JOURNAL OF COMPUTATIONAL PHYSICS 47, 375–386 (1982)

Efficient Variance-Reduction Transformations for the Simulation of a Ratio of Two Means: Application to Quantum Monte Carlo Simulations*

HEATHER L. GORDON, STUART M. ROTHSTEIN,† AND TIMOTHY R. PROCTOR‡

Department of Chemistry, Brock University, St. Catharines, Ontario, Canada L2S 3A1

Received October 16, 1981

Computationally efficient transformations (control variates) are developed and used in conjunction with importance sampling to reduce the variance of the Monte Carlo estimate of the variational energy, $\langle \psi \mathscr{H} \psi \rangle / \langle \psi \psi \rangle$, thereby providing tighter confidence intervals for the estimate. The transformations are applied to calculating the ground state energy of He using an explicitly correlated wavefunction; we report an effective method of importance sampling in Hylleraas coordinates as well. For our largest sample ($N = 60{,}000$), relative to importance sampling only, introduction of control variates reduces the variance of the energy estimate by 92%. This dramatic reduction in the variance was obtained with a cost factor of only 1.07 in CPU time on a Burroughs 6700 computer. In practical terms, the cost of these calculations is in generating the random numbers for the importance sampling and inverting the probability density functions; in comparison, the additional cost of control variates is slight.

I. INTRODUCTION

The motivation for the work reported in this paper arises from our interest in explicitly correlated wavefunctions [1], that is, wavefunctions which involve expansions in terms of the interelectronic coordinates, r_{ij}. These wavefunctions give rise to very complicated integrals; even their normalization integrals are not trivial to evaluate. Monte Carlo techniques for integration [2–4] provide one way of overcoming these difficulties, and indeed, currently there is considerable interest in applying the method to quantum chemical problems [5–13]. However, the Monte Carlo method converges slowly and has a large variance associated with it. The objective of this paper is to develop and apply computationally efficient transformations (control variates) to reduce this variance and to provide tighter confidence intervals for Monte Carlo estimates.

The quantity of interest is the energy expectation value, $\langle E \rangle$. In Section II we

* Work supported in part by the Natural Sciences and Engineering Research Council Canada.

† Author to whom correspondence should be addressed.

‡ Present Address: Chemistry Department, University of New Orleans, New Orleans, La. 70122.

0021-9991/82/090375-12$02.00/0

discuss the error in the Monte Carlo estimate of $\langle E \rangle$ when a quotient of means is employed, and in Section III we report a new, efficient variance reduction technique which employes control variates in order to reduce this error. It is, however, common to estimate $\langle E \rangle$ by a single mean as in the Metropolis method [14, 15], and more current methods (e.g., [6, 7]) estimate the exact energy itself by a single mean. In Section IV we discuss control variates for this case and explain why, for technical reasons, we do not employ them in this paper. Finally, there is a concrete example in Section V: the ground state of helium using an explicitly correlated wavefunction. Therein we present equations for efficient importance sampling in Hylleraas coordinates, and we also report and discuss the variance reductions which arise upon using a variety of control variates together with the CPU time penalties incurred.

II. Ratio Estimator and its Variance

The variational energy $\langle E \rangle$ is a quotient of two expectation values

$$\langle E \rangle \equiv \langle H \rangle / \langle S \rangle, \tag{1}$$

where

$$H \equiv \psi^* \mathscr{H} \psi \tag{2}$$

and

$$S \equiv \psi^* \psi \tag{3}$$

The Monte Carlo estimate of the quotient (using importance sampling [2–4]) is given by

$$\hat{E}_{\mathrm{MC}} \approx n^{-1} \sum_{i=1}^{n} \tilde{H}(R^{(i)}) \Big/ n^{-1} \sum_{i=1}^{n} \tilde{S}(R^{(i)}), \tag{4}$$

where

$$\tilde{H} \equiv g \psi^* \mathscr{H} \psi / \rho_H, \tag{5}$$

$$\tilde{S} \equiv g \psi^* \psi / \rho_S, \tag{6}$$

and $R^{(i)} \equiv (\mathbf{r}_1^{(i)}, \ldots, \mathbf{r}_N^{(i)})$ is the ith member of a set of n points in the configuration space of N electrons. In (5) and (6), g is the Jacobian of transformation from Cartesian coordinates; the densities ρ_H and ρ_S are arbitrary, normalized, positive probability distributions chosen to mimic H and S, respectively. Hence a nearly constant value is obtained for $\tilde{H}$ and $\tilde{S}$, and thus a reduced variance for each sum is also obtained.

The quotient (4) converges to $\langle E \rangle$ as $n \to \infty$; it has a negligible bias for large n [16], and the limiting distribution of (4) is asymptotically normal, subject to mild

conditions on the distributions of $\tilde{H}$ and $\tilde{S}$ [17]. Thus the variance of (4) asymptotically provides a confidence interval for $\langle E \rangle$; e.g.,

$$\text{Prob}(|\langle E \rangle - \hat{E}_{\text{MC}}| < 3\,\text{Var}^{1/2}(\hat{E}_{\text{MC}})) \approx 0.997 \tag{7}$$

The problem of estimating the variance of ratio estimators is still an area of active research (e.g., [18]), but for our purposes, the classical estimate [16] should suffice:

$$\widehat{\text{Var}}(\hat{E}_{\text{MC}}) \approx (n(n-1)\langle S \rangle^2)^{-1} \left(\sum \tilde{H}^2 - 2\hat{E}_{\text{MC}} \sum (\tilde{H}\tilde{S}) + \hat{E}^2_{\text{MC}} \sum \tilde{S}^2 \right) \tag{8}$$

For large n, one replaces the unknown $\langle S \rangle$ by its Monte Carlo estimate, $\bar{\tilde{S}}$, the denominator of (4). With this modification, (8) is equivalent to Coldwell [13, Eq. (12)] or, in terms of the variances of $\tilde{H}$ and $\tilde{S}$ [16]:

$$\widehat{\text{Var}}(\hat{E}_{\text{MC}}) \approx (n\bar{\tilde{S}}^2)^{-1} (\widehat{\text{Var}}(\tilde{H}) - 2\hat{E}_{\text{MC}} \widehat{\text{Cov}}(\tilde{H}, \tilde{S}) + \hat{E}^2_{\text{MC}} \widehat{\text{Var}}(\tilde{S})) \tag{9}$$

III. Control Variates

Consider the Monte Carlo estimate of an integral, say, $\langle g \rangle$

$$\bar{\tilde{g}} = n^{-1} \sum_{i=1}^{n} \tilde{g}(R^{(i)}), \tag{10}$$

where $\tilde{g} = g/\rho_g$, and ρ_g is an arbitrary, normalized and everywhere positive density distribution function of $R^{(i)}$ [2–4]. Let $\tilde{c}$ be a function of R with known expectation value $\langle \tilde{c} \rangle$, and define for each point in configuration space

$$g^T \equiv \tilde{g} + \beta(\tilde{c} - \langle \tilde{c} \rangle), \tag{11}$$

where β is a constant. It follows that $\overline{g^T} = \bar{\tilde{g}}$, and as β is arbitrary, we can pick $\beta = \beta^*$, the value which minimizes $\text{Var}(\overline{g^T})$. It is easy to show that [19]

$$\beta^* = -\,\text{Cov}(\tilde{g}, \tilde{c})/\text{Var}(\tilde{c}), \tag{12}$$

and for this choice of β,

$$\text{Var}(\overline{g^T}) = (1 - \rho^2_{\tilde{g}\tilde{c}})\,\text{Var}(\bar{\tilde{g}}), \tag{13}$$

where $\rho_{\tilde{g}\tilde{c}}$ is the correlation coefficient for $\tilde{g}$ and $\tilde{c}$. As $|\rho_{\tilde{g}\tilde{c}}| \leqslant 1$, a reduction in variance is obtained by using g^T rather than $\tilde{g}$ and hence improved confidence intervals for $\langle g \rangle$, estimated by the average value of g^T. From (13) it is apparent that as the correlation between $\tilde{c}$ and $\tilde{g}$ increases, the variance reduction is enhanced.

The function $\tilde{c}$ is a *control variate* for $\tilde{g}$; several references to applications for variance reduction are cited in [20]. Also of importance is Rosenberg's paper [21] on

multidimensional Monte Carlo integration using both importance sampling and control variates.

Extension of the concept of control variates to the problem at hand, reducing the variance of a quotient of means (4), is not obvious. The naive approach, to control both $\tilde{H}$ (5) and $\tilde{S}$ (6)

$$H^T(\beta_H) \equiv \tilde{H} + \beta_H(\tilde{C}_H - \langle \tilde{C}_H \rangle) \tag{14}$$

$$S^T(\beta_S) \equiv \tilde{S} + \beta_S(\tilde{C}_S - \langle \tilde{C}_S \rangle) \tag{15}$$

with β's given by expressions analogous to (12), is not satisfactory. The equation which results for $\mathrm{Var}(\hat{E}_{\mathrm{MC}})$ is too complex unless $\tilde{C}_H \propto \tilde{C}_S$, and in this case the variance reduction obtained is equal to that using (14) alone, with β_H^* chosen to minimize $\mathrm{Var}(\hat{E}_{\mathrm{MC}})$. ($\beta_H^*$ is given by [27, Eq. (2.8)].) This, of course, is clear upon inspection of (9); to minimize the variance we should control for the covariance of $\tilde{H}$ and $\tilde{S}$.

Iglehart and Lewis [22] also addressed the problem of reducing the variance of a ratio estimator. They concluded that only one control variate seems practical, but also, to be effective, the control must be highly correlated with the difference between the numerator and the denominator. However, we have derived expressions which enable us to use multiple controls effectively.

Consider the function

$$D \equiv H^T(\beta_H) - \hat{E}_{\mathrm{MC}} S^T(\beta_S), \tag{16}$$

where H^T and S^T are given by (14) and (15), respectively, and $\hat{E}_{\mathrm{MC}}$ by (4). It is readily apparent that $\overline{H^T} = \tilde{H}$, $\overline{S^T} = \tilde{S}$, and the value of $\hat{E}_{\mathrm{MC}}^T$,

$$\hat{E}_{\mathrm{MC}}^T \approx n^{-1} \sum_{i=1}^{n} H^T(R^{(i)}; \beta_H) \Big/ n^{-1} \sum_{i=1}^{n} S^T(R^{(i)}; \beta_S), \tag{17}$$

equals that of $\hat{E}_{\mathrm{MC}}$. Further, as we show in Appendix I,

$$\mathrm{Var}(D) = n\langle \tilde{S} \rangle^2 \, \mathrm{Var}(\hat{E}_{\mathrm{MC}}^T). \tag{18}$$

Now, if we choose β_H to minimize $\mathrm{Var}(\tilde{H})$

$$\beta_H^* = -\mathrm{Cov}(\tilde{H}, \tilde{C}_H)/\mathrm{Var}(\tilde{C}_H), \tag{19}$$

we are free, for the chosen β_H^* value, to choose β_S to minimize $\mathrm{Var}(D)$, and hence from (18), $\mathrm{Var}(\hat{E}_{\mathrm{MC}}^T)$. We show in Appendix II that the optimum choice is given by

$$\beta_S^* = [\mathrm{Cov}(\tilde{H}, \tilde{C}_S) + \beta_H^* \, \mathrm{Cov}(\tilde{C}_H, \tilde{C}_S) - \hat{E}_{\mathrm{MC}} \, \mathrm{Cov}(\tilde{S}, \tilde{C}_S)]/(\hat{E}_{\mathrm{MC}} \, \mathrm{Var}(\tilde{C}_S)). \tag{20}$$

Equations (14), (15), (19), and (20) define transformations for both the numerator and the denominator variables which reduce the variance of the simulated quotient of means of these variables.

IV. Remarks on the Metropolis Method

Expressing the variational energy as

$$\langle E \rangle = \int E(\mathbf{r}_1, \ldots, \mathbf{r}_N) P_N(\mathbf{r}_1, \ldots, \mathbf{r}_N) \, d\tau, \tag{21}$$

where

$$E \equiv \mathscr{H}\psi/\psi, \tag{22}$$

the Monte Carlo estimate of $\langle E \rangle$ is a single mean. Accordingly, the transformation analogous to (11) and (12) will reduce the variance of this estimate.

The Metropolis method is the commonly used method of sampling from P_N, thus avoiding the need for a normalized wavefunction. There are, however, important technical considerations with the Metropolis method ([5] and references cited therein), and these are shared, in part, by more recent methods (e.g., [6, 7]). We are concerned especially by the need to remove statistical correlation in the Monte Carlo sample which confounds any analysis of the effectiveness of the variance reduction based upon numerical results.

In applications we will focus upon the reduction of variance for a quotient of means, and the methods we propose easily avoid statistical correlation in the Monte Carlo sample.

V. Applications

Many trial wavefunctions, for which the variational energies are known [1], have been constructed to describe the ground state of atomic helium. We have selected an explicitly correlated Hylleraas He wavefunction as a test case for the application of the variance reduction transformations described above:

$$\psi = (1 + \delta t^2 + \gamma u) \exp(-\beta s), \tag{23}$$

where

$$s = r_1 + r_2, \qquad t = r_1 - r_2, \qquad u = r_{12}. \tag{24}$$

In (24), r_i is the distance of the ith electron from the nucleus, r_{ij} is the interelectronic distance, and

$$0 \leqslant t \leqslant u \leqslant s < \infty \tag{25}$$

After integrating over the Euler angles

$$d\tau = 2\pi^2 u(s^2 - t^2) \, ds \, dt \, du, \tag{26}$$

and allowing that the Hamiltonian is a even function of t, expressions for $\langle \psi \mathscr{H} \psi \rangle$ and $\langle \psi \psi \rangle$ are those given in Bethe and Salpeter [23, Eq. (32.8)].

It is readily apparent that we must generate a set of three random numbers corresponding to values of s, t, and u for each of the terms of the Monte Carlo sums that estimate $\langle \psi \mathscr{H} \psi \rangle$ and $\langle \psi \psi \rangle$. Given that the computer-generated pseudo-random numbers will be drawn from a standard univariate uniform distribution, we first must convert these to ones that are drawn from densities ρ_H and ρ_S used in the importance sampling. The technique used is like that described by Sasaki [24].

Given an r-dimensional probability density $\rho(x_1,..., x_r)$, define reduced density functions

$$\rho_k(x_1,..., x_k) \equiv \int \cdots \int \rho(x_1,..., x_r)\, dx_{k+1} \cdots dx_r, \qquad k = 1,..., r-1, \tag{27}$$

and define the conditional density functions

$$P_{k+1}(x_{k+1} \mid x_1,..., x_k) \equiv \rho_{k+1}(x_1,..., x_{k+1})/\rho_k(x_1,..., x_k). \tag{28}$$

Now since

$$\rho_1(x_1) \prod_{i=1}^{r-1} P(x_{i+1} \mid x_1,..., x_i) = \rho(x_1,..., x_r), \tag{29}$$

r-dimensional random numbers can be generated by first selecting x_1 from $\rho_1(x_1)$ and then selecting successively x_i, $i = 2,..., r$ from $P(x_i \mid x_1,..., x_{i-1})$.

In our case, it is possible to derive analytic expressions for $\rho_1(x_1)$ and $P(x_i \mid x_1,..., x_{i-1})$. Choose the density obtained from an STO

$$\rho_\alpha(s, t, u) = Nu(s^2 - t^2) \exp(-2\alpha s), \tag{30}$$

where α is a constant and N is the normalization constant

$$N^{-1} = 2\pi^2 \int_0^\infty ds \int_0^s du \int_0^u dt\, u(s^2 - t^2) \exp(-2\alpha s). \tag{31}$$

In order to select a set of random numbers from ρ_α (30), the following procedure is carried out:

(i) Choose three random numbers p_{s_0}, p_{t_0}, and p_{u_0} from a standard uniform probability density; these represent the probabilities of finding the two electrons at the point in configuration space with coordinates s_0, t_0, and u_0, respectively.

(ii) Integrate ρ_α (30) over t and u in order to obtain $\rho_1(s)$ and compute s_0 such that

$$p_{s_0} = \int_0^{s_0} ds\, \rho_1(s), \qquad 0 \leqslant s_0 < \infty. \tag{32}$$

(iii) Set $s = s_0$ in (30) and integrate over t to obtain $P_2(u \mid s_0)$. Now compute u_0 such that

$$p_{u_0} = \int_0^{u_0} du\, P_2(u \mid s_0), \qquad 0 \leqslant u_0 \leqslant s_0. \tag{33}$$

(iv) Set $s = s_0$ and $u = u_0$ in (30) to obtain $P_3(t \mid s_0, u_0)$. Compute $t = t_0$ such that

$$p_{t_0} = \int_0^{t_0} dt\, P_3(t \mid s_0, u_0), \qquad 0 \leqslant t_0 \leqslant u_0. \tag{34}$$

It is numerically straightforward to solve for s_0, t_0, and u_0 in (32)–(34), given ρ_α (30). We can express $\rho_1(s)$ as a χ^2-function, allowing s_0 to be computed from an inverse χ^2-function using IMSL [25] subroutine MDCHI. Probabilities $P_2(u|s_0)$ and $P_3(t|s_0, u_0)$ give rise to polynomial expressions in u_0 and t_0, respectively; the largest positive root of each of Eqs. (33) and (34) is used, solved for approximately by Newton's method.

Given a set of (s_0, t_0, u_0), one term in each of the sums in (4) is calculated; the density ρ_α (30) is used for both ρ_H (5) and ρ_S (6). This same set of random numbers is also used to calculate $\tilde{C}_H$ (35) and $\tilde{C}_S$(36) or (37).

$$\tilde{C}_H = g\psi_\alpha \mathscr{H} \psi_\alpha / \rho_\alpha, \tag{35}$$

$$\tilde{C}_S = g\psi_\alpha^2 / \rho_\alpha, \tag{36}$$

$$\tilde{C}_S = g\phi_\beta^2 / \rho_\alpha, \tag{37}$$

where

$$\psi_\alpha = \exp(-\alpha s), \tag{38}$$

$$\phi_\beta = \exp(-\beta(s^2 + t^2)). \tag{39}$$

Controls $\tilde{C}_H$ and $\tilde{C}_S$ correlate highly with $\tilde{H}$ and $\tilde{S}$, respectively. Note that ψ^2 is an STO with orbital parameter α, and that ϕ_β is a GTO with orbital parameter β. The means, variances, and covariances necessary to compute β_H^* (19), β_S^* (20), $H^T(\beta_H^*)$ (14), $S^T(\beta_S^*)$ (15), and hence $\hat{E}_{\mathrm{MC}}^T$ (17), are calculated cumulatively using an efficient moments routine [26].

Table I reports the Monte Carlo estimates of the variational energy and the variances of these estimates for various sample sizes and control variates; included is the variance where importance sampling has been used but no controls have been applied. We wish to demonstrate that using multiple controls (controlling both $\tilde{H}$ and $\tilde{S}$) is more effective than controlling either $\tilde{H}$ or $\tilde{S}$ separately.

Control variates for the denominator, $\tilde{C}_S$, use either an STO (38) or a GTO (39); the former gives consistently more variance reduction, as expected, since it more closely approximates the trial wavefunction (23).

TABLE I

Variance of Monte Carlo Energy Estimate for 1 ^1S He[a]

			Control both H and S[d]		Control H only[g]		Control S only[j]	
N	$\hat{E}_{MC}^{T}$[b]	No Controls[c]	C_S: STO[e]	C_S: GTO[f]	Minimize Var(H)[h]	Minimize Var($\hat{E}_{MC}$)[i]	C_S: STO[e]	C_S: GTO[f]
5,000	−3.010	8.5 (−3)	3.0 (−4)	3.0 (−4)	3.0 (−4)	3.0 (−4)	7.0 (−3)	7.3 (−3)
10,000	−2.947	3.1 (−3)	1.5 (−4)	1.4 (−4)	1.5 (−4)	1.5 (−4)	2.4 (−3)	2.5 (−3)
20,000	−2.933	1.7 (−3)	1.2 (−4)	1.2 (−4)	1.3 (−4)	1.3 (−4)	1.3 (−3)	1.4 (−3)
30,000	−2.893	9.5 (−4)	6.7 (−5)	6.7 (−5)	7.0 (−5)	7.0 (−5)	7.3 (−4)	7.7 (−4)
40,000	−2.894	6.8 (−4)	5.2 (−5)	5.2 (−5)	5.4 (−5)	5.4 (−5)	5.2 (−4)	5.5 (−4)
50,000	−2.899	5.3 (−4)	4.1 (−5)	4.1 (−5)	4.3 (−5)	4.3 (−5)	4.0 (−4)	4.2 (−4)
60,000	−2.908	4.4 (−4)	3.3 (−5)	3.3 (−5)	3.4 (−5)	3.4 (−5)	3.3 (−4)	3.5 (−4)

[a] Energy in a.u.; importance sampling is used for all calculations; $\delta = 0.13$, $\gamma = 0.30$, $\beta = 1.816$ in (23); the number of significant digits reported for the variances is based on the results of six independent runs for the largest sample.

[b] Eq. (17).

[c] Eqs. (4)–(6) describe energy estimate; Eq. (9) gives expression for variance.

[d] Eqs. (14), (15), (19), (20), and (35); Eq. (18) gives expression for variance with (B14).

[e] Eq. (36); $\alpha = 1.6875$ in (30), $\alpha' = 2.0$ in (38).

[f] Eq. (37); $\alpha = 1.6875$ in (30), $\beta = 8/9\pi$ in (39).

[g] Eqs. (14), (6), and (35); Eq. (18) gives expression for variance with (B14); $\alpha = 1.6875$ in (30); $\beta_S = 0$ in (B14).

[h] Eq. (19) used.

[i] [27, Eq. 2.8] used for β_H^*.

[j] Eqs. (5), (15), and analogous expression to [27, Eq. (2.8)] for β_S^*; Eq. (18) gives expression for variance with (B14). $\beta_H^* = 0$ in (B14).

When controls are applied only to $\tilde{S}$, β_S^* is chosen to minimize $\mathrm{Var}(\hat{E}_{\mathrm{MC}})$. When controls are applied only to $\tilde{H}$, the results using β_H^* chosen to minimize $\mathrm{Var}(\tilde{H})$ are compared against those using β_H^* chosen to minimize $\mathrm{Var}(\hat{E}_{\mathrm{MC}})$; these are comparable. It is evident that $\mathrm{Var}(\tilde{H})$ plays a greater role in $\mathrm{Var}(\hat{E}_{\mathrm{MC}})$ than does $\mathrm{Var}(\tilde{S})$. This conclusion is supported by comparing the variance reduction resulting from controlling $\tilde{H}$ versus $\tilde{S}$. This justifies our approach of first specifying β_H^* to minimize $\mathrm{Var}(H)$ and then finding β_S^* (20) to minimize $\mathrm{Var}(D)$ for the chosen β_H^* value.

For our largest sample ($N = 60{,}000$), relative to importance sampling only, introduction of STO control variates for both $\tilde{H}$ and $\tilde{S}$ ((35) and (36), respectively) reduces the variance of the energy estimate by 92%; this corresponds to a 72% reduction in the standard deviation. Indeed, the variance for the *largest* uncontrolled sample ($N = 60{,}000$) even exceeds that of the *smallest* ($N = 5{,}000$) controlled sample.

The 92% reduction in the variance over that for importance sampling solely was obtained with a cost factor of only 1.07 in CPU time on a Burroughs 6700 computer for the $N = 60{,}000$ sample. In practical terms, the cost of these calculations is in generating the random numbers for the importance sampling and in inverting the probability density functions; in comparison, the additional cost of control variates was slight.

Work is now in progress to extend these techniques to more complex systems, in particular, to the problem of He_2. While the sampling used to obtain points in configuration space described in the Applications section appears intractable in Hylleraas coordinates, it is not so in prolate spheroidal coordinates. As a guiding function, the Hartree Fock wavefunction should be of sufficient complexity for sampling purposes.

APPENDIX I: Derivation of Eq. (18)

Given the function D of Eq. (16), we wish to show that the variance of D is given by (18). By definition,

$$\mathrm{Var}(D) \equiv \langle D^2 \rangle - \langle D \rangle^2. \tag{A1}$$

Substituting (16) into (A1),

$$\mathrm{Var}(D) = \langle (H^T)^2 \rangle - 2\hat{E}_{\mathrm{MC}} \langle H^T S^T \rangle + \hat{E}_{\mathrm{MC}}^2 \langle (S^T)^2 \rangle - (\langle H^T \rangle - \hat{E}_{\mathrm{MC}} \langle S^T \rangle)^2. \tag{A2}$$

Upon expanding (A2) and rearranging terms,

$$\mathrm{Var}(D) = \mathrm{Var}(H^T) - 2\hat{E}_{\mathrm{MC}} \,\mathrm{Cov}(H^T, S^T) + \hat{E}_{\mathrm{MC}}^2 \,\mathrm{Var}(S^T). \tag{A3}$$

Now since

$$\langle S^T \rangle = \langle \tilde{S} \rangle, \tag{A4}$$

comparison of (A3) with (9) leads to (18), the desired result, given that

$$\hat{E}_{\mathrm{MC}} = \hat{E}^{T}_{\mathrm{MC}}. \tag{A5}$$

APPENDIX II: Derivation of Eq. (20)

Given the function D of Eq. (16), we wish to solve for β_S^* which minimizes $\mathrm{Var}(D)$ given a chosen β_H^*.

Substituting (14) and (15) for H^T and S^T, respectively, the variance of D becomes, upon setting $\beta_H = \beta_H^*$

$$\mathrm{Var}(D) = \mathrm{Var}(a - b), \tag{B1}$$

where

$$a \equiv \tilde{H} + \beta_H^*(\tilde{C}_H - \langle \tilde{C}_H \rangle) - \hat{E}_{\mathrm{MC}}\tilde{S}, \tag{B2}$$

$$b \equiv \hat{E}_{\mathrm{MC}}\beta_S(\tilde{C}_S - \langle \tilde{C}_S \rangle). \tag{B3}$$

Now,

$$\mathrm{Var}(a - b) = \mathrm{Var}(a) + \mathrm{Var}(b) - 2\,\mathrm{Cov}(a, b). \tag{B4}$$

Similarly,

$$\mathrm{Var}(a) = \mathrm{Var}(c) + \mathrm{Var}(d) + 2\,\mathrm{Cov}(c, d), \tag{B5}$$

where,

$$c \equiv \tilde{H} - \hat{E}_{\mathrm{MC}}\tilde{S}, \tag{B6}$$

$$d \equiv \beta_H^*(\tilde{C}_H - \langle \tilde{C}_H \rangle). \tag{B7}$$

It can be easily shown that

$$\mathrm{Var}(c) = \mathrm{Var}(\tilde{H}) - 2\hat{E}_{\mathrm{MC}}\,\mathrm{Cov}(\tilde{H}, \tilde{S}) + \hat{E}^2_{\mathrm{MC}}\,\mathrm{Var}(\tilde{S}) \tag{B8}$$

$$\mathrm{Var}(d) = \beta_H^{*2}\,\mathrm{Var}(\tilde{C}_H) \tag{B9}$$

$$\mathrm{Cov}(c, d) = \beta_H^*\,\mathrm{Cov}(\tilde{H}, \tilde{C}_H) - \hat{E}_{\mathrm{MC}}\beta_H^*\,\mathrm{Cov}(\tilde{S}, \tilde{C}_S) \tag{B10}$$

Therefore, using (B8)–(B10) in (B5),

$$\begin{aligned}\mathrm{Var}(a) = {} & \mathrm{Var}(\tilde{H}) - 2\hat{E}_{\mathrm{MC}}\,\mathrm{Cov}(\tilde{H}, \tilde{S}) + \hat{E}^2_{\mathrm{MC}}\,\mathrm{Var}(\tilde{S}) \\ & + \beta_H^*\,\mathrm{Var}(\tilde{C}_H) + 2\beta_H^*\,\mathrm{Cov}(\tilde{H}, \tilde{C}_H) - 2\hat{E}_{\mathrm{MC}}\beta_H^*\,\mathrm{Cov}(\tilde{S}, \tilde{C}_S).\end{aligned} \tag{B11}$$

Now,

$$\mathrm{Var}(b) = \hat{E}^2_{\mathrm{MC}}\beta_S^2\,\mathrm{Var}(\tilde{C}_S) \tag{B12}$$

and

$$\mathrm{Cov}(a, b) = \hat{E}_{\mathrm{MC}}\beta_S[\mathrm{Cov}(\tilde{H}, \tilde{C}_S) + \beta_H^* \mathrm{Cov}(\tilde{C}_H, \tilde{C}_S) - \hat{E}_{\mathrm{MC}} \mathrm{Cov}(\tilde{S}, \tilde{C}_S)]. \quad \text{(B13)}$$

Finally, substituting (B11)–(B13) into (B1) and (B4), we obtain

$$\begin{aligned}\mathrm{Var}(D) = {} & \mathrm{Var}(\tilde{H}) - 2\hat{E}_{\mathrm{MC}} \mathrm{Cov}(\tilde{H}, \tilde{S}) + \hat{E}_{\mathrm{MC}}^2 \mathrm{Var}(\tilde{S}) + \beta_H^{*2} \mathrm{Var}(\tilde{C}_H) \\ & + 2\beta_H^* \mathrm{Cov}(\tilde{H}, \tilde{C}_H) \\ & + \hat{E}_{\mathrm{MC}}^2 \beta_S^2 \mathrm{Var}(\tilde{C}_S) - 2\hat{E}_{\mathrm{MC}}\beta_S(\mathrm{Cov}(\tilde{H}, \tilde{C}_S) + \beta_H^* \mathrm{Cov}(\tilde{C}_H, \tilde{C}_S)) \\ & + 2\hat{E}_{\mathrm{MC}}^2 \beta_S \mathrm{Cov}(\tilde{S}, \tilde{C}_S) - 2\hat{E}_{\mathrm{MC}}\beta_H^* \mathrm{Cov}(\tilde{S}, \tilde{C}_H). \quad \text{(B14)}\end{aligned}$$

In order to choose a value for β_S which minimizes $\mathrm{Var}(D)$, we differentiate (B14) with respect to β_S, set the resulting expression to zero, and solve for β_S^*

$$\beta_S^* = [\mathrm{Cov}(\tilde{H}, \tilde{C}_S) + \beta_H^* \mathrm{Cov}(\tilde{C}_H, \tilde{C}_S) - \hat{E}_{\mathrm{MC}} \mathrm{Cov}(\tilde{S}, \tilde{C}_S)]/(\hat{E}_{\mathrm{MC}} \mathrm{Var}(\tilde{C}_S)). \quad \text{(B15)}$$

This expression is (20) in the text.

REFERENCES

1. N. C. HANDY, *in* "Theoretical Chemistry," International Review of Science: Physical Chemistry Series 2, Vol. 1, Butterworths, London, 1975.
2. I. M. SOBOL, "The Monte Carlo Method," Univ. of Chicago Press, Chicago/London, 1974.
3. J. M. HAMMERSLEY AND D. C. HANDSCOMB, "Monte Carlo Methods," Methuen, London, 1965.
4. P. J. DAVIS AND P. RABINOWITZ, "Methods of Numerical Integration," Chap. 5, Academic Press, New York, 1975.
5. Recent reviews include: K. BINDER, *in* "Monte Carlo Methods in Statistical Physics" (K. Binder, Ed.), Springer–Verlag, Berlin/New York, 1979; and D. M. CEPERLEY AND M. H. KALOS, *loc. cit.*
6. P. J. REYNOLDS, D. M. CEPERLEY, B. J. ADLER, AND W. A. LESTER, JR., "Monte Carlo Calculation of Molecular Ground State Energies," presented at American Conference on Theoretical Chemistry, Boulder, Colorado, 23 June, 1981. Sponsored by the National Science Foundation.
7. J. B. ANDERSON, *J. Chem. Phys.* **74** (1981), 6307.
8. R. E. LOWTHER AND R. L. COLDWELL, *Phys. Rev.* **22** (1980), 14.
9. J. OZAKI AND Y. TOMISHIMA, *J. Phys. Soc. Japan* **49** (1980), 1497.
10. J. B. ANDERSON, *J. Chem. Phys.* **73** (1980), 3897.
11. Y. TOMISHIMA AND J. OZAKI, *J. Comput. Phys.* **33** (1979), 382.
12. J. A. BARKER, *J. Chem. Phys.* **70** (1979), 2914.
13. R. L. COLDWELL, *Int. J. Quantum Chem.* **13** (1979), 705.
14. N. METROPOLIS AND S. ULAM, *J. Amer. Stat. Assoc.* **44** (1949), 335.
15. N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER, AND E. TELLER, *J. Chem. Phys.* **21** (1953), 1087.
16. W. A. COCHRAN, "Sampling Techniques," 3rd ed., Chap. 6, Wiley, New York, 1977.
17. H. CRAMÉR, "Mathematical Methods of Statistics," Chap. 17, Princeton Univ. Press, Princeton, 1963.
18. P. HEIDELBERGER AND P. A. W. LEWIS, *Commun. of ACM* **24** (1981), 260.
19. P. HEIDELBERGER, *IBM J. Res. Develop.* **24** (1980), 570.
20. S. S. LAVENBERG AND P. D. WELCH, *Management Science* **27** (1981), 322.

21. L. ROSENBERG, *SIAM J. Numer. Anal.* **4** (1967), 566.
22. D. L. IGLEHART AND P. A. W. LEWIS, *J. ACM* **26** (1979), 271.
23. H. A. BETHE AND E. E. SALPETER, "Quantum Mechanics of One- and Two-Electron Atoms," Springer–Verlag, Berlin, 1957.
24. T. SASAKI, *SIAM J. Numer. Anal.* **15** (1978), 938.
25. International Mathematical and Statistical Libraries, Inc. IMSL, 1980, Version 8, Houston, Texas.
26. C. C. SPICER, *Appl. Stat.* **21** (1972), 226.
27. D. L. IGLEHART AND P. A. W. LEWIS, "Variance Reduction for Regenerative Simulations, I: Internal Control and Stratified Sampling for Queues," Technical Report No. 86-22, Control Analysis Corp., Palo Alto, Calif., 1976.