A Modified Monte-Carlo Quadrature. II*

By Seymour Haber

1. Introduction. In a previous paper ("A modified Monte-Carlo quadrature," Math. Comp., v. 19, 1966, pp. 361-368; I shall refer to it below as "MMC") I proposed a Monte-Carlo quadrature procedure which incorporated a very simple form of stratified sampling and which produces somewhat more accurate estimates of integrals than would be obtained by simple Monte-Carlo. As applied to integration over the s-dimensional unit cube $G_s$ (i.e., $G_s$ is the set of all $x = (x_1, x_2, \ldots, x_s)$ such that $0 \leq x_i \leq 1$ for $i = 1, \ldots, s$) the method was as follows: Let $n_1, n_2, \ldots, n_s$ be positive integers, and partition $G_s$ into $N = n_1 \cdot n_2 \cdot \ldots \cdot n_s$ congruent subintervals by dividing the interval $[0,1]$ on the $x_i$-axis, for each $i$, into $n_i$ equal subintervals. Choose one point at random in each of these $N$ parts of $G_s$, and call those points—in any order—$x_1, x_2, \ldots, x_N$; then take the quantity

$$\frac{1}{N} \sum_{r=1}^{N} f(x_r)$$

as the estimate of the integral over $G_s$ of the function $f$.

In the simplest case the numbers $n_1, n_2, \ldots, n_s$ are all taken equal so that $G_s$ is partitioned into $N = K^s$ subcubes and $N = K^s$ for some positive integer $K$. The present paper proposes a second modification of simple Monte-Carlo quadrature, which makes further use of this partition of $G_s$.

If $N$ is large, so that each of the subcubes into which $G_s$ is divided is quite small one would generally suppose that the integrand would be a monotone function of each $x_i$ in most of the subcubes. Thinking along the lines of Hammersley and Morton’s method of antithetic variates [1], [2], it would seem that if $x$ is a point chosen at random in a subcube and $x'$ is the point symmetrically opposite it in that subcube, the quantity

$$\frac{f(x) + f(x')}{2}$$

should have smaller variance than $f(x)$. We therefore define the following estimate for $\int f$: Dividing $G_s$ into $N = K^s$ subcubes as above, call these subcubes, in some order, $A_1, A_2, \ldots, A_N$. For $1 \leq r \leq N$, let $c_r$ be the center of $A_r$, and choose a single point—call it $x_r$—at random in $A_r$; and let $x_r' = 2c_r - x_r$. Then our estimate is

$$J_2 = J_2(f, N) = \frac{1}{N} \sum_{r=1}^{N} \frac{f(x_r) + f(x_r')}{2} .$$

Received August 28, 1966.

* This work was supported in part by NASA Research Grant NSG-398 to the Computer Science Center, University of Maryland.
(The quantity (1) we shall denote by \( J_1 \), while "\( J_0 \)" will designate the simple Monte-Carlo estimate

\[
\frac{1}{N} \sum_{r=1}^{N} f(y_r)
\]

where the \( N \) points \( y_r \) are chosen at random in \( G_s \).)

2. Error Analysis. The discussion of the error of these estimates is based on the usual assumption of Monte-Carlo theory, that the points \( y_r \) and the points \( x_r \) can be regarded as independent (or at least pairwise independent) random variables; in our case each \( y_r \) is uniformly distributed over \( G_s \), while \( x_r \) is uniformly distributed over \( A_r \). Then \( J_0, J_1, \) and \( J_2 \) are random variables, and it is easy to see that all three have the integral of \( f \) as mean value:

\[
m(J_i) = \int_{G_s} f, \quad i = 0, 1, 2.
\]

(For \( J_2 \) we note that \( x_r' \) is, like \( x_r \), uniformly distributed over \( A_r \); so that

\[
m(J_2) = \frac{1}{N} \sum_{r=1}^{N} \left( \frac{1}{2} (m(f(x_r)) + m(f(x_r'))) \right) = \frac{1}{N} \sum_{r=1}^{N} N \int_{A_r} f = I.
\]

Numerical evaluation of an integral by the formula (1), (2), or (3) then can be regarded as taking a sample value of one of the \( J_i \)'s as an estimate of the mean \( I \); so that the standard deviation of that \( J \) may be taken as a measure of the error to be expected. On the usual assumption that \( J \) is approximately normally distributed, there would then be a probability of \( 1/2 \) that the error of the estimate is less than about (5/8) \( \sigma(J) \), while the chance that the error is greater than 2\( \sigma(J) \) would be less than 1/20.

We set

\[
\sigma_N(J_0) = \sigma(J_0(f,N)) , \quad \sigma_N(J_1) = \sigma(J_1(f,N)) , \quad \sigma_2N(J_2) = \sigma(J_2(f,N)) ,
\]

so that the subscript of \( \sigma \) is equal, in each case, to the number of evaluations of the integrand required for the estimate of \( I \). For the simple Monte-Carlo estimator \( J_0 \) we have the well-known result:

\[
\sigma_N(J_0) = d_0 N^{-1/2}, \quad d_0^2 = \int_{G_s} f^2 - \left( \int_{G_s} f \right)^2.
\]

In MMC I showed that \( \sigma_N(J_1) \leq \sigma_N(J_0) \) for any \( f \in L^2(G_s) \); for continuous \( f \), \( \sigma_N(J_1)/\sigma_N(J_0) \rightarrow 0 \) as \( N \rightarrow \infty \); and if the gradient \( \nabla f \) exists and is continuous on \( G_s \), then

\[
\sigma_N(J_1) = d_1 N^{-(1/2+1/s)} + o(N^{-(1/2+1/s)}) , \quad d_1 = \left( \int_{G_s} |\nabla f|^2 \right)^{1/2}.
\]

The first of these three results has no analogue for \( J_2 \)—it may happen that \( \sigma_2N(J_2) > \sigma_2N(J_0) \). However, for smooth integrands we have the following:

**Theorem.** If \( \partial^2 f/\partial x^i \partial x^j \) is continuous on \( G_s \), for \( 1 \leq i, j \leq s \), then

\[
\sigma_2N(J_2) = (d_2 + o(1)) N^{-(1/2+2/s)}
\]

where
Proof. From the (pairwise) independence of the \( x_r \), we have

\[
\sigma_{2N}^2(J_z) = \frac{1}{N^2} \sum_{i=1}^{N} \left( \frac{f(x_r) + f(x_r')}{2} \right) - \frac{1}{480} \sum_{i=1}^{8} \left( \frac{\partial^2 f}{\partial x^i \partial x^i} \right)^2.
\]

Let

\[
\tau_r = \sigma^2 \left( \frac{f(x_r) + f(x_r')}{2} \right) = N \int_{A_r} \left( \frac{f(x) + f(x')}{2} - N \int_{A_r} \frac{f(x) + f(x')}{2} \right) dx^2
\]

where \( x' \) is, as before, the point of \( A_r \), symmetrically opposite \( x \). For \( x \in A_r \), we may under the assumptions of Theorem, write

\[
f(x) = f(c_r) + \sum_{i=1}^{s} (x^i - c_r^i) \frac{\partial f}{\partial x^i}(c_r)
\]

\[
+ \frac{1}{2} \sum_{i,j=1}^{s} (x^i - c_r^i)(x^j - c_r^j) \frac{\partial^2 f}{\partial x^i \partial x^j}(c_r) + o(N^{-2/4}).
\]

For convenience I shall write \( f(x) \) for \( f(c_r) \), \( f_i^j \) for \( \partial f(c_r)/\partial x^i \partial x^j \) and \( \delta_i^j \) for \( x^i - c_r^i \). Then

\[
f(x) + f(x') = f + \frac{1}{2} \sum_{i,j=1}^{s} \delta_i^j f_i^j + o(N^{-2/4}).
\]

Now \( N \int_{A_r} \delta_i^j \) is zero if \( i \neq j \), while

\[
N \int_{A_r} (\delta_i^j)^2 = \frac{1}{12} N^{-2/4}.
\]

Therefore

\[
N \int_{A_r} \frac{f(x) + f(x')}{2} dx = f + \frac{N^{-2/4}}{24} \sum_{i=1}^{s} f_i + o(N^{-2/4}).
\]

From (10), (12), and (13) we get

\[
\tau = N \int_{A_r} \left( \frac{1}{2} \sum_{i,j=1}^{s} \delta_i^j f_i^j - \frac{N^{-2/4}}{24} \sum_{i=1}^{s} f_i + R \right)^2,
\]

where \( R = o(N^{-2/4}) \).

Temporarily writing \( \sum \) for the total of the two sums in (14) we have

\[
\tau = N \int_{A_r} R \sum^2 + 2N \int_{A_r} R \sum + N \int_{A_r} R^2.
\]

The last of these three terms is clearly \( o(N^{-4/4}) \); the same holds true for the second term, since \( \delta_i^i = O(N^{-1/4}) \) and so \( \sum = O(N^{-2/4}) \), and therefore

\[
\left| N \int_{A_r} R \sum \right| \leq \left( N \int_{A_r} \sum^2 \right)^{1/2} \left( N \int_{A_r} R^2 \right) = O(N^{-2/4}) \cdot o(N^{-2/4}).
\]

It follows that
Calculating the integral in (16) is facilitated by noting that \( \int \delta^i \delta^k \delta^L \) is zero unless \( i, j, k, L \) are equal in pairs, and that

\[
N \int_{A_r} (\delta^i)^2 (\delta^j)^2 = \frac{N^{-4/s}}{144} \text{ if } i \neq j, \quad N \int_{A_r} (\delta^i)^4 = \frac{N^{-4/s}}{80}.
\]

We finally obtain:

\[
\tau_r = \frac{N^{-4/s}}{288} \sum_{1 \leq i, j \leq s} f_{ij}^2 + \frac{N^{-4/s}}{720} \sum_{i=1}^s f_{ii}^2 + o(N^{-4/s}).
\]

From (9), (10) and (17) we obtain

\[
N^{1+4/s} \sigma^2_{2N}(J_2) = \frac{1}{288} \sum_{1 \leq i, j \leq s} \int_{G_s} \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)^2 + \frac{1}{720} \sum_{i=1}^s \int_{G_s} \left( \frac{\partial^2 f}{\partial x_i^2}(x) \right)^2 + o(1).
\]

Each inner sum in (18) is a Riemann sum, which approaches the integral (over \( G_s \)) of the function involved as \( N \to \infty \).

Therefore

\[
N^{1+4/s} \sigma^2_{2N}(J_2) = \frac{1}{288} \sum_{1 \leq i, j \leq s} \int_{G_s} \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)^2 + \frac{1}{720} \sum_{i=1}^s \int_{G_s} \left( \frac{\partial^2 f}{\partial x_i^2}(x) \right)^2 + o(1),
\]

and the theorem follows.

We see then that for smooth functions \( J_2 \) converges to the integral somewhat faster than \( J_0 \) or \( J_1 \). It is interesting to note, also, that \( J_2 \) is exact for linear integrands—so that the quantity \( d_2 \) is not a measure of the deviation of the integrand from constancy, as are \( d_0 \) and \( d_1 \), but of its deviation from linearity.

For less smooth integrands \( J_2 \) may not be superior to \( J_1 \); but for a large class of functions which may even be discontinuous, both \( J_1 \) and \( J_2 \) are asymptotically better than \( J_0 \), and in fact

\[
s_{2N}(J_i) = O(N^{-(1/2+1/2s)}), \quad i = 1, 2.
\]

To see this, let \( f \) be bounded in \( G_s \), and piecewise smooth in the following sense: \( G_s \) may be broken up into a finite number of regions, with smooth boundaries, such that in the interior of each region all the second partial derivatives of \( f \) are continuous and bounded. Then for \( N \) large the number of subcubes \( A_r \), which do not lie wholly inside a single one of these regions will be \( O(N^{-1/s}) \). For these subcubes \( \tau_r \) will be bounded, since \( f \) is bounded on \( G_s \); for the remaining subcubes, \( \tau_r = O(N^{-4/s}) \) as in the proof of Theorem 1. Thus

\[
\sigma^2_{2N}(J_2) = \frac{1}{N^2} \sum_{r=1}^N \tau_r \leq \frac{B_1 N^{1-s} + B_2 N(N^{-4/s})}{N^2}
\]

and (20) follows for \( J_2 \). In an exactly similar manner it follows for \( J_1 \), under the weaker assumption that the first partial derivatives of \( f \) are continuous and bounded in each region, using the proof of Theorem 3 of MMC.
3. Related Methods. Two modifications of the proposed quadrature method were considered, that seem worth some comment. The first is simply a generalization, in which $G_s$ is not necessarily partitioned into subcubes, but more generally into congruent subintervals in the manner described in the first paragraph of this paper. In MMC the estimate $J_1$ (there called $J''$) was treated in this more general manner, and it was seen that there might sometimes be a gain to partitioning the different coordinate axes differently. However, this does not seem to be the case for $J_2$. If, in the general case, we impose the "regularity" condition of MMC on the partitions of $G_s$, and carry through the reasoning of the proof of the Theorem above, we obtain in place of (19):

$$N\sigma_{2N}^2(J_2) = \frac{1}{288} \sum_{1 \leq i \leq s} \frac{1}{n_i^2} \sum_{j=1}^{n_i^2} \int_{\sigma_s} \left( \frac{\partial^2 f}{\partial x^i \partial x^j} \right)^2 \left( \frac{1}{n_i^2} \sum_{j=1}^{n_i^2} \int_{\sigma_s} \left( \frac{\partial^2 f}{\partial x^i \partial x^j} \right)^2 + o(N^{-1/s}) \right).$$

While, for given $N = n_1 n_2 \cdots n_s$, the total of the 2 sums in (19') might be made smaller by choosing the $n_i$ unequal than by choosing them equal, it is not likely that one would know in advance how to do this.

The second modification I considered was an attempt to halve the number of evaluations of the integrand needed for a given accuracy. For each $A_r$, instead of evaluating both $f(x_r)$ and $f(x_r')$, I proposed to choose at random a number $x_r$ between 0 and 1, set $x_r'' = \alpha x_r + (1 - \alpha)x_r'$, and calculate only $f(x_r'')$. However, the quantity

$$\frac{1}{N} \sum_{r=1}^{N} f(x_r'')$$

does not have mean value $I$ in general, since $x_r''$ is not uniformly distributed on $A_r$. This can be compensated for by introducing a weight factor: For $x = (x_1, x_2, \cdots, x_s)$ such that $-N^{-1/s} \leq x_i \leq N^{-1/s}$, let

$$\phi(x) = \left( 2N^{1/s} \max_{1 \leq i \leq s} |x_i| \right)^{-1}.$$

Then the estimator

$$J_4 = \frac{1}{N} \sum_{r=1}^{N} f(x_r'') \phi((2\alpha_r - 1)(x_r - c_r))$$

does have mean value $I$. However, it turns out to have infinite variance (though its expected absolute deviation from the mean is finite). In some numerical experiments on simple integrands, convergence to the integral was not apparent as $N$ was raised up to 65,536; so the method seems worthless.

4. Automatic Error Estimation. One of the attractive features of simple Monte-Carlo is the ease with which the error of the quadrature may be estimated. Once $f(y_r)$ is calculated for each $1 \leq r \leq N$, a single arithmetic operation produces $f^2(y_r)$, so that $\int_{\sigma_s} f^2$, and so $d_0$ and $\sigma_N(J_0)$, can be approximately evaluated in the course of the calculation of $J_0$. In MMC I pointed out a method of evaluating $\sigma_N(J_4)$ by, in effect,
calculating two independent samples of \( J_1 \). This method can be extended to \( J_2 \). In each subcube \( A_r \), one chooses at random two points \( x_r \) and \( z_r \), and finds \( x_r' \) and \( z_r' \) as above. The integral is then estimated by

\[
J_2'(f, N) = \frac{1}{N} \sum_{r=1}^{N} \frac{f(x_r) + f(x_r') + f(z_r) + f(z_r')}{4},
\]

At the same time the quantity

\[
D_z = D_z(f, N) = \frac{1}{2N} \left( \sum_{r=1}^{N} \left( \frac{f(x_r) + f(x_r') - f(z_r) - f(z_r')}{2} \right)^2 \right)^{1/2}
\]

is calculated; the expected value of \( D_z^2 \) is the variance of \( J_2' \), so that \( D_z \) can be taken as an estimate of the standard deviation of \( J_2' \). The calculation of \( J_2' \) involves \( 4N \) evaluations of the integrand \( f \), and

\[
\sigma_{4N}(J_2') = \frac{1}{\sqrt{2}} \sigma_{2N}(J_2)
\]

since \( J_2' \) is just the average of 2 independent samples of \( J_2 \). This implies that using \( J_2' \) instead of \( J_2 \) involves a slight loss of accuracy, since by Theorem 1,

\[
\sigma_{4N}(J_2) = \frac{1}{2^{1/2 + 2/\alpha}} \sigma_{2N}(J_2)
\]

if \( f \) is smooth. However, in my experience the importance of having a good estimate of the error is so great as to justify this loss of accuracy and the extra effort required to calculate \( D_z \).

Usually, in approximately evaluating an integral by the present methods, one would calculate \( J_2(f, N) \) or \( J_2'(f, N) \) for a few values of \( N \), in ascending sequence. When \( J_2' \) is used, one could, for each value of \( N \), also calculate \( D_z N^{1/2 + 2/\alpha} \). If the asymptotic form of \( \sigma_{2N}(J_2) \) given in the Theorem is, for the last two or three values of \( N \) used, a good approximation to the standard deviation, \( D_z N^{1/2 + 2/\alpha} \) will be close to constant and its value (for the last value of \( N \)) may be taken as the value of \( 2^{-1/2} d_2 \). Further calculations, with higher \( N \), may thereafter be made using \( J_2 \), and taking \( d_z N^{-1/2 + 2/\alpha} \) as the standard deviation.

(For \( J_1 \), automatic error estimation is achieved by using the estimator

\[
J_1'(f, N) = \frac{1}{N} \sum_{r=1}^{N} \frac{f(x_r) + f(z_r)}{2}
\]

to approximate \( I \), and using

\[
D_1(f, N) = \frac{1}{2N} \left( \sum_{r=1}^{N} (f(x_r) - f(z_r))^2 \right)^{1/2}
\]

as the estimate of the standard deviation of \( J_1' \). For convenience I shall also denote the above-mentioned estimate of the standard deviation of \( J_0 \), i.e., the quantity

\[
N^{-1/2} \left( \frac{1}{N} \sum_{r=1}^{N} f^2(y_r) - \left( \frac{1}{N} \sum_{r=1}^{N} f(y_r) \right)^2 \right)^{1/2}
\]

by “\( D_0' \)".)
Experiments. In order to test the accuracy of the error estimates above for smooth functions, and to see the behavior of $J_1$ and $J_2$ when the integrand is discontinuous, three 4-dimensional integrals were calculated. Tables I and Ia present the results for the calculation of

---

**Table I**

<table>
<thead>
<tr>
<th>$N$</th>
<th>$D_0$</th>
<th>$D_1$</th>
<th>$D_2$</th>
<th>$D_0N^{1/2}$</th>
<th>$D_1N^{1/2+1/4}$</th>
<th>$D_2N^{1/2+2/4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^4$</td>
<td>.019</td>
<td>.0085</td>
<td>.0030</td>
<td>.078</td>
<td>.066</td>
<td>.048</td>
</tr>
<tr>
<td>$3^4$</td>
<td>.011</td>
<td>.0037</td>
<td>.0085</td>
<td>.102</td>
<td>.102</td>
<td>.069</td>
</tr>
<tr>
<td>$4^4$</td>
<td>.0064</td>
<td>.0014</td>
<td>.0028</td>
<td>.102</td>
<td>.092</td>
<td>.073</td>
</tr>
<tr>
<td>$5^4$</td>
<td>.0044</td>
<td>.00085</td>
<td>.0010</td>
<td>.110</td>
<td>.105</td>
<td>.062</td>
</tr>
<tr>
<td>$8^4$</td>
<td>.0018</td>
<td>.00019</td>
<td>.00016</td>
<td>.115</td>
<td>.098</td>
<td>.068</td>
</tr>
<tr>
<td>$10^4$</td>
<td>.0011</td>
<td>.00010</td>
<td>.000068</td>
<td>.113</td>
<td>.099</td>
<td>.068</td>
</tr>
<tr>
<td>$16^4$</td>
<td>.00044</td>
<td>.000024</td>
<td>.000011</td>
<td>.112</td>
<td>.100</td>
<td>.069</td>
</tr>
</tbody>
</table>

**Table Ia**

<table>
<thead>
<tr>
<th>$N$</th>
<th>$E_0$</th>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$r_0$</th>
<th>$r_1$</th>
<th>$r_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^4$</td>
<td>-.00673</td>
<td>-.00505</td>
<td>.00159</td>
<td>.35</td>
<td>.59</td>
<td>.53</td>
</tr>
<tr>
<td>$3^4$</td>
<td>-.00401</td>
<td>-.00557</td>
<td>-.00039</td>
<td>.36</td>
<td>1.50</td>
<td>.05</td>
</tr>
<tr>
<td>$4^4$</td>
<td>.000988</td>
<td>.00080</td>
<td>-.00082</td>
<td>.16</td>
<td>.57</td>
<td>.29</td>
</tr>
<tr>
<td>$5^4$</td>
<td>.00256</td>
<td>.00097</td>
<td>-.00042</td>
<td>.58</td>
<td>1.14</td>
<td>.42</td>
</tr>
<tr>
<td>$8^4$</td>
<td>-.00086</td>
<td>.000065</td>
<td>-.000088</td>
<td>.48</td>
<td>.34</td>
<td>.55</td>
</tr>
<tr>
<td>$10^4$</td>
<td>-.00124</td>
<td>-.00090</td>
<td>-.0000104</td>
<td>1.10</td>
<td>.90</td>
<td>.15</td>
</tr>
<tr>
<td>$16^4$</td>
<td>-.00023</td>
<td>.000238</td>
<td>.0000089</td>
<td>.52</td>
<td>.99</td>
<td>.81</td>
</tr>
</tbody>
</table>

**Table II**

<table>
<thead>
<tr>
<th>$N$</th>
<th>$D_0$</th>
<th>$D_1$</th>
<th>$D_2$</th>
<th>$D_0N^{1/2}$</th>
<th>$D_1N^{1/2+1/4}$</th>
<th>$D_2N^{1/2+2/4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^4$</td>
<td>.17</td>
<td>.11</td>
<td>0</td>
<td>.68</td>
<td>.91</td>
<td>.00000</td>
</tr>
<tr>
<td>$3^4$</td>
<td>.083</td>
<td>.051</td>
<td>.035</td>
<td>.75</td>
<td>1.30</td>
<td>2.86</td>
</tr>
<tr>
<td>$4^4$</td>
<td>.046</td>
<td>.023</td>
<td>.012</td>
<td>.73</td>
<td>1.48</td>
<td>3.14</td>
</tr>
<tr>
<td>$5^4$</td>
<td>.029</td>
<td>.013</td>
<td>.0056</td>
<td>.72</td>
<td>1.58</td>
<td>3.49</td>
</tr>
<tr>
<td>$8^4$</td>
<td>.011</td>
<td>.0034</td>
<td>.00098</td>
<td>.71</td>
<td>1.74</td>
<td>4.00</td>
</tr>
<tr>
<td>$10^4$</td>
<td>.0070</td>
<td>.0018</td>
<td>.00041</td>
<td>.70</td>
<td>1.75</td>
<td>4.08</td>
</tr>
<tr>
<td>$16^4$</td>
<td>.0028</td>
<td>.00044</td>
<td>.000064</td>
<td>.71</td>
<td>1.80</td>
<td>4.17</td>
</tr>
</tbody>
</table>

**Table IIa**

<table>
<thead>
<tr>
<th>$N$</th>
<th>$E_0$</th>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$r_0$</th>
<th>$r_1$</th>
<th>$r_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^4$</td>
<td>-.09922</td>
<td>.1363</td>
<td>0</td>
<td>.58</td>
<td>1.24</td>
<td>.75</td>
</tr>
<tr>
<td>$3^4$</td>
<td>.07652</td>
<td>.0481</td>
<td>.02627</td>
<td>.92</td>
<td>.94</td>
<td>.68</td>
</tr>
<tr>
<td>$4^4$</td>
<td>-.05043</td>
<td>.0372</td>
<td>.00821</td>
<td>1.10</td>
<td>1.62</td>
<td>1.3</td>
</tr>
<tr>
<td>$5^4$</td>
<td>-.01536</td>
<td>.0261</td>
<td>.000733</td>
<td>.53</td>
<td>2.00</td>
<td>1.16</td>
</tr>
<tr>
<td>$8^4$</td>
<td>-.00198</td>
<td>.00243</td>
<td>-.000158</td>
<td>.18</td>
<td>.71</td>
<td>.16</td>
</tr>
<tr>
<td>$10^4$</td>
<td>-.01132</td>
<td>-.00092</td>
<td>.00020</td>
<td>1.61</td>
<td>.51</td>
<td>.05</td>
</tr>
<tr>
<td>$16^4$</td>
<td>-.0002285</td>
<td>-.00066</td>
<td>.000412</td>
<td>.08</td>
<td>1.50</td>
<td>.64</td>
</tr>
</tbody>
</table>

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
\[ \int_0^1 \int_0^1 \int_0^1 \int_0^1 (\exp [x^1 \cdot x^2 \cdot x^3 \cdot x^4] - 1) \, dx^1 \, dx^2 \, dx^3 \, dx^4 = .0693976 \]

and Tables II and IIa give the corresponding results for
\[ \int_0^1 \int_0^1 \int_0^1 \int_0^1 \sin 2\pi (x^1 + x^2 + x^3 + x^4) \, dx^1 \, dx^2 \, dx^3 \, dx^4 = 0 . \]

Here and below, \( E_0, E_1, \) and \( E_2 \) are the actual errors of \( J_0, J_1, \) and \( J_2 \) respectively; and in each case \( r_i = |E_i|/D_i, \) that is, the ratio of the actual error to the estimated standard deviation.

Table I shows that \( D_0N^{1/2}, D_1N^{1/2+1/2}, \) and \( D_2N^{1/2+2/2} \) were substantially constant for \( K \geq 3, \) indicating that the standard deviations were in close accordance with the predictions of the Theorem above and of Theorem 3 of MMC. The second integrand is more rapidly oscillating, with the result that in Table II \( D_1N^{1/2+1/2} \) and \( D_2N^{1/2+2/2} \) are approximately constant only for \( K \geq 7. \) In both cases the \( D_i \) accurately estimate the standard deviations for moderate values of \( N. \) The gain in accuracy due to \( J_2 \) is clear, though in comparing \( D_0 \) and \( D_1 \) to \( D_2, \) \( D_0 \) should be divided by 2 and \( D_1 \) by 1.4 to compensate for the greater number of function evaluations done in the calculation of \( J_2'. \)

**Table III**

<table>
<thead>
<tr>
<th>( N )</th>
<th>( D_0 )</th>
<th>( D_1 )</th>
<th>( D_2 )</th>
<th>( D_0N^{1/2} )</th>
<th>( D_1N^{1/2+1/2} )</th>
<th>( D_2N^{1/2+2/2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 2^4 )</td>
<td>.108</td>
<td>.076</td>
<td>.044</td>
<td>.43</td>
<td>.43</td>
<td>.25</td>
</tr>
<tr>
<td>( 3^4 )</td>
<td>.051</td>
<td>.032</td>
<td>.017</td>
<td>.46</td>
<td>.49</td>
<td>.26</td>
</tr>
<tr>
<td>( 4^4 )</td>
<td>.029</td>
<td>.015</td>
<td>.0080</td>
<td>.47</td>
<td>.48</td>
<td>.26</td>
</tr>
<tr>
<td>( 5^4 )</td>
<td>.019</td>
<td>.0079</td>
<td>.0046</td>
<td>.47</td>
<td>.44</td>
<td>.26</td>
</tr>
<tr>
<td>( 8^4 )</td>
<td>.0073</td>
<td>.0025</td>
<td>.0014</td>
<td>.46</td>
<td>.46</td>
<td>.24</td>
</tr>
<tr>
<td>( 10^4 )</td>
<td>.0046</td>
<td>.0014</td>
<td>.00077</td>
<td>.46</td>
<td>.44</td>
<td>.24</td>
</tr>
<tr>
<td>( 16^4 )</td>
<td>.0018</td>
<td>.00045</td>
<td>.00023</td>
<td>.46</td>
<td>.46</td>
<td>.24</td>
</tr>
</tbody>
</table>

**Table IIIa**

<table>
<thead>
<tr>
<th>( N )</th>
<th>( E_0 )</th>
<th>( E_1 )</th>
<th>( E_2 )</th>
<th>( r_0 )</th>
<th>( r_1 )</th>
<th>( r_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 2^3 )</td>
<td>.0584</td>
<td>.0272</td>
<td>-.0040</td>
<td>.54</td>
<td>.36</td>
<td>.09</td>
</tr>
<tr>
<td>( 3^3 )</td>
<td>-.0002</td>
<td>.00060</td>
<td>-.00287</td>
<td>0</td>
<td>.02</td>
<td>.17</td>
</tr>
<tr>
<td>( 4^3 )</td>
<td>-.0119</td>
<td>-.00408</td>
<td>-.00408</td>
<td>.41</td>
<td>.27</td>
<td>.51</td>
</tr>
<tr>
<td>( 5^3 )</td>
<td>-.0180</td>
<td>-.00118</td>
<td>-.00198</td>
<td>.95</td>
<td>.15</td>
<td>.43</td>
</tr>
<tr>
<td>( 8^3 )</td>
<td>-.0075</td>
<td>-.00310</td>
<td>-.00029</td>
<td>1.03</td>
<td>1.24</td>
<td>.21</td>
</tr>
<tr>
<td>( 10^4 )</td>
<td>-.00108</td>
<td>.000275</td>
<td>.00020</td>
<td>.23</td>
<td>.20</td>
<td>.26</td>
</tr>
<tr>
<td>( 16^4 )</td>
<td>-.00052</td>
<td>.000114</td>
<td>.000041</td>
<td>.29</td>
<td>.25</td>
<td>.18</td>
</tr>
</tbody>
</table>

Tables III and IIIa present the results of calculating the integral of the function given by

\[ f(x^1, x^2, x^3, x^4) = 1 \text{ if } (x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2 \leq 1 , \]

\[ = 0 \text{ otherwise} . \]
This function is discontinuous in $G_4$, and so in accordance with the discussion leading to (20), we should expect that $D_1$ and $D_2$ would go to zero as $N^{-1/2+1/2\epsilon}$. The approximate constancy of $D_1 N^{1/2+1/2\epsilon}$ and $D_2 N^{1/2+1/2\epsilon}$ in Table III confirms this very well for this simple integrand.

The values of $r$ in Tables Ia, IIa, and IIIa also conform to what is expected on the assumption that $J_0, J_1', J_2'$ are approximately normally distributed. The $r$'s are mostly < 1, and are only rarely as high as 2. In practical calculations, of course, only the $D_i$ are known and information about the actual errors is to be inferred from them; $2D_i$ is then a fairly safe ("5% confidence level") upper bound for $|E_i|$.

The following integral arose in a physical problem [3]:

\[
\frac{10^4}{\pi T^2} \int_0^{A(T)} \int_0^{\pi/2} k^2 \sin \theta \frac{2.07k^2 + .01 \sin^2 \theta}{we^{w/0.138T}} d\theta dk,
\]

where

\[
w = (4.285k^4 + .0414k^2 \sin^2 \theta)^{1/2};
\]

it was to be evaluated to an accuracy of about 1 part in 100 for various values of the parameter $T$. As the region of integration is a rectangle, it was first attempted to do the calculation using the trapezoid rule in each dimension. In one dimension the trapezoid rule approximation converges as $M^{-2}$ where $M$ is the number of points used—if the integrand is sufficiently smooth. In two dimensions the approximations should then converge as $M^{-1}$. In this case however, they were found to converge only as $M^{-1/2}$. This is apparently due to the fact that the second derivative, with respect to $k$, of the integrand is infinite at the origin; and so it could not be corrected by substituting any higher-order quadrature rule for the trapezoid rule. The calculation was then done by the simple and modified Monte-Carlo methods under discussion, and the results (for $T = 10^{-5}$, $A(T) = 10^{-4}$) are given in Table IV.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$J_0$</th>
<th>$J_1'$</th>
<th>$J_2'$</th>
<th>$D_0 N^{1/2}$</th>
<th>$D_1 N$</th>
<th>$D_2 N^{3/2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50²</td>
<td>.755</td>
<td>.592</td>
<td>.582</td>
<td>3.97</td>
<td>34.4</td>
<td>224</td>
</tr>
<tr>
<td>100²</td>
<td>.634</td>
<td>.590</td>
<td>.588</td>
<td>3.53</td>
<td>46.1</td>
<td>346</td>
</tr>
<tr>
<td>200²</td>
<td>.584</td>
<td>.590</td>
<td>.588</td>
<td>3.20</td>
<td>44.6</td>
<td>354</td>
</tr>
<tr>
<td>400²</td>
<td>.590</td>
<td>.588</td>
<td>.588</td>
<td>3.21</td>
<td>44.8</td>
<td>350</td>
</tr>
</tbody>
</table>

In this case the integrand satisfied the hypotheses of Theorem 3 of MMC, but not those of the theorem of the present paper. Thus the approximate constancy of $D_1 N^{1/2+1/4\epsilon}$ was expected, but that of $D_2 N^{1/2+2/4\epsilon}$ was surprising, and indicates a possibility of weakening the hypotheses of the theorem. Practically, the application of $J_2'$ was successful: $J_2'$ achieved the desired accuracy with $M = 10,000$, while with the trapezoid rule (improved by use of Richardson’s “deferred approach to the limit”), it was necessary to go up to $M = 160,000$ and at that the results did not generate much confidence, as Richardson’s extrapolation differed from the last trapezoidal value by about 7%.

I wish to thank Mr. Charles Mesztenyi of the University of Maryland’s Computer Science Center for doing these calculations.
6. Comment. $J'_2$ offers the advantage, over $J'_1$, of faster convergence to the integral. The improvement, for values of $N$ for which the asymptotic error for a given level of effort and expressions are fairly accurate, is given by the factor $(\sqrt{2d_2/d_1}) N^{-1/2}$ by which the standard deviation is multiplied. This may in some cases be no improvement at all; $d_2$ might be sufficiently greater than $d_1$ to make this factor greater than 1 for all reasonable values of $N$. Now in calculating, $J'_2$, $J'_1$ can be obtained simultaneously; and $D_1$ can also be gotten with very little extra effort. In doubtful situations, where it is not known how high an $N$ will have to be used, it is advisable to do this for the first values of $N$ tried, after which it can be seen whether $J'_1$ or $J'_2$ is the better estimator for the specific integral being studied.

Mathematics Division
National Bureau of Standards
Washington, D.C. 20234

3. S. J. Glass & S. Haber, "On the very low temperature behavior of a ferromagnetic." (To appear.)