because it is a linear function of the arguments $\vec{s}$ and $\vec{\sigma}$.

$$\bar{\nabla}^l_{a,b} = \bar{\lambda}^l_{a,b} = \bar{\omega}^l_{a,b} = 0 \tag{5.2.87}$$

With the relation between factorial moments and the (ordinary) central moments given in Eq 5.2.26 and Eq 5.2.30, the central response moment of prompt neutron pairs to a delayed neutron, v is evaluated to be

$$\begin{aligned} v^l_{a,b} &= \bar{\nabla}^l_{a,b} + m^l_a \delta_{a,b} - m^l_a m^l_b \\ &= p_a \delta^l_a \delta_{a,b} - p_a p_b \delta^l_a \delta^l_b = p_a (1 - p_b) \delta^l_a \delta_{a,b} \end{aligned} \tag{5.2.88}$$

With the relation between factorial moments and the (ordinary) central moments given in Eq 5.2.39 and Eq 5.2.43, the central response moment of the delayed neutron pairs to a delayed neutron, λ is evaluated to be

$$\begin{aligned} \lambda^l_{a,b} &= \bar{\lambda}^l_{a,b} + \psi^l_a \delta_{a,b} - \psi^l_a \psi^l_b \\ &= (1 - p_a) \delta^l_a \delta_{a,b} - (1 - p_a)(1 - p_b) \delta^l_a \delta^l_b = (1 - p_a) p_b \delta^l_a \delta_{a,b} \end{aligned} \tag{5.2.89}$$

With the relation between factorial moments and the (ordinary) central moments given in Eq 5.2.49 and Eq 5.2.54, the central response moment of prompt and delayed neutron pairs to a delayed neutron, ω is evaluated to be

$$\omega^l_{a,b} = \bar{\omega}^l_{a,b} - m^l_a \psi^l_b = -p_a (1 - p_b) \delta^l_a \delta^l_b \tag{5.2.90}$$

In addition to the definition in terms of derivatives, a more intuitive interpretation of moments v, λ and ω comes from the equivalent definition in the form of conditional expectations. Given a delayed neutron from region l, the induced neutron for the next generation must remain in the same region. Therefore, the number of prompt neutrons in region l is a Bernoulli random variable with parameter p_l. Consequentially, $v^l_{l,l} =$

$p_l(1-p_l)$. Similarly, the number of delayed neutrons in region l is a Bernoulli random variable with parameter p_l and we have $\lambda^l_{l,l} = (1-p_l)p_l$. Since the neutron cannot be sampled as both prompt and delayed simultaneously for the next generation, the covariance $\omega^l_{l,l}$ only contains the negative product of the expectation term, $-m^l_l\psi^l_l = -p_l(1-p_l)$.

5.2.5.2 Response to a prompt neutron

According the source update scheme with delayed neutrons described in section 5.2.1, a fraction f of neutrons in the fission bank $\vec{Z}^*(n)$ generated from the prompt source bank $\vec{Z}(n)$ will be sampled into the prompt source bank for the next generation. Similarly to the sampling from the delayed source bank discussed in section 5.2.5.1, a slightly different scheme is modelled where sampling is performed on each individual neutron. This yield an expected value of f for the sampled fraction with a variance $\frac{f(1-f)}{s_f}$. We assume s_f is large enough to neglect the variance. In addition, the neutron number normalization is not incorporated into the model.

This section first evaluates the moment generating function of $\vec{Z}(1)$ and $\vec{U}(1)$ given one prompt neutron as defined in Eq 5.2.91.

$$f_l(\vec{s},\vec{\sigma}) = \mathbb{E}\left[\prod_j s_j^{Z_j(1)} \prod_k \sigma_k^{U_k(1)} \middle| Z_i(0) = \delta_{i,l}, U_i(0) = 0\right] \tag{5.2.91}$$

Then we evaluate the response moments by taking the derivative of the moment generating function.

The sampling probability f can be generalized to be region dependent. After the transport process and branching from fission, one prompt neutron induces new neutrons specified by the neutron count vector $\vec{Z}^*(1)$. This process is characterized by the moment generating function $f_a(\vec{s})$ defined in Eq 3.3.22. Then for each induced prompt neutron in region b, it will be sampled as a prompt neutron source for the next generation with probability f_b and as a delayed neutron source for the next generation with probability $1-f_b$. This process can be described similarly to the response to a delayed neutron as discussed in section 5.2.5.1.

The moment generating function can be evaluated by first taking the expectation in the definition conditioned on $\vec{Z}^*(1)$ then evaluating the expectation as a function of $\vec{Z}^*(1)$ with respect to the distribution of $\vec{Z}^*(1)$ conditioned on $\vec{Z}(0)$ and $\vec{U}(0)$. This shows the relation between the moment generation function of prompt and delayed neutron count in response to a prompt neutron $f_l(\vec{s},\vec{\sigma})$ and the moment generating function of prompt neutron count in response to a prompt neutron ($f_l(\vec{s})$ discussed in section 3.3).

First, we evaluate the conditional expectation in the definition of $f_l(\vec{s},\vec{\sigma})$ in Eq 5.2.5 with the intermediate neutrons ($\vec{Z}^*(1)$) in the fission bank after the transport process but before the source sampling: $\mathbb{E}\left[\prod_i s_i^{Z_i(1)}\sigma_i^{U_i(1)}\middle|\vec{Z}^*(1)\right]$.

Since the sampling process does not change the neutron type (region), the product index i must satisfy $Z_i^*(1)>0$, that is

$$\mathbb{E}\left[\prod_i s_i^{Z_i(1)}\sigma_i^{U_i(1)}\middle|\vec{Z}^*(1)\right] = \mathbb{E}\left[\prod_{i\in\{i':Z_{i'}^*(1)>0\}} s_i^{Z_i(1)}\sigma_i^{U_i(1)}\middle|\vec{Z}^*(1)\right] \tag{5.2.92}$$

And within this set, the sum of prompt and delayed neutrons is equal to the count prior to sampling, i.e., $U_i(1)=Z_i^*(1)-Z_i(1)$. Since the sampling in each region i is independent of each other, the expectation of the product can be expanded into a product of expectations.

$$\mathbb{E}\left[\prod_i s_i^{Z_i(1)}\sigma_i^{U_i(1)}\middle|\vec{Z}^*(1)\right] = \prod_{i\in\{i':Z_{i'}^*(1)>0\}} \mathbb{E}\left[s_i^{Z_i(1)}\sigma_i^{Z_i^*(1)-Z_i(1)}\middle|\vec{Z}^*(1)\right] \tag{5.2.93}$$

$Z_i(1)$ can be expressed as a sum of sampling results for each neutron within region i:

$$Z_i(1)=\sum_{\alpha=1}^{Z_i^*(1)}\xi_i^{(\alpha)} \tag{5.2.94}$$

where $\xi_i^{(\alpha)}$ indicates whether the α^{th} prompt neutron in region i is sampled as prompt ($\xi_i^{(\alpha)}=1$) or delayed ($\xi_i^{(\alpha)}=0$) for the next generation. The summation in Eq 5.2.94 becomes a product in Eq 5.2.93 and the independence of the sampling among the

neutrons within the same region simplifies the expectation products into product of expectations:

$$\mathbb{E}\left[\prod_i s_i^{Z_i(1)}\sigma_i^{U_i(1)}\middle|\vec{Z}^*(1)\right] = \prod_{i\in\{i':Z^*_{i'}(1)>0\}}\prod_{\alpha=1}^{Z_i^*(1)}\mathbb{E}\left[s_i^{\xi_i^{(\alpha)}}\sigma_i^{1-\xi_i^{(\alpha)}}\middle|\vec{Z}^*(1)\right] \tag{5.2.95}$$

According to the delayed source update scheme, $\mathbb{P}(\xi_i^{(\alpha)}=1)=f_i$, which evaluates the expectation of $s_i^{\xi_i^{(\alpha)}}\sigma_i^{1-\xi_i^{(\alpha)}}$ to be $s_if_i+(1-f_i)\sigma_i$. In addition, all the $Z_i^*(1)$ neutrons in region i are sampled following the same distribution,

the expectations in Eq 5.2.95 do not depend on α and are thus identical, the product of $Z_i^*(1)$ identical terms is simplified to an expectation of one term to the power of $Z_i^*(1)$:

$$\mathbb{E}\left[\prod_i s_i^{Z_i(1)}\sigma_i^{U_i(1)}\middle|\vec{Z}^*(1)\right] = \prod_{i\in\{i':Z^*_{i'}(1)>0\}}(s_if_i+(1-f_i)\sigma_i)^{Z_i^*(1)} \tag{5.2.96}$$

Eq 5.2.96 still holds even when $Z_i^*(1)=0$, therefore the product constraints $i\in\{i':Z^*_{i'}(1)>0\}$ can be removed without loss of generality.

Finally, we evaluate Eq 5.2.96, which can be viewed as a deterministic function of the random vector $\vec{Z}^*(1)$, with respect to the distribution of $\vec{Z}^*(1)$ given $\vec{Z}(0)$:

$$f_l(\vec{s},\vec{\sigma}) = \mathbb{E}\left[\prod_i(s_if_i+(1-f_i)\sigma_i)^{Z_i^*(1)}\middle|Z_i(0)=\delta_{i,l},U_i(0)=0\right] \tag{5.2.97}$$

Eq 5.2.97 is immediately recognized as the moment generating function of $\vec{Z}^*(1)$ given a prompt neutron source at region l with a new argument vector $\vec{s'}$, defined as

$$s_i' = f_is_i+(1-f_i)\sigma_i \tag{5.2.98}$$

which trivially leads to

$$\frac{\partial}{\partial s_a} s'_i = \delta_i^a f_i \tag{5.2.99}$$

$$\frac{\partial}{\partial \sigma_a} s'_i = \delta_i^a (1 - f_i) \tag{5.2.100}$$

To summarize, in response to a prompt source neutron, the moment generating function of a prompt and delayed neutron is related to that of a prompt neutron without the delayed neutron sampling process through:

$$f_l(\vec{s}, \vec{\sigma}) = f_l(\vec{s'}) \tag{5.2.101}$$

with $\vec{s'}$ defined in Eq 5.2.98.

With the MGF relation in Eq 5.2.101, the moments M,Ψ,$\bar{\nabla}$,$\bar{\Lambda}$, $\bar{\Omega}$ can be written in terms of M_0 and $\bar{\nabla}_0$ as follows.

5.2.5.2.1 First order response moments

First order derivatives of $f_l(\vec{s}, \vec{\sigma})$ evaluated at $\vec{s} = \vec{\sigma} = \vec{1}$ gives the first order response moments.

$$M_a^l \equiv \frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma}) \bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} = {M_0}_a^l f_a \tag{5.2.102}$$

$$\Psi_a^l \equiv \frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma}) \bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} = {M_0}_a^l (1 - f_a) \tag{5.2.103}$$

The derivations of Eq 5.2.102 and Eq 5.2.103 are detailed in Eq A1.4.21 and Eq A1.4.22. Conceptually, one neutron in region l in the prompt source bank contributes ${M_0}_a^l$ neutrons in region a in the fission bank. Among these neutrons, ${M_0}_a^l f_a$ are sampled into the prompt source bank and ${M_0}_a^l (1 - f_a)$ are sampled into the delayed source bank.

5.2.5.2.2 Second order response moments

Second order derivatives of $f_l(\vec{s}, \vec{\sigma})$ evaluated at $\vec{s} = \vec{\sigma} = \vec{1}$ gives the second order response moments.

The expected number of pairs of prompt neutrons in region a and b in response to

one prompt neutron in region l is derived in Eq A1.4.23 and the final result is shown in Eq 5.2.104

$$\bar{\mathbb{V}}^l_{a,b} \equiv \frac{\partial^2}{\partial s_a \partial s_b} f_l(\vec{s},\vec{\sigma})\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} = \bar{\mathbb{V}}^l_{0_{a,b}} f_a f_b \tag{5.2.104}$$

Inserting in the relation between central and factorial moments as described in Eq 5.2.29, yields Eq 5.2.25,

$$\begin{aligned} V^l_{a,b} &= \bar{\mathbb{V}}^l_{a,b} + M^l_a \delta_{a,b} - M^l_a M^l_b \\ &= \bar{\mathbb{V}}^l_{0_{a,b}} f_a f_b + M^l_a \delta_{a,b} - M^l_a M^l_b \\ &= (V^l_{0_{a,b}} + M^l_{0_a} M^l_{0_b} - M^l_{0_a}\delta_{a,b}) f_a f_b + M^l_a \delta_{a,b} - M^l_a M^l_b \\ &= V^l_{0_{a,b}} f_a f_b + M^l_{0_a}\delta_{a,b} f_a (1-f_b) \end{aligned} \tag{5.2.105}$$

where the simplification in the last equality is reached via the relation between M and M_0 in Eq 5.2.102.

The expected number of pairs of delayed neutrons in region a and b in response to one prompt neutron in region l is derived in Eq A1.4.25 and the final result is shown in Eq 5.2.106

$$\bar{\mathbb{\Lambda}}^l_{a,b} = \bar{\mathbb{V}}^l_{0_{a,b}} (1-f_a)(1-f_b) \tag{5.2.106}$$

Inserting the relation between central and factorial moments as described in Eq 5.2.42, gives Eq 5.2.38,

$$\begin{aligned} \Lambda^l_{a,b} &= \bar{\mathbb{\Lambda}}^l_{a,b} + \Psi^l_a \delta_{a,b} - \Psi^l_a \Psi^l_b \\ &= \bar{\mathbb{V}}^l_{0_{a,b}} (1-f_a)(1-f_b) + \Psi^l_a \delta_{a,b} - \Psi^l_a \Psi^l_b \\ &= (V^l_{0_{a,b}} + M^l_{0_a} M^l_{0_b} - M^l_{0_a}\delta_{a,b})(1-f_a)(1-f_b) + \Psi^l_a \delta_{a,b} - \Psi^l_a \Psi^l_b \\ &= V^l_{0_{a,b}} (1-f_a)(1-f_b) + M^l_{0_a}\delta_{a,b}(1-f_a) f_b \end{aligned} \tag{5.2.107}$$

where the simplification in the last equality is reached via the relation between Ψ and M_0 in Eq 5.2.103.

The expected number of pairs of prompt and delayed neutrons in region a and b in response to one prompt neutron in region l is derived in Eq A1.4.24 and the final

result is in Eq 5.2.108

$$\bar{\Omega}^l_{a,b} = \bar{V}^{\,l}_{0_{a,b}} f_a(1-f_b) \tag{5.2.108}$$

Inserting the relation between central and factorial moments as described in Eq 5.2.52, yields Eq 5.2.48,

$$\begin{aligned}
\Omega^l_{a,b} &= \bar{\Omega}^l_{a,b} - M^l_a \Psi^l_b \\
&= \bar{V}^{\,l}_{0_{a,b}} f_a(1-f_b) - M^l_a \Psi^l_b \\
&= (V^l_{0_{a,b}} + M^l_{0_a} M^l_{0_b} - M^l_{0_a} \delta_{a,b}) f_a(1-f_b) - M^l_a \Psi^l_b \\
&= V^l_{0_{a,b}} f_a(1-f_b) - M^l_{0_a} \delta_{a,b} f_a(1-f_b)
\end{aligned} \tag{5.2.109}$$

where the simplification in the last equality is reached via the relation between M and M_0 in Eq 5.2.102 and the relation between Ψ and M_0 in Eq 5.2.103.

5.2.5.3 Simplification of moments evolution

With the spatial response moments evaluated for the special delayed neutron sampling process, the evolution (recursive formula) for the moments μ,ρ,V,v,Λ,λ, Ω, ω can be simplified correspondingly.

Inserting Eq 5.2.102 and Eq 5.2.85 into Eq 5.2.13 gives

$$\mu_a(n) = \rho_a(n-1)p_a + \mu_l(n-1)M^l_a \tag{5.2.110}$$

Inserting Eq 5.2.103 and Eq 5.2.86 into Eq 5.2.14 gives

$$\rho_a(n) = \rho_a(n-1)(1-p_a) + \mu_l(n-1)\Psi^l_a \tag{5.2.111}$$

Inserting Eq 5.2.102, Eq 5.2.105, Eq 5.2.85 and Eq 5.2.88 into Eq 5.2.28 gives

$$\begin{aligned} C_{a,b}(n) = {} & \mu_l(n-1)V^l_{0a,b}f_af_b + M^l_{0a}C_{l,j}(n-1)M^j_{0b}f_af_b \\ & + M^l_{0a}C\Gamma_{l,b}(n-1)f_ap_b + C\Gamma_{l,a}(n-1)M^l_{0b}p_af_b \\ & + \Gamma_{a,b}(n-1)p_ap_b \\ & + \mu_l(n-1)M^l_{0a}\delta_{a,b}f_a(1-f_b) + \rho_a(n-1)\delta_{a,b}p_a(1-p_b) \end{aligned} \tag{5.2.112}$$

Inserting Eq 5.2.103, Eq 5.2.107, Eq 5.2.86 and Eq 5.2.89 into Eq 5.2.41 gives

$$\begin{aligned} \Gamma_{a,b}(n) = {} & \mu_l(n-1)V^l_{0a,b}(1-f_a)(1-f_b) + M^l_{0a}C_{l,j}(n-1)M^j_{0b}(1-f_a)(1-f_b) \\ & + M^l_{0a}C\Gamma_{l,b}(n-1)(1-f_a)(1-p_b) + C\Gamma_{b,l}(n-1)M^l_{0b}(1-p_a)(1-f_b) \\ & + \Gamma_{a,b}(n-1)(1-p_a)(1-p_b) \\ & + \mu_l(n-1)M^l_{0a}\delta_{a,b}(1-f_a)f_b + \rho_a(n-1)\delta_{a,b}(1-p_a)p_b \end{aligned} \tag{5.2.113}$$

Inserting Eq 5.2.102, Eq 5.2.103, Eq 5.2.85, Eq 5.2.86, Eq 5.2.109 and Eq 5.2.90 into Eq 5.2.50 gives

$$\begin{aligned} C\Gamma_{a,b}(n) = {} & \mu_l(n-1)V^l_{0a,b}f_a(1-f_b) + M^l_{0a}C_{l,j}(n-1)M^j_{0b}f_a(1-f_b) \\ & + M^l_{0a}C\Gamma_{l,b}(n-1)f_a(1-p_b) + C\Gamma_{l,a}(n-1)M^l_{0b}p_a(1-f_b) \\ & + \Gamma_{a,b}(n-1)p_a(1-p_b) \\ & - \mu_l(n-1)M^l_{0a}\delta_{a,b}f_a(1-f_b) - \rho_a(n-1)\delta_{a,b}p_a(1-p_b) \end{aligned} \tag{5.2.114}$$

5.2.5.4 Onset of the delayed scheme

The last part of Monte Carlo simulation to be incorporated into the DMBP model is the onset of active generations where the delayed scheme starts. This section discusses how to initialize the moments μ,ρ,C,$C\Gamma$ and Γ for recursive evaluation. Moment initialization was not an issue for the MBP model developed previously because the separation between inactive and active generations is only for tally purposes and does not affect how the neutrons are transported in the simulation or equivalently does not affect how the moments are recursively evaluated. However, under the delayed

neutron scheme, the beginning of active generations also splits the source bank into prompt and delayed source banks and thus initializes the moments.

Denote the source bank before the $\alpha/1-\alpha$ splitting by $\vec{Z}(0)$. Then following the convention in section 3.3, in the source bank, the expected number of neutrons in region a is $\mu_a(0)$ and the expected number of neutron pairs across region a and b is $\complement_{a,b}(0)$. The question is, given $\mu_a(0)$ and $\complement_{a,b}(0)$ (or its ordinary moment counterpart $C_{a,b}(0)$, how do we initialize the moments $\mu_a(1)$,$\rho_a(1)$, $C_{a,b}(1)$, $C\Gamma_{a,b}(1)$ and $\Gamma_{a,b}(1)$.

Suppose that each neutron in the source bank is sampled into the prompt source bank with probability α and into the delayed source bank with probability $1-\alpha$. We then use $\xi_a^{(i)}$ to denote whether the i^{th} neutron in region a is sampled into the prompt source bank. Clearly $\xi_a^{(i)}$ is a Bernoulli random variable with parameter α. $\vec{Z}(1)$ and $\vec{U}(1)$ can be expressed in terms of $\xi_a^{(i)}$ as shown in Eq 5.2.115 and Eq 5.2.116.

$$Z_a(1) = \sum_{i=1}^{Z_a(0)} \xi_a^{(i)} \tag{5.2.115}$$

$$U_a(1) = \sum_{i=1}^{Z_a(0)} (1-\xi_a^{(i)}) \tag{5.2.116}$$

In order to evaluate, $\mu_a(1)$,$\rho_a(1)$, $C_{a,b}(1)$, $C\Gamma_{a,b}(1)$ and $\Gamma_{a,b}(1)$. We first evaluate the conditional expectations given $\vec{Z}(0)$, $\mu_a(1)|_{\vec{Z}(0)}$,$\rho_a(1)|_{\vec{Z}(0)}$, $C_{a,b}(1)|_{\vec{Z}(0)}$, $C\Gamma_{a,b}(1)|_{\vec{Z}(0)}$ and $\Gamma_{a,b}(1)|_{\vec{Z}(0)}$.

The first order moments can be calculated as

$$\begin{aligned} &\mu_a(1)|_{\vec{Z}(0)} = Z_a(0)\mathbb{E}\xi_a^{(i)} = Z_a(0)\alpha \\ &\Rightarrow \mu_a(1) = \mu_a(0)\alpha \end{aligned} \tag{5.2.117}$$

$$\begin{aligned} &\rho_a(1)|_{\vec{Z}(0)} = Z_a(0)(1-\mathbb{E}\xi_a^{(i)}) = Z_a(0)(1-\alpha) \\ &\Rightarrow \rho_a(1) = \mu_a(0)(1-\alpha) \end{aligned} \tag{5.2.118}$$

For $\mu_a(0)$ neutrons in the source bank before splitting, on average, an α fraction of them is sampled into the prompt source bank and the remaining $1-\alpha$ fraction goes into the delayed source bank.

Then we derive the initialization of second order moments. Inserting Eq 5.2.115 into $C_{a,b}(1)|_{\vec{Z}(0)}$,

$$C_{a,b}(1)|_{\vec{Z}(0)} = \mathbb{E}\left[Z_a(1)Z_b(1)|\,\vec{Z}(0)\right] = \mathbb{E}\left[\sum_{i=1}^{Z_a(0)}\xi_a^{(i)}\sum_{j=1}^{Z_b(0)}\xi_b^{(j)}\middle|\,\vec{Z}(0)\right] \tag{5.2.119}$$

Eq 5.2.119 takes different forms for $a = b$ and $a \neq b$. When $a \neq b$, $Z_a(1)$ and $Z_b(1)$ are independent, the expectation of products is simplified to a product of expectations,

$$\begin{aligned} C_{a,b}(1)|_{\vec{Z}(0)} &= \mathbb{E}\left[\sum_{i=1}^{Z_a(0)}\xi_a^{(i)}\middle|\,\vec{Z}(0)\right]\mathbb{E}\left[\sum_{j=1}^{Z_b(0)}\xi_b^{(j)}\middle|\,\vec{Z}(0)\right] \\ &= \sum_{i=1}^{Z_a(0)}\mathbb{E}\left[\xi_a^{(i)}\right]\sum_{j=1}^{Z_b(0)}\mathbb{E}\left[\xi_b^{(j)}\right] \\ &= Z_a(0)\alpha \times Z_b(0)\alpha \end{aligned} \tag{5.2.120}$$

We evaluate the conditional expectation with respect to the distribution of $\vec{Z}(0)$,

$$C_{a,b}(1) = \mathbb{E}[Z_a(0)\alpha \times Z_b(0)\alpha] = \alpha^2 C_{a,b}(0) \tag{5.2.121}$$

When $a = b$, independent terms can still be separated from Eq 5.2.119.

$$\begin{aligned} C_{a,a}(1)|_{\vec{Z}(0)} &= \mathbb{E}\left[\sum_{i=1}^{Z_a(0)}\sum_{j=1}^{Z_a(0)}\xi_a^{(i)}\xi_a^{(j)}\middle|\,\vec{Z}(0)\right] \\ &= \mathbb{E}\left[\sum_{i=1}^{Z_a(0)}{\xi_a^{(i)}}^2 + \sum_{i\neq j}^{Z_a(0)}\xi_a^{(i)}\xi_a^{(j)}\middle|\,\vec{Z}(0)\right] \\ &= \sum_{i=1}^{Z_a(0)}\mathbb{E}\left[{\xi_a^{(i)}}^2\right] + \sum_{i\neq j}^{Z_a(0)}\mathbb{E}\left[\xi_a^{(i)}\xi_a^{(j)}\right] \end{aligned} \tag{5.2.122}$$

Because $\xi_a^{(i)}$ is either 0 or 1, $\xi_a^{(i)} = {\xi_a^{(i)}}^2$ for the first term in Eq 5.2.122. Also, $\xi_a^{(i)}$ and $\xi_a^{(j)}$ are independent for $i \neq j$, thus the expectation of products in the second term

can also be simplified to product of expectations.

$$C_{a,a}(1)|_{\vec{Z}(0)} = Z_a(0)\alpha + Z_a(0)(Z_a(0)-1)\alpha^2 \tag{5.2.123}$$

We evaluate the conditional expectation with respect to the distribution of $\vec{Z}(0)$,

$$C_{a,a}(1) = \mathbb{E}[Z_a(0)^2\alpha^2 + Z_a(0)\alpha(1-\alpha)] = \alpha^2 C_{a,a}(0) + \mu_a(0)\alpha(1-\alpha) \tag{5.2.124}$$

Eq 5.2.121 and Eq 5.2.124 can be combined into Eq 5.2.125.

$$C_{a,b}(1) = \alpha^2 C_{a,b}(0) + \delta_{a,b}\mu_a(0)\alpha(1-\alpha) \tag{5.2.125}$$

Considering the relation between C and $\complement$ (Eq 5.2.22) and $\mu_a(1)$ found in Eq 5.2.117, Eq 5.2.125 is shown to be equivalent to Eq 5.2.126.

$$\complement C_{a,b}(1) = \alpha^2 \complement C_{a,b}(0) \tag{5.2.126}$$

By symmetry, the relation between $\Gamma_{a,b}(1)$ and $C_{a,b}(0)$ can be found by replacing all α with $1-\alpha$ in Eq 5.2.125 or Eq 5.2.126.

$$\Gamma_{a,b}(1) = (1-\alpha)^2 C_{a,b}(0) + \delta_{a,b}\mu_a(0)\alpha(1-\alpha) \tag{5.2.127}$$

Considering the relation between Γ and $\complement\Gamma$ (Eq 5.2.35), the relation between C and $\complement$ (Eq 5.2.22) and $\mu_a(1)$ found in Eq 5.2.117, Eq 5.2.127 is shown to be equivalent to Eq 5.2.128.

$$\complement\Gamma_{a,b}(1) = (1-\alpha)^2 \complement C_{a,b}(0) \tag{5.2.128}$$

Next, we must find $C\Gamma(1)$. Inserting Eq 5.2.115 and Eq 5.2.116 into $C\Gamma_{a,b}(1)|_{\vec{Z}(0)}$,

we get

$$C\Gamma_{a,b}(1)|_{\vec{Z}(0)} = \mathbb{E}\left[Z_a(1)U_b(1)|\,\vec{Z}(0)\right] = \mathbb{E}\left[\sum_{i=1}^{Z_a(0)} \xi_a^{(i)} \sum_{j=1}^{Z_b(0)} (1-\xi_b^{(j)}) \middle| \vec{Z}(0)\right] \tag{5.2.129}$$

When $a \neq b$, independence between $Z_a(1)$ and $U_b(1)$ simplifies Eq 5.2.129 to Eq 5.2.131

$$\begin{aligned} C\Gamma_{a,b}(1)|_{\vec{Z}(0)} &= \mathbb{E}\left[\sum_{i=1}^{Z_a(0)} \xi_a^{(i)} \middle| \vec{Z}(0)\right] \mathbb{E}\left[\sum_{j=1}^{Z_b(0)} (1-\xi_b^{(j)}) \middle| \vec{Z}(0)\right] \\ &= \sum_{i=1}^{Z_a(0)} \mathbb{E}\left[\xi_a^{(i)}\right] \sum_{j=1}^{Z_b(0)} \mathbb{E}\left[(1-\xi_b^{(j)})\right] \\ &= Z_a(0)\alpha Z_b(0)(1-\alpha) \end{aligned} \tag{5.2.130}$$

We then evaluate the conditional expectation with respect to the distribution of $\vec{Z}(0)$,

$$C\Gamma_{a,b}(1) = \mathbb{E}[Z_a(0)\alpha \times Z_b(0)(1-\alpha)] = \alpha(1-\alpha)C_{a,b}(0) \tag{5.2.131}$$

When $a = b$,

$$\begin{aligned} C\Gamma_{a,a}(1)|_{\vec{Z}(0)} &= \mathbb{E}\left[\sum_{i=1}^{Z_a(0)} \sum_{j=1}^{Z_a(0)} \xi_a^{(i)}(1-\xi_a^{(j)}) \middle| \vec{Z}(0)\right] \\ &= \mathbb{E}\left[\sum_{i=1}^{Z_a(0)} \xi_a^{(i)}(1-\xi_a^{(i)}) + \sum_{i\neq j}^{Z_a(0)} \xi_a^{(i)}(1-\xi_a^{(j)}) \middle| \vec{Z}(0)\right] \\ &= \sum_{i=1}^{Z_a(0)} \left(\mathbb{E}\left[\xi_a^{(i)}\right] - \mathbb{E}\left[{\xi_a^{(i)}}^2\right]\right) + \sum_{i\neq j}^{Z_a(0)} \left(\mathbb{E}\left[\xi_a^{(i)}\right] - \mathbb{E}\left[\xi_a^{(i)}\xi_a^{(j)}\right]\right) \end{aligned} \tag{5.2.132}$$

Following the same reasoning as previously, Eq 5.2.132 is simplified to

$$C\Gamma_{a,a}(1)|_{\vec{Z}(0)} = Z_a(0)(Z_a(0)-1)\alpha(1-\alpha) \tag{5.2.133}$$

We then evaluate the conditional expectation with respect to the distribution of $\vec{Z}(0)$,

$$\begin{aligned} C\Gamma_{a,a}(1) &= \mathbb{E}[Z_a(0)(Z_a(0)-1)\alpha(1-\alpha)] \\ &= \alpha(1-\alpha)C_{a,a}(0) - \mu_a(0)\alpha(1-\alpha) \end{aligned} \tag{5.2.134}$$

Eq 5.2.131 and Eq 5.2.134 can be combined into Eq 5.2.135.

$$C\Gamma_{a,b}(1) = \alpha(1-\alpha)C_{a,b}(0) - \delta_{a,b}\mu_a(0)\alpha(1-\alpha) \tag{5.2.135}$$

Considering the relation between C and $\complement C$ (Eq 3.3.33), Eq 5.2.135 is shown to be equivalent to Eq 5.2.136.

$$\complement C\Gamma_{a,b}(1) = \alpha(1-\alpha)\complement C_{a,b}(0) \tag{5.2.136}$$

It is worthwhile to note that the relations between the moments of $\vec{Z}(1)$, $\vec{U}(1)$ and the moments of $\vec{Z}(0)$ are in complete analogy to the response moments derived in section 5.2.5.2. More specifically, starting from the results in section 5.2.5.2, discarding the type of source neutron, l, and setting f_a to the constant number α, we find that Eq 5.2.102 is equivalent to Eq 5.2.117, Eq 5.2.103 is equivalent to Eq 5.2.118, Eq 5.2.104 is equivalent to Eq 5.2.126, Eq 5.2.105 is equivalent to Eq 5.2.125, Eq 5.2.106 is equivalent to Eq 5.2.128, Eq 5.2.107 is equivalent to Eq 5.2.127, Eq 5.2.108 is equivalent to Eq 5.2.136, Eq 5.2.109 is equivalent to Eq 5.2.135.

This is because section 5.2.5.2 and section 5.2.5.4 are actually two derivations of the same problem. In section 5.2.5.2, the task is to find the moments of $\vec{Z}(1)$ and $\vec{U}(1)$ in terms of the moments of $\vec{Z}^*(1)$, while in section 5.2.5.4, the task is to find the moments of $\vec{Z}(1)$ and $\vec{U}(1)$ in terms of the moments of $\vec{Z}(0)$. $\vec{Z}^*(1)$ in section 5.2.5.2 is a state of the prompt source bank induced by one source neutron. $\vec{Z}(0)$ in this section is a state of the prompt source bank induced by inactive generations. However, both the processes from $\vec{Z}^*(1)$ to $\vec{Z}(1)$ and $\vec{U}(1)$ and from $\vec{Z}(0)$ to $\vec{Z}(1)$ and $\vec{U}(1)$ are simply determined by Bernoulli random variables indicating whether each prompt neutron is sampled into the prompt source bank or the delayed source bank.

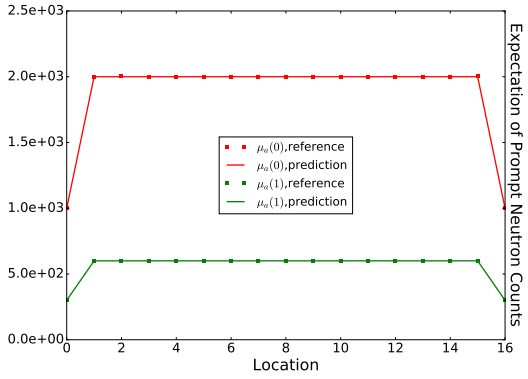
5.2.6 Numerical Results

This section compares reference values for the $TD1$ problem and predictions from the above derivations. Test problem $TD1$ was simulated with $s = 32000$ neutrons per generation. After 20 inactive generations, the source bank is separated using $\alpha = 0.3$, thus approximately 9600 neutrons are sampled into the prompt source bank, the remaining are stored in the delayed source bank. For each active generation, the prompt neutrons are transported and sampled into the prompt source bank for the next generation with a probability $f = 0.7$. The neutrons in the delayed source bank are sampled into the delayed source bank for the next generation with probability $p = 0.129$, calculated from α and f according to Eq 5.2.1. 2000 independent simulations were performed to generate the reference solution. Test problem $TD1$ simulated with the above values of α, f and p parameters is denoted as $TD1 - D1$ (Discrete Test problem 1 with Delayed scheme 1).

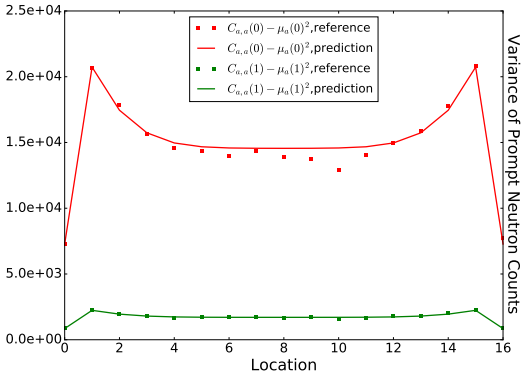
Figure 5-3 plots the transition of moments immediately after the delayed neutron scheme starts. The results at generation 0 in the delayed scheme are actually the source bank generated by the last inactive generation. Therefore, the predictions at generation 0 are produced from the MBP model by evolving a deterministic source to generation 21. The predictions at generation 1 are made according to Eq 5.2.117 (Figure 5-3(a)), Eq 5.2.125 (Figure 5-3(b),Figure 5-3(c),Figure 5-3(d)). The consistence between the predicted and reference values at generation 1 verifies the derivations in section 5.2.5.4.

With the correctly initialized moments $\mu_a(1)$, $C_{a,b}(1)$, $C\Gamma_{a,b}(1)$ and $\Gamma_{a,b}(1)$, the recursive formula derived in section 5.2.3 can be applied to predict the moments for any active generation n. In Figure 5-4, the solid curves of different colors correspond to the prediction at different active generations. The squares correspond to reference values calculated from independent simulations. The prediction curves and reference values share the same color for results at the same generation. The predicted moments of both prompt and delayed neutron count match the references well.

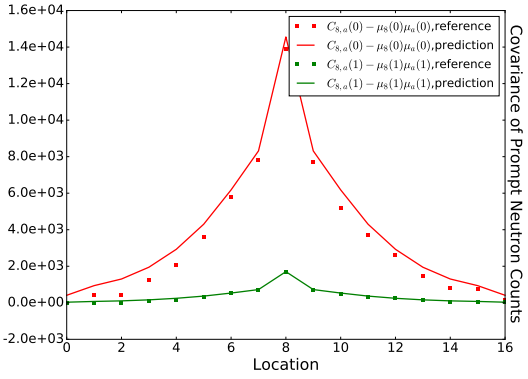
The moments are then used to predict correlation coefficients. Similar to the



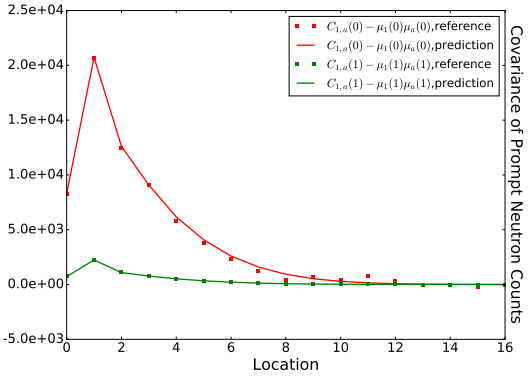
(a) $\mu_a(n)$ for $n = 0, 1$ and references



(b) $C_{a,a}(n) - \mu_a^2(n)$ for $n = 0, 1$ and references

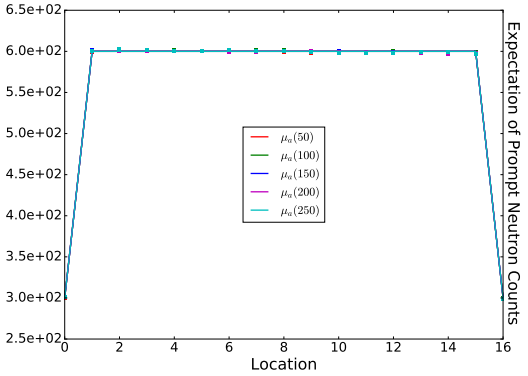


(c) $C_{8,a}(n) - \mu_8(n)\mu_a(n)$ for $n = 0, 1$ and references

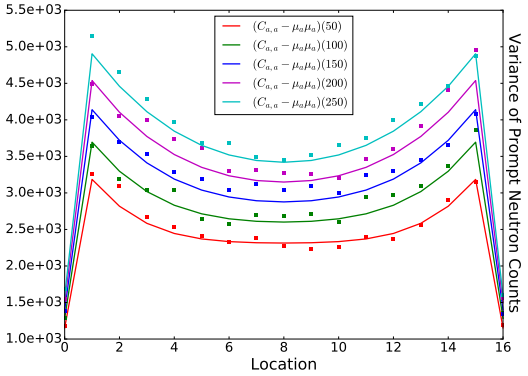


(d) $C_{1,a}(n) - \mu_1(n)\mu_a(n)$ for $n = 0, 1$ and references

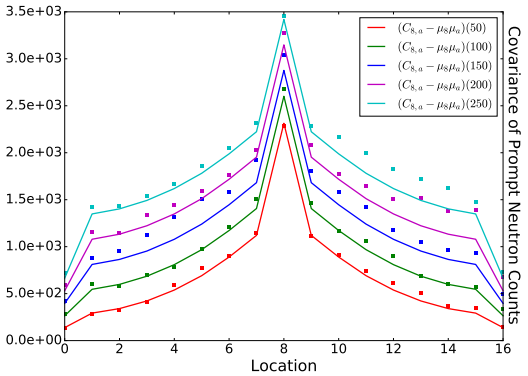
Figure 5-3: Spatial moments on starting delayed scheme, test problem $TD1 - D1_f$



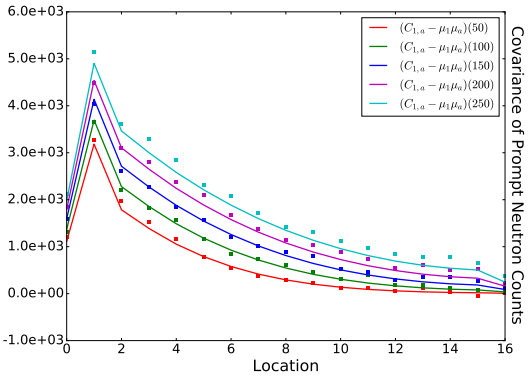
(a) $\mu_a(n)$ and references



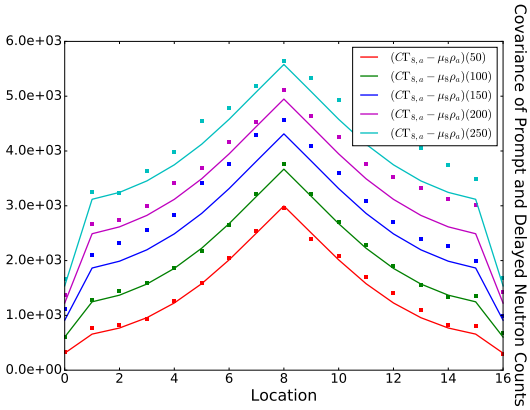
(b) $C_{a,a}(n) - \mu_a(n)^2$ and references



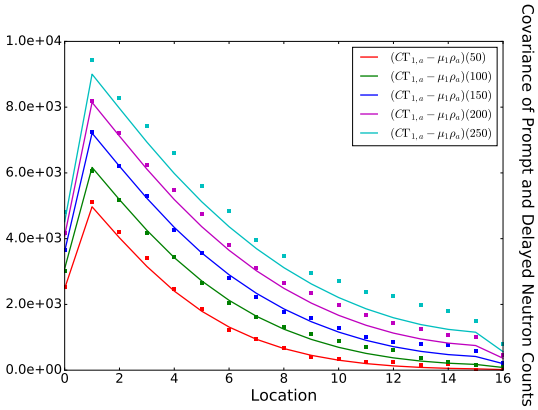
(c) $C_{8,a}(n) - \mu_8(n)\mu_a(n)$ and references



(d) $C_{1,a}(n) - \mu_1(n)\mu_a(n)$ and references



(e) $CT_{8,a}(n) - \mu_8(n)\rho_a(n)$ and references



(f) $CT_{1,a}(n) - \mu_1(n)\rho_a(n)$ and references

Figure 5-4: Spatial moments results at active generation 50, 100, 150, 200 *and* 250, *test problem* $TD1 - D1_f$

discussion in section 3.1.2, the DMBP model is not expected to predict correlation coefficients of Z tallies of prompt neutrons well for simulations with neutron number normalization. Instead, correlation coefficients of X tallies of prompt neutrons predicted by the DMBP model are used as a good approximation. Covariances of X tallies in the same region across different generations (Eq 3.3.70) are still valid for the DMBP model since only prompt neutrons contribute to the tally. The only difference is that the temporal-spatial moment $\mathbb{E}[Z_l(n)Z_j(n+k)]$ required by Eq 3.3.70 should be evaluated according to Eq 5.2.81 rather than Eq 3.3.51 due to the added correlation through the delayed source bank.

The predicted correlation coefficients for $TD1-D1$ are plotted with solid curves in Figure 5-5. Figure 5-5(a) plots correlation coefficients for tally X_1, normalized prompt neutron count in region 1, which is near the boundary of the problem. Figure 5-5(b) plots correlation coefficients for tally X_8, normalized prompt neutron count in region 8, which corresponds to the central region. The reference values for the test problem $TD1D1_f$ are also given with cross symbols. In addition, a variation of the problem with neutron number normalization for both prompt and delayed source bank ($TD1D1_n$) is also used to generate reference correlation coefficients. The reference values of $TD1D1_n$ are plotted with dot symbols. As expected, the correlation of X tallies from $TD1D1_n$ and $TD1D1_f$ are similar to each other and both of them are consistent with the prediction. Figure 5-5 shows that correlation coefficients under the delayed neutron scheme still preserves the feature that the boundary region has higher correlation coefficients than the central regions due to the reflection of neutrons.

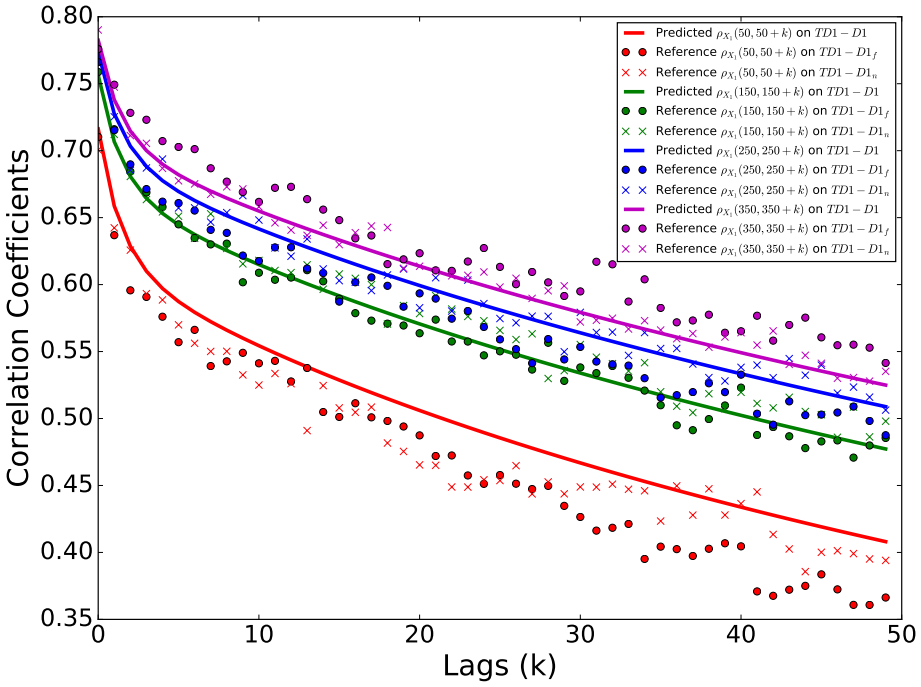
Figure 5-5 plots the correlation coefficients $\rho(n, n+k)$ as a function of k for $n \in \{50, 150, 250, 350\}$. Correlation coefficients $\rho(n, n+k)$ depend on n and grow as n increases until they reach an asymptotic, the difference is significant for boundary tallies. The correlation coefficients are also lower than those on $TD1$ without delayed neutrons (Figure 3-8 vs Figure 5-5).

The n dependence and comparison with the traditional simulation scheme can be explained conceptually. Similar to the observation that the correlation coefficients decay exponentially as a function of the generation lag, we observe that the correlation

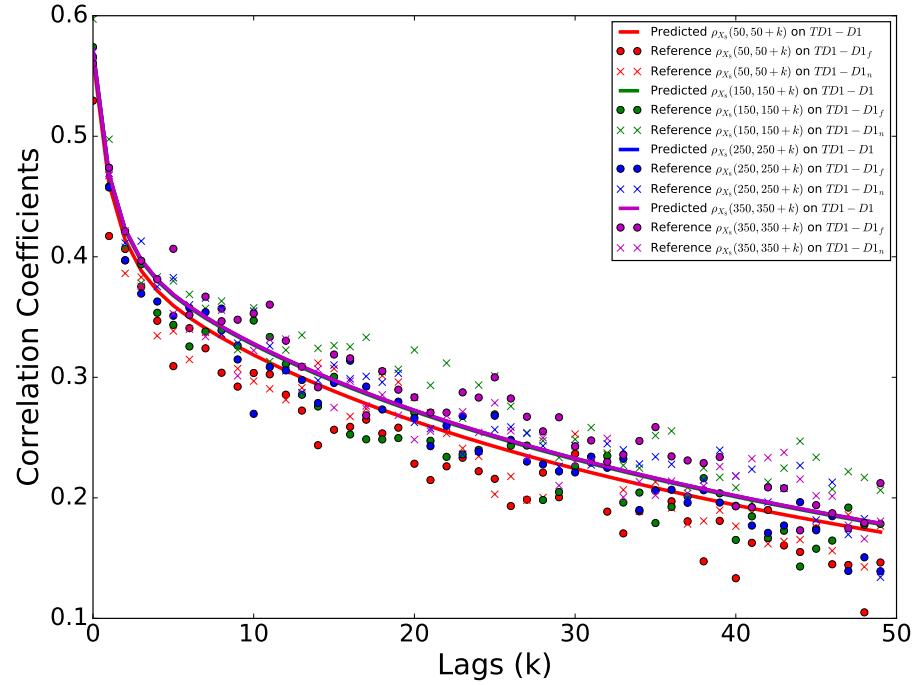
between adjacent generations is dominated by neutrons with an age difference of 1, where neutron age is defined as the number of times a neutron is sampled into the prompt source bank and transported to the absorption site, and the age of a new neutron born out of fission inherits the value of its predecessor. For the prompt source bank across neighboring active generations, the fraction of neutrons with age difference 1 increases as the simulation continues. For generation n and $n+1$, a fraction of roughly f neutrons in the prompt source banks have an age difference of 1 since they were just sampled and transported. However, there are neutrons with an age difference of 1 in the delayed source bank at generation n that can be sampled into the prompt source bank at generation n. This latter contribution increases in further generations and leads to an increase of $\rho(n, n+k)$ as a function of n.

Another important task of the predictive model is to evaluate the variance underestimation ratio. The assumption of stationary correlation coefficients used in predicting variance underestimation ratio $r(N)$ (Eq 2.2.8 from correlation coefficients is not valid. The derivations for a more general situation in section 2.2.2.3 should be used. We use Eq 2.2.36 to calculate the generalized variance underestimation ratio $R(N)$ from variance and covariance of X tallies. The reference values are calculated according to Eq 2.2.47. The prediction values of $R(N)/N$ are plotted in Figure 5-6. Figure 5-6(a) and Figure 5-6(b) show that the boundary region has a higher variance underestimation ratio due to the higher correlation coefficients observed in Figure 5-5. Reference values from both $TD1D1_f$ and $TD1D1_n$ are also plotted. As expected, the correlation behavior of X tallies from simulations with and without neutron number normalization are close to each other. And they are both consistent with the prediction from the DMBP model.

With the accurately predicted $R(N)$, the next task is to correct the biased variance estimator $\widehat{s^2}(N)/N$ (Eq 2.2.21) to provide a better estimate of $\mathrm{Var}[\bar{X}(N)]$. According to the derivation in section 2.2.2.3, $\widehat{\bar{\sigma}^2(N)}/N$ (Eq 2.2.37) is an unbiased estimator of $\mathrm{Var}[\bar{X}(N)]$ given $R(N)$. Therefore, we multiply $\widehat{\bar{\sigma}^2}(N)$ with $\frac{R(N)}{N-R(N)}$. The corrected variance $\widehat{\bar{\sigma}^2}\frac{R(N)}{N-R(N)}$ is then compared to the reference $\mathrm{Var}[\bar{X}(N)]$ calculated from independent simulations. The comparisons are shown in Figure 5-7. Since we have

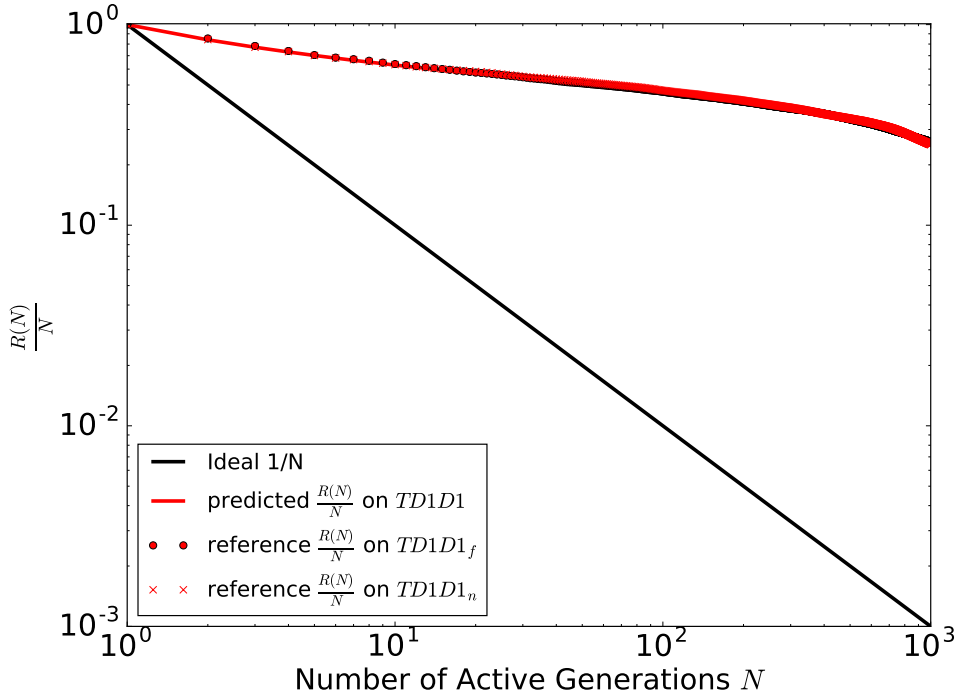


(a) Correlation coefficients of tally X_1

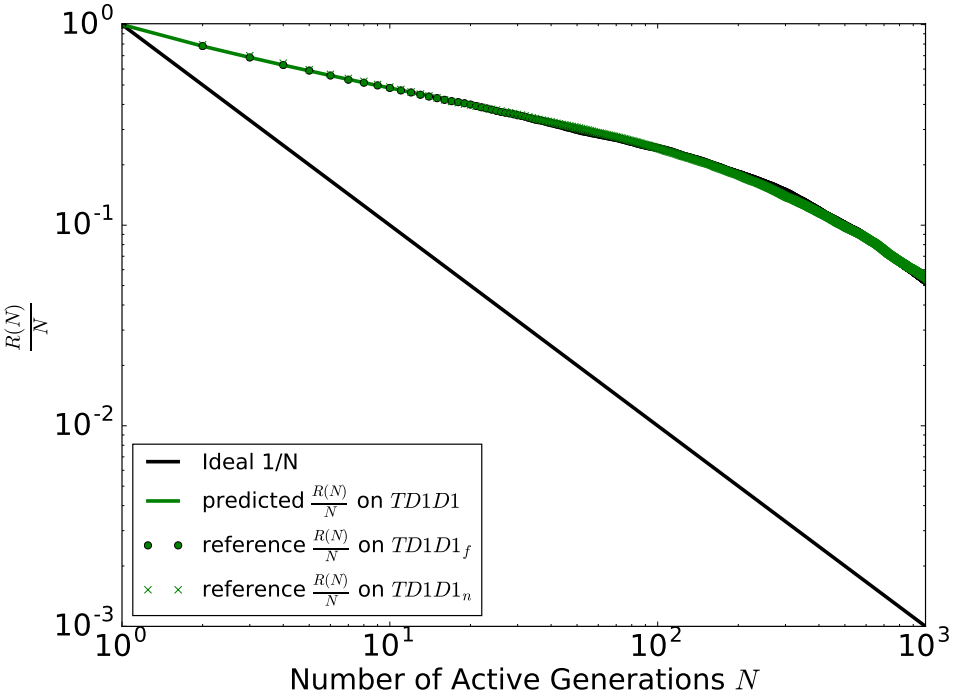


(b) Correlation coefficients of tally X_8

Figure 5-5: Correlation coefficients of normalized tallies on test problem $TD1D1_f$ and $TD1D1_n$



(a) Variance convergence rate $R(N)/N$ of tally X_1



(b) Variance convergence rate $R(N)/N$ of tally X_8

Figure 5-6: Variance convergence rate $R(N)/N$ of normalized tallies on test problem $TD1D1_f$ and $TD1D1_n$

accurately predicted $R(N)$, the corrected variances also match the reference values very well. Figure 5-7 also plots the variance $\frac{\widehat{s^2}(N)}{N}$ that is often incorrectly reported to illustrate once more the impact of correlations.

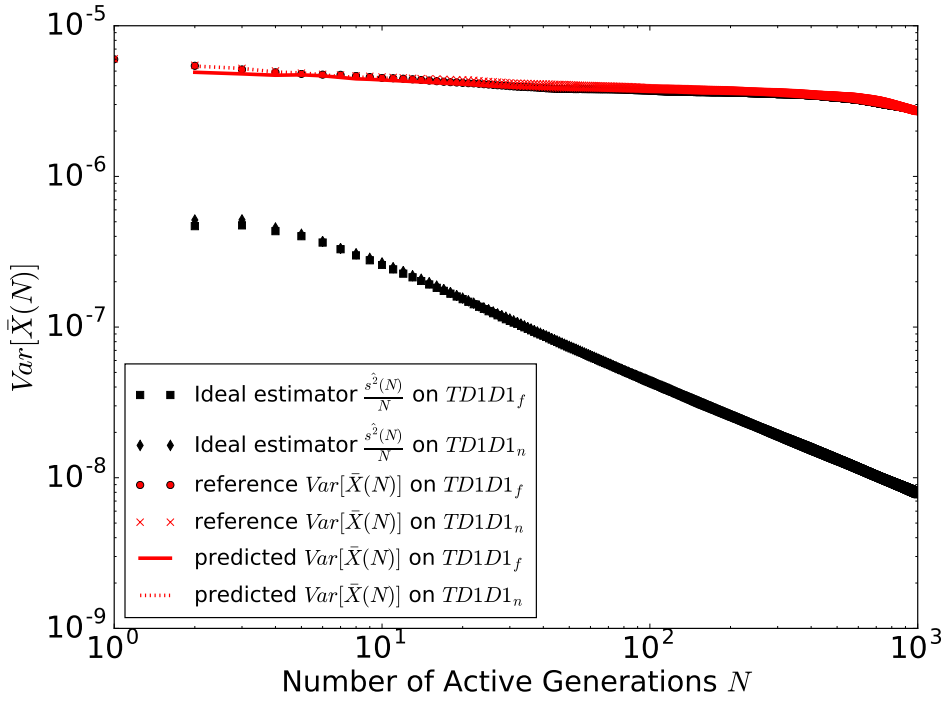
The previous figures verify that the DMBP model is capable of accurately predicting correlation coefficients and real variances for tallies from the delayed neutron simulation method. This now allows us to compare the variance convergence rate with the traditional source method and assess if any benefit can be obtained.

Figure 5-8 plots $\text{Var}[\bar{X}_1]$ and $\text{Var}[\bar{X}_8]$ from both $TD1_n$ and $TD1D1_n$. In $TD1_n$, 1020 generations are simulated with 32000 neutrons in the source bank, the last 1000 generations are tallied. in $TD1D1_n$, 1020 generations are simulated with 106666 neutrons in the source bank for the first 20 generations. From the 21^{st} generation, the source bank is separated into prompt and delayed source bank with $\alpha = 0.3$. In the 1000 active generations, there are 31999 neutrons in the prompt source bank and these neutrons contribute to the tallies. The parameters above are selected such that we can compare results from active generations of equal size (~ 32000) regardless of the size of the inactive generations.

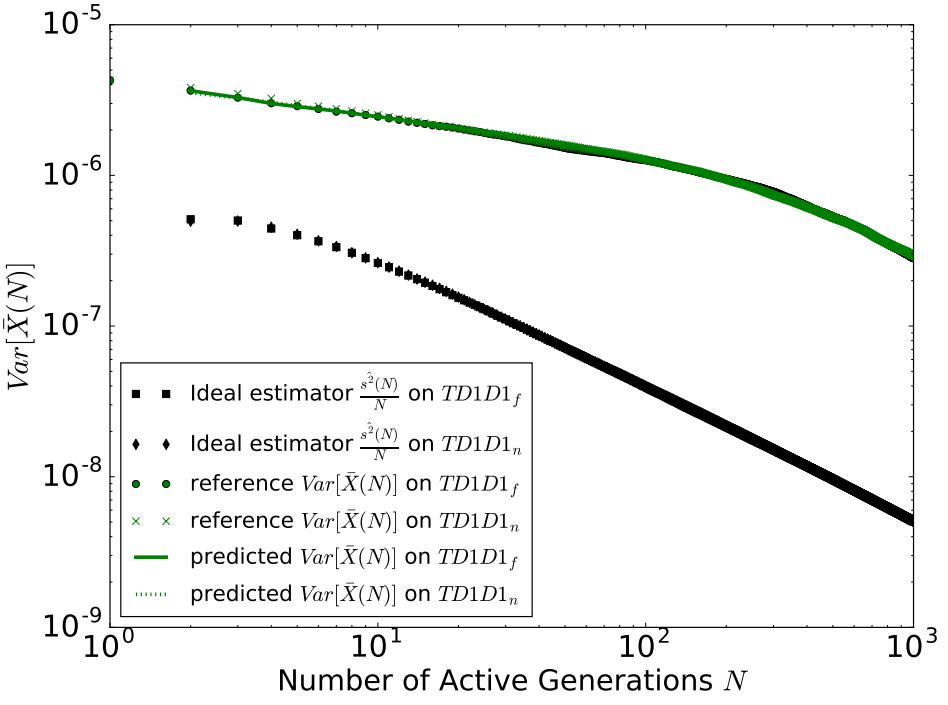
Although both $TD1_n$ and $TD1D1_n$ have the same number of neutrons contributing to tallies (active), $TD1D1_n$ features a significantly lower $\text{Var}[\bar{X}_I(1)]$'s. In $TD1_n$, $X_I(1)$ sees contributions from $s = 32000$ correlated neutrons, while in $TD1D1_n$, $X_I(1)$ contributions come from $s = 31999$ neutrons sampled from 106666 correlated neutrons, the correlations are therefore diluted. This explains the lower $\text{Var}[\bar{X}_I(1)]$. Lower leading variance in $TD1D1_n$ is also reflected in the drop of $\text{Var}[Z_I(1)]$ derived in Eq 5.2.125 and numerically shown in Figure 5-3(b).

Note that in Figure 5-8, the predicted $\text{Var}[\bar{X}_I(N)]$'s are purely determined by the DMBP model. In comparison, predicted values $\text{Var}[\bar{X}_I(N)]$ in Figure 5-7 come from traditional estimators corrected by $R(N)$. Figure 5-7 shew that the DMBP model can predict $R(N)$ accurately, while Figure 5-8 shows that the DMBP model predicts the magnitude of $\text{Var}[\bar{X}_I(N)]$ in addition to the shape determined by $R(N)$.

Numerical results on $TD1$ and $TD1D1$ show the possibility to reach better variance and consequentially lower square error for the same simulation efforts in active

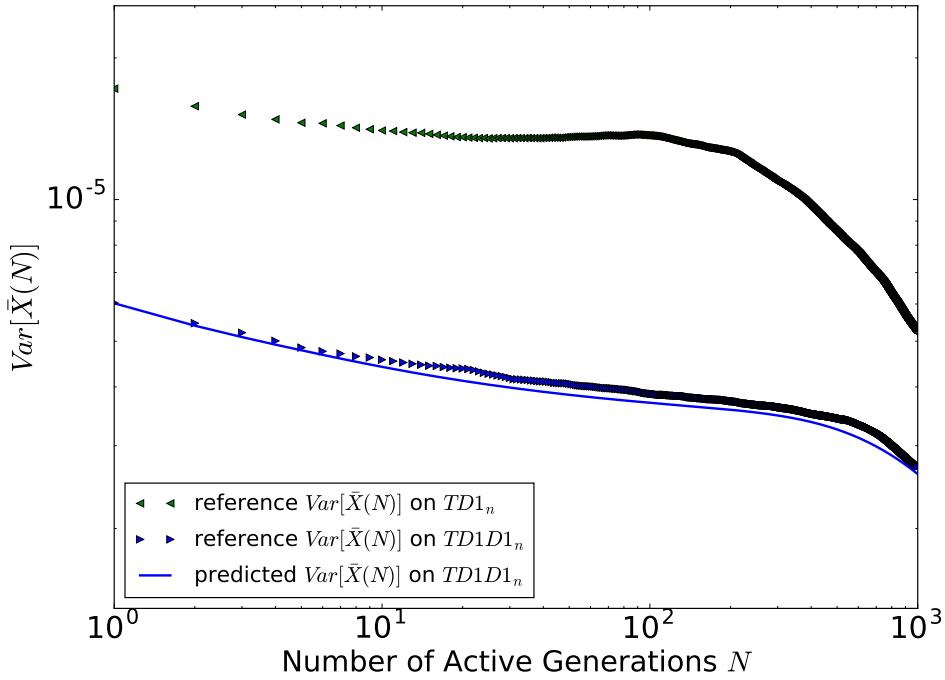


(a) Real variance $\mathrm{Var}\,\bar{X}(N)$ of tally X_1

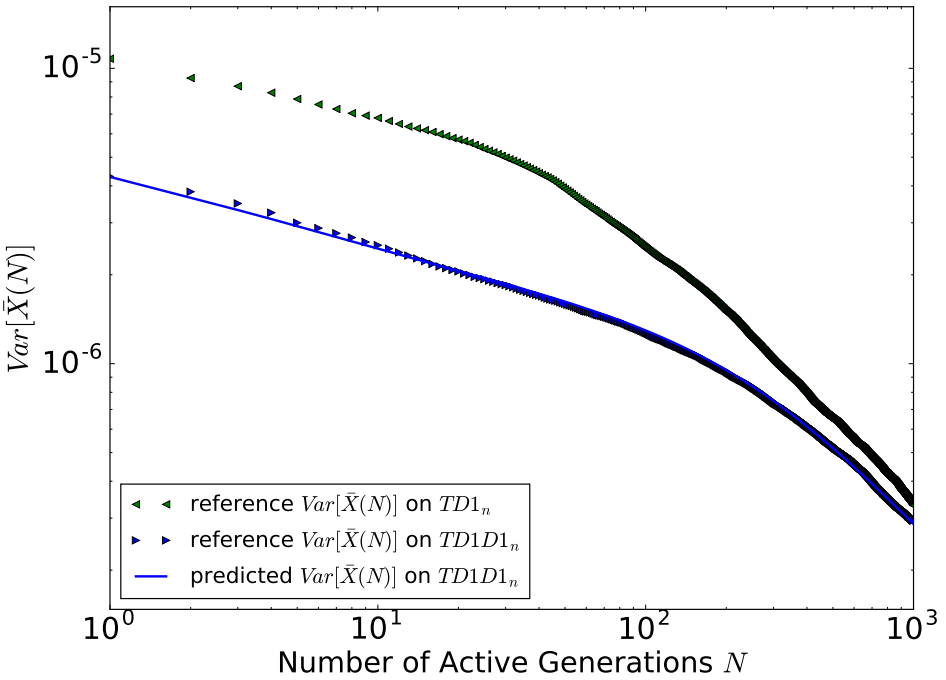


(b) Real variance $\mathrm{Var}\,\bar{X}(N)$ of tally X_8

Figure 5-7: $\mathrm{Var}\,\bar{X}(N)$ *of normalized tallies on test problem* $TD1D1_f$ *and* $TD1D1_n$



(a) Real variance $\mathrm{Var}\,\bar{X}(N)$ of tally X_1



(b) Real variance $\mathrm{Var}\,\bar{X}(N)$ of tally X_8

Figure 5-8: $\mathrm{Var}\,\bar{X}(N)$ *of normalized tallies on test problem* $TD1_n$ *and* $TD1D1_n$

generations, but this comes at the cost of expensive inactive simulations. It is worthwhile to investigate how to select the best α, f and p parameters to reach the error threshold with the least simulation cost or reach the lowest error given at total simulation cost constraint. This is discussed further in the next section.

5.3 Running strategy in delayed neutron simulation

Results from section 5.2.6 show that the recursive formula for moments in DMBP not only can predict the correlation coefficients $\rho(n;k)$ and the variance underestimation ratio $R(N)$ but can also predict the absolute value of $\text{Var}[\bar{X}(N)]$ accurately. Therefore, results of a corresponding DMBP model can be used to find optimal α, f and p parameters for a real problem. More specifically, given the number of inactive generations N_0 and the number of neutrons per inactive generation s, the DMBP model predicts the $\text{Var}[\bar{X}_I(N)]$ for any region I for any tally accumulated up to active generation N as a function of the delayed neutron parameters α, f, p. We now seek to find optimal values of α, f, p which minimize $\text{Var}[\bar{X}_I(N)]$ with given $N_0 s + \alpha N s$ or minimizes $N_0 s + \alpha N s$ for a target $\text{Var}[\bar{X}_I(N)]$.

5.3.1 Explicit form of moments

The task now is to write $\text{Var}[\bar{X}_I(N)]$ explicitly as a function of α, f, p and N_0, s. With the delayed neutron method, $\text{Var}[\bar{X}_I(N)]$ is calculated according to Eq 2.2.38. Since X_I is the normalized tally, $Cov[X_I(n), X_I(n+k)]$ in Eq 4.2.8 should be evaluated with $k = 0, 1, \cdots, N-1$ and $n = 1, \cdots, N-k$.

The temporal-spatial moments $\mathbb{E}[Z_l(n)Z_j(n+k)]$ (Eq 5.2.81) are the key in predicting $Cov[X_I(n), X_I(n+k)]$. In order to evaluate $\mathbb{E}[Z_l(n)Z_j(n+k)]$, the moments in response to one source neutron after k generations, $\mu_b^{(j)}(k)$ and $\mu_{\overline{b}}^{(j)}(k)$ are required. The current method is to start from a single prompt neutron in region j and then recursively calculating $\vec{\mu}(1), \cdots, \vec{\mu}(k)$ and $\vec{\rho}(1), \cdots, \vec{\rho}(k)$ according to Eq 5.2.110 and Eq 5.2.111. The resulting $\vec{mu}(k)$ at each generation is the needed $\mu_b^{(j)}(k)$. The same procedure is carried on for one delayed neutron in region j to calculate $\mu_{\overline{b}}^{(j)}(k)$. These

calculations are reorganized below.

First, we concatenate vector $\vec{\mu}$ (expected number of prompt neutrons) and vector $\vec{\rho}$ (expected number of delayed neutrons) into a single vector $\vec{\pi}$ of length $2M$

$$\pi_i(n) = \begin{cases} \mu_i(n) & i = 1, \cdots, M \\ \rho_{i-M}(n) & i = M+1, \cdots, 2M \end{cases} \tag{5.3.1}$$

Then the recursive formula Eq 5.2.110 and Eq 5.2.111 can be rewritten in terms of $\vec{\pi}$ as

$$\pi(n+1) = \pi(n)G \tag{5.3.2}$$

G in Eq 5.3.2 is a $2M \times 2M$ matrix defined by specifying each of its $M \times M$ blocks

$$G = \left[\begin{array}{c|c} M & \Psi \\ \hline m & \psi \end{array}\right] = \left[\begin{array}{c|c} fM_0 & (1-f)M_0 \\ \hline p\mathbb{1} & (1-p)\mathbb{1} \end{array}\right] \tag{5.3.3}$$

where constant f and p are restricted and Eq 5.2.102 and Eq 5.2.103 are used.

According to Eq 5.3.2, $\mu_b^{(j)}(k)$ is the b^{th} element of $\vec{\pi}(k)$ with $\pi(0)_i = \delta_{i,j}$ and $\mu_{\overline{b}}^{(j)}(k)$ is the $M + b^{th}$ element of $\vec{\pi}(k)$ with $\pi(0)_i = \delta_{i,j+M}$. In matrix form,

$$\mu_b^{(j)}(k) = G^{(k)}(1,1)_{j,b} \tag{5.3.4}$$

$$\mu_{\overline{b}}^{(j)}(k) = G^{(k)}(2,1)_{j,b} \tag{5.3.5}$$

where $G^{(k)}(i_1, i_2)$ denotes the (i_1, i_2) block of the matrix G to the power of k.

Eq 5.3.2 summarizes the recursive formula for first order moments into one equation. Eq 5.3.4 and Eq 5.3.5 give a clean representation of the $k-generation$ response moments required to evaluate the temporal-spatial moments $\mathbb{E}[Z_l(n)Z_j(n+k)]$. Now we can explicitly write the second order moments $C_{a,b}$, $C\Gamma_{a,b}$ and $\Gamma_{a,b}$ from the recursive formula in Eq 5.2.28, Eq 5.2.41 and Eq 5.2.50. For convenience, the recursive formula for second order moments are rewritten in terms of the corresponding covari-

ances defined below.

$$Cov_{a,b}(n) \equiv C_{a,b}(n) - \mu_a(n)\mu_b(n) \tag{5.3.6}$$

$$C\Gamma ov_{a,b}(n) \equiv C\Gamma_{a,b}(n) - \mu_a(n)\rho_b(n) \tag{5.3.7}$$

$$\Gamma ov_{a,b}(n) \equiv \Gamma_{a,b}(n) - \rho_a(n)\rho_b(n) \tag{5.3.8}$$

The definitions above along with the recursive formula for first order moments (Eq 5.2.13 and Eq 5.2.14) transform the recursive formula for second order moments into the recursive formula for second order central moments.

Eq 5.2.28 is transformed into Eq 5.3.9.

$$\begin{aligned} Cov(n)_{a,b} = {} & \mu_l(n-1)V^l_{a,b} + \rho_k(n-1)v^k_{a,b} \\ & + M^l_a Cov(n-1)_{l,j} M^j_b + m^k_a \Gamma ov(n-1)_{k,j} m^j_b \\ & + M^l_a C\Gamma ov(n-1)_{l,k} m^k_b + m^k_a C\Gamma ov(n-1)_{l,k} M^l_b \end{aligned} \tag{5.3.9}$$

Eq 5.2.41 is transformed into Eq 5.3.10.

$$\begin{aligned} \Gamma ov(n)_{a,b} = {} & \mu_l(n-1)\Lambda^l_{a,b} + \rho_k(n-1)\lambda^k_{a,b} \\ & + \Psi^l_a Cov(n-1)_{l,j} \Psi^j_b + \psi^k_a \Gamma ov(n-1)_{k,j} \psi^j_b \\ & + \Psi^l_a C\Gamma ov(n-1)_{l,k} \psi^k_b + \psi^k_a C\Gamma ov(n-1)_{l,k} \Psi^l_b \end{aligned} \tag{5.3.10}$$

Eq 5.2.50 is transformed into Eq 5.3.11.

$$\begin{aligned} C\Gamma ov(n)_{a,b} = {} & \mu_l(n-1)\Omega^l_{a,b} + \rho_k(n-1)\omega^k_{a,b} \\ & + M^l_a Cov(n-1)_{l,j} \Psi^j_b + m^k_a \Gamma ov(n-1)_{k,j} \psi^j_b \\ & + M^l_a C\Gamma ov(n-1)_{l,k} \psi^k_b + m^k_a C\Gamma ov(n-1)_{l,k} \Psi^l_b \end{aligned} \tag{5.3.11}$$

The above three equations can be written in matrix form and then grouped into one matrix equation with the matrices $Cov(n)$,$C\Gamma ov(n)$, $\Gamma ov(n)$, M, Ψ, m, ψ, V,Λ, $Omega$, v, λ and ω as blocks.

$$\left[\begin{array}{c|c} Cov(n) & C\Gamma ov(n) \\ \hline C\Gamma ov(n)^T & \Gamma ov(n) \end{array}\right] = \mu_l(n-1)\left[\begin{array}{c|c} V^l & \Omega^l \\ \hline \Omega^l & \Lambda^l \end{array}\right] + \rho_l(n-1)\left[\begin{array}{c|c} v^l & \omega^l \\ \hline \omega^l & \lambda^l \end{array}\right] + \left[\begin{array}{c|c} M^T & m^T \\ \hline \Psi^T & \psi^T \end{array}\right]\left[\begin{array}{c|c} Cov(n) & C\Gamma ov(n) \\ \hline C\Gamma ov(n)^T & \Gamma ov(n) \end{array}\right]\left[\begin{array}{c|c} M & \Psi \\ \hline m & \psi \end{array}\right] \tag{5.3.12}$$

Assuming $\vec{\mu}(0)$ reaches the fundamental eigenmode distribution, that is the delayed neutron scheme is launched after the inactive generations become stationary, then we have $\vec{\mu}(1)$ and $\vec{\rho}(1)$ according to Eq 5.2.117 and Eq 5.2.118, and Eq 5.3.2 gives $\vec{\mu}(n) = \vec{\mu}(n')$ and $\vec{\rho}(n) = \vec{\rho}(n')\ \forall n, n'$. Similar to the development around Eq 4.1.6 in section 4.1, we use the Feynman slash notation to define the response moments multiplied by the neutron distribution.

$$\not{V}_{i,j} \equiv \mu_l(n) V^l_{j,i} \tag{5.3.13}$$

$$\not{v}_{i,j} \equiv \rho_l(n) v^l_{j,i} \tag{5.3.14}$$

$$\not{\Omega}_{i,j} \equiv \mu_l(n) \Omega^l_{j,i} \tag{5.3.15}$$

$$\not{\omega}_{i,j} \equiv \rho_l(n) \omega^l_{j,i} \tag{5.3.16}$$

$$\not{\Lambda}_{i,j} \equiv \mu_l(n) \Lambda^l_{j,i} \tag{5.3.17}$$

$$\not{\lambda}_{i,j} \equiv \rho_l(n) \lambda^l_{j,i} \tag{5.3.18}$$

Defining the covariance among prompt and delayed neutrons into one matrix Kov,

$$Kov(n) \equiv \left[\begin{array}{c|c} Cov(n) & C\Gamma ov(n) \\ \hline C\Gamma ov(n)^T & \Gamma ov(n) \end{array}\right] \tag{5.3.19}$$

and defining the contracted response moments into one matrix $\not{W}$,

$$\not{W} \equiv \left[\begin{array}{c|c} \not{V} & \not{\Omega} \\ \hline \not{\Omega} & \not{\Lambda} \end{array}\right] + \left[\begin{array}{c|c} \not{v} & \not{\omega} \\ \hline \not{\omega} & \not{\lambda} \end{array}\right] \tag{5.3.20}$$

and recognizing the matrix G defined in Eq 5.3.3, the recursive formula for Kov is simplified to

$$Kov(n) = \not{W} + G^T Kov(n-1)G \quad n \geq 2 \tag{5.3.21}$$

which is analogous to Eq 4.1.3.

Recursively evaluating $Kov(n)$ according to Eq 5.3.21 leads to

$$Kov(n) = \sum_{g=n-1}^{1} (G^T)^{g-1} \not{W} G^{g-1} + (G^T)^{n-1} Kov(1) G^{n-1} \tag{5.3.22}$$

where $Kov(1)$ can be initialized with the results from inactive generations $C_{a,b}(0)$ and $\mu_a(0)$ according to the relations between moments of the first active generation and moments of the last inactive generation derived in section 5.2.5.4.

Finally, we explicitly write the temporal-spatial moment $\mathbb{E}[Z_l(n)Z_j(n+k)]$ but in the form of the corresponding central moments defined in Eq 5.3.23.

$$Cov(n, n+k)_{l,j} = \mathbb{E}[Z_l(n)Z_j(n+k)] - \mu_l(n)\mu_j(n+k) \tag{5.3.23}$$

Inserting Eq 5.2.81 along with the matrix form of the required $k - generation$ response moments in Eq 5.3.4 and Eq 5.3.5, and the definition of $Cov(n)$ and $C\Gamma ov(n)$ in Eq 5.3.6 and Eq 5.3.7, $Cov(n, n+k)$ is simplified to Eq 5.3.24.

$$\begin{aligned} Cov(n, n+k) &= \left(Cov(n) + \mu(n)^T \mu(n)\right) G^{(k)}(1,1) \\ &\quad + \left(C\Gamma ov(n) + \mu(n)^T \rho(n)\right) G^{(k)}(2,1) - \mu(n)^T \mu(n+k) \\ &= Cov(n) G^{(k)}(1,1) + C\Gamma ov(n) G^{(k)}(2,1) \\ &\quad + \mu(n)^T \left(\mu(n) G^{(k)}(1,1) + \rho(n) G^{(k)}(2,1)\right) - \mu(n)^T \mu(n+k) \end{aligned} \tag{5.3.24}$$

Evaluating Eq 5.3.2 k times and indexing for the μ part of the π vector

$$\mu(n+k) = \mu(n) G^{(k)}(1,1) + \rho(n) G^{(k)}(2,1) \tag{5.3.25}$$

Eq 5.3.25 simplifies Eq 5.3.24 to Eq 5.3.26

$$Cov(n, n+k) = Cov(n)G^{(k)}(1,1) + C\Gamma ov(n)G^{(k)}(2,1) \tag{5.3.26}$$

and is recognized as the $(1,1)$ block of size $M \times M$ of the $2M \times 2M$ matrix $Kov(n)G^{(k)}$, that is

$$Cov(n, n+k) = (Kov(n)G^{(k)})(1,1) \tag{5.3.27}$$

It can be shown that the temporal-spatial covariance matrix $Kov(n, n+k)$ defined in Eq 5.3.28

$$Kov(n, n+k) \equiv \left[\begin{array}{c|c} Cov(n, n+k) & C\Gamma ov(n, n+k) \\ \hline C\Gamma ov(n, n+k)^T & \Gamma ov(n, n+k) \end{array}\right], \tag{5.3.28}$$

can be related to the spatial covariance matrix $Kov(n)$ through the response matrix G as in Eq 5.3.29

$$Kov(n, n+k) = Kov(n)G^{(k)}. \tag{5.3.29}$$

The covariance of Z tallies in Eq 5.3.27 can be converted to normalized X tallies by Eq 4.2.8 . Then Eq 2.2.38 summarizes the covariance of X tallies into $\mathrm{Var}[\bar{X}(N)]$. Given matrix $\overline{Cov}(N)$ defined as

$$\overline{Cov(N)} \equiv \sum_{n=1}^{N} \mathrm{Cov}(n) + 2 \sum_{k=1}^{N-1} \sum_{n=1}^{N-k} \mathrm{Cov}(n, n+k) \tag{5.3.30}$$

$\mathrm{Var}[\bar{X}_I(N)]$ can be evaluated as

$$\mathrm{Var}[\bar{X}_I(N)] = \left(\frac{\mu_I(1)}{\mu(1)}\right)^2 \times \left[\frac{\sum\limits_{\substack{i\in W\\ j\in W}} \overline{\mathrm{Cov}(N)}_{i,j}}{\mu(1)^2} - \frac{\sum\limits_{\substack{i\in W\\ j\in I}} \overline{\mathrm{Cov}(N)}_{i,j}}{\mu(1)\mu_I(1)} - \frac{\sum\limits_{\substack{i\in I\\ j\in W}} \overline{\mathrm{Cov}(N)}_{i,j}}{\mu_I(1)\mu(1)} + \frac{\sum\limits_{\substack{i\in I\\ j\in I}} \overline{\mathrm{Cov}(N)}_{i,j}}{\mu_I(1)\mu_I(1)}\right] \tag{5.3.31}$$

where stationarity is assumed to make the summation over active generation and generation lag only act on numerators and I denotes the set of phase space regions contained in tally region I and W denotes the set of all phase space regions.

5.3.2 Numerical Optimization

Even with the recursive moments evaluation formula reorganized in matrix forms, there does not seem to exist a closed form for $\mathrm{Var}[\bar{X}_I(N)]$ with explicit α, f, p and N_0, s dependence. Therefore, in order to find the α, f and p parameters that lead to the optimal variance performance for a real problem the practical way is to try sufficiently many such parameter combinations on a DMBP model constructed for the real problem, then select the optimal parameters for the DMBP model and use these parameters to perform simulation on the real problem.

Before searching for the optimal parameters, a few more simulations schemes (with different α, f and p combinations) on problem $TD1$ are performed to numerically verify the prediction capability of the DMBP model. In addition to $TD1D1_n$ with $\alpha = 0.3$ and $f = 0.7$, the other simulation schemes tested are $\alpha = 0.1, f = 0.9$ (denoted as $TD1D2_n$), $\alpha = 0.1, f = 0.1$ (denoted as $TD1D3_n$), $\alpha = 0.1, f = 0.999$ (denoted as $TD1D4_n$). In all the simulation schemes, there are $s = 32000$ neutrons per active generation and therefore s/α neutrons per inactive generation. 2000 independent simulations were run to generate reference solutions.

References and prediction of $\mathrm{Var}[\bar{X}_I(N)]$ as function of active generation N for $I = 1$ and 8 are plotted in Figure 5-9. The predicted variances match references very

well.

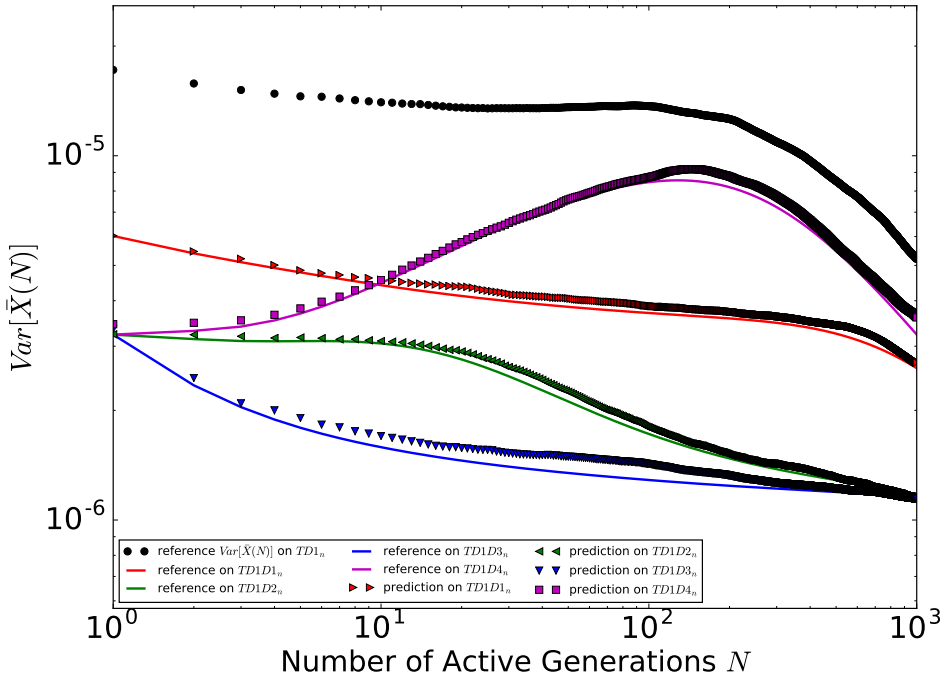
Proving the agreement of $\text{Var}[\bar{X}_I(N)]$ with the DMBP models , more predictions are made in Figure 5-10. First, we observe that $\text{Var}[\bar{X}_I(1)]$ is the same for all simulations with the same α. This is because generation 1 tallies are contributed to by neutrons sampled with probability α from a correlated neutron bank and the variances of the tallies are completely determined by α according to Eq 5.2.125. $\text{Var}[\bar{X}_I(1)]$ decreases for lower α values since all simulations have the same prompt source bank size and lower α corresponds to sampling from a larger source bank in the first active generation.

For the same α value, $\text{Var}[\bar{X}_I(N)]$ increases with higher f. Intuitively, since all simulations have the same prompt neutron bank size, the closer f is to 1, the closer the delayed scheme is to the traditional simulation without delayed neutrons. For the predictions with $f = 0.999$, $\text{Var}[\bar{X}_I(N)]$ increases as a function of N for $N < 200$ to approach the behavior of $\text{Var}[\bar{X}_I(N)]$ of the traditional method. This was also observed with reference values in Figure 5-9. In fact, the increase of $\text{Var}[\bar{X}_I(N)]$ as function of f can be attributed to higher correlation coefficients with a larger f. This implies that more neutrons are sampled from the prompt source bank, simulated and then sampled into the prompt source bank of next generation.

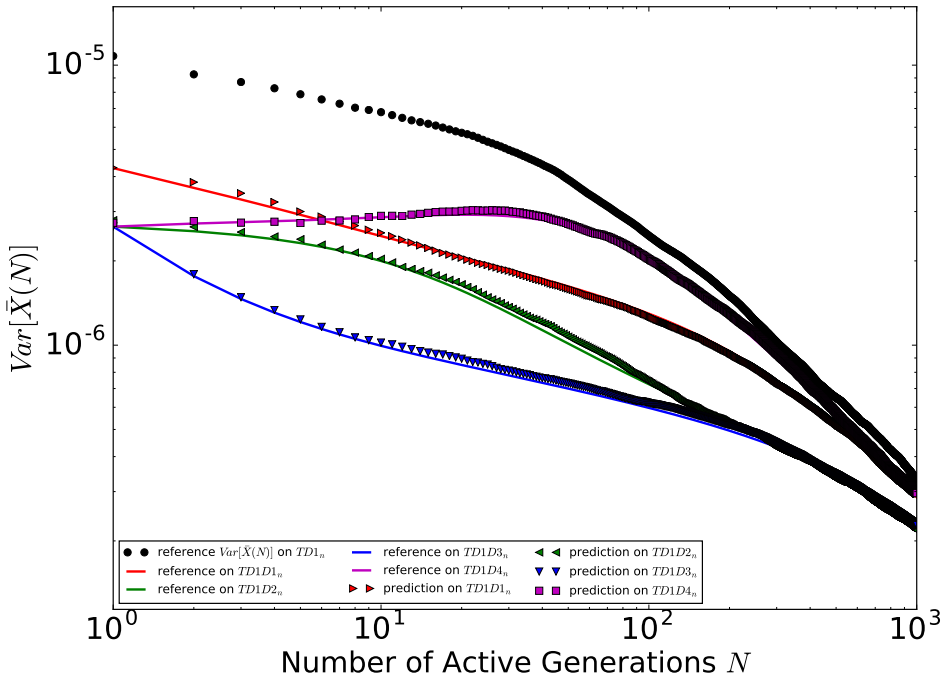
The above discussion shows the preference of low α and low f for better variance behaviors.

However, the issue of bias emerges for extremely low α's. As discussed in section 2.2, expectation of square error measures the performance of tallies in terms of the average distance from the corresponding true value. The variance of tallies have been used since the true value is not available to evaluate the error and the variance is equal to expected square error if the estimator is unbiased (Eq 2.2.9).

The discrete problem $TD1$ is so simple that the issue of bias does not become apparent. We now turn to investigate the performance of delayed methods on the continuous problem $TC1$. On $TC1$ we know the true value of tallies and can thus evaluate real square errors as well as the variance of tallies to view the impact of biases.

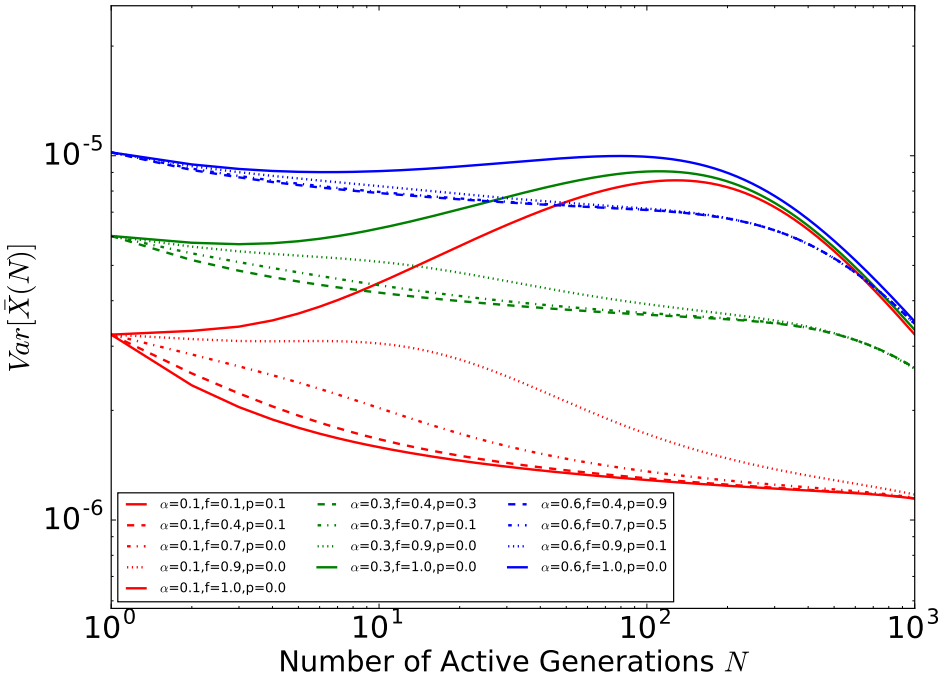


(a) Real variance $\mathrm{Var}\,\bar{X}_1(N)$

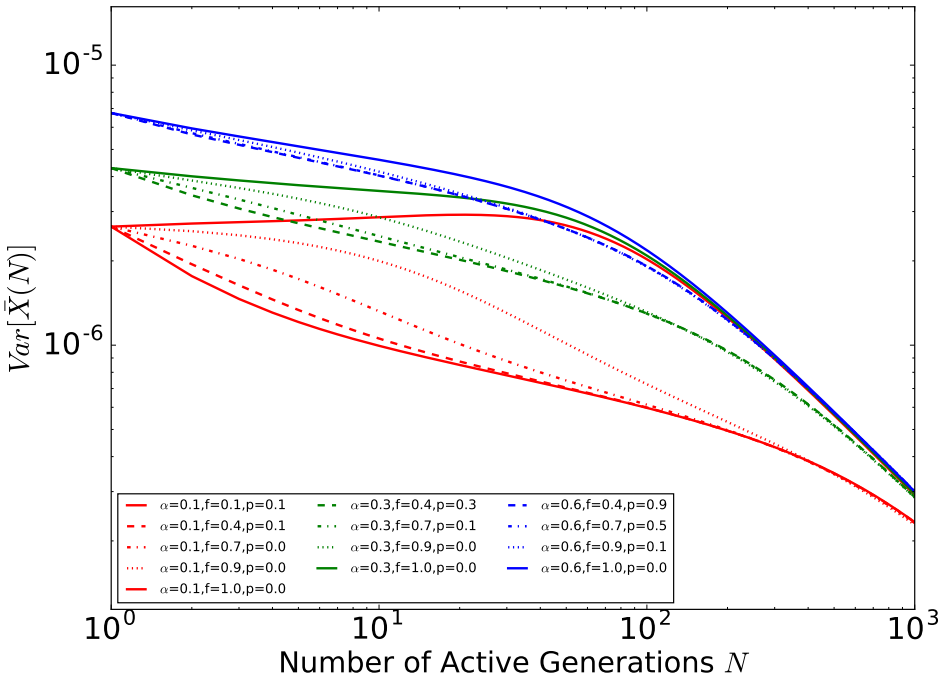


(b) Real variance $\mathrm{Var}\,\bar{X}_8(N)$

Figure 5-9: $\mathrm{Var}\,\bar{X}(N)$ *of normalized tallies on test problem* $TD1$ *with different delayed neutron schemes.*



(a) Real variance Var $\bar{X}_1(N)$



(b) Real variance Var $\bar{X}_8(N)$

Figure 5-10: Var $\bar{X}(N)$ *of normalized tallies on test problem* $TD1$

Three α values 0.1, 0.3 and 0.6 are selected and for each case and the lowest limit of f is used. From these parameters, the p value is now set. Denote these problems as $TC1D1$ ($\alpha = 0.1, f = 0$), $TC1D2$ ($\alpha = 0.3, f = 0$) and $TC1D3$ ($\alpha = 0.6, f = 0.34$). Note for $\alpha > 0.5$, the lowest allowed f is $2 - 1/\alpha$.

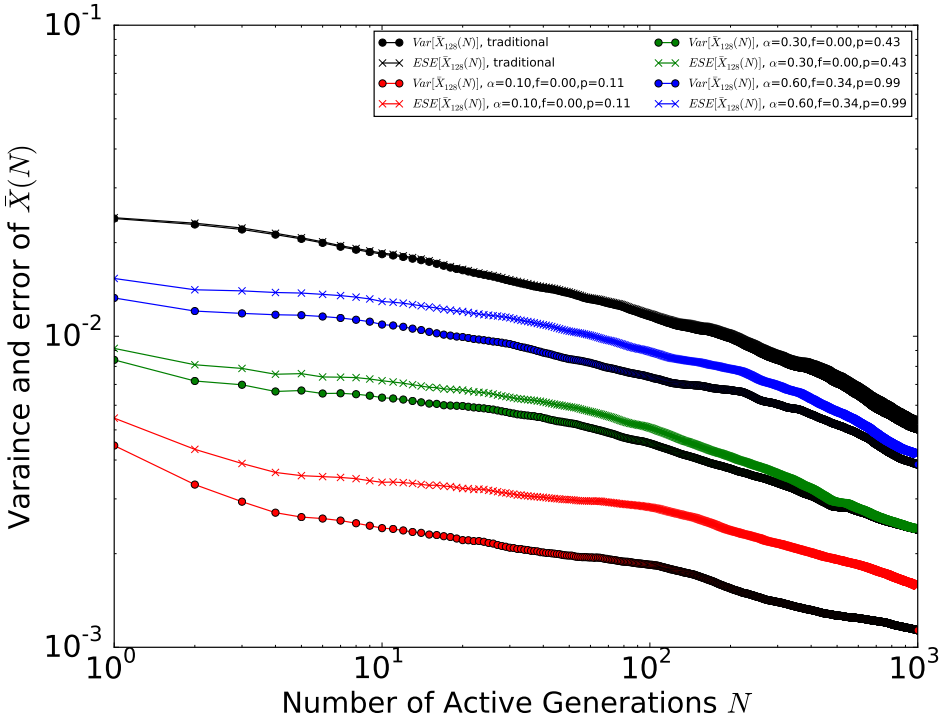


Figure 5-11: Variance and expected Square error (normalized by squared tally reference) of X tallies on test problem $TC1$

Results for the $TC1$ problem, in Figure 5-11, show that low values of α induce a larger bias that cannot be neglected in comparison with the tally variance. Therefore, the bias of the tallies must be used as a constraint when finding an optimal delayed scheme. One potential solution is to use the DMBP model to predict both the variance and the bias of the tallies and perform an optimization on the expectation of square error. However, the predictive model can only evaluate the expectation of tallies and have no knowledge of the true mean needed to evaluate the bias.

Although the bias decreases at an exponential rate and the variance decreases at a power rate asymptotically, in the simulations with low α, the bias can become comparable to the variance at a finite number of generations. In the two sections

below, we propose a way to estimate the bias reduction rate. In section 5.3.2.1, we use the expectation of X tallies and the known exact solution to predict the bias of X tallies. Then based on the predicted and numerical bias behavior over different delayed schemes on the $TD1$ problem, an approximate method that relates the bias of X tallies to the age distribution of neutrons in the prompt source bank is developed in section 5.3.2.2.

5.3.2.1 Bias of $TD1$

For problem $TD1$, the tallies X_l's are estimating the spatial neutron distribution π_l specified in Eq 2.1.6. Therefore, the bias of $\bar{X}_l(N)$ can be predicted as

$$b(\bar{X}_l(N)) = \mathbb{E}\bar{X}_l(N) - \pi_l \tag{5.3.32}$$

with π_l known in advance. Despite $b(X_l(N))$ being a more canonical metric, $b(\bar{X}_l(N))$ is chosen due to the lower statistical noise in calculating numerical references. $b(\bar{X}_l(N))$ can be expressed in terms of $b(X_l(N))$ as

$$b(\bar{X}_l(N)) = \frac{\sum_{n=1}^{N} \left(\mathbb{E}X_l(n) - \pi_l\right)}{N} \tag{5.3.33}$$

Then $\mathbb{E}X_l(n)$ can be predicted according to Eq 3.3.61, Eq 3.3.53, Eq 3.3.54 and Eq 3.3.55 as shown in Eq 5.3.34.

$$\begin{aligned}
\mathbb{E}X_l(n) &= \frac{\mu_l(n)}{\mu(n)} + \frac{1}{2}\sum_{i,j}\left(2\frac{\mu_l(n)}{\mu(n)^3} - \frac{\delta_{l,j}}{\mu(n)^2} - \frac{\delta_{l,i}}{\mu(n)^2}\right)\left(C_{i,j}(n) - \mu_i(n)\mu_j(n)\right) \\
&= \frac{\mu_l(n)}{\mu(n)}\left(1 + \frac{\sum_{i,j} C_{i,j}(n)}{\mu(n)^2}\right) - \frac{\sum_j C_{l,j}(n)}{\mu(n)^2}
\end{aligned} \tag{5.3.34}$$

Eq 5.3.33 and Eq 5.3.34 can provide a prediction of $b(\bar{X}_l(N))$.

Next, we estimate the numerical reference of $b(\bar{X}_l(N))$ from independent simula-

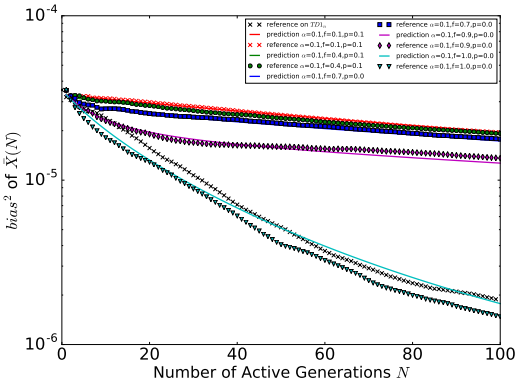
tions. According to Eq 2.2.9, $b(\bar{X}_l(N))$ satisfies

$$b(\bar{X}_l(N)^2 = \mathbb{E}[(\bar{X}_l(N) - \pi_l)^2] - \mathrm{Var}[\bar{X}_l(N)] \tag{5.3.35}$$

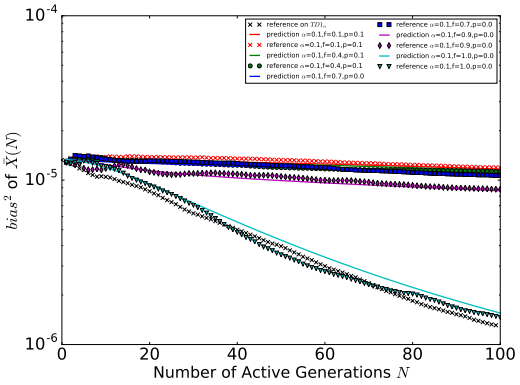
The numerical reference of $b(\bar{X}_l(N))^2$ can be obtained by estimating the numerical references of the expected square error and variance using the independent simulations.

The predicted and reference values of $b(\bar{X}_l(N))$ as a function of N for region 1 and 8 under different delayed neutron schemes are plotted in Figure 5-13. For better readability, results from the delayed scheme with different α's are plotted in different rows. Firstly, comparing figures in each column in Figure 5-13 shows that after the fixed number of inactive generations, $b(\bar{X}_l(1))^2$ is roughly the same for different α's (different source bank sizes during the active generations). Bias does not change with the number of neutrons per generation. Secondly, for the first 100 active generations, the bias of the delayed scheme with $f = 0.999$ resembles that of the traditional simulation method without delayed neutrons. This is because the tallies are dominated by the neutrons that are sampled into the prompt source bank at each generation. Thirdly, although the predicted bias does not match reference values exactly, the relative trend across different delayed schemes are conserved.

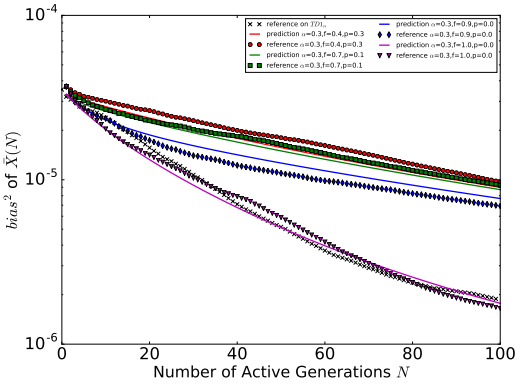
With the good accuracy of the predicted $b(\bar{X}_l(N))$ using the DMBP models demonstrated, additional predictions on $b(X_l(N))$ are made in Figure 5-13. For most of the cases, the bias of $X_l(N)$ decreases exponentially as a function of N. The rate of decrease is smaller for lower α's as fewer neutrons are transported at each active generation. For a given α, the bias decrease rate is faster early on but slower asymptotically at large N for higher f's. At early active generations, for higher f's, the fraction of neutrons that have been sampled at each generation dominates and the behavior of tallies from these neutrons is close to that of a simulation without delayed neutrons. However, as the simulation continues, the neutrons that have been transported for only a few generations accumulate in the delayed source bank and start to contribute to the tally with non-negligible effect. More quantitative analysis is



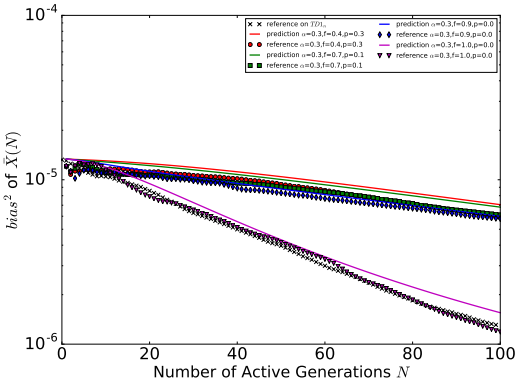
(a) $b(\bar{X}_1(N))^2$ when $\alpha = 0.1$



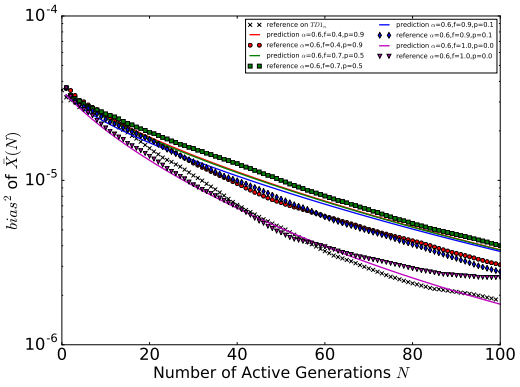
(b) $b(\bar{X}_8(N))^2$ when $\alpha = 0.1$



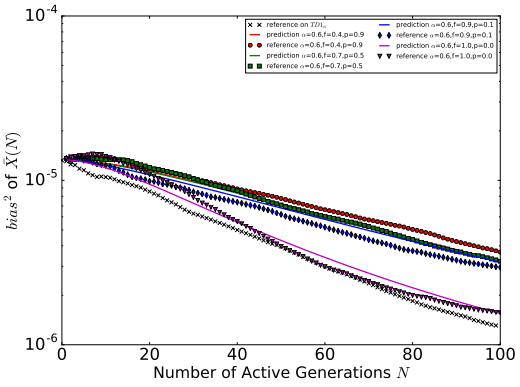
(c) $b(\bar{X}_1(N))^2$ when $\alpha = 0.3$



(d) $b(\bar{X}_8(N))^2$ when $\alpha = 0.3$



(e) $b(\bar{X}_1(N))^2$ when $\alpha = 0.6$



(f) $b(\bar{X}_8(N))^2$ when $\alpha = 0.6$

Figure 5-12: bias of $\bar{X}(N)$ of normalized tallies on test problem $TD1$

given in the next section by investigating the neutron age distribution at each active generation.

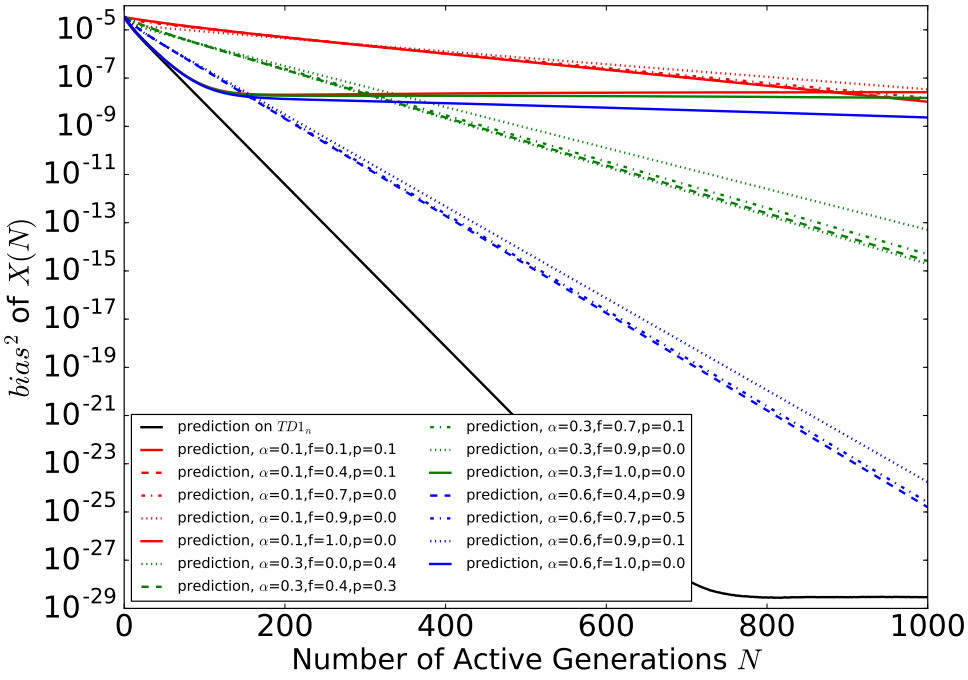
5.3.2.2 Bias and neutron age distribution

All neutrons in the source bank before splitting into prompt and delayed source banks have an age of 0. We denote the fraction of neutrons of age i in the prompt source bank at generation n as $r_{Z,n,i}$ and the fraction of neutrons of age i in the delayed source bank at generation n as $r_{U,n,i}$. We also denote the corresponding expected neutron counts as $s_{Z,n,i}$ and $s_{U,n,i}$ respectively. The age distribution at any generation n can be evaluated recursively given an initial distribution at generation 1. In the derivations below, for simplicity, we will assume that the transport process does not change the source bank size. Therefore, there are always αs neutrons in the prompt source bank and $(1-\alpha)s$ neutrons in the delayed source bank. For the neutrons of age j in the prompt bank at generation n, one part comes from the intermediate fission bank between generations $n-1$ and n with the sampling probability f, these neutrons were at age $j-1$ in the prompt source bank at generation $n-1$. And the other part is sampled with probability p from the delayed source bank at generation $n-1$ with these neutrons already at age j. Therefore, $s_{Z,n,j}$ can be evaluated from $s_{Z,n-1,j-1}$ and $s_{U,n-1,j}$ as shown in Eq 5.3.36.

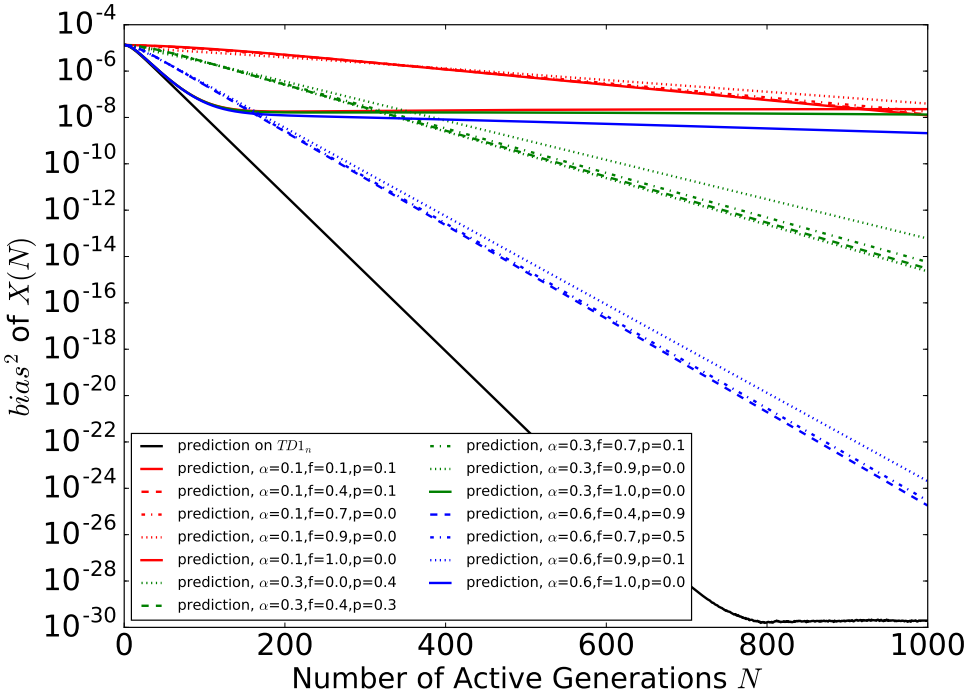
$$s_{Z,n,j} = f s_{Z,n-1,j-1} + p s_{U,n-1,j} \tag{5.3.36}$$

For neutrons of age j in the delayed bank at generation n, one part comes from the intermediate fission bank between generation $n-1$ and n with the sampling probability $1-f$, these neutrons were at age $j-1$ in the prompt bank at generation $n-1$. The other part is sampled with probability $1-p$ from the delayed source bank at generation $n-1$ with these neutrons already at age j. Therefore, $s_{U,n,j}$ can be evaluated from $s_{Z,n-1,j-1}$ and $s_{U,n-1,j}$ as shown in Eq 5.3.37.

$$s_{U,n,j} = (1-f) s_{Z,n-1,j-1} + (1-p) s_{U,n-1,j} \tag{5.3.37}$$



(a) Bias of $X_1(N)$



(b) Bias of $X_8(N)$

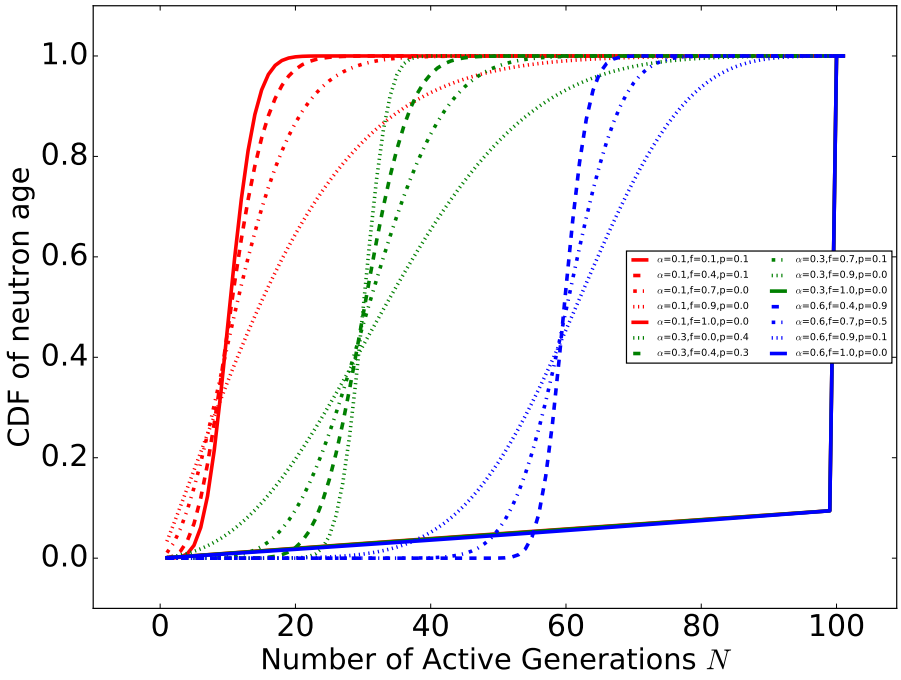
Figure 5-13: bias of $X(N)$ of normalized tallies on test problem $TD1$

$r_{Z,n,i}$ and $r_{U,n,i}$ can be calculated by recursively calling Eq 5.3.36 and Eq 5.3.37 with $s_{Z,1,0} = \alpha s$, $s_{U,1,0} = (1-\alpha)s$ for any s and finally converting $s_{Z,n,i}$ and $s_{U,n,i}$ to $r_{Z,n,i}$ and $r_{U,n,i}$ by Eq 5.3.38.

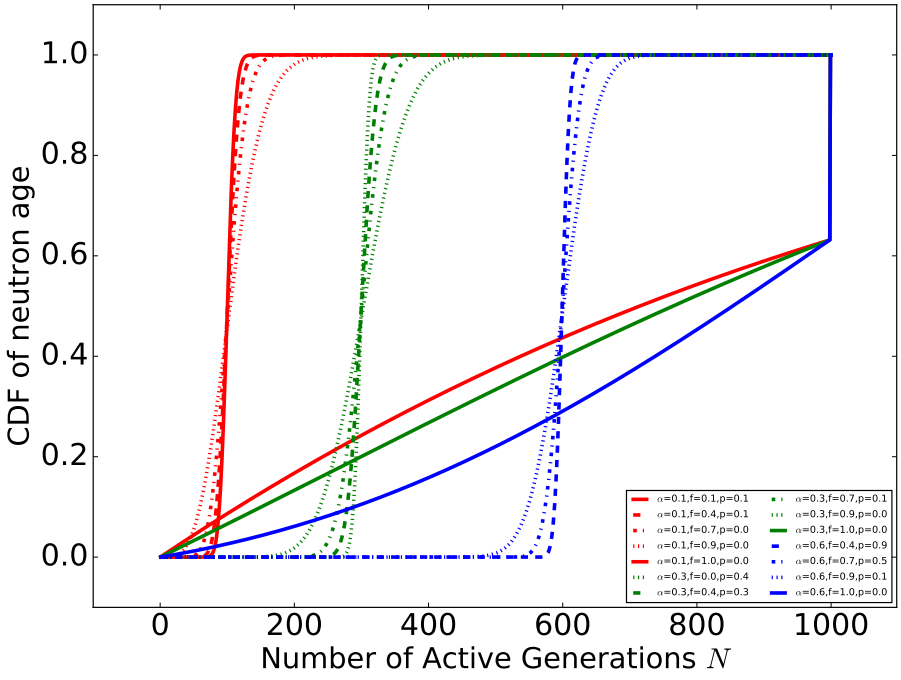
$$\begin{aligned} r_{Z,n,i} &= \frac{s_{Z,n,i}}{\alpha s} \\ r_{U,n,i} &= \frac{s_{U,n,i}}{(1-\alpha)s} \end{aligned} \tag{5.3.38}$$

The cumulative distribution function (CDF) of age distribution for several delayed schemes are plotted in Figure 5-14. CDF's rather than the probability mass function $r_{Z,n,i}$ derived above are plotted because $r_{Z,n,i}$'s are distributions with a single very narrow peak, peak values at different scales from different delayed schemes make it difficult to compare all the α, f, p combinations in the same figure. Figure 5-14(a) plots the cumulative neutron age distributions at generation 100 and Figure 5-14(b) plots at generation 999. The steep increases in the neutron age CDF corresponds to the peak of the neutron age distribution. At generation N, the peak of neutron age distribution is approximately located at αN. Since at each generation, a fraction α of the neutrons are transported, after N generations, the most probable neutron age should be around αN.

Figure 5-14 also shows that for the same α, lower f corresponds to a narrower peak of neutron age distribution and higher f corresponds to a wider peak. With $f \sim 1$, the CDF looks like a straight line up to $n-1$, which implies an almost flat $r_{Z,n,i}$ as a function of i up to $n-1$ followed by a large contribution of neutron of age $n-1$. This means that when f is very close to 1, at any generation n, most neutrons are at age $n-1$ since they were sampled into the prompt source bank during all previous generations. The tradeoff of almost all neutrons being sampled into the prompt fission bank at each generation is that the neutrons that were sampled into the delayed source bank would remain there and contribute to the tally with lower age if sampled into the prompt source bank. As the number of active generations n increases, the dominant neutrons of age $n-1$ occupies a smaller fraction of the prompt source bank and leads to worse bias behavior as observed previously.



(a) CDF of neutron age distribution at generation 100



(b) CDF of neutron age distribution at generation 999

Figure 5-14: CDF of neutron age distribution under different delayed neutron schemes

Now we can provide an estimate for the relation between bias of a tally $X(N)$ contributed by all the $s\alpha$ neutrons in the prompt source bank and the age distribution among the $s\alpha$ neutrons. Suppose the observable of interest estimated by $X(N)$ is actually equal to μ, then the bias of $X(N)$ is defined as

$$b(X(N))_{\{s_{Z,N,i}\}_{i=0}^{N-1}} = \mathbb{E}X(N) - \mu \tag{5.3.39}$$

where the dependence on age distribution is explicitly specified.

Among the αs neutrons, denote the contribution to tally $X(N)$ by th j^{th} neutron on age i as $\xi_i^{(j)}$ and assume tally $X(N)$ can be simply expressed as sum of the $\xi_i^{(j)}$'s.

$$X(N) = \sum_{i=0}^{N-1} \sum_{j=1}^{s_{Z,N,i}} \xi_i^{(j)} \tag{5.3.40}$$

Then, the bias $b(X(N))_{\{s_{Z,N,i}\}_{i=0}^{N-1}}$ expression can be simplified as below.

$$b(X(N))_{\{s_{Z,N,i}\}_{i=0}^{N-1}} = \sum_{i=0}^{N-1} \sum_{j=1}^{s_{Z,N,i}} \mathbb{E}\xi_i^{(j)} - \mu \tag{5.3.41}$$

Since the age distribution at any generation is normalized, that is

$$\sum_{i=0}^{n-1} r_{Z,n,i} = 1 \quad \forall n, \tag{5.3.42}$$

Eq 5.3.41 can be equivalently rewritten as

$$b(X(N))_{\{s_{Z,N,i}\}_{i=0}^{N-1}} = \sum_{i=0}^{N-1} \left(\sum_{j=1}^{s_{Z,N,i}} \mathbb{E}\xi_i^{(j)} - r_{Z,N,i}\mu \right). \tag{5.3.43}$$

Then, we recognize that $s_{Z,N,i} = s\alpha r_{Z,N,i}$, meaning that all the $s_{Z,N,i}$ expectations in Eq 5.3.43 are identical, the above equation is simplified to

$$b(X(N))_{\{s_{Z,N,i}\}_{i=0}^{N-1}} = \sum_{i=0}^{N-1} \left(r_{Z,N,i} s_f \mathbb{E}\xi_i^{(j)} - r_{Z,N,i}\mu \right). \tag{5.3.44}$$

where $s_f = \alpha s$ is the size of prompt source bank. Eq 5.3.44 can be reorganized to be more similar to the definition in Eq 5.3.39 by recognizing the $s_f \mathbb{E}\xi_i^{(j)}$ as contribution from s_f neutrons with age i.

$$b(X(N))_{\{s_{Z,N,i}\}_{i=0}^{N-1}} = \sum_{i=0}^{N-1} r_{Z,N,i} \left(\sum_{j=1}^{s_f} \mathbb{E}\xi_i^{(j)} - \mu \right). \tag{5.3.45}$$

The term within the parenthesis in Eq 5.3.45 is understood as the bias of $X(N)$ if all the s_f neutrons were at age i and is denoted as $b(X(i))_{s_f}$. We can therefore approximate the bias of the tally contributed by s_f neutrons with age distribution $\{r_{Z,N,i}\}_{i=0}^{N-1}$ by the weighted average of the bias of the tally contributed by all s_f neutrons at all the $i = 0, \cdots, N-1$ generations with the age fraction at each generation as the weights.

$$b(X(N))_{\{s_{Z,N,i}\}_{i=0}^{N-1}} = \sum_{i=0}^{N-1} r_{Z,N,i} b(X(i))_{s_f}. \tag{5.3.46}$$

Assuming $b(X(i))_{s_f}^2 = c^i$, Eq 5.3.46 can estimate $b(X(N))_{\{s_{Z,N,i}\}_{i=0}^{N-1}}$ for any specified α, f, p. Figure 5-15 plots the estimated $bias^2(X(N))$ for all the delayed schemes discussed above. All features of the $bias^2(X(N))$'s observed in Figure 5-13 are conserved in Figure 5-15 despite the shift in the magnitude of the curves.

Eq 5.3.46 can now be used to put constraints on the bias of tallies when finding an optimal run strategy as follows. Since $\mathrm{Var}[\bar{X}(N)]$ decreases according to the power law $1/N$ (or slower for finite N's) while $bias^2(X(N))$ decreases exponentially, $bias^2(\bar{X}(N))$ decreases faster than $\mathrm{Var}[\bar{X}(N)]$ asymptotically. However, with delayed neutrons, low α could lead to non-negligible bias, that is larger or comparable in magnitude to the variance. Assuming the inactive generations end up with $bias^2(X(1)) \ll \mathrm{Var}[X(1)]$, with the $\mathrm{Var}[\bar{X}(N)]$ predicted from DMBP model and $bias^2(\bar{X}(N))$ estimated from $bias^2(X(N))$ according to Eq 5.3.46, the constraint can be enforced as that the relative change of bias cannot be smaller than that of the variance, which is stated in Eq 5.3.47.

$$\frac{bias^2(\bar{X}(N))}{bias^2(X(1))} \leq \frac{\mathrm{Var}[\bar{X}(N)]}{\mathrm{Var}[X(1)]} \tag{5.3.47}$$

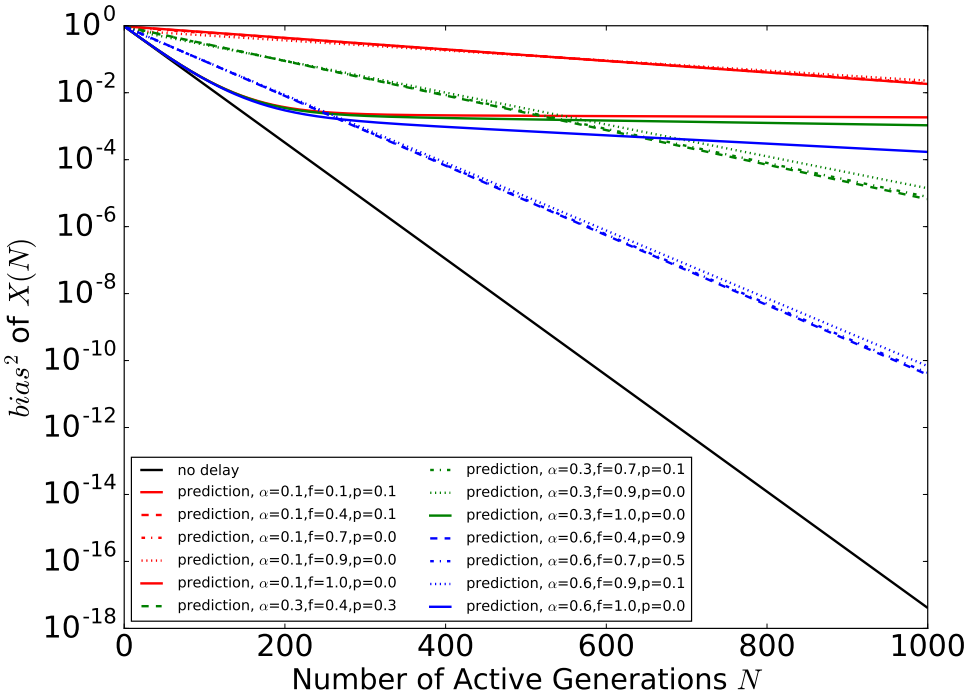


Figure 5-15: bias of normalized tallies and age distribution

5.3.2.3 Steps on finding optimal delayed scheme

The optimal running strategy is found as a constrained optimization problem. We denote the number of inactive generations as N_0, the cost of simulating one neutron in inactive generations as β, (suppose the cost per neutron in active generations is 1), the number of neutrons per active generation as s. Then the total cost after N active generations is

$$N_c = \beta N_0 s/\alpha + Ns. \tag{5.3.48}$$

The variance of a tally for any given region I over the N generations, $\mathrm{Var}[\bar{X}(N)]$, can be predicted according to Eq 5.3.31. Then optimization of either the Type 1 or Type 2 can be performed to identify optimal delayed scheme parameters α, f, p.

5.4 Numerical results on 2D BEAVRS benchmark

Three different delayed schemes are used to perform simulations on the BEAVRS benchmark. We denote these problems as $BEAVRS - D1$ ($\alpha = 0.1, f = 0$), $BEAVRS - D2$ ($\alpha = 0.3, f = 0$) and $BEAVRS - D3$ ($\alpha = 0.6, f = 0.34$). In each case, neutron source counts are tallied on an assembly size mesh from 1000 active generations of size $s = 400,000$. 800 independent simulations are used to generate reference solutions for $\text{Var}[\bar{X}(N)]$. And $\text{Var}[\bar{X}(N)]/\mu_X^2$ calculated from independent simulations is used as a reference for the relative square error (RSE defined in Eq 2.2.12).

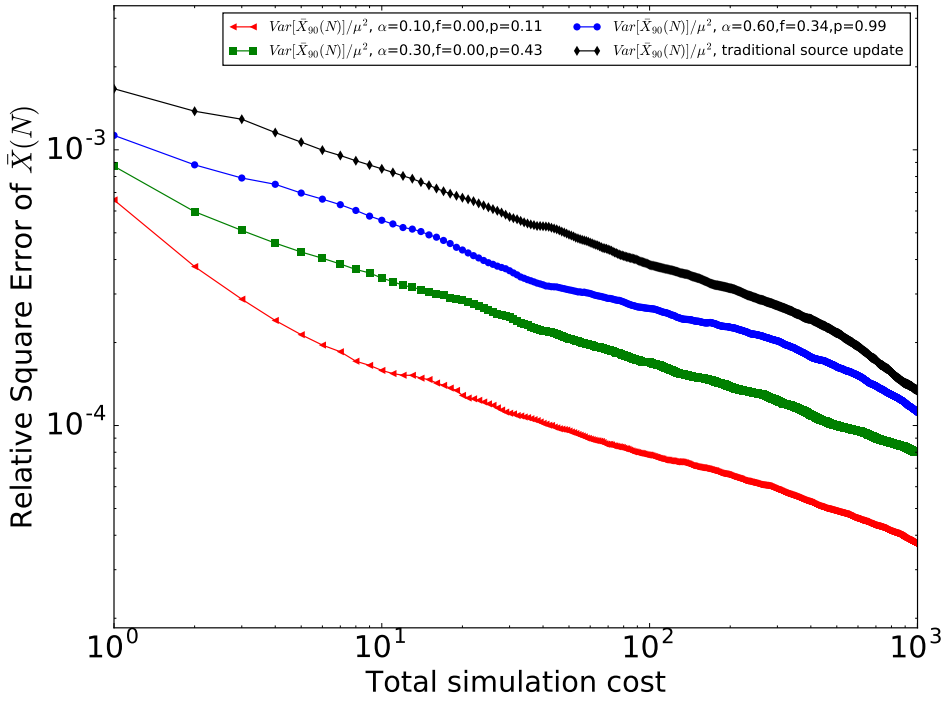
The numerical reference $RSE_I(N)$'s for Assembly 2 and 3 in the BEAVRS benchmark (Figure 3-9) are then plotted vs total computation cost $N_c(N, N_0, s, \alpha, \beta)$ calculated according to Eq 5.3.48.

Figure 5-16 plots the convergence of RSE's with $N_0 = 0$. As expected, the delayed neutron method reaches any RSE threshold with fewer neutrons simulated when neglecting the cost of the inactive generations.

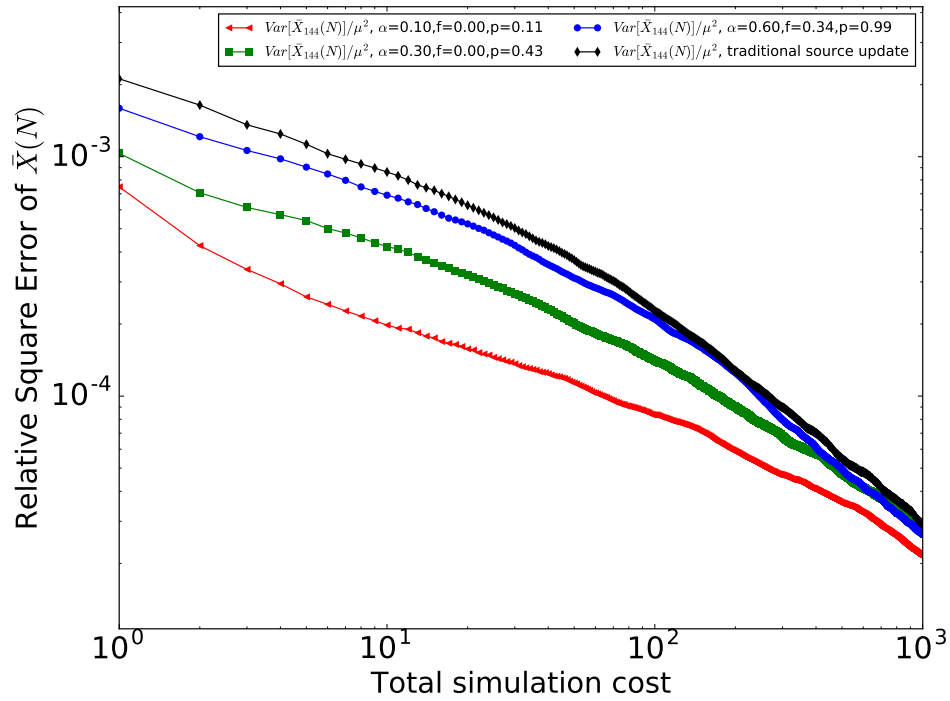
Next, we compare the RSE convergence rate for different values of N_0 and β. Typically, the 2D BEAVRS benchmark is assumed stationary after 200 inactive generations. Employing $CMFD$ could further reduce the number of inactive generations by a factor of 10 [22]. Figures below consider $N_0 \in \{20, 200\}$. Due to the extra work needed in evaluating tallies, the cost per neutron in active generations is usually higher than in inactive generations. Therefore $\beta \leq 1$ and the results below consider $\beta \in \{0.1, 0.5, 1\}$.

$RSE(N)$ in Assembly 2 and 3 over all the delayed schemes and all inactive cost conditions are plotted in Figure 5-17 and Figure 5-18, respectively.

Figure 5-17 and Figure 5-18 show that the performance of the delayed methods depend largely on the tally region cost of inactive generations. In Assembly 3, if $N_0 = 200$, delayed methods reach the 10^{-4} RSE (or %1 RMS) slower than the traditional method; if $N_0 = 20$, the advantage of delayed methods becomes more effective for smaller β's. In Assembly 2, delayed methods reach the 10^{-4} RSE (or %1 RMS) faster

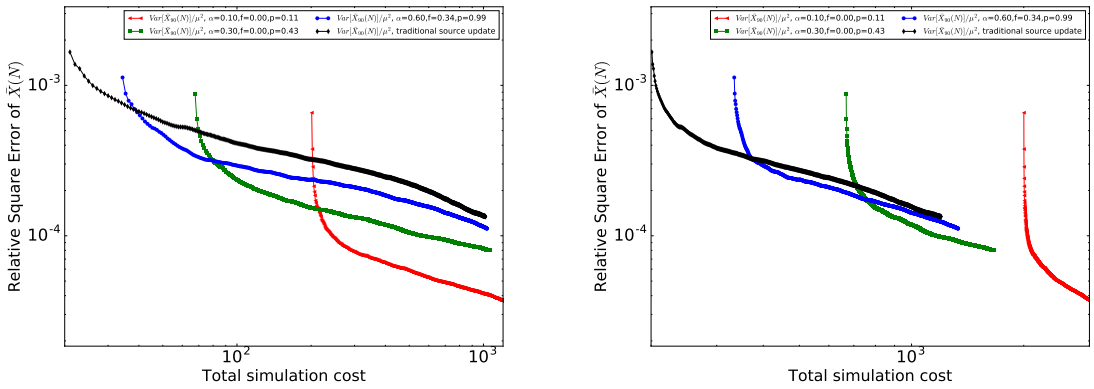


(a) Assembly 2

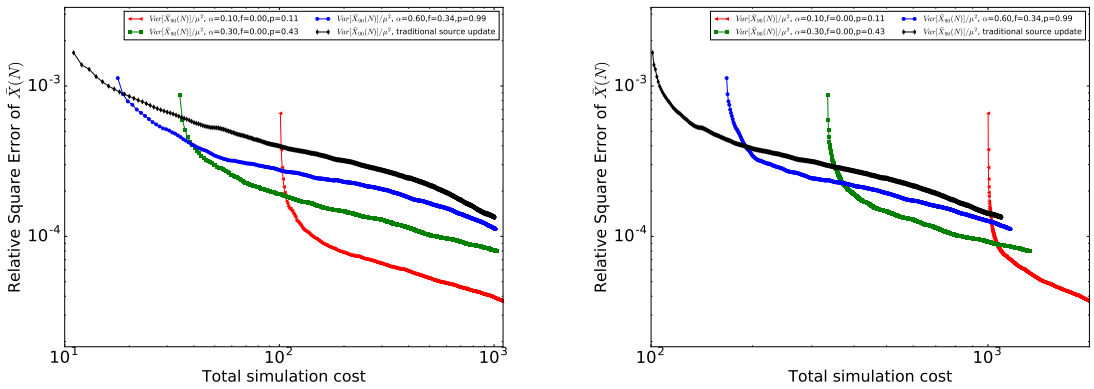


(b) Assembly 3

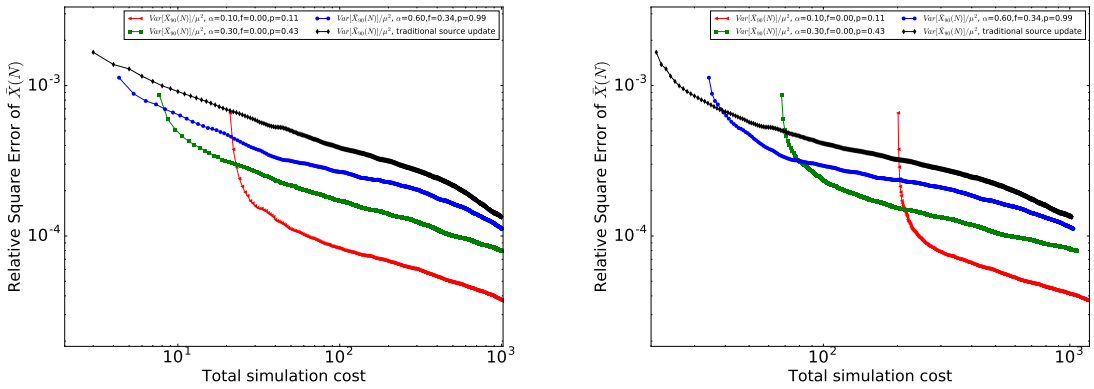
Figure 5-16: $RSE_I(N)$ vs N_c with $N_0 = 0$



(a) $N_0 = 20$,$\beta = 1.0$ (b) $N_0 = 200$,$\beta = 1.0$

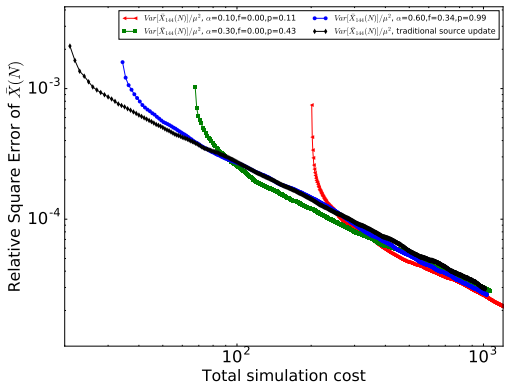


(c) $N_0 = 20$,$\beta = 0.5$ (d) $N_0 = 200$,$\beta = 0.5$

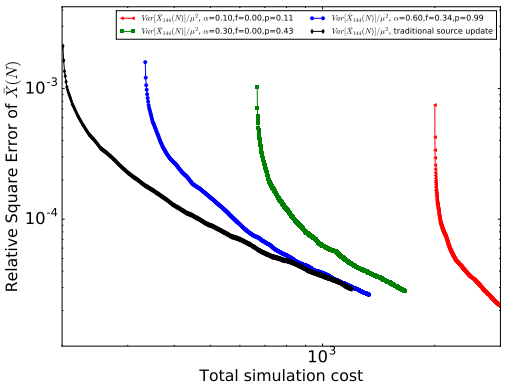


(e) $N_0 = 20$,$\beta = 0.1$ (f) $N_0 = 200$,$\beta = 0.1$

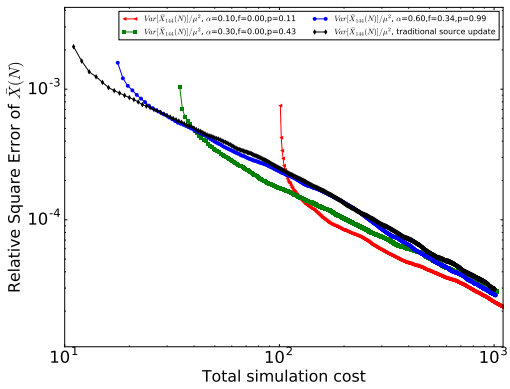
Figure 5-17: $RSE(N)$ vs N_c with $N_0 \in \{20, 200\}$ and $\beta \in \{0.1, 0.5, 1\}$ in Assembly 2



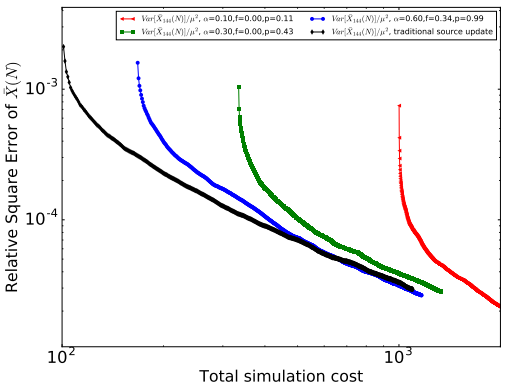
(a) $N_0 = 20$,$\beta = 1.0$



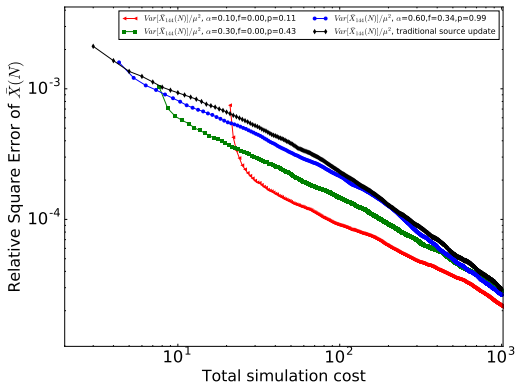
(b) $N_0 = 200$,$\beta = 1.0$



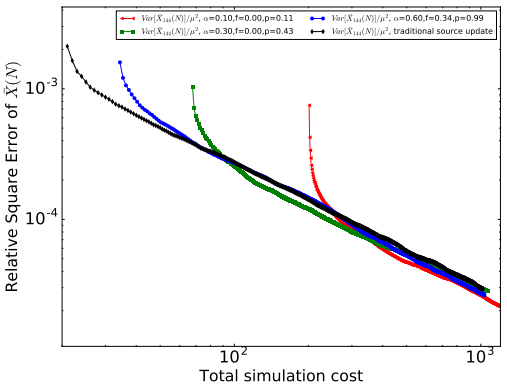
(c) $N_0 = 20$,$\beta = 0.5$



(d) $N_0 = 200$,$\beta = 0.5$



(e) $N_0 = 20$,$\beta = 0.1$



(f) $N_0 = 200$,$\beta = 0.1$

Figure 5-18: $RSE(N)$ vs N_c with $N_0 \in \{20, 200\}$ and $\beta \in \{0.1, 0.5, 1\}$ in Assembly 3

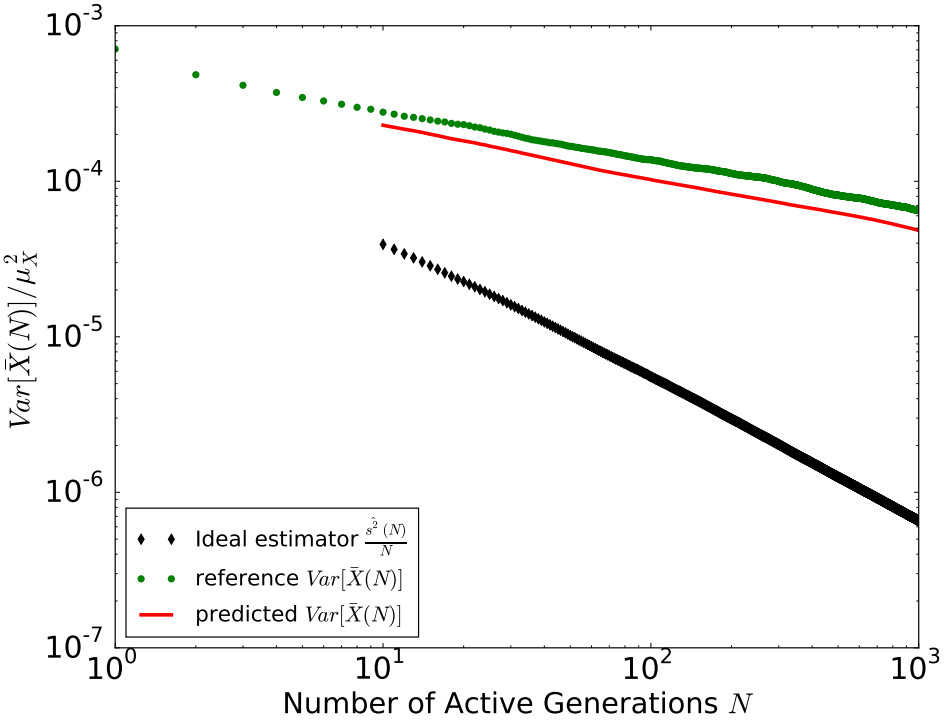
than the traditional method for all the cases of inactive generation cost.

The relative speedup to reach %1 RMS is summarized in Table 5.1.

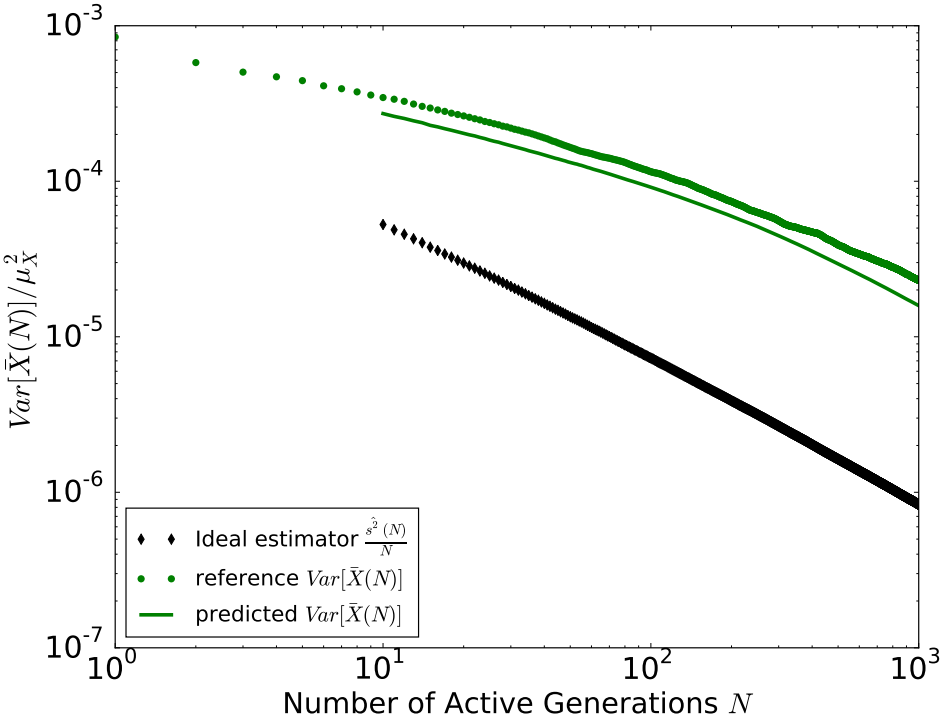
We also show in Figure 5-19 that the variance estimator can be corrected with predicted correlations for the selected delayed scheme $\alpha = 0.3, f = 0.0$.

Table 5.1: Performance of delayed neutron method

Inactive cost		Delayed scheme			speedup to
N_0	β	α	f	p	reach 1% RMS
200	1.0	1.0	1.0	0.0	1.0
200	1.0	0.1	0.0	0.11	~ 1
200	1.0	0.3	0.0	0.43	~ 2
200	1.0	0.6	0.34	0.99	~ 1.3
200	0.5	1.0	1.0	0.0	1.0
200	0.5	0.1	0.0	0.11	~ 2
200	0.5	0.3	0.0	0.43	~ 2.5
200	0.5	0.6	0.34	0.99	~ 2
200	0.1	1.0	1.0	0.0	1.0
200	0.1	0.1	0.0	0.11	~ 8
200	0.1	0.3	0.0	0.43	~ 3
200	0.1	0.6	0.34	0.99	~ 2
20	1.0	1.0	1.0	0.0	1.0
20	1.0	0.1	0.0	0.11	~ 8
20	1.0	0.3	0.0	0.43	~ 3.6
20	1.0	0.6	0.34	0.99	~ 2
20	0.5	1.0	1.0	0.0	1.0
20	0.5	0.1	0.0	0.11	~ 7
20	0.5	0.3	0.0	0.43	~ 4
20	0.5	0.6	0.34	0.99	~ 2
20	0.1	1.0	1.0	0.0	1.0
20	0.1	0.1	0.0	0.11	~ 29
20	0.1	0.3	0.0	0.43	~ 4
20	0.1	0.6	0.34	0.99	~ 2



(a) Var $\bar{X}(N)$ reference and correction on Assembly 2



(b) Var $\bar{X}(N)$ reference and correction on Assembly 3

Figure 5-19: Variance correction under the selected delayed source update scheme. $\mathrm{Var}[\bar{X}(N)]$ *is plotted as function of active generation number* N*. It can be transformed into function of computation cost as in Figure 5-17 by scaling and shifting the x-axis.*

Chapter 6

Conclusion

6.1 Summary of Work

This thesis focuses on the correlation of tallies across active generations in Monte Carlo eigenvalue simulations. This thesis is divided along three main themes focusing on: 1) a quantitative understanding of the correlations and their impact on tally variance, 2) the development of predictive models to determine the impact of correlations and evaluation of accurate variance correction factors, and 3) the development of a new source update scheme based on the concept of delayed neutrons to reduce the inter-generation correlation in a predictable way.

Many studies have in the past looked at the impact of these correlations on reported variance and this work extends the analysis to the observed convergence rate of the tallies, the effect of tally size and the effect of generation size. The analysis was based on the numerically verified assumption that correlation coefficients are function of generation lag only and decreases exponentially. This hypothesis holds for the traditional source update scheme but not when source acceleration methods, like CMFD, are used. A fitting procedure of the auto-correlation coefficients (ACC's) was developed which enabled the prediction of the convergence rate. Thus, knowing the fit of the ACC's, one can determine the number of generations needed to achieve a desired level of accuracy. The fitting process and supporting simulations were also used to demonstrate the independence of the ACC's on the generation size.

Generation size will impact the level of accuracy but not the correlations, thus only additional generations can overcome the correlations and allow convergence of the error to return to $1/\sqrt{N}$. This work demonstrates that realistic cases may require 1000's of generations before the convergence rate reaches the asymptotic behavior, thus indicating that commonly held belief of adding more generations (unless using 1000's of generations) provides little added value when true variance is measured.

This new information on ACC's and how they are impacted by generation size and number of generations brought a very interesting question on what is the optimal run strategy to minimize the tally errors. An analysis was performed under the premise that a converged source distribution was provided using one of the many techniques currently under development [22] [31] [30] [27] [5] [8] [48]. This analysis demonstrates that in the presence of correlations, it is beneficial to opt for large generation sizes and few generations. There are obvious limitations to such a running scheme like the impact on the inactive generations and storage requirements of fission sites. Additionally, common practice often requires many generations in order to provide suitable confidence intervals which also have limited value unless the ACC's are known. This thesis described three techniques to evaluate suitable confidence intervals in the presence of correlations, one based on using history statistics, one using generation statistics and one batching generations to reduce batch-to-batch correlation [13]. The first two require some estimate of the variance underestimation factor evaluated from the ACC's, while batching can do without. The analysis indicates that history statistics are possible as long as one can remove history-to-history correlations often observed in the traditional source update methods and would allow the use of a single active generation if such a converged source bank could be generated. Batching will provide a good estimate of the true variance, but suffers from poor convergence on highly-correlated problems since the scheme favors the use of many generations which is contrary to the optimal run strategy identified in this work. Fitting correlation coefficients to exponentials is a working solution to correct the underestimated variance and predict the variance decrease rate. However, this method requires tallies for several hundreds of active generations to properly estimate

the correlation coefficients.

This thesis then proposes an approach for predicting the underestimation of variance often seen in full core nuclear reactor simulations prior to performing a detailed simulation. Knowing this underestimation will allow for more efficient simulations with truly quantifiable error estimates and stopping criteria. By discretizing the phase space, a Multitype Branching Process was developed to explicitly write the spatial and temporal moments which were then related to the correlation coefficients of tallies. It should be noted that the discretization mesh must be finer than the tally regions of interest, and that the closer to stationary the fission source used to compute the moments, the coarser the mesh can be. The MBP model does not treat the number normalization commonly used in the traditional source update method but can still accurately predict the correlation on normalized tallies.

Numerical results of the MBP model were compared to references calculated from independent simulations and demonstrated the accuracy of the predictive model on assembly size tallies in the 2D full core PWR BEAVRS benchmark. A quarter assembly mesh MBP model was constructed and moment tallies were computed over the first 8 active generations. The fission to absorption matrix and moments of ν_i (new neutrons produced from an absorption site at each phase space cells) were tallied to construct the model. The evolution of the MBP model using the moment responses M (source to source) and V (source to source-covariance) can be related to the ACC's for any coarser tally region at any generation lag. For efficiency, the predicted ACC's can be exponentially fitted using $2 \sim 3$ parameters which can then be used in calculating the variance underestimation. This predictive capability can help us calculate the number of active generations needed to achieve a target variance.

In many cases, however, the correlation effects are quite small (e.g. small system or small tally in large system) and it becomes unnecessary to develop an MBP model that can become somewhat costly. Ideally, a simple metric should be provided to determine whether significant variance underestimation is expected in a given problem with a given tally size. This task was accomplished by developing a continuous space model that can predict correlation coefficients on an homogeneous problem and thus

also provide an estimate of the variance under-prediction for any finite tally region in the system. This model is a generalization of the previous MBP model and an exact solution exists for homogeneous systems with reflective boundaries. Results of the correlation coefficients and variance underestimation on tally region of different sizes were shown to be very accurate for $1D$ and $3D$ homogeneous problems. Additionally, this simple model provided a simple predictive way of determining which tally size would suffer from large variance under-estimation on the heterogeneous BEAVRS benchmark.

In addition to the above work on correlation analysis and variance prediction, this thesis also explores modification to current Monte Carlo eigenvalue simulation procedure in order to reduce the generation-to-generation correlations. The generation-to-generation correlation coefficients can be reduced by delaying neutron source update between active generations, similar to the physical fission process where certain neutrons are delayed by many generations. The dependence of tallies at one generation on previous generations flattens the correlation coefficients as a function of the generation lag. The delayed neutron method thus leads to a different variance reduction profile comparison with the traditional source update method. The delayed neutron method, as proposed, shifts a greater computational cost to the inactive generations in order to build a prompt and delayed fission bank that is sampled during the active generations. This will lead to a greater storage requirement for the fission bank. The performance of the algorithm was demonstrated in the BEAVRS benchmark and shows great promise for cases where inactive generations are much faster than active ones, and for cases where the number of inactive generations can be reduced using an acceleration technique.

In order to find an optimal run strategy balancing the delayed scheme parameters (size of delayed bank, delayed sampling rate, ...), the variance behavior of the delayed scheme must be predicted. The MBP model was further generalized to incorporate the delayed neutron scheme and can be used to calculate the variance for different delayed schemes. Verification of the model was performed on simple benchmarks to demonstrate the predictive capabilities of the model. A demonstration on the

2D BEAVRS benchmark was then performed and showed that an optimal delayed scheme exists. For example, for a typical PWR run on assembly size tallies where inactive generations are twice as fast as active generations, and using an acceleration technique reducing inactive generation to 20, a factor of 10 in runtime could be saved by using an appropriate delay scheme to achieve 1% uncertainty in the tally. The improvement is reduced if inactive and active generations are comparable in runtime, and further reduced in many inactive generations are needed.

6.2 Contributions

- Developed a post-processing method to estimate generation-to-generation correlation coefficients
- Applied well accepted assumptions of correlation coefficients to identify new conclusions in correlation analysis.
  - Correlations not only make $\mathrm{Var}[\bar{X}(N)]$ higher than $\mathrm{Var}[X]/N$, but also make the traditional variance estimator of $\mathrm{Var}[X]$ biased.
  - Considering only the simulation cost in the active generations, fewer active generations with more neutrons per generation is preferred over many smaller generations.
- Quantitatively attributed generation-to-generation correlation of tallies to two parts: neutron source dependence between generations and spatial correlation of tallies.
- Developed the Multitype Branching Process model to predict correlation coefficients and variance correction for Monte Carlo eigenvalue simulations and demonstrated its accuracy on assembly size tallies in the BEAVRS benchmark.
- Developed a correlation diagnosis method to determine whether correlations will have a significant impact for a given tally size and demonstrated its accuracy on the BEAVRS benchmark.
- Developed a new Monte Carlo eigenvalue simulation procedure, the delayed neutron method, for improved variance behavior in the presence of correlations and demonstrated its impact on the BEAVRS benchmark.
- Generalized the Multitype Branching Process model to enable predictive capability of the variance for the delayed neutron method.
- Developed scheme to find optimal delayed neutron run strategy such that the most accurate results can be obtained with less computational expenses.

6.3 Future Work

6.3.1 Predicting correlation coefficients

Presently, the correlation prediction method based on the MBP model requires memory $\propto M^2$, where M is the fineness of the discretized mesh on which MBP is applied. Future work will investigate further approximations that can be made to reduce the memory requirements. The prediction method works by approximating the correlation coefficients of a real problem using the corresponding values of the discrete problem. To predict the variance underestimation ratio at active generation N, the moments of the constructed MBP model must be calculated to generation N. There is no simple closed form that can evaluate the correlations directly at generation N (matrix powers are not recognized as a closed form formula) except when fitting the predicted correlations for early active generations to exponentials. However, fitting to a single exponential term loses accuracy while fitting to a sum of exponentials is an ill-posed problem. It is worthwhile to investigate other approximations more specific to the MBP model that could be made that could accelerate the model without significant accuracy loss.

6.3.2 Efficiently searching optimal delayed neutron scheme

The current strategy to find an optimal delayed scheme is to first build DMBP models for all the searched delayed neutron schemes, then evolve the models to very large numbers of active generations, then compare the final variances after a fixed amount of work, and finally select the delayed scheme corresponding to the best variance among all the tested schemes. This limits the search space of the delayed scheme and may not provide the optimal run strategy. One solution could be to generalize the DMBP method into a continuous space and find an exact solution for the variances of a homogeneous reflective problem similarly to what was done in developing the simple diagnostic. It can be shown that the neutron age distribution developed to control the bias reduction rate is also related with the explicit form of matrix G (Eq 5.3.3).

Therefore, the prediction of variances can still be implemented numerically with the neutron age distribution assumed as constant coefficients, whose calculation cost is negligible in comparison with the moments evolution in the MBP models.

6.3.3 Extended applications of predictive models

The MBP model can be applied to analyze correlation behavior of many quantities that depend on the neutron count such as the analog k_{eff} estimator. The DMBP model can also be generalized to investigate other correlation reduction methods. For example, introducing a transport process of the secondary bank would correspond to the Wielandt shift method.

Appendix A1

Derivations

A1.1 Neutron in homogeneous fissile material

This section investigates independent neutrons in a homogeneous fissile system with reflective boundaries. The quantity of interest is the number of fission events in any tally region induced by s neutrons.

Denote the total cross section, absorption cross section, scattering cross section and fission cross section of the system as Σ_t, Σ_a, Σ_s and Σ_f respectively. Then at any collision, the probability that a neutron will induce scattering or fission is

$$p_s = \frac{\Sigma_s}{\Sigma_t} \tag{A1.1.1}$$

$$p_f = \frac{\Sigma_f}{\Sigma_t} \tag{A1.1.2}$$

The number of fission events induced by one neutron n_f is either 0 or 1. The probability that a neutron eventually induces fission comes from the series of events: { the neutron induces fission at first collision, the neutron is scattered after the first 1 collision and induces fission at second collision, the neutron is scattered after the first 2 collisions and induces fission at third collision, $\cdots$ the neutron is scattered after the

first n collisions and induces fission at $n+1^{st}$ collision, $\cdots\}$. Therefore

$$\begin{aligned}\mathbb{P}(n_f = 1) &= p_f + p_s p_f + p_s^2 p_f + \cdots + p_s^n p_f + \cdots \\ &= \sum_{n=0} p_s^n p_f \\ &= \frac{p_f}{1-p_s} = \frac{\Sigma_f}{\Sigma_t - \Sigma_s} \\ &\equiv P_f \end{aligned} \tag{A1.1.3}$$

The number of fission event induced by one neutron is a Bernoulli random variable with probability P_f [2].

$$n_f \sim Bernoulli(P_f) \tag{A1.1.4}$$

The number of fission events induced by all the s neutrons is

$$N_f = \sum_{a=1}^{s} n_f^{(a)} \tag{A1.1.5}$$

where superscript (a) indexes the s neutrons. Because the s neutrons behave independently, N_f is a binomial random variable with parameters s and P_f

$$N_f \sim B(s, P_f) \tag{A1.1.6}$$

whose moment generating function can be explicitly written as

$$\phi_{N_f}(r) = \mathbb{E}[e^{rN_f}] = [P_f e^r + (1-P_f)]^s \tag{A1.1.7}$$

Eq A1.1.6 fully characterizes the number of fission events induced by the s neutrons in the whole system. Then consider the contributions to a specific tally region. Use $\xi^{(\alpha)}$ to denote whether the α^{th} fission event out of the N_f fissions occurs in the tally region. Reflective boundaries of the homogeneous system make any quantity of interest translational invariant. Thus, any one of the N_f fission events occurs in the system uniformly. Suppose the tally region occupies fraction η of volume of the whole

system.

$$\xi^{\alpha} \sim Bernoulli(\eta) \tag{A1.1.8}$$

whose moment generating function can be explicitly written as

$$\phi_{\xi}(r) = \mathbb{E}[e^{r\xi}] = \eta e^{r} + (1-\eta) \tag{A1.1.9}$$

Finally, the real tally, the number of fission events occurred in tally region with volume fraction η is

$$Z = \sum_{\alpha=1}^{N_f} \xi^{(\alpha)} \tag{A1.1.10}$$

Now the expectation and variance of the tally Z can be derived.

Let $\phi_Z(t)$ be the moment generating function of Z with argument t

$$\phi_Z(t) = \mathbb{E}[e^{tZ}] \tag{A1.1.11}$$

Eq A1.1.11 can be expanded to incorporate the information of N_f and $\xi^{(\alpha)}$.

$$\begin{aligned}
\phi_Z(t) = \mathbb{E}\left[e^{t\sum\limits_{\alpha=1}^{N_f}\xi^{(\alpha)}}\right] &= \mathbb{E}\left[\prod_{\alpha=1}^{N_f} e^{t\xi^{(\alpha)}}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\prod_{\alpha=1}^{N_f} e^{t\xi^{(\alpha)}} \,\middle|\, N_f\right]\right] && \text{(A1.1.12)} \\
&= \mathbb{E}\left[\prod_{\alpha=1}^{N_f} \mathbb{E}\left[e^{t\xi^{(\alpha)}} \,\middle|\, N_f\right]\right] && \text{(A1.1.13)} \\
&= \mathbb{E}\left[\prod_{\alpha=1}^{N_f} \phi_{\xi}(t)\right] && \text{(A1.1.14)} \\
&= \mathbb{E}\left[(\phi_{\xi}(t))^{N_f}\right] \\
&= \phi_{N_f}\left(ln(\phi_{\xi}(t))\right) && \text{(A1.1.15)}
\end{aligned}$$

where Eq A1.1.12 evaluates the expectation with respect to both N_f and the $\xi^{(\alpha)}$'s by first evaluates the expectation conditioned on N_f then evaluate the conditional

expectation as function of N_f with respect to the distribution of N_f; Eq A1.1.13 exchanges the order of inner conditional expectation and product since given N_f, the N_f number of ξ's are independent of each other; Eq A1.1.14 recognizes the inner expectation as the definition of the moment generating function of $\xi^{(\alpha)}$ with t as the argument, since the $\xi^{(\alpha)}$'s are identical, the superscript has been omitted; Eq A1.1.15 recognizes the moment generating function of Z with argument t as the moment generating function of N_f with the argument $ln(\phi_\xi(t))$.

Derivatives of the moment generating functions defined above will yield ordinary rather than factorial moments.

Before evaluating the derivatives at $t = 0$ to obtain moments, the explicit form of the derivatives are derived first.

$$\frac{d\phi_Z(t)}{dt} = \frac{d\phi_{N_f}\left(ln(\phi_\xi(t))\right)}{dln(\phi_\xi(t))}\frac{dln(\phi_\xi(t))}{dt} \tag{A1.1.16}$$

$$\begin{aligned}\frac{d^2\phi_Z(t)}{dt^2} &= \frac{d}{dt}\left(\frac{d\phi_{N_f}\left(ln(\phi_\xi(t))\right)}{dln(\phi_\xi(t))}\frac{dln(\phi_\xi(t))}{dt}\right) \\ &= \left(\frac{d^2\phi_{N_f}\left(ln(\phi_\xi(t))\right)}{dln(\phi_\xi(t))^2}\frac{dln(\phi_\xi(t))}{dt}\right)\frac{dln(\phi_\xi(t))}{dt} \\ &+ \left(\frac{d\phi_{N_f}\left(ln(\phi_\xi(t))\right)}{dln(\phi_\xi(t))}\frac{d^2ln(\phi_\xi(t))}{dt^2}\right)\end{aligned} \tag{A1.1.17}$$

where

$$\frac{dln(\phi_\xi(t))}{dt} = \frac{1}{\phi_\xi(t)}\frac{d\phi_\xi(t)}{dt} \tag{A1.1.18}$$

$$\frac{d^2ln(\phi_\xi(t))}{dt^2} = -\frac{1}{\phi_\xi(t)^2}\left(\frac{d\phi_\xi(t)}{dt}\right)^2 + \frac{1}{\phi_\xi(t)}\frac{d^2\phi_\xi(t)}{dt^2} \tag{A1.1.19}$$

The following equations are needed to be evaluated at $t = 0$ in order to calculate

the first and second order derivatives of the moment generating function of Z.

$$\phi_\xi(t)\big|_{t=0} = 1 \tag{A1.1.20}$$

$$\left.\frac{d\phi_\xi(t)}{dt}\right|_{t=0} = \mathbb{E}\xi = \eta \tag{A1.1.21}$$

$$\left.\frac{d^2\phi_\xi(t)}{dt^2}\right|_{t=0} = \mathbb{E}\xi^2 = \eta \tag{A1.1.22}$$

$$ln(\phi_\xi(t))\big|_{t=0} = 0 \tag{A1.1.23}$$

$$\left.\frac{d\phi_{N_f}\left(ln(\phi_\xi(t))\right)}{dln(\phi_\xi(t))}\right|_{t=0} = \left.\frac{d\phi_{N_f}(l)}{dl}\right|_{l=0} = \mathbb{E}N_f = sP_f \tag{A1.1.24}$$

$$\left.\frac{d^2\phi_{N_f}\left(ln(\phi_\xi(t))\right)}{dln(\phi_\xi(t))^2}\right|_{t=0} = \left.\frac{d^2\phi_{N_f}(l)}{dl^2}\right|_{l=0} = \mathbb{E}N_f^2 = sP_f(1 - P_f + sP_f) \tag{A1.1.25}$$

where Eq A1.1.21 and Eq A1.1.22 are the results of the Bernoulli random variable xi (Eq A1.1.8). Eq A1.1.24 and Eq A1.1.25 are the results of the binomial random variable N_f (Eq A1.1.6).

The above results enable evaluating Eq A1.1.16 and Eq A1.1.17 at $t = 0$.

$$\mathbb{E}Z = sP_f\eta \tag{A1.1.26}$$

$$\begin{aligned}\mathrm{Var}[Z] &= \mathbb{E}Z^2 - (\mathbb{E}Z)^2 = sP_f(1 - P_f + sP_f)\eta\eta + sP_f(-\eta^2 + \eta) - (\mathbb{E}Z)^2 \\ &= sP_f\eta(1 - P_f\eta)\end{aligned} \tag{A1.1.27}$$

Eq A1.1.26 and Eq A1.1.27 implies that

$$Z \sim B(s, P_f\eta) \tag{A1.1.28}$$

In fact, substituting the moment generating function of $\xi^{(\alpha)}$ (Eq A1.1.9), the moment generating function of N_f (Eq A1.1.7) into the moment generating function of Z

(Eq A1.1.15) which incorporates both leads to the explict form of $\phi_Z(t)$ as

$$\begin{aligned}\phi_Z(t) &= [P_f\phi_\xi(t) + (1-P_f)]^s \\ &= \left[P_f\eta e^t + +(1-P_f\eta)\right]^s \end{aligned} \tag{A1.1.29}$$

Eq A1.1.29 is equivalent to Eq A1.1.28.

More specifically, if the whole system is discretized into M equal volume meshes, and the tally region occupies one of them, with $\eta = \frac{1}{M}$, the expectation and variance of Z are simplified to

$$\mathbb{E}Z = \frac{sP_f}{M} \tag{A1.1.30}$$

$$\mathrm{Var}[Z] = s\frac{P_f}{M}\left(1 - \frac{P_f}{M}\right) \tag{A1.1.31}$$

A1.2 Variance of Variance Estimators

Sample 1 $\{x_1, x_2, \cdots, x_M\}$ is independently sampled from an identical distribution. Sample 2 $\{X_1, X_2, \cdots, X_N\}$ is obtained by splitting $\{x_1, \cdots, x_M\}$ into N groups of size $s = M/N$ and calculating the average of each group.

$$X_i = \frac{1}{s}\sum_{j=1}^{s} x_{(i-1)s+j} \tag{A1.2.1}$$

Variance of these two samples σ_1^2 and σ_2^2 satisfies

$$\sigma_2^2 = \frac{\sigma_1^2}{s} \tag{A1.2.2}$$

σ_1^2 can be estimated with

$$S_0^2 = \frac{1}{M-1}\sum_{i=1}^{M}(x_i - \bar{x})^2 \tag{A1.2.3}$$

Alternatively, σ_1^2 can be estimated by estimating the variance of each of the N groups and averaging them

$$S_1^2 = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{s-1}\sum_{j=1}^{s}\left(x_{(i-1)s+j} - X_i\right)^2 \equiv \frac{1}{N}\sum_{i=1}^{N} S_{1,i}^2 \tag{A1.2.4}$$

And σ_2^2 can be estimated with

$$S_2^2 = \frac{1}{N-1}\sum_{j=1}^{N}\left(X_j - \bar{X}\right)^2 \tag{A1.2.5}$$

This section proves:

$$\mathrm{Var}[S_0^2] < \mathrm{Var}[s\ S_2^2] \tag{A1.2.6}$$

$$\mathrm{Var}[S_0^2] < \mathrm{Var}[s\ S_2^2] \tag{A1.2.7}$$

or equivalently

$$\mathrm{Var}[S_0^2] < s^2\,\mathrm{Var}[S_2^2] \tag{A1.2.8}$$

$$\mathrm{Var}[S_1^2] < s^2\,\mathrm{Var}[S_2^2] \tag{A1.2.9}$$

Variance of sample variance can be evaluated as [10]

$$\mathrm{Var}[S_0^2] = \frac{\mu_4(x)}{M} - \frac{\sigma(x)^4}{M}\frac{M-3}{M-1} \tag{A1.2.10}$$

$$\mathrm{Var}[S_{1,i}^2] = \frac{\mu_4(x)}{s} - \frac{\sigma(x)^4}{s}\frac{s-3}{s-1} \tag{A1.2.11}$$

$$\mathrm{Var}[S_2^2] = \frac{\mu_4(X)}{N} - \frac{\sigma(X)^4}{N}\frac{N-3}{N-1} \tag{A1.2.12}$$

where $\mu_4(x)$ is the 4^{th} central moment of x, $\sigma(x)^4$ is the square of variance of x. To compare Eq A1.2.10 (or Eq A1.2.11) with A1.2.12, relation between $\mu_4(x)$ and $\mu_4(X)$, $\sigma^4(x)$ and $\sigma^4(X)$ is needed. Since X is the average of s x,

$$\sigma^4(X) = (\sigma^2(X))^2 = \left(\frac{\sigma^2(x)}{s}\right)^2 = \frac{\sigma^4(x)}{s^2} \tag{A1.2.13}$$

Relation between $\mu_4(X)$ and $\mu_4(x)$ can be derived as below, where $\mu = E[x] = E[X] = E[\bar{x}] = E[\bar{X}]$,

$$\begin{aligned}\mu_4(X) = E[(X-\mu)^4] &= E\left[\left(\frac{\sum_{j=1}^{s} x_j}{s} - \mu\right)^4\right] \\ &= \frac{1}{s^4} E\left[\left(\sum_{j=1}^{s}(x_j - \mu)\right)^4\right]\end{aligned} \tag{A1.2.14}$$

For simplicity, define

$$y_i \equiv x_i - \mu \tag{A1.2.15}$$

therefore

$$E[y_i] = 0 \tag{A1.2.16}$$

$$\begin{aligned}s^4\mu_4(X) &= E\left[\left(\sum_{j=1}^{s} y_j\right)^4\right] \\ &= E\left[\sum_{j=1}^{s} y_j^4 + 3\sum_{j\neq i} y_j^2 y_i^2 + 4\sum_{j\neq i} y_j^3 y_i + 6\sum_{j\neq i\neq k} y_j^2 y_i y_k + \sum_{j\neq i\neq k\neq l} y_j y_i y_k y_l\right] \\ &= \sum_{j=1}^{s} E\left[y_j^4\right] + 3\sum_{j\neq i} E\left[y_j^2 y_i^2\right] + 4\sum_{j\neq i} E\left[y_j^3 y_i\right] + 6\sum_{j\neq i\neq k} E\left[y_j^2 y_i y_k\right] + \sum_{j\neq i\neq k\neq l} E\left[y_j y_i y_k y_l\right]\end{aligned} \tag{A1.2.17}$$

Since the x's are independent,

$$\begin{aligned}s^4\mu_4(X) = &\sum_{j=1}^{s} E[y_j^4] + 3\sum_{j\neq i} E[y_j^2]E[y_i^2] + 4\sum_{j\neq i} E[y_j^3]E[y_i] \\ &+ 6\sum_{j\neq i\neq k} E[y_j^2]E[y_i]E[y_k] + \sum_{j\neq i\neq k\neq l} E[y_j]E[y_i]E[y_k]E[y_l]\end{aligned} \tag{A1.2.18}$$

The above equation can be simplified due to Eq A1.2.16

$$\begin{aligned} s^4\mu_4(X) &= sE[y_j^4] + 3s(s-1)E[y_j^2]E[y_i^2] \\ &= sE[(x_i-\mu)^4] + 3s(s-1)E[(x_j-\mu)^2]^2 \\ &= s\mu_4(x) + 3s(s-1)\sigma^4(x) \end{aligned} \tag{A1.2.19}$$

Therefore $\mu_4(X)$ and $\mu_4(x)$ are related with

$$\mu_4(X) = \frac{1}{s^3}\mu_4(x) + \frac{3(s-1)}{s^3}\sigma^4(x) \tag{A1.2.20}$$

With Eq A1.2.13 and A1.2.20, inequality to be proved (Eq A1.2.8) can be equivalently transformed to

$$\frac{\mu_4(x)}{Ns} - \frac{\sigma^4(x)}{Ns}\frac{Ns-3}{Ns-1} < \frac{\mu_4(x)}{Ns} + \frac{3(s-1)}{Ns}\sigma^4(x) - \frac{\sigma^4(x)}{N}\frac{N-3}{N-1} \tag{A1.2.21}$$

which is equivalent to

$$2\frac{\sigma^4(x)}{Ns}\left(s-1+\frac{s}{N-1}-\frac{1}{Ns-1}\right) > 0 \tag{A1.2.22}$$

It can be verified that if $s=1$, the two estimators S_0^2 and sS_2^2 have the same variance. As long as $s>1$, the above equation holds and S_0^2 has lower variance than S_2^2.

Similarly, with Eq A1.2.13, Eq A1.2.20 and $\mathrm{Var}[S_1^2] = \frac{\mathrm{Var}[S_{1,i}^2]}{N}$, inequality to be proved in Eq A1.2.9 can be equivalently transformed to

$$\frac{1}{N}\left(\frac{\mu_4(x)}{s} - \frac{\sigma^4(x)}{s}\frac{s-3}{s-1}\right) < \frac{\mu_4(x)}{Ns} + \frac{3(s-1)}{Ns}\sigma^4(x) - \frac{\sigma^4(x)}{N}\frac{N-3}{N-1} \tag{A1.2.23}$$

$$\tag{A1.2.24}$$

which is equivalent to

$$2\frac{\sigma^4(x)}{Ns}\left(s-1+\frac{s}{N-1}-\frac{1}{s-1}\right) > 0 \tag{A1.2.25}$$

It can be verified that as long as $s > 1$, the above equation holds and S_1^2 has lower variance than S_2^2.

When comparing variance estimated from history statistics and generation statistics, the uncertainty of the former one corresponds to S_1^2 the latter one corresponds to S_2^2. When comparing variance estimated from generation statistics and batch statistics, the uncertainty of the former one corresponds to S_0^2 the latter one corresponds to S_2^2. Therefore, history-based variance estimator always has lower uncertainty than generation-based variance estimator. And generation-based variance estimator always has lower uncertainty than batch-based variance estimator.

A1.3 Multitype Branching Processes

For the moment generation function for random variable Z with argument s

$$f_Z(s) = \mathbb{E}s^Z \tag{A1.3.1}$$

It is clear that

$$f_Z(1) = 1 \tag{A1.3.2}$$

A1.3.1 temporal-spatial moments

This section derives the second and third order derivative of the joint moment generating function (Eq 3.3.44).

First evaluate the first order derivative with respect to t_b

$$\begin{aligned}
&\frac{\partial}{\partial t_b} F_{n,n+k}\left(\vec{Z}(0); \vec{s}, \vec{t}\right) = \frac{\partial}{\partial t_b} F_n\left(\vec{Z}(0); \vec{u}(k)\right) \\
&= \sum_j \left(\frac{\partial}{\partial u_j} F_n\left(\vec{Z}(0); \vec{u}(k)\right)\right) \frac{\partial u_j}{\partial t_b} \\
&= \sum_j \left(\frac{\partial}{\partial u_j} F_n\left(\vec{Z}(0); \vec{u}(k)\right)\right) s_j \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t})
\end{aligned} \tag{A1.3.3}$$

Then take the second derivative for each summand indexed by j. The derivative with respect to s_a of the j^{th} term in Eq A1.3.3 is

$$\begin{aligned}
&\frac{\partial}{\partial s_a}\left\{\left(\frac{\partial}{\partial u_j} F_n\left(\vec{Z}(0); \vec{u}(k)\right)\right) s_j \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t})\right\} \\
&= \left[\frac{\partial}{\partial s_a}\left(\frac{\partial}{\partial u_j} F_n\left(\vec{Z}(0); \vec{u}(k)\right)\right)\right] s_j \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t}) + \left(\frac{\partial}{\partial u_j} F_n\left(\vec{Z}(0); \vec{u}(k)\right)\right) \frac{\partial s_j}{\partial s_a} \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t})
\end{aligned} \tag{A1.3.4}$$

The $\frac{\partial^2}{\partial s_a \partial u_j}$ term in Eq A1.4.32 need to be expanded using chain rule again.

$$\begin{aligned}\frac{\partial}{\partial s_a}\left(\frac{\partial}{\partial u_j}F_n\left(\vec{Z}(0);\vec{u}(k)\right)\right) &= \sum_i \left(\frac{\partial^2}{\partial u_i \partial u_j}F_n\left(\vec{Z}(0);\vec{u}(k)\right)\right)\frac{\partial u_i}{\partial s_a} \\ &= \sum_i \left(\frac{\partial^2}{\partial u_i \partial u_j}F_n\left(\vec{Z}(0);\vec{u}(k)\right)\right)\delta_i^a f_i^{(k)}(\vec{t}) \\ &= \left(\frac{\partial^2}{\partial u_a \partial u_j}F_n\left(\vec{Z}(0);\vec{u}(k)\right)\right) f_a^{(k)}(\vec{t})\end{aligned} \tag{A1.3.5}$$

Substitute Eq A1.3.5 into Eq A1.3.4, and substitute Eq A1.3.4 into Eq A1.3.3 yields

$$\begin{aligned}&\frac{\partial^2}{\partial s_a \partial t_b}F_n\left(\vec{Z}(0);\vec{u}(k)\right) \\ &= \sum_j \left(\frac{\partial^2}{\partial u_a \partial u_j}F_n\left(\vec{Z}(0);\vec{u}(k)\right)\right) f_a^{(k)}(\vec{t}) s_j \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t}) + \left(\frac{\partial}{\partial u_j}F_n\left(\vec{Z}(0);\vec{u}(k)\right)\right)\delta_j^a \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t}) \\ &= \left(\frac{\partial^2}{\partial u_a \partial u_j}F_n\left(\vec{Z}(0);\vec{u}(k)\right)\right) f_a^{(k)}(\vec{t}) s_j \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t}) + \left(\frac{\partial}{\partial u_a}F_n\left(\vec{Z}(0);\vec{u}(k)\right)\right)\frac{\partial}{\partial t_b} f_a^{(k)}(\vec{t})\end{aligned} \tag{A1.3.6}$$

A1.3.2 Derivation of third order spatial moments

This section derives Eq 3.3.66. Third order derivatives of the moment generating function will provide the third order factorial moments, $\tilde{T}_{k,j,i}(n)$, which can eventually be related to $T_{i,j,i}(n)$ through Eq A1.3.7.

$$\begin{aligned}\tilde{T}_{k,j,i}(n) &= \left.\frac{\partial^3}{\partial s_k \partial s_j \partial s_i}F_n(\vec{r}_0,\vec{s})\right|_{\vec{s}=\vec{1}} \\ &= \mathbb{E}\left[Z_k(n)Z_j(n)Z_i(n) - Z_i(n)Z_k(n)\delta_{i,j} - Z_i(n)Z_j(n)\delta_{i,k} - Z_j(n)Z_i(n)\delta_{j,k} + 2Z_i(n)\delta_{i,j}\delta_{i,k}\right] \\ &= T_{k,j,i}(n) - C_{i,k}(n)\delta_{i,j} - C_{i,j}(n)\delta_{i,k} - C_{j,i}(n)\delta_{j,k} + 2\mu_i(n)\delta_{i,j}\delta_{i,k}\end{aligned} \tag{A1.3.7}$$

$\tilde{T}_{k,j,i}(n+1)$ is first evaluated conditional on $\vec{Z}(n)$ and then expressed as a function of $\tilde{T}_{k,j,i}(n)$. The conditional expectation is evaluated as,

$$
\begin{aligned}
&\tilde{T}_{k,j,i}(n+1)\Big|_{\vec{Z}(n)} = \frac{\partial^3}{\partial s_k \partial s_j \partial s_i} \prod_{k=1}^{m} F_1(k,\vec{s})^{Z_k(n)}\Bigg|_{\vec{s}=\vec{1}} = \frac{\partial^2}{\partial s_k \partial s_j}\left(\sum_{l=1}^{m} F_1(l,\vec{s})^{Z_l(n)-1} Z_l(n) \frac{\partial}{\partial s_i} F_1(l,\vec{s}) \prod_{k\neq l} F_1(k,\vec{s})^{Z_k(n)}\right)\Bigg|_{\vec{s}=\vec{1}} \\
&= \frac{\partial}{\partial s_k}\sum_{l=1}^{m}\frac{\partial}{\partial s_j}\left(F_1(l,\vec{s})^{Z_l(n)-1} Z_l(n)\frac{\partial}{\partial s_i}F_1(l,\vec{s})\right)\prod_{k\neq l}F_1(k,\vec{s})^{Z_k(n)} + F_1(l,\vec{s})^{Z_l(n)-1}Z_l(n)\frac{\partial}{\partial s_i}F_1(l,\vec{s})\frac{\partial}{\partial s_j}\left(\prod_{k\neq l}F_1(k,\vec{s})^{Z_k(n)}\right)\Bigg|_{\vec{s}=\vec{1}} \\
&= \sum_{l=1}^{m}\frac{\partial}{\partial s_k}\left(Z_l(n)(Z_l(n)-1)F_1(l,\vec{s})^{Z_l(n)-2}\frac{\partial}{\partial s_j}F_1(l,\vec{s})\frac{\partial}{\partial s_i}F_1(l,\vec{s})\right)\prod_{k\neq l}F_1(k,\vec{s})^{Z_k(n)} \\
&+ \left(F_1(l,\vec{s})^{Z_l(n)-1}Z_l(n)\frac{\partial^2}{\partial s_j\partial s_i}F_1(l,\vec{s})\right)\prod_{k\neq l}F_1(k,\vec{s})^{Z_k(n)} \\
&+ F_1(l,\vec{s})^{Z_l(n)-1}Z_l(n)\frac{\partial}{\partial s_i}F_1(l,\vec{s})\sum_{h\neq l}\left(Z_h(n)F_1(h,\vec{s})^{Z_h(n)-1}\frac{\partial}{\partial s_j}F_1(h,\vec{s})\prod_{\substack{g\neq h\\ g\neq l}}F_1(g,\vec{s})^{Z_g(n)}\right)\Bigg|_{\vec{s}=\vec{1}} \\
&= \sum_{l=1}^{m}\Bigg\{\Bigg(Z_l(n)(Z_l(n)-1)(Z_l(n)-2)F_1(l,\vec{s})^{Z_l(n)-3}\frac{\partial}{\partial s_k}F_1(l,\vec{s})\frac{\partial}{\partial s_j}F_1(l,\vec{s})\frac{\partial}{\partial s_i}F_1(l,\vec{s}) \\
&+ Z_l(n)(Z_l(n)-1)F_1(l,\vec{s})^{Z_l(n)-2}\frac{\partial^2}{\partial s_k\partial s_j}F_1(l,\vec{s})\frac{\partial}{\partial s_i}F_1(l,\vec{s}) \\
&+ Z_l(n)(Z_l(n)-1)F_1(l,\vec{s})^{Z_l(n)-2}\frac{\partial}{\partial s_j}F_1(l,\vec{s})\frac{\partial^2}{\partial s_k\partial s_i}F_1(l,\vec{s}) \\
&+ Z_l(n)(Z_l(n)-1)F_1(l,\vec{s})^{Z_l(n)-2}\frac{\partial}{\partial s_k}F_1(l,\vec{s})\frac{\partial^2}{\partial s_j\partial s_i}F_1(l,\vec{s}) \\
&+ F_1(l,\vec{s})^{Z_l(n)-1}Z_l(n)\frac{\partial^3}{\partial s_k\partial s_j\partial s_i}F_1(l,\vec{s})\Bigg)\prod_{k\neq l}F_1(k,\vec{s})^{Z_k(n)} \\
&+ \left(Z_l(n)(Z_l(n)-1)F_1(l,\vec{s})^{Z_l(n)-2}\frac{\partial}{\partial s_j}F_1(l,\vec{s})\frac{\partial}{\partial s_i}F_1(l,\vec{s}) + F_1(l,\vec{s})^{Z_l(n)-1}Z_l(n)\frac{\partial^2}{\partial s_j\partial s_i}F_1(l,\vec{s})\right)\times \\
&\sum_{\alpha\neq l}\left(Z_\alpha(n)F_1(\alpha,\vec{s})^{Z_\alpha(n)-1}\frac{\partial}{\partial s_k}F_1(\alpha,\vec{s})\prod_{\substack{\beta\neq\alpha\\ \beta\neq l}}F_1(\beta,\vec{s})^{Z_\beta(n)}\right) \\
&+ \left(Z_l(n)(Z_l(n)-1)F_1(l,\vec{s})^{Z_l(n)-2}\frac{\partial}{\partial s_k}F_1(l,\vec{s})\frac{\partial}{\partial s_i}F_1(l,\vec{s}) + F_1(l,\vec{s})^{Z_l(n)-1}Z_l(n)\frac{\partial^2}{\partial s_k\partial s_i}F_1(l,\vec{s})\right) \\
&\times\sum_{h\neq l}\left(Z_h(n)F_1(h,\vec{s})^{Z_h(n)-1}\frac{\partial}{\partial s_j}F_1(h,\vec{s})\prod_{\substack{g\neq h\\ g\neq l}}F_1(g,\vec{s})^{Z_h(n)}\right) + F_1(l,\vec{s})^{Z_l(n)-1}Z_l(n)\frac{\partial}{\partial s_i}F_1(l,\vec{s})\times \\
&\sum_{h\neq l}\Bigg[\left(Z_h(n)(Z_h(n)-1)F_1(h,\vec{s})^{Z_h(n)-2}\frac{\partial}{\partial s_k}F_1(h,\vec{s})\frac{\partial}{\partial s_j}F_1(h,\vec{s})\prod_{\substack{g\neq h\\ g\neq l}}F_1(g,\vec{s})^{Z_g(n)}\right) \\
&+ \left(Z_h(n)F_1(h,\vec{s})^{Z_h(n)-1}\frac{\partial^2}{\partial s_k\partial s_j}F_1(h,\vec{s})\prod_{\substack{g\neq h\\ g\neq l}}F_1(g,\vec{s})^{Z_g(n)}\right) \\
&+ \left(Z_h(n)F_1(h,\vec{s})^{Z_h(n)-1}\frac{\partial}{\partial s_j}F_1(h,\vec{s})\sum_{\substack{g\neq h\\ g\neq l}}Z_g(n)F_1(g,\vec{s})^{Z_g(n)-1}\frac{\partial}{\partial s_k}F_1(g,\vec{s})\prod_{\substack{\beta\neq g\\ \beta\neq h\\ \beta\neq l}}F_1(\beta,\vec{s})^{Z_\beta(n)}\right)\Bigg]\Bigg\}\Bigg|_{\vec{s}=\vec{1}}
\end{aligned}
\tag{A1.3.8}
$$

Evaluating Eq A1.3.8 at $\vec{s} = \vec{1}$, the moment generating functions become 1, the derivatives become factorial moments.

$$
\begin{aligned}
&\tilde{T}_{k,j,i}(n+1)\Big|_{\vec{Z}(n)} \\
&= \sum_{l=1}^{m} \left(Z_l(n)(Z_l(n)-1)(Z_l(n)-2)M_k^l M_j^l M_i^l + Z_l(n)(Z_l(n)-1)\left(\tilde{C}_{k,j}^l(1)M_i^l + M_j^l \tilde{C}_{k,i}^l(1) + M_k^l \tilde{C}_{j,i}^l(1)\right) + Z_l(n)\tilde{T}_{k,j,i}^l(1)\right) \\
&+ \sum_{l=1}^{m} \left(Z_l(n)(Z_l(n)-1)M_j^l M_i^l + Z_l(n)\tilde{C}_{j,i}^l(1)\right) \times \sum_{\alpha \neq l} Z_\alpha(n) M_k^\alpha \\
&+ \sum_{l=1}^{m} \left(Z_l(n)(Z_l(n)-1)M_k^l M_i^l + Z_l(n)\tilde{C}_{k,i}^l(1)\right) \times \sum_{h \neq l} Z_h(n) M_j^h \\
&+ \sum_{l=1}^{m} Z_l(n) M_i^l \times \sum_{h \neq l} \left[\left(Z_h(n)(Z_h(n)-1)M_k^h M_j^h \right) + Z_h(n)\tilde{C}_{k,j}^h(1) + Z_h(n) M_j^h \sum_{\substack{g \neq h \\ g \neq l}} Z_g(n) M_k^g \right] \\
&= \sum_{l=1}^{m} \left(Z_l(n)(Z_l(n)-1)(Z_l(n)-2)M_k^l M_j^l M_i^l + Z_l(n)(Z_l(n)-1)\left(\tilde{C}_{k,j}^l(1)M_i^l + M_j^l \tilde{C}_{k,i}^l(1) + M_k^l \tilde{C}_{j,i}^l(1)\right) + Z_l(n)\tilde{T}_{k,j,i}^l(1)\right) \\
&+ \sum_{l=1}^{m} \left(Z_l(n)(Z_l(n)-1)M_j^l M_i^l + Z_l(n)\tilde{C}_{j,i}^l(1)\right) \times \sum_{\alpha \neq l} Z_\alpha(n) M_k^\alpha \\
&+ \sum_{l=1}^{m} \left(Z_l(n)(Z_l(n)-1)M_k^l M_i^l + Z_l(n)\tilde{C}_{k,i}^l(1)\right) \times \sum_{h \neq l} Z_h(n) M_j^h \\
&+ \sum_{l=1}^{m} \left(Z_l(n)(Z_l(n)-1)M_k^l M_j^l + Z_l(n)\tilde{C}_{k,j}^l(1)\right) \times \sum_{h \neq l} Z_h(n) M_i^h \\
&+ \sum_{l=1}^{m} Z_l(n) M_i^l \times \sum_{h \neq l} \left[Z_h(n) M_j^h \sum_{\substack{g \neq h \\ g \neq l}} Z_g(n) M_k^g \right]
\end{aligned}
\tag{A1.3.9}
$$

Denote the 1^{st} term, the 2^{nd} to 4^{th} terms and the last term in Eq A1.3.9 as T_1, T_2 and T_3 respectively for simplification in the following derivations. T_3 can be simplified as Eq A1.3.10.

$$
T_3 = \sum_{\substack{l,h,g=1 \\ l \neq h, h \neq g, g \neq l}}^{m} Z_l(n) M_i^l Z_h(n) M_j^h Z_g(n) M_k^g \tag{A1.3.10}
$$

T_1 can be simplified as Eq A1.3.11.

$$
\begin{aligned}
&T_1 = \sum_{l=1}^{m} Z_l(n)\times \\
&\Big[(Z_l(n)-1)(Z_l(n)-2)M_k^l M_j^l M_i^l + (Z_l(n)-1)\left(\tilde{C}_{k,j}^l(1)M_i^l + M_j^l\tilde{C}_{k,i}^l(1) + M_k^l\tilde{C}_{j,i}^l(1)\right) + \tilde{T}_{k,j,i}^l(1)\Big] \\
&= \sum_{l=1}^{m} Z_l(n)\times \\
&\Big[(Z_l(n)^2 - 3Z_l(n))M_k^l M_j^l M_i^l + Z_l(n)\left(\tilde{C}_{k,j}^l(1)M_i^l + M_j^l\tilde{C}_{k,i}^l(1) + M_k^l\tilde{C}_{j,i}^l(1)\right) \\
&+2M_k^l M_j^l M_i^l - \tilde{C}_{k,j}^l(1)M_i^l - M_j^l\tilde{C}_{k,i}^l(1) - M_k^l\tilde{C}_{j,i}^l(1) + \tilde{T}_{k,j,i}^l(1)\Big]
\end{aligned}
\tag{A1.3.11}
$$

The third spatial (central) moment response to a point source of type l (defined in Eq 3.3.67) can be expanded as Eq A1.3.12

$$
\begin{aligned}
W_{k,j,i}^l &= \mathbb{E}(Z_k(1)Z_j(1)Z_i(1)|\vec{r}_0 = \vec{e}_l) \\
&\quad + 2\mathbb{E}(Z_k(1)|\vec{r}_0 = \vec{e}_l)\mathbb{E}(Z_j(1)|\vec{r}_0 = \vec{e}_l)\mathbb{E}(Z_i(1)|\vec{r}_0 = \vec{e}_l) \\
&\quad - \mathbb{E}(Z_i(1)Z_j(1)|\vec{r}_0 = \vec{e}_l)\mathbb{E}(Z_k(1)|\vec{r}_0 = \vec{e}_l) \\
&\quad - \mathbb{E}(Z_i(1)Z_k(1)|\vec{r}_0 = \vec{e}_l)\mathbb{E}(Z_j(1)|\vec{r}_0 = \vec{e}_l) \\
&\quad - \mathbb{E}(Z_k(1)Z_j(1)|\vec{r}_0 = \vec{e}_l)\mathbb{E}(Z_i(1)|\vec{r}_0 = \vec{e}_l) \\
&\quad = T_{k,j,i}^l(1) + 2M_k^l M_j^l M_i^l - C_{i,j}^l(1)M_k^l - C_{k,i}^l(1)M_j^l - C_{k,j}^l(1)M_i^l
\end{aligned}
\tag{A1.3.12}
$$

The third factorial moment response to a point source of type l can be defined by replacing the moments in Eq A1.3.12 with factorial moments.

$$
\tilde{W}_{k,j,i}^l = \tilde{T}_{k,j,i}^l(1) + 2M_k^l M_j^l M_i^l - \tilde{C}_{i,j}^l(1)M_k^l - \tilde{C}_{k,i}^l(1)M_j^l - \tilde{C}_{k,j}^l(1)M_i^l \tag{A1.3.13}
$$

$\tilde{W}^l_{k,j,i}$ defined in Eq A1.3.13 simplifies T_1 to

$$T_1 = \sum_{l=1}^{m} Z_l(n) \left\{ Z_l(n) \left[(Z_l(n)-3) M^l_k M^l_j M^l_i + \tilde{C}^l_{k,j}(1) M^l_i + M^l_j \tilde{C}^l_{k,i}(1) + M^l_k \tilde{C}^l_{j,i}(1) \right] + \tilde{W}^l_{k,j,i} \right\} \tag{A1.3.14}$$

For simplification , the three terms in T_2 are defined as $T_{2,1}$, $T_{2,2}$, $T_{2,3}$. $T_{2,1}$ is simplified in Eq A1.3.15.

$$\begin{aligned} T_{2,1} &= \sum_{l=1}^{m} Z_l(n) \left[Z_l(n) \left(\tilde{C}^l_{j,i}(1) - M^l_j M^l_i + Z_l(n) M^l_j M^l_i \right) \right] \times \sum_{\alpha \neq l} Z_\alpha(n) M^\alpha_k \\ &= \sum_{l=1}^{m} Z_l(n) \left(\tilde{V}^l_{j,i} + Z_l(n) M^l_j M^l_i \right) \times \sum_{\alpha \neq l} Z_\alpha(n) M^\alpha_k \\ &= \sum_{\substack{l=1,\alpha=1 \\ \alpha \neq l}}^{m} Z_l(n) \left(\tilde{V}^l_{j,i} + Z_l(n) M^l_j M^l_i \right) Z_\alpha(n) M^\alpha_k \\ &= \sum_{\substack{l=1,\alpha=1 \\ \alpha \neq l}}^{m} Z_l(n) Z_\alpha(n) \tilde{V}^l_{j,i} M^\alpha_k + Z_l(n)^2 Z_\alpha(n) M^l_j M^l_i M^\alpha_k \end{aligned} \tag{A1.3.15}$$

Simplifying $T_{2,2}$ and $T_{2,3}$ in the same way and combining them to T_2 lead to

$$\begin{aligned} T_2 = &\sum_{\substack{l=1,\alpha=1 \\ \alpha \neq l}}^{m} Z_l(n) Z_\alpha(n) \left(\tilde{V}^l_{j,i} M^\alpha_k + \tilde{V}^l_{k,i} M^\alpha_j + \tilde{V}^l_{j,k} M^\alpha_i \right) \\ &+ Z_l(n)^2 Z_\alpha(n) \left(M^l_j M^l_i M^\alpha_k + M^l_k M^l_i M^\alpha_j + M^l_j M^l_k M^\alpha_i \right) \end{aligned} \tag{A1.3.16}$$

where $\tilde{V}^l_{j,i}$ is defined in a similar way to $\tilde{W}^l_{k,j,i}$ by replacing moments in Eq 3.3.41 with factorial moments.

$$\tilde{V}^l_{j,i} = \tilde{C}^l_{j,i}(1) - M^l_j M^l_i \tag{A1.3.17}$$

Combining T_1(Eq A1.3.14), T_2(Eq A1.3.16) and T_3(Eq A1.3.10) simplifies $\tilde{T}_{k,j,i}(n+$

1)(Eq A1.3.9) to Eq A1.3.18.

$$
\begin{aligned}
&\tilde{T}_{k,j,i}(n+1)\Big|_{\vec{Z}(n)} \\
&= \sum_{l=1}^{m} Z_l(n)\tilde{W}^l_{k,j,i} + Z_l(n)^3 M^l_k M^l_j M^l_i - 3Z_l(n)^2 M^l_k M^l_j M^l_i + Z_l(n)^2\left(\tilde{C}^l_{k,j}(1)M^l_i + M^l_j\tilde{C}^l_{k,i}(1) + M^l_k\tilde{C}^l_{j,i}(1)\right) \\
&+ \sum_{\substack{l=1,\alpha=1\\ \alpha\neq l}}^{m} Z_l(n)Z_\alpha(n)\left(\tilde{V}^l_{j,i}M^\alpha_k + \tilde{V}^l_{k,i}M^\alpha_j + \tilde{V}^l_{j,k}M^\alpha_i\right) + Z_l(n)^2 Z_\alpha(n)\left(M^l_j M^l_i M^\alpha_k + M^l_k M^l_i M^\alpha_j + M^l_j M^l_k M^\alpha_i\right) \\
&+ \sum_{\substack{l,h,g=1\\ l\neq h, h\neq g, g\neq l}}^{m} Z_l(n)M^l_i Z_h(n)M^h_j Z_g(n)M^g_k
\end{aligned}
\tag{A1.3.18}
$$

Denote the sum of terms in the form of Z^3M^3 in Eq A1.3.18 as T_{ZM}.

$$
\begin{aligned}
T_{ZM} &= \sum_{l=1}^{m} Z_l(n)^3 M^l_k M^l_j M^l_i + \sum_{\substack{l,h,g=1\\ l\neq h, h\neq g, g\neq l}}^{m} Z_l(n)M^l_i Z_h(n)M^h_j Z_g(n)M^g_k \\
&+ \sum_{\substack{l=1,\alpha=1\\ \alpha\neq l}}^{m} Z_l(n)^2 Z_\alpha(n)\left(M^l_j M^l_i M^\alpha_k + M^l_k M^l_i M^\alpha_j + M^l_j M^l_k M^\alpha_i\right) \\
&= \sum_{\substack{l,h,g=1\\ l=h=g}}^{m} Z_l(n)Z_h(n)Z_g(n)M^g_k M^h_j M^l_i + \sum_{\substack{l,h,g=1\\ l\neq h, h\neq g, g\neq l}}^{m} Z_l(n)M^l_i Z_h(n)M^h_j Z_g(n)M^g_k \\
&+ \sum_{\substack{l,h,g=1\\ l=h, g\neq l}}^{m} Z_l(n)Z_h(n)Z_g(n)M^h_j M^l_i M^g_k + \sum_{\substack{l,h,g=1\\ l=g, g\neq h}}^{m} Z_l(n)Z_h(n)Z_g(n)M^g_k M^l_i M^h_j + \sum_{\substack{l,h,g=1\\ g=h, g\neq l}}^{m} Z_l(n)Z_h(n)Z_g(n)M^h_j M^g_k M^l_i \\
&= \sum_{l,h,g=1}^{m} Z_l(n)Z_h(n)Z_g(n)M^l_i M^h_j M^g_k
\end{aligned}
\tag{A1.3.19}
$$

Denote the sum of terms in the form of Z^2M^3 or $Z^2M\tilde{C}$ in Eq A1.3.18 as T_{ZC}.

$$
\begin{aligned}
T_{ZC} &= \sum_{l=1}^{m} Z_l(n)^2\left(\tilde{C}^l_{k,j}(1)M^l_i + M^l_j\tilde{C}^l_{k,i}(1) + M^l_k\tilde{C}^l_{j,i}(1)\right) - 3Z_l(n)^2 M^l_k M^l_j M^l_i \\
&= \sum_{l=1}^{m} Z_l(n)^2\left(\tilde{C}^l_{k,j}(1)M^l_i - M^l_k M^l_j M^l_i + M^l_j\tilde{C}^l_{k,i}(1) - M^l_k M^l_j M^l_i + M^l_k\tilde{C}^l_{j,i}(1) - M^l_k M^l_j M^l_i\right) \\
&= \sum_{l=1}^{m} Z_l(n)^2\left(M^l_i\tilde{V}^l_{k,j} + M^l_j\tilde{V}^l_{k,i} + M^l_k\tilde{V}^l_{i,j}\right)
\end{aligned}
\tag{A1.3.20}
$$

where the last line uses the definition of $\tilde{V}^l_{j,i}$ in Eq A1.3.17. Denote the sum of T_{ZC} (Eq A1.3.20) and terms in the form of $Z^2\tilde{V}M$ in Eq A1.3.18 as T_{ZV}.

$$\begin{aligned} T_{ZV} &= \sum_{l=1}^{m} Z_l(n)^2 \left(M_i^l \tilde{V}_{k,j}^l + M_j^l \tilde{V}_{k,i}^l + M_k^l \tilde{V}_{i,j}^l \right) + \sum_{\substack{l=1,\alpha=1 \\ \alpha \neq l}}^{m} Z_l(n) Z_\alpha(n) \left(\tilde{V}_{j,i}^l M_k^\alpha + \tilde{V}_{k,i}^l M_j^\alpha + \tilde{V}_{j,k}^l M_i^\alpha \right) \\ &= \sum_{l,\alpha=1}^{m} Z_l(n) Z_\alpha(n) \left(\tilde{V}_{j,i}^l M_k^\alpha + \tilde{V}_{k,i}^l M_j^\alpha + \tilde{V}_{j,k}^l M_i^\alpha \right) \end{aligned} \tag{A1.3.21}$$

T_{ZM}, T_{ZV} simplify $\tilde{T}_{k,j,i}(n+1)\Big|_{\vec{Z}(n)}$ to

$$\begin{aligned} \tilde{T}_{k,j,i}(n+1)\Big|_{\vec{Z}(n)} &= \sum_{l,h,g=1}^{m} Z_l(n) Z_h(n) Z_g(n) M_i^l M_j^h M_k^g \\ &\quad + \sum_{l,\alpha=1}^{m} Z_l(n) Z_\alpha(n) \left(\tilde{V}_{j,i}^l M_k^\alpha + \tilde{V}_{k,i}^l M_j^\alpha + \tilde{V}_{j,k}^l M_i^\alpha \right) + \sum_{l=1}^{m} Z_l(n) \tilde{W}_{k,j,i}^l \end{aligned} \tag{A1.3.22}$$

Then taking the expected value of $\tilde{T}_{k,j,i}(n+1)|_{\vec{Z}(n)}$ removes the dependence of $\vec{Z}(n)$ and yields

$$\begin{aligned} &\tilde{T}_{k,j,i}(n+1) = \\ &\qquad T_{l,h,g}(n) M_i^l M_j^h M_k^g + C_{l,\alpha}(n) \left(\tilde{V}_{j,i}^l M_k^\alpha + \tilde{V}_{k,i}^l M_j^\alpha + \tilde{V}_{j,k}^l M_i^\alpha \right) + \mu_l(n) \tilde{W}_{k,j,i}^l \end{aligned} \tag{A1.3.23}$$

Substituting the relation between $\tilde{T}_{k,j,i}$ and $T_{k,j,i}$ (Eq A1.3.7) simplifies Eq A1.3.23 to Eq 3.3.66.

A1.3.3 Simplification of covariance condensation

This section simplifies Eq 3.1.3 with Eq 3.3.65. The simplification is performed for the $\mathcal{O}\left(\mathbb{E}\epsilon^2\right)$ terms and the $\mathcal{O}\left(\mathbb{E}\epsilon^3\right)$ terms separately. That is

$$\begin{aligned}
&\mathrm{Cov}[Y_I(n), Y_I(n+k)] \\
&= \mathrm{Cov}^{(2)}[Y_I(n), Y_I(n+k)] + \mathrm{Cov}^{(3)}[Y_I(n), Y_I(n+k)] \\
&+ o\left(\mathbb{E}\epsilon^2\right)
\end{aligned} \tag{A1.3.24}$$

First sum out the dummy index i and j in $\mathrm{Cov}^{(2)}[Y_I(n), Y_I(n+k)]$ in all phase space regions,

$$\begin{aligned}
&\mathrm{Cov}^{(2)}[Y_I(n), Y_I(n+k)] \\
&= \sum_{I_i, I_j \in I} \left(\frac{\delta^i_{I_i}}{\mu(n)} - \frac{\mu_{I_i}(n)}{\mu(n)^2} \right) \left(\mathbb{E} Z_i(n) Z_j(n+k) - \mu_i(n)\mu_j(n+k) \right) \left(\frac{\delta^j_{I_j}}{\mu(n+k)} - \frac{\mu_{I_j}(n+k)}{\mu(n+k)^2} \right) \\
&= \sum_{I_i, I_j \in I} \frac{1}{\mu(n)} \left(\mathbb{E} Z_{I_i}(n) Z_{I_j}(n+k) - \mu_{I_i}(n)\mu_{I_j}(n+k) \right) \frac{1}{\mu(n+k)} \\
&\quad - \frac{1}{\mu(n)} \left(\mathbb{E} Z_{I_i}(n) Z(n+k) - \mu_{I_i}(n)\mu(n+k) \right) \frac{\mu_{I_j}(n+k)}{\mu(n+k)^2} \\
&\quad - \frac{\mu_{I_i}(n)}{\mu(n)^2} \left(\mathbb{E} Z(n) Z_{I_j}(n+k) - \mu(n)\mu_{I_j}(n+k) \right) \frac{1}{\mu(n+k)} \\
&\quad + \frac{\mu_{I_i}(n)}{\mu(n)^2} \left(\mathbb{E} Z(n) Z(n+k) - \mu(n)\mu(n+k) \right) \frac{\mu_{I_j}(n+k)}{\mu(n+k)^2}
\end{aligned} \tag{A1.3.25}$$

where any term indexed with i is changed to index I_i after being summed with $\delta^i_{I_i}$, any term indexed with j is changed to index I_j after being summed with $\delta^i_{I_j}$, and other summations count the total number of neutrons in the system from the count in all phase space regions using Eq 3.3.7 and Eq 3.3.56.

Then sum out the index I_i and I_j through the range $\{I_i | I_i \in I\}$.

$$
\begin{aligned}
&\mathrm{Cov}^{(2)}[Y_I(n), Y_I(n+k)] \\
&= \frac{1}{\mu(n)} \left(\mathbb{E} Z_I(n) Z_I(n+k) - \mu_I(n)\mu_I(n+k)\right) \frac{1}{\mu(n+k)} \\
&\quad - \frac{1}{\mu(n)} \left(\mathbb{E} Z_I(n) Z(n+k) - \mu_I(n)\mu(n+k)\right) \frac{\mu_I(n+k)}{\mu(n+k)^2} \\
&\quad - \frac{\mu_I(n)}{\mu(n)^2} \left(\mathbb{E} Z(n) Z_I(n+k) - \mu(n)\mu_I(n+k)\right) \frac{1}{\mu(n+k)} \\
&\quad + \frac{\mu_I(n)}{\mu(n)^2} \left(\mathbb{E} Z(n) Z(n+k) - \mu(n)\mu(n+k)\right) \frac{\mu_I(n+k)}{\mu(n+k)^2} \\
&= \frac{\mu_I(n)}{\mu(n)} \mathbb{E}\left[\left(\frac{Z(n)}{\mu(n)} - \frac{Z_I(n)}{\mu_I(n)}\right)\left(\frac{Z(n+k)}{\mu(n+k)} - \frac{Z_I(n+k)}{\mu_I(n+k)}\right)\right] \frac{\mu_I(n+k)}{\mu(n+k)}
\end{aligned}
\tag{A1.3.26}
$$

where the summation in I adds the number of neutrons in region phase space region contained within region I into the count of region I according to Eq 3.1.1 and its expectation.

$\mathrm{Cov}^{(3)}[Y_I(n), Y_I(n+k)]$ contains two terms which are symmetric by interchanging $l1 \leftrightarrow l2$ and $(n) \leftrightarrow (n+k)$. It is sufficient to derive the simplification for the first term. Similar to the $\mathcal{O}\left(\mathbb{E}\epsilon^2\right)$ terms above, the summation is evaluated on the dummy

index i and j over all phase space regions.

$$
\begin{aligned}
&\mathrm{Cov}_1^{(3)}[Y_I(n), Y_I(n+k)] \\
&= \sum_{I_i, I_j \in I} \frac{1}{2}\left(\frac{\delta^h_{I_i}}{\mu(n)} - \frac{\mu_{I_i}(n)}{\mu(n)^2}\right)\left(-\frac{\delta^i_{I_j}}{\mu(n+k)^2} - \frac{\delta^j_{I_j}}{\mu(n+k)^2} + 2\frac{\mu_{I_j}(n+k)}{\mu(n+k)^3}\right) \times \\
&\qquad \mathbb{E}(Z_h(n) - \mu_h(n))(Z_i(n+k) - \mu_i(n+k))(Z_j(n+k) - \mu_j(n+k)) \\
&= \frac{1}{2}\sum_{I_i, I_j \in I} \frac{1}{\mu(n)}\frac{-1}{\mu(n+k)^2}\mathbb{E}(Z_{I_i}(n) - \mu_{I_i}(n))(Z_{I_j}(n+k) - \mu_{I_j}(n+k))(Z(n+k) - \mu(n+k)) \\
&+ \frac{1}{\mu(n)}\frac{-1}{\mu(n+k)^2}\mathbb{E}(Z_{I_i}(n) - \mu_{I_i}(n))(Z(n+k) - \mu(n+k))(Z_{I_j}(n+k) - \mu_{I_j}(n+k)) \\
&+ \frac{1}{\mu(n)}2\frac{\mu_{I_j}(n+k)}{\mu(n+k)^3}\mathbb{E}(Z_{I_i}(n) - \mu_{I_i}(n))(Z(n+k) - \mu(n+k))(Z(n+k) - \mu(n+k)) \\
&+ \frac{\mu_{I_i}(n)}{\mu(n)^2}\frac{1}{\mu(n+k)^2}\mathbb{E}(Z(n) - \mu(n))(Z_{I_j}(n+k) - \mu_{I_j}(n+k))(Z(n+k) - \mu(n+k)) \\
&+ \frac{\mu_{I_i}(n)}{\mu(n)^2}\frac{1}{\mu(n+k)^2}\mathbb{E}(Z(n) - \mu(n))(Z(n+k) - \mu(n+k))(Z_{I_j}(n+k) - \mu_{I_j}(n+k)) \\
&- \frac{\mu_{I_i}(n)}{\mu(n)^2}2\frac{\mu_{I_j}(n+k)}{\mu(n+k)^3}\mathbb{E}(Z(n) - \mu(n))(Z(n+k) - \mu(n+k))(Z(n+k) - \mu(n+k))
\end{aligned}
\tag{A1.3.27}
$$

where any term indexed with h is changed to index I_i after being summed with $\delta^h_{I_i}$, any term indexed with i is changed to index I_j after being summed with $\delta^i_{I_j}$, any term indexed with j is changed to index I_j after being summed with $\delta^j_{I_j}$, and other summations count the total number of neutrons in the system from the count in all phase space regions using Eq 3.3.7 and Eq 3.3.56.

Identical terms are recognized in Eq A1.3.27 and

$$
\begin{aligned}
&\mathrm{Cov}_1^{(3)}[Y_I(n), Y_I(n+k)] \\
&= \sum_{I_i, I_j \in I} \frac{1}{\mu(n)} \frac{-1}{\mu(n+k)^2} \mathbb{E}(Z_{I_i}(n) - \mu_{I_i}(n))(Z_{I_j}(n+k) - \mu_{I_j}(n+k))(Z(n+k) - \mu(n+k)) \\
&+ \frac{1}{\mu(n)} \frac{\mu_{I_j}(n+k)}{\mu(n+k)^3} \mathbb{E}(Z_{I_i}(n) - \mu_{I_i}(n))(Z(n+k) - \mu(n+k))(Z(n+k) - \mu(n+k)) \\
&+ \frac{\mu_{I_i}(n)}{\mu(n)^2} \frac{1}{\mu(n+k)^2} \mathbb{E}(Z(n) - \mu(n))(Z_{I_j}(n+k) - \mu_{I_j}(n+k))(Z(n+k) - \mu(n+k)) \\
&- \frac{\mu_{I_i}(n)}{\mu(n)^2} \frac{\mu_{I_j}(n+k)}{\mu(n+k)^3} \mathbb{E}(Z(n) - \mu(n))(Z(n+k) - \mu(n+k))(Z(n+k) - \mu(n+k))
\end{aligned}
\tag{A1.3.28}
$$

Then sum out the index I_i and I_j through the range $\{I_i | I_i \in I\}$.

$$
\begin{aligned}
&\mathrm{Cov}_1^{(3)}[Y_I(n), Y_I(n+k)] \\
&= \frac{1}{\mu(n)} \frac{-1}{\mu(n+k)^2} \mathbb{E}(Z_I(n) - \mu_I(n))(Z_I(n+k) - \mu_I(n+k))(Z(n+k) - \mu(n+k)) \\
&+ \frac{1}{\mu(n)} \frac{\mu_I(n+k)}{\mu(n+k)^3} \mathbb{E}(Z_I(n) - \mu_I(n))(Z(n+k) - \mu(n+k))(Z(n+k) - \mu(n+k)) \\
&+ \frac{\mu_I(n)}{\mu(n)^2} \frac{1}{\mu(n+k)^2} \mathbb{E}(Z(n) - \mu(n))(Z_I(n+k) - \mu_I(n+k))(Z(n+k) - \mu(n+k)) \\
&- \frac{\mu_I(n)}{\mu(n)^2} \frac{\mu_I(n+k)}{\mu(n+k)^3} \mathbb{E}(Z(n) - \mu(n))(Z(n+k) - \mu(n+k))(Z(n+k) - \mu(n+k)) \\
&= -\mathbb{E}\left[\left(\frac{Z(n)}{\mu(n)} \frac{\mu_I(n)}{\mu(n)} - \frac{Z_I(n)}{\mu(n)} \right) \left(\frac{Z(n+k)}{\mu(n+k)} - 1 \right) \left(\frac{Z(n+k)}{\mu(n+k)} \frac{\mu_I(n+k)}{\mu(n+k)} - \frac{Z_I(n+k)}{\mu(n+k)} \right) \right]
\end{aligned}
\tag{A1.3.29}
$$

The second term in $\mathrm{Cov}^{(3)}[Y_I(n), Y_I(n+k)]$ can be achieved by interchanging $(n) \leftrightarrow (n+k)$.

$$
\begin{aligned}
&\mathrm{Cov}_2^{(3)}[Y_I(n), Y_I(n+k)] \\
&= -\mathbb{E}\left[\left(\frac{Z(n+k)}{\mu(n+k)} \frac{\mu_I(n+k)}{\mu(n+k)} - \frac{Z_I(n+k)}{\mu(n+k)} \right) \left(\frac{Z(n)}{\mu(n)} - 1 \right) \left(\frac{Z(n)}{\mu(n)} \frac{\mu_I(n)}{\mu(n)} - \frac{Z_I(n)}{\mu(n)} \right) \right]
\end{aligned}
\tag{A1.3.30}
$$

Eq A1.3.29 and Eq A1.3.30 leads to

$$\begin{aligned}&\mathrm{Cov}^{(3)}[Y_I(n), Y_I(n+k)]\\&= -\frac{\mu_I(n)}{\mu(n)}\mathbb{E}\left[\left(\frac{Z(n+k)}{\mu(n+k)} - \frac{Z_I(n+k)}{\mu_I(n+k)}\right)\left(\frac{Z(n)}{\mu(n)} - 1 + \frac{Z(n+k)}{\mu(n+k)} - 1\right)\left(\frac{Z(n)}{\mu(n)} - \frac{Z_I(n)}{\mu_I(n)}\right)\right]\frac{\mu_I(n+k)}{\mu(n+k)}\end{aligned} \tag{A1.3.31}$$

A1.3.4 Branching processes within generations

To treat the (n, xn) processes accurately, instead of one fission-to-absorption probability matrix, four matrices are needed

1. P_{fi}^{l}: the probability of a neutron born from fission at region l to being absorbed at region i

2. Q_{fi}^{l}: the probability of a neutron born from fission at region l to inducing a (n, xn) reaction at region i

3. P_{si}^{l}: the probability of a neutron born from (n, xn) at region l to being absorbed at region i

4. Q_{si}^{l}: the probability of a neutron born from (n, xn) at region l to inducing a (n, xn) reaction at region i

Similarly, rather than one random variable ν, moments of two random variables are needed

1. ν_i: number of new neutrons out of fission events from absorption at phase space region i, for simplicity, the moments of ν_i are denoted as below

$$\begin{aligned}\mathbb{E}\nu_i &= \lambda_i\\ \mathbb{E}\nu_i^2 &= \sigma_i^2\end{aligned} \tag{A1.3.32}$$

2. ξ_i: number of new neutrons out of (n, xn) at phase space region i , $\xi_i \geq 2$, for

simplicity, the moments of ξ_i are denoted as below

$$\begin{aligned} \mathbb{E}\xi_i &= \chi_i \\ \mathbb{E}\xi_i(\xi_i - 1) &= \tau_i^2 \end{aligned} \tag{A1.3.33}$$

Similarly to the case without (n, xn) processes discussed above, all the transport processes P_f,Q_f,P_s,Q_s and the branching processes ν, ξ are assumed to be independent of each other. In other words, ν_i and ξ_i is assumed to be function of phase space region i only (not dependent on whether the incoming neutron is born from fission or (n, xn)). Therefore, the expectation of the composite processes can be expressed as the product of the expectation of each process.

A1.3.4.1 1^{st} order moment response M_i^l

The lowest order component of M_i^l corresponds to the process where a neutron is born at cell l and finally induces fission events at cell i without any (n, xn) processes in between. The expected number of new neutrons would be

$$M_i^{l(0)} = {P_f}_i^l \lambda_i \tag{A1.3.34}$$

The next order component would come from the process where a neutron is born at cell l, then encounter (n, xn) reaction at cell j, and finally induces fission events at cell i. Since it is assumed branching and transport are independent and the neutrons out of (n, xn) behave independently, the expected number of new neutrons of this process would be

$$M_i^{l(1)} = {Q_f}_j^l \chi_j {P_s}_i^j \lambda_i \tag{A1.3.35}$$

The above process continues with infinitely higher order terms and can be conveniently written in matrix forms. Denote the moments of new neutrons from absorption and

(n, xn) in the diagonal of matrices Λ and X respectively.

$$\Lambda_{i,j} = \lambda_i \delta_{i,j} \tag{A1.3.36}$$

$$X_{i,j} = \chi_i \delta_{i,j} \tag{A1.3.37}$$

$$\Lambda'_{i,j} = \sigma_i^2 \delta_{i,j} \tag{A1.3.38}$$

Then, the moment response M can be expressed as the series

$$\begin{aligned} M &= M^{(0)} + M^{(1)} + M^{(2)} + \cdots \\ &= P_f\Lambda + Q_f X P_s \Lambda + Q_f X Q_s X P_s \Lambda + \cdots \\ &= P_f\Lambda + Q_f X (P_s\Lambda + Q_s X (P_s\Lambda + Q_s X (\cdots))) \\ &= P_f\Lambda + Q_f X (1 + Q_s X + (Q_s X)^2 + \cdots) P_s \Lambda \end{aligned} \tag{A1.3.39}$$

The series of $Q_s X$ must converge and thus sums to $(1 - Q_s X)^{-1}$. Therefore the infinite contributions of M sum to

$$M = P_f\Lambda + Q_f X (1 - Q_s X)^{-1} P_s \Lambda \tag{A1.3.40}$$

A1.3.4.2 2^{nd} order moment response $V^l_{i,j}$

By Eq 3.3.72, the undetermined part of $V^l_{i,j}$ so far is $\mathbb{E}Z^l_i(1)Z^l_j(1)$. The lowest order contribution of $V^l_{i,j}$ come from the neutrons that were born at phase space cell l and induced new neutrons into i, j without (n, xn) processes in between.

$$V^l_{i,j}{}^{(0)} = P_{fi}{}^l \sigma_i^2 \delta_{i,j} \tag{A1.3.41}$$

Higher order terms result from (n, xn) neutrons. Define the number of new neutrons at phase space cell i induced by the β^{th} neutron among the ξ_k neutrons out of the (n, xn) reaction at phase space cell k as $\rho^k_{\beta,i}$. Then contribution from (n, xn) neutrons

in $V_{i,j}^{l}$ can be written as

$$V_{i,j}^{l\ (1)} = Q_f {}_k^l \mathbb{E} \sum_{\beta=1}^{\xi_k} \rho_{\beta,i}^k \sum_{\beta=1}^{\xi_k} \rho_{\beta,j}^k \tag{A1.3.42}$$

The expectation part can be solved following the processes below

$$\begin{aligned} D_{i,j}^k &\equiv \mathbb{E}\left(\sum_{\beta=1}^{\xi_k} \rho_{\beta,i}^k \rho_{\beta,j}^k + \sum_{\beta\neq\gamma}^{\xi_k} \rho_{\beta,i}^k \rho_{\gamma,j}^k\right) \\ &= \mathbb{E}\left(\mathbb{E}(\xi_k \rho_{\beta_i}^k \rho_{\beta_j}^k | \xi_k) + \mathbb{E}(\xi_k(\xi_k-1)\rho_{\beta_i}^k \rho_{\gamma_j}^k | \xi_k)\right) \\ &= \left(\mathbb{E}\xi_k \mathbb{E}\rho_{\beta_i}^k \rho_{\beta_j}^k + \mathbb{E}(\xi_k(\xi_k-1))\mathbb{E}\rho_{\beta_i}^k \rho_{\gamma_j}^k\right) \\ &= \left(\chi^k \mathbb{E}\rho_{\beta_i}^k \rho_{\beta_j}^k + \tau^{2k} \mathbb{E}\rho_{\beta_i}^k \mathbb{E}\rho_{\gamma_j}^k\right) \end{aligned} \tag{A1.3.43}$$

where the ξ_k neutrons out of a (n, xn) reaction are assumed to behave independently. For simplicity of below equations, expectations in Eq A1.3.43 are denoted as

$$R_i^k \equiv \mathbb{E}\rho_{\beta_i}^k \tag{A1.3.44}$$

$$F_{i,j}^k \equiv \mathbb{E}\rho_{\beta_i}^k \rho_{\beta_j}^k \tag{A1.3.45}$$

$$G_{i,j}^k \equiv \chi^k F_{i,j}^k \tag{A1.3.46}$$

$$S_{i,j}^k \equiv \tau^{2k} R_i^k R_j^k \tag{A1.3.47}$$

$$D_{i,j}^k = G_{i,j}^k + S_{i,j}^k \tag{A1.3.48}$$

The first order moment response of (n, xn) reaction denoted as R_i^k (Eq A1.3.44) can be solved the same way as the above analysis for M_i^l and turned out to be in a very similar form.

$$R = (1 - Q_s X)^{-1} P_s \Lambda \tag{A1.3.49}$$

The second order moment response of (n, xn) reaction denoted as $F_{i,j}^k$ (Eq A1.3.45) can be solved recursively. Similarly to $V_{i,j}^l$, decompose $F_{i,j}^k$ into zeroth order contributions (neutrons out of (n, xn) were absorbed before any (n, xn) reactions) and high

order terms (neutrons out of (n, xn) were absorbed after some (n, xn) reactions).

$$\begin{aligned} F^k_{i,j} &= P_{s}{}^{k}_{i,j}\sigma_i^2\delta_{i,j} + Q_{s}{}^{k}_{k'}\mathbb{E}\sum_{\beta=1}^{\xi_{k'}}\rho^{k'}_{\beta,i}\sum_{\beta=1}^{\xi_{k'}}\rho^{k'}_{\beta,j} \\ &= P_{s}{}^{k}_{i,j}\sigma_i^2\delta_{i,j} + Q_{s}{}^{k}_{k'}D^{k'}_{i,j} \end{aligned} \tag{A1.3.50}$$

Plugging the solution of $D^k_{i,j}$ in the form of $F^k_{i,j}$ (Eq A1.3.43) into the above equation gives another equation where $F^k_{i,j}$ is the only unknown term.

$$F^k_{i,j} = P_{s}{}^{k}_{i}\sigma_i^2\delta_{i,j} + Q_{s}{}^{k}_{k'}(\chi^{k'}F^{k'}_{i,j} + S^{k'}_{i,j}) \tag{A1.3.51}$$

which can be easily converted to an equation of $D^k_{i,j}$

$$D^k_{i,j} = P_{s}{}^{k}_{i}\chi^k\sigma_i^2\delta_{i,j} + S^k_{i,j} + Q_{s}{}^{k}_{k'}\chi^k D^{k'}_{i,j} \tag{A1.3.52}$$

and $D^k_{i,j}$ is solved as

$$D^k_{i,j} = \left((1 - Q_sX)^{-1}\right)^k_{k'}(P_{s}{}^{k'}_{i}\chi^{k'}\sigma_i^2\delta_{i,j} + S^{k'}_{i,j}) \tag{A1.3.53}$$

In summary, the 2^{nd} order moment response $V^l_{i,j}$ is found as

$$\begin{aligned} V^l_{i,j} &= V^{l\ (0)}_{i,j} + V^{l\ (1)}_{i,j} - M^l_iM^l_j \\ V &= P_f \circ \Lambda' + Q_f(1 - Q_sX)^{-1} * (P_sX \circ \Lambda' + S) - M \cdot M \end{aligned} \tag{A1.3.54}$$

where four types of tensor product are implied

$$(AB)^i_j \equiv A^i_kB^k_j \tag{A1.3.55}$$

$$(A * B)^k_{i,j} \equiv A^k_{k'}B^{k'}_{i,j} \tag{A1.3.56}$$

$$(A \circ B)^k_{i,j} \equiv A^k_iB_{i,j} \tag{A1.3.57}$$

$$(A \cdot B)^k_{i,j} \equiv A^k_iB^k_j \tag{A1.3.58}$$

A1.4 Delayed Multitype Branching Processes

A1.4.1 responses

This section summarizes the first and second order derivatives of different parts of the moment generating function and their evaluation at $\vec{s} = \vec{\sigma} = \vec{1}$.

A1.4.1.1 First order derivatives

1. Expand the derivative $\frac{\partial}{\partial s_a} \prod_i f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)}$:

$$\begin{aligned}\frac{\partial}{\partial s_a} \prod_i f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} &= \sum_l \left(\frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma})^{Z_l(n-1)} \right) \prod_{i \neq l} f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} \\ &= \sum_l Z_l(n-1) f_l(\vec{s}, \vec{\sigma})^{Z_l(n-1)-1} \left(\frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma}) \right) \prod_{i \neq l} f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)}\end{aligned} \tag{A1.4.1}$$

Evaluate the above equation at $\vec{s} = \vec{\sigma} = \vec{1}$:

$$\begin{aligned}\frac{\partial}{\partial s_a} \prod_i f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} &= \sum_l Z_l(n-1) \frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma}) \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\ &= \sum_l Z_l(n-1) M_a^l\end{aligned} \tag{A1.4.2}$$

where the first equation uses Eq A1.3.2, the second equation uses the definition in Eq 5.2.8.

2. Expand the derivative $\frac{\partial}{\partial s_a} \prod_l g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)}$:

$$\begin{aligned}\frac{\partial}{\partial s_a} \prod_l g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} &= \sum_k \left(\frac{\partial}{\partial s_a} g_k(s_k, \sigma_k)^{U_k(n-1)} \right) \prod_{l \neq k} g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} \\ &= \sum_k U_k(n-1) g_k(s_k, \sigma_k)^{U_k(n-1)-1} \left(\frac{\partial}{\partial s_a} g_k(s_k, \sigma_k) \right) \prod_{l \neq k} g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)}\end{aligned} \tag{A1.4.3}$$

Evaluate the above equation at $\vec{s} = \vec{\sigma} = \vec{1}$:

$$\begin{aligned}\frac{\partial}{\partial s_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} &= \sum_k U_k(n-1)\frac{\partial}{\partial s_a}g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\ &= \sum_k U_k(n-1)m_a^k\end{aligned} \tag{A1.4.4}$$

where the first equation uses Eq A1.3.2, the second equation uses the definition in Eq 5.2.9.

3. Expand the derivative $\frac{\partial}{\partial\sigma_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}$:

$$\begin{aligned}\frac{\partial}{\partial\sigma_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} &= \sum_l\left(\frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)}\right)\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\ &= \sum_l Z_l(n-1)f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\left(\frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\right)\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\end{aligned} \tag{A1.4.5}$$

Evaluate the above equation at $\vec{s} = \vec{\sigma} = \vec{1}$:

$$\begin{aligned}\frac{\partial}{\partial\sigma_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} &= \sum_l Z_l(n-1)\frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\ &= \sum_l Z_l(n-1)\Psi_a^l\end{aligned} \tag{A1.4.6}$$

where the first equation uses Eq A1.3.2, the second equation uses the definition in Eq 5.2.11.

4. Expand the derivative $\frac{\partial}{\partial\sigma_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}$:

$$\begin{aligned}\frac{\partial}{\partial\sigma_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} &= \sum_k\left(\frac{\partial}{\partial\sigma_a}g_k(s_k,\sigma_k)^{U_k(n-1)}\right)\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\ &= \sum_k U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\left(\frac{\partial}{\partial\sigma_a}g_k(s_k,\sigma_k)\right)\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\end{aligned} \tag{A1.4.7}$$

Evaluate the above equation at $\vec{s} = \vec{\sigma} = \vec{1}$:

$$\begin{aligned}\frac{\partial}{\partial \sigma_a} \prod_l g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} \Big|_{\vec{s}=\vec{\sigma}=\vec{1}} &= \sum_k U_k(n-1) \frac{\partial}{\partial \sigma_a} g_k(s_k, \sigma_k) \Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\ &= \sum_k U_k(n-1) \psi_a^k\end{aligned} \tag{A1.4.8}$$

where the first equation uses Eq A1.3.2, the second equation uses the definition in Eq 5.2.12.

A1.4.1.2 Second Order Derivatives

1. Expand the derivative $\frac{\partial^2}{\partial s_b \partial s_a} \prod_i f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)}$:

$$\begin{aligned}&\frac{\partial^2}{\partial s_b \partial s_a} \prod_i f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} \\ &= \frac{\partial}{\partial s_b} \sum_l Z_l(n-1) f_l(\vec{s}, \vec{\sigma})^{Z_l(n-1)-1} \frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma}) \prod_{i \neq l} f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} \\ &= \sum_l Z_l(n-1) \frac{\partial}{\partial s_b} f_l(\vec{s}, \vec{\sigma})^{Z_l(n-1)-1} \times \frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma}) \times \prod_{i \neq l} f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} \\ &+ Z_l(n-1) f_l(\vec{s}, \vec{\sigma})^{Z_l(n-1)-1} \times \frac{\partial^2}{\partial s_b \partial s_a} f_l(\vec{s}, \vec{\sigma}) \times \prod_{i \neq l} f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} \\ &+ Z_l(n-1) f_l(\vec{s}, \vec{\sigma})^{Z_l(n-1)-1} \frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma}) \times \frac{\partial}{\partial s_b} \prod_{i \neq l} f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} \\ &= \sum_l Z_l(n-1)(Z_l(n-1)-1) f_l(\vec{s}, \vec{\sigma})^{Z_l(n-1)-2} \frac{\partial}{\partial s_b} f_l(\vec{s}, \vec{\sigma}) \times \frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma}) \times \prod_{i \neq l} f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} \\ &+ Z_l(n-1) f_l(\vec{s}, \vec{\sigma})^{Z_l(n-1)-1} \times \frac{\partial^2}{\partial s_b \partial s_a} f_l(\vec{s}, \vec{\sigma}) \times \prod_{i \neq l} f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)} \\ &+ Z_l(n-1) f_l(\vec{s}, \vec{\sigma})^{Z_l(n-1)-1} \times \frac{\partial}{\partial s_a} f_l(\vec{s}, \vec{\sigma}) \times \sum_{j \neq l} Z_j(n-1) f_i(\vec{s}, \vec{\sigma})^{Z_j(n-1)-1} \frac{\partial}{\partial s_b} f_j(\vec{s}, \vec{\sigma}) \prod_{\substack{i \neq l \\ i \neq j}} f_i(\vec{s}, \vec{\sigma})^{Z_i(n-1)}\end{aligned} \tag{A1.4.9}$$

where the first equation uses the first derivative expanded in Eq A1.4.1.

Evaluation of the above derivative at $\vec{s} = \vec{\sigma} = \vec{1}$ will be useful for the conditional

expectations later:

$$
\begin{aligned}
&\frac{\partial^2}{\partial s_b \partial s_a} \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_l Z_l(n-1)(Z_l(n-1)-1)\, \frac{\partial}{\partial s_b} f_l(\vec{s},\vec{\sigma}) \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \times \frac{\partial}{\partial s_a} f_l(\vec{s},\vec{\sigma}) \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ Z_l(n-1) \times \frac{\partial^2}{\partial s_b \partial s_a} f_l(\vec{s},\vec{\sigma}) \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ Z_l(n-1) \times \frac{\partial}{\partial s_a} f_l(\vec{s},\vec{\sigma}) \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \times \sum_{j\neq l} Z_j(n-1)\, \frac{\partial}{\partial s_b} f_j(\vec{s},\vec{\sigma}) \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_l \left\{ Z_l(n-1)(Z_l(n-1)-1) M_b^l M_a^l + Z_l(n-1) \bar{\mathbb{V}}_{a,b}^l + Z_l(n-1) M_a^l \sum_{j\neq l} Z_j(n-1) M_b^j \right\} \\
&= \sum_l Z_l(n-1)(\bar{\mathbb{V}}_{a,b}^l - M_b^l M_a^l) + \sum_{l,j} Z_l(n-1) M_a^l Z_j(n-1) M_b^j
\end{aligned}
\tag{A1.4.10}
$$

where the first equation uses Eq A1.3.2, the second equation uses definition of M in Eq 5.2.8 and definition of $\bar{\mathbb{V}}$ in Eq 5.2.19, the third equation combines the summand indexed with l and summation over $\{j\}_{j\neq l}$ into a non-constrained summation over j.

2. Expand the derivative $\frac{\partial^2}{\partial s_b \partial s_a} \prod_i g_i(s_i, \sigma_i)^{U_i(n-1)}$:

$$
\begin{aligned}
&\frac{\partial^2}{\partial s_b \partial s_a} \prod_l g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} \\
&= \frac{\partial}{\partial s_b} \sum_k U_k(n-1) g_k(s_k, \sigma_k)^{U_k(n-1)-1} \frac{\partial}{\partial s_a} g_k(s_k, \sigma_k) \prod_{l \neq k} g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} \\
&= \sum_k U_k(n-1) \frac{\partial}{\partial s_b} g_k(s_k, \sigma_k)^{U_k(n-1)-1} \times \frac{\partial}{\partial s_a} g_k(s_k, \sigma_k) \times \prod_{l \neq k} g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} \\
&+ U_k(n-1) g_k(s_k, \sigma_k)^{U_k(n-1)-1} \times \frac{\partial^2}{\partial s_a s_b} g_k(s_k, \sigma_k) \times \prod_{l \neq k} g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} \\
&+ U_k(n-1) g_k(s_k, \sigma_k)^{U_k(n-1)-1} \frac{\partial}{\partial s_a} g_k(s_k, \sigma_k) \times \frac{\partial}{\partial s_b} \prod_{l \neq k} g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} \\
&= \sum_k U_k(n-1)(U_k(n-1)-1) g_k(s_k, \sigma_k)^{U_k(n-1)-2} \frac{\partial}{\partial s_b} g_k(s_k, \sigma_k) \times \frac{\partial}{\partial s_a} g_k(s_k, \sigma_k) \times \prod_{l \neq k} g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} \\
&+ U_k(n-1) g_k(s_k, \sigma_k)^{U_k(n-1)-1} \times \frac{\partial^2}{\partial s_a \partial s_b} g_k(s_k, \sigma_k) \times \prod_{l \neq k} g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)} \\
&+ U_k(n-1) g_k(s_k, \sigma_k)^{U_k(n-1)-1} \times \frac{\partial}{\partial s_a} g_k(s_k, \sigma_k) \times \\
&\quad \sum_{j \neq k} U_j(n-1) g_j(s_j, \sigma_j)^{U_j(n-1)-1} \frac{\partial}{\partial s_b} g_j(s_j, \sigma_j) \prod_{\substack{l \neq k \\ l \neq j}} g_l(\vec{s}, \vec{\sigma})^{U_l(n-1)}
\end{aligned}
\tag{A1.4.11}
$$

where the first equation uses the derivative expanded in Eq A1.4.3.

Evaluation of the above derivative at $\vec{s} = \vec{\sigma} = \vec{1}$ will be useful for the conditional

expectations later:

$$
\begin{aligned}
&\frac{\partial^2}{\partial s_b \partial s_a} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_k U_k(n-1)(U_k(n-1)-1)\ \frac{\partial}{\partial s_b} g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \times \frac{\partial}{\partial s_a} g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ U_k(n-1) \times \frac{\partial^2}{\partial s_a \partial s_b} g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ U_k(n-1) \times \frac{\partial}{\partial s_a} g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \times \sum_{j\neq k} U_j(n-1)\ \frac{\partial}{\partial s_b} g_j(s_j,\sigma_j)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_k \left\{ U_k(n-1)(U_k(n-1)-1) m_b^k m_a^k + U_k(n-1)\bar{\nabla}_{a,b}^k + U_k(n-1) m_a^k \sum_{j\neq k} U_j(n-1) m_b^j \right\} \\
&= \sum_k U_k(n-1)(\bar{\nabla}_{a,b}^k - m_b^k m_a^k) + \sum_{k,j} U_k(n-1) m_a^k U_j(n-1) m_b^j
\end{aligned}
\tag{A1.4.12}
$$

where the first equation uses Eq A1.3.2, the second equation uses definition of m in Eq 5.2.9 and definition of $\bar{\nabla}$ in Eq 5.2.20, the third equation combines the summand indexed with l and summation over $\{j\}_{j\neq l}$ into a non-constrained summation over j.

3. Expand the derivative $\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}$:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&= \frac{\partial}{\partial\sigma_b}\sum_l Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&= \sum_l Z_l(n-1)\frac{\partial}{\partial\sigma_b}f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\times\frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\times\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&+ Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\times\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\times\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&+ Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\times\frac{\partial}{\partial\sigma_b}\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&= \sum_l Z_l(n-1)(Z_l(n-1)-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-2}\frac{\partial}{\partial\sigma_b}f_l(\vec{s},\vec{\sigma})\times\frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\times\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&+ Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\times\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\times\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&+ Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\times\frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\times\sum_{j\neq l} Z_j(n-1) f_i(\vec{s},\vec{\sigma})^{Z_j(n-1)-1}\frac{\partial}{\partial\sigma_b}f_j(\vec{s},\vec{\sigma})\prod_{\substack{i\neq l\\ i\neq j}} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}
\end{aligned}
\tag{A1.4.13}
$$

where the first equation uses the derivative expanded in Eq A1.4.5.

Evaluation of the above derivative at $\vec{s}=\vec{\sigma}=\vec{1}$ will be useful for the conditional

expectations later:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_l Z_l(n-1)(Z_l(n-1)-1)\,\frac{\partial}{\partial\sigma_b}f_l(\vec{s},\vec{\sigma})\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \times \frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ Z_l(n-1)\times\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ Z_l(n-1)\times\frac{\partial}{\partial\sigma_a}f_l(\vec{s},\vec{\sigma})\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \times \sum_{j\neq l} Z_j(n-1)\,\frac{\partial}{\partial\sigma_b}f_j(\vec{s},\vec{\sigma})\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_l \left\{ Z_l(n-1)(Z_l(n-1)-1)\Psi^l_b\Psi^l_a + Z_l(n-1)\bar{\not\Lambda}^l_{a,b} + Z_l(n-1)\Psi^l_a\sum_{j\neq l}Z_j(n-1)\Psi^j_b \right\} \\
&= \sum_l Z_l(n-1)(\bar{\not\Lambda}^l_{a,b}-\Psi^l_b\Psi^l_a) + \sum_{l,j} Z_l(n-1)\Psi^l_a Z_j(n-1)\Psi^j_b
\end{aligned}
\tag{A1.4.14}
$$

where the first equation uses Eq A1.3.2, the second equation uses definition of Ψ in Eq 5.2.11 and definition of $\bar{\not\Lambda}$ in Eq 5.2.32, the third equation combines the summand indexed with l and summation over $\{j\}_{j\neq l}$ into a non-constrained summation over j.

4. Expand the derivative $\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}\prod_i g_i(s_i,\sigma_i)^{U_i(n-1)}$:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&= \frac{\partial}{\partial\sigma_b}\sum_k U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\frac{\partial}{\partial\sigma_a}g_k(s_k,\sigma_k)\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&= \sum_k U_k(n-1)\frac{\partial}{\partial\sigma_b}g_k(s_k,\sigma_k)^{U_k(n-1)-1}\times\frac{\partial}{\partial\sigma_a}g_k(s_k,\sigma_k)\times\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&+ U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\times\frac{\partial^2}{\partial\sigma_a\partial\sigma_b}g_k(s_k,\sigma_k)\times\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&+ U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\frac{\partial}{\partial\sigma_a}g_k(s_k,\sigma_k)\times\frac{\partial}{\partial\sigma_b}\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&= \sum_k U_k(n-1)(U_k(n-1)-1)g_k(s_k,\sigma_k)^{U_k(n-1)-2}\frac{\partial}{\partial\sigma_b}g_k(s_k,\sigma_k)\times\frac{\partial}{\partial\sigma_a}g_k(s_k,\sigma_k)\times\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&+ U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\times\frac{\partial^2}{\partial\sigma_a\sigma_b}g_k(s_k,\sigma_k)\times\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&+ U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\times\frac{\partial}{\partial\sigma_a}g_k(s_k,\sigma_k)\times \\
&\quad\sum_{j\neq k} U_j(n-1)g_j(s_j,\sigma_j)^{U_j(n-1)-1}\frac{\partial}{\partial\sigma_b}g_j(s_j,\sigma_j)\prod_{\substack{l\neq k\\ l\neq j}} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}
\end{aligned}
\tag{A1.4.15}
$$

where the first equation uses the first derivative expanded in Eq A1.4.7.

Evaluation of the above derivative at $\vec{s}=\vec{\sigma}=\vec{1}$ will be useful for the conditional

expectations later:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&=\sum_k U_k(n-1)(U_k(n-1)-1)\ \frac{\partial}{\partial\sigma_b}g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}}\times\ \frac{\partial}{\partial\sigma_a}g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+U_k(n-1)\times\ \frac{\partial^2}{\partial\sigma_a\sigma_b}g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+U_k(n-1)\times\ \frac{\partial}{\partial\sigma_a}g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}}\sum_{j\neq k}U_j(n-1)\ \frac{\partial}{\partial\sigma_b}g_j(s_j,\sigma_j)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&=\sum_k\left\{U_k(n-1)(U_k(n-1)-1)\psi_b^k\psi_a^k+U_k(n-1)\bar{\lambda\!\!\!\lambda}_{a,b}^k+U_k(n-1)\psi_a^k\sum_{j\neq k}U_j(n-1)\psi_b^j\right\} \\
&=\sum_k U_k(n-1)(\bar{\lambda\!\!\!\lambda}_{a,b}^l-\psi_b^k\psi_a^k)+\sum_{k,j}U_k(n-1)\psi_a^kU_j(n-1)\psi_b^j
\end{aligned}
\tag{A1.4.16}
$$

where the first equation uses Eq A1.3.2, the second equation uses definition of ψ in Eq 5.2.12 and definition of $\bar{\lambda\!\!\!\lambda}$ in Eq 5.2.33, the third equation combines the summand indexed with l and summation over $\{j\}_{j\neq l}$ into a non-constrained summation over j.

5. Expand the derivative $\frac{\partial^2}{\partial\sigma_b\partial s_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}$:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\sigma_b\partial s_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&= \frac{\partial}{\partial\sigma_b}\sum_l Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\frac{\partial}{\partial s_a}f_l(\vec{s},\vec{\sigma})\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&= \sum_l Z_l(n-1)\frac{\partial}{\partial\sigma_b}f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\times\frac{\partial}{\partial s_a}f_l(\vec{s},\vec{\sigma})\times\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&+ Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\times\frac{\partial^2}{\partial\sigma_b\partial s_a}f_l(\vec{s},\vec{\sigma})\times\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&+ Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\frac{\partial}{\partial s_a}f_l(\vec{s},\vec{\sigma})\times\frac{\partial}{\partial\sigma_b}\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&= \sum_l Z_l(n-1)(Z_l(n-1)-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-2}\frac{\partial}{\partial\sigma_b}f_l(\vec{s},\vec{\sigma})\times\frac{\partial}{\partial s_a}f_l(\vec{s},\vec{\sigma})\times\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&+ Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\times\frac{\partial^2}{\partial\sigma_b\partial s_a}f_l(\vec{s},\vec{\sigma})\times\prod_{i\neq l} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \\
&+ Z_l(n-1) f_l(\vec{s},\vec{\sigma})^{Z_l(n-1)-1}\times\frac{\partial}{\partial s_a}f_l(\vec{s},\vec{\sigma})\times\sum_{j\neq l} Z_j(n-1) f_i(\vec{s},\vec{\sigma})^{Z_j(n-1)-1}\frac{\partial}{\partial\sigma_b}f_j(\vec{s},\vec{\sigma})\prod_{\substack{i\neq l\\ i\neq j}} f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}
\end{aligned}
\tag{A1.4.17}
$$

where the first equation uses the derivative expaned in Eq A1.4.1.

Evaluation of the above derivative at $\vec{s}=\vec{\sigma}=\vec{1}$ will be useful for the conditional

expectations later:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\sigma_b\partial s_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_l Z_l(n-1)(Z_l(n-1)-1)\,\frac{\partial}{\partial\sigma_b}f_l(\vec{s},\vec{\sigma})\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \times \frac{\partial}{\partial s_a}f_l(\vec{s},\vec{\sigma})\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ Z_l(n-1)\times\frac{\partial^2}{\partial\sigma_b\partial s_a}f_l(\vec{s},\vec{\sigma})\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ Z_l(n-1)\times\frac{\partial}{\partial s_a}f_l(\vec{s},\vec{\sigma})\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \times \sum_{j\neq l} Z_j(n-1)\,\frac{\partial}{\partial\sigma_b}f_j(\vec{s},\vec{\sigma})\bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_l \left\{ Z_l(n-1)(Z_l(n-1)-1)\Psi_b^l M_a^l + Z_l(n-1)\bar{\Omega}_{a,b}^l + Z_l(n-1)M_a^l\sum_{j\neq l}Z_j(n-1)\Psi_b^j \right\} \\
&= \sum_l Z_l(n-1)(\bar{\Omega}_{a,b}^l - \Psi_b^l M_a^l) + \sum_{l,j} Z_l(n-1)M_a^l Z_j(n-1)\Psi_b^j
\end{aligned}
\tag{A1.4.18}
$$

where the first equation uses Eq A1.3.2, the second equation uses the definition of M in Eq 5.2.8, the definition of Ψ in Eq 5.2.11 and the definition of $\bar{\Omega}$ in Eq 5.2.45, the third equation combines the summand indexed with l and summation over $\{j\}_{j\neq l}$ into a non-constrained summation over j.

6. Expand the derivative $\frac{\partial^2}{\partial\sigma_b\partial s_a}\prod_i g_i(s_i,\sigma_i)^{U_i(n-1)}$:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\sigma_b\partial s_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\\
&=\frac{\partial}{\partial\sigma_b}\sum_k U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\frac{\partial}{\partial s_a}g_k(s_k,\sigma_k)\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\\
&=\sum_k U_k(n-1)\frac{\partial}{\partial\sigma_b}g_k(s_k,\sigma_k)^{U_k(n-1)-1}\times\frac{\partial}{\partial s_a}g_k(s_k,\sigma_k)\times\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\\
&+U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\times\frac{\partial^2}{\partial s_a\partial\sigma_b}g_k(s_k,\sigma_k)\times\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\\
&+U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\frac{\partial}{\partial s_a}g_k(s_k,\sigma_k)\times\frac{\partial}{\partial\sigma_b}\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\\
&=\sum_k U_k(n-1)(U_k(n-1)-1)g_k(s_k,\sigma_k)^{U_k(n-1)-2}\frac{\partial}{\partial\sigma_b}g_k(s_k,\sigma_k)\times\frac{\partial}{\partial s_a}g_k(s_k,\sigma_k)\times\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\\
&+U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\times\frac{\partial^2}{\partial s_a\sigma_b}g_k(s_k,\sigma_k)\times\prod_{l\neq k} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\\
&+U_k(n-1)g_k(s_k,\sigma_k)^{U_k(n-1)-1}\times\frac{\partial}{\partial s_a}g_k(s_k,\sigma_k)\times\\
&\quad\sum_{j\neq k}U_j(n-1)g_j(s_j,\sigma_j)^{U_j(n-1)-1}\frac{\partial}{\partial\sigma_b}g_j(s_j,\sigma_j)\prod_{\substack{l\neq k\\ l\neq j}} g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}
\end{aligned}
\tag{A1.4.19}
$$

where the first equation uses the derivative expaned in Eq A1.4.4.

Evaluation of the above derivative at $\vec{s}=\vec{\sigma}=\vec{1}$ will be useful for the conditional

expectations later:

$$
\begin{aligned}
&\frac{\partial^2}{\partial\sigma_b\partial s_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&=\sum_k U_k(n-1)(U_k(n-1)-1)\ \frac{\partial}{\partial\sigma_b}g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}}\times\frac{\partial}{\partial s_a}g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+U_k(n-1)\times\frac{\partial^2}{\partial s_a\sigma_b}g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+U_k(n-1)\times\frac{\partial}{\partial s_a}g_k(s_k,\sigma_k)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}}\times\sum_{j\neq k}U_j(n-1)\ \frac{\partial}{\partial\sigma_b}g_j(s_j,\sigma_j)\Big|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&=\sum_k\left\{U_k(n-1)(U_k(n-1)-1)\psi_b^k m_a^k+U_k(n-1)\bar{\omega}_{a,b}^k+U_k(n-1)m_a^k\sum_{j\neq k}U_j(n-1)\psi_b^j\right\} \\
&=\sum_k U_k(n-1)(\bar{\omega}_{a,b}^l-\psi_b^k m_a^k)+\sum_{k,j}U_k(n-1)m_a^k U_j(n-1)\psi_b^j
\end{aligned}
\tag{A1.4.20}
$$

where the first equation uses Eq A1.3.2, the second equation uses the definition of m in Eq 5.2.9, the definition of ψ in Eq 5.2.12 and the definition of $\bar{\omega}$ in Eq 5.2.46, the third equation combines the summand indexed with l and summation over $\{j\}_{j\neq l}$ into a non-constrained summation over j.

A1.4.1.3 Relation between response moments with and without delayed neutron

Before derive the relation between moments, this section first derive the relation between first and second order derivatives. The relation between MGF (Eq 5.2.101) and the relation between MGF arguments (Eq 5.2.98) are needed. The derivation first uses the chain rule to expand the derivative of MGF with respect to $\vec{s}$, $\vec{\sigma}$ into the derivative of MGF with respect to $\vec{s'}$ (using Eq 5.2.101) and the derivative of $\vec{s'}$ to $\vec{s}$ and $\vec{\sigma}$ (using Eq 5.2.99 and Eq 5.2.100).

$$
\frac{\partial}{\partial s_a}f_l(\vec{s},\vec{\sigma})=\sum_j\frac{\partial}{\partial s'_j}f_l(\vec{s'})\frac{\partial s'_j}{\partial s_a}=\sum_j\frac{\partial}{\partial s'_j}f_l(\vec{s'})\delta_j^a f_j=\frac{\partial}{\partial s'_a}f_l(\vec{s'})f_a \tag{A1.4.21}
$$

$$\frac{\partial}{\partial \sigma_a} f_l(\vec{s}, \vec{\sigma}) = \sum_j \frac{\partial}{\partial s'_j} f_l(\vec{s'}) \frac{\partial s'_j}{\partial \sigma_a} = \sum_j \frac{\partial}{\partial s'_j} f_l(\vec{s'}) \delta^a_j (1 - f_j) = \frac{\partial}{\partial s'_a} f_l(\vec{s'})(1 - f_a) \tag{A1.4.22}$$

$$\begin{aligned} \frac{\partial^2}{\partial s_b \partial s_a} f_l(\vec{s}, \vec{\sigma}) &= \frac{\partial}{\partial s_b} \frac{\partial}{\partial s'_a} f_l(\vec{s'}) f_a \\ &= \sum_j \frac{\partial}{\partial s'_j} \frac{\partial}{\partial s'_a} f_l(\vec{s'}) f_a \frac{\partial s'_j}{\partial s_b} \\ &= \sum_j \frac{\partial}{\partial s'_j} \frac{\partial}{\partial s'_a} f_l(\vec{s'}) f_a \delta^b_j f_j \\ &= \frac{\partial^2}{\partial s'_b \partial s'_a} f_l(\vec{s'}) f_a f_b \end{aligned} \tag{A1.4.23}$$

where the first equality uses Eq A1.4.21.

$$\begin{aligned} \frac{\partial^2}{\partial \sigma_b \partial s_a} f_l(\vec{s}, \vec{\sigma}) &= \frac{\partial}{\partial \sigma_b} \frac{\partial}{\partial s'_a} f_l(\vec{s'}) f_a \\ &= \sum_j \frac{\partial}{\partial s'_j} \frac{\partial}{\partial s'_a} f_l(\vec{s'}) f_a \frac{\partial s'_j}{\partial \sigma_b} \\ &= \sum_j \frac{\partial}{\partial s'_j} \frac{\partial}{\partial s'_a} f_l(\vec{s'}) f_a \delta^b_j (1 - f_j) \\ &= \frac{\partial^2}{\partial s'_b \partial s'_a} f_l(\vec{s'}) f_a (1 - f_b) \end{aligned} \tag{A1.4.24}$$

where the first equality uses Eq A1.4.21.

$$
\begin{aligned}
\frac{\partial^2}{\partial\sigma_b\partial\sigma_a}f_l(\vec{s},\vec{\sigma}) &= \frac{\partial}{\partial\sigma_b}\frac{\partial}{\partial s'_a}f_l(\vec{s'})(1-f_a) \\
&= \sum_j \frac{\partial}{\partial s'_j}\frac{\partial}{\partial s'_a}f_l(\vec{s'})(1-f_a)\frac{\partial s'_j}{\partial\sigma_b} \\
&= \sum_j \frac{\partial}{\partial s'_j}\frac{\partial}{\partial s'_a}f_l(\vec{s'})(1-f_a)\delta^b_j(1-f_j) \\
&= \frac{\partial^2}{\partial s'_b\partial s'_a}f_l(\vec{s'})(1-f_a)(1-f_b)
\end{aligned}
\tag{A1.4.25}
$$

where the first equality uses Eq A1.4.22.

A1.4.2 Evolution of moments

A1.4.2.1 first order moments

A1.4.2.1.1 count of prompt neutrons

$$
\begin{aligned}
&\mathbb{E}\left[Z_a(n)\middle|\vec{Z}(n-1),\vec{U}(n-1)\right] \equiv \mu_a(n)|_{\vec{Z}(n-1),\vec{U}(n-1)} \\
&= \frac{\partial}{\partial s_a}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right]\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \left\{\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\frac{\partial}{\partial s_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} + \frac{\partial}{\partial s_a}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right]\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right\}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \frac{\partial}{\partial s_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} + \frac{\partial}{\partial s_a}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right]\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_l U_l(n-1)m^l_a + Z_l(n-1)M^l_a
\end{aligned}
\tag{A1.4.26}
$$

A1.4.2.1.2 count of delayed neutrons

$$
\begin{aligned}
&\mathbb{E}\left[U_a(n)\left|\vec{Z}(n-1),\vec{U}(n-1)\right.\right] \equiv \rho_a(n)|_{\vec{Z}(n-1),\vec{U}(n-1)} \\
&= \frac{\partial}{\partial\sigma_a}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right]\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \left\{\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\frac{\partial}{\partial\sigma_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} + \frac{\partial}{\partial\sigma_a}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right]\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right\}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \frac{\partial}{\partial\sigma_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} + \frac{\partial}{\partial\sigma_a}\left[\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right]\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \sum_l U_l(n-1)\psi_a^l + Z_l(n-1)\Psi_a^l
\end{aligned}
\tag{A1.4.27}
$$

A1.4.2.2 second order moments

A1.4.2.2.1 Product of count of prompt neutrons

$$
\begin{aligned}
&\mathbb{E}\left[Z_a(n)Z_b(n) - Z_a(n)\delta_{a,b} | \, \vec{Z}(n-1), \vec{U}(n-1) \right] \equiv \complement_{a,b}(n)|_{\vec{Z}(n-1),\vec{U}(n-1)} \\
&= \frac{\partial^2}{\partial s_a \partial s_b} \left(\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \right) \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \Bigg\{ \left(\frac{\partial^2}{\partial s_a \partial s_b} \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \right) \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&+ \left(\frac{\partial}{\partial s_a} \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \right) \left(\frac{\partial}{\partial s_b} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \right) \\
&+ \left(\frac{\partial}{\partial s_b} \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \right) \left(\frac{\partial}{\partial s_a} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \right) \\
&+ \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \frac{\partial^2}{\partial s_a \partial s_b} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \Bigg\} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \frac{\partial^2}{\partial s_a \partial s_b} \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ \frac{\partial}{\partial s_a} \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \frac{\partial}{\partial s_b} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ \frac{\partial}{\partial s_b} \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \frac{\partial}{\partial s_a} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ \frac{\partial^2}{\partial s_a \partial s_b} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= Z_l(n-1)(\bar{\complement}^l_{a,b} - M^l_b M^l_a) + Z_l(n-1) M^l_a Z_j(n-1) M^j_b \\
&+ Z_l(n-1) M^l_a U_k(n-1) m^k_b + U_k(n-1) m^k_a Z_l(n-1) M^l_b \\
&+ U_k(n-1)(\bar{\psi}^k_{a,b} - m^k_b m^k_a) + U_k(n-1) m^k_a U_j(n-1) m^j_b
\end{aligned}
\tag{A1.4.28}
$$

where the second from last equation uses Eq A1.3.2, the last equation uses Eq A1.4.2, Eq A1.4.10, Eq A1.4.4 and Eq A1.4.12.

A1.4.2.2.2 Product of count of prompt and delayed neutrons

$$
\begin{aligned}
&\mathbb{E}\left[Z_a(n)U_b(n)|\, \vec{Z}(n-1), \vec{U}(n-1)\right] \equiv C\Gamma_{a,b}(n)|_{\vec{Z}(n-1),\vec{U}(n-1)} \\
&= \frac{\partial^2}{\partial s_a \partial \sigma_b}\left(\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right)\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \Bigg\{\left(\frac{\partial^2}{\partial s_a \partial \sigma_b}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right)\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&+ \left(\frac{\partial}{\partial s_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right)\left(\frac{\partial}{\partial \sigma_b}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right) \\
&+ \left(\frac{\partial}{\partial \sigma_b}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right)\left(\frac{\partial}{\partial s_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right) \\
&+ \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)} \frac{\partial^2}{\partial s_a \partial \sigma_b}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg\}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \frac{\partial^2}{\partial s_a \partial \sigma_b}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ \frac{\partial}{\partial s_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \frac{\partial}{\partial \sigma_b}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ \frac{\partial}{\partial \sigma_b}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \frac{\partial}{\partial s_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ \frac{\partial^2}{\partial s_a \partial \sigma_b}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= Z_l(n-1)(\bar{\Omega}^l_{a,b} - \Psi^l_b M^l_a) + Z_l(n-1)M^l_a Z_j(n-1)\Psi^j_b \\
&+ Z_l(n-1)M^l_a U_k(n-1)\psi^k_b + U_k(n-1)m^k_a Z_l(n-1)\Psi^l_b \\
&+ U_k(n-1)(\bar{\varpi}^l_{a,b} - \psi^k_b m^k_a) + U_k(n-1)m^k_a U_j(n-1)\psi^j_b
\end{aligned}
\tag{A1.4.29}
$$

where the second from last equation uses Eq A1.3.2, the last equation uses Eq A1.4.2,Eq A1.4.6, Eq A1.4.18, Eq A1.4.4, Eq A1.4.8 and Eq A1.4.20.

A1.4.2.2.3 Product of count of delayed neutrons

$$
\begin{aligned}
&\mathbb{E}\left[U_a(n)U_b(n) - U_a(n)\delta_{a,b}|\, \vec{Z}(n-1), \vec{U}(n-1)\right] \equiv \Upsilon_{a,b}(n)|_{\vec{Z}(n-1),\vec{U}(n-1)} \\
&= \frac{\partial^2}{\partial\sigma_a\partial\sigma_b}\left(\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right)\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \Bigg\{\left(\frac{\partial^2}{\partial\sigma_a\partial\sigma_b}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right)\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)} \\
&+ \left(\frac{\partial}{\partial\sigma_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right)\left(\frac{\partial}{\partial\sigma_b}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right) \\
&+ \left(\frac{\partial}{\partial\sigma_b}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\right)\left(\frac{\partial}{\partial\sigma_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\right) \\
&+ \prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\frac{\partial^2}{\partial\sigma_a\partial\sigma_b}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg\}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= \frac{\partial^2}{\partial\sigma_a\partial\sigma_b}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ \frac{\partial}{\partial\sigma_a}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}}\frac{\partial}{\partial\sigma_b}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ \frac{\partial}{\partial\sigma_b}\prod_i f_i(\vec{s},\vec{\sigma})^{Z_i(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}}\frac{\partial}{\partial\sigma_a}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&+ \frac{\partial^2}{\partial\sigma_a\partial\sigma_b}\prod_l g_l(\vec{s},\vec{\sigma})^{U_l(n-1)}\Bigg|_{\vec{s}=\vec{\sigma}=\vec{1}} \\
&= Z_l(n-1)(\bar{\Lambda}^l_{a,b} - \Psi^l_b\Psi^l_a) + Z_l(n-1)\Psi^l_a Z_j(n-1)\Psi^j_b \\
&+ Z_l(n-1)\Psi^l_a U_k(n-1)\psi^k_b + U_k(n-1)\psi^k_a Z_l(n-1)\Psi^l_b \\
&+ U_k(n-1)(\bar{\lambda}^l_{a,b} - \psi^k_b\psi^k_a) + U_k(n-1)\psi^k_a U_j(n-1)\psi^j_b
\end{aligned}
\tag{A1.4.30}
$$

where the second from last equation uses Eq A1.3.2, the last equation uses Eq A1.4.6, Eq A1.4.14, Eq A1.4.8 and Eq A1.4.16.

A1.4.2.3 temporal-spatial moments

This section derives the second order derivative of the joint moment generating function (Eq 5.2.55) with respect to s_a and s_b. For simplicity, the parameters $\vec{Z}(0)$ and

$\vec{U}(0)$ are hidden in the joint moment generating function.

First evaluate the first order derivative

$$\begin{aligned}
&\frac{\partial^2}{\partial s_a \partial t_b} F_n\left(\vec{u}(k), \vec{\xi}(k)\right) \\
&= \frac{\partial}{\partial s_a}\left\{\sum_j \left(\frac{\partial}{\partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) \frac{\partial u_j}{\partial t_b} + \left(\frac{\partial}{\partial \xi_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) \frac{\partial \xi_j}{\partial t_b}\right\} \\
&= \frac{\partial}{\partial s_a}\left\{\sum_j \left(\frac{\partial}{\partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) s_j \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t}, \vec{\theta}) + \left(\frac{\partial}{\partial \xi_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) \sigma_j \frac{\partial}{\partial t_b} g_j^{(k)}(\vec{t}, \vec{\theta})\right\}
\end{aligned} \tag{A1.4.31}$$

Then take the second derivative for each summand indexed by j. The derivative with respect to s_a of the first j^{th} term in Eq A1.4.31 is

$$\begin{aligned}
&\frac{\partial}{\partial s_a}\left\{\left(\frac{\partial}{\partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) s_j \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t}, \vec{\theta})\right\} \\
&= \left[\frac{\partial}{\partial s_a}\left(\frac{\partial}{\partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right)\right] s_j \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t}, \vec{\theta}) + \left(\frac{\partial}{\partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) \frac{\partial s_j}{\partial s_a} \frac{\partial}{\partial t_b} f_j^{(k)}(\vec{t}, \vec{\theta})
\end{aligned} \tag{A1.4.32}$$

The $\frac{\partial^2}{\partial s_a \partial u_j}$ term in Eq A1.4.32 need to be expanded using chain rule again.

$$\begin{aligned}
&\frac{\partial}{\partial s_a}\left(\frac{\partial}{\partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) \\
&= \sum_i \left(\frac{\partial^2}{\partial u_i \partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) \frac{\partial u_i}{\partial s_a} + \left(\frac{\partial^2}{\partial \xi_i \partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) \frac{\partial \xi_i}{\partial s_a} \\
&= \sum_i \left(\frac{\partial^2}{\partial u_i \partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) \delta_i^a f_i^{(k)}(\vec{t}, \vec{\theta}) + \left(\frac{\partial^2}{\partial \xi_i \partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) \times 0 \\
&= \left(\frac{\partial^2}{\partial u_a \partial u_j} F_n\left(\vec{u}(k), \vec{\xi}(k)\right)\right) f_a^{(k)}(\vec{t}, \vec{\theta})
\end{aligned} \tag{A1.4.33}$$

The derivative with respect to s_a of the second j^{th} term in Eq A1.4.31 is

$$\begin{aligned}&\frac{\partial}{\partial s_a}\left\{\left(\frac{\partial}{\partial \xi_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)\sigma_j\frac{\partial}{\partial t_b}g_j^{(k)}(\vec{t},\vec{\theta})\right\}\\&=\left[\frac{\partial}{\partial s_a}\left(\frac{\partial}{\partial \xi_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)\right]\sigma_j\frac{\partial}{\partial t_b}g_j^{(k)}(\vec{t},\vec{\theta})\end{aligned} \tag{A1.4.34}$$

The $\frac{\partial^2}{\partial s_a \partial \xi_j}$ term in Eq A1.4.34 need to be expanded using chain rule.

$$\begin{aligned}&\frac{\partial}{\partial s_a}\left(\frac{\partial}{\partial \xi_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)\\&=\sum_i\left(\frac{\partial^2}{\partial u_i\partial \xi_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)\frac{\partial u_i}{\partial s_a}+\left(\frac{\partial^2}{\partial \xi_i\partial \xi_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)\frac{\partial \xi_i}{\partial s_a}\\&=\sum_i\left(\frac{\partial^2}{\partial u_i\partial \xi_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)\delta_i^a f_i^{(k)}(\vec{t},\vec{\theta})+\left(\frac{\partial^2}{\partial \xi_i\partial \xi_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)\times 0\\&=\left(\frac{\partial^2}{\partial u_a\partial \xi_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)f_a^{(k)}(\vec{t},\vec{\theta})\end{aligned} \tag{A1.4.35}$$

Substitute Eq A1.4.33 into Eq A1.4.32, substitute Eq A1.4.35 into Eq A1.4.34 and substitute Eq A1.4.32 and Eq A1.4.34 into Eq A1.4.31 yields

$$\begin{aligned}&\frac{\partial^2}{\partial s_a\partial t_b}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\\&=\sum_j\left(\frac{\partial^2}{\partial u_a\partial u_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)f_a^{(k)}(\vec{t},\vec{\theta})s_j\frac{\partial}{\partial t_b}f_j^{(k)}(\vec{t},\vec{\theta})\\&+\left(\frac{\partial}{\partial u_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)\delta_j^a\frac{\partial}{\partial t_b}f_j^{(k)}(\vec{t},\vec{\theta})\\&+\left(\frac{\partial^2}{\partial u_a\partial \xi_j}F_n\left(\vec{u}(k),\vec{\xi}(k)\right)\right)f_a^{(k)}(\vec{t},\vec{\theta})\sigma_j\frac{\partial}{\partial t_b}g_j^{(k)}(\vec{t},\vec{\theta})\end{aligned} \tag{A1.4.36}$$

Bibliography

[1] Forman S Acton. Numerical methods that work. mathematical association of america, washington, dc, usa, 1990. Technical report, ISBN 0-88385-450-3. xx+ 549 pp. LCCN QA297.

[2] Dimitri P Bertsekas and John N Tsitsiklis. *Introduction to probability*, volume 1. Athena Scientific Belmont, MA, 2002.

[3] R.J. Brissenden and A.R. Garlick. Biases in the estimation of Keff and its error by Monte Carlo methods. *Annals of Nuclear Energy*, 13(2):63–83, January 1986.

[4] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of Markov Chain Monte Carlo*. CRC press, 2011.

[5] F. B. Brown, S. E. Carney, B. C. Kiedrowski, and W. R. Martin. *Fission matrix capability for MCNP, Part I - Theory*. American Nuclear Society - ANS; La Grange Park (United States), Jul 2013.

[6] Forrest Brown et al. Wielandt acceleration for mcnp5 monte carlo eigenvalue calculations. In *Joint International Topical Meeting on Mathematics & Computation and Supercomputing in Nuclear Applications (M&C+ SNA 2007), Monterey, California*, 2007.

[7] Forrest B Brown. Fundamentals of monte carlo particle transport. *Los Alamos National Laboratory, LA-UR-05-4983*, 2005.

[8] S. E. Carney, F. B. Brown, B. C. Kiedrowski, and W. R. Martin. *Fission matrix capability for MCNP, Part II - Applications*. American Nuclear Society - ANS; La Grange Park (United States), Jul 2013.

[9] George Casella and Roger L Berger. *Statistical inference*, volume 2. Duxbury Pacific Grove, CA, 2002.

[10] George Casella and Roger L Berger. *Statistical inference*, volume 2. Duxbury Pacific Grove, CA, 2002.

[11] Richard W Conway. Some tactical problems in digital simulation. *Management science*, 10(1):47–61, 1963.

[12] Daryl J Daley and David Vere-Jones. *An introduction to the theory of point processes: volume I: elementary theory and method*. Springer Science & Business Media, 2002.

[13] Stephen C. Wilson Daniel J. Kelly, Thomas M. Sutton. MC21 ANALYSIS OF THE NUCLEAR ENERGY AGENCY MONTE CARLO PERFORMANCE BENCHMARK PROBLEM. In *PHYSOR 2012 âĂŞ Advances in Reactor Physics âĂŞ Linking Research, Industry, and Education*, pages 1–17, Knoxville, Tennessee, USA, 2002.

[14] L Demaret, A Nouri, L Carraro, and O Jacquet. Accurate determination of confidence intervals in monte carlo eigenvalue calculations. In *Proc. 6th Int. Conf. Nuclear Criticality Safety*, pages 20–24, 1999.

[15] James J Duderstadt and Louis J Hamilton. *Nuclear reactor analysis*, volume 1. Wiley New York, 1976.

[16] Jan Dufek and Kaur Tuttelberg. Monte carlo criticality calculations accelerated by a growing neutron population. *Annals of Nuclear Energy*, 94:16–21, 2016.

[17] E. Dumonteil, F. Malvagi, a. Zoia, a. Mazzolo, D. Artusio, C. Dieudonné, and C. De Mulatier. Spatial Correlations in Monte Carlo Criticality Simulations. *SNA + MC 2013 - Joint International Conference on Supercomputing in Nuclear Applications + Monte Carlo*, 03505:03505, June 2014.

[18] Eric Dumonteil, Fausto Malvagi, Andrea Zoia, Alain Mazzolo, Davide Artusio, Cyril Dieudonné, and Clélia De Mulatier. Particle clustering in Monte Carlo criticality simulations. *Annals of Nuclear Energy*, 63:612–618, January 2014.

[19] Eric Dumonteil and Fausto Mavlagi. Variance estimation in monte carlo criticality simulations application to the axial study of a pwr cell with tripoli-4. In *Proceedings of the Second International Conference on Physics and Technology of Reactors and Applications (PHYTRA2), Fez, Morocco*, 2011.

[20] EM Gelbard and R Prael. Computation of standard deviations in eigenvalue calculations. *Progress in Nuclear Energy*, 24(1):237–241, 1990.

[21] Bryan R Herman, Benoit Forget, Kord Smith, Paul K Romano, Thomas M Sutton, Daniel J Kelly, and Brian N Aviles. ANALYSIS OF TALLY CORRELATION IN LARGE LIGHT WATER REACTORS. In *PHYSOR 2014 - The Role of Reactor Physics toward a Sustainable Future*, pages 1–14, The Westin Miyako, Kyoto, Japan, 2014.

[22] Bryan Robert Herman. *Monte Carlo and Thermal Hydraulic Coupling using Low-Order Nonlinear Diffusion Acceleration* . PhD thesis, Massachusetts Institute of Technology, 2014.

[23] N Horelik, B Herman, B Forget, and K Smith. Benchmark for Evaluation and Validation of Reactor Simulations (BEAVRS), v1. 0.1. In *Proc. Int. Conf. Mathematics and Computational Methods Applied to Nuc. Sci. & Eng*, 2013.

[24] Bahram Houchmandzadeh. Theory of neutral clustering for growing populations. *Physical Review E*, 80(5):051920, November 2009.

[25] Il'dar Abdullovich Ibragimov. A note on the central limit theorems for dependent random variables. *Theory of Probability & Its Applications*, 20(1):135–141, 1975.

[26] Galin L Jones. On the Markov chain central limit theorem. *Probability surveys*, 1(299-320):1–5, 2004.

[27] D. J. Kelly Iii, B. N. Aviles, and B. R. Herman. *MC21 analysis of the MIT PWR benchmark: Hot zero power results*. American Nuclear Society - ANS; La Grange Park (United States), Jul 2013.

[28] Maurice George Kendall and Stuart Alan. *THE ADVANCED THEORY OF STATISTICS*, volume 2. CHARLES GRIFFIN. COMPANY LIMITED, 42. DRURY LANE. LONDON, W.C.2, 1961.

[29] JR Lamarsh and Anthony J Baratta. Introduction to nuclear engineering. 3 rd prentice hall. *New Jersey*, 2001.

[30] M. J. Lee, H. G. Joo, D. Lee, and K. Smith. *Monte Carlo reactor calculation with substantially reduced number of cycles*. American Nuclear Society - ANS; La Grange Park, IL (United States), Jul 2012.

[31] Min Jae Lee, Han Gyu Joo, Deokjung Lee, and Kord Smith. Coarse mesh finite difference formulation for accelerated monte carlo eigenvalue calculation. *Annals of Nuclear Energy*, 65:101 – 113, 2014.

[32] Lulu Li et al. *Acceleration methods for Monte Carlo particle transport simulations*. PhD thesis, Massachusetts Institute of Technology, 2017.

[33] Zhaoyuan Liu, Kord Smith, Benoit Forget, and Javier Ortensi. Cumulative migration method for computing rigorous diffusion coefficients and transport cross sections from monte carlo. *Annals of Nuclear Energy*, 112:507–516, 2018.

[34] Brian Peacock Merran Evans, Nicholas Hastings. *Statistical Distributions*. Wiley Series in Probability and Statistics. Wiley-Interscience, 3 edition, 2000.

[35] Jilang. Miao, Benoit. Forget, and Kord. Smith. Analysis of correlations and their impact on convergence rates in monte carlo eigenvalue simulations. *Annals of Nuclear Energy*, 92:81–95, 2016.

[36] Ignas Mickus and Jan Dufek. Optimal neutron population growth in accelerated monte carlo criticality calculations. *Annals of Nuclear Energy*, 117:297–304, 2018.

[37] Shoichiro Nakamura. Computational methods in engineering and science, with applications to fluid dynamics and nuclear systems. 1977.

[38] Paul Reuss. *Neutron physics*. EDP sciences, 2012.

[39] Paul K Romano and Benoit Forget. The OpenMC Monte Carlo particle transport code. *Annals of Nuclear Energy*, 51:274–281, 2013.

[40] Qicang Shen, William Boyd, Benoit Forget, and Kord Smith. Tally Precision Triggers for the OpenMC Monte Carlo Code. *transaction of american nuclear society*, 112:637–640, 2015.

[41] Hyung Jin Shim. Bias of sample variance and real variance estimation in monte carlo eigenvalue calculations, July 2015. Presentation by Shim Hyung Jin at the Workshop on Advanced Monte Carlo Method, Kyoto University Research Reactor Institute, July 21-22, 2015.

[42] Thomas M Sutton. APPLICATION OF A DISCRETIZED PHASE SPACE APPROACH TO THE ANALYSIS OF MONTE CARLO UNCERTAINTIES. In *ANS MC2015 - Joint International Conference Mathematics and Computation (M&C), Supercomputing in Nuclear Applications (SNA) and the Monte Carlo (MC) Method*, number Mc, pages 1–17, Nashville, TN, 2015.

[43] X-5 Monte Carlo Team. Mcnp–a general monte carlo n-particle transport code, version 5. *MCNP user manual. Report LA–UR–03–1987. Los Alamos, CA: Los Alamos National Laboratory*, 5, 2003.

[44] Taro Ueki. Standard deviation of local tallies in global monte carlo calculation of nuclear reactor core. *Journal of nuclear science and technology*, 47(8):739–753, 2010.

[45] Taro Ueki. Fractal dimension analysis for run length diagnosis of monte carlo criticality calculation. *Journal of Nuclear Science and Technology*, 53(3):312–322, 2016.

[46] Taro UEKI, Forrest B. BROWN, D. Kent PARSONS, and Drew E. KORNREICH. Autocorrelation and dominance ratio in Monte Carlo criticality calculations. *Nuclear science and engineering*, 145(3):279–290, 2003.

[47] Taro Ueki, Forrest B Brown, D Kent Parsons, and Drew E Kornreich. Autocorrelation and dominance ratio in monte carlo criticality calculations. *Nuclear science and engineering*, 145(3):279–290, 2003.

[48] Jeffrey A Willert, CT Kelley, DA Knoll, and H Park. A hybrid approach to the neutron transport k-eigenvalue problem using nda-based algorithms. In *Proceedings of International Conference on Mathematics and Computational Methods Applied to Nuclear Science and Engineering (M&C 2013), Sun Valley, ID. American Nuclear Society*, 2013.

[49] Akio Yamamoto, Kotaro Sakata, and Tomohiro Endo. Behavior of Higher Order Fission Source Distribution in Monte-Carlo Calculations. *transaction of american nuclear society*, 109(1):1361–1364, 2013.

[50] Akio Yamamoto, Kotaro Sakata, and Tomohiro Endo. Prediction on Underestimation of Variance for Fission Rate Distribution in Monte-Carlo Calculation. *transaction of american nuclear society*, 110(1):515–518, 2014.

[51] A. Zoia, E. Dumonteil, A. Mazzolo, C. de Mulatier, and A. Rosso. Clustering of branching Brownian motions in confined geometries. *Physical Review E*, 90(4):042118, October 2014.