

TR 200534

**Improved Monte Carlo Methods with Application to
Borehole Logging Simulations**

Part of this work was carried out within the framework of the “Improved Radiation Transport Modelling for Borehole Application” project, which was financed by the 5th framework programme of the European Union

On cover: image based on the petals of a Monte Carlo tulip

The research described in this thesis was performed in the section Physics of Nuclear Reactors of the department Radiation, Radionuclides and Reactors, Faculty of Applied Sciences, Delft University of Technology, Mekelweg 15, 2629JB Delft

Improved Monte Carlo Methods with Application to Borehole Logging Simulations

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.dr.ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen

op dinsdag 29 november 2005 om 13:00 uur

door

Dávid LÉGRÁDY

okleveles geofizikus,
Eötvös Loránd University of Sciences, Boedapest, Hongarije

geboren te Boedapest, Hongarije.

Dit proefschrift is goedgekeurd door de promotoren:

Prof.dr.ir. H. van Dam

Prof.dr.ir. T.H.J.J. van der Hagen

Toegevoegd promotor:

Dr.ir. J.E. Hoogenboom

Samenstelling promotiecommissie:

Rector Magnificus,

Prof.dr.ir. H. van Dam,

Prof.dr.ir. T.H.J.J. van der Hagen,

Dr.ir. J.E. Hoogenboom,

Prof.dr.ir. C.W.E van Eijk,

Prof.dr. W.Th.F. den Hollander,

Prof.dr. C.K. Harris,

Dr. A. Hogenbirk,

voorzitter

Technische Universiteit Delft, promotor

Technische Universiteit Delft, promotor

Technische Universiteit Delft, toegevoegd promotor

Technische Universiteit Delft

Technische Universiteit Eindhoven

Shell E&P Company Houston, University of London

Nuclear research and Consultancy Group (NRG), Petten

Published and distributed by: DUP Science

DUP Science is an imprint of

Delft University Press

P.O. Box 98

2600 MG Delft

The Netherlands

Telephone: +31 15 27 85 678

Telefax: +31 15 27 85 706

E-mail: info@library.tudelft.nl

ISBN 90-407-2614-0

Keywords: Monte Carlo, variance reduction, adjoint, borehole logging

Copyright © 2005 by Dávid Légrády

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the publisher:

Delft University Press

Printed in The Netherlands

Voor mijn vader (1947-2003)

Voor Gábor (1977-2005)



CONTENTS

Introduction	1
Time Dependent Midway Method and Borehole Logging	5
1.1 Theory of the Time Dependent Midway Method.....	6
1.1.1 The time dependent forward response form.....	6
1.1.2 The time dependent adjoint response form	7
1.1.3 The time dependent Midway response form	8
1.1.4 The time dependent Midway response form for coupled neutron-photon calculations.....	10
1.1.5 Virtual surface source and the black absorber technique	12
1.2 Monte Carlo estimation of the Midway response	14
1.2.1 The Monte Carlo Midway Response Estimator	14
1.2.2 Propagation of the statistical error	15
1.2.3 Realisation of the time dependent Monte Carlo Midway method	16
1.3 Basic concepts of nuclear borehole logging.....	18
1.3.1 The borehole environment	19
1.3.2 Neutron-gamma tool geometries.....	21
1.3.3 C/O logging	23
1.3.4 Sigma logging	24
1.3.5 Straight-forward application of the time dependent Midway method to sigma logging.....	26
Analysis of the Midway Coupling with Phase-Space Segmentation.....	29
2.1 Analysis of the response estimate	29
2.1.1 Coupling on Discretised Base Functions	29
2.1.2 One-Dimensional Considerations	31
2.2 Analysis of the Variance Estimate	36
2.3 Technical details of the Midway coupling	41
2.3.1 Three forms of time convolution.....	41
2.3.2 Midway Segmentation for a Borehole Logging Case	45
2.3.3 Some Special Algorithms of the Coupling.....	46
2.4 Summary	50
Midway Coupling with Legendre Polynomials	53
3.1 Function Expansion Using Monte Carlo samples.....	53
3.1.1 Function Expansion on Orthogonal Bases	53
3.1.2 Function Expansion using Legendre Polynomials	56
3.1.3 Numerical examples.....	57
3.2 Midway Coupling with Function Expansion coefficients.....	62
3.2.1 Midway Coupling on Complete Orthogonal Bases.....	62
3.2.2 Convergence of the Midway coupling using Legendre function base.....	63
3.2.3 General estimator for the error of the Midway response estimate	65
3.2.4 Variance estimator for the Legendre Midway coupling.....	68
3.2.5 Applying the Legendre Midway coupling for a borehole logging calculation.....	70

3.3	General Remarks on the Midway Coupling Possibilities.....	74
3.3.1	Expansion of the linking function	74
3.3.2	Optimal Midway coupling	77
Adjoint Sampling of a Pulse Height Distribution.....		79
4.1	Scintillation Detectors, Pulse Height Distribution and non-Boltzmann Tallies...	80
4.1.1	Simulation of scintillation detectors.....	80
4.1.2	The Pulse Height Distribution.....	81
4.1.3	Pulse Height Distribution and non-Boltzmann tallies.....	82
4.2	Feasibility of the Adjoint Sampling of the Pulse Height Distribution	83
4.2.1	Monte Carlo Method for Boltzmann responses	83
4.2.2	Monte Carlo Method for non-Boltzmann responses	86
4.2.3	Collision-wise response of the flux.....	88
4.2.4	Adjoint Sampling of the Pulse Height Distribution	89
4.2.5	Numerical example	94
4.3	Summary	96
Efficiency of the Time Dependent Midway Method Applied to Borehole Logging		97
5.1	Optimisation Options	97
5.1.1	Determining the Midway efficiency	97
5.1.2	One dimensional considerations.....	98
5.1.3	Response flow and Monte Carlo efficiency	101
5.1.4	Contribution flow for a time dependent borehole calculation.....	106
5.2	Enhancement of the Midway Efficiency	
	Using the Weight Window Technique.....	109
5.2.1	The weight window technique	109
5.2.2	Importance functions for the Midway response.....	112
5.2.3	Performance of the WW technique with time dependent weight windows	114
5.2.4	Performance of the WW technique with energy dependent deterministically generated weight windows	117
Conclusions		123
List of Symbols and Abbreviations.....		126
References		127
List of Publications.....		131
Summary		133
Samenvatting		135
Acknowledgements.....		137
Curriculum Vitae		139





INTRODUCTION

This thesis covers the development and assessment of the time dependent Monte Carlo Midway method for application to nuclear borehole logging with the primary aim of providing an acceleration tool for simulating responses on a detailed model parameter space.

Oil detection techniques are becoming more and more sophisticated with the decrease of available oil reserves; both when locating underground geological formations where oil could accumulate, and when drilling **boreholes** for closer inspection of a possible reservoir. Tools lowered into the borehole can yield information on many parameters of the rock formation by measuring physical properties like electrical conductivity or the speed of shock wave propagation. Such measurements allow the identification of the formation type, porosity, oil content etc. Significant amounts of undetected oil reservoirs may be found at sites that were already surveyed but not assessed promising; with improved measuring techniques overlooked reservoirs may be discovered.

Some of the frequently used borehole logging tools apply radiation sources and detectors. Tools with a gamma ray source yield information on the average density of the formation; tools with a neutron source and neutron detectors are applied to identify elements that excessively capture neutrons (e.g. chlorine). This thesis focuses on a **neutron-gamma tool** (a tool equipped with a neutron source and gamma photon detectors) that is often used to deliver information on the ratio of carbon and oxygen abundances in the formation. The importance of this tool is pronounced when revisiting former oil (survey) wells. To prevent the many hundred meters deep borehole from collapsing usually a metal casing is put alongside the borehole wall. This artificial structure (as an example) poses a serious problem for many conventional measuring techniques: it dampens shock waves and shadows electromagnetic fields. For the neutron-gamma tool such an obstacle does not seriously damage the information content, allowing successful measurements even under these circumstances.

Measured raw data must be translated into the quantities of interest, here into formation properties. This **data interpretation** is carried out by comparing the measurements to how the tool responds for various known formations. These reference responses are obtained by measuring in designated test pits, specifically built for this purpose, comprising different rock types, porosities, oil saturations etc. These responses can also be calculated by modelling the underlying physical phenomena. The accuracy of the modelling and the variety of the corresponding formation properties determine the quality of the data interpretation. To lessen the reported insufficiencies and inaccuracies of the data interpretation for the neutron-gamma tool, -as part of the European Commission 5th Framework Programme- the Improved Radiation Transport Modelling for Borehole Applications project was initiated. The investigations presented in this thesis are part of this research.

Our numerical approach the **Monte Carlo method** is an essential tool to solve radiation transport problems, such as nuclear borehole logging simulations. Laws of particle transport are given in a stochastic sense; we cannot predict the exact path that an individual particle will take, but we know the probability of the possible events happening during the life of a particle. Transport Monte Carlo calculations are numerical simulations of particles that obey the same probabilistic laws that a real particle in nature would follow. The simulation is done

sequentially for many particles using random numbers that determine where and when a particle is born, in which direction it travels and when it suffers a collision, what interaction takes place and how its energy and direction get modified, until the particle finally gets absorbed or escape from the system.

As unlikely as it is that two real particles will travel exactly the same way, as unlikely it is that a simulated particle follows the same path as a real particle has followed. It is however rarely of interest to characterise individual particles, it is mostly average quantities such as the average number of interactions per starting particle we are interested in determining. **Deterministic transport calculations** -unlike Monte Carlo simulations- utilise the transport behaviour of the average of the particle population, in other words the behaviour of the flux. The average particle population is calculated over finite-sized subdivisions in the phase space of position, angle, energy and time. This is a result of discretising the governing integro-differential equation, the Boltzmann equation. This discretisation of the deterministic transport calculations introduces approximations that are not present for Monte Carlo simulations limiting the accuracy of the results. On another note, some quantities of interest may not be calculated as a function of the average particle population, and Monte Carlo simulations are the only available option.

To obtain sensible computer times, the number of simulated particles is considerably less in practice than the number of real particles in a measurement. Estimates of measured responses might be given as an average of a few simulated particles giving counts if the detector is small, remotely situated from the source or difficult to reach by the transported particles, resulting in **slow convergence** to statistically stable estimates. To speed up the statistical convergence of the estimated average quantity, the Monte Carlo simulations are often modified or replaced by another Monte Carlo calculation that yields the same response, but deviates from simulating transport analogous to nature. The altered simulation should be statistically equivalent to the original one in terms of yielding the same average value for the detector reading, but it should reduce the resulting statistical variance of the average (while consuming the same computation time).

The most commonly used **variance reduction techniques** focus on increasing the number of particles travelling towards the detector, and on discouraging transport in other directions, while compensating the introduced transport bias by assigning a weight to the particles which changes proportionally. In this way the distance between collisions can be stretched or be shrunk (path stretching), the angle after a scattering (angular biasing) may be differently sampled or the population of the particles can be changed by splitting up particles into independent fragments (splitting) or by rarefying them (Russian Roulette).

A more radical alteration to the regular flow of a Monte Carlo calculation is **adjoint Monte Carlo**. Here particles travel in reverse, starting from the detector instead of the source. These pseudo-particles gain instead of losing energy during collisions, and score at the source. As stated by the general reciprocity theorem, the response obtained in this way equals the response obtained from a regular, "forward" Monte Carlo calculation. This approach can be useful when the source size is considerably larger than that of the detector; thereby the chance for a score is higher using adjoint Monte Carlo. The population averages of the adjoint pseudo-particles are related to the adjoint function, which is also called the importance.

Relocation of the scoring domain to the source is not the only option. It can be proven that the detector response can also be calculated on an arbitrary surface that separates source and detector by a surface integral of the importance and the flux over space, energy and angle. The Monte Carlo interpretation of this response form is called the **Midway method**. It comprises a forward calculation with particles starting from the source, and an adjoint calculation starting from the detector. The scores of both calculations on the surface midway between source and detector are registered and the two calculation results are coupled to estimate the surface integral. Since the scoring domains of both calculations have been replaced by a domain more extensive in size, thus the chance of a hit is increased for both calculations, this method is more effective than a single forward or a single adjoint calculation. Based on these arguments, I. Serov in his PhD research work [9] proposed its application as a general variance reduction tool, and demonstrated its capabilities for numerous radiation transport problems, with many of them considered to be challenging for Monte Carlo simulations.

To obtain the borehole logging tool responses by calculation for measurement interpretation, the required accuracy necessitate Monte Carlo simulations. Beyond the huge variety of formation properties for which calculations should be carried out, Monte Carlo simulation of the neutron-gamma tool in itself is a challenging task because of its slow convergence. This thesis is dedicated to the investigations of applying the Midway method to speed up the Monte Carlo simulation of the neutron-gamma tool. The results are presented in five chapters; each focuses on a different element of the research as described in the following paragraphs.

The neutron-gamma tool is equipped with a time dependent source that is capable of emitting neutrons in short bursts. The photons created by the neutrons in the formation are measured time (and energy) dependently. Therefore the Midway method has been extended to handle time dependent responses. The necessary theoretical developments are treated in **Chapter 1**. Also, the same chapter provides a short introduction to the operation modes and working principles of the neutron-gamma tool. The first chapter concludes with a preliminary application of the time dependent Midway method to a rudimentary model of the tool based on the realisation of the method as developed by I. Serov.

For our application the Midway surface integral has to be carried out in space, angle, energy and time, based on forward and adjoint Monte Carlo scores. This is approximated by discretising the integral into finite elements of the phase-space in all dimensions, and the average flux and the adjoint function is calculated on each segment from the respective scores. Assuming constant behaviour of the flux and the adjoint function on each segment, the integral can be calculated. **Chapter 2** gives the description of this technique and describes improvements for the variance calculation, for the practical realisation of the method and for handling the time variable.

Chapter 3 introduces a different way of carrying out the surface integral by using orthogonal function expansion of the flux and the adjoint function. It is demonstrated that such expansion is feasible using Monte Carlo samples, and its application is shown for the Midway integral with special attention to the statistical variance calculation.

The neutron-gamma tool is equipped with scintillation detectors. The response that this type of detector gives cannot be expressed using the flux and the adjoint function. As both the

reciprocity theorem and the Midway response form are based on the concept of the flux and the adjoint function, it is questionable whether an adjoint or a Midway Monte Carlo calculation would give the same response as the forward one. In **Chapter 4** it is demonstrated, that an adjoint Monte Carlo game is feasible for such situations in a non-multiplying system.

The way, in which the Midway method works, as a variance reduction technique is essentially different from conventional methods, as the Midway method delivers an efficiency gain by altering the transport of neither the adjoint nor the forward Monte Carlo calculations. In **Chapter 5** a demonstration is given, that on top of the efficiency gain of the Midway method, further improvement can be achieved by applying a combination of splitting and Russian Roulette for the forward and adjoint Midway calculations resulting in an efficiency improvement up to a factor 50.

Chapter 1

TIME DEPENDENT MIDWAY METHOD AND BOREHOLE LOGGING

The basic concept of the Midway method is the estimation of the detector response at a surface bounding completely either the source or the detector. In a conventional Monte Carlo simulation, particles arriving at this surface would proceed, and depending on the probability of reaching the detector, possibly yield a score. This probability could be estimated by an adjoint calculation: a transport in reverse; by particles originating from the detector, flying backwards in space and time, while gaining instead of losing energy at collisions. Scoring could be relocated then to this surface, and the response estimation would include a forward and an adjoint calculation. When this technique is applied to enhance the efficiency of the calculation, it forms the Midway method.

The theory of the time independent Midway response estimation is given by *Serov et al.* [1], although the same response form has been also given earlier by *Hayashida et al.* [2], and most comprehensively by *Williams* [3], but with a different derivation. Relocation of the scoring domain to a volumetric domain instead, has been studied by *Cramer* [4] for Monte Carlo applications. *Ueki et al.* [5] provided a comprehensive study of the possibilities of combinations of forward and adjoint Monte Carlo games. *Aboughantous* [6] devised a special Monte Carlo algorithm based on a combined quantity of flux and adjoint function, called the Contribution Monte Carlo Method. *Densmore* [7] used forward-adjoint coupling on a volumetric scoring domain using perturbation theory as grounds, focusing mostly on criticality calculations, but using deterministic calculation to obtain the adjoint function. In a broader sense, replacing the actual detector scoring process by first collision or surface estimators for scintillation detectors has been studied by *Mosher* [8], referred to as the expected value technique. Based on intuition rather than rigorous derivations, countless Monte Carlo practitioners have been using the same set of particles registered on a surface around the detector to investigate the behaviour of different types and sizes of counters by just restarting the calculation from the surface for each scenario (see for example Ref. [14]). Pulsed sources and time dependent detectors have been considered only by *Williams* [3], but have been disregarded in the surface response form.

The Monte Carlo Midway method as a general variance reduction technique is essentially based on the doctorate research of *Serov* [9] whereas the description of the theory of the time independent Midway method is given, together with its Monte Carlo interpretation, and application to some test problems that are considered challenging for Monte Carlo simulations. Our first chapter consists of three parts: the extension of the Midway method for time dependent responses; description of the realisation of the time dependent Midway method; and small introduction to nuclear borehole logging measurements together with results of straight-forward application of the Midway method to a rudimentary tool model.

1.1 Theory of the Time Dependent Midway Method

1.1.1 The time dependent forward response form

The derivation of the time dependent Midway response form begins with the time dependent Boltzmann equation [10]:

$$\begin{aligned} \frac{1}{v} \frac{\partial}{\partial t} \phi(\underline{r}, E, \underline{\Omega}, t) + \nabla \underline{\Omega} \phi(\underline{r}, E, \underline{\Omega}, t) + \Sigma_t(\underline{r}, E) \phi(\underline{r}, E, \underline{\Omega}, t) = \\ \int_{E' \Omega'} \int \Sigma_s(\underline{r}, E' \rightarrow E, \underline{\Omega}' \rightarrow \underline{\Omega}) \phi(\underline{r}, E', \underline{\Omega}', t) d\Omega' dE' + S(\underline{r}, E, \underline{\Omega}, t) \end{aligned} \quad (1.1)$$

where $\phi(\underline{r}, E, \underline{\Omega}, t)$ denotes the flux, as function of the spatial position \underline{r} , energy E , direction $\underline{\Omega}$ and time t ; $\Sigma_t(\underline{r}, E)$ stands for the total cross section and $\Sigma_s(\underline{r}, E' \rightarrow E, \underline{\Omega}' \rightarrow \underline{\Omega})$ for the differential scattering cross section, and $S(\underline{r}, E, \underline{\Omega}, t)$ for the source function. The integrals are carried out on the whole domain of energy and angle and this convention will be applied throughout this thesis: if integration boundaries are not specified the whole model domain should be regarded. It should be noted that time dependence is meant to describe the time dependent behaviour of the particle population, and not of the model parameters. In other words it is not a dynamic calculation: the cross sections do not change as a function of time. This holds for all time dependent forms of this text.

To obtain a unique solution to this equation we need to specify the boundary and initial conditions. In a finite model volume V that is bounded by a surface A , with surface normal \underline{n} directed outwards the boundary condition could be specified as:

$$\phi(\underline{r}, E, \underline{\Omega}, t) = 0 \text{ for } \underline{r} \in A \text{ and } \underline{n} \cdot \underline{\Omega} < 0 \quad (1.2)$$

and an initial condition for an initial time t_i (before the source is turned on) as

$$\phi(\underline{r}, E, \underline{\Omega}, t) = 0 \text{ for } t \leq t_i \quad (1.3)$$

A response R , most commonly the reading of a detector, is defined as a linear functional of the flux by

$$R = \iiint \iiint D(\underline{r}, E, \underline{\Omega}, t) \phi(\underline{r}, E, \underline{\Omega}, t) d\underline{r}^3 dE d\Omega dt \quad (1.4)$$

where $D(\underline{r}, E, \underline{\Omega}, t)$ is the detector function. This form will be referred to as the *time dependent forward response form*.

In practice, a real detector would observe a particle if it had a reaction within, therefore the detector function would be zero outside the detector and proportional inside to the cross section of some reaction the detector is sensitive to. This spatial limitation of the detector function to a finite volume is essential for the derivation of the Midway response form.

Not every counter used for measurements behaves in a way that can be described by Eq.(1.4), as the flux might not contain enough information to allow calculation of the actual detector reading. Such responses are called *non-Boltzmann tallies*. In Chapter 4, we will discuss models for such counters. The D detector function is not strictly related to detectors, more to *detecting* particles; and also the detector response R might also stand for any desired abstract quantity, as long as it can be described as a linear functional of the flux.

The Monte Carlo interpretation of the forward response estimation will be outlined in Chapter 4, heuristically we can say here that it consists of simulating particles originating from the source, following their transport through chains of collisions that generally decrease their energy, until the particles get captured or escape from the system; meanwhile tallying their contribution to the response at the detector.

1.1.2 The time dependent adjoint response form

In a given model geometry, the probability that a particle reaches a certain (phase-space) position is determined by the source characteristics and it is independent of the detector function. In the same way we can define a quantity that is independent of the source function and relates to the detector, namely the probability that a particle at a certain (phase-space) position would contribute to the detector reading, i.e. the *importance* of a position regarding a contribution. This quantity is called the *adjoint function* $\phi^+(\underline{r}, E, \underline{\Omega}, t)$, and the equation it obeys, is adjoint to Eq.(1.1) [11]:

$$-\frac{1}{v} \frac{\partial}{\partial t} \phi^+(\underline{r}, E, \underline{\Omega}, t) - \nabla \cdot \underline{\Omega} \phi^+(\underline{r}, E, \underline{\Omega}, t) + \Sigma_t(\underline{r}, E) \phi^+(\underline{r}, E, \underline{\Omega}, t) = \int \int_{E' \Omega'} \Sigma_s(\underline{r}, E \rightarrow E', \underline{\Omega} \rightarrow \underline{\Omega}') \phi^+(\underline{r}, E', \underline{\Omega}', t) d\Omega' dE' + D(\underline{r}, E, \underline{\Omega}, t) \quad (1.5)$$

The boundary condition is usually given by

$$\phi^+(\underline{r}, E, \underline{\Omega}, t) = 0 \text{ for } \underline{r} \in A \text{ and } \underline{n} \cdot \underline{\Omega} > 0 \quad (1.6)$$

and the initial condition now for a time t_f (some time after the detector stopped measuring) is

$$\phi^+(\underline{r}, E, \underline{\Omega}, t) = 0 \text{ for } t \geq t_f \quad (1.7)$$

To shorten the formulae we proceed now using P as a symbol for all phase space variables except time. This will be followed throughout this thesis: P will stand for all phase space

variables except those that are explicitly written out. In the same way, if the spatial integration domain is not specified as a surface, it is meant to be a volumetric integral.

If we multiply Eq.(1.1) by ϕ^+ and subtract Eq.(1.5) multiplied by ϕ and integrate over all phase space variables in the model volume, and from time t_i to t_f , we get

$$\begin{aligned} & \int_{t_i}^{t_f} \int_V \frac{1}{v} \frac{\partial}{\partial t} (\phi(P, t) \phi^+(P, t)) dP dt + \int_{t_i}^{t_f} \int \nabla \cdot (\underline{\Omega} \phi(P, t) \phi^+(P, t)) dP dt = \\ & = \int_{t_i}^{t_f} \int S(P, t) \phi^+(P, t) dP dt - \int_{t_i}^{t_f} \int D(P, t) \phi(P, t) dP dt \end{aligned} \quad (1.8)$$

Note, that the second term on the RHS of Eq. (1.8) is nothing else than the response R .

The LHS of Eq.(1.8) can be rewritten by evaluating the time integral for the first term, and transforming the volumetric integral on V of the second term into a surface integral on A using the Gauss theorem:

$$\int_V \frac{1}{v} (\phi(P, t_f) \phi^+(P, t_f) - \phi(P, t_i) \phi^+(P, t_i)) dP dt + \int_A \underline{n} \cdot \underline{\Omega} \phi(P, t) \phi^+(P, t) dP dt \quad (1.9)$$

The first term of Eq.(1.9) will vanish because of the initial conditions (Eq. (1.3) and Eq.(1.7)), and the second term because of the boundary conditions (Eq.(1.2) and Eq. (1.6)). Finally we obtain the *time dependent adjoint response form*:

$$R = \int_{t_i}^{t_f} \int S(P, t) \phi^+(P, t) dP dt \quad (1.10)$$

Again, in a source-detector system, S will be nonzero only on a domain where the source is located, and the volumetric integration is reduced to an integral on the source domain. The adjoint response form can be interpreted as an integral of the source function with the probability of scoring from a phase space point. The equality between the forward and adjoint response forms is often referred to as the *general reciprocity theorem*.

The Monte Carlo interpretation of the adjoint equation and response would be further investigated in Chapter 4, but it is worthwhile to mention that during an adjoint Monte Carlo game adjoint pseudo particles start at the detector (the adjoint source), gain energy at their collisions, while time is flowing backwards, and contribution to the response will be tallied at the source.

1.1.3 The time dependent Midway response form

The adjoint response form enables the relocation of the integration domain of the response to the detector volume, but that is not the only option. If both detector and source functions are such that they are nonzero in finite volumes and the source and detector domains are disjoint, the response R can be relocated to a surface fully encompassing either the source or the detector or to the volume bounded by this surface.

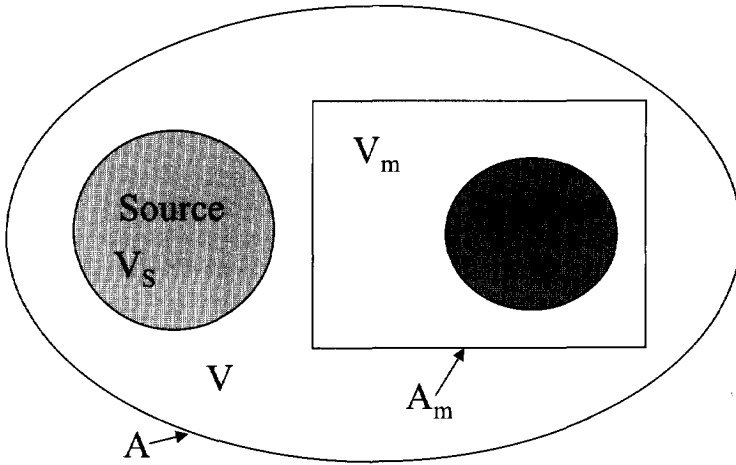


Figure 1. : Integration domains for the Midway response form

This enclosure is sketched on Fig.1, now encompassing the detector. We will refer to it as the *Midway enclosure*. Most of the theorems to be applied are interchangeable regarding forward and adjoint, and respectively source and detector: the enclosure around the source yields similar theorems as enclosure of the detector. This freedom of choice will not be mentioned later on, but regarded as obvious. If we repeat the procedure to arrive to Eq.(1.8), but we integrate only on the volume V_m of the Midway enclosure, we obtain

$$\int_{V_m} \int_{t_i}^{t_f} \nabla \cdot (\underline{\Omega} \phi(P, t) \phi^+(P, t)) dP dt = - \int_{V_m} \int_{t_i}^{t_f} D(P, t) \phi(P, t) dP dt \quad (1.11)$$

as the integral of the time derivatives will vanish again, and also - as the Midway enclosure does not contain any parts of the source domain,- the term containing S disappears too. The LHS of Eq.(1.11) is again an alternative expression for the response R , as the RHS is the forward response form. Such a form is not simple to interpret in Monte Carlo terms, and is not frequently used. Promising attempts have been presented by *Ueki et al* [5] and *Cramer* [4]. Volumetric forward-adjoint integrals as estimators for the response have been investigated by *Densmore* [7], using deterministic adjoint solutions.

Applying the Gauss theorem to the LHS of Eq.(1.11) yields the *time dependent Midway response form*:

$$\int_{A_m} \int_{t_i}^{t_f} n \underline{\Omega} \phi(P, t) \phi^+(P, t) dP dt = - \int_{V_m} \int_{t_i}^{t_f} D(P, t) \phi(P, t) dP dt \quad (1.12)$$

where A_m denotes the surface bounding V_m . If the detection is restricted to a single time moment, for example, the detector function would be $D(P, t) = D(P) \delta(t - t_m)$, and the response at time t_m would become

$$R(t_m) = \int_{V_D} D(P)\phi(P, t_m) dP = - \int_{t_i}^{t_f} \int_{A_m} n\Omega\phi(P, t)\phi^+(P, t) dP dt \quad (1.13)$$

where the adjoint function is calculated according to an adjoint source which is active at time t_m (see Fig. 2).

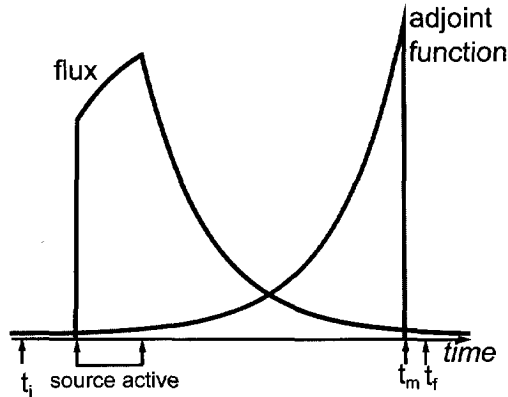


Figure 2. : Sketch of adjoint and forward functions in time

The time dependence of the adjoint function is relative to the starting time t_m . If $\phi^+(P, t)$ is known, the response at t_R can be expressed as

$$R(t_R) = \int_{t_i}^{t_f} \int_{A_m} n\Omega\phi(P, t)\phi^+(P, t + (t_m - t_R)) dP dt \quad (1.14)$$

Forms for utilisation of these results will be shown in Chapter 2, Section 3.

Monte Carlo interpretation of the Midway response form would consist of an adjoint and a forward game, where scores are made at the surface A_m , and the two results will be folded to give an estimate of the surface integral, and with that, to the response.

1.1.4 The time dependent Midway response form for coupled neutron-photon calculations

The Midway response form is the same for neutrons and photons, so are the adjoint and forward equations. In a coupled neutron-photon case, with a neutron source and photon detector, two pairs of adjoint and forward transport equations can be specified, and for each pair additional source specifications would connect the equations. To simplify formulation we write the Boltzmann equation in operator form with B_n and B_p neutron and photon, and with

B_n^+ and B_p^+ adjoint neutron and photon Boltzmann operators, excluding the streaming term and the time derivative:

$$\frac{1}{v_n} \frac{\partial}{\partial t} \phi_n + \underline{\Omega} \nabla \phi_n + B_n \phi_n = S_n \quad (1.15)$$

$$\frac{1}{v_p} \frac{\partial}{\partial t} \phi_p + \underline{\Omega} \nabla \phi_p + B_p \phi_p = S_p \quad (1.16)$$

$$-\frac{1}{v_n} \frac{\partial}{\partial t} \phi_n^+ - \underline{\Omega} \nabla \phi_n^+ + B_n^+ \phi_n^+ = D_n \quad (1.17)$$

$$-\frac{1}{v_p} \frac{\partial}{\partial t} \phi_p^+ - \underline{\Omega} \nabla \phi_p^+ + B_p^+ \phi_p^+ = D_p \quad (1.18)$$

The forward equations are connected by the definition of the photon source term:

$$S_p(P, t) = \iint \Sigma_{pr}(\underline{r}, E' \rightarrow E, \underline{\Omega}' \rightarrow \underline{\Omega}) \phi_n(\underline{r}, \underline{E}', \underline{\Omega}', t) dE' d\Omega' \quad (1.19)$$

with Σ_{pr} the production cross section for photons. The adjoint equations are coupled by an equation adjoint to Eq.(1.19):

$$D_n(P, t) = \iint \Sigma_{pr}(\underline{r}, E \rightarrow E', \underline{\Omega} \rightarrow \underline{\Omega}') \phi_p^+(\underline{r}, \underline{E}', \underline{\Omega}', t) dE' d\Omega' \quad (1.20)$$

In words, the source of photons is coming from photon production at neutron interactions, and the adjoint neutrons are generated at adjoint interactions of adjoint photons. The boundary conditions stay the same for both types of particles, and the initial conditions should also read the same for both. The response is given by

$$R_p = \int_{t_i}^{t_f} \int D_p(P, t) \phi_p(P, t) dP dt \quad (1.21)$$

Eq.(1.15) multiplied by ϕ_n^+ and subtracted from Eq.(1.17) multiplied by ϕ_n , then integrated over all phase space variables on the Midway domain, yields:

$$\int_{t_i}^{t_f} \int n \underline{\Omega} \phi_n \phi_n^+ dP dt = \int_{t_i}^{t_f} \int S_n \phi_n^+ dP dt - \int_{t_i}^{t_f} \int D_n \phi_n dP dt \quad (1.22)$$

Repeating the same procedure for the photon equations give:

$$\int_{t_i}^{t_f} \int n \underline{\Omega} \phi_p \phi_p^+ dP dt = \int_{t_i}^{t_f} \int S_p \phi_p^+ dP dt - \int_{t_i}^{t_f} \int D_p \phi_p dP dt \quad (1.23)$$

The second term on the RHS of Eq.(1.22) equals the first term of the RHS of Eq.(1.23) because of Eq.(1.19) and Eq.(1.20); and S_n is zero in the current Midway enclosure, therefore adding Eq.(1.22) to Eq.(1.23) yields

$$R_p = - \int_{t_i}^{t_f} \int n \Omega \phi_n \phi_n^+ dP dt - \int_{t_i}^{t_f} \int n \Omega \phi_p \phi_p^+ dP dt \quad (1.24)$$

Eq.(1.24) is the time dependent Midway response form for neutron-photon problems, expressing a dual sum for the response that involves both photons and neutrons for a photon response estimation.

1.1.5 Virtual surface source and the black absorber technique

As the theories to be mentioned here will be used later, it is worthwhile to show their derivations. Both theorems, as described by *Serov* [9] are based on the fact, that the derivation of the Midway response forms make use of an integral on the Midway enclosure alone.

Firstly, we consider the Midway response perturbation theory, providing basis for the black absorber technique. When deriving the Midway response form with the operator notation like in Eq.(1.15), we make use of the fact, that

$$\int_{V_m} \phi^+ \mathbf{B} \phi dP = \int_{V_m} \phi \mathbf{B}^+ \phi^+ dP \quad (1.25)$$

This holds also for the whole system domain, but the equality is only required in the Midway enclosure. Repeating the steps of the derivation yielding the time dependent Midway response estimate, would not be affected by changing \mathbf{B} or \mathbf{B}^+ (i.e. changing the cross sections) outside the Midway enclosure. These alterations, however, would have an effect on the flux and the adjoint function. To obtain the detector response on the RHS, like in Eq.(1.12), we must keep the flux unperturbed, meaning that the \mathbf{B} operator should be unchanged; but we can perturb \mathbf{B}^+ outside the Midway domain resulting in the $\tilde{\phi}^+$ perturbed adjoint function. We obtain then Eq.(1.26), with the detector response on the RHS, and the Midway response estimate with the perturbed adjoint function, i.e. the Midway response form would still give the response correctly:

$$\int_{A_m} n \Omega \phi \tilde{\phi}^+ dP = - \int D \phi dP \quad (1.26)$$

Changing the cross section outside the system domain can be arbitrarily radical; it could be even replaced by a black absorber. In practice it means, that one of the two (forward or adjoint) calculations is performed only in half of the geometry, allowing saving computing effort. It has been also proposed to introduce an ideal reflector instead. The Midway perturbation theory seems to contradict intuition, and even hard to believe, though numerical calculations never questioned its validity.

The second theorem we discuss, concerns a rather theoretical detail. Although it will be used in Chapter 5, its derivation is repeated only for the sake of completing the formulation given

by Serov [9]. We now consider replacing the actual source with a surface source on a surface bounding a Midway enclosure V_v in a form of:

$$S_v = -n\Omega\phi(P) \text{ if } \underline{r} \in A_v \quad (1.27)$$

while keeping the boundary conditions and cross sections intact. This obviously would not affect the adjoint equation at all. The flux will change, and the solution of the transport equations with the new surface source term will be denoted by $\tilde{\phi}$. The equivalence of adjoint and forward responses states:

$$\int D(P)\tilde{\phi}(P)dP = \int S_m(P)\phi^+(P)dP = \int_{A_v} n\Omega\phi(P)\phi^+(P)dP \quad (1.28)$$

The last term on the RHS is the Midway response; it is equal to the original detector response. Eq.(1.28) expresses that $\tilde{\phi}$ equals ϕ in the detector domain as the response in the detector using the original and the perturbed flux are equal for an arbitrary detector function. Now let us define a new Midway enclosure V_m around the detector that is inside A_v . If we regard now S_v as our new source, we can write its Midway response form:

$$\int D(P)\tilde{\phi}(P)dP = \int_{A_m} n\Omega\tilde{\phi}(P)\phi^+(P)dP \quad (1.29)$$

Comparing Eq.(1.28) and Eq. (1.29), and recall that the new Midway enclosure is still an enclosure for the original system with the original source, we see that

$$\int_{A_m} n\Omega\phi(P)\phi^+(P)dP - \int_{A_m} n\Omega\phi(P)\phi^+(P)dP = 0 \quad (1.30)$$

for an arbitrary domain, therefore in volume of V_v with an open boundary:

$$\phi \equiv \tilde{\phi} \text{ if } \underline{r} \in V_v \quad (1.31)$$

Eq.(1.31) verifies that the real source can be replaced by a surface source on a Midway enclosure, and this will keep the flux unchanged within the Midway volume.

1.2 Monte Carlo estimation of the Midway response

Particles starting at the source have to hit the detector to give a contribution to the response estimate, and if the chance for it is small, relocation of the scoring domain should be considered. The adjoint response form offers a possibility to choose the source as a scoring domain and play adjoint Monte Carlo, the Midway response form allows scoring at an arbitrary large surface while needing scores from both adjoint and forward games.

The potential of the time independent Midway method to increase the efficiency of scoring, therefore to increase the efficiency of the Monte Carlo game has been discovered and attempted to put in practice in the early 1980's by *Hayasida et al* [2], and its capabilities as a variance reduction technique investigated. Their way of realising the Midway coupling assumed isotropic azimuthal angular dependency. In the 1990's *Serov and Hoogenboom* [1] refined the Midway coupling model by taking properly into account the angular variables, and incorporated it into the general Monte Carlo code MCNP [12] The formulae to be mentioned here, are extensions of their findings for time dependent situations.

1.2.1 The Monte Carlo Midway Response Estimator

Time dependent Midway response form (Eq. (1.12)) poses a challenge for Monte Carlo interpretation. In general, as to be further described later on, a Monte Carlo estimator is in a form of

$$R = \int \wp(x) h(x) dx \quad (1.32)$$

\wp represents here a probability density function (*pdf*) of x , and h stands for some arbitrary scoring function. If N number of (w_j, x_j) samples are drawn from \wp , where w stands for a statistical weight, and x for some coordinate, the estimator for R may be given by:

$$R \approx \frac{1}{N} \sum_{j=1}^N w_j h(x_j) \quad (1.33)$$

When playing a Monte Carlo game, each forward particle reaching the Midway surface can be considered as a sample from an implicitly given pdf, and the scoring function would be the adjoint function, which is yet also undetermined. This implies that some form of an approximation is necessary. *Hayashida* [2] and also *Serov* [1] made use of a discretisation scheme, meaning the division of the phase-space domain into small segments, and on each the functions as integral averages are calculated.

Using this phase-space segmentation technique (we will refer to it as the segmentation technique from now on) relies on the following approximation:

$$R = \int_{A_m} n \Omega \phi \phi^+ dP = \int_{A_m} J \phi^+ dP \approx \sum_i^M \widehat{J}_i \widehat{\phi}_i^+ \Delta P_i = \sum_i^M \int_{\Delta P_i} J dP_i \int_{\Delta P_i} \phi^+ dP_i / \Delta P_i = R_m \quad (1.34)$$

where $J = n\Omega\phi$ denotes the radiation current, ΔP_i stands for the i^{th} phase space segment size on the surface, and $\Delta P_i = \Delta E_j \Delta A_k \Delta \mu_l \Delta \theta_m \Delta t_n$, where ΔE stands for the width of the energy group, ΔA for the size of the surface, $\Delta \mu$ for the polar, $\Delta \theta$ for the azimuthal angle interval width, and t for the time. This approximation might be crude sometimes, especially in energy, when the spectrum is often a discontinuous and irregular function. An analysis on the accuracy of this estimate is given in Chapter 2. An alternative approximation for the response is discussed in Chapter 3. As indicated earlier, in accordance with Eq (1.14), the time dependent Midway response allows the utilisation of the arbitrariness of the starting time of particle: an adjoint and a forward calculation yields not only one, but possibly a series of time dependent responses. For that, three possible forms is given in Chapter 2.

The Monte Carlo surface crossing estimator for the current density is given by

$$\hat{J}_j = \frac{1}{N} \sum_{i=1}^N \frac{w_{i,j}}{\Delta P_j} \quad (1.35)$$

where $w_{i,j}$ denotes the weight of the sample that falls in the j^{th} interval, coming from the i^{th} sample out of N total number of samples. The adjoint function is estimated by

$$\hat{\phi}_j^+ = \frac{1}{N^+} \sum_{i=1}^{N^+} \frac{w_{i,j}^+ / \mu_i}{\Delta P_j} \quad (1.36)$$

where $w_{i,j}^+$ stands for the weight of the i^{th} adjoint particle out of the N^+ total samples, that crosses the Midway surface with μ_i polar angle within the j^{th} segment. Using Eq.(1.35) and Eq.(1.36) the necessary quantities can be obtained for the response estimate. This separate calculation of the adjoint and forward quantities is not always the most effective way, Chapter 2 will comment on the alternatives.

1.2.2 Propagation of the statistical error

Statistical error is always associated with Monte Carlo estimates, and it is also subject to estimation. Of the several forms available, the most widely used is an estimator of the relative variance [12]:

$$r^2(Q) = \frac{\sum_{i=1}^N q_i^2}{\left(\sum_{i=1}^N q_i\right)^2} - \frac{1}{N} \quad (1.37)$$

where q_i is a sample for the quantity Q . For the current at the j^{th} segment ($Q:=J_j$) the sample is $q_i = w_{ij} / \Delta P_j$, and for the adjoint function $q_i = w_{ij}^+ / \mu_i / \Delta P_j$. The statistical error of the Midway response estimate then is caused by the statistical error of the individual estimates of the quantities per mesh.

As indicated by *Hayasida et al* [2] and *Serov* [1], the relative error of the product of the adjoint function and the current at mesh j is obtained by summing up their relative variances, as the adjoint and forward calculations are independent:

$$r^2(\hat{J}_j \hat{\phi}_j^+) = r^2(\hat{J}_j) + r^2(\hat{\phi}_j^+) \quad (1.38)$$

The variance (D^2) of R_m might be calculated from the variance of the response per mesh by

$$D^2(R_m) = \sum_j r^2(\hat{J}_j \hat{\phi}_j^+) \hat{J}_j^2 \hat{\phi}_j^{+2} \Delta P_j^2 \quad (1.39)$$

Here, the sum of the segment-wise variances should also include for a coupled calculation possible neutron and photon responses.

The relative variance of the final response is

$$r^2(R_m) = D^2(R_m) / R_m^2 \quad (1.40)$$

Critics to this statistical model will be formulated in Chapter 2.

The Figure of Merit (FOM) was originally used to measure how trustworthy a Monte Carlo calculation is: if the value of FOM stabilised and did not fluctuate with the number of starting samples, the estimator was converging to a creditable estimate. Later it became also a measure of the efficiency of a simulation.

$$FOM = \frac{1}{r^2 T} \quad (1.41)$$

As Eq.(1.41) shows, the FOM is inversely proportional to the computer time (T) spent on a calculation and to the relative variance. As computer time is proportional to the number of samples started (N), and the relative variance usually shows a $1/N$ dependency as predicted by the central limit theorem, the FOM should be around a constant value, but this constant specific to the given computer system and Monte Carlo simulation.

1.2.3 Realisation of the time dependent Monte Carlo Midway method

The time independent Midway method has been incorporated into the widely used general Monte Carlo code MCNP (*Briesmeister* [12]) by *Serov* and *Hooogenboom* [1] in a user-friendly manner. Subsequent versions of MCNP have been also modified the same way from version 4A to 4C. The latest implementations included the time dependent Midway response form.

The MCNP Midway versions needed modification of several subroutines. The Midway method worked together with most of the standard MCNP options, and the free Midway parameters could be controlled from the generic input file. The segmentation of the phase space on the Midway surface could make use of the built in tally segmentation options of MCNP, except the azimuthal angle which needed additional modifications. The spatial

discretisation of the Midway surface into small facets was far from being automatic. Before version 4C only meshes of regular second order surfaces with the Midway surface could be provided to define the segments (i.e. there was no way to cover the surface e.g. by squares), and also the areas of such facets had to be provided. Serov developed a so-called fish-scale technique, using the intersections of cylinders with the surface forming small areas bounded by segments of circles somewhat similar to scales of a fish (see Fig.3).

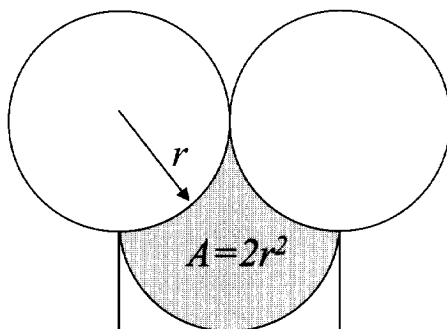


Figure 3. : *Area segment resembling a fish-scale*

The spatial segmentation had an additional rule set by MCNP, that an already covered area could not be segmented any more, and belonged to the facet that first claimed it. As MCNP up till the current version 5 is not capable of continuous energy treatment of the adjoint calculation, the segmentation in energy was determined by the multigroup energy structure of the cross section library. The angular divisions were set according to the cosine of the surface normal for the polar angle, and like the azimuthal angle, were divided into small bins. The later developed time dependent Midway version required equidistant binning in time, allowing the calculation of multiple time dependent responses.

The results of the first (forward or adjoint) calculation for the segment were stored in a file together with their relative variance, and the second calculation was responsible for the coupling of the forward and adjoint results, reading in the file from the previous run. Control parameters were introduced to MCNP to distinguish between the first and the second run, and the type (forward or adjoint) calculation determined the type of function (current or flux) to be estimated. This utilisation ensured a general applicability of the method, but it required the modification of each new version of MCNP. Moreover for the time dependent calculations the number of segments grew higher than the computer could afford in memory at that time. The memory for the segment bins needed to be allocated together with the memory space for cross sections, posing a serious limitation for Midway set-ups. To allow an easier spatial division and to handle the memory problem, from MCNP version 4C the time dependent Midway method obtained a different realisation scheme.

Utilising the so-called PTRAC (particle trajectory) files of MCNP, the Midway quantities are calculated independently of MCNP from the raw score data. The coordinates of the particles at a surface crossing are printed out into the PTRAC files in binary format by a standard instruction in the input file, and a separate program does the sorting of the scores into phase space bins, and estimates the Midway response. This coupling of scores happens after both

runs are done. Such independent coupling allows the investigation of many details of the Midway response estimate, and creates an easy platform for spatial segmentation. With this realisation portability of the coupling code to other operating systems and possibly to other Monte Carlo codes, became possible. Printing out raw scores twice takes more computer time than printing out one coupling file as before, but the difference was found to be insignificant. Also, a score file might be used with different segmentation structures to analyse their behaviour, if the accuracy of the Midway response estimate is questioned. And last but not least it has proven to be a research tool to investigate the Midway coupling characteristics.

Further analysis and refining of the time dependent Midway response calculation induced fundamental alterations to the coupling code; therefore it will be discussed more in detail later on, in Chapter 2.

1.3 Basic concepts of nuclear borehole logging

Borehole logging, or wireline logging is a geophysical exploration technique to investigate formation properties around a borehole. It is an essential tool for exploration of mineral reservoirs such as oil. Based on different physical disciplines (static electricity, electromagnetic waves, seismic waves, etc.), many types of tools have been constructed and a fair share is given for nuclear techniques. Such a tool may be equipped with a neutron or photon source, and neutron or photon detectors, referred to as neutron-neutron, gamma-gamma, neutron-gamma ($n-\gamma$) tools.

The $n-\gamma$ tool, sometimes referred to as the RMT (Reservoir Monitoring Tool), as the RST (Reservoir Saturation Tool) or as the RPM (Reservoir Performance Monitoring) tool, contains a time dependent fast neutron source and multiple gamma detectors. Photons generated at *inelastic collisions* of neutrons have a characteristic energy specific to the hit nucleus. Energy dependent measurement of the generated photons (the so-called *C/O operation mode*) provides information on the elemental composition of the surroundings of the borehole. If used to account for photons emitted at *neutron capture*, the tool detects high capture cross section elements in the vicinity of the borehole (*Sigma operation mode*).

The information on the formation properties should be extracted from the measured data. The inversion procedure that translates measurement results to model parameters, compares simulations (or measurements) on known models to the acquired data. Many simulations have to be performed to cover the model space of the $n-\gamma$ measurements. The quality of model parameter prediction relies on the resolution of the simulations; in case of a Monte Carlo estimation, this might be a challenging task. The basic design of the tool is directed towards high spatial resolution and good signal-to-noise ratio; in other words only informative particle paths are encouraged to enter the detector. This is a relatively small subset of the particles generated, and it causes the Monte Carlo simulation to be highly ineffective, and implies the use of some variance reduction method, the time dependent Midway method for one.

This section is aiming at an introduction to borehole logging concepts and models, and most importantly at providing an enumeration of key problems that are specific to this application of the time dependent Midway method, and need to be handled.

1.3.1 The borehole environment

The most important model constituent of the borehole geometry is the lithology. Major oil-bearing matrices are sand (SiO_2), limestone (CaCO_3) or dolomite (MgCO_3). Hydrocarbons accumulate in free spaces within the formation, such as pores, -or less frequently- discontinuities (fractures or karst dissolutions), and taking up 10-30% of the volume of the rock. The upward migrating hydrocarbons are blocked by non-permeable formations, like clay, silt or shale: formations of fine grain size, mostly silicates. For nuclear transport modelling the atomic constituents are of importance. The most abundant element is oxygen, followed by Si, Ca, Mg, Al, and Fe. Hydrocarbon $(\text{CH}_2)_n$ presence can be detected by the increased C content, though carbon is also present in carbonates. The formation pores if not with oil or gas, are filled with saline water, where Cl is the most important indicator for nuclear measurements.

The drilling process has a considerable influence on the borehole environment. Special borehole fluid (brine) is circulated in the hole during drilling, to transport the fractioned rock to the surface, and to keep the layer pressure in balance. The borehole fluid is a brine solution with fine grain size solids and heavy salts (e.g. BaSO_4). If indeed a porous material is being drilled, the brine infiltrates the rock leaving some of its solid content on the borehole wall that sometimes forms a so-called mudcake. The brine invades the formation, and forces the original pore fluids to withdraw, forming the invaded zone.

Wireline logging is performed usually after a new length of hole is lowered. To prevent the hole from collapsing and to stop the exchange of fluids, the borehole is covered from inside by a metal casing and the gap between casing and formation is filled with cement, before the drilling proceeds. When an oil deposit is found, casing and cementing are perforated for the production phase at convenient spots to access the reservoir.

The production phase would inevitably change the fluid content of the reservoir, and the place of the mined oil will be taken gradually by groundwaters. This is an elaborate process to be controlled and therefore monitored. For that, a separate system of tubes is put in the borehole, where measurement tools can pass. The measurements of the monitoring should account for the many artificial elements in the hole, disqualifying many conventional tools. A metal casing for example shields all electrical and magnetic tools, and even seismic devices, increasing the importance of nuclear tools.

The time dependent Midway method has been applied to two borehole models, and all results presented later are based on these two. The first one we will refer to as the "*Shell*" model, the second one as the "*Generic*" Model. **Table 1** shows the respective parameters.

		"Shell" Model	"Generic" Model
Brine	Inner radius	0	0
	Outer radius	7.74	7.858
	constituents	H ₂ O+ 120mg/l NaCl	H ₂ O+ 120mg/l NaCl
	density	1.0	1.0
Casing	Inner radius	7.74	7.858
	Outer radius	8.89	8.89
	constituents	Fe	Fe
	density	7.86	7.56
Cement	Inner radius	<i>absent</i>	8.89
	Outer radius		10.16
	constituents		Class C Cement
	density		3.14
Formation	Inner radius	8.89	10.16
	Outer radius	80	88.9
	constituents	SiO ₂ +H ₂ O+NaCl (120g/l)	SiO ₂ +H ₂ O+NaCl (120g/l)
	density	2.3	2.3
	porosity	23%	19%
Model bounds	height	200	168
	radius	80	88.9

Table 1. : *Model parameters (distances are in cm, densities in g/cm³)*

1.3.2 Neutron-gamma tool geometries

When a potential oil field is being explored, the decision of drilling a borehole is made on the *reasonable* minimisation of the chance of missing an oil reservoir and not on the maximisation of the number of successful drillings. As oil prices are constantly growing with the decreasing amount of oil left in the deposits, the word 'reasonable' in the above sentence is changing its interpretation too. Exploration sites declared empty now may be revisited and measurements might be redone, but again, the majority of the tools cannot cope with cased boreholes. Neutron-gamma tools are powerful anyway, but hard to substitute in these situation.

All major vendors provided the oil industry with their version of the neutron gamma tool, but exact designs are not public, though some vendors publish [15] more details than others, but as far as it can be known, the tools share several common features. The tool is contained in about half a centimetre thick tube, of length about a meter. The tools are equipped with a D-T neutron source of 14MeV, with the neutron burst lasting for some times $10\mu\text{s}$. Slim tool designs have an outer radius of around 3cm, including half a centimetre thick metal housing.

Usually two (sometimes three) photon detectors measure the signals; the *near detector* is around 20, the *far detector* is around 40 (up to 70) cm away from the source. The detector crystals are BGO (Bismuth Germanate) or NaI scintillators. Because of the better energy resolution, BGO detectors are superior to NaI scintillators, but being very sensitive to temperature, they require constant cooling. Keeping in mind that borehole temperatures above 100°C are not uncommon, the insulation for BGO's should be taken into account when modelling the nuclear transport. Detector crystal radii are about 1.5cm, far below the size commonly used for laboratory measurements. Each detector has a phototube attached to their ends on the opposite side of the source, and at the other ends - to avoid many direct scores from the source- thick metal shieldings are placed.

From the available geometry data, two models have been constructed for use in Monte Carlo simulations. The simpler "Shell" tool model served diagnostic purposes for the application of the Midway method, the more sophisticated "Generic" tool model (see Fig. 4) aimed at approximating actual measurement situations. Table 2 gives the tool parameters for both settings.

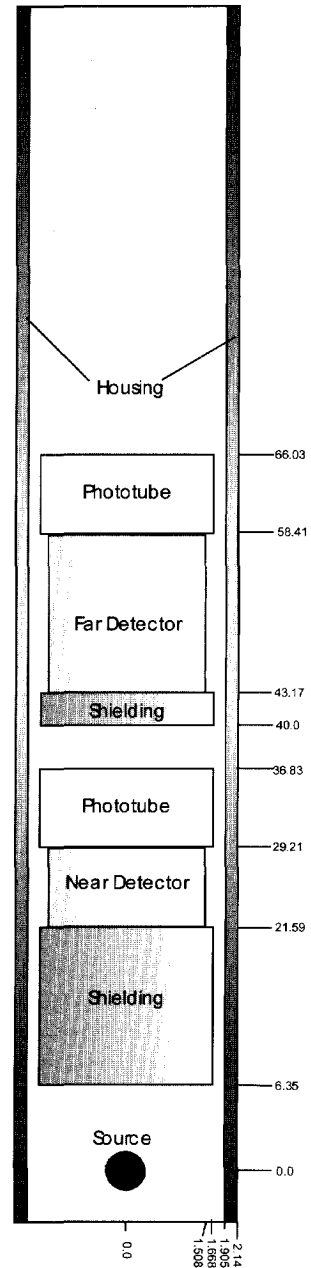


Figure 4.

The generic tool model

		Shell tool model	Generic tool Model
Source	radius	1.27	0
	height	5	0
	distance from (centre of the) source	-2.5	0
	material constituent	vacuum	0
	density	0	0
Near shielding	radius	1.27	1.667
	height	22	15.24
	distance from (centre of the) source	2.5	6.35
	material constituent	W	Fe
	density	19.3	7.86
Near detector	radius	1.27	1.508
	height	5.0	7.62
	distance from (centre of the) source	27.5	21.59
	material constituent	Nal	Nal
	density	3.67	3.67
Near phototube	radius		1.667
	height		7.62
	distance from (centre of the) source	<i>absent</i>	29.21
	material constituent		void
	density		0
Far shielding	radius	1.27	1.667
	height	45	3.175
	distance from (centre of the) source	32.5	40
	material constituent	W	Fe
	density	19.3	7.86
Far detector	radius	1.27	1.508
	height	5.0	15.24
	distance from (centre of the) source	77.5	43.18
	material constituent	Nal	Nal
	density	3.67	3.67
Far phototube	radius		1.667
	height		7.62
	distance from (centre of the) source	<i>absent</i>	58.42
	material constituent		void
	density		0
Housing	radius		1.667->2.143
	height		119.7
	distance from (centre of the) source	<i>absent</i>	-2.5
	material constituent		Fe
	density		7.86
Non-accounted Space Inside tool	material constituent	<i>absent</i>	Fe
	density		3.93

Table 2. : "Generic" and "Shell" Tool Model parameters
(distances in cm, densities in g/cm³)

1.3.3 C/O logging

The first of the two basic operation modes of the neutron-gamma tool is the C/O (Carbon/Oxygen) logging, based purely on photons generated at inelastic collisions, it could also be named as the inelastic mode. Inelastic neutron scattering is dominant for high-energy neutrons ($>1\text{-}2\text{MeV}$). The 14MeV source neutrons slow down in a couple of microseconds to that energy, therefore the inelastic scattering gamma measurement is done during, or shortly after the neutron burst. The emitted gamma rays are specific to the nucleus, for example to Carbon or Oxygen.

Most important inelastic gamma energies of oxygen are the 6.14 , 6.92 and 7.12MeV lines, and 4.43 for carbon. For formation evaluation, mostly the *Sigma* operation mode is used, but inelastic gamma energies of formation qualifier Si (1.8MeV), Ca (3.73 , 3.9MeV) and Mg (1.37MeV) are also detected, sometimes referred as the Ca/Si logging, though it is not a separate operation mode.

The spectral measurement of the photon energies is not an ideal process. Only interactions on the detector can deliver information on the photons. Therefore if the energy of a photon is to be measured, it must release its full energy in the detector. The requirement of a slim tool for post-drilling-phase measurements, limits the detector size, and this increases the chance for partial energy release of the photons. In other words this increases the escape chances of photons from the detector crystal. The crystals applied here are unusually small, and the measured spectrum is far from being restricted to discrete lines. Additionally, photons arriving to the detector might have lost from their energy between their birth place and the detector, and even if measured ideally, would not contribute to the frequency count of the C/O energy lines.

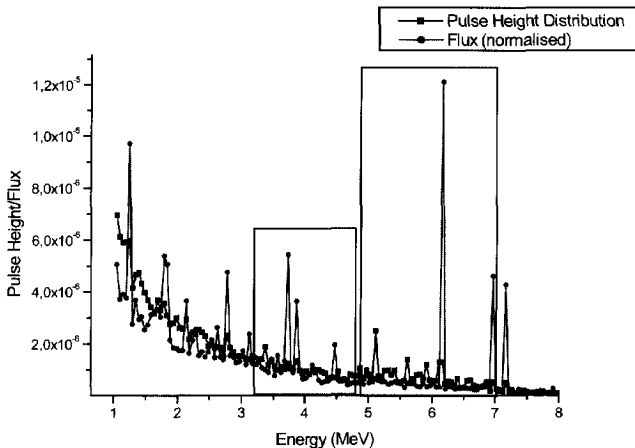


Figure 5. *Simulated scintillation detector spectra*

Fig.5 shows comparison of an ideal photon energy spectrum and a scintillation detector spectrum for the Generic tool model (normalised for better comparison). The prominent oxygen-line at 6.13MeV of the photon flux indicates a significant number of photons arriving to the detector with this energy, but the scintillation crystal is hardly detecting it at all. In practice in place of identifying peaks, broad energy ranges are defined to pick up the signatures of carbon and oxygen. On Fig 5. we indicated with rectangles a possible choice for the carbon (3.21-4.77MeV) and oxygen energy windows (4.8-7.05MeV). The peaks in a realistic measurement are so much smeared out, that the actual information is only a few percent of the measured data. Data can be extracted only with at low signal to noise ratio, where the actual information is a variation of less than 10%. Simulations, therefore, has to be accurate within a few percent.

Theoretical formulation of energy dependent photon detection responses of such small scintillation crystals is far from being obvious. First of all, the response of a scintillation detector is a typical example of a non-Boltzmann response, and because of the small crystal size, not many of the Boltzmann-type detector functions approximate it well enough. The Monte Carlo methods may cope with more difficult responses, but lack of formulation denies any possibilities for variance reduction application, and also of the Midway method. Many attempts have been reported that constructed a well functioning estimator in Boltzmann terms, but even the Monte Carlo estimator gives an inaccurate estimate when comparing it with measured data. Chapter 4 provides a detailed explanation of this problem, and attempts to lay down the foundations of adjoint sampling of the scintillation detectors.

1.3.4 *Sigma logging*

The other operation mode of a neutron-gamma tool is the *sigma* mode, where sigma stands for the capture cross-section of the formation. This operation mode measures a time series of responses, and sometimes utilises energy dependent windows too. Capture of neutrons can also be associated with emission of photons that are characteristic to the nuclei. The rate in which the neutron population decreases yields an estimate to the total capture cross section of the formation. The time evolution of the neutrons can be further characterised by the domains where they are located, when their capture induces a photon: early capture events (<200 μ s) are located mostly in the borehole and the tool, later responses come from domains in the vicinity of the borehole wall (<400-500 μ s), and the latest events happen mostly in the formation domain.

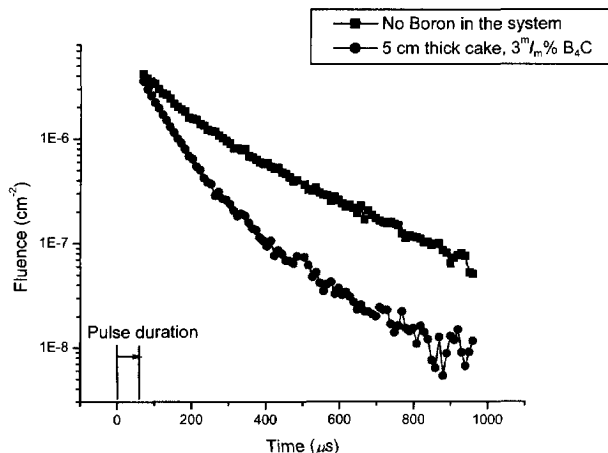


Figure 6. *Time evolution of capture photons: faster decrease indicates presence of high capture cross section elements in the formation*

Figure 6 shows a comparison of two simulated sigma measurements for the near detector in a lin-log plot with a model containing a 5 cm thick mudcake and in one of the simulations this mudcake contains 3^m% (mass-percent) boron carbide. In a homogenous medium the plot should show a near linear decrease, in both graphs a more complicated effect can be seen. The response when boron is not present in the system is apparently not a single linear function, the slope changes around 200-300 μ s. Adding the heavy neutron-absorber Boron to the mudcake, at least three different slopes belonging to different domains can be identified with radically changing cross sections.

The sigma operation mode is a primary tool for lithology evaluation, especially when combined with energy dependent windows accounting for the separate elements. If simulation of the energy spectrum is not needed, the special behaviour of the scintillation detectors pose less of a problem.

Both graphs of Fig 6. show a definite increase of statistical spread for the high time bins even though a vast amount (10^9) of particles were used for the simulation. Keeping in mind that these results were obtained for the near detector, which usually gathers ten times more scores than the far one, some form of an efficiency increasing technique is a must for realistic calculations.

1.3.5 Straight-forward application of the time dependent Midway method to sigma logging

As reported by *Serov and Hoogenboom* [13], application of the time independent Midway method to a neutron-neutron logging tool is advantageous in terms of calculation efficiency. The above described neutron-gamma tool is a more sophisticated problem of the same sort. Similarities in the geometry are substantial, and application of the time dependent Midway method is destined to induce an efficiency improvement. This is demonstrated here, with the intention of pointing out key features and problems that is investigated in the remaining chapters of this thesis.

The results presented here, have been calculated using the time dependent extension of the Midway method as incorporated in MCNP version 3. The Shell model was used with a $100\mu\text{s}$ long source pulse with 20 measuring intervals, up to $500\mu\text{s}$. The Midway surface was selected as a plane perpendicular to the model axis, meshing the near shielding 15cm away from the source centre. 155 fish-scale segments formed the spatial segmentation, 10 and 15 bins the angular divisions, and the standard MCNP ENDF-VI multigroup library structure provided the 30 neutron and 12 photon energy boundaries.

The Midway response estimates were compared to a full forward reference calculation both for accuracy and efficiency. For the reference calculation, only implicit capture, and weight cut-off (i.e. Russian Roulette below a certain weight) was used. The FOM of the Midway method in units of the FOM of the reference calculation (Efficiency Gain) for the near detector can be seen on Fig.7.

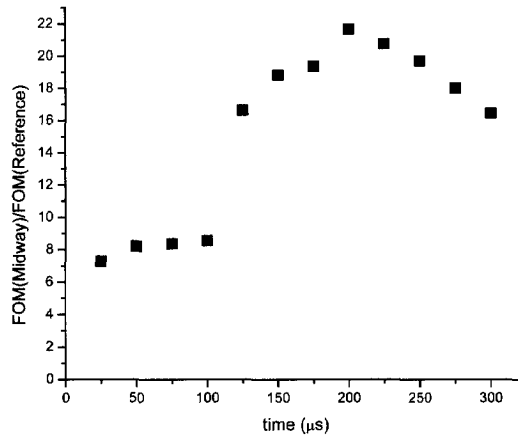


Figure 7. *Efficiency Gain of the Midway method for the Near detector: for higher times the efficiency of the Midway method is 20 times higher than that of an unoptimised forward calculation*

The results show a promising minimum efficiency gain of a factor 7, i.e. the Midway Response estimate converges 7 times faster than an autonomous forward calculation. The gain varies with time, due to a reason to be clarified in Chapter 5. The results were only presented until the 13th time bin, as the variance of the reference calculation for the remaining time bins exhibited the limit where results can be considered reliable.

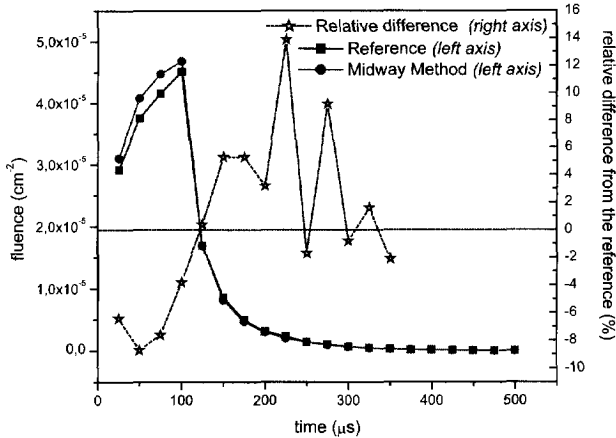


Figure 8. Comparison of response estimates

As promising the efficiency gain the Midway method showed, as discouraging is the comparison of the responses, as can be seen in Fig.8. Symbols connected with solid line are flux estimates at the near detector. Associated statistical errors are not shown, as error bars would be smaller than the symbols representing results. On a relative scale, the statistical errors are below one percent for the first ten bins for the Midway method, and below 2 percents for the reference calculation. The dashed line shows the deviation of the Midway results relative to the reference, and it is obviously out of the range of statistical uncertainty. Deviations of such magnitude are of course not acceptable, and the reasons for it must be found and eliminated, if possible. Chapter 2 is mainly concerned with that problem, and Chapter 3 aims at a solution of this deviation.



Chapter 2

ANALYSIS OF THE MIDWAY COUPLING WITH PHASE-SPACE SEGMENTATION

The time dependent Midway method is a variance reduction technique that estimates the detector response by coupling forward and adjoint Monte Carlo simulations. Of the many questions that can be asked about this sentence, three will be dealt with in this chapter, focusing on the words *estimate*, *variance* and *coupling*. The previous chapter provided formulas on the techniques: how to calculate the estimate, how to calculate the variance and how to do the coupling. Here we investigate the quality of those techniques: how good the estimate for the response is, how good the statistical model of the variance is, and how the computing technique of the coupling can be improved.

It is not only the technical nature of these investigations, that explains why to cover them in the same chapter. As we have seen in Section 1.4.5, results of the application of the time dependent Midway method showed discrepancies when compared to unbiased reference results, beyond statistical error. Reasons for those discrepancies could originate from an erroneous response estimate or an erroneous variance estimate or both; while any modification to these estimators should be feasible in terms of programming, without radically increasing the computing effort.

The topics of this chapter have been researched neither by *Serov*[1,9], nor *Hayasida et al.* [2]. Regarding the accuracy of the Midway responses estimate, the need thereof was not apparent or not reported, as deviations of such extent were not present in the time-independent form. As, clearly, introducing a new dimension for the surface integral increases the inaccuracy of the Midway response estimator.

2.1 Analysis of the response estimate

2.1.1 Coupling on Discretised Base Functions

We have seen in Chapter 1 that the Midway response form is given by

$$R = \int_{A_n} J(P)\phi^+(P)dP \quad (2.1)$$

where both J and ϕ^* are unknown, but respectively (P_i, w_i) and (P_i^+, w_i) samples of them are given. P_i stands for the phase-space coordinates of the particle when crossing the Midway surface. An analogue forward calculation can be considered as direct simulation of possible paths of real particles. In this case, a sample would have a weight (w) of 1 or zero. When the analogue calculation is replaced by a modified, but statistically equivalent game, w might get any positive value, but usually between 0 and 1.

The adjoint and forward Monte Carlo calculations are statistically independent; we might note this by writing the integral of Eq.(2.1) in the form of

$$R_m = \int_{A_m} \int_{A_m} J(P) L(P, P') \phi^+(P') dP' dP \quad (2.2)$$

where the linking function $L(P, P')$ stands for a yet non specified function, introduced to emphasise that the connection of the two quantities must be established. The only way to have $R_m = R$ is to choose $L(P, P') = \delta(P - P')$. That means that the phase-space positions of particles of a forward simulation crossing the Midway surface should exactly be the same as for the adjoint calculation: they have to cross the Midway surface at the exact same position, with the same energy, time, and angle (note that the flight direction of the adjoint particles is reversed). This will normally not happen at all. Given a set of forward crossing coordinates, both the forward and the adjoint particles should be forced to cross at the same positions or a different h (as in Eq.1.32) function should be chosen. The response can be estimated conveniently by independent Monte Carlo calculations, if L is separable in its two variables, therefore if it can be written in a form of

$$L(P, P') = f(P) g(P') \quad (2.3)$$

To put in this term the coupling technique as described in Chapter 1, let us divide the integration domain into K small segments, using a characteristic function for the i^{th} segment, $\Pi_i(P)$, which function is *unity* on the i^{th} interval, and *zero* otherwise; and let us further specify that these divisions are completely covering the integration domain, with each two of them being disjoint, i.e.:

$$\int \sum_{i=1}^K \Pi_i(P) dP = \int dP \quad \text{and} \quad \int \Pi_i(P) \Pi_j(P) dP = \delta_{ij} \int \Pi_i^2(P) dP \quad (2.4)$$

Now we define L as

$$L(P, P') = \sum_{i=1}^K \frac{\Pi_i(P) \Pi_i(P')}{\int \Pi_i^2(P) dP} \quad (2.5)$$

and substitute in Eq.(2.2),

$$R_m = \int_{A_m} \int_{A_m} J(P) \sum_{i=1}^K \frac{\Pi_i(P) \Pi_i(P')}{\int \Pi_i(P) dP} \phi^+(P') dP' dP = \sum_{i=1}^K \frac{\int J(P) dP \int \phi^+(P') dP'}{\Delta P_i} \quad (2.6)$$

R_m takes a similar form to Eq.(1.34). Introducing this notation, might seem more sophisticated than necessary, but unifies the formulation with some later expressions. We will refer to this form of the response estimate as the *segmentation technique*.

The Monte Carlo interpretation is now obvious: it is enough now that P and P' fall in the same segment. It is also trivial, that such a choice of L would only yield exact results if $\Delta P_i \rightarrow 0$ and with that $K \rightarrow \infty$, as long as ϕ and ϕ^+ can be integrated:

$$\lim_{\Delta P_i \rightarrow 0} R_m = \lim_{\Delta P_i \rightarrow 0} \sum_{i=1}^K \int \frac{J}{\Delta P_i} dP \int \frac{\phi^+}{\Delta P_i} dP \Delta P_i = \int J(P) \phi^+(P) dP \quad (2.7)$$

The response estimate is exact with finite K if for $J(P)$ holds that

$$J(P) = \sum_{i=1}^K \int \frac{J(P')}{\Delta P_i} \Pi_i(P') dP' \Pi_i(P) \quad (2.8)$$

or a similar equation for ϕ^+ . In another way of saying if the orthogonal system (as established by Eq.(2.4)) of $\Pi_i(P)$'s would be complete. Eq.(2.8) only holds for constant functions, therefore if both functions have higher order components, the estimate of Eq.(2.6) is not exact. This error is referred to as the truncation error.

It is also possible to interpret L as a function describing the spatial sensitivity of the coupling. An adjoint and a forward sample are regarded the same, as long as they fall in the same segment, and the correction of this approximation is given by the division by the interval width.

If we represent $J(P)$ as given by Eq.(2.8), the better this representation is, the better the quality of the segmented coupling estimate will be. Given a number of samples crossing the Midway surface, the best estimate is given by the highest number of intervals as long as the statistical estimate on each remains valid. These two criteria need to be balanced: a certain accuracy of the response estimate requires a certain amount of segments, and a certain amount of segment requires enough scores

2.1.2 One-Dimensional Considerations

A known deviation of the response can be handled in two basic ways: the residual can be estimated or it can be minimised. As the response estimate is exact if both functions are constant, the first term that causes a deviation is the linear term. To investigate the first non-vanishing error term, let us consider one-dimensional linear functions with the constant terms B and B^+ being actually zero:

$$\phi(x) = Ax + B; \text{ and } \phi^+(x) = A^+x + B^+ \quad (2.9)$$

having A and A^+ the slope of the functions. Taking an interval from a point a to c gives for the exact response estimate:

$$R = \int_a^c \phi(x) \phi^+(x) dx = \frac{AA^+}{3} (c^3 - a^3) \quad (2.10)$$

Using the segmentation technique, if we do not divide this interval, the estimate is given by

$$R_{m1} = \frac{1}{c-a} \int_a^c \phi(x) dx \int_a^c \phi^+(x) dx = \frac{AA^+}{4} \frac{(c^2 - a^2)^2}{c-a} \quad (2.11)$$

and therefore the residual read:

$$\Delta R_m = \frac{AA^+}{3} (c^3 - a^3) - \frac{AA^+}{4} \frac{(c^2 - a^2)^2}{c-a} \quad (2.12)$$

The Monte Carlo estimator for R_{m1} , given (x_i, w_i) and (x_i^+, w_i^+) samples from $\phi(x)$ and $\phi^+(x)$ falling in $[a, c]$, reads:

$$\hat{R}_{m1} = \frac{1}{N} \sum_{i=1}^N w_i \frac{1}{N^+} \sum_{j=1}^{N^+} w_j^+ \quad (2.13)$$

A satisfactory description of the estimate would require the estimate of the residual in first order, therefore and according to Eq. (2.12), the estimate of the slopes. Both as a representation of the slope of the linear term, and as a quantity that can be estimated from Monte Carlo samples, using the response function h (in the sense of Eq.1.32) that will now be given by

$$h(x) = x - \frac{a+c}{2} \quad (2.14)$$

, yields a useful estimate, if applied separately for the two functions. For arbitrary functions, identification of the slopes, which would represent the actual functions, requires already a model, and some information on the nature of those functions. For arbitrary B the integral of h with ϕ yields:

$$\begin{aligned} \int_a^c \phi(x) h(x) dx &= \int_a^c (Ax + B) \left(x - \frac{a+c}{2} \right) dx = A \left[\frac{c^3 - a^3}{3} - \frac{(c^2 - a^2)(c+a)}{4} \right] = \\ &= A \left(\frac{c-a}{2} \right)^3 \frac{2}{3} \end{aligned} \quad (2.15)$$

Such an integral describes a choice of a first order estimate of an arbitrary function, in the sense of *Legendre* polynomial expansion. A Monte Carlo estimator for A is written by:

$$\hat{A} = \frac{1}{N} \sum_{i=1}^N w_i \left(x_i - \frac{a+c}{2} \right) / \left[\left(\frac{c-a}{2} \right)^2 \frac{2}{3} \right] \quad (2.16)$$

As every Monte Carlo estimate, the estimator for A is associated with a relative error, and as a rule of thumb, we can say, that a relative error above 0.1 indicates an unreliable estimate. It can be expected, that the Monte Carlo estimation of the residual requires more samples than of R_{ml} , since the variation of the weights is increased by a variation in x . The required number of samples for a reliable estimate on the response might not suffice in case of the estimator of the slope. To demonstrate this, the convergence of a simple response of the integral of $\phi(x)$ has been compared to the convergence of the estimate of A , having chosen $B=1$, and $A=1$ (see Fig. 9).

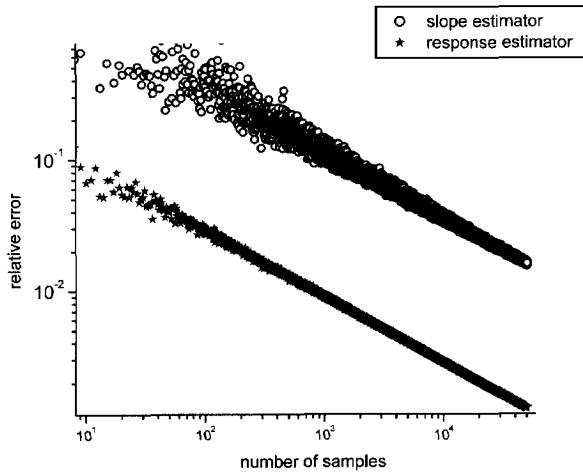


Figure 9. *Convergence of estimators for the response and for the linear component The Monte Carlo estimate for the slope of a function requires considerably more samples than for the integral average of the function.*

If we believe this case to be representative for a general problem, the estimate for the slope converges at least ten times slower than the estimate for the response. In other words the slope-estimator requires ten times more samples for a stable estimate. In a usual Midway application, especially with time dependence, this requirement is far-fetched. Having 10 divisions in each of the independent coordinates, means 10^6 divisions in total, and requires approximately 10^7 scores for a stable response estimate, and 10^9 for estimating the slope. For a borehole logging case, at least for the near detector, 10^9 starting samples give enough scores in the real detector already for a stable response estimate, and as a matter of interest (with the current generally available computing capacities) it would take about three days.

The second path to take, is to minimise the residual, i.e. choosing the distribution of a given number of intervals such, that the deviation is minimal. Taking again the linear functions of

Eq.(2.9), with zero B and B^+ constant terms, we can write the integral in the segmentation approximation on two intervals, inserting a point b between a and c :

$$R_{m_2}(b) = \frac{\int_a^b \phi dx \int_a^b \phi^+ dx}{(b-a)} + \frac{\int_b^c \phi dx \int_b^c \phi^+ dx}{(c-b)} = \frac{AA^+}{4} \left[\frac{(c^2 - b^2)^2}{c-b} + \frac{(b^2 - a^2)^2}{b-a} \right] \quad (2.17)$$

Setting the partial derivative with respect to b of the residual with the 2-segment response to zero, yields an optimal position for the insertion point b :

$$b_{opt} = \frac{a+c}{2} \quad (2.18)$$

The optimised residual become

$$\Delta R_2^{opt} = -\frac{AA^+}{48} (a^3 - c^3 + 3ac^2 - 3a^2c) \quad (2.19)$$

If, in any manner, a higher approximation of the functions than the first order is given, and knowing –as we will also see in Chapter 3– such an approximation is feasible from Monte Carlo samples, an optimisation scheme follows from the linear approximation of the residual. First we divide the interval (a,c) into two parts by inserting b_{opt} . The second step is to check which of the two sections should be divided again, and that decision is made on the residual by filling in the estimated A and A^+ into Eq.(2.19): we choose the interval where the optimum residual is higher. This yields the third interval, by inserting a division point in the middle of the selected segment.

Such an optimisation scheme is hardly feasible when having two-dimensional non-separable functions to integrate. Even if all the linear terms can be identified, the minimisation scheme now would not regard a single point to insert in order to increase the number of divisions, but a continuous closed line. The oversimplified model of linear terms might not yield a more difficult boundary than a second-order geometrical object, but further subdivisions would require sophisticated accounting of an unstructured mesh. This definitely puts the algorithm in this form beyond hope for application in all dimensions. It is possible, however, to treat the functions as separable, and proceed with the iteration scheme likewise, thereby disregarding the essential nature of introducing a second dimension to a function.

One of the functions that will not suit the segmentation technique is the exponential function. Selecting ϕ as decreasing exponential and ϕ^+ as x^{-2} , and selecting the slopes as finite differentials on an interval, we can test this procedure. In this one-dimensional system, the procedure showed an increased rate of convergence to the actual result. Applying the method to two-dimensional functions, where one function is a decreasing exponential in both dimensions (separable), while the other is given by $\phi^+(x,y)=x^{-2}y^{-2}$ and the additional divisions were introduced by selecting also in which dimension it will take place, showed a definitely more impressive increase in convergence. Figure 10. shows the results normalised to 1, with increasing number of segments.

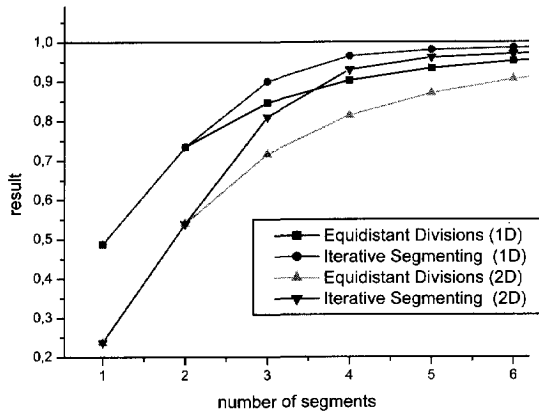


Figure 10. *Response estimates with increasing number of segments using simple equidistant distribution of intervals, and using an optimisation scheme*

It is not surprising, that any alteration to the equidistant division of interval widths provide an increased convergence rate. For any given function, such iteration can be carried out, and it might be advantageous for checking intuition-based divisions. To arrive at a fully optimal segmentation of the phase space requires numerical iteration far more sophisticated than this, and investigations of such have not been able to establish so far a close connection between the truly optimal segmentation into K and the truly optimal segmentation into $K-1$ divisions. Given the strongly approximate nature of the algorithm, it has not been attempted for a Midway calculation further than helping intuition to familiarise with the system.

A simple rule of thumb is given by rewriting the residual according to the discrete representation of J (see Eq. (2.8)):

$$\Delta R_m = \int \left[J(P) - \sum_{i=1}^K \frac{J(P^i)}{\Delta P_i} \Pi_i(P^i) dP^i \Pi_i(P) \right] \phi^+(P) dP \quad (2.20)$$

The differences of the true J to its segmented representation would be weighted by ϕ^+ , therefore choosing more segments at those regions where ϕ^+ is high, gives chance for decreasing the truncation error.

The Monte Carlo estimates should be statistically stable enough for each segment to yield a reliable final response value. It is not impossible, that loosely given estimates of J_i and ϕ_i^+ might still result in a creditable sum, as long as samples fall in a certain segment from both adjoint and forward calculations. The estimate of the response then might not converge properly, and also the estimate of the respective variances might be unsatisfactory. If the total response does converge to a certain extent, we might expect an indication of this undersampling of the estimate by the variance. If most of the segments have no scores at all from a forward run, the adjoint particles will contribute much less in average to the response

as they should when hitting such an unscored segment. If scores are made in each, but the estimate is not converging per segment, the response contribution of adjoint samples might be too high or too low.

As a measure of this statistical instability, the behaviour of the variance with increasing number of segments can be seen in Fig. 11. The functions to integrate by two independent Monte Carlo processes using the segmentation estimate are chosen as

$$\phi = e^{-x} \text{ and } \phi^+ = e^{-x'} \text{ where } x, x' \in [1, 5] \quad (2.21)$$

The number of samples was 10000 for both functions, and the number of segments increased from 1 to 10000 in each dimension. The analytical estimate for the variance is calculated using Eq.(3.43). The results can be seen on Fig.11.

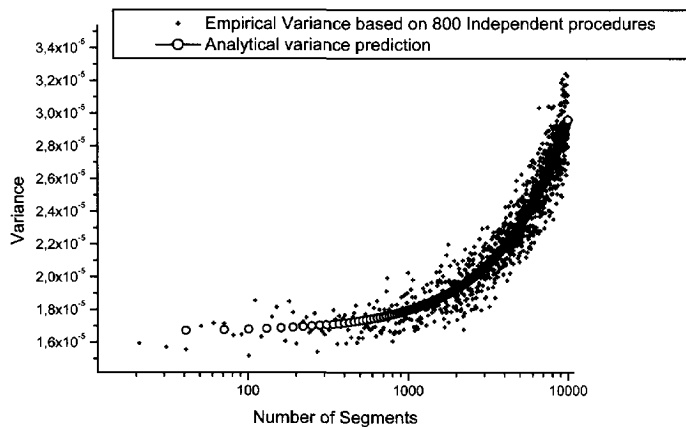


Figure 11. *Estimates of the relative error with increasing number of segments for the same number of samples. If too many intervals are created for a given number of samples, the variance of the Midway estimate grow rapidly*

There are two basic conclusions to formulate about this numerical example. First of all, the response estimate is relatively stable for low segment/sample ratio. Secondly, the variance increases with the number of segments from an initial constant value.

2.2 Analysis of the Variance Estimate

The method of the estimate of the statistical variance of the Midway Monte Carlo response was outlined by Eq.(1.37-1.40). To refine and to correct those formulas, a different statistical model has been established. If we regard the (x_i, P_i) and (x_i^+, P_i^+) samples taken from the

implicitly given pdf's $\wp(P)$ and $\wp^+(P')$ respectively, we can write the Midway integral in the form of an expected value, similar to Eq.(2.2):

$$R_m = \iint \wp(P)\wp^+(P')l(P)l^+(P')L(P,P')dPdP' \quad (2.22)$$

Here l and l^+ are functions that take care of transforming the samples into the current and the adjoint function, e.g. division by the angle cosine for the adjoint function. Such a model is not entirely complete, as we will see in Chapter 4, but for the major differences we will account later on. Its limitation is based on the fact that a particle can score several times even if the Monte Carlo modelling does not apply physical or statistical splitting, but such difference is very much pronounced when dealing with coupled neutron-photon cases. That prohibits using the same symbols as in Eq.(2.2).

If we substitute in Eq. (2.22) for the linking function L the form as given by Eq.(2.5), and using E as a symbol for the expected value, we can write:

$$R_m = \sum_{i=1}^K \frac{E(l(P)\Pi_i(P))E(l^+(P')\Pi_i(P'))}{\Delta P_i} \quad (2.23)$$

given P and P' are independent random variables. This form is equivalent to Eq.(2.6). To shorten the formulas we substitute

$$X_i = l(P)\Pi_i(P) \text{ and } Y_i = l^+(P')\Pi_i(P')/\Delta P_i \quad (2.24)$$

with X_i and Y_i being still independent random variables. The variance D^2 of the response estimate then read:

$$\begin{aligned} D^2\left(\sum_{i=1}^K X_i Y_i\right) &= E\left(\sum_{i=1}^K X_i Y_i\right)^2 - \left[E\left(\sum_{i=1}^K X_i Y_i\right)\right]^2 = \\ &= \sum_{i=1}^K E(X_i Y_i)^2 - \sum_{i=1}^K E^2(X_i Y_i) + \sum_{i=1}^K \sum_{\substack{j=1 \\ j \neq i}}^K E(X_i X_j)E(Y_i Y_j) - E^2(X_i Y_i) \end{aligned} \quad (2.25)$$

The first two terms of the RHS of Eq.(2.25) can be identified as the sum of the variance of the combined $Z_i = X_i Y_i$ random variable, as would be given if all the Z_i 's were independent; and the second term stands for the covariance term.

First we will handle the term of the sum of variances. It can be stated, that

$$\begin{aligned} D^2(X_i Y_i) &= E(X_i Y_i - E(X_i Y_i))^2 = \\ &= E\left(\left(X_i^2 - E^2(X_i)\right)\left(Y_i^2 - E^2(Y_i)\right) + E^2(X_i)E(Y_i^2) + E^2(Y_i)E(X_i^2) - 2E^2(Y_i)E^2(X_i)\right) \end{aligned} \quad (2.26)$$

and easily verified by expanding the first term of the RHS. If we divide the variance by the expected value of Z_i , we will get a simple formula for the relative variance r^2 of each segment:

$$\begin{aligned}
 r^2(X_i Y_i) &= \frac{D^2(X_i Y_i)}{E^2(X_i Y_i)} = \frac{D^2(X_i) D^2(Y_i)}{E^2(X_i) E^2(Y_i)} + \frac{D^2(X_i)}{E^2(X_i)} + \frac{D^2(Y_i)}{E^2(Y_i)} = \\
 &= r^2(X_i) r^2(Y_i) + r^2(X_i) + r^2(Y_i)
 \end{aligned} \tag{2.27}$$

The difference of Eq.(2.27) and Eq.(1.38) is in the second order, and if we accept the requirement of keeping the relative error r below 0.1 for accepting an estimate, this correction term does not play a significant role. However, if many of the interval responses are undersampled, this so-far neglected term gives a significant contribution, and helps identifying such a situation.

The Monte Carlo estimate is an average of many (N) samples, each taken from the same implicit distribution, meaning X_i should be substituted by $\sum x_{ij}/N$ yielding the well known estimator of Eq.(1.37) for r^2 , with $q_j = x_{ij}$ and analogously the same should be done for Y_i . Before we proceed with these substitutions (before handling the covariance term), in order to ease identifying the role of some of the coming transformations, we give first the derivation of Eq.(1.37). The derivation relies on proving the theorem that the variance of the sum of independent samples equals the average of their variance. The derivation can be done much simpler, it is only given here to serve as an analogy for the handling of the covariance term. Let us have z_i independent samples from the same distribution, and calculate the variance of their average:

$$\begin{aligned}
 D^2\left(\frac{\sum_{i=1}^N z_i}{N}\right) &= \frac{1}{N^2} \left[E\left(\sum_{i=1}^N z_i\right)^2 - E^2\left(\sum_{i=1}^N z_i\right) \right] = \\
 &= \frac{1}{N^2} \left[E\left(\sum_{i=1}^N z_i + \sum_{i=1}^N \sum_{j \neq i}^N z_i z_j\right) - \sum_{i=1}^N E\left(\sum_{i=1}^N z_i\right)^2 \right] = \\
 &= \frac{1}{N^2} \left[NE(z_i)^2 + N(N-1)E^2(z_i) - N^2 E^2(z_i) \right] = \\
 &= \frac{1}{N} \left[E(z_i)^2 - E^2(z_i) \right] = \frac{1}{N} D^2(z_i)
 \end{aligned} \tag{2.28}$$

The last line of Eq.(2.28) contains two expected values to be estimated, and that might be given by the average of z_i^2 and z_i . Filling in those estimates into Eq.(2.28), the estimate for the variance reads:

$$\widehat{D^2} = \frac{1}{N^2} \sum_{i=1}^N z_i^2 - \frac{1}{N^3} \left(\sum_{i=1}^N z_i \right)^2 \tag{2.29}$$

Eq.(2.29) divided by the square of the average over all samples, gives Eq.(1.37). For large N number of starting samples the second term will be far less significant than the first one.

Now we can proceed to handle the covariance term. The covariance term of Eq. (2.25) is a negative contributor to the variance. This is based on the fact, that if a sample falls in an

interval, it cannot fall in another one too. Given that the Π_i characteristic functions are mutually disjoint, the first term of the covariance vanishes as it contains products of only disjoint characteristic functions, but the second term is non-zero, giving:

$$\text{Cov} = - \sum_{i=1}^K \sum_{\substack{j=1 \\ j \neq i}}^K E^2(X_i Y_j) \quad (2.30)$$

The covariance term does not simplify so easily, when considering X and Y as an average of samples drawn from the same distribution. The covariance reads if N and N^+ samples are taken for the forward and adjoint calculations respectively:

$$\frac{1}{(NN^+)^2} \left[\sum_{i=1}^K \sum_{\substack{j=1 \\ j \neq i}}^K \left\{ E \left(\sum_{m=1}^N x_{i,m} \sum_{n=1}^N x_{j,n} \right) E \left(\sum_{m=1}^{N^+} y_{i,m} \sum_{n=1}^{N^+} y_{j,n} \right) - E^2 \left(\sum_{m=1}^{N^+} y_{i,m} \sum_{n=1}^N x_{j,n} \right) \right\} \right] \quad (2.31)$$

Samples are independent, therefore we know, that

$$E(x_{i,m} x_{j,n}) = (1 - \delta_{mn}) E^2(x_{i,m}) \quad (2.32)$$

meaning that the first term of Eq. (2.31) vanishes only (provided the same equality is written for the $y_{i,m}$ random variables), if the same sample is considered. The sums can be then evaluated, yielding

$$\sum_{i=1}^K \sum_{\substack{j=1 \\ j \neq i}}^K E(x_{i,m}) E(x_{i,m}) E(y_{i,m}) E(y_{i,m}) \frac{1}{(NN^+)^2} \left[N(N-1)N^+(N^+-1) - (NN^+)^2 \right] \quad (2.33)$$

We can rewrite the Eq. (2.33) in a more interpretable form:

$$\text{Cov} = \left(R^2 - \sum_i E^2(x_i) E^2(y_i) \right) \left(\frac{1}{NN^+} - \frac{1}{N} - \frac{1}{N^+} \right) \quad (2.34)$$

We see, that the first term of the product is very similar to a form of calculating the variance, except here not a statistical variation is measured, but the variation of the interval responses. This term is weighted by a negative coefficient, which is a function of the number of starting samples, very much analogously to the $1/N$ weighting of the variance of the single random variables in Eq.(2.28). If we fill in the estimators for the expected values, we get in the first term of the RHS of Eq.(2.34), the weighting term with the number of starting samples will be $1/N^3 N^{+2}$, $1/N^2 N^{+3}$ and $1/N^3 N^{+3}$; i.e. this term will not likely be a significant contributor to the variance estimator, if huge numbers of adjoint and forward samples are taken.

Unfortunately, the negative correlation caused by the segmentation of the phase space is not the only covariance term to take into account, as it is also not true, that a source particle will cause only one contribution at the Midway surface. This might even occur in an analogue Monte Carlo calculation. As it has been indicated in Chapter 1, in a coupled neutron-photon Midway calculation both the neutrons and both the photons deliver scores. As the neutron

should not necessarily be captured to result in photon emission, it often happens that both the neutron and the photons it generated reach the Midway surface. As they both belong to the same starting particle, they belong statistically to the same sample. The same thing happens, when a single particle re-crosses the Midway surface. The Monte Carlo games are rarely played without any modifications to the analogue sampling methods, sometimes resulting in many fragments of a particle walking different paths. Accounting for this effect on the variance would require the estimation of the correlation matrix of all the segments, requiring unreasonable storage capacity.

If we resort to the description of the variance as it has been done in Eq. (1.39), in the way of the notations of Eq.(2.27), we might write:

$$D^2(R) \approx \sum_{i=1}^K E^2(X_i) D^2(Y_i) + E^2(Y_i) D^2(X_i) \quad (2.35)$$

In words, the variance of the response can be approximated by the variance of the forward calculation per segment weighted by the adjoint segment-wise response squared, summed with the same term but respective forward and adjoint quantities swapped. If we obtain two fragments x_i and x_j of a forward score, one falling in interval i the other in interval j , the sum can be calculated by calculating the variance contribution as

$$D^2(x_i Y_i + x_j Y_j) \approx D^2(x_i E^2(Y_i) + x_j E^2(Y_j)) + D^2(Y_i) E^2(x_i) + D^2(Y_j) E^2(x_j) \quad (2.36)$$

Estimating the correlation of both X and Y fragments would require the full covariance matrix, except if both fragments fall in the same segment.

To summarise the findings in brief, the former error calculation model was not exact, and two of the missing terms can be incorporated with relative ease. It was also seen that a major impact can only be expected from the correction of Eq.(1.38), and is only significant, if the number of segments are in the order of the number of samples. With heavily splitted samples, their covariance might also be of importance.

To demonstrate and to quantify these effects, a Midway borehole logging calculation for the Generic tool model for the carbon window of the near detector has been repeated 50 times, using every time the same forward samples (1.2 million), and different and independent sets of 60000 adjoint samples. Fig. 12 shows the estimates with their variance. The dashed lines are put in a distance of two times the standard deviation based on the average of the predicted error using the former statistical model. The data show apparently a higher spread then accounted for. The empirical deviation of the results, as estimated from the 50 calculations, was $1.92 \cdot 10^{-6} \pm 2.7 \cdot 10^{-7}$ (9% on relative scale). The old statistical method predicted $7.11 \cdot 10^{-7} \pm 1.7 \cdot 10^{-7}$ on average (3.4% on relative scale), while a calculation that incorporated the above identified additional terms, resulted in $1.87 \cdot 10^{-6} \pm 3.7 \cdot 10^{-7}$ clearly better tallying with the experimental spread. The number of adjoint samples was intentionally selected rather low, therefore such result does not mean that previous estimates of the Midway applications are all untrustworthy. It is, though, to be questioned, whether relative errors of a couple of percent (at least above 3%) obtained with the former error calculation model should be accepted as

reliable estimates, and those calculations may better be redone. Also, it can be stated, that the improved model does indicate higher variance resulted from undersampling of the integral.

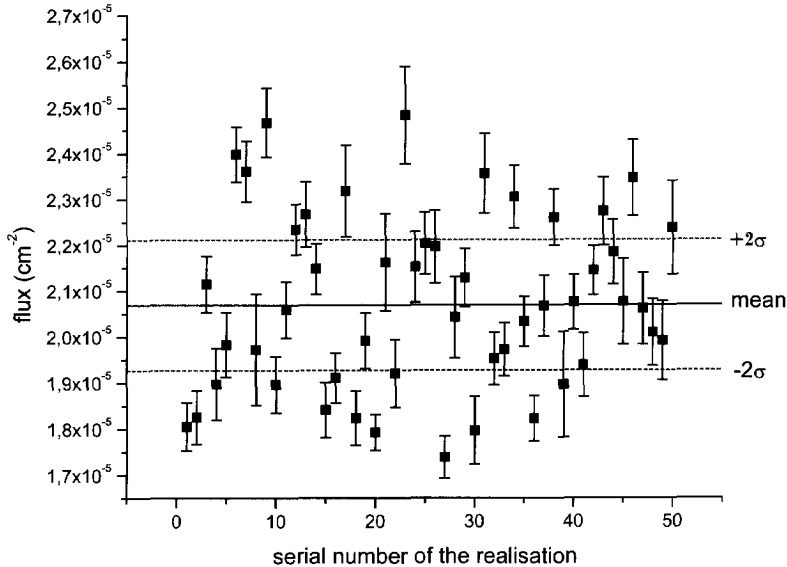


Figure 12. *Midway estimates for the response and the standard deviation (old variance model). The uncorrected variance formula results in an underestimate of the variance.*

2.3 Technical details of the Midway coupling

2.3.1 Three forms of time convolution

A single forward calculation can deliver many responses, and each needs its own detector function to be defined. If the responses are estimated by an adjoint game, each detector function serves as an adjoint source, and each determines different adjoint functions. In the same way, multiple source functions can be integrated with an adjoint function, delivering again multiple responses from the same calculation. It is more common to define several responses for a model than several sources. Different detector responses do not always require a completely separate adjoint calculation, it is often possible to run the adjoint calculation with many adjoint sources, and just account for the different responses by proper bookkeeping. Let us say, if two energy ranges are specified for the detection at the same detector volume, it is common to sample both energy ranges in the same time in the adjoint calculation, and register for each particle the starting energy to distinguish between the two detector functions.

In case of time dependent Monte Carlo calculations, multiple time responses might be estimated from a single adjoint run, without any adjoint source bookkeeping. Let us divide a time interval equidistantly into L segments. Let the source be active during the first interval, and the response calculated at the third one. It is equivalent to the set-up, when the source is active only on the second interval, and the detector on the fourth. The same idea works for the adjoint, just the other way around. Therefore, the set of responses obtained from the forward calculation (for different detector functions) would be the same as the set of responses obtained from the adjoint calculation (for different source functions).

Having a detector only active before time t , the adjoint function will be zero after that time, and nonzero before t ; and depending on our choice of time zero, it might still give values at negative times. It eases Monte Carlo bookkeeping (and it is done so in MCNP) to account for the elapsed time of adjoint particles, as they were forward particles, i.e. the time passes forward instead of backward. If the detector reading is at time t_m , the corresponding adjoint function ϕ_m^+ can be expressed with an adjoint function ϕ_0^+ corresponding to a detector reading at $t=0$, with the time bookkeeping reversed:

$$\phi_m^+(P, t) = \phi_0^+(P, t_m - t) \quad (2.37)$$

Then the Midway response would take a form of a convolution in time:

$$R(t_m) = \int \int_{A_m} n \Omega \phi(P, t) \phi_0^+(P, t_m - t) dt dP \quad (2.38)$$

The dependency in time of the detector function is conveniently given by a characteristic function like $\Pi(t)$, i.e. the detector is switched on at a certain time, and switched off later, while constant time behaviour is given in between. We might associate t_m with the beginning of the reading. If we shift t_m to different time points, the calculated adjoint function that corresponds to the same shape of Π but with a starting time $t=0$, provides enough information for all those responses. Each shift would require a different segmentation in time for the calculation of the Midway integral. If the segmentation is equidistant in time, and we shift time t_m to correspond with some of the boundaries of the segmentation, we do not have to recalculate the interval response estimates, but use the very same discretisation results. This obviously simplifies the calculation, but limits the possible responses to calculate. The segmented Midway estimate now gives

$$R_m = \sum_{j=1}^K \sum_{i=1}^m J_{j,i} \phi_{j,m-(i-1)}^+ / \Delta P_j \Delta t_i \quad (2.39)$$

where the adjoint function is meant in the reversed bookkeeping sense, with m segments in time, and K segments in other phase-space variables. This establishes the *first type* of time convolution. The major disadvantage of this formula comes from a possible increase in truncation errors.

The *second type* of convolution is an extension of the first one; it uses its technique but adds an extra possibility for obtaining more scores per sample from the same calculation. If we

define the source time dependency by a function $S(t)$ but calculate the flux by using a $\delta(t)$ function instead, obtaining $\phi_\delta(t)$ solution, the flux with the source S can be obtained by

$$\phi(P, t) = \int S(P, t-t') \phi_\delta(P, t') dt' \quad (2.40)$$

Then the integral on the i^{th} time interval for the segmentation technique and the j^{th} other phase space segment is given by

$$\begin{aligned} \phi_{j,i} &= \int_{\Delta P_j} \Pi_j(P) \int_t \phi(P, t) \Pi_i(t) dt dP = \\ &= \int_{\Delta P_j} \Pi_j(P) \int_{t'} \phi_\delta(P, t') \int_t S(P, t-t') \Pi_i(t) dt dt' dP \end{aligned} \quad (2.41)$$

The same would hold of course for the adjoint function, with the detector function in place of S . In terms of Monte Carlo, it means, that we might start the forward particles all at time 0, and whenever they reach the Midway surface, we check the result of the integral of the source function and the characteristic function of the segmentation, and add such a score. If the source is –let us say– takes the same simple shape that is to be described by a Π function, the integral simply means, that we need to check all the time segments after the time coordinate of a score and give score fragments to each that are affected, proportional to their overlap. This might be done with both calculations, as long as the integral of S and Π can be explicitly given. The final result is still obtained by using Eq.(2.39)

The *third type* of time convolution establishes a completely different type of response estimation. If we write the adjoint equivalent of Eq.(2.40), we obtain:

$$\phi^+(P, t) = \int D(P, t-t') \phi_\delta^+(P, t') dt' \quad (2.42)$$

We can now write the Midway response time interval in a different form:

$$R = \int_{A_m} dP \iint J_\delta(P, t') \phi_\delta^+(P, t'') \int D(P, t-t'') S(P, t-t') dt dt' dt'' \quad (2.43)$$

The integral on t can possibly be worked out beforehand, and the particles use that function as a representation of the linking function L (See Eq.(2.5)) regarding the time variables:

$$L(t', t'') = \int D(t-t'') S(t-t') dt \quad (2.44)$$

This linking function can be illustrated by having both detector and source functions behaving as a Π -function. If we take a forward particle scoring at t' and adjoint particle scoring at t'' , we make a score if the two respective source functions overlap, proportionally to the distance they both cover. Given the linking function is different, the calculation of such response needs major alteration to the coupling algorithm. The effort is certainly worth it, as such a coupling is exact, and multiple response calculation of the same coupling can be achieved without resorting to limitations in fitting segmentation structure to response structure. In other words, the coupling works for arbitrary adjoint and forward sources, as long as their integral as in

Eq.(2.44) can be explicitly given. One problem -though investigated thoroughly- has not been satisfactorily solved and would await future studies, is the variance calculation of such estimate. The basic problem lies in the amount of positive correlation that is not easy to account for. Each adjoint sample falling in a certain non-time phase-space segment, may be coupled to same, or different forward samples, therefore each estimate may or may not be strongly correlated with each other, and devising a simple variance calculation scheme, has not yet been successful so far.

We might mention here, that it is possible to use only one of the source functions to create an exact linking function in the time variable. This might be necessary, if the product of the adjoint and forward source functions are not easily integrated. The forward source function, for example, can be sampled in the Monte Carlo calculation, and upon hit at the Midway surface checked against a score from the adjoint calculation that started from a Dirac-delta source, its time variable convolved with the adjoint source function only at the coupling

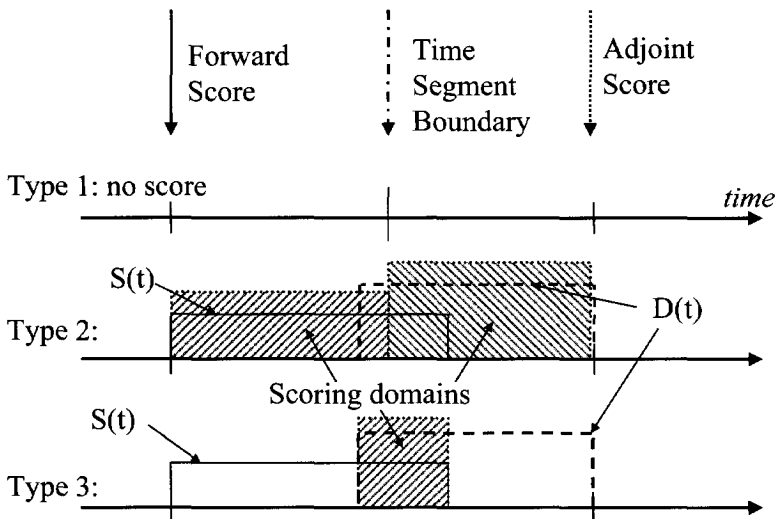


Figure 13. Sketch of contributions of particles, using different estimators for the coupling in the time variable

To compare the type 3 and type 1 time dependent response estimators, an equidistant set of time intervals has been set, coinciding for the first type with the time segmentation. Fig. 13 shows a sketch of this scenario for all three types. A single adjoint and a single forward particle fall in adjacent time segments. The type 1 time convolution will not yield any score, though for a different time reading the adjoint score will be shifted into the interval of the forward score, and for that reading they will fall into the same segment, and a contribution will be given. The type 2 convolution would distribute both particle's weight in both intervals, proportionally to their overlap with the segments. The type 3 convolution accounts for the overlapping of the detector and source functions.

As we accept the type 3 convolution to be exact, we see that the type 1 convolution uses less scores for the response, and as the type 3 method, it is a biased estimate. The type 2 convolution uses more scores, but the response is still biased.

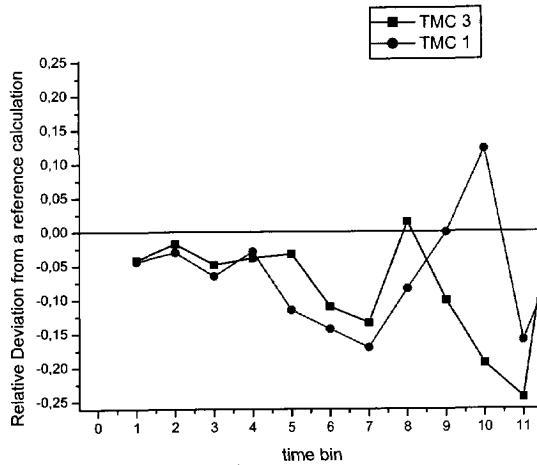


Figure 14. Comparison of time convolution type 1 (biased estimator) and type 3 (unbiased estimator)

The truncation errors are given in a relative scale as deviations from a reference autonomous forward calculation in Fig.14. Source and detector functions were both describable by a II function, with width of $50 \mu\text{s}$, the same width as for the response shift and time segment length. The calculations used the Generic tool model. The time convolution of type 3 has not been analysed to distinguish between (stochastic) undersampling and truncation error reduction. It is safe to say, that the first couple of bins clearly show a superiority of the 3rd type time convolution, and this method, with proper error calculation (if feasible) should form an integral part of the time dependent Midway method.

2.3.2 Midway Segmentation for a Borehole Logging Case

For a borehole logging application, the actual segmentation of the phase-space has been realised on a plane perpendicular to the axis of the cylindrical geometry. To close the midway surface, the model boundaries are used, and those are set far enough from the source to prevent giving contributions. The phase-space segmentation was always carried out on a structured mesh.

The spatial discretisation (see Fig. 15) comprised a number of radial divisions, forming annuli around the axis. The annuli were cut into sectors by lines through the intersection point of the plane with the geometry axis, and were placed to form an equidistant distribution of the sectors. The solid angle has been segmented according to an azimuthal angle and the surface cosine.

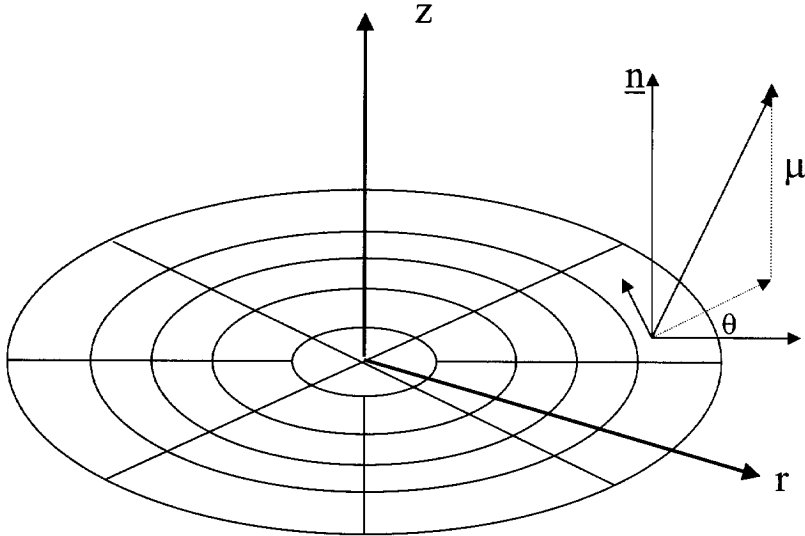


Figure 15. Sketch of the segmentation of the Midway surface for a borehole logging geometry

The discretisation in energy was linked to the energy divisions of the multigroup cross section library. Such limitations were posed by MCNP, as the continuous energy treatment of the adjoint calculation is not incorporated. Such a treatment for coupled neutron-photon calculations is feasible [16], though proper segmentation in energy would open a new chapter in the Midway coupling research. If the segmentation technique is to be used, the response estimate groups the particles by energy, but that is equivalent to a Monte Carlo calculation based on a multigroup cross section library. It has not been attempted, but would not contradict any previous assumptions, to still simulate the forward particles with continuous energy treatment. In the same time we know, that photon energy dependency is not a continuous function, and needs special treatment at characteristic energy lines.

The discretisation in time followed the requirements of the time convolution type 1.

2.3.3 Some Special Algorithms of the Coupling

When not using the modified, Midway version of MCNP, the scores (the crossings) occurring on the Midway surface by adjoint and forward simulations, were written out to a binary file, together with the information on which starting sample caused each. A coupling program *MICOU* has been written to couple the forward and adjoint scores to estimate the Midway response and its variance. Some details of the programming is described in this section.

Section 1.3.1 described how the Monte Carlo Midway estimate should be calculated. Using the simplified symbols of Section 2.2, and replacing the $x_{i,j}$ with the j^{th} Monte Carlo sample delivering $s_{i,j}$ score from the forward calculation falling in segment i , actually meaning $s_{i,j} = w_{i,j}/\Delta P_i$ and also renaming $s^+_{i,j} = w^+_{i,j}/\mu_j$ the Midway response estimate reads:

$$R_m = \sum_{i=1}^K \frac{1}{N} \sum_{j=1}^N (s_{i,j}) \frac{1}{N^+} \sum_{j=1}^{N^+} (s_{i,j}^+) \quad (2.45)$$

Though Eq.(2.45) suggests it, it is not necessary to store the segment-wise averages for both calculations. If the forward calculation results have been processed, and the estimates of J_i are given, it is enough to tally the contribution the adjoint particles would immediately make. If $s_{i,j}^+$ is indeed made in segment i , its c_j contribution is given by

$$c_j = s_{i,j}^+ \frac{1}{N} \sum_{j=1}^N s_{i,j} \quad (2.46)$$

and the response is an average of these contributions:

$$R_m = \frac{1}{N^+} \sum_{j=1}^{N^+} c_j \quad (2.47)$$

This form of calculating a response eliminates the need for two possibly huge arrays of segment-wise responses, as long as the variance calculation can also be achieved. In other words this form is the discrete representation of $J(P)$ as given by Eq.(2.8) as a response function h of Eq.(1.32).

If each adjoint score gives a contribution immediately when first processed, a way should be devised to have an immediate contribution to the variance estimate as well. To devise such a formula, we have to substitute the proper s and s^+ scores to first Eq.(1.37), then that to Eq.(2.27), correct it with the negative covariance calculated by Eq.(2.34) with the proper substitution of score symbols, and proceed to have a variance estimate according to Eq.(1.39) and Eq.(1.40). Having an $s_{i,j}^+$ adjoint sample again, the result of these substitutions would yield the following expression (omitting an $(NN^+)^{-2}$ factor):

$$D^2(R_m) = \sum_{j=1}^{N^+} (s_{i,j}^+)^2 \left[\left(\sum_{j=1}^N s_{i,j} \right)^2 \left(1 - \frac{1}{N} \right) + \sum_{j=1}^N s_{i,j}^2 \right] + \left(\sum_{j=1}^{N^+} s_{i,j}^+ \right)^2 \left[\left(\sum_{j=1}^N s_{i,j} \right)^2 \left(1 - \frac{1}{N^+} \right) \right] \quad (2.48)$$

, having added the second term of Expression (2.34), but temporarily neglected the first one. The first term of the RHS of Eq.(2.48) is easy to interpret in terms of a contribution to the variance per score. The second term however contains a summation of all the adjoint scores squared, and that needs waiting until all scores are made. We can rewrite the summation of all adjoint scores by

$$\left(\sum_{j=1}^N s_{i,j}^+ \right)^2 = \sum_{j=1}^N s_{i,j}^+ \left(s_{i,j}^+ + 2 \sum_{k=1}^{j-1} s_{i,k}^+ \right) \quad (2.49)$$

It is apparent, that for $N=2$ the equality holds. One way of proving it in general, we need to demonstrate that the difference of summing to N and to $N+1$ is the same for both sides:

$$\begin{aligned} \left(\sum_{j=1}^{N+1} s_{i,j}^+ \right)^2 - \left(\sum_{j=1}^N s_{i,j}^+ \right)^2 &= \left(\sum_{j=1}^N s_{i,j}^+ + s_{i,K+1}^+ \right)^2 - \left(\sum_{j=1}^N s_{i,j}^+ \right)^2 = 2s_{i,K+1}^+ \sum_{j=1}^N s_{i,j}^+ + (s_{i,N+1}^+)^2 \\ \sum_{j=1}^{N+1} s_{i,j}^+ \left(s_{i,j}^+ + 2 \sum_{k=1}^{j-1} s_{i,k}^+ \right) - \sum_{j=1}^N s_{i,j}^+ \left(s_{i,j}^+ + 2 \sum_{k=1}^{j-1} s_{i,k}^+ \right) &= s_{i,N+1}^+ 2 \sum_{j=1}^N s_{i,j}^+ + (s_{i,N+1}^+)^2 \text{ q.e.d.} \end{aligned} \quad (2.50)$$

The form of the summation of Eq.(2.49) would allow to have the new sample immediately contribute to the variance, but the sum of the previous adjoint scores should be stored for each segment. Now we can construct a v_j contribution to the variance by individual samples:

$$\begin{aligned} v_j &= (s_{i,j}^+)^2 \left[\left(\sum_{j=1}^N s_{i,j} \right)^2 \left(1 - \frac{1}{N} \right) + \sum_{j=1}^N s_{i,j}^2 \right] + \\ &+ s_{i,j}^+ \left(s_{i,j}^+ + 2 \sum_{k=1}^{j-1} s_{i,k}^+ \right) \left[\left(\sum_{j=1}^N s_{i,j} \right)^2 \left(1 - \frac{1}{N^+} \right) \right] \end{aligned} \quad (2.51)$$

The final relative variance would read, incorporating the first term of Eq.(2.34):

$$r^2(R_m) = \frac{\sum_{j=1}^{N^+} v_i}{R_m^2} - \left[\frac{1}{N} + \frac{1}{N^+} - \frac{1}{NN^+} \right] \quad (2.52)$$

This means, that out of the sum of score and the sum of score-squares for each calculation, the sum of squares for the adjoint (in general: the second) calculation can be spared, and each adjoint score might directly yield a contribution to the final response and the final variance, giving chance for tallying and monitoring the behaviour of the convergence of the variance and the score. We can also attempt to account for the correlation of scores coming from the same starting sample, i.e. coming from the same history. We need to distinguish two cases: when scores are made in the same segment and when they hit different ones. For the first case, it can be handled exactly by taking multiple scores of the same history in the same segment as belonging to the same score, therefore each history has to be processed and sorted into the segments before any contribution is tallied. This part has been calculated correctly for the old variance model as well. For the second term, if we have, let us say, two fragments $s_{i,j1}^+$ and $s_{k,j2}^+$ of the same history hitting different i and k segments, we might modify the first term of the variance contribution slightly by

$$\left(s_{i,j1}^+ \right)^2 \left(\sum_{j=1}^N s_{i,j} \right)^2 + \left(s_{k,j2}^+ \right)^2 \left(\sum_{j=1}^N s_{k,j} \right)^2 \rightarrow \left(s_{i,j1}^+ \sum_{j=1}^N s_{i,j} + s_{k,j2}^+ \sum_{j=1}^N s_{k,j} \right)^2 \quad (2.53)$$

This can only be applied for the second calculation, and this form is not exact, but might approximate the variance contribution better. The correlation of this type should be further investigated especially when handling neutron-photon cases. There, the scores are made by both neutrons and photons, and each contribution from the same history should be considered as belonging to the same statistical entity.

The same system of direct contributions can be set-up for the time convolution *type 3*. We have seen, that the linking function $L(t', t'')$ of Eq.(2.44) requires the t' time coordinate of the forward particle, and t'' time coordinate of the adjoint. If the forward scores have been processed, and the appropriate phase-space interval has been found in energy, angle and space, still the time coordinate and weight of the particle should be stored, and the average quantities and sums of the weights will not be calculated. Therefore, each score from the forward calculation would only be processed to a stage, when the segment index i , the time coordinate t' and the score s are given. For each particle this set of numbers has to be stored.

As it is unknown beforehand how much space is required for storage, and i ($i = 1 \dots K$) might run up to a million, a linked list system has been evoked. Using the FORTRAN 90 feature of dynamic pointer allocation and creation of embedded type variables, a null-sized array is created with K indices. An element of this array is created of a storage place for the t' , another one for s and finally a pointer is added that is pointing to itself. If a score arrives in segment i , such a set of data is stored. The second score will be put right before the first one to a new storage element, and the pointer this element will now point to the first entity. With such a system, only the required amount of memory will be allocated.

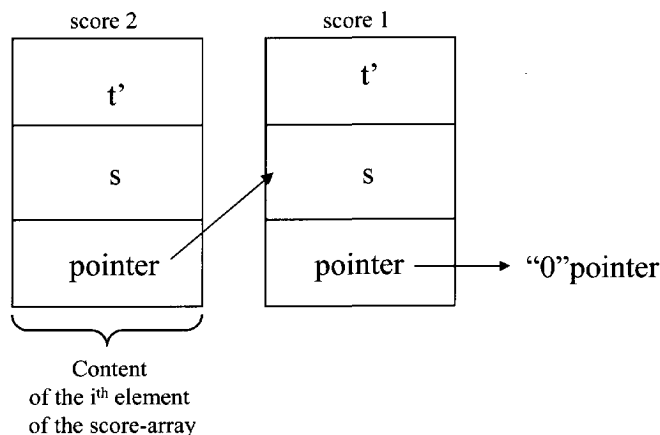


Figure 16. *Linked list system for time convolution type 3*

When processing of the forward calculation is finished, the adjoint scores will directly score. When the index i has been determined, the code checks if it contains an element. If so, it checks if the Linking function L would provide a score, and adds the contribution to the response, and its square to the variance. To be precise, the following step is to check if L belonging to other detector functions would yields a score, and account for the contributions. When done, the code checks if the pointer points to a next element. This feature gives the name for the construct: the elements are linked, and it is not known how long the list actually is, and we have to go through all the links until the list ends. It would be, of course, possible to store all the scores in such a system, and keep their coordinates registered. This means that an adjoint score could only give all its contributions by browsing through all of the forward scores with their coordinates, and that would make the coupling highly inefficient. However, if a better linking function for the other coordinates, or some different statistical estimator is

found, programming of such with even some million of particles is feasible. Note, that this special treatment of scores is only necessary, because the linking function L is not separable.

2.4 Summary

Given that the details of the Midway coupling with the segmentation technique have been investigated, it is time to assess the factors that contribute to the discrepancies. The error calculation has been corrected, and we concluded, that the quantity on which the reliability of the results had been assessed in the past, had been incorrectly given. We have seen that the source for deviations could also come from the truncation error, having the thereby defined linking function L not exact. In the time variable it is possible to give an exact form for the coupling, and that indeed increases the accuracy of the results, although the variance calculation is not straightforward.

We can now tune our calculation to deliver acceptable results in terms of accuracy, by changing the segmentation structure fitting the distribution of the flux or the adjoint function, and run enough samples to get a definitely converged estimate. Figure 17. shows the result of such an attempt, calculated on the Generic model set-up, as was given for the calculation for Fig. 14, but now changing intuitively the segmentation structure, and mainly, starting ten times more particles. The accuracy of the results clearly improved, and the deviations from the reference calculation decreased below a couple of percents (at least for the first few time responses). We can conclude that in general the simplest factor to tune is the number of starting particles: it should reach a level when the estimate becomes stable. If we are confident that a certain segmentation structure would give a good enough approximation for the Midway integral, with this choice we have already decided on the number of particles to run. For a regular Monte Carlo calculation we usually decide the number of particles by the level of statistical inaccuracy we want to reach. For a Midway calculation, the variance cannot be targeted directly, a given segmentation would already decide on the minimum amount of particles to reach the Midway surface. This difference becomes very much pronounced, when a sufficiently accurate estimate would require a large number of segments, and the number of particles this segmentation requires might give already an acceptable estimate in the real detector. This does not mean that the autonomous forward/adjoint calculation would also be more effective, when we reach the number of samples giving a stable Midway estimate, we might end up a very low variance, but the minimum computation time is already set by the method of doing the Midway coupling.

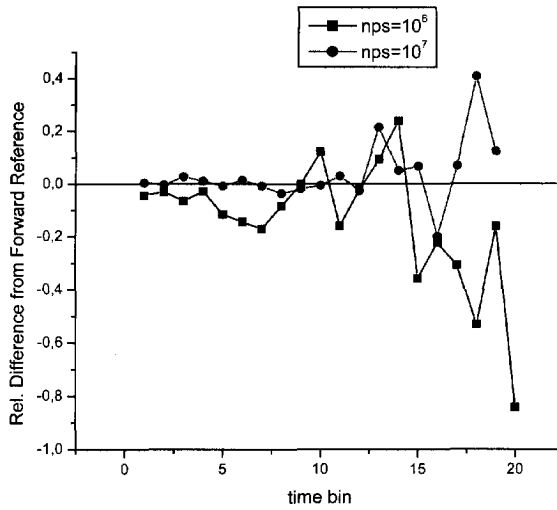


Figure 17. *Difference of the discrepancies when applying a higher number of samples. Seemingly systematic underprediction changes to a statistical variation around the mean, when increasing the number of samples*

Further investigations of the coupling method should target the use of more information out of the adjoint and forward scores, i.e. trying to increase the adjoint and forward scores that would together contribute to the response.



Chapter 3

MIDWAY COUPLING WITH LEGENDRE POLYNOMIALS

The Midway method is not the single Monte Carlo application that would benefit from representing a response as a function instead of a (set of) integral values. It is often desirable to describe distribution of quantities in the phase-space, let that be even just for a reason of better understanding a modelled system. This chapter describes methods for estimating (phase-) space dependency of quantities based on Monte Carlo samples, and how this technique might be used for estimating the Midway integral.

The idea of estimating functional expansion coefficients by Monte Carlo has been introduced by *Beers et al* [17] and later applied to shielding problems by *Noel et al* [18]. The usefulness of such a technique has not been debated since, though it has not become a standard feature of the major Monte Carlo codes. At the same time with the author of this thesis (and independently from each other), *Griesheimer et al.* [19] revived the method and extended it to use track length estimators [20], and later provided an investigation on the convergence properties of the functional expansion techniques. *Densmore* [7] applied the expansion technique to couple Monte Carlo calculations with deterministic adjoint solutions, in the first two angular moments. Here we will demonstrate the capabilities of Monte Carlo functional expansion and its application to the Midway coupling.

3.1 Function Expansion Using Monte Carlo samples

3.1.1 Function Expansion on Orthogonal Bases

Segmenting the phase space to obtain estimators for representing a function of the random variables of a Monte Carlo calculation is not the only, unique technique: neither in the way of obtaining the representation, nor in the possibility of estimating it by Monte Carlo. If we construct on the \mathbb{L}_2 space a set of orthogonal $g_i(x)$ functions, i.e.:

$$\int g_i(x) g_j(x) dx = \delta_{ij} \int g_i^2(x) dx \quad (3.1)$$

and this set is complete, meaning that there is no other function that is orthogonal to all of the basis functions, any arbitrary function can be expressed by their weighted sum:

$$f(x) = \sum_{j=0}^{\infty} d_j g_j(x) \quad (3.2)$$

The coefficients are obtained by multiplying both sides by $g_i(x)$ and integrating over x :

$$d_i = \frac{\int g_i(x) f(x) dx}{\int g_i^2(x) dx} \quad (3.3)$$

As a side note we mention, that if we choose now $f(x)$ to be the radiation current, and the basis functions as the Π_i characteristic functions of the segmentation technique, as given by Eq.(2.4), we obtain

$$d_i = \int \frac{J(P')}{\Delta P_i} \Pi_i(P') dP' \quad (3.4)$$

and by inserting it to Eq.(3.2) we obtain an equation similar to Eq.(2.8).

The Monte Carlo estimates of the d_i coefficients are obtained from the (x_i, w_i) samples of $f(x)$:

$$\hat{d}_j = \frac{1}{N} \sum_{i=1}^N w_i g_j(x_i) / \int g_j^2(x) dx \quad (3.5)$$

For the reconstruction of the function we need to truncate the function series, and estimate only the first K coefficients:

$$\hat{f}(x) = \sum_{j=1}^K \hat{d}_j g_j(x) \quad (3.6)$$

The nature of this truncation is not to be confused with the truncation of the segmentation technique. Given the set of Π_i functions is not a complete function base, the summation includes all of the elements of the function base, and the residual is given by the incompleteness of the set. The truncation regarding a complete set means neglecting some of the elements of the complete function base.

For the reconstruction of $f(x)$, we need K coefficients, and they might all be estimated by the same run. According to the formula of Eq.(3.5), the same (x_i, w_i) sample could give contributions to the estimates of each d_j , yielding a heavily correlated series of coefficient estimates. The variance of the recreated function would be given, therefore, at x as:

$$\begin{aligned}
 D^2(\hat{f}(x)) &= D^2\left(\sum_{i=0}^K \hat{d}_i g_i(x)\right) = \\
 &= \sum_{i=0}^K D^2(\hat{d}_i) g_i(x) + \sum_{i=0}^K \sum_{\substack{j=0 \\ j \neq i}}^K \text{Cov}(\hat{d}_i, \hat{d}_j) g_i(x) g_j(x)
 \end{aligned} \tag{3.7}$$

In case of the segmentation technique, the covariance term vanishes, given the Π_i 's are pairwise disjoint. This is not the case for a general set of basis functions, moreover such a covariance term might not be insignificant, especially if we consider that now d_i can also be negative. If we have x_i independent position samples, and we keep in mind that now the statistical quantities are generated from the independent position samples, the variance estimate of the recreated function at y will give:

$$\widehat{D}^2(\hat{f}(x)) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=0}^K \sum_{k=0}^K \frac{w_i^2 g_j(x_i) g_k(x_i)}{\int g_j^2(x) dx \int g_k^2(x) dx} - \frac{1}{N^3} \left(\sum_{i=1}^N \sum_{j=0}^K \frac{w_i g_j(x_i)}{\int g_j^2(x) dx} \right)^2 \tag{3.8}$$

If we calculate the relative error, we obtain

$$r^2(\hat{f}(x)) = \frac{\sum_{i=1}^N \sum_{j=0}^K \sum_{k=0}^K \frac{w_i^2 g_j(x_i) g_k(x_i)}{\int g_j^2(x) dx \int g_k^2(x) dx}}{\left(\sum_{i=1}^N \sum_{j=0}^K \frac{w_i g_j(x_i)}{\int g_j^2(x) dx} \right)^2} - \frac{1}{N} \tag{3.9}$$

The necessity of evaluating Eq. (3.9) for an error estimate, means, that not only the individual variances of the estimates of d_j are informative on the quality of the reconstructed function. Moreover that is the only term that gives always a positive contribution to the variance of the reconstructed function at x

The choice of the basis function set must be done carefully and meaningfully. *Beers et al* [17] claim, that for a yet unknown $f(x)$ the arbitrariness of the choice of the basis function limits the general applicability of the method, mentioning the example of the unfortunate situation of having $f(x) = g_{K+1}(x)$, meaning that the truncation is done before the first important term. Such a blindfolded selection of the function base would indeed undermine the method, and therefore (as it is done generally in all branches of physics) such a set of functions should be selected, that is meaningful for the interpretation, and by no way should be such a function base selected, when the effect possible of major contributions from the $K+1$, $K+2$, ... terms are not well understood, and with a reason disregarded. It seems, however apriori (before starting any Monte Carlo calculation) more informative to assess the possible effect of truncating a function series at K , than apriori seeing the difference that creating K instead of $K+1$ segments would make. It should also be noted that in particle transport the functions to reconstruct are usually strictly positive, and could only coincide with the first element of the function series.

Function expansion is a common tool for solving differential equations. The Monte Carlo estimation of the expansion coefficients is also useful when comparing analytical solutions to simulations.

3.1.2 Function Expansion using Legendre Polynomials

One meaningful choice for the basis functions could be the Legendre polynomials, as one of the simplest of the Gegenbauer polynomials. Though the base functions are not given explicitly, it does not pose a limitation to their application for this purpose, as a recursion formula available for the successive generation of the functions. These polynomial representations are usually valid on a preset interval (a,b) , and sometimes require the definition of a weight function $\rho(x)$ for the orthogonality:

$$\int_a^b g_i(x)g_j(x)\rho(x)dx = \delta_{ij} \int_a^b g_i^2(x)\rho(x)dx \quad (3.10)$$

instead of Eq.(3.1). The recursion formula is then given [21] by

$$g_n(x) = (A_n x + B_n)g_{n-1}(x) - C_n g_{n-2}(x) \quad (3.11)$$

where A_n , B_n and C_n are constants and characteristic to the polynomial function base. If a Monte Carlo sample scored at x_i , the contribution to the coefficients can be calculated using previous estimates by

$$g_n(x_i) = (A_n x_i + B_n)g_{n-1}(x_i) - C_n g_{n-2}(x_i) \quad (3.12)$$

to fill in Eq.(3.5). The recurrence formula is not always numerically stable, therefore not always allowing the iterative calculation the base function values at a certain position. With the choice of the Legendre polynomials, we get the recursion formula of

$$\begin{aligned} (n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) &= 0; \\ P_0(x) &= 1; \\ P_1(x) &= x; \end{aligned} \quad (3.13)$$

Eq.(3.13) is numerically stable, therefore it can be used for successive approximation of contributions to the different coefficients. The orthogonality of the Legendre polynomials is given on the interval $(-1,1)$. The norm of the polynomials is given by:

$$\int_{-1}^1 P_l^2(x)dx = \frac{2}{2l+1} \quad (3.14)$$

To apply the above formulae for Monte Carlo, we need to transform the coordinates of the function to the $(-1,1)$ interval. If the function $f(x)$ was given on (a,b) , the coefficient of the expansion is given by

$$d_i = \frac{2i+1}{2} \int_{-1}^1 f\left(x \frac{b-a}{2} + \frac{b+a}{2}\right) P_i(x) \frac{b-a}{2} dx \quad (3.15)$$

It is also possible to transform the Legendre polynomials to (a,b) , by

$$d_i = \frac{2i+1}{2} \int_a^b f(x) P_i\left(\left(x - \frac{b+a}{2}\right) \frac{2}{b-a}\right) \frac{2}{b-a} dx \quad (3.16)$$

If we happen to calculate the first Legendre coefficient ($P_1(x)=x$), and reconstruct the linear term of the function we get

$$d_1 \left(x - \frac{b+a}{2}\right) \frac{2}{b-a} = \frac{3}{2} \left(\frac{2}{b-a}\right)^3 \left(x - \frac{b+a}{2}\right) \int_a^b f(x) \left(x - \frac{b+a}{2}\right) dx \quad (3.17)$$

It follows therefore, that

$$d_1 \frac{2}{3} \left(\frac{b-a}{2}\right)^3 = \int_a^b f(x) \left(x - \frac{b+a}{2}\right) dx \quad (3.18)$$

We see, that Eq.(3.18) is equivalent to Eq.(2.15), proving that the choice of the scoring function of Eq.(2.14) calculates indeed the first Legendre term for an arbitrary $f(x)$. Although it is not practical (and mentioning it is just for the sake of curiosity), we might select the transformed Legendre functions with the coordinate system origin set at the middle of (a,b) , and set the orthogonal functions on (a,b) instead of $(-1,1)$. The norm will now differ from Eq.(3.14), for the first Legendre term it is given by

$$\int_a^b P_1^2 \left(x - \frac{b+a}{2}\right) dx = \frac{2}{3} \left(\frac{b-a}{2}\right)^3 \quad (3.19)$$

in agreement with Eq.(3.18). If $f(x)$ is given as a pdf, the meaning of the RHS of Eq.(3.18) is the expected value of the deviation of x from the middle point of the interval, and as we have seen, in Chapter 2., it might be used as a direct measure for the first order inaccuracy of the segmented Midway coupling. The zeroth coefficient obviously gives the integral average of $f(x)$.

3.1.3 Numerical examples

The equivalence of the formulae Eq.(3.18) and Eq.(2.15) means that the convergence properties of the estimator of d_1 are the same as for the slope estimator on Fig. 9. Reformulating our conclusions in Legendre-terms, the amount of scores needed for a stable estimation of d_1 is about ten times higher than for estimating d_0 . We have also seen, that when reconstructing $f(x)$, the relative variance is not as informative: we might expect a successful function expansion even when having coefficients with high relative errors.

To assess the quality of Monte Carlo Legendre expansions, we compared it to a function expansion by segmentation using a simple model geometry. The model volume was z -axis, 40 cm radius cylinder, filled with SiO_2 ; with 20cm radius 5cm thick cylindrical 14MeV neutron source. We calculated the flux on a surface, which was located 5cm below the bottom of the source cylinder, perpendicular to the z -axis. The bottom of the geometry was set 10cm lower than this “mapping” surface. The scores were calculated by MCNP, and printed out to a PTRAC file, therefore both the Legendre expansion and the segmentation representation made use of the same samples.

For the segmented flux estimator, the mapping surface was divided into 80 sectors, and the flux values were given as integral averages. As the system is rotationally symmetric, only the radial dependence is shown on Fig.18. Also on Fig.18, the function expansion using Legendre polynomials is shown, using 20 and 80 coefficients.

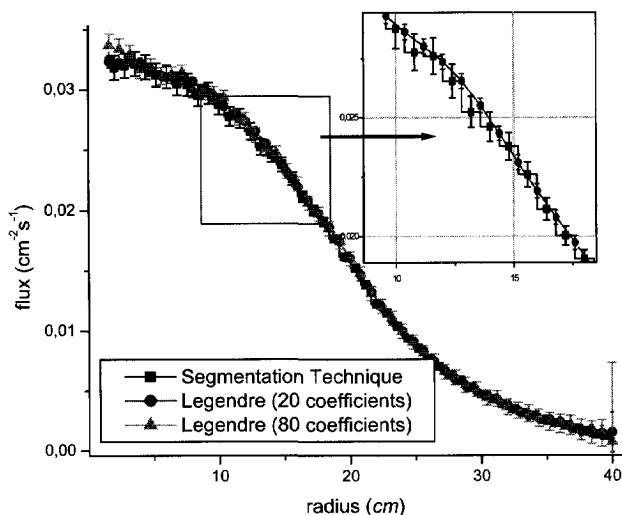


Figure 18. Flux as a function of the radius calculated by dividing the phase-space into small blocks, and applying Legendre polynomial expansion

The first conclusion is to draw, that even 20 Legendre coefficients approximate the flux accurately (see also the magnification on Fig.17). When applying only 20 coefficients, also statistical error is lower than for the segmentation representation. To make the comparison fairer we should compare the variances when using the same number of segments as coefficients. In that case, using, the 80th order Legendre expansion, the variance exceeds of that of the segmentation technique. A practically useful feature of the Legendre expansion, that the calculation of 80 coefficients allows the post-calculation filtering of the coefficients, a 20th order expansion is a partial result of the 80th order calculation; if we find that we have chosen too many coefficients, we already have the means of calculating the 20th order

expansion, without having to recalculate the expansion coefficients. Such a process is obviously a problem for the segmentation technique, though it is possible merge some of the segments, and average their scores, the segment boundaries cannot be modified any longer. This feature of the Legendre reconstruction solves the problem of selecting more segments than we have information on. In the opposite case, when having more information we could still extract, having kept the scores, the coefficients and their covariance matrix, we can simply calculate the contribution of additional terms. Also, by such a process we obtain information on the importance of a certain Legendre term to the reconstructed function, giving an indication of the accuracy of the process. The cost for these advantages is to be paid in computing time and computer memory. An apparent weakness of the Legendre expansion is, the convergence at the edges of the intervals, where the convergence is not point-wise. This is much more pronounced when using 80 coefficients, moreover the function values show an artefact, that most likely not correspond to the actual flux.

It might be possible to devise a set of segments by an unstructured mesh that describes the flux using just 20 segments as good as using 20 coefficients. That might be very well the case, when the targeted function is not continuous, or when the function is not smooth. For the purpose of comparing the methods for less smooth functions, another model geometry has been chosen, using the same model volume and source, but filling the model with water instead of SiO_2 , and placing three cylinders in the model each with radius of 10cm oriented towards the z axis, one filled with sand, one filled with an absorber and one was left void. The sand cylinder was modelled with $2.3\text{g}/\text{cm}^3$ density, while the absorber as an $80\text{g}/\text{cm}^3$ lead. The scores were calculated by MCNP, using 10^6 starting samples.

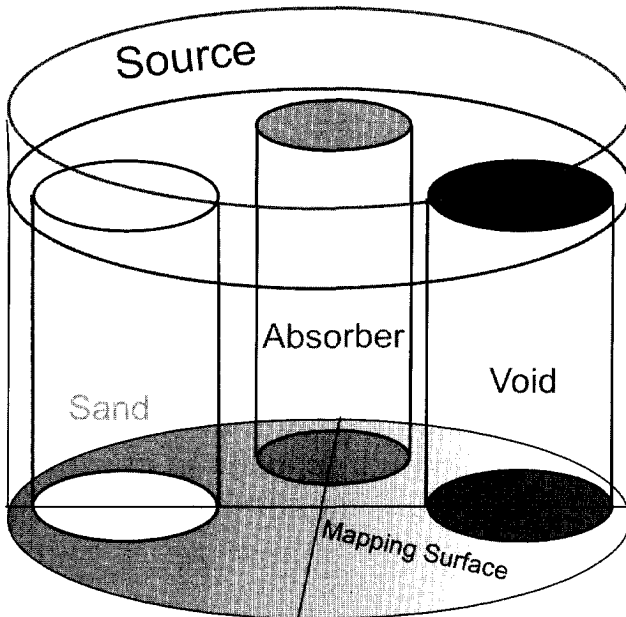


Figure 19. Model geometry for 2D flux mapping

For the Legendre expansions, we have now two spatial variables to account for, and such an expansion can be devised for the x and y variables as

$$d_{i,j} = \frac{\iint P_i(x)P_j(y)f(x,y)dydx}{\int P_i^2(x)dx \int P_j^2(y)dy} \quad (3.20)$$

and

$$\hat{f}(x,y) = \sum_{j=0}^L \sum_{i=0}^K d_{i,j} P_i(x) P_j(y) \quad (3.21)$$

We note here, that such an expansion does not mean that the function is handled as a separable one in x and y . For the variables the radius, and the spatial azimuthal angle has been chosen, therefore both the segmentation and the Legendre expansions were devised according to the same variables. When transforming the regular x and y dimensions to an (r,θ) cylindrical system, we have to take care of the Jacobian determinant (here r for the radius) when calculating the integrals of Eq.(3.20).

About the spatial distribution of the flux, we can predict that within the absorber cylinder the flux should be very low, surrounded by a very steep depression almost an abrupt change, and at the boundaries the derivatives should be discontinuous. The void cylinder should have the same effect but the other way around, while it is hard to foresee whether the change to higher flux values around the boundaries will be limited to a small region. . It is also hard to foresee, whether the sand cylinder will have a noticeable effect. Outside the three cylinders the flux should be decreasing with the radius.

Fig.20 and Fig.21 show the flux as a function of the spatial coordinates x and y , for the segmentation and the Legendre expansions, respectively. The segment-wise representation of the flux is obviously compromising the shape of the absorber intersection we should clearly see a circle; and segments that are half covering the perimeter blur the edges. The centre of the geometry comprises very small segments, resulting in yet non-converged, unrepresentative estimates. In the void cylinder the image shows increased flux, though it is hard to see how large domain is affected. The sand cylinder is not giving a noticeable effect.

This model has been created to give a very hard task to the Legendre method. Though it seems, that the Legendre expansion yields the same positive results as the segmentation representation, somewhat more in accordance with the expectations. Introduced artefacts are somewhat more disturbing now: low r values below 5-10 centimetres are very doubtful, and the at spatial azimuth starting and ending coordinates the function values do not seem to match. Here it is impossible to assess whether these artefacts make it impossible to discover the effect of the sand cylinder, or it is indeed negligible. All in all, we can say, that the Legendre method delivers good results even for such an unsuitable problem, at least in a sense of visualising the flux. Numerically the artefacts are important enough to raise doubts on applying the Legendre expansion for such an unsuitable problem. Meanwhile *Griesheimer et al.* [22] demonstrated, that the convergence of the Legendre expansion technique may be faster even for only piecewise smooth functions.

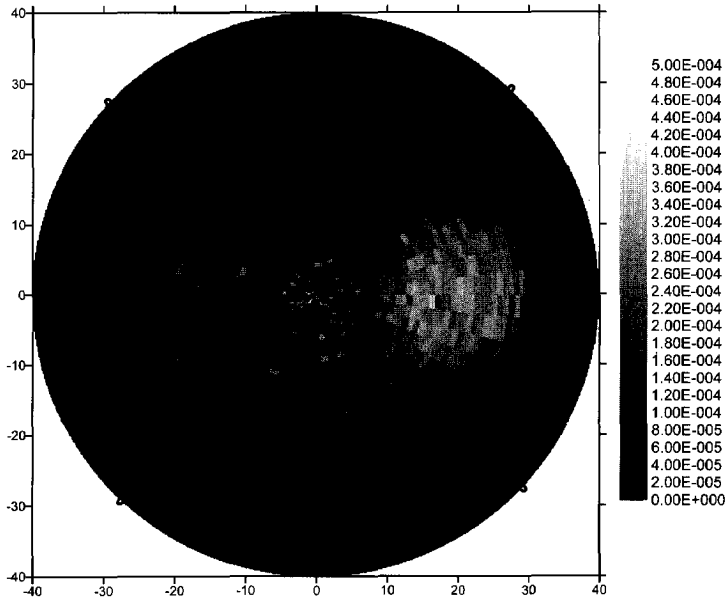


Figure 20. *Segmentation expansion of the flux (dimensions in centimetres, tone scale in $\text{cm}^{-2}\text{s}^{-1}$)*

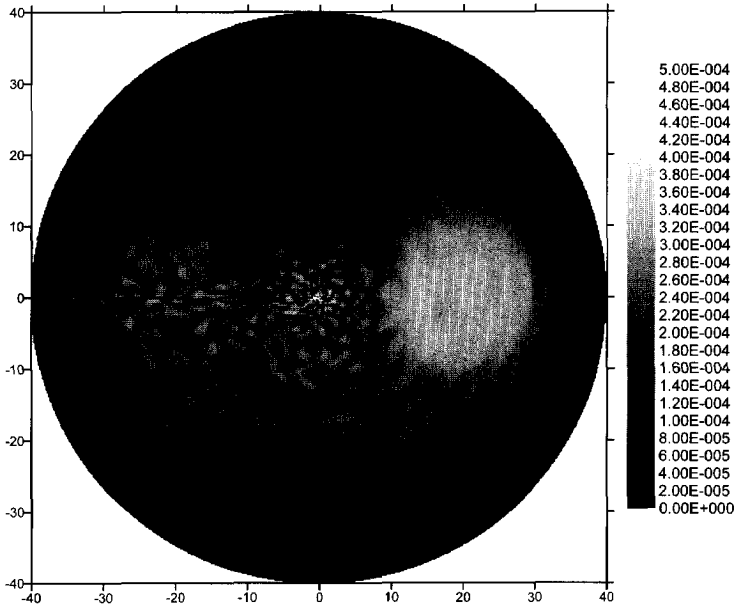


Figure 21. *Legendre expansion of the flux (dimensions in centimetres, tone scale in $\text{cm}^{-2}\text{s}^{-1}$)*

Apart from visualising some quantities of a Monte Carlo calculation, the function expansion technique gives all the advantages of treating a function as sum of analytical functions. Having the coefficients calculated, we can give the function shape on any domain or line; we can also give the maximum and minimum points even locally, together with the statistical confidence intervals of the domain; might calculate the derivative of the flux or the primitive function, etc.; and these are obviously problematic for the segmented representation.

3.2 Midway Coupling with Function Expansion coefficients

3.2.1 Midway Coupling on Complete Orthogonal Bases

The basic aim of recreating the function is to calculate the Midway response in an alternative form. If J (or ϕ^+) of Eq.(2.1) is represented by an orthogonal expansion truncated at K , i.e.

$$\hat{J}(P) = \sum_{i=0}^K d_i g_i(P) \quad (3.22)$$

the Midway response estimate is given using orthogonal function expansion by

$$R = \int \hat{J}(P) \phi^+(P) dP \quad (3.23)$$

where the approximation of J is now given at every P , and the Monte Carlo interpretation is

$$R_m = \frac{1}{N^+} \sum_{i=1}^{N^+} s_i^+ \hat{J}(P_i^+) \quad (3.24)$$

in line with the notations of Eq.(2.45), regarding the symbol s_i^+ . The only main difference, or odd behaviour of the explicit scoring to a general scoring function $h(x)$, that this one is associated with a variance at every P_i^+ .

If we insert Eq.(3.3) into Eq.(3.22), and then to Eq.(3.23) we obtain:

$$R_m = \sum_{i=0}^K \frac{\int \phi^+(P) g_i(P) dP \int J(P') g_i(P') dP'}{\int g_i^2(P) dP} \quad (3.25)$$

We define the adjoint expansion coefficients by

$$d_i^+ = \int \phi^+(P) g_i(P) dx \quad (3.26)$$

to obtain

$$R_m = \sum_{i=0}^K d_i d_i^+ \quad (3.27)$$

The Monte Carlo interpretation of the integral for d_i^+ is obvious. The linking function L would then be given as

$$L(P, P') = \frac{\sum_{i=0}^K g_i(P) g_i(P')}{\int g_i^2(P) dP} \quad (3.28)$$

We can expect two main advantages of the function expansion coupling over the segmentation technique. First, an increase in accuracy (smaller truncation error), at least when the flux and the adjoint function are smooth enough. Secondly, as each score gives contributions to each of the coefficients, we might achieve an increase in the amount of information that is actually used for estimating the Midway response; and with this, we might obtain a better statistical estimate. The obvious drawbacks are the increase in computation time and in computer memory requirements.

3.2.2 Convergence of the Midway coupling using Legendre function base

To apply the function expansion technique to the Midway coupling, we need to define an orthogonal set of functions on the whole phase-space. Using Legendre polynomials, the extension for multiple dimensions is done in the same way as for Eq.(3.20), and analogously for the adjoint coefficients. We will proceed to show results for two-dimensional Legendre expansions, as extending to further dimensions is straightforward.

The convergence [34] of the coefficients of the Legendre Midway estimate is algebraic, given by

$$|d_i d_i^+| = O\left(\frac{2i+1}{2} \left(\frac{1}{i}\right)^{(\kappa+\kappa^+)}\right) \quad (3.29)$$

here κ and κ^+ stand for a measure of the smoothness [34] of J and ϕ^+ , and equals the number of derivatives that are square integrable. If we compare Eq.(2.12) with Eq. (2.15) or with Eq. (3.18) we can give the first-order estimate of the truncation error for the segmentation technique, having K equally spaced segments:

$$\Delta R_m = \sum_{i=1}^K O(\Delta P_i^3) = O\left(\frac{1}{K^2}\right) \quad (3.30)$$

κ and κ^+ are characteristic of J and ϕ^+ , therefore it is not possible to assess which method would converge faster. For analytic functions, the Legendre coupling technique would converge exponentially with the number of coefficients, while the segmentation technique improves in the order of $1/K^2$. Fig. 21 shows such a case. For Fig.21, the integral of the following function has been estimated using Monte Carlo samples:

$$\phi^+(x, y) = \frac{1}{45} \frac{x^2}{y} \text{ and } J(x, y) = \frac{y}{x^2}; \quad x \in [1, 6] \quad y \in [1, 10] \quad (3.31)$$

For these analytic functions the Legendre method showed an extremely fast convergence to the true results, while the segmented approximation converged as $1/K^4$ as now we varied the segments simultaneously in both dimensions (fitting a general hyperbole the power was estimated as 4.44).

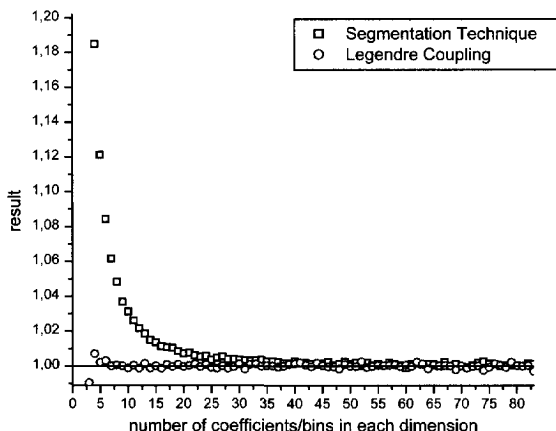


Figure 22. *Convergence of the segmented and Legendre Monte Carlo estimates of the coupling integral. The Legendre coupling converges considerably faster than the segmentation technique*

If we compare the amount of computation time we have to invest in the estimates, in two dimensions, a structured segmented mesh would require a searching procedure to find the proper mesh for a score, and a simple mathematical operation of calculating the bin-wise function results. For the continuous base function the contribution to each coefficient would require K times more mathematical operations per score. The amount of calculation would increase with the total number of coefficients used. Finding a segment in a structured mesh requires more operation than K operations, and can be done in $K^{1/2}$ on average per dimension. This obviously results (in realistic cases) in a huge difference, and the advantage that the function expansion coupling would offer, has to worth the extra time.

Griesheimer *et al.* [22] provided an analytical and numerical study on the convergence of a selected continuous, piecewise smooth function having an algebraic index (κ) 2.5 of the expansion based on Monte Carlo samples. The selected function had two points where the derivatives were discontinuous, and showed that the convergence of the Legendre expansion was faster than of the segmentation technique. They have also shown for arbitrary functions, that when the amount of coefficients or segments became undersampled, i.e. when we select too many coefficients/bins with regard to the available scores, the increase in statistical spread followed the same pattern for both methods, namely $O(K/N)$. As it will from the derivation of the next section, the variance of the Midway coupling grows in the order of $K/(NN^+)$ for the

segmentation technique. Though rigorously not studied, we can assume in the analogy of *Griesheimer et al.* [22], that the same order of variance increase can be observed for the Legendre expansion technique.

3.2.3 General estimator for the error of the Midway response estimate

The analysis of the variance estimate of the Legendre coupling may follow the same scheme as in Section 2.2. However, we will choose another path: having a complete function base for the coupling allows a simpler way of giving an overview of the variance calculation. Let us choose an integral of

$$I = \int f(x) f^+(x) dx \quad (3.32)$$

Now using the Legendre polynomials (this choice of the basis function is not essential to all the further steps in the derivation), we can write this integral in the form of

$$I = \sum_{i=1}^{\infty} \frac{2i+1}{2} d_i d_i^+ \quad (3.33)$$

where now the d_i coefficients (and analogously the d_i^+ coefficients) are given:

$$d_i = \int f(x) P_i(x) dx \quad (3.34)$$

differing only from the previous definitions by not including the normalisation factor for the polynomials. The coefficients are estimated from two independent Monte Carlo simulations, by sampling the pdf's $\wp(x)$ and $\wp^+(x)$. The estimator for d_i reads:

$$\hat{d}_i = \int \wp(x) \left(\frac{f(x)}{\wp(x)} \right) P_i(x) dx = E \left\{ \left(\frac{f(x)}{\wp(x)} \right) P_i(x) \right\} = E(e_i) \quad (3.35)$$

For the estimate of the coefficient we draw x_j samples from $\wp(x)$, and average the scores made on the response function (now P_i), each weighted by the term in brackets. In a non-analogue simulation the pdf will not be explicitly known, while for an analogue simulation the weight is reduced to 1 meaning that the pdf is actually $f(x)$. The symbol e_i Equation (3.35) was only introduced for shortening the formulas.

The variance of I now reads:

$$\begin{aligned} D^2(I) &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} \text{Cov}(e_i e_j^+, e_j e_i^+) = \\ &= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} E(e_i e_j) E(e_i e_j^+) - \left(\sum_{i=0}^{\infty} E(e_i) E(e_i^+) \right)^2 \end{aligned} \quad (3.36)$$

Now we did not make a separation of the covariances to diagonal and non-diagonal elements, as now the non-diagonal elements cannot be further simplified, or neglected. Similarly to the derivation with the relative variances in Eq.(2.26), we can write the covariance terms by

$$D^2(I) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} \begin{pmatrix} +\text{Cov}(e_i, e_j) E(e_i^+) E(e_j^+) + \\ +\text{Cov}(e_i^+, e_j^+) E(e_i) E(e_j) + \\ +\text{Cov}(e_i, e_j) \text{Cov}(e_i^+, e_j^+) \end{pmatrix} \quad (3.37)$$

The first term of the summation yields:

$$\begin{aligned} & \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} \text{Cov}(e_i, e_j) E(e_i^+) E(e_j^+) = \\ & = \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} [E(e_i e_j) - E(e_i) E(e_j)] E(e_i^+) E(e_j^+) \end{aligned} \quad (3.38)$$

We can evaluate exactly the RHS. Taking its first term gives:

$$\begin{aligned} & \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} E(e_i e_j) E(e_i^+) E(e_j^+) = \\ & = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} \left(\int \wp(x) \left(\frac{f(x)}{\wp(x)} \right)^2 P_i(x) P_j(x) dx \right) E(e_i^+) E(e_j^+) = \\ & = \int \wp(x) \sum_{i=0}^{\infty} \frac{2i+1}{2} \left(\frac{f(x)}{\wp(x)} \right) P_i(x) E(d_i^+) \sum_{j=0}^{\infty} \frac{2j+1}{2} \left(\frac{f(x)}{\wp(x)} \right) P_j(x) E(d_j^+) dx = \\ & = \int \wp(x) \left(\left(\frac{f(x)}{\wp(x)} \right) f^+(x) \right)^2 dx \end{aligned} \quad (3.39)$$

The second term yields:

$$\begin{aligned} & \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} E(e_i) E(e_j) E(e_i^+) E(e_j^+) = \\ & = \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} \int \wp(x) \left(\frac{f(x)}{\wp(x)} \right) P_i(x) dx \int \wp(x) \left(\frac{f(x)}{\wp(x)} \right) P_j(x) dx \times \\ & \times E(e_i^+) E(e_j^+) = \left(\int \wp(x) \left(\frac{f(x)}{\wp(x)} \right) f^+(x) dx \right)^2 \end{aligned} \quad (3.40)$$

Summarising Eq.(3.39) and Eq.(3.40), we get an expression in a form of a variance, namely the variance of the function of a random variable, and this is now taken with regard to $\wp(x)$.

$$\begin{aligned} & \int \wp(x) \left(\frac{f(x)}{\wp(x)} f^+(x) \right)^2 dx - \left(\int \wp(x) \frac{f(x)}{\wp(x)} f^+(x) dx \right)^2 = \\ & = E_{\wp} \left(\frac{f(x)}{\wp(x)} f^+(x) \right)^2 - E_{\wp}^2 \left(\frac{f(x)}{\wp(x)} f^+(x) \right) = D_{\wp}^2 \left(\frac{f(x)}{\wp(x)} f^+(x) \right) \end{aligned} \quad (3.41)$$

Following the same steps, with the other "single-covariance" term, we obtain for the variance of I :

$$\begin{aligned} D^2(I) &= D_{\wp}^2 \left(\frac{f(x)}{\wp(x)} f^+(x) \right) + D_{\wp^+}^2 \left(\frac{f^+(x)}{\wp^+(x)} f(x) \right) + \\ &+ D_{\wp \wp^+}^2 \left(\sum_i \left(\frac{f^+(x^+)}{\wp^+(x^+)} P_i(x^+) - c_i^+ \right) \left(\frac{f(x)}{\wp(x)} P_i(x) - c_i \right) \right) \end{aligned} \quad (3.42)$$

The third term of the RHS of Eq.(3.42) cannot be further simplified (to the author's knowledge), and remains the only term, that resembles the coupling technique. The lower indices of the variance of this term are meant to express, that the last variance term should be calculated according to the joint probability density function of the two random processes.

If we have N and N^+ samples from both pdf's, the population means will give:

$$\begin{aligned} D^2(I) &= \frac{1}{N} D_{\wp}^2 \left(\frac{f(x)}{\wp(x)} f^+(x) \right) + \frac{1}{N^+} D_{\wp^+}^2 \left(\frac{f^+(x)}{\wp^+(x)} f(x) \right) + \\ &+ \frac{1}{N} \frac{1}{N^+} D_{\wp \wp^+}^2 \left(\sum_i \left(\frac{f^+(x^+)}{\wp^+(x^+)} P_i(x^+) - c_i^+ \right) \left(\frac{f(x)}{\wp(x)} P_i(x) - c_i \right) \right) \end{aligned} \quad (3.43)$$

The third term of the RHS, we now see, will decrease very fast with the number of starting samples, and we will neglect it from on. This term diverges when increasing the number of coefficients, and is responsible for the increase of variance when too many coefficients are selected. For the segmentation estimate, the 3rd term shows a linear increase of the variance, if the segment divisions are equally spaced, the unstable part equals IK/NN^+ .

This formula for the variance, gives a recipe for estimating the asymptotic (in terms of truncation error) variance based on Monte Carlo samples. The actual form of the variance estimator depends on the exact linking function L , that establishes a biased, but feasible connection between the two independent random processes. Changing Eq.(3.32) to a form of Eq.(2.22) is changing only the notations.

3.2.4 Variance estimator for the Legendre Midway coupling

Estimating I is relatively simple using Legendre polynomials. A more difficult task is to give estimates to the second moments of the variance calculation (as in Eq.(3.39)), if we intend to run the adjoint and the forward calculations in succession. The cross-variance term (the third term in Eq.(3.43)) we will neglect here, as already devising an unbiased estimator for the variance, as will be demonstrated here, is not practical for a real Midway calculation.

The first term we treat is the end-result of Eq.(3.39), having assumed that the adjoint calculation has been done first, and we obtained the (x_i, w_i) samples of the second (now forward) calculation. The Monte Carlo estimator read:

$$\int \varphi(x) \left(\frac{f(x)}{\varphi(x)} f^+(x) \right)^2 dx \approx \frac{1}{N} \sum_{k=1}^N w_k^2 \sum_{i=0}^K \sum_{j=0}^K \frac{2i+1}{2} \frac{2j+1}{2} P_i(x_k) P_j(x_k) \hat{d}_i^+ \hat{d}_j^+ \quad (3.44)$$

The adjoint coefficients are obviously stored for the computation of I ; therefore this term does not require allocating additional computer memory. Calculation-wise, however, this formula needs $K \times K$ calculations for each score of the second run. The more problematic term is the adjoint counterpart of the result of Eq.(3.39):

$$\int \varphi^+(x) \left(\frac{f^+(x)}{\varphi^+(x)} f(x) \right)^2 dx \approx \frac{1}{N^+} \sum_{k=1}^{N^+} (w_k^+)^2 \sum_{i=0}^K \sum_{j=0}^K \frac{2i+1}{2} \frac{2j+1}{2} P_i(x_k^+) P_j(x_k^+) \hat{d}_i \hat{d}_j \quad (3.45)$$

The least memory consuming way of calculating this estimate is to obtain first the estimators for the d_i coefficients, and then process the adjoint scores. Altogether it would require handling the forward and the adjoint scores twice: first calculating the d_i^+ coefficients and afterwards processing the forward scores according to Eq.(3.44), in the meanwhile obtaining the estimates for the d_i coefficients and the other unaccounted terms of Eq.(3.40). Then we would need to reprocess the adjoint scores to calculate Eq.(3.45). The inconvenience it causes is not only the increased computation time, but also it basically makes it impossible to incorporate such a process in a regular flow of a Monte Carlo calculation. We can solve this problem by storing the C_{ij} elements of the first term of the covariance matrix of the adjoint calculation, calculated as

$$C_{i,j} = \frac{1}{N^+} \sum_{k=1}^{N^+} (w_k^+)^2 \frac{2i+1}{2} \frac{2j+1}{2} P_i(x_k^+) P_j(x_k^+) \quad (3.46)$$

The estimator of Eq.(3.45) reads:

$$\int \varphi^+(x) \left(\frac{f(x)}{\varphi^+(x)} f(x) \right)^2 dx \approx \sum_{i=0}^K \sum_{j=0}^K \frac{1}{N} \sum_{k=1}^N w_k P_i(x_k) \frac{1}{N} \sum_{k=1}^N w_k P_j(x_k) \hat{C}_{i,j} \quad (3.47)$$

The way of programming this estimator to be calculable in the second (now forward) score processing, we need to apply the formula of Eq.(2.49). For a realistic borehole logging simulation this form of the estimator is still not practically helpful, considering the number of

dimensions of the integration. For the above reasons, when applying the Legendre Midway coupling for borehole logging calculations, we neglected the covariance term. Such an approximation is not per se an underestimate.

We need, therefore, to devise an approximate scheme for the variance calculation, and quantify the effect of neglecting the covariance term. The covariance of the coefficients expresses their statistical interdependency. With that it gives an indication whether superfluous information is stored in the coefficients on the underlying statistical process: if the correlation is non-zero, the data is used above their information content. If the coefficients are completely independent, the covariance term should be zero, and the estimator of Eq.(3.47) could be neglected. If we fill in the statistical estimators for the variance of the population mean into Eq.(3.41) for the truncated Legendre series, we obtain

$$\begin{aligned} & \frac{1}{N^+} \left\{ \int \phi^+(x) \left(\frac{f^+(x)}{\phi^+(x)} f(x) \right)^2 dx - \left[\int f^+(x) f(x) dx \right]^2 \right\} \approx \\ & \approx \frac{1}{N^2} \frac{1}{N^{+2}} \sum_{i=0}^K \sum_{j=0}^K \sum_{k=1}^N w_k P_i(x_k) \sum_{k=1}^N w_k P_j(x_k) \sum_{k=1}^{N^+} (w_k^+)^2 \frac{2i+1}{2} \frac{2j+1}{2} P_i(x_k^+) P_j(x_k^+) - \\ & \quad - \frac{1}{N^2} \frac{1}{N^{+3}} \left[\sum_{i=0}^K \sum_{k=1}^N w_k P_i(x_k) \sum_{k=0}^{N^+} w_k^+ \frac{2i+1}{2} P_i(x_k^+) \right]^2 \end{aligned} \quad (3.48)$$

The second term of the RHS vanishes N^+ times faster than the first term, and we will neglect it from now on.

Neglecting the covariance of the coefficients means that the significant term of Eq.(3.48) will take the form of

$$\begin{aligned} & \int \phi^+(x) (\phi(x))^2 dx = \int \phi^+(x) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} d_i d_j P_i(x) P_j(x) dx \approx \\ & \approx \int \phi^+(x) \sum_{i=0}^{\infty} \left(\frac{2i+1}{2} d_i P_i(x) \right)^2 dx \end{aligned} \quad (3.49)$$

i.e. we estimate the square of a function with the sum of its squared Legendre components. In case of truncated series, the most conservative estimator for the error is given by

$$\begin{aligned} & \sum_{i=0}^K \sum_{j=0}^K \frac{2i+1}{2} \frac{2j+1}{2} d_i d_j P_i(x) P_j(x) = \\ & \leq \frac{1}{2} \sum_{i=0}^K \sum_{j=0}^K \left(d_i \frac{2i+1}{2} P_i(x) \right)^2 + \left(\frac{2j+1}{2} d_j P_j(x) \right)^2 = (K+1) \sum_{i=0}^K \left(d_i \frac{2i+1}{2} P_i(x) \right)^2 \end{aligned} \quad (3.50)$$

and for the integral of Eq. (3.49) this means that it is less than $K+1$ times the independent second moment. For further investigation we write the exact expression for the function squared:

$$\phi^2(x) = \sum_{i=0}^{\infty} \left(\frac{2i+1}{2} d_i P_i(x) \right)^2 + \sum_{i=0}^{\infty} \sum_{j \neq 0}^{\infty} \frac{2i+1}{2} \frac{2j+1}{2} d_i d_j P_i(x) P_j(x) = f(x) + g(x) \quad (3.51)$$

where $f(x)$ and $g(x)$ are introduced to simplify the notations. As $g(x)$ contains only the cross products of the Legendre terms, the integral of $g(x)$ is zero. This means that for constant ϕ^+ the approximation is exact. We also see that $f(x)$ contains all the terms that are always positive. As $f(x) + g(x)$ describes the square of a real function, it must be positive for every x . If we separate the integration domain Γ into the parts where g is positive (Γ_+) and where g is negative (Γ_-), we see, that

$$\left| \int_{\Gamma_-} g(x) dx \right| = \int_{\Gamma_+} g(x) dx \quad (3.52)$$

and $g(x)$ on Γ_- is bounded by

$$|g(x)| \leq f(x) \text{ for } x \in \Gamma_- \quad (3.53)$$

If the sizes of Γ_+ and Γ_- equals, $f + g \leq 2f$ for every x in the whole integration domain, meaning that the exact variance is less or equal than two times the estimated one. It is most likely that the integration domains will be different in size, and the positive part of $g(x)$ then might, but still not surely, exceed $f(x)$, while Eq.(3.52) should still hold: if g exceeds f the size of the domain where it happens is inversely proportional to their ratio. To have, therefore, an underestimate of the variance by neglecting $g(x)$ (beyond a factor 2 difference) is possible, but it would require special ϕ^+ which is low on Γ_- and high on Γ_+ . In general it is more likely to have an overestimate instead, i.e. to have the neglected covariance term negative.

3.2.5 Applying the Legendre Midway coupling for a borehole logging calculation

Application of the Legendre coupling in each dimension of the Midway integration is not practical. First of all, if we apply it for the time variable, the time convolution formulae cannot be used any more, and each time response have to be calculated separately by a separate coupling. For energy, the multigroup calculation results should not be expressed by continuous functions, leaving four variables for consideration: two independent space- and two independent angle coordinates. We should also take into account that the number of required computational operations would increase with $K_{\Omega_1} K_{\Omega_2} K_{r_1} K_{r_2}$ where K with the different subscripts stands for the number of coefficients used for the two independent angular and the two independent spatial dimensions respectively, while for the segmented coupling it increases with $K_{\Omega_1} + K_{\Omega_2} + K_{r_1} + K_{r_2}$. If we would implement a proper error calculation the required computer time would increase to its square. The required time for a segmented coupling is dependent also on the number of scores, for our applications it took around 1/10-1/20 of the time of the Monte Carlo calculations. For different applications such a rule might not be valid, but here we should keep the computation time of the coupling below the efficiency gain times the computer time of the Monte Carlo calculations. For our borehole

logging simulations this would give the chance to use 1-3 Legendre coefficients per dimensions, and that would clearly not suffice. We should also choose a dimension in which the functions are smooth enough, preferably a dimension where the functions behave completely smoothly. On another note, we might replace the coupling in a dimension where the approximation with the segmented coupling can hardly be set optimally using a structured mesh.

One choice to fulfil all the above requirements is to select the angular variables for the Legendre coupling. The flux in the azimuthal part of the solid angle is strongly correlated with the flux dependent on the angle (the "spatial angle") of the cylindrical coordinate system: with the tool model placed in the middle of the geometry the rotational symmetry of the system is not ideally utilised by setting the origin of the azimuthal angle fixed to a certain direction. At each spatial angle the maximum flux might be expected in the direction of the source, while that direction vector is represented in a coordinate system with an origin that does not rotate with the spatial angle. Naturally we could change the origin of the azimuthal angle to point always at the source, but to keep the method general, we will avoid this. In the simple case of a tool centred in the middle of the borehole we can simply neglect both of these angles (they can not be neglected separately) if we keep this coordinate system. Though we do not present the results of such simulations in this thesis, but most commonly the tool is pushed to the borehole wall, and to maintain the generality of the conclusions here, we keep both dimensions of the functions to be integrated. We can therefore expect a better performance of the Legendre function in the azimuthal angle direction.

With this choice of variables we achieve that the application of the Legendre coupling became feasible in terms of keeping the computer memory need reasonable according to current standards, but we introduce a major unreliability in the comparison of the methods by still having a segmentation structure in the two spatial dimensions, with truncation error unknown. We have also learned in Chapter 2, that the amount of scores reaching the Midway surface has a major effect on the accuracy, until having enough scores for stable segment-wise estimates. We selected therefore a Midway set-up with a very low number of segments by creating 6x6 spatial bins and 6x4 (cosine times azimuthal) angular bins for 10^7 starting particles of both calculations, moreover we present the calculations for the near detector, whereas the scoring efficiency at the Midway surface is high. The Generic Tool Model has been used, and both techniques used the same set of scores.

The segment boundaries have been "tuned" according to a trial-and-error optimisation scheme that allowed good agreement with the Midway and the reference calculation results. Such segmentation is possible, as the adjoint function and the flux decrease with increasing radius, resulting in an overestimate of the response having both function's derivatives negative; while with increasing time, the flux decreases and the adjoint function increases, resulting in an underestimate of the response. With proper segment boundaries the two effects cancels, and the inaccuracy will not exceed the statistical fluctuations. The comparison, therefore, cannot result in a better estimate of the response using the Legendre coupling, but the results show, that the Legendre estimates are as accurate as for a fine-tuned segmented coupling. Fig. 23 shows the results, together with two additions to the calculation settings: a change from the fine-tuned angular boundaries to an equally spaced azimuthal and cosine boundary system still having 6x4 divisions, and secondly an additional calculation for inserting some additional radial divisions. The error bars are given by 1σ .

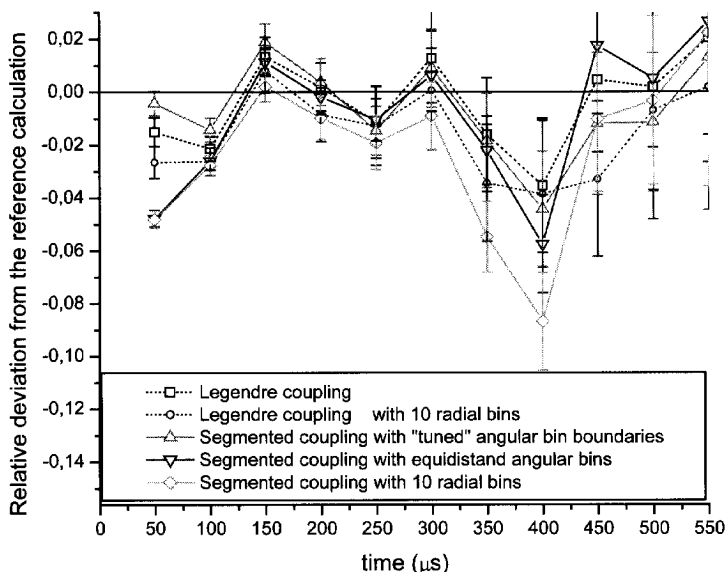


Figure 23. Legendre and segmented coupling compared in accuracy. The Legendre coupling achieves high accuracy without laborious tuning of the coupling parameters.

As it can be seen in Fig. 23, all results are within statistical inaccuracy (2σ), moreover these confidence intervals contain the zero deviation from the reference results. Exceptions are at the first two points, caused by the applied time convolution (type 1), whereas only one or two time segments are used for the coupling, and the derivatives are the highest there for both functions in absolute value, and are of opposite sign, altogether resulting in large negative (first order) deviation. There is a fair agreement of the Legendre coupling with the segmentation with fine-tuned angular bin boundaries. A systematical deviation (though still within statistics) can be observed when changing the angular binning to an equidistant one. When changing the spatial segmentation by randomly inserting three additional bin boundaries, an apparently systematic change can be observed in the segmented response tending to a slight underestimate (lowest solid line), indicating that the “fine-tuned” segmentation was somewhat specific also to the radial bin structure. This phenomenon appears to be also valid in sense of the estimated confidence intervals for the 350 and 400 μ s time readings.

The selection of 6×4 angular coefficients and segments was a result of “tuning” the Legendre Midway coupling. This comprised a calculation with 7×7 coefficients, and checking the difference of the result when eliminating the higher coefficients. The lowest number of important contributions has been kept. To keep the comparison in some sense fair, we selected the same number of angular segments and “tuned” the segment boundaries. Naturally

we could have selected another rule for a fair comparison, but this one included the least amount of human input, while allowing the formulation of general statements.

To experimentally check the validity of the simplified Legendre error approximation, we carried out the same calculation 50 times for the Generic tool model, with the same settings as for Fig.12. We can see, that the predicted error bounds exceed the statistical fluctuation, and our error estimates are an overprediction, according to the empirical variance figure by around 10%.

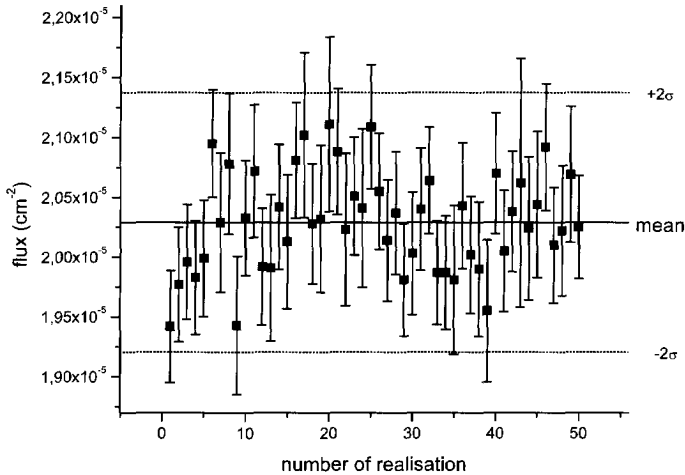


Figure 24. *Empirical spread of the Legendre Midway estimates for a borehole logging calculation within the predicted standard deviation (σ).*

It was not the intention of this calculation to show that for all of the time dependent responses the variance estimator of the Legendre coupling is an overestimate, its purpose is mainly implying that the given error bounds are representative in a qualitative sense.

Although the efficiency of the time dependent Midway method will be the subject of Chapter 5, a comparison will be shown here of the above calculations in terms of efficiency, whereas the biggest factor differentiating between the coupling schemes was the relative error, as the coupling calculations themselves took less than 1/10 of the Monte Carlo simulations. Fig. 25 shows the efficiency of the Midway coupling measured by their FOM. The solid lines are representing the Midway coupling results for the above time dependent cases, the dashed line shows the efficiency of the autonomous forward calculation as obtained from 10^9 starting particles. Here we draw two general conclusions. The first one is that the Legendre coupling is not more effective than the segmented coupling, even though it couples more adjoint and forward scores from the same set. The second one is that the efficiency of the time dependent Midway method with low number of “tuned” segments is heavily dependent on the phase space segmentation structure.

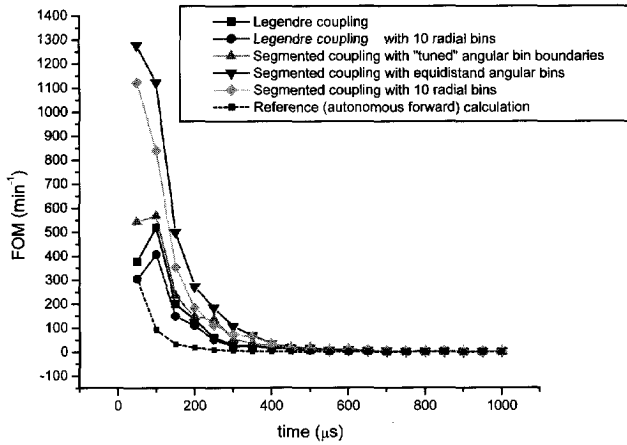


Figure 25. Comparison of efficiency of the Legendre and the segmented couplings with the reference calculation

We can conclude, that it is less efficient than coupling with the segmentation technique, at least if we rely on a simplified error calculation scheme. We have seen that the calculation of the necessary quantities for the response and the variance require considerably more computing effort, and depending on the way of programming, possibly more computer memory. The drawbacks can be balanced by the fact, that for an unknown system, the Legendre coupling provides a simple scheme, that does not require deep insight into the problem before the calculation, and using more or less coefficients does not need the adjustment of many, possibly interlinked quantities like setting new segment boundaries. It is therefore possible to do the coupling for a number of coefficients, and observe the effect of omitting some of them, thereby obtaining an indicator on the accuracy of the coupling. If having enough coefficients, we can also give a simple estimate of the truncation error, as it is in the order of the last retained coefficient (for coefficients with algebraic convergence).

3.3 General Remarks on the Midway Coupling Possibilities

3.3.1 Expansion of the linking function

The derivation of alternative Midway coupling based on complete function bases given in Section 3.2.1 did not rely on the Legendre polynomials, or any polynomials at all. It is also possible to use Chebyshev, Laguerre or other polynomials, to do the coupling by discrete Fourier transform, using complete orthogonal function bases in more dimensions like spherical harmonics, etc. Each function system has its well-established theory and each would allow special simplifications of the calculations. The Fourier transform for example would offer, perhaps, an alternative to the *type 1* time convolution, and some of the second moments coincide with other first moments. These techniques all share however the same basic

approximation by replacing the Dirac-delta function $\delta(x-x')$ by a function representation which we denoted earlier with $L(x,x')$.

In case of the Legendre polynomials, the linking function of Eq. (3.28) can be further simplified using the Christoffel summation formula [21], to

$$L(x,x') = \sum_{i=0}^K \frac{2i+1}{2} P_i(x) P_i(x') = \frac{K+1}{2} \frac{P_K(x) P_{K+1}(x') - P_{K+1}(x) P_K(x')}{x'-x} \quad (3.54)$$

This formula allows a score-wise coupling scheme, in the same way as for time coordinates the time convolution *type 3* coupling in Chapter 2 by the calculation of only two Legendre terms. The linking function, being a representation of the Dirac-delta function in each dimensions can be easily extended to multiple dimensions, let us say for a new y coordinate it takes the form of $L(x,x') L(y,y')$. In a concrete representation, as we have seen earlier, this does not result in assuming the flux and the adjoint function separable in x and y .

Given the Monte Carlo estimate for the Midway integral can be written by

$$\hat{R} = \frac{1}{N^+} \frac{1}{N} \sum_{i=1}^{N^+} \sum_{j=1}^N w_i^+ w_j L(x_j, x_i^+) = \frac{1}{N^+} \frac{1}{N} \sum_{i=1}^{N^+} \sum_{j=1}^N c_{i,j} \quad (3.55)$$

where c_{ij} stands for the contribution that is delivered by the i^{th} adjoint sample combined with the j^{th} forward one. While for the time variable it is possible to give an exact form of the linking function, in the other variables this form is an approximation, and the order of estimate depends on the spatial distribution of the particles crossing the Midway surface.

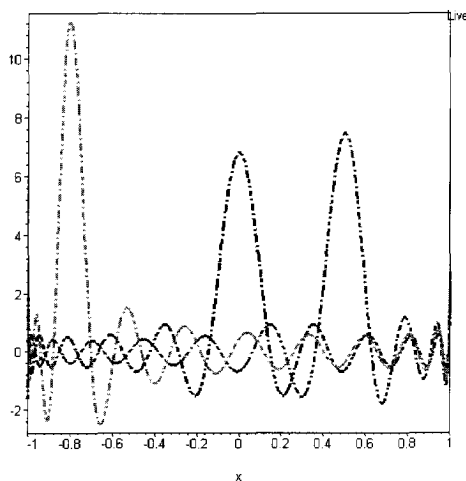


Figure 26. 20th order Legendre expansion of the $L(x,x')$ Linking function at $x'=0$, $x'=0.5$, $x'=-0.8$

It may seem that this definition of the linking function differs mainly from the segmentation technique by using all the scores available, but that is not the case. As it can be seen in Fig. 26, the Legendre linking function is also effectively limited to small spatial domains, and the sizes of the domains of surely significant and positive contributions are dependent on the order of Legendre expansions. Fig. 25 shows the linking function in the x variable for three x' (adjoint crossing) coordinates: -0.8 , 0.0 and 0.5 respectively. The effective size of the domain, where a forward score would give a reasonable contribution to the response, is limited, and varies even with the actual coordinate x' . We may attempt to establish a correspondence between a certain number of segments and a certain order of Legendre expansion based on the sizes of these effective domains. The essential difference of the two methods is given by the fact, that for the segmentation technique, the boundaries are fixed, while for the Legendre coupling they vary with the surface crossing coordinates.

Eq. (3.54) offers a simple manner of calculating the response from the raw score coordinates like for the time convolution type 3. This coupling would not require the calculation of the expansion coefficients at all: only the K^{th} and $(K+1)^{\text{th}}$ Legendre polynomials should be calculated, but would require the storage of all continuous independent coordinates of scores of the first processed calculation. This, again, would require further studies on the calculation of the variance. Having enough scores to estimate the response, this scheme increases the number of calculations to the order of NN^{r} , and would hardly be advantageous. It offers, however, the possibility to combine the segmentation technique with the Legendre method, or rather a segmentation scheme can be applied to the formula of Eq. (3.55). For a selected order of Legendre polynomials we can set up a rough segmentation of the phase space by using some times the effective domain of L . This imposes an additional filtering of the scores and decreases the calculation effort, together with the memory need. If the phase space is not oversegmented by this coarse mesh, we can ensure that many particles will be coupled, while the memory requirement is still kept reasonably low, allowing the storage of both forward and adjoint particles as it is necessary for a proper error calculation. With such a method, the scores can be processed in batches: a certain subset of forward scores can be coupled with a certain subset of adjoint scores, selected according to the memory capacity of the given computer. When one batch is done, another pair of (uncorrelated) sets can be chosen.

This procedure is meant to be given as an example only; such approximation schemes are already devised in other fields of physics and mathematics. If we look at the Monte Carlo scores as data obtained from an unknown process, we can incorporate the techniques of digital signal processing. In fact, the difference of the segmentation and the Legendre techniques is the same in many details as the difference of finite elements and spectral methods; many devised schemes of both fields could be applied, if the statistical estimators for every relevant quantity can be formulated. Adaptation of Gauss quadratures offers the chance to represent the flux or the adjoint function in a point somewhere in the segment instead of representing the function with the integral average on the whole segment domain, whereas the function value could be calculated at that point by 2 Legendre coefficients of the desired order according to (3.54). In a statistical way of approaching the problem, data sorting algorithms, e.g. Delunay triangulation, could determine the proper segment boundaries.

3.3.2 Optimal Midway coupling

The optimal Midway coupling should be an optimum according to the following factors: the computation effort, computer memory requirements, the requirement in human expertise and effort for the Midway method settings, accuracy of the coupling scheme even without much a priori knowledge on the model system, robust statistical estimate of the response, robust and programmable statistical estimate of the variance, and low final statistical variance. The last item on this list, the variance outcome will determine whether the Midway method can indeed be used as a variance reduction technique, and as we have seen on Fig. 26, the variance calculated from the same set of scores may depend on how we set the free parameters of the coupling technique.

	Segmentation Technique	Legendre Coupling
Accuracy of the coupling	Low/apriori unknown	High
Computation effort	Low	High
Stable statistical estimate	No	Yes
Stable variance estimate	No	-
Variance estimator	Exact, programmable	Major difficulties in programming
Variance of a sample set	Increases with number of segments	Increases with polynomial order

Table 3. : *General properties of the coupling schemes*

Summarising our findings of Chapter 2 and 3, we can say that the segmentation technique can deliver Midway-coupling results by a relatively simple and fast calculation, while having an easily programmable variance estimator. The accuracy of the coupling is unknown, and the estimation of the residual is not feasible in a statistical sense. The samples on the response as given by some forward and adjoint scores of the same segment is not homogeneous, the number of outliers increases with decreasing number of samples, and if the number of samples is comparable to the number of segments the estimator for the response might comprise only outliers. The accuracy of the coupling increases with the number of segments and depends on the distribution of segment boundaries, and a high accuracy of the coupling requires a certain minimum amount of segments. Different segmentation structures might give different variance estimates even if the number of outliers is small. For an unknown system the minimum number of segments must be guessed, and the corresponding number of scores must be simulated until the estimates become stable. We might select the segment boundaries to result in a lower variance afterwards, while the result stays within statistical error bounds. The analysis using different segmentation boundaries is necessary for the segmentation technique for each calculation, as the segmentation technique is a trial-and-error method,

without an immediate indication of the quantity of the error, or the decrease thereof when applying more segments.

The Legendre technique offers a statistically well behaving solution to all the above problems, to the cost of strongly increasing coupling times and/or memory need and probably increasing variance (and sophisticated variance formulas). The Legendre coupling if applied in all continuous dimensions gives a well-behaving statistical set even with low number of scores, and it is possible to give estimates of the response based on a couple of adjoint and forward scores. The model parameter here is the number of coefficients, and the response estimate is built up from estimates of increasing parameter number, giving the possibility of checking the response of lower polynomial order to inspect the convergence of the estimate.

To cast a picture of an optimal coupling, we can say that the Legendre method applied in all continuous dimensions delivers an accurate and well-behaving estimate. It fails, when we wish to use the Midway coupling as an automatic variance reduction technique: the calculation time and the memory needs are too high. Therefore a feasible technique for variance reduction purposes should be a combination of these two methods. A possible solution would be to combine segment-wise the Legendre and segmentation techniques: to estimate in a statistically stable way the first few Legendre coefficients per segment. This procedure would result in a similar method as finite elements interpolatory schemes. But already then, the number of dimensions for the integration would increase the memory requirements by a factor K^5 for K higher order components for the 5 dimensions (of continuous dependency). The increase in memory and computation time is more alarming for the variance estimates.

Spectral methods are only more advantageous than the segmentation technique, as long as few coefficients would suffice for a successful function representation. Even with the increased computer resource requirements well chosen functions for the expansion (Chebyshev rather than Legendre polynomials, spherical harmonics for the angles, Laguerre for the radius and exact coupling in time) may be considered. These choices might improve the Midway coupling relating more or less specifically to a borehole logging application.

As long as we wish to apply the time dependent Midway method as a variance reduction technique, incorporation of every further higher order term in each dimension would cost around at least a factor 2^5 increase in the necessary computer resources, compromising extremely easily the aim of increasing computation efficiency. We should therefore decrease the complexity of the Midway integral by selecting a Midway surface fitting somehow the particle distribution in the phase space, or reducing the complexity by omitting one dimension of the integration (e.g. time). Without the necessary addition of the accuracy estimation and statistical stability check, the method would not possess the statistical properties that would qualify the time dependent Midway method as a proper general and automatic Monte Carlo variance reduction tool.

Chapter 4

ADJOINT SAMPLING OF A PULSE HEIGHT DISTRIBUTION

A Monte Carlo calculation does not necessarily require stating equations and defining detector functions, as long as the calculation really simulates particle transport as it could also occur in nature i.e. the simulation is *analogue* to nature. These simulated detector responses can be very close to what we measure with a real tool, down to the processes of data acquisition. However when we replace the simulation with another calculation technique, we need to prove that they are statistically equivalent, and this requires mathematical formulation. There are responses of real detectors that are particularly hard to put in mathematical form, the Pulse Height Distribution (PHD) for one. These responses share the common feature that they cannot be described as a linear functional of the flux using a single detector function. This obviously poses a problem for an adjoint calculation, being its source term the detector function. Scintillation detectors of the neutron-gamma tool are measuring a pulse height distribution of the photons in energy. The Midway calculations utilised the flux at the detector so far, hence for a complete application of the Midway method to borehole logging calculations, we need to find a way of sampling such a detector response for the adjoint calculation. This chapter is devoted to this effort: finding the theoretical framework that both satisfactorily describes the detector response, and can be sampled for the adjoint calculation to allow the formulation of the Midway coupling.

Mathematical formulation for the PHD has not been devised yet (to the author's knowledge), and the lack thereof invoked much dedicated research to prove that a certain calculation is statistically equivalent to an analogue simulation. Especially for neutron-gamma borehole logging applications the usual Monte Carlo simulations do not approximate the PHD satisfactorily, not even by an analogue calculation. Many attempts targeted an approximation of the PHD by a detector function [8, 23, 26] especially with regards to borehole calculations. Booth [24] and Shuttleworth [25] devised a scheme that allows using some form of variance reduction technique with the PHD. To reduce the discrepancies between simulations and measurements especially for coupled neutron-photon problems, even special Monte Carlo codes have been written [8, 23, 26]. The extent of these discrepancies sometimes violates successful data interpretation. Therefore sometimes measurement designs are altered to allow their easier simulation. The references cited here are a small few of the many publications; scintillation counters are very frequently used in all fields of particle transport. The references are selected according to their relevance to time and energy dependent neutron-gamma Monte Carlo borehole logging simulations. Especially slim tool designs show all the special

phenomena for the detection that can cause a problem for Monte Carlo simulations, and every mentioned approach applied some sort of an approximation.

4.1 Scintillation Detectors, Pulse Height Distribution and non-Boltzmann Tallies

4.1.1 Simulation of scintillation detectors

Scintillation detectors are very often used for photon detection. Our aim here is by no means to give a detailed description of the detection process, but to outline the basic characteristics that verify and characterise the approximations to be applied. It suffices to say here, that the scintillation detector (although we use it for photon detection) is an electron detector [27] that can indirectly detect photons and their energies. A scintillation detector essentially consists of a transparent crystal (NaI, BGO, etc.), a photoelectrode and a photomultiplier tube. The principle of the detection is that free electrons excite the bounded electrons in the crystals, which (while returning to their normal state) will produce visible photons, i.e. scintillation will occur. The number of excitations is proportional to the energy of the electron. There are many interactions in the electron transport that we will neglect here, and assume that all of the energy of the electron is transformed into visible light. This assumption is certainly valid, as long as the detector crystal is big enough for minimal energy loss by electron escape, practically meaning a size at least in the order of a centimetre. The produced visible light is transformed into electric signal after the electron signal has been amplified in the photomultiplier tube.

Measurement of a gamma photon is actually the measurement of the free electrons it produced. Incident gamma photons may undergo some interactions (e.g. Compton scatter, photoelectric effect, pair production) in the detector crystal, creating free electrons with energy equal to the energy the photon has lost. Through the process of scintillations the released energy can be measured. The lifetime of the photon is far less than the process of the measurement of its energy (exciton decay, data acquisition), therefore if a photon had more than one interaction in the detector, the corresponding energy measurements cannot distinguish between these interactions, and the measured energy will be proportional to the total energy release of the photon in the scintillation crystal. The measurement of the energy is not an ideal process; it is strongly dependent on the crystal properties (material constituents, purity etc.) and on the temperature.

As described in Chapter 1, the C/O operation mode of the neutron-gamma tools aims at distinguishing between photons emitted at inelastic collisions of neutrons on carbon and on oxygen. As we have seen in Fig.4, the number of photons arriving to the detector with their characteristic energy still intact is high, while the detected energy distribution hardly resembles the incoming one. The reason of this is *the partial energy release of photons* caused by the escape of gamma photons; thereby identifying the first important process to take into account for the simulations.

The difference of the time scales of detection- and photon lifetimes results in the joint detection of all interactions of the photon. For neutron-photon problems a high intensity neutron source such as the one applied for borehole logging measurements, may produce so

many photon, that interactions of separate photons will not be distinguishable, and energy releases of different photons are accounted as a total energy release of the same photons. This will result in pile-up peaks, at energies not corresponding to energies of single photons of the system. This constitutes the second important process: *coincident detection of photons*. Monte Carlo simulation of such phenomenon for coupled neutron-photon borehole logging problems is very problematic, given that the neutron lifetime ($\sim 100\mu\text{s}$), the photon lifetime ($\sim \text{ns}$) and the scintillation decay times ($\sim \mu\text{s}$), are on different time scales by orders of magnitudes.

The third effect to consider is *physical branching* of particle histories. The simplest and most important example is pair production, when a gamma photon creates an electron-positron pair. After the positron has slowed down it will annihilate, creating two 0.511MeV photons flying exactly in the opposite direction. If this process happens in the detector, both or one of these photons can escape from the crystal, resulting in 1.022 or 0.511MeV lower measured energies. From a Monte Carlo point of view, this process can be simplified (if we neglect the electron transport) to a branch in the particle trajectory (with the energies properly accounted for).

4.1.2 The Pulse Height Distribution

The above-described procedure yields an energy spectrum that is not directly proportional to the energies of the incident photons. This spectrum is called the (differential) Pulse Height Spectrum, or the Pulse Height Distribution. If the detector crystal is large enough to prevent any escape, the energy of the photons would be fully measured, and the PHD transforms almost to an energy spectrum of the measured photons (except for the pile-up peaks). The other case when the PHD can be described as a simple function of the flux, is when we omit the energy discrimination, and tally the particles if they had at least one interaction in the detector.

In case of a small detector and a neutron source of high activity and an optically thick medium between source and detector, the coincident detection of photons is hardly solvable by an analogue Monte Carlo simulation, for to obtain enough particles for real coincidences is not in the range of common computing capacities. It is possible [29] to use the results of many simulations (detection time and energy release), and estimate pile-up in a post-simulation calculation. A used -but not likely valid- approximation is calculating pile-up peaks for any physical branching in the system, e.g. every photon generated at neutron interactions would be regarded as being detected coincidentally.

An essentially easy implementation of the other two processes is available in most Monte Carlo codes: coincident detection of all the energy releases of the same photon and its progeny in the detector. The tallying process essentially differs from a response defined using a response function, in one crucial point. If we count the particles that had an interaction in the detector with energy in a certain ΔE_i energy range, instead of calculating the energy loss, each particle getting into the detector with the energy in ΔE_i would give a contribution independently of its afterlife, and any subsequent collisions in the detectors might give contributions independently of their previous history if they still have an energy in ΔE_i . For the PHD on the contrary, we need to wait until the particle (and all its progeny) is terminated to calculate the total energy release, it is not enough to wait until the simulated particle leaves the detector, as it might later return and modify the total energy release count.

If it indeed is impossible to describe the PHD response by a linear functional of the flux, what is the missing information that is not contained in the flux? The answer is the correlation of the flux at a phase space point with the flux at another phase space point. If we know the flux at every point of the system, we can calculate the probability of an energy release of a given extent at every point. However if we know the flux at point a and at point b we do not know the probability of transmission from point a to point b , therefore we cannot determine the probability of a particle releasing a certain total energy when having a collision at a and at b .

4.1.3 Pulse Height Distribution and non-Boltzmann tallies

The PHD is not the only quantity that cannot be given as a linear functional of the flux. These quantities are called *non-Boltzmann tallies* (perhaps a better name would be *non-Boltzmann responses*) after the publication of Booth [24], although a proper definition was not given there. Their common feature and finally definition is, that they are dependent on characteristics of multiple events of a particle history, or of separate particle histories. A Boltzmann-tally can then be defined as a response that is sufficiently well describable as a linear functional of a quantity that describes an expected (average) value of a function of the particle population at a certain phase-point.

Simple examples of non-Boltzmann quantities that have been investigated for Monte Carlo calculations are quantities of reactor noise analysis [29, 30] (time correlation), coincidence counters (direct correlation of progeny), and segment-wise response of the Midway method (spatial correlation). This last example covers a useful concept for understanding source-detector systems. The quantity there defined is the probability that a particle reaches a certain phase-space point/region of the model domain and scores afterwards. This quantity can be constructed from Boltzmann-responses, but still cannot be given as a simple linear functional of the flux. Not as long as we use only one Boltzmann-quantity, (as we have seen at the derivation of the Midway method,) it is a bilinear integral (or a product) of two Boltzmann-concepts: the flux and the adjoint function. It is possible to detach the calculations into a first one that calculates the flux at the given intermediate phase-space domain, then calculate the second part (detector response of particles started at the specified intermediate domain) by assigning the flux of the first calculation as a source term for the second one. Either ways we can precompute by another stochastic process a detector function i.e. the probability of reaching the detector from the given domain, and transform the original non-Boltzmann response to a linear functional of the detector function.

Such a transformation for the PHD would solve our problem of defining an adjoint source for the PHD sampling. We could precompute a response function let us say on the boundary surface of the detector and use it as a detector function for the flux (more precisely for the current) [23, 26], and finally use it as an adjoint source. This attempt would be problematic in two ways. First of all, such a detector function would be dependent on the surroundings of the detector as photons might re-enter the detector and that would need special non-Boltzmann handling again. Secondly, more than one progenies entering or re-entering the detector would again result in incorrect estimates. If we contain the detector in an envelope that -very much alike the midway enclosure- contains completely the detector, we might set this surface far enough to minimise both effects. The problems we gain in exchange would be the same as for the Midway method but in one more spatial dimension.

4.2 Feasibility of the Adjoint Sampling of the Pulse Height Distribution

The following sections focus on a single special feature of the PHD: the sampling of the total energy release in the detector without coincidence counting or any physical branching. First we will show that it is not necessary to state the adjoint Boltzmann equation to play an adjoint Monte Carlo game for Boltzmann responses. After that we develop a formalism based on the random walk process that allows the definition of the response function of the PHD. By showing the form the Boltzmann-response function takes using this formalism, we will point out why an adjoint Monte Carlo game essentially differs from playing a simulation in reverse. Afterwards we develop a form of the PHD that can be sampled entirely from the detector, and an adjoint game that can be played entirely according to adjoint transport kernels. The last section will show a numerical example to illustrate that the developed theory is valid.

4.2.1 Monte Carlo Method for Boltzmann responses

To arrive at a demonstration that adjoint Monte Carlo is not necessarily related to the adjoint transport equation, we briefly review the integral equation formalism of Monte Carlo simulations. The integral equation formalism behind transport Monte Carlo is based on the following form [33]:

$$\psi(P) = S(P) + \int K(P' \rightarrow P) \psi(P') dP' \quad (4.1)$$

The detector response can be described as:

$$R = \int h(P) \psi(P) dP \quad (4.2)$$

Here $K(P' \rightarrow P) dP$ is the probability that a particle has coordinate P in dP at a collision if its previous collision was at P' , and $\psi(P)$ represents the number of particles at a collision. Depending on the actual definition of $K(P' \rightarrow P)$ (i.e. if it describes first a free flight then a collision or vice versa), $\psi(P)$ is interpreted as the number of particles with coordinate P , *leaving* a collision or *entering* a collision. Note, that the choice of this definition changes the exact form of $S(P)$ as well: for the (entering) collision density it transform to first collision points.

The equation adjoint to Eq. (4.1) is

$$\psi^+(P) = h(P) + \int K(P \rightarrow P') \psi^+(P') dP' \quad (4.3)$$

and a response is to be found in the form of

$$R^+ = \int \psi^+(P) g(P) dP \quad (4.4)$$

The reciprocity theorem states that if $g(P) = S(P)$ then $R = R^+$.

The Monte Carlo sampling of such integral equations will be briefly summarized now. Integrals of functions can be calculated by the Monte Carlo method. To arrive at the solution of integrals of functions defined by an integral equation, we may consider the following: let $\psi_1(P)$ be an unknown function, $\psi_0(P)$ a known pdf, $K(P' \rightarrow P)$ a nonnegative function, and let

$$\psi_1(P) = \int K(P' \rightarrow P) \psi_0(P') dP' \quad (4.5)$$

The aim is to determine the integral

$$I_1 = \int h(P) \psi_1(P) dP \quad (4.6)$$

This problem can be solved by Monte Carlo, by defining the conditional pdf

$$k(P|P') = \frac{K(P' \rightarrow P)}{\int K(P' \rightarrow P'') dP''} \quad (4.7)$$

and selecting P' from $\psi_0(P')$, selecting P according to the conditional pdf $k(P|P')$ and calculating

$$I_1 = E \left[\frac{1}{N} \sum_{i=1}^N h(P_i) \cdot \int K(P_i \rightarrow P'') dP'' \right] \quad (4.8)$$

If $\psi(P)$ of Eq.(4.1) can be written as a sum of functions that are connected recursively like those in Eq.(4.5), then the response R of Eq.(4.2) can also be expressed as a sum of integrals like the one shown in Eq.(4.5).

If we wish to reverse the process and draw samples from the scoring function $h(P)$, the feasibility of such a game can be proven. A simple "adjoint" (reversed) method for calculating the same integral can be formulated from the previous equations. If we put the expression for $\psi_1(P)$ into the expression for I_1 , we get from Eq.(4.6):

$$I_1 = \int h(P) \int K(P' \rightarrow P) \psi_0(P') dP' dP = \int \psi_0(P) \int h(P') K(P \rightarrow P') dP' dP \quad (4.9)$$

If $h(P)$ is a nonnegative function, selecting P' from $h(P') / \int h(P'') dP''$, and P from

$$k^+(P|P') = \frac{K(P \rightarrow P')}{\int K(P'' \rightarrow P') dP''}, \quad (4.10)$$

the final expression can be estimated as

$$\hat{I} = \frac{1}{N} \sum_{i=1}^N \psi_0(P_i) \cdot \int K(P' \rightarrow P_i) dP' \cdot \int h(P'') dP'' \quad (4.11)$$

since the expected value of this expression is

$$\begin{aligned}
 E\left[\hat{I}_j\right] &= E\left[\frac{1}{N}\sum_{i=1}^N\psi_0(P_i)\cdot\int K(P'\rightarrow P_i)dP'\cdot\int h(P'')dP''\right]= \\
 &= \int\int h(P)k^+(P|P')\psi_0(P')\int K(P\rightarrow P')dP'dP = \int h(P)\psi_1(P)dP
 \end{aligned}
 \tag{4.12}$$

In the forward case, when a chain of functions is constructed by repeated application of Eq.(4.5),

$$\psi_j(P) = \int K(P' \rightarrow P)\psi_{j-1}(P')dP', \tag{4.13}$$

This recursion expresses nothing else but the collision density after j collisions.

With the 0th (source) function known, the integral

$$I_j = \int h(P)\psi_j(P)dP \tag{4.14}$$

can be estimated by Monte Carlo for any j in the following way. First select an initial coordinate P_0 from the first function of the chain $\psi_0(P)$, select the next co-ordinate P_1 from

$$k(P|P_0) = \frac{K(P_0 \rightarrow P)}{\int K(P_0 \rightarrow P')dP'} \tag{4.15}$$

and set a weight factor to $w_0 = \int K(P_0 \rightarrow P')dP'$. After this, keep repeating the process of selecting P_i from $k(P_i|P_{i-1})$ and of setting $w_i = \int K(P_{i-1} \rightarrow P')dP'$ until j is reached. Then a sample of the estimator for I_j is obtained:

$$\hat{I}_j = h(P_j) \prod_{i=0}^{j-1} w(P_i) \tag{4.16}$$

and an average of these samples is the final estimate of I_j . The proof for this can be found in Ref. 33. It can be also shown that a solution for an integral $R = \int h(P)\psi(P)dP$ with $\psi(P)$ defined by an integral equation like Eq.(4.1) is given by the sum

$$R = \sum_{j=0}^{\infty} I_j \tag{4.17}$$

Provided the series converges, and the integral exists, the integral I_j can be estimated by adjoint Monte Carlo as well by changing the order of integrations in the chain of recursive functions like in Eq.(4.9):

$$\begin{aligned}
 I_j &= \int h(P_j) \int K(P_{j-1} \rightarrow P_j) \dots \int K(P_0 \rightarrow P_1) S(P_0) dP_0 \dots dP_j = \\
 &= \int S(P_0) \int h(P_0) K(P_1 \rightarrow P_0) \dots \int K(P_j \rightarrow P_{j-1}) dP_0 \dots dP_j.
 \end{aligned}
 \tag{4.18}$$

If a sample is selected from the detector function $h(P)$, and the next samples are drawn from the kernel k^+ of Eq.(10), then, if setting the adjoint weight factor to $w_i^+ = \int K(P \rightarrow P_{i-1}) dP$, an adjoint sample for I_j can be obtained as

$$\hat{I}_j = S(P_j) \prod_{i=0}^{j-1} w^+(P_i) \quad (4.19)$$

Such a form is often derived from the adjoint equation (Eq.(4.3)) following the derivation for the forward Monte Carlo solution. With this line of thought, however, we can arrive at a derivation that does not need the statement of the adjoint equation, but verifies the adjoint Monte Carlo process with results coincident with the solution of the adjoint transport equation. The step is important, as the PHD cannot be defined by a single detector function; therefore the source of the adjoint equation is undefined in a Boltzmann-sense. We might now have the chance to formulate an adjoint game, without stating an adjoint equation.

4.2.2 Monte Carlo Method for non-Boltzmann responses

The continuous random walk process can also be described by a single pdf that is a function of the coordinates of all collisions (of all particles).

This pdf consists of a multiplication of conditional pdf's $\wp(P_i | P_{i-1})$, expressing the probability of a particle reaching a phase space coordinate P_i if the previous collision happened at P_{i-1} , is given by:

$$\wp(P_0, P_1, \dots, P_\infty) = \wp(\underline{P}) = \prod_{i=1}^{\infty} \wp(P_i | P_{i-1}) \quad (4.20)$$

Usually we want to estimate an expected value, which can be expressed as:

$$E[h(P_0, P_1, \dots, P_\infty)] = \int \dots \int h(P_0, P_1, \dots, P_\infty) \prod_{i=1}^{\infty} \wp(P_i | P_{i-1}) dP_\infty \dots dP_0 \quad (4.21)$$

for an n -long chain, and these expectations should be summed for arbitrary n . In certain cases this may have to be specified further: the scoring function dependent on the phase space coordinates may also depend on the actual process leading to P_i (absorption, free flight, scattering). Then the pdf must be broken up into other pdf's defining, for example, the pdf of travelling from one collision point to another and the pdf of post-collision coordinates assuming the pre-collision ones are known. Often the scoring function h can be transformed into simpler forms when it depends on one phase-space point, rather than on multiple collisions together. In those cases a corresponding score equations can be formulated, or it can be well represented by partially unbiased estimators. In the most general case the response can be dependent on several independent chains like in reactor noise measurement simulations [35], where a different statistical model needs to be set up. If independent particle histories can interact, it will change the form of the expected value calculation to

$$E[h(\underline{P}', \underline{P}'' \dots)] = \int \int \dots \int h(\underline{P}', \underline{P}'' \dots) \wp(\underline{P}') \wp(\underline{P}'') \dots d\underline{P}' d\underline{P}'' \dots \quad (4.22)$$

but the pdf's describing the individual particle histories are independent. If we have a multiplying game, meaning that a single particle can deliver multiple progeny, the expression of the pdf will change. However, we will restrict ourselves to handling non-multiplying, single chains. An important irreducible chain is when h is setting a constraint for the integration, which constraint depends on the indexed collision points P_i .

With the commonly used notations, we might formulate this expected value integral to be solved by Monte Carlo by the following form:

$$I = \int \dots \int S(P_0) h(P_0) \prod_{i=1}^{\infty} K(P_{i-1} \rightarrow P_i) h(P_1, \dots, P_{\infty}) dP_{\infty} \dots dP_0, \quad (4.23)$$

where the scoring function h can be split into a function defining the domain of the integration and a function that accounts for the magnitude of the score. K , in an analogue, non-multiplying case, is normalised to 1 in the outgoing coordinates, therefore it can be related to the conditional pdf's. Naturally the chains should not be infinitely long, and the particles will terminate after some collisions.

Calculating such an integral by forward Monte Carlo includes sampling from the source function, selecting the next coordinates from the transport kernel, and scoring according to the scoring functions. In an adjoint mode the particles should *start from the termination probability* in the system, and their travel should be followed (sampling the kernel K backwards) and their contribution to be tallied in the scoring function.

Later on we will investigate an equivalent form of Eq. (4.23) where the integral is defined using exactly n -long chains:

$$I = \sum_{n=1}^{\infty} \int \dots \int S(P_0) \prod_{i=1}^{n-1} K_s(P_{i-1} \rightarrow P_i) K_a(P_{n-1} \rightarrow P_n) h_n(P_0, \dots, P_n) dP_n \dots dP_0, \quad (4.24)$$

where $K_a(P \rightarrow P')$ is the transport kernel where the collision results in an absorption:

$$K_a(P \rightarrow P') = T(r \rightarrow r' | E, \Omega) \frac{\Sigma_a(r', E)}{\Sigma_t(r', E)} \delta(E') \delta(\Omega' - \Omega_a) \quad (4.25)$$

with Ω_a an arbitrary direction, and $K_s(P \rightarrow P')$ is the transport kernel where the collision results in a scattering

$$K_s(P \rightarrow P') = T(r \rightarrow r' | E, \Omega) \frac{\Sigma_s(r', E \rightarrow E', \Omega \rightarrow \Omega')}{\Sigma_t(r', E)}. \quad (4.26)$$

This form of the definition of the Kernels relate to the emission density representation of the collision density.

4.2.3 Collision-wise response of the flux

We have stated, that in this generalized description of the response the adjoint calculation (a simulation played according to adjoint kernels) is a reversed calculation starting from termination probabilities across the geometry. First we have to show that there is an expression of $h(P_0 \dots P_\infty)$ that transforms this formulation to the original integral equations, and in case of such a response the adjoint calculation can be started from the detector. The derivation presented here is equivalent to the one in Ref. 32.

For such an estimator as the flux (or a linear functional of the flux), the scoring function h will be chosen according to

$$h(P_0, \dots, P_\infty) = S(P_0) \sum_{j=1}^{\infty} h(P_j) \quad (4.27)$$

When we substitute this into Eq.(4.23) we can write:

$$\begin{aligned} I &= \int \dots \int S(P_0) \sum_{j=1}^{\infty} h(P_j) \prod_{i=1}^{\infty} K(P_{i-1} \rightarrow P_i) dP_\infty \dots dP_0 = \\ &= \sum_{j=1}^{\infty} \left\{ \int \dots \int S(P_0) h(P_j) \prod_{i=1}^j K(P_{i-1} \rightarrow P_i) \prod_{i=j+1}^{\infty} K(P_{i-1} \rightarrow P_i) dP_\infty \dots dP_0 \right\}. \end{aligned} \quad (4.28)$$

Next, we define the function

$$\psi_j(P_j) = \int \dots \int S(P_0) \prod_{k=1}^j K(P_{k-1} \rightarrow P_k) dP_{j-1} \dots dP_0 \quad j = 1, 2, \dots \quad (4.29)$$

to get

$$I = \sum_{j=1}^{\infty} \int \psi_j(P_j) h(P_j) \left\{ \int \dots \int \prod_{i=j+1}^{\infty} K(P_{i-1} \rightarrow P_i) dP_\infty \dots dP_{j+1} \right\} dP_j \quad (4.30)$$

Eq.(4.30) expresses, that the integral is a sum of the collision densities at the j^{th} collision multiplied by the response function. The term in the bracelets express the probability that something will happen to the particle (as it will be formally shown in the next paragraph), and the probability of that is 1. The most important feature of Eq.(4.30) (that will not hold for the PHD), that the score made at the j^{th} collision is not affected by any previous or later scores, and after this collision the particle is free to travel in the whole model geometry without making a further score for the j^{th} collision contribution. This will give basis to start the adjoint particles from the detector.

The joint kernel K for an analogue game is normalized to unity for the outgoing coordinates in an infinite medium. Therefore $\int \int K(P_1 \rightarrow P_2) K(P_2 \rightarrow P_3) dP_2 dP_3 = \int K(P_1 \rightarrow P_2) dP_2 = 1$.

From this we can conclude that the part in brackets of Eq.(4.30) is unity, as $\int \dots \int \prod_{i=j+1}^n K(P_{k-1} \rightarrow P_k) dP_\infty \dots dP_{j+1} = 1$ for arbitrary n . The integral finally takes the form of

$$I = \sum_{j=1}^{\infty} \int \psi(P_j) h(P_j) dP_j \quad (4.31)$$

This form is the same as Eq.(4.14) and Eq.(4.17) combined, and the definition in Eq.(4.29) is equivalent to the application of Eq.(4.5) for higher indices. Such an integral can be solved by adjoint Monte Carlo as well, as stated earlier. Taking Eq.(4.30) and the number of scattering collisions n , we can change the numbering of n going backward from n to zero. Moreover, let us set

$$\begin{aligned} I &= \sum_{j=1}^{\infty} \int \psi(P_j) h(P_j) dP_j = \\ &= \sum_{j=1}^{\infty} \int \dots \int S(P_0) h(P_j) \prod_{i=1}^j K(P_{i-1} \rightarrow P_i) dP_j \dots dP_0 = \\ &= \sum_{l=1}^{\infty} \int \dots \int S(P_l) h(P_0) \prod_{i=1}^l K(P_i \rightarrow P_{i-1}) dP_l \dots dP_0 \end{aligned} \quad (4.32)$$

The last expression gives the adjoint scheme, with samples taken from the scoring function $h(P_0)$ (usually only non-zero in the detector domain), and sample successive coordinates from the kernel K with variables interchanged. The scoring is made by the source function S , and the contributions of a particle should be summed.

4.2.4 Adjoint Sampling of the Pulse Height Distribution

If we were able to define $h(P_0 \dots P_\infty)$ for the PHD, we could basically simulate the reversed Monte Carlo: starting from the termination probability throughout the system. This way of sampling is highly inefficient, and hardly useful: it would basically mean a highly inefficient way of rejection sampling. However, we know that after the last collision of a particle in the detector the remaining random walk will not change the response. We should be able to find a method that allows sampling the adjoint particle density from the detector domain.

The generalised response function for the energy release probability could be formulated regarding n -long scattering chains ending in absorption. The kernels determine the energy release, noted as $\Delta E_i = E_{i-1} - E_i$ at the i^{th} collision. The space coordinate \mathbf{r} determines whether the collision happened in the detector. We will use again the symbol $\Pi_{E_d, \Delta E_d}(E)$ for the function that is zero if E is not within the interval $[E_d, E_d + \Delta E_d]$ and 1 otherwise. $\Pi_{\mathbf{r} \in V_d}(\mathbf{r})$ is zero if the \mathbf{r} is not in the detector domain, 1 otherwise. With these notations the constrain for total energy release in the detector for an n -long chain looks like:

$$h(P_0, \dots, P_n) = \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^n \Delta E_i \Pi_{\mathbf{r} \in V_d}(\mathbf{r}_i) \right) \quad (4.33)$$

We can insert Eq.(4.33) into (4.24), obtaining:

$$I = \sum_{n=1}^{\infty} \int \dots \int K_a(P_{n-1} \rightarrow P_n) S(P_0) \prod_{E_d, \Delta E_d} \left(\sum_{i=1}^n \Delta E_i \prod_{r \in V_d} (r_i) \right) \prod_{j=1}^{n-1} K_s(P_{j-1} \rightarrow P_j) dP_n \dots dP_0 \quad (4.34)$$

Scoring is not troublesome in the forward game, especially if a full, detailed spectrum is to be estimated. However, it becomes harder to bias the game for higher efficiency if only a certain part of the spectrum is important, or to play an adjoint game. The difficulty of the latter lies in the fact, that if we start an adjoint sample in the detector, we could account for h in the response, but this contribution should only be made, if we calculate the probability that a particle travelling forward from the sampled point will not deposit any more energy in its path, including the first escape, and the probability of not returning to the detector. An obviously feasible idea is to start a forward particle at the phase-space position of the adjoint source sample, and let it "finish" the history; naturally its contribution should also be tallied on h together with the adjoint sample. Implementation of such a process needs major adjustments to the normal flow of Monte Carlo simulations, and such a development in common Monte Carlo codes is not expected in the near future. An expected value estimator could replace this forward part of the simulation, as we have seen it earlier, not requiring major alterations in the computation flow of a general Monte Carlo code, but introducing the need for separate expected value estimators. This approach seems the most applicable for practical cases, especially for borehole logging. It remains, though, of interest whether we can limit the sampling to the detector domain using one purely adjoint calculation flow.

Our aim is now to eliminate the calculation of the probability that a particle collided in the detector will never return to it. This can be done by estimating this probability one-by-one starting from the last collision in the history. The last collision of the particle will either be in the detector volume and its energy will contribute to the sum of deposited energies, or will be out of the detector volume and its energy is irrelevant for the detector-reading. It can be expressed as:

$$\begin{aligned} h(P_1, \dots, P_n) &= \prod_{E_d, \Delta E_d} \left(\sum_{i=1}^n \Delta E_i \prod_{r \in V_d} (r_i) \right) = \\ &= \prod_{r \in V_d} (r_n) \prod_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-1} \Delta E_i^s \prod_{r \in V_d} (r_i) + \Delta E_n^a \right) + \\ &+ \prod_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-1} \Delta E_i^s \prod_{r \in V_d} (r_i) \right) (1 - \prod_{r \in V_d} (r_n)) = \\ &= \prod_{r \in V_d} (r_n) \prod_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-1} \Delta E_i^s \prod_{r \in V_d} (r_i) + \Delta E_n^a \right) \\ &- \prod_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-1} \Delta E_i^s \prod_{r \in V_d} (r_i) \right) \prod_{r \in V_d} (r_n) + \prod_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-1} \Delta E_i^s \prod_{r \in V_d} (r_i) \right) \end{aligned} \quad (4.35)$$

where the superscripts a and s are distinguishing the energy release of an absorption from a scattering event, respectively. The first two terms of the three we obtained in Eq.(4.35), express the sampling a point from the detector volume. The third term is a sampling a point

from somewhere in the model region. The third term cannot be sampled thus from the detector domain, and needs to be expressed by other terms. Decomposing it in the same way as for Eq.(4.35):

$$\begin{aligned} & \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-1} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) \right) = \\ & \Pi_{r \in V_d}(\mathbf{r}_{n-1}) \left\{ \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-2} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) + \Delta E_{n-1}^s \right) - \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-2} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) \right) \right\} \\ & + \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-2} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) \right). \end{aligned} \quad (4.36)$$

The last term can be decomposed, yielding a term the same as Eq.(4.36) with the indices reduced by one. This process can be repeated, leading to the following form of the response function:

$$\begin{aligned} & h(P_1, \dots, P_n) = \\ & = \Pi_{r \in V_d}(\mathbf{r}_n) \left\{ \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-1} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) + \Delta E_n^a \right) - \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{n-1} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) \right) \right\} \\ & + \sum_{k=2}^{n-1} \Pi_{r \in V_d}(\mathbf{r}_{k-1}) \left\{ \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{k-1} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) + \Delta E_k^a \right) - \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{k-1} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) \right) \right\} \end{aligned} \quad (4.37)$$

The term in the second line is still not suitable for the adjoint sampling as it contains terms, where samples should be taken from the whole model volume weighed by the absorption probability. Putting this form (second term of Eq. (4.37)) into the transport integral of Eq.(4.24), we obtain the following:

$$\begin{aligned} & \int \dots \int S(P_0) \sum_{n=1}^{\infty} K_a(P_{n-1} \rightarrow P_n) \\ & \times \sum_{k=1}^{n-1} \Pi_{r \in V_d}(\mathbf{r}_{k-1}) \left\{ \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{k-1} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) + \Delta E_k^a \right) - \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{k-1} \Delta E_i^s \Pi_{r \in V_d}(\mathbf{r}_i) \right) \right\} \\ & \times \prod_{i=1}^{n-1} K_s(P_{i-1} \rightarrow P_i) dP_n \dots dP_0. \end{aligned} \quad (4.38)$$

The summation over n is a summation of all possible lengths of chains, the index k expresses the number of collisions until the last event in the detector. The last event in the detector (the k^{th}) is followed by $n-k$ collisions that will not contribute to the response estimator independently of their phase-space positions. For our purpose it is more convenient to change the order of the summations, selecting the ordinal numeral of the last event in the detector (k) multiplied by the sum of terms expressing the lengths of the chains from k to infinity, and also multiplied by the corresponding scattering and absorption kernels. To avoid a very lengthy formulation, and to show the index changes in a simplified form, we set s_0 for the source expression, a_k for the response functions, b_i for the scattering kernel terms using i for the outgoing coordinates, and c_n for the absorption kernel. Omitting the integrations we obtain:

$$s_0 \sum_{n=1}^{\infty} \sum_{k=1}^{n-1} \sum_{i=1}^{n-1} a_k b_i c_n = s_0 \sum_{k=1}^{\infty} \sum_{n=k+1}^{\infty} \prod_{i=1}^{n-1} a_k b_i c_n = s_0 \sum_{k=1}^{\infty} a_k \left\{ \prod_{i=1}^k b_i \underbrace{\sum_{n=k}^{\infty} c_{n+1} \prod_{j=k+1}^n b_j}_{\text{Eq.(4.40)}} \right\} \quad (4.39)$$

The k^{th} term of the response function will therefore be multiplied by the following product of kernels:

$$\sum_{n=k}^{\infty} K_a(P_n \rightarrow P_{n+1}) \prod_{j=k+1}^n K_s(P_{j-1} \rightarrow P_j) \quad (4.40)$$

From k to n the chain of kernels are integrated without constraint for the detector domain, but now the result can be calculated with relative ease. To show this, let us integrate the $n = k$ term over P_{k+1} and the $n = k+1$ term over P_{k+1} and P_{k+2} , using the fact that the kernels are normalized to 1 on the outgoing co-ordinate, and that the absorption probability and the scattering probability add up to 1:

$$\begin{aligned} (n = k): & \int T(r_k \rightarrow r_{k+1} | E_k, \Omega_k) \frac{\Sigma_a(E_k, r_{k+1})}{\Sigma_t(E_k, r_{k+1})} dr_{k+1} \\ (n = k+1): & \iiint \int T(r_k \rightarrow r_{k+1} | E_k, \Omega_k) \frac{\Sigma_s(E_k \rightarrow E_{k+1}, \Omega_k \rightarrow \Omega_{k+1} | r_{k+1})}{\Sigma_t(E_k, r_{k+1})} \times \\ & T(r_{k+1} \rightarrow r_{k+2} | E_{k+1}, \Omega_{k+1}) \frac{\Sigma_a(E_{k+1}, r_{k+2})}{\Sigma_t(E_{k+1}, r_{k+2})} dr_{k+2} d\Omega_{k+1} dE_{k+1} dr_{k+1} = \\ & = \int T(r_k \rightarrow r_{k+1} | E_k, \Omega_k) \frac{\Sigma_s(E_k, r_{k+1})}{\Sigma_t(E_k, r_{k+1})} dr_{k+1} \\ & - \iiint \int T(r_k \rightarrow r_{k+1} | E_k, \Omega_k) \frac{\Sigma_s(E_k \rightarrow E_{k+1}, \Omega_k \rightarrow \Omega_{k+1} | r_{k+1})}{\Sigma_t(E_k, r_{k+1})} \\ & \times T(r_{k+1} \rightarrow r_{k+2} | E_{k+1}, \Omega_{k+1}) \frac{\Sigma_s(E_{k+1}, r_{k+2})}{\Sigma_t(E_{k+1}, r_{k+2})} dr_{k+2} d\Omega_{k+1} dE_{k+1} dr_{k+1} \end{aligned} \quad (4.41)$$

The first term ($n = k$) of Eq.(4.41) added to the first term of the ($n = k+1$)-expression will give 1 as a result, leaving behind a term with both the k^{th} and the $(k+1)$ -th terms are scatterings. If the same procedure is made with the ($n = k+2$)-th integral, we will obtain the same form except the indices are shifted. The remaining term will cancelled out by a part the ($n=k+1$)-th term leaving behind the form where the $(k+2)$ -th, the $(k+1)$ -th and the k^{th} terms are all scatters. This can be applied recursively, and finally the infinite sum of these integrals yields 1.

The transport integral looks like (having changed the direction of the summation, and having repeated the same manipulation with the absorption term):

$$\begin{aligned}
& \int \dots \int \sum_{k=1}^{\infty} S(r_k) \Pi_{r \in V_d}(\underline{r}_k^a) \times \\
& \times \left\{ \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{k-1} \Delta E_i^s \Pi_{r \in V_d}(\underline{r}_i) + \Delta E_0^a \right) - \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{k-1} \Delta E_i^s \Pi_{r \in V_d}(\underline{r}_i) \right) \right\} \\
& \times K_a(P_1 \rightarrow P_0) \prod_{i=1}^{k-1} K_s(P_{i+1} \rightarrow P_i) dP_k \dots dP_0 + \\
& + \int \dots \int \sum_{k=1}^{\infty} S(r_k) \Pi_{r \in V_d} \times \\
& \times (\underline{r}_k^s) \left\{ \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{k-1} \Delta E_i^s \Pi_{r \in V_d}(\underline{r}_i) + \Delta E_0^s \right) - \Pi_{E_d, \Delta E_d} \left(\sum_{i=1}^{k-1} \Delta E_i^s \Pi_{r \in V_d}(\underline{r}_i) \right) \right\} \\
& \times \prod_{i=0}^{k-1} K_s(P_{i+1} \rightarrow P_i) dP_k \dots dP_0
\end{aligned} \tag{4.42}$$

Now the sources of all terms are in the detector domain. For the complete adjoint simulation we have to cast Eq.(4.42) in the form described in Ref.31 using the transition and collision kernels for the adjoint particle. If we look first to the integrand of the first term of Eq.(4.42) for a certain value of k we can write this term after integration over E_0 and Ω_0 as

$$\begin{aligned}
& S(P_k) \prod_{i=1}^{k-1} T(r_{i+1} \rightarrow r_i, E_{i+1}, \Omega_{i+1}) C_s(r_i, E_{i+1} \rightarrow E_i, \Omega_{i+1} \rightarrow \Omega_i) \mathcal{I}(r_i \rightarrow r_0, E_1, \Omega_1) \frac{\Sigma_a(r_0, E_1)}{\Sigma_i(r_0, E_1)} \\
& = \frac{S(P_k)}{\Sigma_i(P_k)} T^+(r_{k-1} \rightarrow r_k, E_k, \Omega_k) \prod_{i=2}^{k-1} P^+(r_i, E_i) C^+(r_i, E_i \rightarrow E_{i+1}, \Omega_i \rightarrow \Omega_{i+1}) T^+(r_{i-1} \rightarrow r_i, E_i, \Omega_i) \\
& \times P^+(r_i, E_i) C^+(r_i, E_i \rightarrow E_2, \Omega_i \rightarrow \Omega_2) T^+(r_0 \rightarrow r_1, E_1, \Omega_1) \Sigma_a(r_0, E_1)
\end{aligned} \tag{4.43}$$

with T^+ the adjoint displacement kernel, C^+ the adjoint scattering kernel and P^+ the adjoint weight factor [16]. To start the adjoint simulation we have to select r_0 uniformly within the detector volume and Ω_1 isotropically. The energy E_1 should be sampled proportional to $\Sigma_a(r_0, E_1)$, but it is more efficient to sample E_1 uniformly from the total energy range of interest and to introduce a weight factor $\Sigma_a(r_0, E_1)$. The functions $\Pi_{E_d, \Delta E_d}$ determine the deposited energy that is recorded during that history and thus determine the bin of the pulse height spectrum to which a score is made. A score is obtained when the adjoint particle has a collision in the original particle source (or crosses the source) with magnitude $S(P_k)/\Sigma_i(P_k)$. Due to the Π term with a minus sign also a negative score must be recorded, but for a different deposited energy, as the arguments of the Π function differ.

The second term of Eq.(4.42) can be sampled almost the same as the first term, except that the start of the simulation is different. The adjoint source term is determined by

$$T^+(r_0 \rightarrow r_1, E_1, \Omega_1) P^+(r_0, E_0) C^+(r_0, E_0 \rightarrow E_1, \Omega_0 \rightarrow \Omega_1) \Sigma_i(r_0, E_0) \tag{4.44}$$

This means that r_0 is again selected uniformly within the detector volume, Ω_0 is selected isotropically, and E_0 is sampled uniformly from the total energy range of interest with a weight factor $\Sigma(r_0, E_0)$. Then E_1 and Ω_1 are sampled from the adjoint collision kernel, and so on.

It is worthwhile to mention that for the derivation of Eq.(4.42) we did not use the fact that the chosen response function belongs to the PHD. The only special feature we applied, is the expression of the response when one less collision is giving the response.

4.2.5 Numerical example

To demonstrate the feasibility of such a game, a simplified form of the photon transport has been considered with only Compton-scattering and the photoelectric effect. The geometric model (see Fig. 27) of the system consists of a 5-cm radius cylinder of 15-cm axis length. The photon source is a sphere of 2-cm radius emitting photons with a flat spectrum between 1.5 and 2 MeV. Its centre is located along the system axis, 5 cm from one outer end of the system. The detector is a 0.25-cm radius sphere with its centre 6 cm away from the other outer end of the system. The whole system has a uniform material composition. The Compton cross section is given by the Klein-Nishina formula with an amplitude of 0.1 cm^{-1} . The cross section for the photoelectric effect is taken inversely proportional to the third power of the photon energy with a value of 0.01 cm^{-1} at 1 MeV.

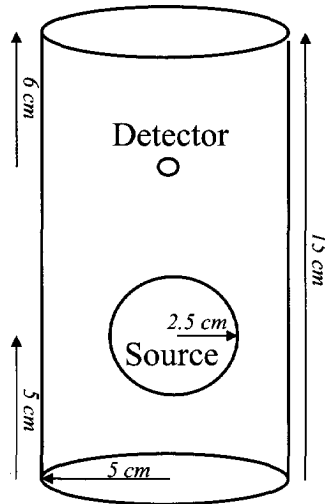


Figure 27. Model geometry for numerical verification

To check the results of the adjoint calculation a program was written to perform a forward calculation for this system. As the probability for a source photon to have an interaction in the detector is low, 10^9 photons were started at the source to obtain an accurate pulse height distribution of energy deposited in the detector.

The adjoint calculation was carried out based on the proposed methods in [16], also using 10^9 particle histories. The results of the forward and adjoint calculations can be seen in Fig. 28. The pulse height distribution is registered in 50 energy bins of 0.04 MeV width up to 2 MeV.

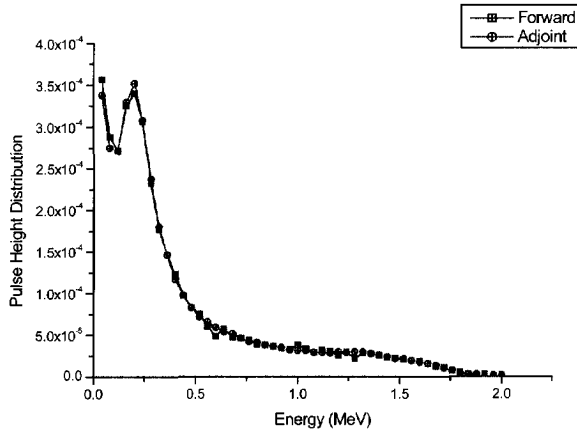


Figure 28. *Adjoint and Forward simulations of a PHD*

The results (Fig. 28) show good agreement between the adjoint and forward calculations, which demonstrates the validity of the theory presented here for adjoint estimation of a pulse height distribution.

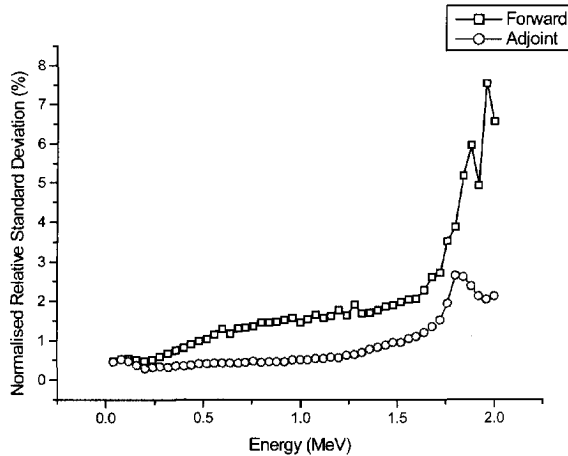


Figure 29. *Comparison of efficiency of the forward and the adjoint calculation*

Figure 29 shows the relative standard deviation in each energy bin, both for the forward and the adjoint calculation. The standard deviations are normalised to a CPU time of 300 min both for the forward and the adjoint calculation. From Fig. 28 it can be seen that the adjoint calculation is more efficient than the forward calculation, mainly because of the difference in detector and source sizes.

4.3 Summary

According to this study, the fully adjoint sampling of the PHD is possible, even though there is not any adjoint equation stated, or Boltzmann detector function given. Out of the three important effects of the PHD calculation only one has been handled. Most importantly the physical branching of pair production is not incorporated in the model, and this should be the first topic for further investigations. A Midway coupling would still be valid, as long as the Midway surface is far enough from the detector. The sampling routine in a general Monte Carlo code could be feasible in terms of incorporating a full photon adjoint Monte Carlo calculation, and this would do all the calculations in the detector domain, overwriting the source energy at every re-exit of the particle.

The partly forward simulation offers the best approximation, as incorporation of pair production is easily done. Incorporation to a general code, however, will meet major difficulties.

Using an expected value function for the detector response needs separate adjoint source calculations before the actual sampling. Its implementation to a general code is relatively easy.

Any of the three methods would also meet another problem, in case of using the MCNP code. The multigroup adjoint calculation in itself would require a very fine energy structure for proper estimation of the PHD.

Chapter 5

EFFICIENCY OF THE TIME DEPENDENT MIDWAY METHOD APPLIED TO BOREHOLE LOGGING

So far we have investigated many details of the Midway coupling. One topic has not been treated yet in detail, namely why the Midway method could serve variance reduction purposes, and how well it would perform for borehole logging calculations. The efficiency of the Midway Method depends on many factors: the quality of the technique of the coupling, the way of sampling, the geometry of the Midway surface, but mostly on the flux and the adjoint function. Apart from these topics, this chapter is also dedicated to a comparison with another variance reduction technique, which is widely applied for Monte Carlo borehole logging simulations.

Borehole logging problems are considered challenging for variance reduction techniques. The practical need for fast simulations (mainly for oil detection tools) has inspired many researchers to improve the methodology of convergence enhancement. Many special-purpose Monte Carlo codes have been written for borehole logging simulations, and many of them used special variance reduction options for borehole logging simulations. Presently the most widely acknowledged variance reduction method is the weight window technique with frequent application to borehole logging simulations [26,35-39]. The second part of this chapter (Section 5.2) will target at the comparison of the Midway method with such attempts, while the first one (Section 5.1) aims at clarifying concepts related to the efficiency of the Midway method in general.

5.1 Optimisation Options

5.1.1 *Determining the Midway efficiency*

A Monte Carlo variance reduction technique, as a terminology, stands for the *efficiency improvement* instead of the literal reading of *reduction of the variance* of the final response. To match the two terms we should interpret 'variance reduction technique' as variance improvement by the same calculation time. In other words maximising the Figure Of Merit. For the Midway method, the score processing increases the calculation time, but no time increase is introduced at the simulation phase. If we can keep the coupling time reasonably

low, the Midway efficiency, when comparing to autonomous forward or adjoint simulations, is mainly dependent on the variance.

We concluded in Section 3.2.3 that the inherent variance of the Midway method is principally given by

$$D^2(I) \approx \frac{1}{N} D_{\wp}^2 \left(\frac{f(x)}{\wp(x)} f^+(x) \right) + \frac{1}{N^+} D_{\wp^+}^2 \left(\frac{f^+(x)}{\wp^+(x)} f(x) \right) \quad (5.1)$$

for N forward and N^+ adjoint starting particles, with f and f^+ integrands (to be later related to the adjoint function and the flux), and \wp and \wp^+ probability density functions, which are used for the sampling. Here we neglected the variance increase caused by the increased number of coupling coefficients as this term is clearly only characteristic for the coupling process.

The number of starting particles N and N^+ only affects the Midway efficiency until we reach the limit of stable statistical estimates. Beyond that, the variance is inversely proportional to the computer time, and the FOM does not decrease when increasing the number of starters.

The functions f and f^+ represent the current and the adjoint function, and are determined by the model parameters of the system. However, when we relate the actual ϕ^+ and J to f^+ and f , we should take into account that the flux and the current are taken on the Midway surface, in a scalar product with the surface normal. The shape and position of the Midway surface, therefore, changes f and f^+ , and can be subjected to optimisation.

The \wp and \wp^+ probability density functions are not known explicitly, but they are usually show similarities to the flux and the adjoint function, respectively. Implicit information is given by the weight of the particles reaching the Midway surface. In case of an analogue simulation, the pdf's are proportional to the flux and the adjoint function, and the weight reduces to one for the forward simulation. The adjoint calculation does not represent physical particles; therefore we can hardly play the adjoint game as an analogue one. It is possible to force the particles to have $w^+ = 1$ by e.g. path stretching.

5.1.2 One dimensional considerations

It is instructive to see how the above-identified factors influence the variance in a one-dimensional case. For this purpose, the integral of $f(x)f^+(x)$ will be calculated using the segmentation technique:

$$I \approx \sum_{k=1}^K \frac{1}{|\Delta_k|} \int_{\wp^+} \frac{f^+}{\wp^+} \Pi_k dx \int_{\wp} \frac{f}{\wp} \Pi_k dy \approx \sum_k \frac{1}{N} \frac{1}{N^+} \frac{1}{|\Delta_k|} \sum_j^{N^+} w_j^+ \Pi_k(x_j^+) \sum_i^N w_i \Pi_k(x_i) \quad (5.2)$$

Two situations are studied here, each sampled according to two different methods. The first situation assumes exponential decrease (e^{-x}) for both f and f^+ on the interval (1,5), while for the second case one of the function changes to an increasing exponential ($e^{-(5-x)}$). The functions have been normalised by their integral on (1,5), to allow handling them as pdf's. For

every case the segmentation was an equally spaced grid. For the pdf's the first method of sampling comprised uniform sampling in x , and the weights have been calculated by simply taking the function value. The second sampling method is similarly to an analogue simulation, applies unit weights, meaning the samples are taken from f and f^* as pdf's. N and N^* were set both to 10^5 . All quantities (segment-wise means and variances) have been calculated according to their expected value analytically. Such a scenario resembles the integral in the radial coordinate, for the Midway borehole logging simulations where both adjoint and forward calculation decreases steeply with the radius. The results for the two decreasing exponentials are shown in Fig. 30.

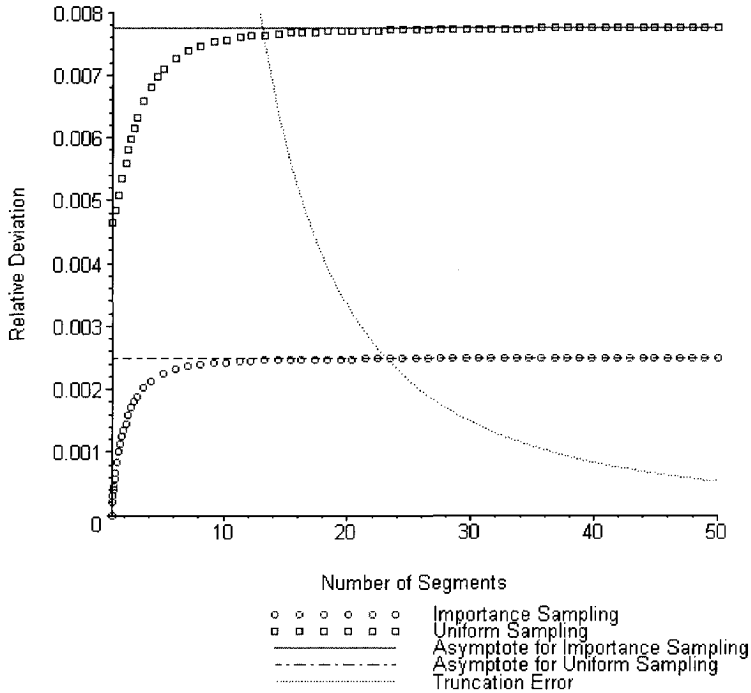


Figure 30. *Relative deviation of the integral of two decreasing exponentials using the segmentation technique, according to uniform and analogue (the latter referenced as "Importance") samplings.*

The asymptotes were calculated using Eq.(5.1). The variance, as stated in Section 3.2.5, indeed depends on the choice of the segmentation structure but there is an asymptotic lower limit value determined only by the integrands and the sampling method. Results of Monte Carlo sampling would differ from Fig.30 when K and N are in the same magnitude (see also Fig. 11); when the Midway integral becomes undersampled for a given number of segments. In case of real simulation results, significantly varying variance estimates indicate either under- or oversegmentation.

If we take only one of the functions to integrate, and apply “analogue” sampling, the variance would be zero (all weights are equal). Hence we indicated this sampling as the importance sampling. We see, that if both functions are sampled like this, it still not ensures a zero asymptotic variance of the estimated I . Uniform sampling would introduce weight fluctuations much like in a biased transport simulation, and from the samples themselves the function and its sampling pdf is not separable directly.

The fact that the variance increases with increasing number of segments is not a general behaviour of the method (as we will see now), nor is the superiority of the “Importance Sampling”. For our second example, f^+ has been changed to $e^{-(5-x)}$, normalised by its integral.

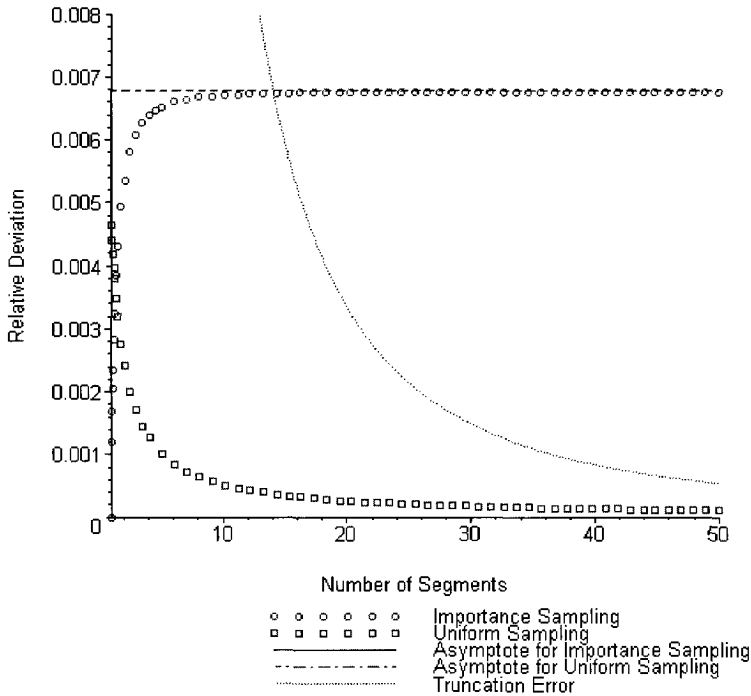


Figure 31. *Integral of the product of increasing and decreasing exponentials using the segmentation approximation. The uniform sampling establishes asymptotically a zero-variance sampling.*

As Fig. 31 shows, now the uniform sampling delivers the better estimate. Moreover the uniform sampling establishes a zero-variance solution. Now the evolution of the variance is decreasing rather than increasing. Such a scenario is somewhat similar of the time dependency of the Midway integrands for a borehole simulation. It is also important to note, that the convergence of the integral approximation, (the decrease of the truncation error), did not become faster by obtaining a zero-variance solution, even though the product of the two integrands are now everywhere constant.

The zero-variance solution would require, in general, the same sampling pdf's for both adjoint and forward processes, namely

$$\wp(x) = \wp^+(x) = \frac{f(x)f^+(x)}{\int f(x)f^+(x) dx} \quad (5.3)$$

as

$$D_{\wp}^2(I) = \int \wp(x) \left(\frac{f(x)f^+(x)}{\wp(x)} \right)^2 dx - \left(\int f(x)f^+(x) dx \right)^2 = \int f f^+ dx - I^2 = 0 \quad (5.4)$$

And the proof is identical for the adjoint function. The respective weights will be different, though:

$$w = \frac{f}{\wp} = \frac{\int f f^+}{f^+} \quad \text{and} \quad w^+ = \frac{f^+}{\wp^+} = \frac{\int f f^+}{f} \quad (5.5)$$

The analogue of these one-dimensional integrals to a real 6 dimensional Midway Monte Carlo simulation fails by taking the pdf's on the whole domain of integration. That leaves out the chance of particles not reaching the Midway surface, or in a one-dimensional analogue, that the pdf ranges beyond the integral boundaries. A zero variance solution in general would require a score from every starting particle.

5.1.3 Response flow and Monte Carlo efficiency

The reason for the efficiency gain using the Midway method must be related to characteristics of the sampling pdf's. *Serov* showed [9], that in a one-dimensional system with purely absorbing medium, the Midway efficiency is varying with the Midway surface position. He assumed that the variance behaves as with analogue sampling of both forward and adjoint calculations, and showed, that the optimal Midway position is halfway the distance of source and detector. Indeed, the population of the samples in the phase space is one of the principal components of the Midway efficiency gain. Particles arriving at the detector must cross the Midway surface first; therefore the amount of particles contributing to the Midway response must be higher than the number of particles arriving to the detector. For analogue sampling, when the variance is strictly related to the number of particles reaching the scoring domain, the Midway method will show lower variance due to particles getting lost (absorption or geometrical spread) after crossing the Midway surface and before reaching the detector.

For a multidimensional system, even with analogue sampling, the amount of scores arriving at the Midway surface does not uniquely determine the variance, owing to the variation in space of the contributions made by different histories at crossing. Still, the amount of nonzero contributions is higher on the Midway surface, than at the detector. The efficiency of utilisation of the information carried by the Midway surface crossing particles is dependent on the coupling scheme, but as we have seen in the previous section, the asymptotic value of the variance is not dependent on the segmentation structure. If we disregard the 'transient'

variance behaviour, we can still find, that a vast amount of particle crossings does not guarantee a highly efficient response estimate.

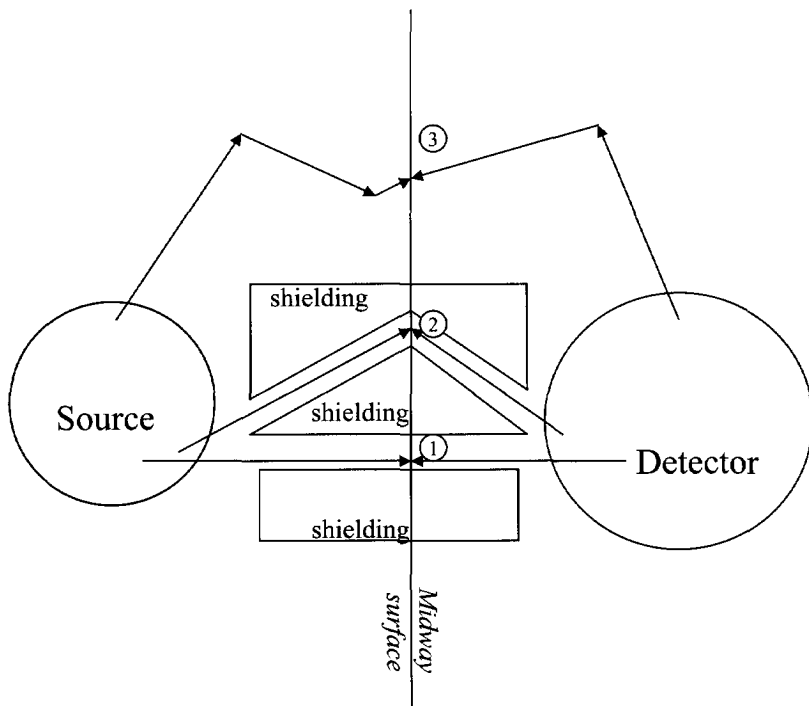


Figure 32. *Sketch of Midway coupling scenarios*

Fig.32 shows some relevantly different coupling scenarios. The sketched particle paths are drawn according to the particle flying direction, and not their solid angle. Scenario (1) is the simplest case to determine the Midway efficiency gain. The particles are travelling in a channel connecting source and detector, and the Midway method is more efficient if the particle elimination probability is high.

Scenario (2) is hardly an effective application of the Midway method. Obviously the low particle transmission probability from source to detector is mainly caused by the bend in the duct, and less by the transmission to this particular phase-space domain. Even though many scores might hit the Midway surface, the adjoint and forward particles would rarely match in angle, in that sense the phase space has a bottleneck. Obviously the highest efficiency increase would be obtained by forcing the particles to take the bend. It might be still advantageous to apply the Midway method, though there is an obvious pitfall for the segmentation structure. A loosely given discretisation structure might erroneously match forward and adjoint scores.

Scenario (3) is a typically good application field for the Midway method: the chance of transmission of particles from the Midway surface to the source or detector is low and the

typical particle paths are not limited to a narrow domain that would decrease the likeliness of possible forward-adjoint coordinate matching.

All three scenarios are sketched in space only, though the other phase-space variables should also be considered. Even though the detector is considerably large, only a limited part of it would receive hits. The size of the detector in itself is not what only matters, though it usually helps an effective Monte Carlo game. All in all, it is not a simple matter of optical width or detector size that determines the efficiency ratios of the adjoint, forward and Midway response estimates.

For a practical calculation, the channels of typical particle flow from source and detector can be identified by the concept of the contributon flow described by *Williams* [3]. A pseudoparticle called a 'contributon' is defined as being "*one of the particles (or its progeny) that manages to survive the journey from the particle source through the geometric configuration and then finally contributes to the response of interest*" [3]. These contributons are a subset of the particle population that successfully arrives at the detector. The contributon transport can be described by the *response flux* $C(P)$:

$$C(P) = \phi(P)\phi^+(P) \quad (5.6)$$

The response flux obeys its special transport equation, and can be subjected to a Monte Carlo calculation [6]. As a consequence of the definition, the contributon pseudoparticles cannot get absorbed or escape from the system. The streamlines of the response flow would originate in the source and end in the detector.

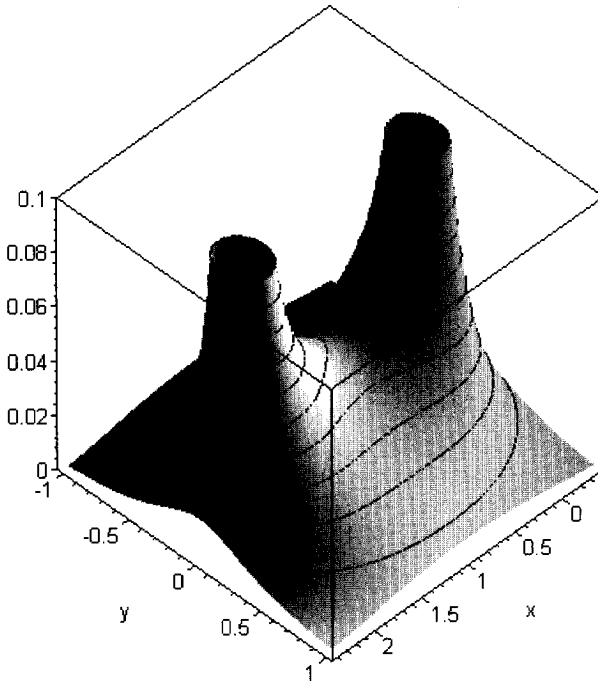


Figure 33. *Response flux for a point source-point detector dipole.*

Using the solution for the flux/adjoint function from diffusion theory for point sources located 2 cm away from each other, the response flow is calculated and plotted in Fig.33 ($\Sigma_t=2.0\text{cm}^{-1}$, $\Sigma_a=0.5\text{ cm}^{-1}$, $\Sigma_s=1.5\text{cm}^{-1}$, $\mu_0=1/3$). The angular dependency has not been taken into account. Around the source and the detector, the flux of the scoring particles are concentrated, while in between source and detector the response flux is 'smeared' out in the phase space. The channel of streaming is limited to a relatively small domain surrounding the source and the detector.

In terms of Monte Carlo sampling, simulating particles outside the response flow channel (i.e. simulating particles that would hardly score) is a waste of computer time; therefore a successful biasing of the sampling scheme should be transporting particles within this phase-space domain. If we take a fully forward analogue simulation, meaning that the sampled pdf is proportional to the flux, most of the particles would not take the response channel. In fact, zero variance results can be obtained by biasing the flux proportional to the adjoint function [41], obtaining an equation identical [3,6] with the contribution flow integral transport equation [see also Eq.(5.3)]. For a zero-variance solution every history should deliver the same contribution, and this means (if we do not want to distort the particle population in the

phase space) that the flow of causative particles (of particles that contribute to the response) should be sampled. The neglected non-scoring particles do not make a difference, though to keep the game fair, the possibility of non-scoring should be accounted for. More precisely, the expected value of the flux should still be kept unbiased at every phase-space position, achieved by an adjusted weight depending only on the phase-space position. If the transport equation for $C(P) = \phi(P) \phi^+(P)$ is sampled, the weight must be given at a phase space point by $w(P) = 1/\phi^+(P)$, or rather just proportionally [41,38]. The cited references give a detailed account of the exact formulation.

Phase-space domains where causative particles travel form therefore the essence of the transport, and those domains must be fully and adequately sampled. This fact is independent of the sampling scheme, or type of variance reduction technique. Also impact of contributions made on the Midway surface would only be high as long as they are in the response flow domain, and such mismatch of adjoint and forward scores as by scenario 2 of Fig.33 does not occur. Unbiased (not more than applying implicit capture) transport sampling matches a response flow more adequately, if the response flow broadens, as the chance of particles travelling in the right direction is higher, while a narrowing channel is less frequently sampled. For the point source-point detector geometry the response flow is the broadest in the middle, while passing mid-distance, the particles should converge, and this does not match at all the unbiased sampling paths. Low variance Midway surface positions should be therefore in that region.

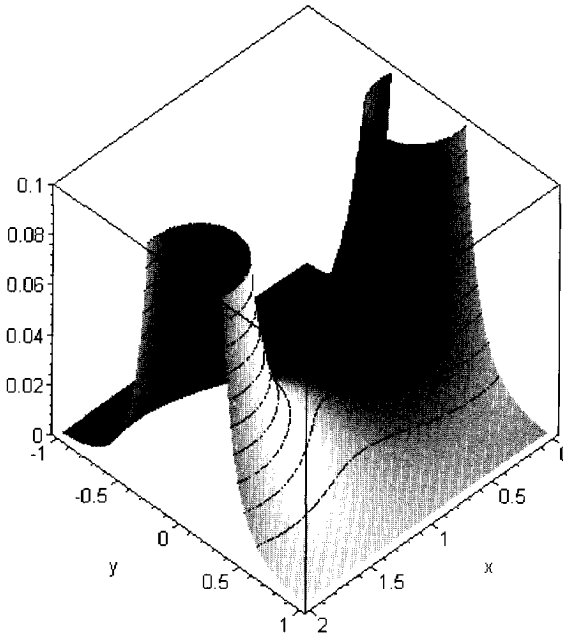


Figure 34. *Quantitative distribution of the variance of the Midway coupling*

Fig. 34 shows the distribution of the phase space points where the second moment of the Midway coupling -assuming analogue sampling- is low (given by $\phi^2\phi^+ + \phi\phi^{+2}$). Omitting the angular dependency is a serious approximation, therefore constant variance Midway surfaces only qualitatively agree with the constant lines of the second moment. This is most significant around the respective adjoint and forward detectors, where the particle directions will least likely match. As accounting of the segmentation structure on a surface of sophisticated shape would be rather cumbersome in practice, such investigations have low practical relevance and it will never be a single closed surface regarding all of phase space variables. On the other hand, choosing the Midway plane coincident with constant flux or adjoint function surfaces might ease an accurate coupling.

5.1.4 Contribution flow for a time dependent borehole calculation

We can extend the theory for the contribution transport to time dependent cases. Defining the same way as for other variables the contributions as the particles scoring in a time dependent detector in a specified time interval, would allow plotting the spatial channels where they travel. For the Generic tool model stochastic adjoint and forward calculation results have been combined to estimate the response flux. The geometry has been divided into small cells, and the adjoint and forward results have been integrated with respect to angle and energy. Given the geometry has rotational symmetry; where also the "spatial" angle of the cylindrical coordinate system has been integrated out. For all 20 time readings (50 μ s long each) the response flux has been calculated. The response flux of the first (0-50 μ s) and the 4th (200-250 μ s) detector reading can be seen by an r - z plot on Fig. 35 and Fig. 36 respectively; for both cases for the near detector. The response flux is calculated as the product of the neutron and the photon response fluxes obtained from cell-wise track length estimators. Lighter shade indicates higher response flow, except at the periphery, where far from source and detector the response flow is extremely low (<0.005% of the maximum value).

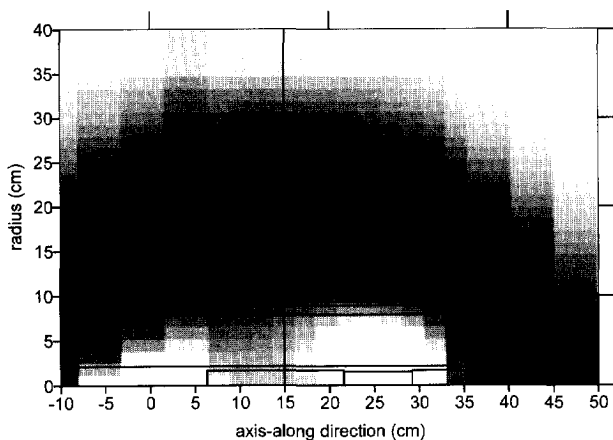


Figure 35. Response flux for the 0-50 μ s time bin

In the figures the boundaries of the model geometry are indicated. The source is located at (0,0). The boundaries of the casing, the pressure housing, cementing and the borehole

constitute the horizontal lines, while the superimposed boxes are the near shielding, the near detector and the near phototube. The vertical line is the position of the Midway plane for most calculations.

Fig. 35 resembles largely a dipole system. The response flow decreases strongly with the radius, and it is localised into a small domain, mainly to the borehole itself. This is a consequence of the small time allowed for the scores to travel: as the contributions should arrive within $50\mu\text{s}$. This would allow only a few neutron scatterings. The transport of contributing particles is localised mainly to the borehole. Recalling that the gamma photons from inelastic scatterings arrive during the source burst, we can say, that the neutrons are mainly creating photons in the borehole, as long as the tool is centred to the middle of the hole. The tool should be definitely pushed to the borehole wall to yield useful information on the formation, but even then the majority of the contributions are not passing through the formation. The slim tool designs allowing operation in secondary tubing in the borehole are strongly separated from the formation, and the sensitivity of the tool for formation constituents is small, at least not for the near detector. In practice, therefore, the near detector response is usually used for obtaining correction factors that measures the effects of the borehole itself.

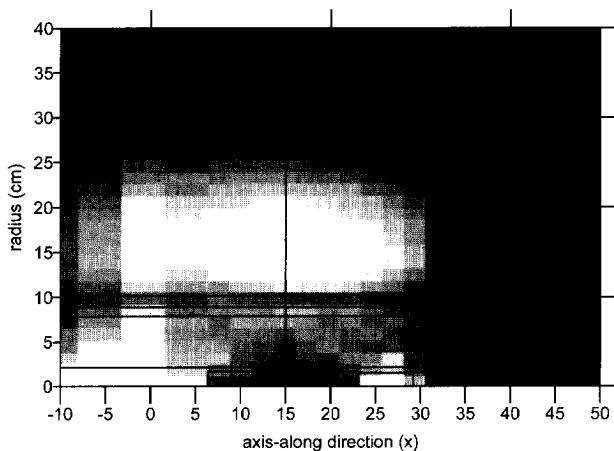


Figure 36. *Response flux distribution for the 200-250 μs time bin*

As Fig.36 shows, the contribution transport for the 4th time bin is not happening any more in the borehole. The particles penetrate the formation, and (as seen in Chapter 1, Fig.5) the detector reading starts to reflect the formation properties. The iron casing at the borehole wall is clearly an effective shielding for the response flow.

The consequences of the response behaviour for the time dependent Midway method are as follows. Regarding a Midway plane positioned at 15cm on the axial coordinate, the response distribution shows different effective domains. Therefore a segmentation system that suffices for higher time readings, might fail for the first time response intervals. The response flow is broader closer to the detector, therefore an effective scoring from both calculations might optimise closer to the near detector. Repositioning the midway plane should also affect the

optimal settings of the segmentation system. The effective scoring domain on the Midway plane moves closer to the axis for positions close to the detector or the source. The angular dependency of the adjoint function, however, will be more anisotropic and might narrow the effective scoring domain.

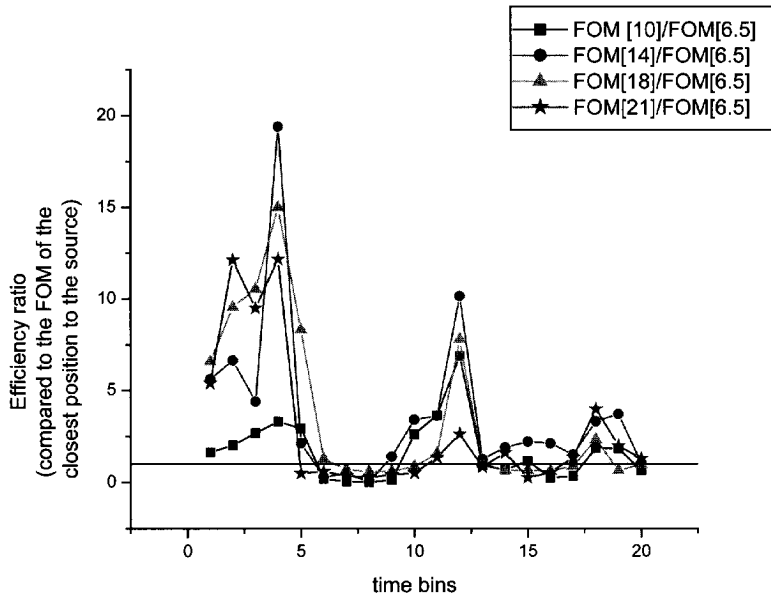


Figure 37. Efficiency change of the Midway method for various Midway plane positions

Fig. 37 shows the efficiency variation for different Midway plane positions. The Midway planes were all positioned perpendicular to the geometry axis, at 6.5, 10, 14, 18 and 21 cm. The FOM-values were normalised by the FOM of the calculation with the Midway plane closest to the detector. The segmentation structure was kept the same. The highest efficiency for the first time bins is achieved by calculations at least halfway source and detector. For higher time bins the efficiency does not change consistently, due to the loose variance estimates (old variance model). It is however apparent, that a Midway plane position (and segmentation) optimal for one time reading, will be less efficient for other ones. In the meantime, the truncation error is also changing for the different Midway planes as shown in Fig. 38.

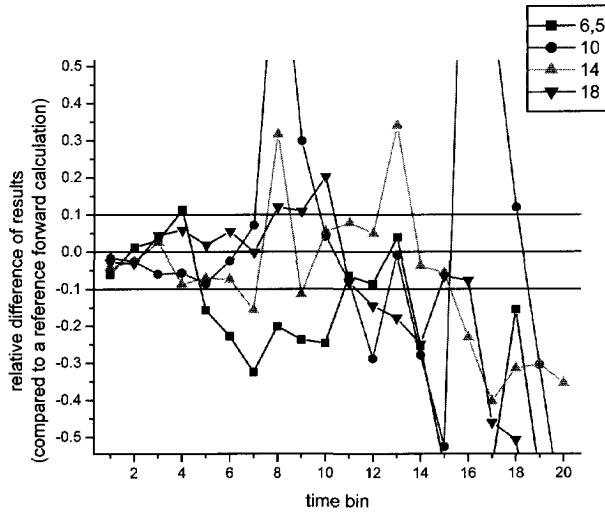


Figure 38. *Midway accuracy for different Midway plane positions*

The least stable and correct estimates can be seen with Midway positions near the source region (6.5, 10 cm).

5.2 Enhancement of the Midway Efficiency Using the Weight Window Technique

The difficult particle transport calculation problem that a borehole logging geometry constitutes, does not only give a problem to the time dependent Midway method. Many variance reduction attempts required additional techniques for significantly improving the efficiency. Borehole logging geometries are typical test problems for general variance reduction techniques. This section is dedicated to a comparison of the so-called Weight Window (WW) technique applied to the same borehole logging geometry with the Midway method. We have seen so far, that the time dependent Midway method has many free parameters to set, and many factors have an influence on its accuracy and efficiency. For the comparison of the Midway method with the WW technique, we have selected one setting that was tuned to deliver accurate results.

5.2.1 The weight window technique

The WW technique [12, 32] is a general variance reduction technique. It achieves an efficiency increase by controlling the particle population travelling through the model geometry. In MCNP the phase space is subdivided into a number of cells, where the controlling parameters are set energy- or time-dependent. For each cell a lower and a higher WW boundary should be specified. When a particle arrives to a cell boundary, its weight is tested against the upper and lower bounds. If the particle weight falls between the two, there

is no action taken. If the particle weight (w) is lower than the lower bound (w_l), Russian Roulette is played: with a probability w/w_s the particle survives and its weight will be set to w_s ; while with probability $1 - w/w_s$ the particle is terminated. The choice on w_s should fall within the WW bounds. If the weight is higher than the upper bound, the particle is split into fragments. From that point on the fragments would walk separately, with reduced weights. The splitting procedure should be made to have all fragments' weights falling within the WW bounds.

For setting the WW bounds properly, we can say, that particles having travelled through the geometry and successfully arrived at the neighbourhood of the detector, should be split to ensure scoring (by at least one of its fragment progeny) thereby avoiding to lose the invested computation time. The amount of splitting should be therefore roughly inversely proportional to the distance to the detector, and this distance might be measured by the importance function. The amount of splittings should be limited by the computer time spent on following the created progeny. If the particle splitted too heavily, many fragments may reach the detector. These progenies belong to the same sample, and the delivered score is the sum of their contributions to the detector response, therefore the same final contribution could have been achieved by following fewer particles.

When particles travel far from the detector, their population should be reduced by Russian Roulette to decrease the wasted computer time. This thinning process can only increase the variance of the final estimate but it effectively decreases the computer time. The further away we are from the detector (the distance is measured again by the importance) the higher the lower WW bound should be set. In MCNP, the lower weight limits are set to half of the response/importance value (more precisely the importance at the source divided by the importance at the given location), and 5 times this value gives the upper WW bound. Experimental studies indicated that the efficiency gain is sensitive to the value of the lower limit [12], and less sensitive to the ratio of the upper and lower limits. The WW technique is particularly useful if there is weight change introduced to the system e.g. by applying implicit capture.

The required importance values can be obtained by deterministic calculations, or it can even be estimated from a forward Monte Carlo calculation. The latter is the so-called WW generator of MCNP, which estimates the importance of a cell by dividing the number of particles entering the cell by the number of them later scoring in the detector. The WW generator can hardly deliver stable estimate for the energy/time dependent cell importances until the calculation is efficient enough to deliver many scores.

For applying the WW technique for Midway calculations, we need to establish the theoretical framework for the calculation of relevant importance values. Unfortunately setting the WW bounds is more practice than theory, as it is not a direct approximation of a zero variance scheme, though *Haghighat et al.* [38] presented a derivation of the WW bounds connected to zero-variance schemes, by the name of Consistent Adjoint Driven Importance Sampling. Their attempt is the same as the above-presented argumentation of the zero variance schemes but starting from the other end of the line of thoughts. If we ensure that the particle weights are proportional to the inverse of the importance, a fair biased game should compensate this by an increased neutron population that must resemble the population of the causal particle

population. In somewhat mathematical terms, contributions to the detector response coming from a phase space domain around P can be approximated by

$$R(P) = \phi(P) \phi^+(P) \quad (5.7)$$

where the argument P means the response R contributed by those particles that have passed through around P . If every particle starts at the detector with unit weight, and the low variance estimate should comprise particles with roughly the same weight, the particle weight should decrease during its journey. If a particle arrives to a phase space point with weight w and the probability of scoring is $\phi^+(P)$, and the process of transporting particles from P to the detector is an ideal process that will not induce further fluctuations in the weight, the particles should arrive to the detector with weight $w\phi^+(P)$. (Such an ideal process can be estimated by splitting the particle into very many fragments for example). The final average should equal the total response R , and for low variance all contributions should equal to that value; if Eq.(5.7) holds, and the probability of the contribution is independent of the weight at P it will impose a constraint on w , that is

$$w(P) \sim R / \phi^+(P) \quad (5.8)$$

This argumentation hardly holds around and inside the detector. Even if every particle at a phase space position at the boundary of the detector arrives with equal weights, the variance introduced by simply the scoring process can be significant, and it is not improved by weight control in the detector. Also, this theory does not account for multiple scores of the same particle history.

With the WW technique it can be ensured, that particles in a phase space region possess a weight around $w(P)$, if we set the weight window bounds around $R/\phi^+(P)$. However in this sense the WW technique is far from being ideal, split fragments must vanish but one of them before reaching the detector to indeed score finally with weight around R . When the weight decreased below R , the technique of statistically ending the particle history, does not reduce the variance any more. The population of the particles is controlled therefore, and indeed the particle population must resemble the distribution of the response flow, though history-wise. The parameters to tune experimentally are numerous: normalising the weight bounds at the source might achieve a better "importance sampling", the saved computer time by Russian Roulette is dependent on the full weight distribution at a phase space point and the expected computer time of a history; for splitting we should avoid superfluous fragments. Nevertheless the technique works very well in practice.

Rigorous optimisation that takes into account the computer time per history is hardly possible. A treatable approximation is to set the computer time proportional to the number of collisions. At a certain phase-space point the drawback and advantage of manipulating the particle population is also dependent on the number of collisions before and after passing P , and such a simple formulation as Eq.(5.8) will not be a good constraint for optimisation. A measure of the average number of collisions per history could at least serve as a guideline for the population control, but it is hard to estimate.

Neutron-photon transport has many special characteristics that might be utilised for a more effective bias that suffers the lack of proper formulation. Given the far lower computer effort

required for gamma than neutron calculations, a population control that increases the number of simulated photons, will not increase the total running time proportionally.

The MCNP Weight Window Generator (WWG) [41] estimates the importances for a forward calculation from a forward stochastic process. The geometry is subdivided into spatial and energy-/time-dependent segments. Particles passing through such a subdomain are registered, together with their final score. The importance is calculated as the total score relative to the weight of a particle entering a subdomain divided by the entering particles, in line with the interpretation of the adjoint function as the probability of scoring from a certain phase-space location. The importance values are later normalised to ensure that particles starting at the highest importance source cell would not be immediately splitted or rouletted out. If the source is subdivided into energy/time intervals, and the source energy/time spans over more than one bin, the importances are normalised according to the maximum importance in an energy/time interval. This latter is not in line with the normalisation constrain of (5.8), as it is not a source-weighted average of the different importance values.

Values are only generated by the WWG if scores are made in the detector. To solve this problem an iterative scheme is proposed [12, 42], or replacement of the actual source by an arbitrary source somewhere closer to the detector. The latter would require a proper normalisation according to Eq.(5.8), and that would require some approximation of the final outcome. It is worthwhile to mention, that there is currently many research efforts directed towards automatising the weight window generation through deterministic codes, that would require minor user input (12,39,43, etc.).

5.2.2 Importance functions for the Midway response

To calculate the importance function for the WW application for a Midway calculation, we need to define first the importance function for the Midway response parts. The respective detector function belonging to the Midway adjoint calculation is given by

$$D_M^+(P) = \underline{n\Omega}\phi(P) \text{ while } P \in A_m \quad (5.9)$$

The detector function for the forward Midway calculation is

$$D_M(P) = \underline{n\Omega}\phi^+(P) \text{ while } P \in A_m \quad (5.10)$$

For the forward calculation, the weight windows should be generated using the adjoint function ϕ_{FM}^+ belonging to this detector response, where D_M functions as the adjoint source. Taking a Midway enclosure around the source, according to the second theorem of Section 1.2.5 of Chapter 1 (while applying it to an adjoint calculation), $\phi_{FM}^+ = \phi^+$ within the enclosure.

For the adjoint calculation, the importance of the adjoint Midway calculation should be calculated, an adjoint to the adjoint calculation $(\phi_{AM}^+)^+$, with the source term of Eq.(5.9). The adjoint of the adjoint function will give the flux, and using the same theorem as above, the importance for accelerating the adjoint will be equal to the flux of the real source, within the source enclosure. These equalities are shown in Fig. 39

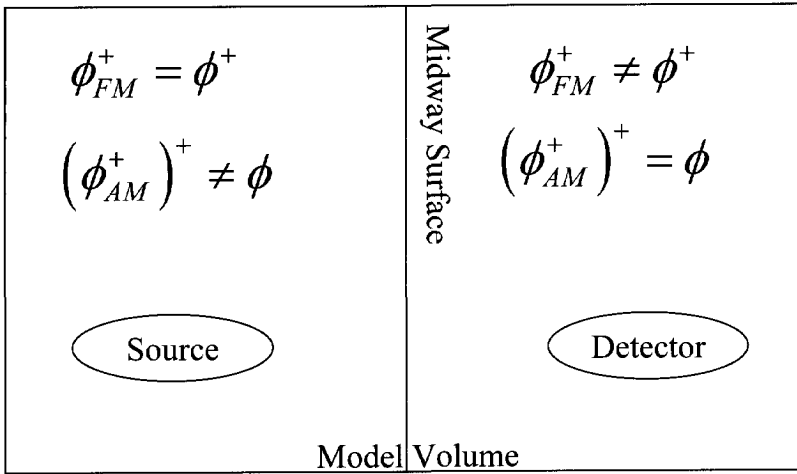


Figure 39. Importance equalities for accelerating the Midway adjoint and Midway forward calculations.

When applying the black absorber technique (for the adjoint calculation), the flux on the source-side domain will vanish. In the detector enclosure the adjoint function would change (as a result of changing the geometry) to $\tilde{\phi}^+$, while we know that disregarding the difference ($\tilde{\phi}^+ - \phi^+$) does not influence the Midway response. Therefore the importance of the adjoint will still be given by the flux corresponding to the real source.

When using a deterministic code to generate the importances, it can be rather troublesome to use Eq.(5.9) and Eq.(5.10) for the source terms, especially when using the black absorber technique. The above equalities allow using the solutions of the original source-detector system, except for the domain outside the Midway enclosure; there special treatment is needed. This special part of the Midway importance regards particles that crossed the Midway surface three times, and the likelihood of such events is relatively small. The WW technique would hardly offer an efficient biasing for those particles as this case is a heavily direction-dependent transport.

The MCNP WWG may be set up targeting at the original source and detector functions. In order to increase the efficiency of the stochastic weight window generation, a simplified Midway coupling procedure has been incorporated into MCNP by a single modified subroutine. From previous forward and adjoint simulations first segment-wise estimates of the Midway detector functions are calculated. The respective Midway detector function is used by the coupling subroutine for scoring. The subroutine returns a series contributions for all of the time bins. It is rarely the case in the general practice that a tally gives a contribution for both photons and neutrons in common Monte Carlo calculations. Here a cell-heating estimator has been utilised, defined in an extremely thin cell. This realisation of the Midway coupling differs from the previous procedures in three major points: it uses less segments, as the data has to share its memory needs with MCNP's; it cannot produce the response functions for the coupling; and the variance calculation is not proper. This realisation allows

an efficient WW generation by increasing the scoring samples and it offers a solution to the needed special treatment outside the Midway enclosure for the non-black absorber part of the Midway calculation.

Two applications of the Midway method have been considered for a comparison of the WW technique with and without the Midway method: time dependent detector reading (Sigma operation mode) with time dependent weight windows and with the WWG; and an energy dependent response (C/O) mode with energy dependent importances generated by the DORT finite elements code.

5.2.3 Performance of the WW technique with time dependent weight windows

When using the MCNP WWG for both the autonomous forward calculation and for the Midway response estimate, setting an equal challenge, i.e. devising a fair comparison in every respect, is hardly possible. Especially when generating the importance map, the quality of the weight windows is not straightforward to compare. Also, comparison with the autonomous unbiased forward calculation might be unfair as the unbiased forward simulation estimates multiple detector responses in the same calculation. The simulations with WW will be optimised to a single response, and a well behaving estimate most certainly means a spoiled estimate for the other responses. A Midway response estimate for two physical detectors needs at least two adjoint and one forward calculation. The time-dependent Midway response estimate provides efficiency increase for many detector responses, and selecting a single optimisation target would deteriorate the variances of the estimates for other responses. A fully optimised speed-up should regard therefore all the information that would be later extracted from the measurements, i.e. it should regard the technique of the data inversion process. It should be emphasized, that the series of time responses are usually tallied without an energy spectrum. Replacement of the detector response by a Pulse Height Distribution estimator is not necessary. Further, the inversion process usually utilises the rate of decrease rather than absolute values of the time dependent readings, therefore this estimate is very close to a realistic simulation except for using multigroup cross-sections.

Our choice for the comparison targets are the near detector and the 5th (200-250 μ s) time bin. For this time bin, the reference autonomous forward calculation provides a statistically stable estimate, but it is a relatively ineffective calculation. This provided a sensible comparison of the FOM's relying on well-estimated quantities, while the inefficiency of the calculation gave space for efficiency improvement. We have selected a segmentation structure, which provided a fairly accurate response estimate, and an original Midway Efficiency Gain of 7.8. The Efficiency Gain is the ratio of the FOM of the simulation in question and the FOM of the reference autonomous calculation.

Knowing the answer from the long-run forward reference calculation allowed the replacement of the physical source for the WWG by a somewhat larger source. The source has been set to a half centimetre thick cylinder around the real source, with a radius of 15cm. Fig. 40 shows the generated importance map, for the forward Midway, non-black-absorber part. The response flows mainly through the formation ($r > 10$ cm). The importance decreases steeply after the Midway surface. Note that scores are only made if the forward particles cross the Midway surface in the direction from the source. To direct the comparison towards comparing

methods, rather than comparing the ease of their optimisations, the full forward weight windows have been generated using 10^9 particles, that would already allow for the 5th time reading to have less than 2% relative error without any biasing.

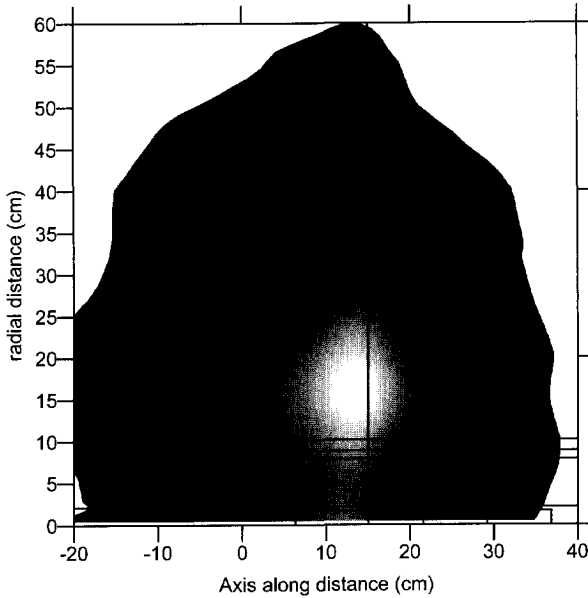


Figure 40. Normalised neutron importance map for the speed-up of the forward Midway (non-black-absorber part) for the 5th time bin, with the model geometry superimposed. The darker shade indicates lower importances, very low values are wiped blank at the periphery.

The same importance map is shown on Fig.41, but with the WW mesh system superimposed.

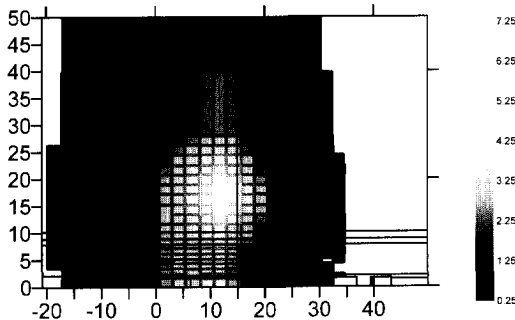


Figure 41. Normalised neutron importance map for the speed-up of the forward Midway (non-black-absorber part) for the 5th time bin, with the WW mesh geometry superimposed. The darker shade indicates lower importances, very low values are wiped blank at the periphery.

The importance values are taken as the inverse of the set WW lower bounds. The results of the calculation with the WW technique are shown on Fig. 42.

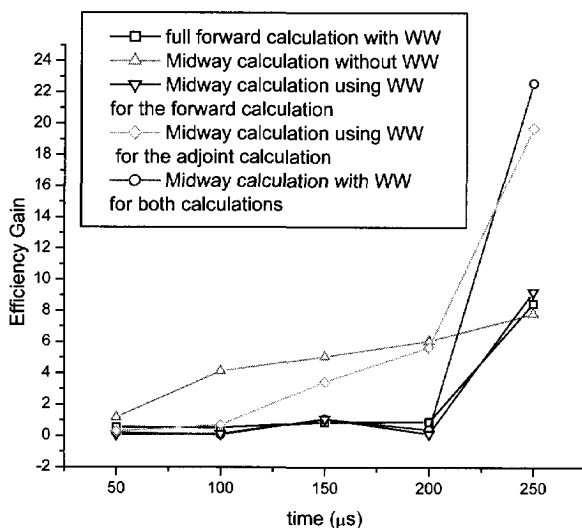


Figure 42. Efficiency Gains (ratios of the Figure of Merit with a forward calculation not using WW) for different optimisation schemes, all targeting the 5th time reading. The unoptimised Midway calculation is almost as efficient as a regular forward calculation with Weight Windows, but shows efficiency increase at other than the targeted time bins.

The efficiency of several optimisation schemes is compared in Fig.42 to an unoptimised forward calculation. The Midway method without the WW technique shows almost the same efficiency increase as the WW technique for the targeted (5th) time bin. For lower bins, the time dependent Midway method shows a considerable efficiency gain. When applying the Midway method together with the WW technique, the efficiency gain increases for the targeted bin, and radically decreases for other responses. A successful application of the WW is applying it to the adjoint Midway calculation, achieving an efficiency gain of a factor of 19.7. This efficiency increase improves further (22.6) by applying the WW technique also to the forward Midway calculation. This last step finally degrades all other response estimates. The least successful biasing scheme belongs to the forward Midway calculation, when the improvement is slight. The reason to this disappointing performance of the WW technique most probably lies in the unsatisfactory treatment of the particles having crossed the Midway surface. These possible contributors should recross the Midway surface for scoring, meaning a strong angular dependency requirement that is not present in the current WW realisation. Regardless of their angle the particles are kept alive though travelling in the wrong direction (further from the Midway surface), though their transport is hardly worth the calculation time. On the other hand they are split into many fragments when travelling back to the surface, while they cannot score until they turn back again. With some form of an angular biasing this unsatisfactory performance could probably be improved.

5.2.4 Performance of the WW technique with energy dependent deterministically generated weight windows

The energy-dependent WW application to borehole logging with and without the Midway method is the subject of this section. Energy-dependent responses are handled only, time dependence is neglected. This gives a simpler Midway integral, where decreasing the number of dimensions offer a considerably more accurate and effective coupling. Also, the required computer memory requirements decreases, allowing a more detailed coupling. This application aims at estimating a somewhat realistic C/O ratios, except neither the Pulse Height Distribution nor the inelastic gamma separation has been considered.

Instead of the previously used MENDF5/6 [44, referring to 45] cross section library, another data set has been used, that was specially designed for borehole logging. This evaluation [46] comprised a complete set of elements that are frequently present in borehole logging environment, with a more detailed energy group structure (175 neutron, 45 gamma groups). This data library based on the ENDF/B-VI (version 8) using the VITAMIN-J structure extended by 9 gamma groups in the range of the Carbon and Oxygen windows.

The absence of time dependency allowed a stable Midway estimate significantly less dependent on the segmentation structure. The Midway coupling in energy relied on the multigroup structure of the used library both for photons and neutrons. The importance function was generated using the discrete ordinates code DORT [47] in a 2D r - z geometry using the same cross section library. The energy structure of the WW settings used the complete 175/45 group structure. For a full forward or adjoint calculation this process can be achieved without further theoretical development to deterministic source-detector calculations. For the Midway method, the generated importances must be corrected beyond the Midway plane (from the perspective of the source). Such an input for DORT is not straightforward; as such an adjoint source should be specified on the Midway surface that is limited to positive direction cosines with regard to the surface normal. Instead of following this path, the MCNP WWG has been invoked to generate the importances beyond the Midway plane, and only in a corrective sense, i.e. replacing the importance values generated by DORT at positions beyond the Midway plane. For this WW generation step, the existing DORT-importances have been used to accelerate convergence. Again the simplified Midway coupling subroutine has been used. For utilisation of these DORT weight windows minor modifications have been made to MCNP. The simplified Midway coupling within MCNP naturally required a previous Midway calculation.

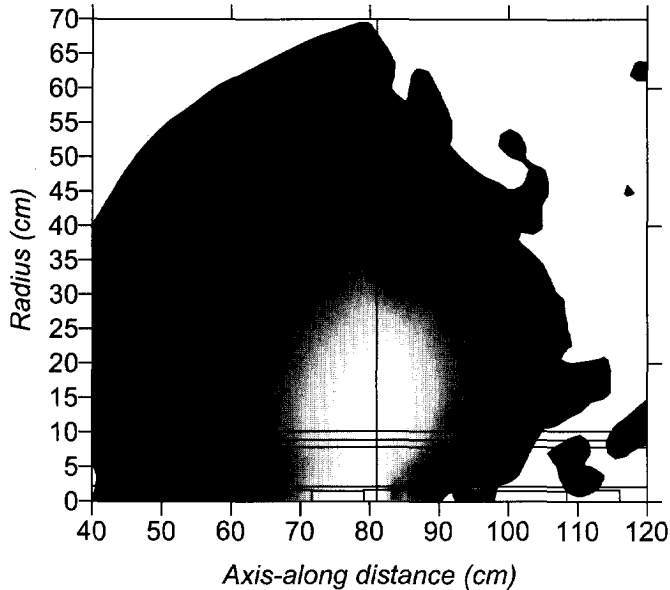


Figure 43. *Neutron importance map for the forward Midway calculation (80th energy interval): stochastically and deterministically generated importance maps are merged at the Midway surface. The targeted response is the far detector (94-109cm axial direction) carbon window. The neutron source is located at $z = 50.8\text{cm}$. Very low importance values are wiped blank at the periphery.*

Fig. 43 shows the neutron importance map for the forward Midway calculation for the far carbon window merged with a stochastic importance generation beyond (right-hand side) the Midway plane. This map belongs to the 80th neutron energy interval of the multigroup library (0.166-0.174MeV). The two maps reasonably match at the Midway plane, though the stochastic part shows higher values in general than expected. Its quality is determined by the statistical nature of the importance generator, giving an inconsistent spatial distribution at many positions.

The investigated responses are the near and far detector readings in the Carbon and Oxygen windows, forming four different responses. The Midway calculations could be set up using a single Midway plane located between the source and the near detector. However, for higher efficiency, two different midway planes have been set separately for the two detectors. The near Midway surface was placed 15cm, the far Midway surface 31cm away from the source. If we choose the adjoint calculation for applying the black-absorber technique, the unoptimised Midway calculation requires one forward and four adjoint calculations. When applying the WW technique, separate WW settings should be applied to each of the detector responses. This means fundamentally different importance maps for the optimised Midway forward runs. For the adjoint calculations, given a single source, the respective Midway importances are provided by the flux, belonging to the same, the actual physical source. Therefore the importances for the responses only differ in a normalisation factor, given by the respective response. To decrease the total workload, only one optimised forward calculation

has been performed, with importances belonging to the far carbon window. This is obviously not optimal for every response, but will the least deteriorate the efficiency of all the estimates.

As mentioned in Section 5.2.1, increasing the number of generated photons might be advantageous for the overall efficiency. In the standard MCNP version at every neutron collision a photon weight equal to $\sigma_{\text{prod}}/\sigma_t$ is generated. This weight will normally be (much) less than unity and a Russian roulette procedure is used to determine whether a photon is obtained with weight equal to the neutron weight at the collision or the photon will be eliminated. As the CPU-time for processing a generated photon until the end of its history is in general much less than that of a neutron, it is a major disadvantage that the Russian roulette for photon survival works very harsh due to the generally small fraction of photon weight generated. It would be better to use the photon importance (when available) to determine what value of the photon weight should be accepted in the Russian roulette procedure. As importances are not always available and, more importantly, the importances do normally not account for variations in the CPU-time expected for the rest of the particle history, we introduced an additional option in MCNP which allowed for the entrance of an input multiplication factor for the generated photon weight at a neutron collision. If this multiplication factor is chosen, for instance, as 5, the generated photon weight will be 5 times as large and (if still smaller than 1) the Russian roulette works 5 times less harsh. Therefore, many more photons will survive the generation process and will contribute to the photon detector response. Based on some test calculations, both a factor 10 was used for both Midway and autonomous forward calculations.

Another possibility of increasing the efficiency of the Monte Carlo calculation is to use the photon importance values to select the energy of the photon. This is possible when a rather fine distribution of (photon) importances is available as obtained from a deterministic code like DORT with a sufficient number of energy groups. An analogous situation is the adjoint neutron generation from adjoint photons, although the procedure for generating adjoint neutrons in MCNP differs somewhat from the above described procedure for generating ('forward') photons in a neutron collision.

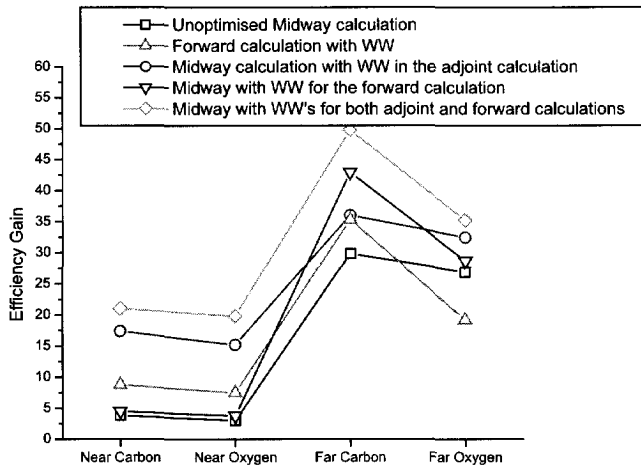


Figure 44. *Efficiency Gain with different energy dependent optimisation settings. (The lines connecting symbols indicate matching values, rather than continuous quantities)*

The efficiency gains (compared to an unoptimised forward calculation) are shown in Fig. 44. The Figure of Merit values have been calculated separately for each response. Calculation times for all detector responses neither in the case of the Midway nor in the case of the autonomous forward calculation have been considered. Now the WW application is significantly more advantageous than applying the simple Midway method. The WW settings of the forward Midway run are not optimal for every reading, and it only aids the Midway calculation at the far carbon window. Acceleration of the adjoint Midway calculation is again the most successful, with efficiency gains above the full forward calculation. We should keep in mind, that the forward Midway calculation is not fully optimised for each response, though the Midway calculations is at least a factor 2 more efficient than an optimised full forward calculation. The efficiency gain is higher for the far detector responses (because of the higher penetration depth), while further improvement with the WW technique proved more effective for the near detector.

The autonomous unoptimised calculation gives estimates for all possible responses. Applying the WW technique deteriorates the responses other than targeted. It is possible to change the optimisation settings for a set of responses but different physical detectors should be calculated by separate runs. Attempts for optimisation for a set of responses are not surprisingly scarce, as a truly optimal setting can only be achieved for a single response. *Mickael* [36] claims, that his optimisation scheme considers such a combination, although no explanation is given in his publication about how and why this is achieved. For the Midway method, separate adjoint calculations are a must for different physical detectors. For energy/time bins of a single detector the time convolution options or a proper bookkeeping could eliminate the need for completely separate runs for each response as we have discussed earlier. It is worthwhile to note, that very many responses of the same physical detector are

not a particularly easy task for the Midway method because of accuracy problems of the coupling and the increased coupling times.

Testing the obtained efficiency figures against values found in literature cannot be extensively done, mainly because the lack thereof, but also the published few results rarely gave numerical values and mainly used substantially different geometries. *Maučec et al.* [48] used the same geometry and cross-section library, moreover the same operation modes. In their publication, the sigma mode has been simulated by an analogy to energy dependency, mainly to Cl capture photons. The C/O ratio simulations regarded only the truly inelastic gammas, but their conclusion of lower efficiency gain for the near detector and an order of a magnitude for the far detector tallies with the results presented here. *Mickael* [36] and *Maučec et al.*[48] showed efficiency measure but taking the PHD into account, requiring all the energy bins to have lower than 5% variance. Their conclusions of the efficiency gains differ by an order of a magnitude at least, mainly as *Mickael* [36] could not obtain then a statistically well converged reference calculation.

		Near Carbon	Near Oxygen	Far Carbon	Far Oxygen
Forward with WW	S	8,8	7,5	35,3	19,2
	C	1,9	1,7	5,2	10,3
Basic Midway	S	3,8	3,0	29,8	26,9
	C	1,9	1,4	14,4	11,5
Midway-Adjoint WW	S	17,4	15,2	36,0	32,4
	C	11,3	9,3	23,5	20,4
Midway-Forward WW	S	4,5	3,8	42,9	28,6
	C	2,5	1,8	21,9	12,7
Midway- Forward and Adjoint WW	S	21	19,8	49,8	35,2
	C	14,3	12,6	34,3	23,1

Table 4. : Efficiency Gain of different detectors compared to the efficiency of the unoptimised forward calculation, obtained by accounting for the computer time a certain response calculation consumes (S), and for the cumulative computer time all the response calculations together would take (C)

The Midway efficiencies for multiple response estimates, therefore, should be compared using the total running time of all calculations. Table 4. shows the efficiency gains when calculating the FOM using the cumulative computer time of all sub-calculations (except the weight window generation), and the previously shown (Fig.44) efficiency gains. It is apparent, that the Midway method without any further variance reduction technique applied is at least a factor 2 more efficient than a regular calculation, although higher values appear for the far detector. Despite the lower gains for the near detector, the Midway method still stays ahead of applying a simple WW technique in terms of efficiency.



CONCLUSIONS

The integral of the flux and the adjoint function on a surface separating source and detector provides an alternative form for calculating a detector response. The Midway method, the Monte Carlo interpretation of this response form, had been proposed [9] as a general variance reduction tool. This thesis analysed the capabilities of the Midway method extended to handle time dependent responses, for application to borehole logging calculations with the aim of providing an accurate and fast technique for simulations on a detailed parameter space.

Quality analysis of Monte Carlo simulation results is based on measures of the statistical properties of the response estimate. In contrast, deterministic transport calculations may yield inaccurate results to an unknown extent due to insufficient discretisation of the transport equations. The Midway method couples two stochastic processes by a discretised surface integral in five (for time dependent cases: six) dimensions. This conceptual change of the Monte Carlo estimator introduces truncation errors caused by the discretisation of the integral while the statistical behaviour of the response estimate should still be monitored for sufficiently low variance; and further: refinement of the discretisation spoils statistical convergence properties. This obvious drawback of having double error source is balanced by the fact that Midway Monte Carlo calculations have been shown to be more effective than autonomous forward Monte Carlo games; for borehole logging simulations presented in this thesis the efficiency increase has even reached a factor of 50.

In this thesis, the Midway method has been extended to handle time-dependent neutron-photon problems, and has been applied to a borehole logging tool equipped with a time-dependent neutron source and photon detectors. It has been found that the Midway method provides considerable efficiency gain compared to a non-optimised autonomous forward calculation. The responses, however, showed deviations from the reference calculations beyond statistical uncertainty. A further problem arose by the computer resource requirements of the Midway coupling as it reached impractical extents for current computing standards, while the energy discretisation was limited to only a few energy groups.

One reason for the aforementioned discrepancies has been found in the inadequate estimation of the variance. The variance estimation has been corrected and the approximations - necessitated by keeping the calculation flow simple - has been shown and treated. The behaviour of the variance has been found to be insensitive of the discretisation structure until enough adjoint and forward scores hit the Midway surface. If too many phase-space segments are taken in comparison with the number of scores crossing the Midway surface, the variance increases and the estimator becomes unstable. These situations are better indicated by the corrected variance estimator, though they cannot always be clearly identified. The coupling algorithm has been tweaked to decrease the computer memory needs.

For the improvement of the accuracy (for the decrease of the bias in the Midway response estimator) two techniques have proven useful. Firstly, an analytical integration of source and detector functions allows exact coupling in the time variable. Secondly, coupling using expansion on orthogonal function bases (tested here only for the solid angle) resulted in far better accuracy. Both techniques lack, however, a simple way of estimating the variance if the computational effort and/or the memory requirements are to be kept low enough for using the

Midway method as a general variance reduction tool. Biased estimators for the variance can be given for both techniques, and in cases when this suffices both could and should be used.

The scintillation detectors of the simulated tool pose challenge for Monte Carlo simulations. Since their energy dependent response is not describable with quantities of the Boltzmann equation, regular formulation of an adjoint game is not applicable. It has been shown for a non-multiplying system, that adjoint sampling of this detector response is feasible; thus it is possible to set up a Midway calculation.

The efficiency gain of the Midway method has been compared to optimised autonomous forward calculations as well. The commonly used Weight Window (WW) technique has been applied, both in energy and in time dependent forms. For both cases the Midway method provides comparable efficiency gain to the optimised forward calculation. It was observed that for the WW technique the efficiency gain is only present at targeted responses while efficiency of estimators of other responses (other time or energy bin, other detector) decreases. For multiple responses the Midway method on the contrary, provided efficiency increase for every response. Also, the theory for application of the Weight Window technique for the adjoint and forward simulation parts of the Midway method has been presented, and resulted in further efficiency improvement, though the capability of increasing efficiency for multiple responses -with this optimisation step- vanished.

The application of the Midway method for time dependent borehole logging simulations – even for a limited set of energy groups– required optimal usage of computer resources for acceptable accuracy and stability. For realistic simulations, therefore, application of the time dependent Midway method is discouraged. As for different borehole logging scenarios the required number of particles for statistical stability would not change drastically, it is possible to set the number of simulated histories by preliminary numerical tests, and to use the time convolution and function expansion coupling improvements without exact estimate of the variance (or with just observing the statistical fluctuations of some sets of histories) to result in an accurate response without accurate variance estimate.

The energy dependent borehole response without time dependence provided a reasonable challenge for the Midway method, both for accuracy and statistical stability, while the efficiency gain –especially with applying further variance reduction for the adjoint calculation– was appreciably high. For realistic cases, implementation of the adjoint sampling of the detector response is a must, and comparison with measurements should be further investigated.

In general we can say, that the applicability of the Midway method is mainly determined by the complexity of the surface integral, that is the number of dimensions to account for and the ease of describing the flux and the adjoint function by a discretised/expanded function representation. The statistical stability of the Midway response estimate is dependent on the coupling scheme: a trade-off exists between refined function representation and statistical convergence of the estimators. The application of the Midway method is better rewarded, therefore, when the number of dimensions of the surface integral can be decreased or when either the adjoint or the flux is simple to represent.

The mentioned difficulties around the realisation of the Midway method as a general variance reduction technique can all be circumvented on the cost of increased computing effort, user-friendliness, memory requirement or fitting calculation flow to current Monte Carlo codes. By investing considerably more computing effort in processing the adjoint and forward scores (also repeatedly), for example, setting the balance of stability and accuracy can be automated, the variance calculations can be made exact, the truncation errors quantified. The techniques and analysis presented in this thesis can be applied and further developed to forward-adjoint coupling calculations where the efficiency increase is not based mainly on the reduction of the variance, but rather on the decoupling of model domains, like measurement (or tool) design, radiation treatment planning, or sensitivity analysis.

LIST OF SYMBOLS AND ABBREVIATIONS

$+$	adjoint
B	Boltzmann-operator without the streaming term and the time derivative
c	contribution
$C(P \rightarrow P')$	collision kernel
d	function expansion coefficient
D (as a function)	detector function
D^2 (as an operator)	variance
$\delta(x)$	Dirac-delta function
δ_{ij}	Kronecker-delta symbol
E (as a variable)	energy
E (as an operator)	expected value
FOM	Figure Of Merit
ϕ	flux
ϕ^+	adjoint function
h	general scoring function
J	radiation current
K (as a scalar)	number of used coefficients
$K(P \rightarrow P')$	transport kernel
L	linking function
N	number of 'forward' particles started in a simulation
N^+	number of adjoint particles started in a simulation
$\underline{\Omega}$	solid angle
O	order
PHD	Pulse Height Distribution
P (as a variable)	phase-space coordinate
P_l (as a function)	Legendre polynomial
\wp	probability density function
Π	interval selector characteristic function
ψ	collision density
\underline{r} (as a vector)	spatial coordinate
r^2 (as a scalar)	relative variance
R	response
s	score
s^+	adjoint score
S	source function
Σ	cross section
t	time
T (as a scalar)	computer time
$T(P \rightarrow P')$	displacement kernel
w	particle weight (forward simulation)
w^+	particle weight (adjoint simulation)

REFERENCES

1. I.V. Serov, T.M. John and J.E. Hoogenboom, "A Midway Forward-Adjoint Coupling Method for Neutron and Photon Monte Carlo Transport", *Nucl. Sc. Eng.* **133**, pp. 55-72 (1999).
2. Y. Hayashida, Kawai M., Nakai M. "Forward-Adjoint Monte Carlo Coupling Technique Applicability to Neutron Streaming Analysis" *Journal of Nuclear Science and Technology* **27**, 912-921 (1990)
3. Williams M.L., "Generalized Contribution Response Theory" *Nucl. Sc. Eng.* **108**, 335-383 (1991)
4. Cramer S.N. "Forward-Adjoint Monte Carlo Coupling with No Statistical Error Propagation" *Nucl. Sc. Eng.* **124**, 398-416 (1996)
5. Ueki T., J.E. Hoogenboom and J.L. Kloosterman: "Analysis of Correlated Coupling of Monte Carlo Forward and Adjoint Histories" *Nucl. Sc. Eng.* **137**, 117-145 (2001)
6. C.H. Aboughantous "A Contribution Monte Carlo Method" *Nucl. Sc. Eng.* **118**, 160-177 (1994)
7. Densmore D., "Variational Variance Reduction for Monte Carlo Reactor Analysis", *PhD thesis, University of Michigan* (2002)
8. S. Mosher, M. Maućec, A. Badruzzaman, C. Chedester, M. Evans, L. Gadeken, "Expected-Value Techniques for Monte Carlo Modelling of Well Logging Problems", Proceedings of ANS Radiation Protection and Shielding Division, 12th Biennial Topical Meeting, Radiation Serving Society, April 14-18, 2002, Santa Fe (on CD-ROM)
9. I.V. Serov "Estimation of Detector Response by Midway Forward-Adjoint Monte Carlo Coupling in Nuclear Systems" *PhD thesis, Delft University of Technology* (1996)
10. Bell G.I. and S. Gladstone, "Nuclear Reactor Theory", Litton Educationa Publishing, New York (1970)
11. Lewins J., "Importance: The Adjoint Function", Pergamon Press, New York, (1970)
12. J. Briesmeister, "MCNP – a General Monte Carlo N-Particle Transport Code," Los Alamos National Laboratory report LA-13709-M, April 2000
13. Serov I.V., T.M. John and J.E. Hoogenboom, "A New Effective Monte Carlo Midway Coupling Method in MCNP Applied to a Well Logging Problem" *Applied Radiation and Isotopes* **49**, No.12, 1737-1744 (1998)
14. M. Kawai, Hayasida Y., Nishi H. "Application of Forward and Adjoint Monte Carlo coupling Technique to Detector System Designs" *Progress in Nuclear Energy* **24**. (1990)
15. G.A.Simpson, Jacobson L.A.: "A new Small-Diameter, High-Performance Reservoir Monitoring Tool" *SPE53736, Society of Petroleum Engineers Conference, Caracas, Venezuela, 21-23 April 1999*
16. J. E. Hoogenboom, "Methodology of Continuous Energy Adjoint Monte Carlo for Neutron, Photon, and Coupled Neutron-Photon Transport," *Nucl. Sc. Eng.* **143** (2003).
17. B.L. Beers and V. W. Pine, "Functional Expansion Technique for Monte Carlo Electron Transport Calculations" *IEEE Transactions on Nuclear Science*, **23**, No. 6 (1976)
18. A. Noel and H.S. Wio, "A New Series-Expansion Approach in Monte Carlo: Application to Neutron Shielding", *Annals of Nuclear Energy*, Vol. **11**, No.5 pp 225-227 (1984)
19. D.P. Griesheimer, W.R. Martin, "Monte Carlo Based Angular Flux Distribution with Orthogonal Function Expansion" *Trans. Am. Nucl. Soc.*, 89 (2003)

20. D.P Griesheimer, W.R. Martin, "Two Dimensional Functional Expansion Tallies for Monte Carlo Calculations" PHYSOR 2004 – The Physics of Fuel Cycles and Advanced Nuclear Systems: Global Developments, (2004) on CD-ROM
21. I.S. Gradshteyn, I.M. Ryzhik, Alan Jeffrey (editor), "Table of Integrals, Series, and Products", Academic Press, London, (1994)
22. D.P Griesheimer, W.R. Martin, J.P. Holloway, "Convergence Properties of Monte Carlo Functional Expansion Techniques" *submitted to Journal of Computational Physics* (2004)
23. T. E. Booth: A Monte Carlo Variance Reduction Approach for Non-Boltzmann Tallies, *Nucl. Sc. Eng.* **116**, pp.113-124 (1994)
24. E. Shuttleworth: The Pulse Height Distribution Tally in MCBEND, Proceedings of the International Conference on Radiation Shielding ICRS9, Tsukuba, Japan. (October 1999)
25. P. Guo, R.P. Gardner and L. Liu, "Development of Monte Carlo Codes and Gamma-Ray Spectroscopy Unfolding Algorithms for Nuclear Logging Techniques" Technical Progress Report in support of the Advanced Computational Technology Initiative Project, Los Alamos National Laboratory, (1997)
26. G.F. Knoll, "Radiation Detection and Measurement", John Wiley & Sons, New York, (1979)
27. R.P. Gardner and S.H. Lee, "Monte Carlo Simulation of Pulse Pile-up", 46th Annual Denver X-ray Conference Proceedings, Steamboat Springs, Colorado, Aug.4-8, (1997)
28. M.Szieberth and J. L. Kloosterman, "New Methods for the Monte Carlo Simulation of Neutron Noise Experiments", American Nuclear Society Topical Meeting on Mathematics & Computation, Gatlinburg, TN, 2003
29. T.E.Valentine, "MCNP-DSP users manual",ORNL/TM-13334,1997
30. J. E. Hoogenboom, "Methodology of Continuous Energy Adjoint Monte Carlo for Neutron, Photon, and Coupled Neutron-Photon Transport," *Nucl. Sc. Eng.* **143** (2003)
31. J. Spanier and E. M. Gelbard: Monte Carlo principles and neutron transport problems, 1969 Addison-Wesley Publishing
32. I. Lux and L. Koblinger: "Monte Carlo Particle Transport Methods: Neutron and Photon calculations", CRC press, Boca Raton, Florida (1991)
33. J.P. Boyd, "Chebyshev and Fourier Spectral Methods", Dover Publications, New York (2000)
34. M.M.R. Williams, "Random Processes in Nuclear Reactors", Pergamon Press, Oxford (1974)
35. R.A. Forester, R.C. Little, J.F. Briesmeister and J.S.Hendricks: "MCNP Capabilities for Nuclear Well Logging Calculations" IEEE Transactions on Nuclear Science, Vol. **37**, No.3 (1990)
36. M. W. Mickael, "A Fast, Semideterministic Weight Windows Generator for MCNP", *Nucl. Sc. Eng.***119** (1995)
37. T.M. Evans and T.A.Wareing: "The Solution of Well Logging Problems Using Hybrid Transport Methods on Unstructured Meshes" Proceedings of the International Conference on Mathematics & Computation, Reactor Physics, and Environmental Analysis of Nuclear Applications, Madrid, Spain, (1999)
38. A Haghighat, J.C. Wagner "Monte Carlo Variance Reduction with Deterministic Importance Functions", *Progress in Nuclear Energy*, Vol. **42**, No.1 (2003)
39. K. A. van Riper, T.J.Urbatsch, P.D. Soran, D.K. Parsons, J.E. Morel, G.W. McKinney, S.R. Lee, L.A. Crotzer, F.W. Brinkley, J.W. Anderson, and R.E. Alcouffe "AVATAR – Automatic Variance Reduction in Monte Carlo Calculations" *Proc. Joint. Int. Conf.* on

- Mathematical Methods and Supercomputing in Nuclear Applications, 1, Saragota Springs, NY, (1997)
40. J.E. Hoogenboom, "Optimum Biasing of Integral Equations in Monte Carlo Calculations", Technical Note, *Nucl. Sc. Eng.* **70**, (1979)
 41. T.E. Booth, and J.E. Hendricks, "Importance Estimation in Forward Monte Carlo Calculations" *Nucl. Technol./Fusion*, **5** (1984)
 42. T.E. Booth, "The Weight Window Generator Output is All Zeroes", Section working title, *LANL report in progress* (2005)
 43. A. Haghghat, H. Hiruta, B. Petrovic and J.C Wagner, "Performance of the Automated Adjoint Accelerated MCNP (A3MCNP) for Simulation of a BWR Core Shroud Problem", Proceedings of the International Conference on Mathematics & Computation, Reactor Physics, and Environmental Analysis of in Nuclear Applications, Madrid, Spain, (1999)
 44. J.C. Wagner, E.L. Redmond II, S.P. Palmtag, J.S. Hendricks, "MCNP: Multigroup/Adjoint Capabilities", Los Alamos National Laboratory report LA-12704 (1994)
 45. R.C. Little, R.E. Seamon, "New MNDF5 and MENDF5G", Los Alamos National Laboratory internal memorandum, X-6:RCL-87-642 (Dec. 1987)
 46. I. Kodeli, D. Adama, P.F.A. Leege, D. Legrady, J. E. Hoogenboom, and P. Cowan, "Generation and Performance of a Multigroup Coupled Neutron-Gamma Cross-Section Library for Deterministic and Monte Carlo Borehole Logging Analysis" PHYSOR 2004 – The Physics of Fuel Cycles and Advanced Nuclear Systems: Global Developments, (2004) on CD-ROM
 47. W. A. Rhoades, "DOORS 3.2 One-, Two-, Three-Dimensional Discrete Ordinates Neutron/Photon Transport Code System, CCC-650", RSICC, ORNL (1998)
 48. M. Maučec, "Implementation of variance-reduction techniques for Monte Carlo nuclear logging calculations with neutron sources", 10th International Conference on Radiation Shielding (ICRS-10) and 13th Topical Meeting on Radiation Shielding (RPS-2004), Funchal, Madeira Island, Portugal, (2004)

LIST OF PUBLICATIONS

1. J.E. HOOGENBOOM and D.LEGRADY: "A Critical Review of the Weight Window Generator in MCNP", Proceedings of The Monte Carlo Method: Versatility Unbound in a Dynamic Computing World, Chattanooga, Tennessee, April 17-21, ANS, 2005
2. D. LEGRADY and J.E. HOOGENBOOM, "Feasibility of an Adjoint Monte Carlo Pulse Height Spectrum Calculation ", Proceedings of Nuclear Mathematical and Computational Sciences Gatlinburg, Tennessee, 2003, on CD-ROM
3. D. LEGRADY, J.E. HOOGENBOOM and J.L. KLOOSTERMAN, "The Time Dependent Monte Carlo Midway Method for Application to Borehole Logging," Proceedings of Nuclear Mathematical and Computational Sciences Salt Lake City, Utah, 2001, on CD-ROM
4. J.E.HOOGENBOOM and D. LEGRADY: "Coupling of Forward and Adjoint Monte Carlo Calculations by Next Event Estimation", Proceedings of PHYSOR-2002, New Frontiers of Nuclear Technology: Reactor Physics, Safety and High-Performance Computing, October 7-10, 2002, Seoul, Korea
5. D. LEGRADY and J.E.HOOGENBOOM: "Midway Monte Carlo Forward-Adjoint Coupling with Legendre Polynomials", Proceedings of PHYSOR-2004, The Physics of Fuel Cycles and Advanced Nuclear Systems: Global Developments, 2004
6. D. LEGRADY and J.E.HOOGENBOOM: "Visualization of Space-Dependency of Responses of Monte Carlo Calculations Using Legendre Polynomials", Proceedings of PHYSOR-2004, The Physics of Fuel Cycles and Advanced Nuclear Systems: Global Developments, 2004
7. I. KODELI, D. ALDAMA, P. F.A. DE LEEGE, D. LEGRADY and J.E. HOOGENBOM: "Generation and Performance of a Multigroup Coupled Neutron-Gamma Cross-Section Library for Deterministic and Monte Carlo Borehole Logging Analysis", Proceedings of PHYSOR-2004, The Physics of Fuel Cycles and Advanced Nuclear Systems: Global Developments, 2004
8. D. LEGRADY and J.E. HOOGENBOOM: "Report on the Investigation of the Possibilities to Detect the Presence or Absence of a Boron Containing Filter Cake in a Borehole", 2002, Report for Shell International Exploration and Production, Rijswijk, (Contract Number: 4600000859)
9. D. Legrady, J.E. Hoogenboom, "The time dependent Midway method with application to borehole logging" to be submitted to Nucl. Sc. Eng.
10. D. Legrady, J.E. Hoogenboom, "Adjoint sampling of a Pulse Height Tally" to be submitted to Nuclear Instrumentation and Methods

11. S. Nievaart, D. Legrady, T.H.J.J. van der Hagen and H. van Dam, "BNCT adjoint treatment planning using Legendre function expansion" submitted to Journal of Medical Physics

SUMMARY

As detection and monitoring devices, nuclear borehole logging tools are frequently applied in (oil) survey and production wells. For its versatile applicability in identification and quantification of elements in the geological formation, the neutron-gamma tool has been recognised as a primary device especially for already cased wells. The tool operates with a pulsed neutron source and time and energy dependent scintillation photon detectors. Simulations of the tool response for the combinations of possible formation parameters are necessary for the interpretation of the measured data and the required accuracy necessitate Monte Carlo calculations. The calculation time for obtaining statistically converged simulations, however, can be impractically high.

The integral of the flux and the adjoint function on a surface separating source and detector provides an alternative form for calculating a detector response. The Midway method, the Monte Carlo interpretation of this response form, had been proposed [9] as a general variance reduction tool. This thesis analyses the capabilities of the Midway method extended to handle time dependent responses, for application to borehole logging calculations with the aim of providing a fast and accurate technique for simulations on a detailed parameter space.

In this thesis the derivation of the time dependent Midway method is given. Application of the method for borehole logging calculations using the general Monte Carlo code MCNP is shown to provide considerable efficiency gain compared to a non-optimised autonomous forward calculation. The responses, on the other hand, show deviations from the reference calculations beyond statistical uncertainty. A further problem arises by the computer resource requirements of the time dependent Midway coupling as it reaches impractical extents for current computing standards, even with a limited energy discretisation of only a few energy groups.

The Midway method combines two stochastic processes by a five (for time dependent cases: six) dimensional bilinear surface integral for calculating the detector response. The integration is carried out by discretising the phase-space into small phase-space segments and the average adjoint function and the flux is estimated on each facet. This approximation is shown to be the primary source for the observed discrepancies. Two techniques are developed to enhance the accuracy of the method: coupling using orthogonal function expansion and a special utilisation of the invariance of samples for a shift in time. Both methods are shown to decrease the inaccuracies to an acceptable level without utilising a priori information, though the calculation of the statistical variance cannot be incorporated in the regular flow of the coupling.

Another source of the discrepancies beyond the variance is shown to originate from inadequate estimate of the statistical uncertainty. An improved model is described and demonstrated to result in a better estimate. The behaviour of the variance is analysed, and shown to be virtually independent of the discretisation structure until enough adjoint and forward scores hit the Midway surface, and to grow linearly with increasing number of segments for low number of samples. If too many phase-space segments are taken in comparison to the number of scores crossing the Midway surface, the variance increases and the estimator becomes unstable.

The scintillation detectors of the simulated tool pose challenge for Monte Carlo simulations. Since their energy dependent response is not describable with quantities of the Boltzmann equation, regular formulation of an adjoint game is not applicable. The theoretical framework for the adjoint handling of such response is developed and it is shown for a non-multiplying system that adjoint sampling of this detector response is feasible; thus it is possible to set up a Midway calculation.

The efficiency gain provided by the Midway method for borehole logging simulations is also compared to optimised Monte Carlo simulations. The often-used Weight Window (WW) technique that achieves efficiency increase by controlling the particle population migrating in the model geometry is shown not to outperform the Midway method and that they both deliver comparable efficiency gains while the Midway method increases the efficiency also for multiple detector responses. The necessary theoretical derivations are given for applying the WW technique for the adjoint and forward calculation parts of the Midway method. It is demonstrated that further efficiency gain can be achieved by simultaneously using both techniques, even reaching a factor of 50 improvement in efficiency.

SAMENVATTING

De in boorgaten gebruikte nucleaire log-instrumenten worden vaak als detectie- en monitoringmethode toegepast in oliedetectie- en -productiegaten. Het neutron-gamma instrument wordt algemeen erkend als meest geschikte methode voor reeds verkende bronnen vanwege zijn veelzijdige toepassingen bij de identificatie en kwantificering van elementen in de geologische formatie. Het instrument werkt met een pulserende neutronenbron en met tijden energieafhankelijke fotonscintillatiedetectoren. Simulaties van de uitgangssignalen van het instrument voor combinaties van mogelijke formatieparameters zijn nodig voor de interpretatie van de gemeten data. De vereiste nauwkeurigheid maakt het gebruik van montecarloberekeningen noodzakelijk. De tijd die nodig is om statistisch geconvergeerde simulaties te verkrijgen kan echter onpraktisch groot zijn.

De integraal van de flux en van de adjoint functie op een oppervlak dat de bron en de detector scheidt, voorziet in een alternatieve methode om een detectorresponsie te berekenen. De Midway-methode, de montecarlointerpretatie van dit type responsie, is voorgesteld [9] als een algemeen instrument voor variantiereductie. Dit proefschrift analyseert de mogelijkheden van de Midway-methode, uitgebreid voor tijdafhankelijke systemen, toegepast op boorgatberekeningen met het doel om in een snelle en nauwkeurige techniek te voorzien voor simulaties met een gedetailleerde parameterruimte.

In dit proefschrift wordt de afleiding van de tijdsafhankelijke Midway-methode gepresenteerd. Toepassing van de methode voor boorgat-logberekeningen, gebruikmakend van de algemene montecarlocomputerprogramma MCNP, blijkt aanzienlijke efficiëntiewinst op te leveren in vergelijking met een niet geoptimaliseerde autonome berekening. Anderzijds tonen de responsies statistisch relevante afwijkingen van de referentieberekeningen. Een bijkomend probleem treedt op bij de eisen die aan de computer gesteld moeten worden om de tijdafhankelijke Midway-koppeling uit te voeren. Deze eisen overschrijden de huidige computergrenzen zelfs als gebruik gemaakt wordt van een beperkt aantal energiegroepen.

De Midway-methode combineert twee stochastische processen d.m.v. een vijfdimensionale (voor tijdafhankelijke gevallen: zesdimensionale) bilineaire oppervlakte-integraal om de detectorresponsie te berekenen. De berekening van de integraal wordt uitgevoerd door de faseruimte te verdelen in kleine segmenten en de gemiddelde adjoint functie en flux voor elk segment te schatten. Er wordt aangetoond dat deze benadering de primaire oorzaak is voor de gevonden afwijkingen. Er zijn twee technieken ontwikkeld om de nauwkeurigheid van de methode te verbeteren: het koppelen door gebruik te maken van een orthogonale functie-expansie, en een speciaal gebruik van de invariantie van samples voor een verschuiving in de tijd. Het blijkt dat door beide methoden de onnauwkeurigheden tot een acceptabel niveau afnemen zonder dat a-priori informatie wordt gebruikt, hoewel de berekening van de statistische variantie niet kan worden opgenomen in de reguliere wijze van de koppeling.

Een andere bron van de discrepantie is afkomstig van een inadequate schatting van de statistische onzekerheid. Een verbeterd model wordt beschreven waarbij wordt aangetoond

dat dit model resulteert in een betere schatting. Het gedrag van de variantie is geanalyseerd en blijkt onafhankelijk te zijn van de discretisatiestructuur totdat het Midway-oppervlak geraakt wordt door voldoende adjoint en voorwaartse scores. Als teveel faseruimte segmenten worden gebruikt in vergelijking met het aantal scores op het Midway-oppervlak neemt de variantie toe en wordt de schatter onstabiel.

De scintillatiedetectoren van het gesimuleerde instrument vormen een uitdaging voor de monte-carlosimulaties. Omdat hun energieafhankelijke responsie niet valt te beschrijven met grootheden van de Boltzmann-vergelijking, is de reguliere formulering van een adjoint simulatie niet toepasbaar. Het theoretische raamwerk voor het adjoint behandelen van zo'n responsie is in dit proefschrift ontwikkeld waarbij wordt aangetoond dat het adjoint bemonsteren van deze detectorresponsie mogelijk is voor een zich niet vermenigvuldigend systeem. Hieruit blijkt dat het mogelijk is om een Midway-berekening uit te voeren.

De efficiëntiewinst die bereikt wordt door de Midway-methode toe te passen voor het simuleren van boorgat-logs wordt ook vergeleken met geoptimaliseerde monte-carlosimulaties. Het blijkt dat de vaak gebruikte Weight Window (WW) -techniek, waarbij een efficiëntietoename wordt bereikt door de deeltjespopulatie te beheersen die in de modelgeometrie migreert, niet beter presteert dan de Midway-methode en dat ze vergelijkbare efficiëntiewinsten bereiken terwijl de Midway-methode ook de efficiëntie vergroot in geval van een meervoudige detectorresponsie. In dit proefschrift worden de noodzakelijke theoretische afleidingen gegeven om de WW-techniek toe te passen op de adjoint en de voorwaartse berekeningen van de Midway-methode. Er wordt aangetoond dat nog meer efficiëntiewinst kan worden bereikt door beide technieken tegelijkertijd te gebruiken: hierdoor kan zelfs een 50-voudige efficiëntieverbetering worden bereikt.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Eduard Hoogenboom first, for his dedicated effort of guiding me through my research, for the rigorous checks of my mathematical derivations, for his openness for new ideas and for his encouragement throughout these years. Also for his endless patience when helping me fix bugged programs or when once my melting ice-cream nearly destroyed his suit.

I am very grateful to my promoters Tim van der Hagen and Hugo van Dam for their tutoring, for their help in learning how to follow right research paths and how to write scientific texts, for their very valuable comments on both science and the way of its presentation, and for their creative attention when decoding my never near-perfect Dutch.

I might have not got engaged with this PhD research had that not been for Sándor Fehér of the TU Budapest, who encouraged me to apply for this research position. He had followed my path ever since and I deeply grateful to him for his continuous fatherly help. It is also due to him that I accepted the research grant of the System International Foundation in Budapest. For this grant I would also like to acknowledge Gyula Csom.

The staff members of the PNR group have always given me loads of support, and for this, for all of you, many thanks. I would like especially to thank the department-mother and secretary Riny Purmer for her attention and caring when I arrived at the group; Dick de Haas and Camiel Kaaijk for always fixing instantly my computer when it went haywire and this happened frequently enough; Piet de Leege for providing always cross section files in any format at any time and for teaching me what the AMPX-MASTER is; Jan Leen Kloosterman for following my research and giving many useful comments. Danny Lathouwers was always at hand for scientific discussions when I got stuck, I want to thank him for the many hours we spent together in front of the blackboard explaining ideas and facts.

I want to thank Katalin Pjeczka, Annalisa Manera, Sander Nievaart, Ruben Abellón, Jessica Kroeze, Candice Johns, Alexander Schmets, Niels Boerema, Mark Frahn, Devika Raktoc and Lucy Mathers, who in one way or another made me really feel at home in the Netherlands. Thanks all of you!

I would like to especially thank, the support that my mother, my late father, sisters and brothers provided.

CURRICULUM VITAE

The author of this thesis was born on the 20th of March, 1977 in Budapest, Hungary. He concluded his studies at the Eötvös Loránd University of Sciences, Budapest by finishing his master thesis in the topic of “Monte Carlo Simulation of Natural Gamma Radiation Spectral Measurements in Boreholes” in 2000. He started his PhD studies at the Reactor Physics Department of the Interfaculty Reactor Institute (now known as the Department of Radiation, Radionuclides and Reactors of the Faculty of Applied Sciences of the Delft University of Technology) in 2000 under the supervision of dr. ir. J. E. Hoogenboom and of his promoters prof. dr. ir. H. van Dam and prof. dr. ir. T.H.J.J. van der Hagen. He interrupted his PhD studies for half a year in 2004 during which period he conducted research funded by the System International Foundation, Hungary. After returning to the Netherlands in fall 2004, he finished this thesis on the Time Dependent Midway Monte Carlo method with application to borehole logging.



