DOE/ER/01 359--T2

FINAL PROGRESS REPORT
SUBMITTED TO

MASTER

U.S. DEPARTMENT OF ENERGY

DOE/ER/01359--T2

DE82 007194

by

CLAREMONT GRADUATE SCHOOL
CLAREMONT, CALIFORNIA 91711

ER 01359

June 11, 1980

Project Title: Systematic Efficiency Enhancement
in Monte Carlo Applications

Principal Investigator: Jerome Spanier
Professor of Mathematics

Duration: July 1, 1976 - January 31, 1980

Total Amount Received of DOE: \$ 45,852

Jerome Spanier, Principal Investigator

DISCLAIMER

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

leg

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

DOE/ER/01359--T2

This constitutes the final report of progress made under grant number EG-77-S-03-1359 (altered to # DE-AS03-77ER01359). The next section contains a brief summary of achievements made during this 3½ years of ERDA/DOE support, both by the principal investigator and three of his Ph.D. students, and the references section includes publications relating to grant support.

Summary of Achievements

Research performed under the grant period has been undertaken as part of the principal investigator's long-term efforts to develop new, more efficient estimators for application to a wide variety of practical problems. Two rather different approaches have characterized the work: (1) the use of a multistage analysis [1] to optimize the efficiencies (variances) of families of estimating random variables in a traditional statistical Monte Carlo framework, and (2) the development of parallel quasi-random sampling techniques and corresponding deterministic error bounds.

A. Statistical Monte Carlo

In area (1), the principal investigator succeeded in generalizing the weighted uniform sampling technique introduced by Powell and Swann [2] for estimating finite dimensional integrals. The generalization is in two directions: (a) from applicability only to finite dimensional integral estimation to random walk processes useful for estimating functionals associated with certain integral equations;

(b) from the original Powell-Swann formulation to a two-parameter family of estimators which includes the Powell-Swann estimator, as well as a variety of other useful estimators, as special cases.

The formulation of this generalization, which we have termed the generalized weighted analog (GWAN) estimator, is in terms of analog measure μ on the space Ω of all random walk histories, a parametrized family $\nu(c)$ of probability measures on Ω with $\nu(o) = \mu$, and any random variable ξ which provides an unbiased estimator of some prescribed parameter θ with respect to analog measure (see [3] for a more complete treatment). With these

conventions, the GWAN estimator is

$$(1) \quad \xi_G = \sum_{i=1}^N \xi(\omega_i) \frac{d\mu}{dv(c_1)}(\omega_i) / \sum_{i=1}^N \frac{dv(c_2)}{dv(c_1)}(\omega_i),$$

where the random walks $\omega_1, \dots, \omega_N$ are generated according to the measure $v(c)$ and the required Radon-Nikodym derivatives are assumed to exist. Thus, Eq(1) defines a two-parameter family of random variables which, while biased for general (c_1, c_2) , represents a true generalization of many well-known unbiased estimators as well as of the estimator of [2]. For example, the choice $c_1 = c_2 = 0$ causes ξ_G to reduce to a "crude" Monte Carlo estimator, $c_1 = c_2 = c \neq 0$ produces an unbiased importance sampling estimator, and $c_1 = 0, c_2$ arbitrary produces the estimator of [2] generalized to random walk problems. The parameter c_1 controls the measure used to generate random walk histories, while the parameter c_2 affects only the "weight" carried by such histories.

Analysis by the principal investigator of the family of estimators embodied in Eq(1) is presented in [4]. In order to state some of the theorems established in [4] we need to introduce some notation. Our concern is with the estimation of the integral

$$(2) \quad \theta = \int_{\Gamma} \Sigma(P) \varphi(P) dP$$

where Γ is a finite dimensional euclidean phase space, $\Sigma(P)$ is a known function, and $\varphi(P)$ satisfies an integral equation

$$(3) \quad \varphi(P) = \int_{\Gamma} K(P, P') \varphi(P') dP' + S(P).$$

Estimation of θ is accomplished via the probabilistic model of [3] in which it is shown how random walks in phase space may be constructed through various probability measures on the space Ω of all random walk histories. In particular, the analog random walk process, which is defined in terms of the source S and kernel K of the integral equation (3), gives rise to the analog measure μ on Ω .

With these preliminaries the ideas of Powell and Swann [2] may be applied to the estimation of

$$\theta = \int_{\Omega} \xi(\omega) d\mu(\omega)$$

where μ is the analog measure on Ω and ξ is any random variable which is an unbiased estimator of θ with respect to μ .

In general, we define the mean value of an estimator $\xi(\omega)$ by

$$T \equiv \int_{\Omega^N} \xi(W) dM(W)$$

where $W = (\omega_1, \dots, \omega_N) \in \Omega^N$ and $M(W)$ is an appropriate product measure on Ω^N .

For example, the "crude" Monte Carlo estimator of θ is defined by

$$(4) \quad \xi_A(\omega_1, \dots, \omega_N) \equiv \frac{1}{N} \sum_{i=1}^N \xi(\omega_i),$$

unbiased importance sampling results from

$$(5) \quad \xi_B(\omega_1, \dots, \omega_N) \equiv \frac{1}{N} \sum_{i=1}^N \eta_i(\omega_i) = \frac{1}{N} \sum_{i=1}^N \xi(\omega_i) \frac{d\mu}{dv}(\omega_i),$$

and a generalized Powell-Swann estimator is defined by

$$(6) \quad \xi_C(\omega_1, \dots, \omega_N) \equiv \sum_{i=1}^N \xi(\omega_i) / \sum_{i=1}^N \frac{dv}{d\mu}(\omega_i).$$

In (4) and (6), histories ω_i are generated according to analog measure μ , while in (5) histories are generated according to a probability measure v which is to be chosen to emphasize events of "importance" in the estimation of θ .

As usual, the variance of ξ is

$$V \equiv \int_{\Omega^N} [\xi(w) - T]^2 dM(w)$$

and its bias is

$$B \equiv \theta - T.$$

Defining the mean square error of ξ by

$$U \equiv \int_{\Omega^N} [\xi(w) - \theta]^2 dM(w),$$

It follows that

$$V = U - B^2.$$

From this we see that the estimator ξ_c will provide a satisfactory error only if the square of its bias is small relative to its variance.

A measure of the "error" in the definition of ξ_c is given by

$$(7) \quad H(\omega) \equiv \xi(\omega) - \theta \frac{dv}{d\mu}(\omega)$$

since the choice $H(\omega) = 0$ clearly gives rise to a ξ_c with zero error. The random variable H provides the key to the asymptotic behavior of the mean square error, U_c , of ξ_c :

Theorem 1. If $H(\omega)$ is bounded and $\frac{dv}{d\mu}(\omega)$ is bounded away from zero, then

$$U_c \sim \frac{1}{N} \int_{\Omega} H^2(\omega) d\mu(\omega) \quad (N \rightarrow \infty).$$

The proof may be found in [4]. This knowledge of the large sample behavior of U_c helps establish

Theorem 2.

$$B_c^2/U_c = O(N^{-1}).$$

Thus one is assured that the bias associated with the estimator ξ_c is satisfactorily small for sufficiently large sample sizes.

Results completely analogous to Theorems 1 and 2 are established in [4] for the GWAN estimator defined by Eq(1).

A careful analysis of the behavior of this GWAN family is presented in [4]. Roughly speaking, two complementary sets of model problems are used to explore the efficiency of these estimators as a function of the parameters c_1 and c_2 . On the one hand, deep penetration problems are investigated to ascertain how GWAN estimators behave in extreme situations for which importance sampling of some sort is always required. On the other hand, model "cell calculations", for which knowledge of the particle population everywhere in phase space is required, are also examined.

Through examination of such a variety of model problems, a new class of biased random variables with uniformly low mean square errors is identified in [4]. This new one-parameter family is defined by specializing to $c_2 = 0$ in Eq(1), which produces the estimator

$$(8) \quad \hat{\theta} = \frac{\sum_{i=1}^N \xi(\omega_i) \frac{d\mu}{d\nu}(\omega_i)}{\sum_{i=1}^N \frac{d\mu}{d\nu}(\omega_i)}$$

in which random walks are generated according to the arbitrary measure ν . We have interpreted (8) as a biased importance sampling estimator in which the measure ν is chosen, as usual, to emphasize events of interest in the estimation of θ , but the resulting weighted estimator is adjusted through division by the sum of these weights, rather than by the number of random walk histories. This adjustment appears to compensate for histories of either abnormally low or high weight in a remarkably effective way. A combination of exact model problem analysis and the multistage analysis of [1], adapted to biased random variables, is used in [4] to arrive at accurate information about the para-

metric dependence of mean square errors of estimators.

Further work on GWAN estimation using pseudo-random methods has been performed by one of the principal investigator's Ph.D. students, Jean Wang Ju. As part of her doctoral dissertation [5], Ms. Ju has studied the applicability of GWAN estimators to practical problems such as those arising in nuclear reactor theory. To this end she has studied both "terminal" estimators (i.e., those random variables ξ which record non-zero contributions to θ only upon the final collision of each history) and "path length-type" estimators (which record non-zero contributions continuously along the history's flight path), as well as optimally selected linear combinations of these. The rationale for taking such linear combinations is that path length estimators are observed to have much lower statistical uncertainty than terminal estimators when used in optically thin regions, whereas the advantage may swing to terminal estimators in very large, highly absorbing regions. Therefore, a practical design might be to combine estimators of both types with a combination coefficient which depends on the optical thickness of a region.

While the analysis is fairly straightforward for unbiased estimators, the combining of possibly biased estimators seems to complicate the task. However, a suitable extension of the theory in the unbiased case does, in fact, provide the correct result, namely:

Theorem 3. If ξ_1 and ξ_2 are any two estimators of θ with mean square errors U_1 , U_2 , respectively, then the mean square error of linear combinations

$$\xi = a\xi_1 + (1-a)\xi_2$$

is minimized over choices of a by

$$a_{\min} = \frac{U_1 - U_{12}}{U_1 + U_2 - 2U_{12}},$$

where

$$U_{12} = E[(\xi_1 - \theta)(\xi_2 - \theta)]$$

is a "co-error" associated with the pair ξ_1, ξ_2 (U_{12} reduces to the covariance if ξ_1, ξ_2 are unbiased.) Furthermore, a necessary and sufficient condition for

$$0 < a_{\min} < 1 \text{ is } U_1 - U_{12} > 0 \text{ and } U_2 - U_{12} > 0.$$

Once again, multistage sampling provides a practical mechanism for verifying these nearly optimal combination coefficients.

Some preliminary results have also been obtained by Ms. Ju in her dissertation which are aimed at building an importance sampling theory for GWAN estimators. Since GWAN estimation is strictly more general than unbiased importance sampling, the intriguing possibility exists that an importance sampling theory which is also more general than the classical one ([6], [7]) will provide new insights into low error GWAN choices. The theory appears to be much more difficult to expose for the (possibly) biased GWAN estimators, and only asymptotic results (for large sample sizes) are available so far. Preliminary results seem to indicate optimal choices for importance functions which are identical to those used for unbiased importance sampling. This result is disappointing but may not be too surprising. It is still quite possible that substantial improvements over unbiased importance sampling are available through GWAN estimation for finite sample sizes, but that the advantage becomes negligible as the sample size approaches infinity. Work aimed at understanding this situation is underway at the present time.

B. Quasi-Random Monte Carlo

In the second area of research, which deals with the development of quasi-random sampling techniques and corresponding deterministic error bounds, considerable progress has also been made. The first such work, undertaken by the principal investigator and Paul Chelson, was concerned with establishing error bounds for quasi-random estimators based on unbiased importance sampling. The first known results of this sort are reported in [8] and apply to the estimation either of multidimensional integrals or to (infinite dimensional) random walk processes. When used in conjunction with low discrepancy (rather than pseudo-random) generating sequences, these bounds establish asymptotically more efficient estimators than can be expected with pseudo-random generators.

The above theory of importance sampling for quasi-random methods is complex and the proofs are long, but the following results illustrate how such a theory may be used advantageously.

Theorem 4. (Importance sampling for quasi-random integration).

Let f be a real-valued function on $[a,b]$ and let g be a probability density function on $[a,b]$ with associated cumulative distribution function G . Let $Q = \{r_i\}$ be any sequence in $[0,1)$ and define a new sequence

$$Q^1 = \{s_i\} = \{G^{-1}(r_i)\}.$$

Then, assuming that f/g has bounded variation,

$$\left| \delta_N(f/g) \right| = \left| \frac{1}{N} \sum_{i=1}^N \frac{f(s_i)}{g(s_i)} - \int_a^b f(x) dx \right| \leq D_N(Q) V(f/g)$$

where D_N is the discrepancy (see [9], for example) and V is the total variation.

Instead of statistical estimates of error of order $K/\sqrt{N}$ where K is the (1-sample) standard deviation, Theorem 4 produces deterministic error bounds of order $\tilde{K}D_N(Q)$, where $\tilde{K} = V(f/g)$. Selection of low discrepancy sequences Q will guarantee that $\delta_N(f/g) = O(\log N/N)$ as $N \rightarrow \infty$, giving such estimates a very large advantage in convergence rate for sufficiently large N . Furthermore, choice of an appropriate density function g will also reduce the constant multiplier, $\tilde{K}$, without degrading the asymptotic convergence rate.

The extension of Theorem 4 to finite dimensional integrals as well as to the infinite dimensional case of random walk processes has also been established. The results are too technical to reproduce here but may be found in [8]. Making use of such quasi-random importance sampling methods, very substantial improvements in convergence have been observed in model transport problems involving deep penetration.

More work in the direction of finding deterministic upper bounds for more general families of quasi-random estimators has been the subject of the doctoral dissertation of a current student, Earl Maize. Mr. Maize has succeeded in obtaining error bounds for GWAN estimators which extend the results of Chelson, at least for the estimation of finite-dimensional integrals. We state one such result as follows:

Theorem 5. (Quasi-random weighted uniform sampling) Let $\omega = \{\underline{x}^{(n)}\}_{n=1}^{\infty}$

be a uniformly distributed sequence in the s -dimensional hypercube Q^s and let $\omega^+ = \{\underline{y}^{(n)}\}_{n=1}^{\infty}$ be the sequence obtained from ω by setting $y_1^{(n)} = G_1^{-1}(x_1^{(n)})$,

$y_2^{(n)} = G_2^{-1}(x_2^{(n)}), \dots, y_s^{(n)} = G_s^{-1}(x_s^{(n)})$ where the $G_j, 1 \leq j \leq s$ are the conditional distribution functions of the density of g . Let h satisfy

$$\int_{Q^s} h(\underline{x}) d\underline{x} = 1.$$

If f/g and h/g are of bounded variation in the sense of Hardy and Krause on Q^s , then, for sufficiently large N ,

$$\left| \frac{\sum_{n=1}^N f(\underline{y}^{(n)}) / g(\underline{y}^{(n)})}{\sum_{n=1}^N h(\underline{y}^{(n)}) / g(\underline{y}^{(n)})} - \int_{Q^s} f(\underline{x}) d\underline{x} \right|$$

$$\leq \sum_{p=1}^s \sum_{1, \dots, s; p}^* D_N(\omega_{p+1, \dots, s}) V^{(p)} \left(\frac{f - \theta h}{g} (x_1, \dots, x_p, 1, \dots, 1) \right) + O(D_N(\omega)),$$

where the symbol $\sum^*$ sums over all possible partitions of the variables into p and $s-p$ elements, D_N is the extreme discrepancy associated with projection of the sequence ω onto the first p dimensions of Q^s , $V^{(p)}$ is the p -dimensional variation in the sense of Vitali, and

$$\theta = \int_{Q^s} f(\underline{x}) d\underline{x}$$

is the unknown integral being estimated. The proof of Theorem 5 provides deterministic error bounds for the Powell-Swann type estimator. We conjecture that a result like Theorem 5 must be true for infinite-dimensional random walk processes as well. Examples which reveal the potential superiority of the estimator appearing in Theorem 5 are presented in [10].

Finally, Maize has studied in [10] the possibility of extending known results which describe the integration error when used in conjunction with smooth, periodic integrands and very specially chosen evaluation points. For example, for smooth periodic integrands error bounds of order $O(N^{-1})$, N = number of integrand evaluations, may be obtained but such estimates depend upon sampling the fractional parts of vectors of irrational numbers. Maize is attempting to obtain similar error estimates for certain dyadic sequences which are easily generated on a computer. Maize is also investigating the possibility of extending such techniques to the study of GWAN estimators.

REFERENCES

1. J. Spanier, "A New Multistage Procedure for Systematic Variance Reduction in Monte Carlo", SIAM J. Numer. Anal., 8, 548, (1971).
2. M.J.D. Powell and J. Swann, "Weighted Uniform Sampling -- A Monte Carlo Technique for Reducing Variance", J. Inst. Math Appl., 2, 228, (1966).
3. J. Spanier, "Some Results on Transport Theory and Their Applications to Monte Carlo Methods", J. Math. Anal. Appl., 17, 549, (1967).
- 4.* , "A New Family of Estimators for Random Walk Problems", J. Inst. Maths Applics., 23, 1, (1979).
- 5.* Jean Wang Ju, untitled Ph.D. Dissertation, Claremont Graduate School, in preparation.
6. J. Spanier and E.M. Gelbard, Monte Carlo Principles and Neutron Transport Problems, Addison-Wesley Pub. Co., Inc., 1969.
7. R.R. Coveyou, V.R. Cain, and K.J. Yost, "Adjoint and Importance in Monte Carlo Application", Nu. Sci. Eng., 27, 219, (1967).
- 8.* Paul O. Chelson, "Quasi-Random Techniques for Monte Carlo Methods", Ph.D. Dissertation, Claremont Graduate School, 1976.
9. L. Kuipers and H. Niederreiter, Uniform Distribution of Sequences, John Wiley and Sons, New York, 1974.
- 10.* Earl H. Maize, "Contributions to the Theory of Error Reduction in Quasi-Monte Carlo Methods", Ph.D. Dissertation, Claremont Graduate School, in preparation.

References marked with an asterisk were prepared with at least partial support under the subject grants.