Variance reduction by means of deterministic computation: Collision estimate

Stefan Heinrich

Universität Kaiserslautern
Fachbereich Informatik
Postfach 30 49
67653 Kaiserslautern
email: heinrich@informatik.uni-kl.de

Abstract

We study the collision estimate of Monte Carlo methods for the solution of integral equations. A new variance technique is proposed and analyzed. It consists in the separation of the main part by constructing a neighboring equation based on deterministic numerical methods.

Keywords: Monte Carlo method, variance reduction, collision estimate.

1991 *Mathematics Subject Classification.* Primary 65C05

1 Introduction

In a recent paper, Heinrich (1995), a new variance reduction technique was introduced for the Monte Carlo solution of Fredholm integral equations. The idea, based on work in complexity theory, Heinrich and Mathé (1993), Heinrich (1994), consists in constructing a new equation sufficiently close to the original one and then applying standard schemes to both equations simultaneously. So the approach is a special case of the separation of main part (also called control variate) technique. As shown in Heinrich (1995), neighboring equations can be constructed by exploiting the system and the approximate solution of deterministic schemes of solving the equation. The gain in variance reduction can be controlled by the discretization error. Hence, by applying the Monte Carlo method with n samples, the overall error is essentially the deterministic discretization error multiplied by the classical Monte Carlo rate $n^{-1/2}$ (see Heinrich and Mathé (1993), Heinrich (1994) for a theoretical foundation of this statement). Considerable improvements are possible this way as experiments in Heinrich (1995) showed.

In Heinrich (1995), exact expressions for the variance of the method as well as estimates in terms of proximity of the two equations were obtained for one of the classical Monte

Carlo algorithms - the absorption scheme, under assumptions, typical for the situation of radiation transport. As a rule (compare Ermakov (1971), Mikhailov (1991a)), the absorption estimate is the technically simpler case. The question arises what happens for the other classical method - the collision scheme. This is the theme of the present paper. We present and analyze the new technique for the collision scheme and prove that the variance is dominated by the square of the proximity of the respective kernels and right hand sides in some function space norms. This means the results of Heinrich (1995) carry over to the collision estimate. The analysis is different and more complicated, but will be based on the results of Heinrich (1995), and we shall work in a similar framework. Nevertheless, we try to keep the present paper selfcontained by recalling all needed notions and results.

Once the variance analysis of Heinrich (1995) is extended to the collision estimate, we can apply the other results of that paper: The Galerkin method can be used to construct a neighboring equation and the proximity of kernels and right-hand sides can be estimated from certain parameters of that method. We do not repeat this here, but refer to Heinrich (1995) and Heinrich (1996) instead.

General references for Monte Carlo methods are Spanier and Gelbard (1969), Ermakov (1971), Sobol (1973), Ermakov and Mikhailov (1982), Kalos and Whitlock (1986), Ermakov et al. (1989), Mikhailov (1991a), Mikhailov (1991b), Sabelfeld (1991). Predecessors of our approach can be found in Ermakov (1971), ch. 6.2.5, Spanier (1979), Ermakov and Sipin (1985), Mikhailov (1991b), §5.10, Sabelfeld (1991), ch. 2.2.3, Lafortune and Willems (1994).

2 The Algorithm

We consider the Fredholm integral equation of the second kind

$$u(x) \quad = \quad \int_X K(x,y)u(y)\ d\mu(y) + f(x). \tag{1}$$

Here X is a non-empty set, endowed with a σ-algebra Σ of subsets and a finite positive, σ-additive measure μ on (X, Σ). If $1 \le s < \infty$, $L_s(X) = L_s(X, \Sigma, \mu)$ stands for the space of s-integrable functions and if $s = \infty$ for the space of essentially bounded functions. We assume that $f \in L_\infty(X)$ and that K is $\Sigma \times \Sigma$ measurable and satisfies

$$\|K\|_{L_\infty(L_1)} := \operatorname*{ess\,sup}_{x \in X} \int_X |K(x,y)|\ d\mu(y) < \infty.$$

This is equivalent to saying that the integral operator T_k defined for $g \in L_\infty(X)$ by

$$(T_K g)(x) = \int_X K(x,y)g(y)\ d\mu(y)$$

acts continuously in $L_\infty(X)$. Moreover,

$$\|T_K : L_\infty(X) \to L_\infty(X)\| = \|K\|_{L_\infty(L_1)}.$$

Our aim is to compute the value (u, Φ) of a functional $\Phi \in L_1(X)$ at the solution u of (1). Let $p_0(x)$ and $p(x,y)$ be non-negative measurable functions on (X, Σ) and $(X \times X, \Sigma \times \Sigma)$, respectively, satisfying

$$\int_X p_0(x)\ d\mu(x) = 1$$

and

$$\int_X p(x,y)\ d\mu(y) \le 1 \qquad\qquad (x \in X).$$

We shall assume that

$$\begin{array}{rcl} \mu\{\Phi(x) \neq 0 & \text{and} & p_0(x) = 0\} = 0 \\ \mu \times \mu\{K(x,y) \neq 0 & \text{and} & p(x,y) = 0\} = 0. \end{array}$$

Define φ and k by

$$\begin{array}{rcl} \Phi(x) & = & \varphi(x) p_0(x) \\ K(x,y) & = & k(x,y) p(x,y). \end{array}$$

We consider an absorbing Markov chain on X with density of initial distribution $p_0(x)$ and density of probability $p(x,y)$ of transition from x to y. We assume that the spectral radius of T_p in $L_\infty(X)$ is less than 1, hence almost all trajectories of the Markov chain are of finite length. Let

$$\xi = (x_0, x_1, \ldots, x_m)$$

be such a trajectory. The classical collision estimate of the von Neumann Ulam scheme is defined as

$$\eta(k, f, \xi) = \sum_{\ell=0}^{m} \varphi(x_0) k(x_0, x_1) \ldots k(x_{\ell-1}, x_\ell) f(x_\ell), \tag{2}$$

and it is well-known that if the spectral radius of T_{kp} is less than one, then

$$\mathbb{E}\eta(k, f, \xi) = \left((I - T_{kp})^{-1} f, \Phi\right) = (u, \Phi). \tag{3}$$

The approximation to (u, Φ) is then obtained by averaging N independent realizations of η. Given another pair of functions (h, g) (assumed to be close to (k, f), see below), such that the exact solution v of

$$v = T_{hp} v + g$$

is known, we define a new random variable

$$\zeta(k, f, h, g, \xi) = (v, \Phi) + \eta(k, f, \xi) - \eta(h, g, \xi). \tag{4}$$

Observe that

$$\mathbb{E}\zeta(k, f, h, g, \xi) = (v, \Phi) + \left((I - T_{kp})^{-1} f, \Phi\right) - \left((I - T_{hp})^{-1} g, \Phi\right) = (u, \Phi).$$

Averaging over N independent trajectories, we approximate (u, Φ) by

$$\frac{1}{N} \sum_{i=1}^{N} \zeta(k, f, h, g, \xi_i) = (v, \Phi) + \frac{1}{N} \sum_{i=1}^{N} (\eta(k, f, \xi_i) - \eta(h, g, \xi_i)). \tag{5}$$

Hence $\eta(h, g, \xi)$ serves as the main part (control variate). The quality of the new scheme is determined by the variance of ζ:

$$\mathbb{E}((u, \Phi) - \frac{1}{N} \sum_{i=1}^{N} \zeta(k, f, h, g, \xi_i))^2 = \frac{\text{Var } (\zeta)}{N}.$$

In the sequel we shall study this variance. We shall assume that

$$\varphi^2 p_0 \in L_1(X)$$

and that the functions k, h, f, g belong to the following classes: Fix $\alpha > 0, 0 < \gamma < 1$, $n_0 \in \mathbb{N}$ and define $\mathcal{K}(\alpha, \gamma, n_0)$ to be the set of all $k \in L_\infty(X^2)$ with

$$\|k\|_{L_\infty(X^2)} \leq \alpha$$

and

$$\|T_{k^2p}^{n_0} : L_\infty(X) \to L_\infty(X)\| \le \gamma.$$

For $\theta > 0$ we let

$$\mathcal{F}(\theta) = \{f \in L_\infty(X);\ \|f\|_{L_\infty(X)} \le \theta\}.$$

For $k, h \in \mathcal{K}(\alpha,\gamma,n_0)$ and $f, g \in \mathcal{F}(\theta)$ the collision estimate η and the new scheme ζ are well-defined random variables with finite second moment. (This is well-known, compare also Lemma 1 below, which we recall from Heinrich (1995) for the sake of completeness.

Lemma 1. *Let $\alpha > 0$, $0 < \gamma < 1$, $n_0 \in \mathbb{N}$. If $k \in \mathcal{K}(\alpha,\gamma,n_0)$, then*

(i) $I - T_{k^2p}$ is invertible in $L_\infty(X)$ and $\|(I - T_{k^2p})^{-1} : L_\infty(X) \to L_\infty(X)\| \le \beta_0$, where $\beta_0 = (1-\gamma)^{-1}\sum_{j=0}^{n_0-1}\alpha^{2j}$,

(ii) $\|kp\|_{L_\infty(L_1)} \le \|k^2p\|_{L_\infty(L_1)}^{1/2} \le \alpha$,

(iii) $\|T_{kp}^{n_0} : L_\infty(X) \to L_\infty(X)\| \le \gamma$,

(iv) $I - T_{kp}$ is invertible in $L_\infty(X)$ and $\|(I - T_{kp})^{-1} : L_\infty(X) \to L_\infty(X)\| \le \beta_1$, where $\beta_1 = (1-\gamma)^{-1}\sum_{j=0}^{n_0-1}\alpha^{j}$.

If $k, h \in \mathcal{K}(\alpha,\gamma,n_0)$, then

(v) $\|T_{khp}^{n_0} : L_\infty(X) \to L_\infty(X)\| \le \gamma$, and $\|(I - T_{khp})^{-1} : L_\infty(X) \to L_\infty(X)\| \le \beta_0$.

3 Variance of the new scheme

The first theorem provides a complicated, but exact expression for the variance of the new estimators ζ. Later on, we shall derive simpler forms yielding upper bounds for the variance.

Proposition 2. *Let $\alpha, \theta > 0$, $0 < \gamma < 1$, $n_0 \in \mathbb{N}$. Suppose that $k, h \in \mathcal{K}(\alpha,\gamma,n_0)$ and $f, g \in \mathcal{F}(\theta)$. Then the variance of $\zeta(k,f,h,g,\xi)$ can be expressed by the following formula:*

$$
\begin{aligned}
\mathrm{Var}(\zeta) \;=\; & \Big((I-T_{k^2p})^{-1}(2f(I-T_{kp})^{-1}f-f^2) \\
& -(I-T_{khp})^{-1}(2f(I-T_{hp})^{-1}g-fg) \\
& -(I-T_{khp})^{-1}(2g(I-T_{kp})^{-1}f-fg) \\
& +(I-T_{h^2p})^{-1}(2g(I-T_{hp})^{-1}g-g^2),\varphi^2p_0\Big) \\
& -\Big((I-T_{kp})^{-1}f-(I-T_{hp})^{-1}g,\varphi p_0\Big)^2 .
\end{aligned}
$$

Proof. We represent the random variable $\zeta-(v,\Phi)$ by an infinite sum of random variables

$$
\zeta-(v,\Phi)=\eta(k,f,\xi)-\eta(h,g,\xi)=\sum_{n=0}^{\infty}\zeta_n \tag{6}
$$

where

$$
\zeta_n=\varphi(x_0)(k(x_0,x_1)\cdots k(x_{n-1},x_n)f(x_n)-h(x_0,x_1)\cdots h(x_{n-1},x_n)g(x_n))
$$

if $n\le m(\xi)$ (the length of ξ) and $\zeta_n=0$ if $n>m(\xi)$. Hence the sum (6) is in fact almost surely finite. For arbitrary $m,n\in\mathbb{N}$ with $n\ge m$ we have

$$
\begin{aligned}
& \mathbb{E}\zeta_m\zeta_n \\
&= \int_{X^{n+1}}\varphi(x_0)\left(k(x_0,x_1)\cdots k(x_{m-1},x_m)f(x_m)-h(x_0,x_1)\cdots h(x_{m-1},x_m)g(x_m)\right) \\
&\quad \times\varphi(x_0)\left(k(x_0,x_1)\cdots k(x_{n-1},x_n)f(x_n)-h(x_0,x_1)\cdots h(x_{n-1},x_n)g(x_n)\right) \\
&\quad \times p(x_0)p(x_0,x_1)\cdots p(x_{n-1},x_n)\;d\mu(x_0)\cdots d\mu(x_n) \\
&= (T^m_{k^2p}(fT^{n-m}_{kp}f)-T^m_{khp}(fT^{n-m}_{hp}g)-T^m_{hkp}(gT^{n-m}_{kp}f)+T^m_{h^2p}(gT^{n-m}_{hp}g)\;,\;\varphi^2p_0).
\end{aligned} \tag{7}
$$

From (6) we get

$$
\mathbb{E}\left(\sum_{n=0}^{\infty}\zeta_n\right)^2=\mathbb{E}\sum_{m,n=0}^{\infty}\zeta_m\zeta_n=2\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\mathbb{E}\zeta_m\zeta_n-\sum_{m=0}^{\infty}\mathbb{E}\zeta_m^2, \tag{8}
$$

where the operations on infinite series are justified, since the series (6) is absolutely convergent in the square mean (i.e. in the norm of L_2 over the probability space of the Markov chain). This follows from (7), the assumptions of the theorem and Lemma 1. Now we combine (7) and (8). We apply the summation of (8) to each summand of (7). For the first one, we obtain

$$\sum_{m=0}^{\infty}\sum_{n=m}^{\infty} T_{k^2p}^m (f T_{kp}^{n-m} f) = \sum_{m=0}^{\infty} T_{k^2p}^m \left(f \sum_{\ell=0}^{\infty} T_{kp}^{\ell} f \right) = (I - T_{k^2p})^{-1}(f(I - T_{kp})^{-1} f)$$

and

$$\sum_{m=0}^{\infty} T_{k^2p}^m (f^2) = (I - T_{k^2p})^{-1}(f^2).$$

The remaining terms of (7) can be handled analogously. Hence we get

$$\begin{aligned} \mathbb{E}(\eta(k,f,\xi) - \eta(h,g,\xi))^2 \quad = \quad & ((I - T_{k^2p})^{-1}(2f(I - T_{kp})^{-1} f - f^2) \\ & -(I - T_{khp})^{-1}(2f(I - T_{hp})^{-1} g - fg) \\ & -(I - T_{khp})^{-1}(2g(I - T_{kp})^{-1} f - fg) \\ & +(I - T_{h^2p})^{-1}(2g(I - T_{hp})^{-1} g - g^2), \varphi^2 p_0) \end{aligned} \qquad (9)$$

which together with (6) and (3) proves the proposition. □

Next we introduce some notation, which we need in the sequel:

$$\begin{aligned} \delta(k,h) &= \|T_{kp} - T_{hp} : L_\infty(X) \to L_\infty(X)\| \\ &= \operatorname*{ess\,sup}_{x \in X} \int_X |k(x,y) - h(x,y)| p(x,y)\, d\mu(y) \\ \varepsilon(k,h)^2 &= \|T_{(k-h)^2p} : L_\infty(X) \to L_\infty(X)\| \\ &= \operatorname*{ess\,sup}_{x \in X} \int_X (k(x,y) - h(x,y))^2 p(x,y)\, d\mu(y). \end{aligned}$$

Hölder's inequality implies

$$\delta(k,h) \le \varepsilon(k,h).$$

Theorem 3. *Let $\alpha, \theta > 0$, $0 < \gamma < 1$, and $n \in \mathbb{N}$. Then there exists a constant $c > 0$ such that for all $k, h \in \mathcal{K}(\alpha, \gamma, n_0)$ and $f, g \in \mathcal{F}(\theta)$*

$$\operatorname{Var}(\zeta(k,f,h,g,\xi)) \le c(\varepsilon(k,h)^2 + \|f - g\|_{L_\infty(X)}^2).$$

Proof. We have according to (4) and (9)

$$\begin{aligned}\mathrm{Var}(\zeta) &\leq \mathbb{E}(\eta(k,f,\xi)-\eta(h,g,\xi))^2\\ &\leq 2\ \mathbb{E}(\eta(k,f,\xi)-\eta(h,f,\xi))^2+2\ \mathbb{E}(\eta(h,f,\xi)-\eta(h,g,\xi))^2\\ &= 2(w_1+w_2,\varphi^2 p_0)\end{aligned} \tag{10}$$

with

$$\begin{aligned}w_1 &= (I-T_{k^2p})^{-1}(2f(I-T_{kp})^{-1}f-f^2)\\ &\quad -(I-T_{khp})^{-1}(2f(I-T_{hp})^{-1}f-f^2)\\ &\quad -(I-T_{khp})^{-1}(2f(I-T_{kp})^{-1}f-f^2)\\ &\quad +(I-T_{h^2p})^{-1}(2f(I-T_{hp})^{-1}f-f^2)\end{aligned} \tag{11}$$

and

$$\begin{aligned}w_2 &= (I-T_{h^2p})^{-1}(2f(I-T_{hp})^{-1}f-f^2\\ &\qquad -(2f(I-T_{hp})^{-1}g-fg)\\ &\qquad -(2g(I-T_{hp})^{-1}f-fg)\\ &\qquad +2g(I-T_{hp})^{-1}g-g^2)\\ &= (I-T_{h^2p})^{-1}(2(f-g)(I-T_{hp})^{-1}(f-g)-(f-g)^2).\end{aligned} \tag{12}$$

It is convenient to rewrite w_1 as

$$w_1 = w_{11}+w_{12}$$

with

$$w_{11} = ((I-T_{k^2p})^{-1}-2(I-T_{khp})^{-1}+(I-T_{h^2p})^{-1})(2f(I-T_{kp})^{-1}f-f^2) \tag{13}$$

and

$$\begin{aligned}w_{12} &= (I-T_{khp})^{-1}(2f(I-T_{kp})^{-1}f-2f(I-T_{hp})^{-1}f)\\ &\quad +(I-T_{h^2p})^{-1}(2f(I-T_{hp})^{-1}f-2f(I-T_{kp})^{-1}f)\\ &= ((I-T_{khp})^{-1}-(I-T_{h^2p})^{-1})(2f((I-T_{kp})^{-1}-(I-T_{hp})^{-1})f)\\ &= (I-T_{khp})^{-1}(T_{khp}-T_{h^2p})(I-T_{h^2p})^{-1}\\ &\quad (2f(I-T_{kp})^{-1}(T_{kp}-T_{hp})(I-T_{hp})^{-1}f)\\ &= (I-T_{khp})^{-1}T_{(k-h)hp}(I-T_{h^2p})^{-1}\\ &\quad (2f(I-T_{kp})^{-1}T_{(k-h)p}(I-T_{hp})^{-1}f).\end{aligned} \tag{14}$$

We need relation (27) of Heinrich (1995):

$$\|T_{(k-h)hp} : L_\infty(X) \to L_\infty(X)\| = \|(k-h)hp\|_{L_\infty(L_1)} \le \alpha\delta(k,h).$$

With the notation of Lemma 1 we get from (12) and (14)

$$\begin{aligned}
\|w_2\|_{L_\infty(X)} &\le \beta_0(2\beta_1+1)\|f-g\|^2_{L_\infty(X)} \\
\|w_{12}\|_{L_\infty(X)} &\le \beta_0\alpha\delta(k,h)\beta_0 2\theta\beta_1\delta(k,h)\beta_1\theta \\
&= 2\alpha\beta_0^2\beta_1^2\theta^2\delta(k,h)^2 \le 2\alpha\beta_0^2\beta_1^2\theta^2\varepsilon(k,h)^2.
\end{aligned}$$

For the estimate of (13) we recall relations (26) and (28) from Heinrich (1995)

$$\begin{aligned}
&(I-T_{k^2p})^{-1} - 2(I-T_{khp})^{-1} + (I-T_{h^2p})^{-1} \\
= \; &(I-T_{k^2p})^{-1}(T_{(k-h)^2p} + T_{(h^2-k^2)p}(I-T_{h^2p})^{-1}T_{(h-k)hp})(I-T_{khp})^{-1} \qquad (15)
\end{aligned}$$

and

$$\|T_{(h^2-k^2)p} : L_\infty(X) \to L_\infty(X)\| \le 2\alpha\delta(k,h).$$

This gives

$$\begin{aligned}
\|w_{11}\| &\le \beta_0\left(\varepsilon(k,h)^2 + 2\alpha^2\beta_0\delta(k,h)^2\right)\beta_0\|2f(I-T_{kp})^{-1}f - f^2\|_{L_\infty(X)} \\
&\le \beta_0^2\left(\varepsilon(k,h)^2 + 2\alpha^2\beta_0\delta(k,h)^2\right)(1+2\beta_1)\theta^2 \\
&\le \beta_0^2\theta^2(1+2\alpha^2\beta_0)(1+2\beta_1)\varepsilon(k,h)^2.
\end{aligned}$$

This completes the proof. □

Finally, we show that L_2-estimates can be obtained for the variance, once slightly stronger assumptions are imposed. We suppose that $\varphi^2 p_0 \in L_\infty(X)$ and

$$\operatorname*{ess\,sup}_{y\in X} \int_X p(x,y)\, d\mu(x) \le 1 \qquad (16)$$

which is the case, in particular, if $p(x,y)$ is symmetric (as e.g. the choice of the transition probability for the radiance equation in Heinrich (1995)). Define for $\alpha > 0$, $0 < \gamma, \gamma_1 < 1$, $n_0, n_1 \in \mathbb{N}$

$$\mathcal{K}^*(\alpha,\gamma,\gamma_1,n_0,n_1) = \left\{k \in \mathcal{K}(\alpha,\gamma,n_0) : \|T^{n_1}_{k^2p} : L_1(X) \to L_1(X)\| \le \gamma_1\right\}.$$

Lemma 4. *Let $\alpha > 0,\ 0 < \gamma, \gamma_1 < 1,\ n_0, n_1 \in \mathbb{N}$. If $k, h \in \mathcal{K}(\alpha, \gamma, \gamma_1, n_0, n_1)$, then*

(i) $I - T_{k^2p}$ is invertible in $L_1(X)$ and $\|(I - T_{k^2p})^{-1} : L_1(X) \to L_1(X)\| \le \beta_2$ with $\beta_2 = (1-\gamma_1)^{-1} \sum_{j=0}^{n_1-1} \alpha^{2j}$,

(ii) $I - T_{kp}$ is invertible in $L_1(X)$ and $\|(I - T_{kp})^{-1} : L_1(X) \to L_1(X)\| \le \beta_3$, with $\beta_3 = (1-\gamma_1)^{-1} \sum_{j=0}^{n_1-1} \alpha^{j}$,

(iii) $I - T_{khp}$ is invertible in $L_1(X)$ and $\|(I - T_{khp})^{-1} : L_1(X) \to L_1(X)\| \le \beta_2$.

Proof. For an integral operator T_K we have

$$\|T_K : L_1(X) \to L_1(X)\| = \|T_{K^*} : L_\infty(X) \to L_\infty(X)\|,$$

where $K^*(x,y) = K(y,x)$. Hence, if $k, h \in \mathcal{K}(\alpha, \gamma, \gamma_1, n_0, n_1)$, then $k^*, h^* \in \mathcal{K}_{p^*}(\alpha, \gamma_1, n_1)$ where $\mathcal{K}_{p^*}$ denotes the class $\mathcal{K}$, based on p^* instead of p. Now Lemma 4 is a direct consequence of Lemma 1. □

Finally, we denote

$$\sigma(k,h)^2 \;=\; \int_X \int_X (k(x,y) - h(x,y))^2 p(x,y)\ d\mu(x) d\mu(y).$$

Theorem 5. *Given $\alpha, \theta > 0,\ 0 < \gamma, \gamma_1 < 1,\ n_0, n_1 \in \mathbb{N}$, there exists a constant $c > 0$ such that for $k, h \in \mathcal{K}^*(\alpha, \gamma, \gamma_1, n_0, n_1)$ and $f, g \in \mathcal{F}(\theta)$,*

$$\mathrm{Var}(\zeta(k,f,h,g,\xi)) \le c \left(\sigma(k,h)^2 + \|f - g\|^2_{L_2(X)}\right).$$

Proof. Lemma 1 and 4 yield

$$\begin{aligned} &\|(I - T_{h^2p})^{-1} : L_\infty(X) \to L_\infty(X)\| \le \beta_0, \text{and} \\ &\|(I - T_{h^2p})^{-1} : L_1(X) \to L_1(X)\| \le \beta_2 \end{aligned}$$

and hence, by the Riesz-Thorin interpolation theorem (see Triebel (1978)),

$$\|(I - T_{h^2p})^{-1} : L_2(X) \to L_2(X)\| \le (\beta_0 \beta_2)^{1/2}. \tag{17}$$

Using $(10-14)$ of the previous proof, we obtain

$$\mathrm{Var}(\zeta) \le 2 \left(\|w_{11}\|_{L_1(X)} + \|w_{12}\|_{L_1(X)} + \|w_2\|_{L_1(X)}\right) \|\varphi^2 p_0\|_{L_\infty(X)}.$$

First we estimate $\|w_{11}\|_{L_1(X)}$ on the basis of relations (13) and (15). We use (17) of the present paper and the following inequalities (35) and (37) of Heinrich (1995)

$$\|T_{(k-h)^2p} : L_\infty(X) \to L_1(X)\| \leq \sigma(k,h)^2$$
$$\|T_{(h^2-k^2)p}(I - T_{h^2p})^{-1}T_{(h-k)hp} : L_\infty(X) \to L_1(X)\| \leq 2\alpha^2(\beta_0\beta_2)^{1/2}\sigma(k,h)^2.$$

This gives

$$\|(I - T_{k^2p})^{-1} - 2(I - T_{khp})^{-1} + (I - T_{h^2p})^{-1} : L_\infty(X) \to L_1(X)\| \\ \leq \beta_0\beta_2(2\alpha^2(\beta_0\beta_2)^{1/2} + 1)\sigma(k,h)^2.$$

Moreover, from Lemma 1,

$$\|2f(I - T_{kp})^{-1}f - f^2\|_{L_\infty(X)} \leq (2\beta_1 + 1)\theta^2.$$

The last two relations combined with (13) yield

$$\|w_{11}\|_{L_1(X)} \leq \beta_0\beta_2\theta^2(2\beta_1 + 1)(2\alpha^2(\beta_0\beta_2)^{1/2} + 1)\sigma(k,h)^2.$$

Next we deal with w_{12}. From (14) we infer

$$\begin{aligned} \|w_{12}\|_{L_1(X)} &\leq \|(I - T_{khp})^{-1}T_{(k-h)hp}(I - T_{h^2p})^{-1} : L_2(X) \to L_1(X)\| \\ &\quad \times\|2f(I - T_{kp})^{-1}T_{(k-h)p}(I - T_{hp})^{-1}f\|_{L_2(X)}. \end{aligned} \tag{18}$$

We shall use the following relations, which are direct consequences of Hölder's inequality as shown in Heinrich (1995), inequality (36): For any $s \in L_\infty(X^2)$

$$\max\left(\|T_{sp} : L_\infty(X) \to L_2(X)\|\ ,\ \|T_{sp} : L_2(X) \to L_1(X)\|\right) \\ \leq \left(\int_{X\times X} s(x,y)^2 p(x,y)\ d\mu(x)\ d\mu(y)\right)^{1/2}. \tag{19}$$

This gives

$$\|T_{(k-h)hp} : L_2(X) \to L_1(X)\| \leq \alpha\sigma(k,h).$$

Consequently, the first factor of (18) satisfies

$$\begin{aligned}
&\|(I-T_{khp})^{-1}T_{(k-h)hp}(I-T_{h^2p})^{-1} : L_2(X)\to L_1(X)\| \\
&\quad\le\ \|(I-T_{khp})^{-1} : L_1(X)\to L_1(X)\|\cdot\|T_{(k-h)hp} : L_2(X)\to L_1(X)\| \\
&\qquad\times\|(I-T_{h^2p})^{-1} : L_2(X)\to L_2(X)\| \\
&\quad\le\ \alpha\beta_0^{1/2}\beta_2^{3/2}\sigma(k,h).
\end{aligned}$$

To estimate the second factor of (18), we note that (19) gives

$$\|T_{(k-h)p} : L_\infty(X)\to L_2(X)\|\le\sigma(k,h).$$

Moreover, arguing in the same way as for the derivation of (17), we conclude from Lemmas 1 and 4

$$\|(I-T_{kp})^{-1} : L_2(X)\to L_2(X)\|\le(\beta_1\beta_3)^{1/2}.$$

Hence,

$$\begin{aligned}
&\|2f(I-T_{kp})^{-1}T_{(k-h)p}(I-T_{hp})^{-1}f\|_{L_2(X)} \\
&\quad\le\ 2\theta^2\|(I-T_{kp})^{-1}T_{(k-h)p}(I-T_{hp})^{-1} : L_\infty(X)\to L_2(X)\| \\
&\quad\le\ 2\theta^2\|(I-T_{kp})^{-1} : L_2(X)\to L_2(X)\|\|T_{(k-h)p} : L_\infty(X)\to L_2(X)\| \\
&\qquad\times\|(I-T_{hp})^{-1} : L_\infty(X)\to L_\infty(X)\| \\
&\quad\le\ 2\theta^2\beta_1^{3/2}\beta_3^{1/2}\sigma(k,h).
\end{aligned}$$

So we get

$$\|w_{12}\|_{L_1(X)}\le 2\alpha\beta_0^{1/2}\beta_1^{3/2}\beta_2^{3/2}\beta_3^{1/2}\theta^2\sigma(k,h)^2.$$

Finally, we turn to w_2. According to (12),

$$\begin{aligned}
\|w_2\|_{L_1(X)}\ &\le\ \|(I-T_{h^2p})^{-1} : L_1(X)\to L_1(X)\| \\
&\qquad\times\|2(f-g)(I-T_{hp})^{-1}(f-g)-(f-g)^2\|_{L_1(X)} \\
&\le\ \beta_2\left(2\|(I-T_{hp})^{-1} : L_2(X)\to L_2(X)\|\|f-g\|_{L_2(X)}^2+\|f-g\|_{L_2(X)}^2\right) \\
&\le\ \beta_2(2(\beta_1\beta_3)^{1/2}+1)\|f-g\|_{L_2(X)}^2.
\end{aligned}$$

This completes the proof. □

In conclusion, let us mention a few consequences of the results proved above. They are of the same form (up to obvious modifications) as those in section 3 of Heinrich (1995). It follows that all consequences drawn in that paper hold true also for the collision estimate. In particular, Corollary 4 is valid, which shows that the optimal rate obtained in Heinrich and Mathé (1993) for the absorption estimate is also true for the collision estimate. Moreover, the analysis of section 4 of Heinrich (1995) carries over: Using some approximate deterministic Galerkin solution, one can construct neighboring h and g of increasing precision, resulting in decreasing variance of the new scheme based on the collision estimate. To estimate the variance, one can use Propositions 6 and 7 of Heinrich (1995). Finally, the applications to the radiance equation of computer graphics hold true as well. Details can be found in Heinrich (1995) and Heinrich (1996). With these variance reductions at hand, it would be interesting to study their consequences for the reduction of computational cost as done in Heinrich and Mathé (1993) for a model class of smooth kernels.

References

Ermakov, S. M. (1971). *The Monte Carlo Method and Related Problems (in Russian).* Nauka, Moscow.

Ermakov, S. M. and Mikhailov, G. A. (1982). *The Theory of Statistical Trials* (in Russian). Nauka, Moscow.

Ermakov, S. M., Nekrutkin, W. W., and Sipin, A. S. (1989). *Random Processes for Classical Equations of Mathematical Physics*, volume 34 of Math. Appl. (Soviet Ser.). Kluwer, Dodrecht.

Ermakov, S. M. and Sipin, A. S. (1985). A new scheme of the Monte Carlo method for solving problems of mathematical physics. *Soviet Phys. Doklady*, 30(11):935 – 936.

Heinrich, S. (1994). Random approximation in numerical analysis. In Bierstedt, K. D., Pietsch, A., Ruess, W. M., and Vogt, D., editors, *Functional Analysis*, pages 123 – 171. Marcel Dekker.

Heinrich, S. (1995). Variance reduction for Monte Carlo methods by means of deterministic numerical computation. *Monte Carlo Methods and Appl.*, 1:251 – 277.

Heinrich, S. (1996). Using the Galerkin method for variance reduction. In *2nd. St. Petersburg Workshop on Simulation*, pages 56 – 59. Publishing House of St. Petersburg University.

Heinrich, S. and Mathé, P. (1993). The Monte Carlo complexity of Fredholm integral equations. *Math. of Computation*, 60:257 – 278.

Kalos, M. H. and Whitlock, P. A. (1986). *Monte Carlo Methods, Vol. I: Basics.* Wiley, New York.

Lafortune, E. P. and Willems, Y. D. (1994). The ambient term as a variance reducing technique for Monte Carlo ray tracing. In *Proc. 5th Eurographics Workshop.* Darmstadt.

Mikhailov, G. A. (1991a). *Minimization of Computational Costs of Non-Analogue Monte Carlo Methods.* World Scientific, Singapore.

Mikhailov, G. A. (1991b). *Optimization of Weighted Monte Carlo Methods.* Springer.

Sabelfeld, K. K. (1991). *Monte Carlo Methods in Boundary Value Problems.* Springer, Berlin-Heidelberg-New York.

Sobol, I. M. (1973). *Computational Monte Carlo Methods* (in Russian). Nauka, Moscow.

Spanier, J. (1979). A new family of estimators for random walk problems. *J. Inst. Math. Appl.*, 23:1 – 31.

Spanier, J. and Gelbard, E. M. (1969). *Monte Carlo Principles and Neutron Transport Problems.* Addison-Wesley, Reading, Massachusetts.

Triebel, H. (1978). *Interpolation Theory, Function Spaces, Differential Operators.* Deutscher Verlag der Wissenschaften, Berlin.