Error Analysis of Sequential Monte Carlo Methods for Transport Problems

Rong Kong & Jerome Spanier

Department of Mathematics
Claremont Graduate University
925 N. Dartmouth Ave.
Claremont, CA 91711, USA
Email: rong.kong@cgu.edu & jerome.spanier@cgu.edu

Abstract. In 1962, Halton introduced a sequential correlated sampling algorithm for the efficient solution of certain matrix problems. We have extended Halton's method to the solution of certain simple transport problems and the resulting algorithm is capable of producing geometric convergence for these problems. In our algorithm, random walks are processed in groups, called stages, and the result of each stage is a correction that is added to the solution at the previous stage. It is then of interest to determine conditions that guarantee strict error reduction at each stage for various transport problems. Specifically, if $\Phi(x)$ is the true transport solution and $\tilde{\Phi}^{n-1}(x)$ and $\tilde{\Phi}^n(x)$ are the estimated solutions from the $(n-1)^{th}$ and $n^{th}$ stages, respectively, we demonstrate the existence of a number $\lambda$, $0 < \lambda < 1$, which is independent of the stage number $n$, such that

$$\|\Phi^n(x) - \Phi(x)\| \leq \lambda \|\tilde{\Phi}^{n-1}(x) - \Phi(x)\| + \varepsilon$$

in a certain probabilistic sense, where $\varepsilon$ is an error term that tends to zero as both the number of terms representing the global solution and the number of random walks per stage tend to infinity. We will indicate how to find such a $\lambda$, which is defined in terms of the number of random walks per stage and the coefficients of the transport problem in a rather natural way.

1 Introduction

In this paper, we will consider a family of monodirectional model transport problems:

$$\begin{cases}
\frac{d\Phi}{dx} + \frac{\Sigma_t}{\cos \theta} \Phi(x) = \frac{\Sigma_s}{\cos \theta} \Phi(x), \\
\Phi(0) = Q_0,
\end{cases} \quad 0 < x \leq T,$$

(1)

where

$\Phi(x)$ is the expected number of the particles per second passing the point $x$;

$\Sigma_t$ is the total macroscopic cross section;

$\Sigma_s$ is the total scattering cross section;

$\Sigma_a = \Sigma_t - \Sigma_s$ is the total absorption cross section;
\( \theta \) is the fixed angle made by the direction of motion of partic positive \( x \)-axis;

\( Q_0 \) is the particle source at the left boundary;

\( T \) is the slab thickness.

Our objective here is to demonstrate that the geometric com have observed in [1] when solving this problem with sequential sampling methods can be justified theoretically. We are also inter whether estimates of the error can be obtained from our analysis we only discuss monodirectional transport problems in this paper that the analysis can also be applied to more complicated transp

First, we introduce the function space in which most of our w done. Specifically, the Hilbert space

\[
L^2[0,T] = \left\{ \varphi(x) \mid \| \varphi(x) \|_{L^2} = \left[ \int_0^T (\varphi(x))^2 \, dx \right]^{1/2} < \infty \right\}
\]

consisting of square-integrable functions on \( 0 \leq x \leq T \) is appro needs. We make the basic assumption that there exists a solution theoretically established or physically implied, and also that

\[
\Phi(x), \quad \frac{d\Phi}{dx} \in L^2[0,T].
\]

We have the following basic result concerning a rather natur the space \( L^2[0,T] \):

**Lemma 1** The set of all Legendre polynomials \( \{ P_i\left(\frac{2x}{T} - 1\right)\} \) complete orthogonal basis of the space and

\[
\int_0^T P_i\left(\frac{2x}{T} - 1\right) P_j\left(\frac{2x}{T} - 1\right) \, dx = \begin{cases} \frac{T}{2i+1}, & \text{if } j = i, \\ 0, & \text{otherwise.} \end{cases}
\]

The proof can be found in [2]. Thus, if \( \Phi(x) \) is a solution of \( L^2[0,T] \) and we have the following representation\(^1\)

\[
\Phi(x) = \sum_{i=0}^{\infty} a_i P_i\left(\frac{2x}{T} - 1\right),
\]

where

\[
a_i = \frac{2i + 1}{T} \int_0^T \Phi(x) P_i\left(\frac{2x}{T} - 1\right) \, dx.
\]

\(^1\) Convergence in the space \( L^2 \) allows the possibility that the equality hold at a set of points \( x \) of measure zero in the interval \([0,T]\) (for the endpoints \(0, T\). However, we believe that the arguments we pr paper do not require strengthened assumptions about the represent
In particular, for any \( \varepsilon > 0 \), there exists an \( N > 0 \), such that

\[
\left\| \sum_{i=N+1}^{\infty} a_i P_i \left( \frac{2x}{T} - 1 \right) \right\|_{L^2} < \varepsilon. \tag{7}
\]

Let

\[
\Phi_N(x) = \sum_{i=0}^{N} a_i P_i \left( \frac{2x}{T} - 1 \right) \tag{8}
\]

be the \((N+1)^{st}\) order term truncation of the solution and let

\[
r_N(x) = \Phi(x) - \Phi_N(x) \tag{9}
\]

denote the error.

In this paper, we will not discuss how to estimate \( r_N(x) \); that is thoroughly covered in books that treat orthogonal expansion of functions in a Hilbert space - for example, [3]. Here our principal concern is with the error that occurs when we estimate the \((N+1)^{st}\) order truncation of the solution, \( \Phi_N(x) \), using sequential Monte Carlo methods.

To make our perspective clearer, we review in outline form the sequential correlated sampling strategy that we have used for solving (1) which was first introduced by J. Halton [4] for matrix problems.

Choose an initial approximation \( \Psi_0(x) \) (we usually take \( \Psi_0(x) \equiv 0 \)) of (1) and let \( \Phi(x) = \Psi_0(x) + \varphi^0(x) \), where \( \varphi^0(x) \) is the solution of the \( 0^{th} \) stage problem as follows

\[
\begin{cases}
\frac{d \varphi^0}{dx} + \frac{\Sigma_1}{\cos \theta} \varphi^0(x) = \frac{\Sigma_2}{\cos \theta} \varphi^0(x) + S_0(x), & 0 < x \leq T, \\
\varphi^0(0) = Q_0 - \Psi_0(0) = Q_0^0.
\end{cases} \tag{10}
\]

where

\[
S_0(x) = -\left( \frac{d \Psi_0}{dx} + \frac{\Sigma_a}{\cos \theta} \Psi_0(x) \right). \tag{11}
\]

Assuming

\[
\varphi^0(x) = \sum_{i=0}^{\infty} a_i^0 P_i \left( \frac{2x}{T} - 1 \right) \tag{12}
\]

then it follows that

\[
a_i^0 = \frac{2i + 1}{T} \int_0^T \varphi^0(x) P_i \left( \frac{2x}{T} - 1 \right) dx. \tag{13}
\]

Estimating (13) by Monte Carlo methods, we obtain an approximation

\[
\bar{\varphi}_N(x) = \sum_{i=0}^{N} a_i^0 P_i \left( \frac{2x}{T} - 1 \right). \tag{14}
\]
This completes the 0th stage.

Assume we have completed the first \((n - 1)\) stages and let \(\bar{\varphi}_N^0, \ldots, \bar{\varphi}_N^{N-1}(x)\) be the solutions for each stage. For an unknown \(\varphi^n\)

\[
\Phi(x) = \Psi^0(x) + \bar{\varphi}_N^0(x) + \bar{\varphi}_N^1(x) + \cdots + \bar{\varphi}_N^{N-1}(x) + \varphi^n(x)
\]

so that each \(\varphi^n(x)\) is just a kind of correction of the approxim: obtained from previous stages. Substituting (15) into (1), we obt
stage problem to determine \(\varphi^n(x)\):

\[
\begin{cases}
    \frac{d\varphi^n}{dx} + \frac{\bar{\sigma}_1}{\cos \theta} \varphi^n(x) = \frac{\bar{\sigma}_d}{\cos \theta} \varphi^n(x) + S^n(x), & 0 < x \leq T \\
    \varphi^n(0) = Q^0_0 - \Psi^0(0) - \bar{\varphi}_N^0(0) - \cdots - \bar{\varphi}_N^{N-1}(0) \equiv Q^n_0,
\end{cases}
\]

where, inductively,

\[
S^n(x) = - \left( \frac{d\bar{\varphi}_N^{N-1}}{dx} + \frac{\bar{\sigma}_d}{\cos \theta} \bar{\varphi}_N^{N-1}(x) \right) + S^{N-1}(x).
\]

Assuming

\[
\varphi^n(x) = \sum_{i=0}^{\infty} a_i^n P_i(2x / T - 1)
\]

then

\[
a_i^n = \frac{2i + 1}{T} \int_0^T \varphi^n(x) P_i(2x / T - 1) dx.
\]

Estimating (19) by Monte Carlo methods (see [1] for details), we approximation

\[
\bar{\varphi}_N^n(x) = \sum_{i=0}^{N} \bar{a}_i^n P_i(2x / T - 1).
\]

For this algorithm, we want to know how close the approxim after stage \(n\)

\[
\bar{\Phi}_N^n(x) \equiv \Psi^0(x) + \bar{\varphi}_N^0(x) + \bar{\varphi}_N^1(x) + \cdots + \bar{\varphi}_N^{N-1}(x) + \bar{\varphi}_N^n(x)
\]

is to the true solution \(\Phi(x)\), or equivalently, how the differences

\[
\left\| \Phi(x) - \bar{\Phi}_N^n(x) \right\|_{L^2}
\]

vary according to \(n\).

2 Construction of the Random Variables

To begin our analysis, without loss of generality, we let \(\Psi^0(x) \equiv \Phi^0(x)\) duce two other groups of functions, defined as follows: if \(\varphi^n(x)\)
solution of (16), the \( n \)th stage problem, \( \varphi_n^N(x) \) is defined by truncating \( \varphi^n(x) \) at its first \( (N + 1) \) terms. Namely, if
\[
\varphi^n(x) = \sum_{i=0}^{\infty} a^n_i P_i \left( \frac{2x}{T} - 1 \right)
\]
then
\[
\varphi_n^N(x) = \sum_{i=0}^{N} a^n_i P_i \left( \frac{2x}{T} - 1 \right).
\] (24)

The following lemma shows the relation between \( \Phi_N(x) \) and \( \Phi_N(x) \) (defined by (8).)

**Lemma 2** Assuming \( \Phi_N(x) \), defined by (8), is the truncated \((N + 1)\) term expansion of the true solution \( \Phi(x) \) of (1) and \( \Phi_N(x) \), defined by (24), is the truncated \((N + 1)\) term expansion of the true solution \( \varphi^n(x) \) of (16), then
\[
\Phi_N(x) = \varphi_n^0(x) + \varphi_n^1(x) + \cdots + \varphi_n^{N-1}(x) + \varphi_n^N(x)
\] (25)
where \( \varphi_n^N(x) \) for \( n = 0, 1, \cdots \) are defined by (20).

Proof. Equating the projections of the two sides of (15) on the finite space spanned by
\[
\left\{ P_i \left( \frac{2x}{T} - 1 \right) \right\}_{i=0}^{N}
\] (26)
leads to (25). \( \Box \)

The following lemma expresses a basic technique important in our later discussion.

**Lemma 3** Using the notations of Lemma 2, we have
\[
\| \Phi(x) - \Phi_N(x) \|_{L^2} = \| r_N(x) \|_{L^2} + \| \varphi_n^N(x) - \varphi_n^N(x) \|_{L^2}.
\] (27)

Proof. We have
\[
\Phi(x) - \Phi_N(x) = \Phi(x) - \Phi(x) + \Phi(x) - \Phi_N(x)
= r_N(x) + \varphi_n^N(x) - \varphi_n^N(x).
\] (28)

Since \( r_N(x) \) is orthogonal to \( \varphi_n^N(x) - \varphi_n^N(x) \), we obtain (27). \( \Box \)

Lemma 3 says that in order to estimate \( \| \Phi(x) - \Phi_N(x) \|_{L^2} \), we only need to estimate the error \( \| \varphi_n^N(x) - \varphi_n^N(x) \|_{L^2} \) of the "corrections" at each sequential stage.

Let us first estimate \( \| \varphi_0^N(x) - \varphi_0^N(x) \|_{L^2} \), which is the error of the solution of the 0th stage. The key to obtain this estimate is the construction of an appropriate random variable to estimate each \( a_i \).

Consider the following problem for \( k(x,y) \), with a parameter \( y \in [0,T] \):
\[
\begin{cases}
\frac{dk}{dz} + \frac{\theta}{\cos \theta} k(x,y) = \frac{x}{\cos \theta} k(x,y) + \delta(x-y), & 0 \leq x \leq T, \\
k(0,y) = 0
\end{cases}
\] (29)
and the problem for \( k(x,0) \):

\[
\begin{align*}
\frac{dk}{dz} + \frac{\Sigma_k}{\cos \theta} k(x,0) &= \frac{\Sigma_k}{\cos \theta} k(x,0), & 0 \leq z \leq T, \\
k(0,0) &= 1.
\end{align*}
\]

The reason that we use the notation \( k(x,0) \) is that (30) can be as a special case of (29) for \( y = 0 \). We write down (30) here for Now, let us construct a random variable used to estimate integra \( k(x,y) \).

Assume \( \xi \) and \( \eta \) are two random variables that have exponential distributions respectively

\[\xi \sim \frac{\Sigma_k}{\cos \theta} e^{-\frac{\Sigma_k}{\cos \theta} z}, \quad z > 0; \]
\[\eta \sim UNIF[0,1].\]

Then we have

**Theorem 1** Suppose that \( \xi_1, \xi_2, \xi_3, \ldots \) and \( \eta_1, \eta_2, \eta_3, \ldots \) are independent, identically distributed random variables of exponential type, respectively, and \( f(x) \in L^2[0,T] \). Then the following variable for \( y \in [0,T] \):

\[
\beta(y) = \int_0^y \left( \sum_{k=1}^{\infty} \left[ \prod_{m=1}^k X_{[0,\xi_m]}(\eta_m) \right] \frac{T - \sum_{m=1}^k \xi_m - y}{\sum_{m=1}^k \xi_m + y} f(x) \right) dx
\]

has the expectation

\[
E_\beta[\beta(y)] = \int_0^T f(x) k(x,y) dx
\]

and the variance

\[
V_\beta[\beta(y)] = \frac{\Sigma_k}{\cos \theta} e^{\frac{\Sigma_k}{\cos \theta} y} \int_0^T e^{\frac{\Sigma_k}{\cos \theta} z} \left( \int_0^T f(x) e^{-\frac{\Sigma_k}{\cos \theta} x} dx \right)^2 dz
\]

where \( X_{[0,c]}(x) \) is the characteristic function of the set \([0,c]\) and \( I \). Heaviside function defined by

\[
H(x) = \begin{cases} 
1, & x \geq 0, \\
0, & \text{otherwise}.
\end{cases}
\]

The proof of this theorem will be given in Appendix A.

**Remarks**

1. \( \beta(y) \) defined in (32) is the generalized track length estimator monodirectional model transport problem. A discussion about the estimator can be found in [5].
(2) \( V_{\beta} [\beta(y)] \) does not depend on \( \Sigma_s \) and \( \Sigma_t \).

From this theorem, we can draw some conclusions about the initial error (i.e., the error resulting from the 0th stage) of our sequential correlated sampling method:

**Theorem 2** Assume \( f(x) = \frac{2i + 1}{T} P_i \left( \frac{2x}{T} - 1 \right) \) (in this case we write \( \beta_i = \beta(x) \) in Theorem 1) and let

\[
\varphi^0_i = Q_{0}^0 \Delta(0).
\]  

Then for any \( \varepsilon > 0 \) and \( \delta > 0 \), there exists a \( W > 0 \), such that

\[
P \left\{ \left\| \frac{1}{W} \sum_{w=1}^{W} \varphi_i^{0} - \varphi_i^{0} \right\| \geq \varepsilon \right\} < \delta,
\]  

where \( \varphi_i^{0} \) is defined by (13) and \( (\varphi_{i1}^{0}, \varphi_{i2}^{0}, \cdots, \varphi_{iW}^{0}) \) are \( W \) samples of the estimator \( \varphi_i^{0} \). Namely, if we use the Monte Carlo method to estimate each coefficient \( \varphi_i^{0} \), the probability that the inequality \( \left\| \varphi_i^{0} - \varphi_i^{0} \right\| \geq \varepsilon \) is less than \( \delta \).

Proof. First, according to Theorem 1, \( \varphi_i^{0} \) is an unbiased estimator of \( \varphi_i^{0} \), i.e.

\[
E[\varphi_i^{0}] = \varphi_i^{0}.
\]  

Therefore, according to Chebyshev's inequality (please refer to [6]), we have

\[
P \left\{ \left\| \frac{1}{W} \sum_{w=1}^{W} \varphi_i^{0} - E[\varphi_i^{0}] \right\| \geq \varepsilon \right\} < \frac{\text{Var}[\varphi_i^{0}]}{\varepsilon^2 W}.
\]  

The proof of (37) is completed if we choose \( W > \frac{\text{Var}[\varphi_i^{0}]}{\varepsilon^2 \delta} \), where, according to Theorem 1,

\[
V[\varphi_i^{0}] = (Q_{0}^0)^2 \cos \theta \int_0^T e^{\frac{\Delta(t)}{4}} \left( \frac{2i + 1}{T} \int_t^T P_i \left( \frac{2x}{T} - 1 \right) e^{-\frac{x^2}{2}} dx \right)^2 dt
\]  

Such a \( W \) always exists since \( V[\varphi_i^{0}] < +\infty \).

The following is a probabilistic estimate of \( \| \Phi(x) - \tilde{\Phi}_N(x) \|_{L^2} \).

**Theorem 3** For any \( \varepsilon > 0 \) and \( \delta > 0 \), there exists a \( W > 0 \) such that

\[
P \left( \left\| \Phi(x) - \tilde{\Phi}_N(x) \right\|_{L^2} \geq \varepsilon \right) \leq \delta,
\]  

where \( W \) is the number of samples of random variables \( \varphi_i^{0} \) (i = 1, 2, \cdots) used for estimating the coefficients of \( \Phi_N(x) \), which is related to \( \tilde{\Phi}_N(x) \) by (28).

Proof. According to Lemma 3, we only need to estimate \( \| \varphi_N^0(x) - \tilde{\varphi}_N(x) \|_{L^2} \).

Using the notations of Theorem 2 and the orthogonality of the Legendre polynomials, we obtain

\[
\| \varphi_N^0(x) - \tilde{\varphi}_N(x) \|_{L^2} = \left\{ \sum_{i=0}^{N} \frac{T}{2i + 1} \left\| \frac{1}{W} \sum_{w=1}^{W} \varphi_i^{0} - a_i^{0} \right\|^2 \right\}^{1/2}.
\]  

(42)
Thus, for any $\epsilon > 0$ and $\delta > 0$, we first can find an $N > 0$ such that

$$\|r_N(x)\|_{L^2} < \frac{\epsilon}{2}.$$ 

For such an $N$, according to Theorem 2 and (42), we can find a $W$ such that

$$P\left(\|\varphi_N^0(x) - \varphi_N^1(x)\|_{L^2} < \frac{\epsilon}{2}\right) > 1 - \delta.$$ 

Thus for such a $W$, we have

$$P\left(\left\|\varphi(x) - \varphi_N^0(x)\right\|_{L^2} < \epsilon\right)$$

$$\geq P\left(\left(\|r_N(x)\|_{L^2} < \frac{\epsilon}{2}\right) \cap \left(\|\varphi_N^0(x) - \varphi_N^1(x)\|_{L^2} < \frac{\epsilon}{2}\right)\right)$$

$$= P\left(\|\varphi_N^0(x) - \varphi_N^1(x)\|_{L^2} < \frac{\epsilon}{2}\right) > 1 - \delta.$$ 

This completes the proof. \(\square\)

**Remark** Theorem 3 actually provides error estimates for c (i.e., non-sequential) Monte Carlo methods used in conjunction with the expansion technique, as our method for expanding the solution linear combination of basis functions is sometimes called. On the other hand, it also serves to describe the initial error when the sequential N method is applied.

Now we turn to the $n^{th}$ stage assuming that we have obtained $\|\varphi_{n-1}^n(x) - \varphi_{n-1}^n(x)\|_{L^2}$. The equation for the $n^{th}$ stage is (1). We have

**Theorem 4** Assume that $\beta_1(y) = \beta(y)$ where $f(x) = \frac{2n+1}{T} F$ defined by (32). Then the following estimator $\zeta^n_1$

$$\zeta^n_1 = Q_0^n \beta_1(0) + \int_0^T S^n(y) \beta_1(y) dy$$

is an unbiased estimator of $a^n_1$, i.e.,

$$E[\zeta^n_1] = a^n_1,$$

and its variance satisfies

$$V[\zeta^n_1] \leq (Q_0^n)^2 V[\beta_1(0)] + \int_0^T (S^n(y))^2 dy \cdot \int_0^T V[\beta_1(y)]$$

where $a^n_1$ is defined by (19).

**Remark** Since we use $\zeta^n_1$ to estimate $a^n_1$, (48) tells us roughly how this estimation process is. Thus (48) can be useful in providing error estimates in transport problems.

**Proof.** It is an algebraic calculation to prove that

$$\varphi^n(x) = Q_0^n k(x, 0) + \int_0^T k(x, y) S^n(y) dy$$
satisfies (16), where, \( k(x, y) \) and \( k(x, 0) \) are solutions of (29) and (30), respectively, and therefore the coefficients \( a^*_n \) can be expressed by

\[
a^*_n = Q^n_0 2^{n+1} \int_0^T k(x, 0) P_i \left( \frac{2^x}{n} - 1 \right) dx + 2^{n+1} \int_0^T S^n(y) dy \int_0^T k(x, y) P_i \left( \frac{2^y}{n} - 1 \right) dx.
\]

(50)

Thus, by taking the expectation of (46) and using Theorem 1 we obtain

\[
E[\zeta^*_n] \equiv E \left[ Q^n_0 \beta_i(0) + \int_0^T S^n(y) \beta_i(y) dy \right]
= Q^n_0 E[\beta_i(0)] + \int_0^T S^n(y) E[\beta_i(y)] dy
= Q^n_0 2^{n+1} \int_0^T k(x, 0) P_i \left( \frac{2^x}{n} - 1 \right) dx + 2^{n+1} \int_0^T S^n(y) dy \int_0^T k(x, y) P_i \left( \frac{2^y}{n} - 1 \right) dx.
\]

(51)

Equation (48) can then be obtained by direct estimation (note that \( \beta_i(0) \) and \( \beta_i(y) \) are independent) using the Cauchy-Schwarz inequality

\[
V[\zeta^*_n] \equiv V \left[ Q^n_0 \beta_i(0) + \int_0^T S^n(y) \beta_i(y) dy \right]
= V \left[ Q^n_0 \beta_i(0) \right] + V \left[ \int_0^T S^n(y) \beta_i(y) dy \right]
\leq (Q^n_0)^2 V[\beta_i(0)] + E \left( \int_0^T S^n(y) \beta_i(y) \right)^2 dy
\leq (Q^n_0)^2 V[\beta_i(0)] + \left( \int_0^T (S^n(y))^2 dy \right) \left( \int_0^T (\beta_i(y) - E[\beta_i(y)])^2 dy \right)
= (Q^n_0)^2 V[\beta_i(0)] + \left( \int_0^T (S^n(y))^2 dy \right) \int_0^T V[\beta_i(y)] dy.
\]

(52)

\[\Box\]

3 Geometric Convergence

In the previous section, we constructed the basic random variables used in our algorithm. In this section, we will use these random variables to construct estimates of the error of the form

\[
\|\Phi(x) - \tilde{\Phi}_N(x)\|_{L^2} \leq \lambda(W, N) \|\Phi(x) - \tilde{\Phi}_N^{-1}(x)\|_{L^2} + \rho(W, N),
\]

(53)

for the solution of problem (1), where \( \rho(W, N) \) and \( \lambda(W, N) \) only depend on the data of the transport problem, the integer \( N \) is the number of basis functions used to represent the solution, and the integer \( W \) is the number of random walks generated in each stage. The term \( \rho(W, N) \) can be made arbitrarily small, and the error reduction factor \( \lambda \) can be made to satisfy \( 0 < \lambda(W, N) < 1 \). We refer to inequalities such as (53) as exhibiting modified geometric convergence because of the appearance of the small additional term \( \rho(W, N) \).
As preparation for the demonstration of (53), we define some functions which will be used in the following theorems:

$$c_{ij} = \begin{cases} 2j + 1, & \text{if } j = i - 1, i - 3, i - 5, \ldots, \\ 0, & \text{otherwise}, \end{cases}$$

$$b_0 = \frac{2\pi}{|\cos \theta|},$$

$$b_i = (2i + 1)^\frac{1}{2} \sum_{j=0}^{i-1} \left( \left( \frac{2j+1}{2i+1} \right)^\frac{1}{2} \frac{2j}{2i+1} c_{ij} \right) + \frac{\sqrt{2\pi}}{|\cos \theta|}, \quad i = 1, 2, \ldots, N.$$  

$$R_W(y) = \sum_{i=0}^{N} \left( \frac{T}{2i+1} \right)^\frac{1}{2} \int_{0}^{T} \frac{\partial}{\partial y} \sum_{w=1}^{W} \beta_{tw}(y) - \frac{2i+1}{T} k(x, y) P_{i}(\frac{2\pi}{T})$$

$$D_{RN}(x) = \frac{\partial r_{N}(x)}{\partial x} + \frac{\alpha}{\cos \theta} f_{RN}(x).$$

where $\beta_{tw}(y)$ denotes the $w^{th}$ sample of $\beta_i(y)$.

**Theorem 5** Suppose that $\tilde{\Phi}_N(x)$ is the approximate solution the $n^{th}$ stage obtained by using the sequential Monte Carlo method (53) holds with two constants $\lambda(W, N)$ and $\rho(W, N)$ defined by

$$\lambda(W, N) = \sqrt{\sum_{i=0}^{N} \frac{2i+1}{2i+1} R_W(0) + \left( \sum_{i=0}^{N} u_i^2 \right)^\frac{1}{2}} \sup_{0 < y < T} R_W(y)$$

$$\rho(W, N) = \left[ r_N(0) R_W(0) + \sqrt{2} \| D_{RN}(x) \|_{L^2} \right] \sup_{0 < y < T} R_W(y)$$

$$+ (1 - \lambda(W, N)) \| r_N(x) \|_{L^2}.$$  

\( \square \)

The proof will be given in Appendix B.

The following theorem establishes the modified geometric convergence making use of Theorem 5.

**Theorem 6** Using the notations of Theorem 5, for any $\epsilon > 0$, $\delta > 0$, there exists an $N > 0$ and a $W > 0$ such that

$$P \left\{ \{ \lambda(W, N) < h \} \cap \{ \rho(W, N) \leq \epsilon \} \right\} > 1 - \delta$$

and

$$P \left\{ \left\| \Phi(x) - \tilde{\Phi}_N(x) \right\|_{L^2} \leq h \left\| \Phi(x) - \tilde{\Phi}_N^{-1}(x) \right\|_{L^2} + \epsilon \right\} > 1 - \delta$$

where $W$ is the total number of samples of the random variable for each coefficient $a_n$.

**Proof.** According to Theorem 5, we only need to prove (56).

Choose a small number $\epsilon'$ that will be determined later such that $N' > 0$, when $N > N'$,

$$\| r_N(x) \|_{L^2} < \epsilon'.$$

According to (3), there must be an $M > 0$, such that

$$\| D_{RN}(x) \|_{L^2} \leq M$$

for all $N$.  

First, let us make $\lambda(W, N) < 1$. For any $\varepsilon_1 > 0$, applying Chebyshev's inequality to $\beta_i(y)$, we obtain

$$
P \left\{ \left| \frac{1}{W} \sum_{w=1}^{W} \beta_{iw}(y) - \frac{2i+1}{T} \int_0^T k(x, y) P_i \left( \frac{2x}{T} - 1 \right) dx \right| < \varepsilon_1 \right\} > 1 - \frac{\text{Var}[\beta_i(y)]}{\varepsilon_1^2 W}.
$$

(60)

Thus, according to (54),

$$
P \left\{ R_W(y) < \varepsilon_1 \sum_{i=0}^{N} \left( \frac{T}{2i+1} \right)^{\frac{1}{2}} \right\} 
\geq P \left\{ \prod_{i=0}^{N} \left\{ \left| \frac{1}{W} \sum_{w=1}^{W} \beta_{iw}(y) - \frac{2i+1}{T} \int_0^T k(x, y) P_i \left( \frac{2x}{T} - 1 \right) dx \right| < \varepsilon_1 \right\} \right\} 
\geq \prod_{i=0}^{N} \left( 1 - \frac{\text{Var}[\beta_i(y)]}{\varepsilon_1^2 W} \right).
$$

(61)

Therefore, according to (55), for any $h < 1$ we have

$$
P \{ \lambda(W, N) \leq h \}
= P \left\{ \sqrt{\sum_{i=0}^{N} \frac{2i+1}{T} R_W(0) + \left( \sum_{i=0}^{N} b_i^2 \right)^{\frac{1}{2}} \sup_{0 < \varphi \leq T} R_W(\varphi) \leq h \} \right\}
\geq P \left\{ \left\{ \sqrt{\sum_{i=0}^{N} \frac{2i+1}{T} R_W(0) \leq \frac{h}{2} \right\} \cap \left\{ \left( \sum_{i=0}^{N} b_i^2 \right)^{\frac{1}{2}} \sup_{0 < \varphi \leq T} R_W(\varphi) \leq \frac{h}{2} \} \right\} \right\}
= P \left\{ R_W(0) \leq \frac{h}{2} \left( \sum_{i=0}^{N} b_i^2 \right)^{-\frac{1}{2}} \right\} \cdot P \left\{ \sup_{0 < \varphi \leq T} R_W(\varphi) \leq \frac{h}{2} \left( \sum_{i=0}^{N} b_i^2 \right)^{-\frac{1}{2}} \right\}
\geq \prod_{i=0}^{N} \left( 1 - 4 \frac{\text{Var}[\beta_i(y)]}{h^2 W} \left( \sum_{i=0}^{N} 2i+1 \right) \left( \sum_{i=0}^{N} \left( \frac{T}{2i+1} \right)^{\frac{1}{2}} \right)^2 \right)
\cdot \prod_{i=0}^{N} \left( 1 - 4 \sup_{0 < \varphi \leq T} \frac{\text{Var}[\beta_i(y)]}{h^2 W} \left( \sum_{i=0}^{N} b_i^2 \right)^2 \left( \sum_{i=0}^{N} \left( \frac{T}{2i+1} \right)^{\frac{1}{2}} \right)^{-2} \right).
$$

(62)

From (62) we can easily see that there must exist a $W_1$ such that when $W \geq W_1$

$$
P \{ \lambda(W, N) < h \} \geq 1 - \frac{\delta}{2}.
$$

(63)
Now, let us make \( \rho(W, N) \leq \varepsilon \). According to (55) and (59), we have

\[
P \{ \rho(W, N) \leq \varepsilon \} = P \{ |r_N(0)| R_W(0) + \sqrt{2} ||D的安全性_N(x)||_{L^2} \sup_{0 \leq \tau \leq T} R_W(y) + (1 - \lambda(W, N)) ||r_N(x)||_{L^2} \leq \varepsilon \} \\
\geq P \{ \{ |r_N(0)| R_W(0) \leq \frac{\varepsilon}{3} \} \cap \{ \sqrt{2} ||D的安全性_N(x)||_{L^2} \sup_{0 \leq \tau \leq T} R_W(y) \leq \frac{\varepsilon}{3} \} \} \\
\cap \{ (1 - \lambda(W, N)) ||r_N(x)||_{L^2} \leq \frac{\varepsilon}{3} \} \\
= P \{ R_W(0) \leq \frac{\varepsilon}{3|r_N(0)|} \} \cdot P \{ \sup_{0 \leq \tau \leq T} R_W(y) \leq \frac{\varepsilon}{3\sqrt{2}||D的安全性_N(x)||_{L^2}} \} \\
\cdot P \{ (1 - \lambda(W, N)) ||r_N(x)||_{L^2} \leq \frac{\varepsilon}{3} \} \\
\geq \prod_{i=0}^{N} \left( 1 - 9 \frac{\text{Var}[eta_i(y)]}{\varepsilon^2 W} |r_N(0)|^2 \left( \sum_{i=0}^{N} \left( \frac{T}{2i+1} \right)^{\frac{1}{2}} \right)^2 \right) \\
\cdot \prod_{i=0}^{N} \left( 1 - 18 \frac{\text{Var}[eta_i(y)]}{\varepsilon^2 W} ||D的安全性_N(x)||_{L^2}^2 \left( \sum_{i=0}^{N} \left( \frac{T}{2i+1} \right)^{\frac{1}{2}} \right)^2 \right) \\
\cdot P \{ ||r_N(x)||_{L^2} \leq \frac{\varepsilon}{3(1-h)} \}.
\]

It can be easily seen that if we choose \( \varepsilon' = \frac{\varepsilon}{3(1-h)} \), then

\[
P \left\{ \left| |r_N(x)||_{L^2} \leq \frac{\varepsilon}{3(1-h)} \right\} = 1.
\]

As for the first two factors, obviously, we can always choose \( W \) so that their products will be arbitrarily close to 1. Thus, there must exist \( W_2 \) such that when \( W \geq W_2 \)

\[
P \{ \rho(W, N) \leq \varepsilon \} > 1 - \frac{\delta}{2}.
\]

Finally, if we choose \( W = \max \{ W_1, W_2 \} \), (63) and (66) hold and is then completed. \( \square \)

From (62) and (64) we see that there are some parameters that affect the geometric convergence. These are the total number of to \( N \); the slab thickness, \( T \); and the angle \( \theta \) of motion involved in the of \( b_i \). As for how \( \Sigma_a \) affects the convergence, it is not very obvious. \( \text{Var}[eta_i(y)] \) is not a monotone function of \( \Sigma_a \). However, we can see that the convergence does not depend on \( \Sigma_a \) and \( \Sigma_t \) individually on their difference. This is to be expected for this simple transport since Equation (1) only depends on \( \Sigma_t - \Sigma_a = \Sigma_a \). More complex dependencies on the problem data can be expected for more general transport problems.

From our probabilistic error analysis, it is clear that in order to \( \lambda(W, N) \) by a factor of, say 10, one must increase the number \( W \) walks by a factor of one hundred (on the average). This is not since the reduction in \( \lambda \) occurs simply as a result of increasing the number of random walks per adaptive stage, and is therefore subject to convergence rates as conventional (pseudorandom) Monte Carlo. More interesting, perhaps, is the question of how to determine a smallest \( W \)
that produces stable geometric convergence. The theory we have developed guarantees that such a minimal \( W \) exists, though it is equally clear from the complexity of the formulas governing it that it will not be an easy matter to compute it in terms of the problem input parameters.

Theorems 5 and 6 can be used for error analysis in the sense that they roughly provide an error bound for the solutions after each stage. However, because we have relied on the Chebyshev inequality heavily, we must anticipate that the error bounds will be rather loose. In order to obtain sharpened error bounds, more careful statistical analysis of the distribution of the random variables used should be carried out.

There are many other issues relating to error analysis that remain. One point that we have glossed over here is that the use of (46) for estimating \( a_n \) may be complicated to carry out. Instead, we may have to resolve the integral in (46) first by using some other methods (our experience shows that quasi-Monte Carlo methods are very good candidates to address this issue). In our later work, we will apply this idea to some more complicated transport problems.

References


4 Appendix A: The proof of Theorem 1

We first prove

\[
E[\beta(y)] = \int_0^T k(x, y) f(x) dx \quad \text{for} \quad y \in [0, T].
\]

(67)

We have

\[
E[\beta(y)] = \{ E \left[ \int_{\eta_m}^{\eta_{m+1}} f(x) dx \right] + \sum_{k=1}^{\infty} E \left[ \prod_{m=1}^{k} X_{\left[ \eta_m, \frac{\eta_m}{2}, \frac{\eta_m}{2} \right]} \right] \}
\]

\[
E \left[ H \left( T - \sum_{m=1}^{k} \xi_m - y \right) \int_{\eta_m}^{\eta_{m+1}} f(x) dx \right].
\]

(68)
Therefore,

\[
E[\beta(y)] = \int_0^\infty \frac{E_\theta}{\cos \theta} e^{-\frac{E_\theta}{\cos \theta} z_1} dz_1 \int_{z_1}^{\infty} f(x) dx + \sum_{k=1}^\infty \left( \frac{E_\theta}{\cos \theta} \right)^k \int_0^\infty \frac{E_\theta}{\cos \theta} e^{-\frac{E_\theta}{\cos \theta} z_1} \cdots \frac{E_\theta}{\cos \theta} e^{-\frac{E_\theta}{\cos \theta} z_k} dz_k+1 \cdot \int_{z_1}^{\infty} \cdots \int_{z_k}^{\infty} f(x) dx,
\]

\[
= \int_0^T \frac{E_\theta}{\cos \theta} e^{-\frac{E_\theta}{\cos \theta} z_1} dz_1 \int_{z_1}^{\infty} f(x) dx + \int_0^T \frac{E_\theta}{\cos \theta} e^{-\frac{E_\theta}{\cos \theta} z_1} dz_1 \int_{z_1}^{\infty} f(x) dx + \sum_{k=1}^\infty \left( \frac{E_\theta}{\cos \theta} \right)^k \int_0^T \cdots \int_{z_1}^{\infty} \cdots \int_{z_k}^{\infty} f(x) dx.
\]

Integration by parts leads to

\[
E[\beta(y)] = \int_T^\infty e^{-\frac{E_\theta}{\cos \theta}(z_1 - y)} f(x_1) dx_1 + \sum_{k=1}^\infty \left( \frac{E_\theta}{\cos \theta} \right)^k \int_T^\infty (z_1 - y)^k e^{-\frac{E_\theta}{\cos \theta}(z_1 - y)} f(x_1) dx_1.
\]

Now the computations are straightforward but tedious, and we fin

\[
E[\beta(y)] = \int_T^\infty e^{-\frac{F_\theta}{\cos \theta}(z_1 - y)} f(x_1) dx_1 + \sum_{k=1}^\infty \left( \frac{F_\theta}{\cos \theta} \right)^k \int_T^\infty (z_1 - y)^k e^{-\frac{F_\theta}{\cos \theta}(z_1 - y)} f(x_1) dx_1.
\]

On the other hand, a simple manipulation gives us the integr.

\[
k(x, y) = \frac{E_\theta}{\cos \theta} \int_0^\infty e^{-\frac{E_\theta}{\cos \theta}(x-z_1)} k(x_1, y) dx_1 + \int_0^\infty e^{-\frac{E_\theta}{\cos \theta}(z-x_1)} \delta(x_1),
\]

which leads to

\[
k(x, y) = H(x - y) e^{-\frac{E_\theta}{\cos \theta}(x-y)} + \sum_{k=1}^\infty \left( \frac{E_\theta}{\cos \theta} \right)^k \int_0^\infty H(x - x_1) e^{-\frac{E_\theta}{\cos \theta}(x-x_1)} dx_1 + \cdots + \left( \frac{E_\theta}{\cos \theta} \right)^k \int_0^\infty dx_1 \int_0^{x_2} H(x - x_1) e^{-\frac{E_\theta}{\cos \theta}(x-x_1)} dx_2 \cdots \int_0^{x_k} H(x - x_1) e^{-\frac{E_\theta}{\cos \theta}(x-x_1)} dx_k
\]

or

\[
k(x, y) = \sum_{k=0}^\infty \left( \frac{E_\theta}{\cos \theta} \right)^k \frac{(x-y)^k}{k!} H(x - y) e^{-\frac{E_\theta}{\cos \theta}(x-y)}.
\]
Multiplying both sides of (75) by \( f(x) \) and then integrating from 0 to \( T \) gives

\[
\int_0^T k(x, y) f(x) dx = \sum_{k=0}^{\infty} \left( \frac{\Sigma_x}{\cos \theta} \right)^k \int_y^T f(x) \frac{(x-y)^k}{k!} e^{-\frac{\Sigma_x}{\cos \theta} (x-y)} dx. \tag{76}
\]

Comparing (72) to (76), we obtain (67).

Next, we prove the variance formula (34). Define

\[
\sigma_t = \frac{\Sigma_t}{\cos \theta}, \quad \sigma_a = \frac{\Sigma_a}{\cos \theta}, \quad \sigma_s = \frac{\Sigma_s}{\cos \theta}. \tag{77}
\]

We will use two formulas that can be found in [8]: for any random variable \( X \) and \( Y \):

\[
E_X [X] = E_Y [E_X [X|Y]] \tag{78}
\]

and

\[
V_X [X] = E_Y [V_X [X|Y]] + V_Y [E_X [X|Y]]. \tag{79}
\]

The idea to prove (34) is to condition on the first track \( \xi_1 \) and apply the scattering/absorption random variable \( \eta_1 \) successively; this results in an ordinary differential equation satisfied by \( V[\beta(y)] \).

By conditioning on the first track \( \xi_1 \), we obtain

\[
V[\beta(y)] = E_\xi [V[\beta(y)|\xi_1]] + V_\xi [E_\beta [\beta(y)|\xi_1]], \tag{80}
\]

where, and hereafter, \( V_\theta \equiv V_{\beta(\cdot)}, E_\theta \equiv E_{V_{\beta(\cdot)}}, E_\xi \equiv E_{\xi_1}, \) and so on. Since \( \xi_1 \) is sampled from an exponential distribution, we obtain

\[
V[\beta(y)] = \int_0^{T-y} V[\beta(y)|\xi_1 = t] P(\xi_1 = t) dt + \int_{T-y}^{\infty} V[\beta(y)|\xi_1 = t] P(\xi_1 = t) dt + E_\xi [E_\beta [\beta(y)|\xi_1]]^2 - \{E_\xi [E_\beta [\beta(y)|\xi_1]]\}^2. \tag{81}
\]

On the right hand side, the second term vanishes because under the condition \( \xi_1 = t \) \((t \geq T-y)\), \( \beta(y) = \int_y^T f(x) dx \) is deterministic, while for the last term, \( E_\xi [E_\beta [\beta(y)|\xi_1]] = E_\beta [\beta(y)] \) according to (78). Applying (79) to \( V[\beta(y)|\xi_1 = t] \) with \( Y \equiv \eta_1 \), we obtain

\[
V[\beta(y)] = \int_0^{T-y} \{E_\beta [V[\beta(y)|\xi_1 = t, \eta_1]] \}
+ \int_0^{T-y} \{E_\beta [V[\beta(y)|\xi_1 = t]] \} \cdot \sigma_t e^{-\sigma_t t} dt
+ \int_0^{\infty} \{E_\beta [V[\beta(y)|\xi_1 = t]] \}^2 \cdot \sigma_t e^{-\sigma_t t} dt
+ \int_{T-y}^{\infty} \{E_\beta [V[\beta(y)|\xi_1 = t]] \}^2 \cdot \sigma_t e^{-\sigma_t t} dt
- \int_0^{T-y} \{E_\beta [E[\beta(y)|\xi_1 = t, \eta_1]] \} \cdot \sigma_t e^{-\sigma_t t} dt
- \int_0^{T-y} \{E_\beta [E[\beta(y)|\xi_1 = t]] \}^2 \cdot \sigma_t e^{-\sigma_t t} dt
- \int_{T-y}^{\infty} \{E_\beta [E[\beta(y)|\xi_1 = t]] \}^2 \cdot \sigma_t e^{-\sigma_t t} dt. \tag{82}
\]
According to (78), the fourth and fifth terms cancel out. The second vanishes since $\beta(y)$ is no longer random if the particle terminates within segment $t$. As for the first term, since the scattered after its first track, we know that it will then occupy another segment $y + t$. Thus, we have

$$
[\beta(y)|\xi_1 = t, \eta_1 < \frac{\sigma_x}{\sigma_t}] = \int_y^{y+t} f(x)dx + \beta(y + t).
$$

The third term of (82) will be calculated directly from the definition of the expectation, while in the sixth term, since the particle passes through only a single track $\xi_1$, the integrand does not depend on $t$ so the integral can be easily carried out:

$$
V_\beta[\beta(y)] = \sigma_x \int_0^{T-y} V_\beta[\beta(y + t)] e^{-\sigma t}dt + \sigma_t \int_0^{T-y} \left\{ \frac{\sigma_x}{\sigma_t} \left[ E_\beta[\beta(y)|\xi_1 = t, \eta_1 < \frac{\sigma_x}{\sigma_t}] \right] \right\}^2 e^{-\sigma t}dt + \frac{\sigma_x}{\sigma_t} \left[ E_\beta[\beta(y)|\xi_1 = t, \eta_1 > \frac{\sigma_x}{\sigma_t}] \right] \right\}^2 e^{-\sigma t}dt + \sigma_x \int_T^{T+y} e^{-\sigma t}dt \left\{ E_\beta[\beta(y)|\xi_1 = t, \eta_1 > \frac{\sigma_x}{\sigma_t}] \right\}^2 - \left\{ E_\beta[\beta(y)|\xi_1 = t, \eta_1 < \frac{\sigma_x}{\sigma_t}] \right\}^2 e^{-\sigma t}dt + \sigma_x \int_T^{T+y} \left\{ \int_y^{y+t} f(x)dx + E_\beta[\beta(y + t)] \right\}^2 e^{-\sigma t}dt + \sigma_x \int_T^{T+y} \left\{ \int_y^{y+t} f(x)dx \right\}^2 e^{-\sigma t}dt + e^{-\sigma(t-v)} \left\{ \int_y^{y+v} f(x)dx \right\}^2 - \left\{ E_\beta[\beta(y)|\xi_1 = t, \eta_1 < \frac{\sigma_x}{\sigma_t}] \right\}^2.
$$

Multiplying both sides of (84) by $e^{-\sigma t}$ and then moving the first right-hand side to the left-hand side, we obtain an integral for $V_\beta[\beta(y)] e^{-\sigma t}$:

$$
V_\beta[\beta(y)] e^{-\sigma t} - \sigma_x \int_T^{T+y} e^{-\sigma t}dt + \sigma_x \int_T^{T+y} \left\{ \int_y^{y+v} f(x)dx + E_\beta[\beta(y)|\xi_1 = t, \eta_1 < \frac{\sigma_x}{\sigma_t}] \right\}^2 e^{-\sigma t}dt + \sigma_x \int_T^{T+y} \left\{ \int_y^{y+v} f(x)dx \right\}^2 e^{-\sigma t}dt + e^{-\sigma(T-v)} \left\{ \int_T^{T+y} f(x)dx \right\}^2 - e^{-\sigma t} \left\{ E_\beta[\beta(y)|\xi_1 = t, \eta_1 < \frac{\sigma_x}{\sigma_t}] \right\}^2.
$$

Taking the derivative of both sides, we obtain an ordinary differential for $V_\beta[\beta(y)] e^{-\sigma t}$:

$$
\frac{d}{dy} (V_\beta[\beta(y)] e^{-\sigma t}) + \sigma_x V_\beta[\beta(y)] e^{-\sigma t} = -\sigma_x e^{-\sigma t} \left\{ E_\beta[\beta(y)] \right\}^2 - 2f(y) \int_y^{y+v} f(x) e^{-\sigma t}dx - 2\sigma_x f(y) \int_y^{y+v} E_\beta[\beta(y)] e^{-\sigma t}dx - \frac{d}{dy} \left( e^{-\sigma t} \left\{ E_\beta[\beta(y)] \right\}^2 \right).
$$
Solving this equation gives the variance

\[ V_\beta [\beta(y)] = 2e^{\sigma_\beta^2} \int_y^T f(t)e^{\sigma_\beta t} dt \int_t^T f(x)e^{-\sigma_\beta x} dx \]
\[ + 2\sigma_\beta e^{\sigma_\beta y} \int_y^T f(t)e^{\sigma_\beta t} dt \int_t^T E_\beta [\beta(x)] e^{-\sigma_\beta x} dx \]
\[ - \{E_\beta [\beta(y)]\}^2. \tag{87} \]

Returning to the original parameters, we obtain

\[ V_\beta [\beta(y)] = 2 \left( \frac{2^{\frac{3+i}{T}}}{T} \right)^2 e^{\frac{E_{\alpha^2}}{\Gamma}} \]
\[ \cdot \int_y^T P_t \left( \frac{2}{T} - 1 \right) e^{\frac{E_{\alpha^2}}{\Gamma}} dt \int_t^T P_t \left( \frac{2}{T} - 1 \right) e^{\frac{E_{\alpha^2}}{\Gamma}} dx \]
\[ + 2 \frac{E_{\alpha^2}}{\Gamma} \left( \frac{2^{\frac{3+i}{T}}}{T} \right)^2 e^{\frac{E_{\alpha^2}}{\Gamma}} \]
\[ \cdot \int_y^T P_t \left( \frac{2}{T} - 1 \right) e^{\frac{E_{\alpha^2}}{\Gamma}} dt \int_t^T E_\beta [\beta(x)] e^{\frac{E_{\alpha^2}}{\Gamma}} dx \]
\[ - \{E_\beta [\beta(y)]\}^2. \tag{88} \]

Using the result

\[ E_\beta [\beta(y)] = \int_y^T f(t)e^{-\sigma_\alpha (t-y)} dt, \tag{89} \]

(88) can be simplified to

\[ V_\beta [\beta(y)] = \sigma_\alpha e^{\sigma_\alpha y} \int_y^T e^{\sigma_\alpha t} \left( \int_t^T f(x)e^{-\sigma_\alpha x} dx \right)^2 dt. \tag{90} \]

The proof is completed.

5 Appendix B: The proof of Theorem 5

By generating \( W \) samples \( (\zeta^n_1, \zeta^n_2, \ldots, \zeta^n_W) \) of \( \zeta^n_1 \) (which can be generated from the same number of samples of \( \beta_1(y) \)), let us consider the difference

\[ \frac{1}{W} \sum_{w=1}^W \zeta^n_{iw} - a^n_1 \]
\[ = Q_0^n \frac{1}{W} \sum_{w=1}^W \beta_{iw}(0) - a^n_1 + \frac{1}{W} \sum_{w=1}^W \int_0^T S^n(y)\beta_{iw}(y) dy - a^n_1 \]
\[ = Q_0^n \left( \frac{1}{W} \sum_{w=1}^W \beta_{iw}(0) - \frac{2^{\frac{3+i}{T}}}{T} \int_0^T k(x,0)P_t\left( \frac{2}{T} - 1 \right) dx \right) \]
\[ + \int_0^T S^n(y) dy \left( \frac{1}{W} \sum_{w=1}^W \beta_{iw}(y) - \frac{2^{\frac{3+i}{T}}}{T} \int_0^T k(x,y)P_t\left( \frac{2}{T} - 1 \right) dx \right). \tag{91} \]

Hence,

\[ \|\varphi^n_N(x) - \overline{\varphi^n_N}(x)\|^2_{L^2} \]
\[ = \sum_{i=0}^N \frac{T}{2^i + 1} \left( \frac{1}{W} \sum_{w=1}^W \zeta^n_{iw} - a^n_1 \right)^2 \]
\[ = \sum_{i=0}^N \frac{T}{2^i + 1} \left[ Q_0^n \left( \frac{1}{W} \sum_{w=1}^W \beta_{iw}(0) - \frac{2^{\frac{3+i}{T}}}{T} \int_0^T k(x,0)P_t\left( \frac{2}{T} - 1 \right) dx \right) \right. \]
\[ + \left. \int_0^T S^n(y) dy \left( \frac{1}{W} \sum_{w=1}^W \beta_{iw}(y) - \frac{2^{\frac{3+i}{T}}}{T} \int_0^T k(x,y)P_t\left( \frac{2}{T} - 1 \right) dx \right) \right]^2. \tag{92} \]
Use of the Minkowski (please refer to [7]) and Cauchy-Schwarz i gives

$$
\| \varphi_N^R(x) - \tilde{\varphi}_N^R(x) \|_{L^2} \\
\leq \sum_{i=0}^N \left( \frac{T}{2T+1} \right)^{\frac{3}{2}} \left| Q_0^R \left( \frac{1}{W} \sum_{u=1}^W \beta_{iu}(0) - 2i + 1 \int_0^T k(x,0) P_i(\frac{2\pi}{T}) \right) \right| + \int_0^T S^N(y) dy \left( \frac{1}{W} \sum_{u=1}^W \beta_{iu}(y) - 2i + 1 \int_0^T k(x,y) P_i(\frac{2\pi}{T}) - 1 \right) dx \right| \\
\leq |Q_0^R| \sum_{i=0}^N \left( \frac{T}{2T+1} \right)^{\frac{3}{2}} \left| \frac{1}{W} \sum_{u=1}^W \beta_{iu}(0) - 2i + 1 \int_0^T k(x,0) P_i(\frac{2\pi}{T}) \right| + \left( \int_0^T |S^N(y)|^2 \right)^{\frac{1}{2}} \\
\leq |Q_0^R| \sum_{i=0}^N \left( \frac{T}{2T+1} \right)^{\frac{3}{2}} \left| \frac{1}{W} \sum_{u=1}^W \beta_{iu}(y) - 2i + 1 \int_0^T k(x,y) P_i(\frac{2\pi}{T}) \right| + \left( \int_0^T |S^N(y)|^2 \right)^{\frac{1}{2}} \\
\leq |Q_0^R| \sum_{i=0}^N \left( \frac{T}{2T+1} \right)^{\frac{3}{2}} \left| \frac{1}{W} \sum_{u=1}^W \beta_{iu}(y) - 2i + 1 \int_0^T k(x,y) P_i(\frac{2\pi}{T}) \right| + \left( \int_0^T |S^N(y)|^2 \right)^{\frac{1}{2}}.$$

Or using (54), (93) can be written as

$$
\| \varphi_N^R(x) - \tilde{\varphi}_N^R(x) \|_{L^2} \leq |Q_0^R| R_w(0) + \left( \int_0^T |S^N(y)|^2 \right)^{\frac{1}{2}} \sup_{0 \leq y \leq T} \\
\sum_{i=0}^N \left( \frac{T}{2T+1} \right)^{\frac{3}{2}} \left| \frac{1}{W} \sum_{u=1}^W \beta_{iu}(y) - 2i + 1 \int_0^T k(x,y) P_i(\frac{2\pi}{T}) \right| + \left( \int_0^T |S^N(y)|^2 \right)^{\frac{1}{2}}.$$

At this point, we need to estimate \( |Q_0^R| \) and \( \int_0^T |S^N(y)|^2 \ dy \). estimate \( Q_0^R \). From

$$
\varphi^{n-1}(0) - Q_0^{n-1} = 0,
$$

obtaining

$$
\varphi^{n-1}(0) - \varphi_N^{n-1}(0) + \varphi_N^{n-1}(0) - \varphi_N^{n-1}(0) + \varphi_N^{n-1}(0) - Q_0^{n-1} = 0
$$

or

$$
\tau_N(0) + \varphi_N^{n-1}(0) - \varphi_N^{n-1}(0) - Q_0^R = 0.
$$

That is,

$$
Q_0^R = \varphi_N^{n-1}(0) - \varphi_N^{n-1}(0) + \tau_N(0).
$$

Thus,

$$
Q_0^R = \sum_{i=0}^N \left( a_i^{n-1} - \tilde{a}_i^{n-1} \right) P_i(-1) + \tau_N(0),
$$

or

$$
Q_0^R = \sum_{i=0}^N (-1)^i \left( a_i^{n-1} - \tilde{a}_i^{n-1} \right) + \tau_N(0)
$$

$$
= \sum_{i=0}^N (-1)^i \left( \frac{2i+1}{T} \right) \left( \frac{2i+1}{2T+1} \right)^{\frac{1}{2}} \left( \frac{T}{2T+1} \right)^{\frac{3}{2}} \left( a_i^{n-1} - \tilde{a}_i^{n-1} \right) + \tau_N(0).
$$
Therefore,

\[ |Q_0^n| \leq \sqrt{\sum_{i=0}^{N} \frac{2i+1}{T} \cdot \sum_{i=0}^{N} \frac{T}{2i+1} (a_i^{n-1} - \bar{a}_i^{n-1})^2 + |r_N(0)|} \]

\[ = \sqrt{\sum_{i=0}^{N} \frac{2i+1}{T} \cdot \| \varphi_N^{n-1}(x) - \bar{\varphi}_N^{n-1}(x) \|_{L^2} + |r_N(0)|}. \]  
(101)

Next, we estimate \( |S_n(y)| \). Since

\[ S_n(z) = - \left( \frac{d\varphi_N^{n-1}}{dx} + \frac{\Sigma_\theta}{\cos \theta} \varphi_N^{n-1}(x) \right) + S^{n-1}(x) \]
and

\[ 0 = - \left( \frac{d\varphi_N^{n-1}}{dx} + \frac{\Sigma_\theta}{\cos \theta} \varphi_N^{n-1}(x) \right) + S^{n-1}(x), \]  
(102)

subtracting (103) from (102) we obtain

\[ S_n(z) = - \left( \frac{d(\bar{\varphi}_N^{n-1} - \varphi^{n-1})}{dx} + \frac{\Sigma_\theta}{\cos \theta} (\bar{\varphi}_N^{n-1}(x) - \varphi^{n-1}(x)) \right) \]

\[ = - \frac{a_i^{n-1}}{\Sigma_\theta} \sum_{i=0}^{N} (\bar{a}_i^{n-1} - a_i^{n-1}) P_i \left( \frac{2\pi}{T} \right) - 1 \]

\[ - \frac{\Sigma_\theta}{\cos \theta} \sum_{i=0}^{N} (\bar{a}_i^{n-1} - a_i^{n-1}) P_i \left( \frac{2\pi}{T} \right) - 1 \]

\[ - \left( \frac{d(\varphi_N^{n-1} - \varphi^{n-1})}{dx} + \frac{\Sigma_\theta}{\cos \theta} (\varphi_N^{n-1}(x) - \varphi^{n-1}(x)) \right) \]

\[ = - \frac{a_i^{n-1}}{T} \sum_{i=0}^{N} (\bar{a}_i^{n-1} - a_i^{n-1}) P_i \left( \frac{2\pi}{T} \right) - 1 \]

\[ - \frac{\Sigma_\theta}{\cos \theta} \sum_{i=0}^{N} (\bar{a}_i^{n-1} - a_i^{n-1}) P_i \left( \frac{2\pi}{T} \right) - 1 \]

\[ + \left( \frac{d\varphi_N(x)}{dx} + \frac{\Sigma_\theta}{\cos \theta} \varphi_N(x) \right). \]

Since

\[ P_i \left( \frac{2\pi}{T} - 1 \right) = \sum_{j=0}^{i-1} c_{ij} P_j \left( \frac{2\pi}{T} - 1 \right), \]  
(105)

where \( c_{ij} \) are defined in (54), we have

\[ S_n(z) = - \frac{a_i^{n-1}}{T} \sum_{i=0}^{N} (\bar{a}_i^{n-1} - a_i^{n-1}) \sum_{j=0}^{i-1} c_{ij} P_j \left( \frac{2\pi}{T} - 1 \right) \]

\[ - \frac{\Sigma_\theta}{\cos \theta} \sum_{i=0}^{N} (\bar{a}_i^{n-1} - a_i^{n-1}) P_i \left( \frac{2\pi}{T} - 1 \right) + D_r(x) \]

\[ = - \frac{a_i^{n-1}}{T} \sum_{i=0}^{N} \left[ \sum_{j=i+1}^{N} (\bar{a}_j^{n-1} - a_j^{n-1}) c_{ji} \right] P_i \left( \frac{2\pi}{T} - 1 \right) \]

\[ - \frac{\Sigma_\theta}{\cos \theta} \sum_{i=0}^{N} (\bar{a}_i^{n-1} - a_i^{n-1}) P_i \left( \frac{2\pi}{T} - 1 \right) + D_r(x) \]

\[ = - \frac{a_i^{n-1}}{T} \sum_{i=0}^{N} \left[ \sum_{j=i+1}^{N} (\bar{a}_j^{n-1} - a_j^{n-1}) c_{ji} \right] P_i \left( \frac{2\pi}{T} - 1 \right) \]

\[ + \frac{\Sigma_\theta}{\cos \theta} (\bar{a}_i^{n-1} - a_i^{n-1}) P_i \left( \frac{2\pi}{T} - 1 \right) \]

\[ - \frac{\Sigma_\theta}{\cos \theta} (\bar{a}_i^{n-1} - a_i^{n-1}) P_i \left( \frac{2\pi}{T} - 1 \right) + D_r(x). \]

(106)

Noticing that \( D_r(x) \) (defined in (54)) is not orthogonal to the other terms in the expression of \( S_n(z) \) in general, we use the following inequality when we integrate the square of \( S_n(z) \)

\[ (a + b)^2 \leq 2 (a^2 + b^2). \]  
(107)
Thus we have (using the Minkowski inequality again)

\[
\left( \int_0^T |S^n(y)|^2 \, dy \right)^{\frac{1}{2}} \leq \frac{2}{\sqrt{T}} \sum_{i=0}^{N-1} \left\{ \frac{2}{T} \sum_{j=i+1}^N (\tilde{a}^n_j - a^n_j) \right\}^{\frac{1}{2}} + \frac{2T}{2N+1} \left\{ \frac{2}{|\cos \theta|} \left( \tilde{a}^n_N - a^n_N \right) \right\}^{\frac{1}{2}} + 2 \|D_rN(x)\| \\
\leq \frac{2}{\sqrt{T}} \sum_{i=0}^{N-1} \left\{ \frac{2}{T} \sum_{j=i+1}^N (\tilde{a}^n_j - a^n_j) \right\}^{\frac{1}{2}} + 2 \|D_rN(x)\| \\
\leq \sum_{i=0}^{N-1} \left( \frac{2T}{2i+1} \right)^{\frac{1}{2}} \left\{ \frac{2}{|\cos \theta|} \left( \tilde{a}^n_i - a^n_i \right) \right\}^{\frac{1}{2}} \left( \tilde{a}^n_N - a^n_N \right) + \sqrt{2} \|D_rN(x)\|
\]

Rearranging the terms in the summation gives

\[
\left( \int_0^T |S^n(y)|^2 \, dy \right)^{\frac{1}{2}} \leq \sum_{i=0}^N \left( \sum_{j=0}^{i-1} \left( \frac{2T}{2j+1} \right)^{\frac{1}{2}} \frac{1}{|\cos \theta|} c_{ij} \right) \left( \tilde{a}^n_i - a^n_i \right) + \sqrt{2} \|D_rN(x)\|
\]

Using Hölder's inequality we obtain

\[
\left( \int_0^T |S^n(y)|^2 \, dy \right)^{\frac{1}{2}} \leq \sum_{i=0}^N \left( \sum_{j=0}^{i-1} \left( \frac{2T}{2j+1} \right)^{\frac{1}{2}} |\tilde{a}^n_i - a^n_i| \right)^{\frac{1}{2}} + \sqrt{2} \|D_rN(x)\|_{L^2} \\
= \left( \sum_{i=0}^N \left( \sum_{j=0}^{i-1} \left( \frac{2T}{2j+1} \right)^{\frac{1}{2}} |\tilde{a}^n_i - a^n_i| \right)^{\frac{1}{2}} \right)^{\frac{1}{2}} \|\nu^{i-1}_N(x) - \tilde{\nu}^{i-1}_N(x)\|_{L^2} + \sqrt{2} \|D_rN(x)\|_{L^2},
\]

where \( b_t \)'s are defined in (54).
272

From (94), (101) and (110) we obtain

\[
\| \tilde{\varphi}_N^n(x) - \varphi_N^n(x) \|_{L^2} \\
\leq \left( Q_0^T R_W(0) + \left( \int_0^T |S_n(y)|^2 \, dy \right)^{\frac{1}{2}} \cdot \sup_{0 < y \leq T} R_W(y) \cdot \frac{1}{N} \right) \cdot \tilde{\varphi}_N^n(x) \\
\leq \left( \sqrt{\sum_{i=0}^{N-1} \| \varphi_N^{-1}(x) - \tilde{\varphi}_N^{-1}(x) \|_{L^2}^2} + |r_N(0)| \right) \cdot R_W(0) \\
+ \left( \left( \sum_{i=0}^{N-1} b_i^2 \right)^{\frac{1}{2}} \cdot \| \varphi_N^{-1}(x) - \tilde{\varphi}_N^{-1}(x) \|_{L^2} + \sqrt{2} \cdot \| Dr_N(x) \|_{L^2} \right) \\
\cdot \sup_{0 < y \leq T} R_W(y) \\
= \left( \sqrt{\sum_{i=0}^{N-1} \frac{2^{i+1}}{N} R_W(0) + \left( \sum_{i=0}^{N-1} b_i^2 \right)^{\frac{1}{2}} \sup_{0 < y \leq T} R_W(y) \right) \\
\cdot \| \varphi_N^{-1}(x) - \tilde{\varphi}_N^{-1}(x) \|_{L^2} \\
+ |r_N(0)| \cdot R_W(0) + \sqrt{2} \cdot \| Dr_N(x) \|_{L^2} \sup_{0 < y \leq T} R_W(y) \right). 
\]

(111)

Using Lemma 3 we obtain

\[
\| \tilde{\varphi}_N^n(x) - \varphi_N^n(x) \|_{L^2} \leq \lambda(W, N) \left\| \phi(x) - \tilde{\varphi}_N^{-1}(x) \right\|_{L^2} + \rho'(W, N), 
\]

(112)

...here

\[
\lambda(W, N) = \sqrt{\sum_{i=0}^{N-1} \frac{2^{i+1}}{N} R_W(0) + \left( \sum_{i=0}^{N-1} b_i^2 \right)^{\frac{1}{2}} \sup_{0 < y \leq T} R_W(y), \\
\rho'(W, N) = |r_N(0)| \cdot R_W(0) + \sqrt{2} \cdot \| Dr_N(x) \|_{L^2} \sup_{0 < y \leq T} R_W(y) \\
\cdot \| \tilde{\varphi}_N^n(x) - \varphi_N^n(x) \|_{L^2} \right). 
\]

(113)

Appealing to Lemma 3 again finally completes the proof. □