Proceedings of the 1991 Winter Simulation Conference
Barry L. Nelson, W. David Kelton, Gordon M. Clark (eds.)

COMBINING RANDOM NUMBER GENERATORS

Lih-Yuan Deng

Department of Mathematical Sciences
Memphis State University
Memphis, TN 38152

Yu-Chao Chu

Department of Preventive Medicine
University of Tennessee – Memphis
Memphis, TN 38163

ABSTRACT

We provide some support for the combination generator from a statistical theory viewpoint. The combination generator is generated from taking the fractional part of the sum of several random number generators. It is shown that combining several independent generators improves not only the uniformity but also the independence over each component generators.

1 INTRODUCTION

Combining two or more pseudorandom sequences into a "more uniform" random sequence was recommended by many authors over the classical Lehmer (1951) congruential generator or the Tausworthe (1965) shift-register generator. Wichmann and Hill (1982) suggested adding three simple multiplicative congruential generators and taking the fractional part. They claimed, with a simple example but no general proof, that this procedure "ironed out" the imperfections in the component variates. Marsaglia (1985) empirically compared several popular generators and concluded that the combination generator is superior to others. Other authors such as Collings (1987), L'Ecuyer (1988) and Anderson (1990) also recommended the combination generator.

Some of the theoretical justification is given in Horton (1948), Horton and Smith (1949), Brown and Solomon (1979), Lécuyer (1988), Marsaglia (1985) and Deng and George (1990). Brown and Solomon (1979) and Marsaglia (1985) proved that the combination generator will yield a distribution which is closer, or at least no worse than, either of the individual generators. Deng and George (1990) provided some additional theoretical justification by showing the combination generator should improve upon the uniformity. In section 2, we prove a k-dimensional extension of Deng and George (1990)'s result. We show that the combination generator should improve upon the uniformity as well as the independence over individual generators. In section 3, a more detailed comparison between our results and those by Brown and Solomon (1979) and Marsaglia (1985) is given. We also give some comments on the relationship between the combination generators and some of the recently proposed generators.

2 MAIN RESULTS

Let $X_{11}, X_{21}, X_{31}...$ and $X_{12}, X_{22}, X_{32}...$ be any two sequence of random variables representing two separate RNG's. Our main objective is to study the combination sequence $Y_1, Y_2, ...$ where $Y_i = X_{i1} + X_{i2} \bmod 1$. In particular, for any positive integer k, we want to investigate the joint probability distribution of the k-dimensional random vector $\mathbf{Y} = \mathbf{X_1} + \mathbf{X_2} \bmod 1$, where $\mathbf{X_1} = (X_{i_1 1}, X_{i_2 1}, ..., X_{i_k 1})'$ $\mathbf{X_2} = (X_{i_1 2}, X_{i_2 2}, ..., X_{i_k 2})'$ with $i_1 < i_2 < ... < i_k$.

Several notations will be introduced here. Let $[0,1]^k = \{(x_1, x_2, ..., x_k)' | 0 \le x_i \le 1, 1 \le i \le k\}$ and $(0,1)^k = \{(x_1, x_2, ..., x_k)' | 0 < x_i < 1, 1 \le i \le k\}$ denote the closed, open sets of k- dimensional cubes, respectively. For any real number x, $x \bmod 1 = x - [x]$, where $[x]$ is the greatest integer $\le x$. For a vector $\mathbf{x} = (x_1, x_2, ..., x_k)'$, let $\mathbf{x} \bmod 1 = (x_1 \bmod 1, x_2 \bmod 1, ..., x_k \bmod 1)'$. Let δ_j be the k-vector vertex in $[0,1]^k$ corresponding to the binary representation of j for $j = 0, 1, 2, ..., 2^k - 1$. Let $\mathbf{\Delta_k} = \{\delta_j | j = 0, 1, 2, ..., 2^k - 1\}$ be the set of vertices in $[0,1]^k$. For each $\mathbf{y} \in (0,1)^k$, there are exactly 2^k partitions of $[0,1]^k$, $\{A_{j,\mathbf{y}}, j = 0, 1, 2, ..., 2^k - 1\}$, where $A_{j,\mathbf{y}}$ corresponds to the sub-cube of the partition containing the j-th vertex δ_j. For example, $k = 2$, $\mathbf{y} = (y_1, y_2)'$, $A_{0,\mathbf{y}} = [0, y_1] \times [0, y_2]$, $A_{1,\mathbf{y}} = [0, y_1] \times [y_2, 1]$, $A_{2,\mathbf{y}} = [y_1, 1] \times [0, y_2]$, $A_{3,\mathbf{y}} = [y_1, 1] \times [y_2, 1]$.

The following lemma lists some simple properties about $A_{j,\mathbf{y}}$.

Lemma 1. *Let $A_{j,\mathbf{y}}, j = 0, 1, 2, ...2^k - 1$ be defined as above. Then*

1. *$\mathbf{x} \in A_{j,\mathbf{y}}$ if and only if $\delta_j + \mathbf{y} - \mathbf{x} \in A_{j,\mathbf{y}}$.*
2. *$\bigcup_{j=0}^{2^k-1} A_{j,\mathbf{y}} = [0,1]^k$.*

3. $\int_{\mathbf{x}\in(A_{i,\mathbf{y}}\bigcap A_{j,\mathbf{y}})} d\mathbf{x} = 0$, *for* $i \neq j$.

Proof. (1) follows easily from the fact that $A_{j,\mathbf{y}}$ is symmetric around $(\delta_j + \mathbf{y})/2$. Proofs of (2) and (3) are trivial. □

To show our main result, we need the following lemma to find the p.d.f. of the fractional part of the sum of two independent random vectors.

Lemma 2. *Let* $\mathbf{X}_1, \mathbf{X}_2$ *be any two independent and continuous random vectors over* $[0,1]^k$, *with the p.d.f.* $f_{\mathbf{X}_1}(\mathbf{x})$ *and* $f_{\mathbf{X}_2}(\mathbf{x})$. *Let* $\mathbf{Y} = \mathbf{X}_1 + \mathbf{X}_2$ *mod* 1, *and* $f_{\mathbf{Y}}(\mathbf{y})$ *be the p.d.f. of* $\mathbf{Y}$. *Then for each* $\mathbf{y} \in (0,1)^k$,

$$f_{\mathbf{Y}}(\mathbf{y}) = \sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} f_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x}) f_{\mathbf{X}_1}(\mathbf{x}) d\mathbf{x}$$

Proof. The p.d.f. of $\mathbf{Z} = \mathbf{X}_1 + \mathbf{X}_2$ can be expressed as

$$h(\mathbf{z}) = \int_{\mathbf{x}\in[0,1]^k} f_{\mathbf{X}_2}(\mathbf{z} - \mathbf{x}) f_{\mathbf{X}_1}(\mathbf{x}) d\mathbf{x}, \quad \mathbf{z} \in [0,2]^k.$$

Lemma 2 follows from this and that the p.d.f. of $\mathbf{Y} = \mathbf{Z}$ mod 1 is

$$f_{\mathbf{Y}}(\mathbf{y}) = \sum_{j=0}^{2^k-1} h(\delta_j + \mathbf{y}), \quad \mathbf{y} \in (0,1)^k. \square$$

Let $\mathbf{X}$ be a random vector over $[0,1]^k$, with the p.d.f. $f_{\mathbf{X}}(\mathbf{x})$ representing any k-dimensional realization of a RNG. This assumption is not realistic because any RNG can only generate finite points in $[0,1]$. Any theory based on the assumption of existing RNG with a p.d.f. is only approximate. However, it can greatly reduce the need for the exact computations with discrete values and it should shed some light on the justification of the combination generator.

The following theorem shows that the fractional part of the sum of two independent "nearly" uniform random vectors will produce a distribution whose p.d.f. is closer to uniform distribution.

Theorem 1. *Let* $\mathbf{X}_1, \mathbf{X}_2$ *be two independent random vectors over* $[0,1]^k$, *with the p.d.f.* $f_{\mathbf{X}_1}(\mathbf{x})$ *and* $f_{\mathbf{X}_2}(\mathbf{x})$. *Let* $\mathbf{Y} = \mathbf{X}_1 + \mathbf{X}_2$ *mod* 1, *and* $f_{\mathbf{Y}}(\mathbf{y})$ *be the p.d.f. of* $\mathbf{Y}$. *If* $|f_{\mathbf{X}_1}(\mathbf{x}) - 1| \leq \epsilon_1$ *and* $|f_{\mathbf{X}_2}(\mathbf{x}) - 1| \leq \epsilon_2$, *then* $|f_{\mathbf{Y}}(\mathbf{y}) - 1| \leq \epsilon_1 \cdot \epsilon_2$.

Proof. Let

$$f_{\mathbf{X}_i}(\mathbf{x}) = 1 + g_{\mathbf{X}_i}(\mathbf{x}), \quad \text{for } i = 1,2, \qquad (A)$$

where $g_{\mathbf{X}_i}(\mathbf{x})$ is the "deviation" of the p.d.f. of $\mathbf{X}_i$ from the uniform p.d.f. By the assumption of Theorem 1, we have

$$|g_{\mathbf{X}_i}(\mathbf{x})| \leq \epsilon_i, \quad \text{for } i = 1,2.$$

Lemma 2 states that the p.d.f. of $\mathbf{Y} = \mathbf{X}_1 + \mathbf{X}_2$ mod 1 is

$$f_{\mathbf{Y}}(\mathbf{y}) = \sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} f_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x}) f_{\mathbf{X}_1}(\mathbf{x}) d\mathbf{x}.$$

Substituting (A) into the formula above, for each term $j = 0, 1, 2, \ldots, 2^k - 1$,

$$\begin{aligned}
&\int_{\mathbf{x}\in A_{j,\mathbf{y}}} f_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x}) f_{\mathbf{X}_1}(\mathbf{x}) d\mathbf{x} \\
&= \int_{\mathbf{x}\in A_{j,\mathbf{y}}} (1 + g_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x}))(1 + g_{\mathbf{X}_1}(\mathbf{x})) d\mathbf{x} \\
&= \int_{\mathbf{x}\in A_{j,\mathbf{y}}} 1 d\mathbf{x} + \int_{\mathbf{x}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_1}(\mathbf{x}) d\mathbf{x} \\
&\quad + \int_{\mathbf{x}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x}) d\mathbf{x} \\
&\quad + \int_{\mathbf{x}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x}) \cdot g_{\mathbf{X}_1}(\mathbf{x}) d\mathbf{x}.
\end{aligned}$$

Letting $\mathbf{u} = \delta_j + \mathbf{y} - \mathbf{x}$ in the third term and using Lemma 1, we have

$$\int_{\mathbf{x}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x}) d\mathbf{x} = \int_{\mathbf{u}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_2}(\mathbf{u}) d\mathbf{u}.$$

Combining the all parts of $f_{\mathbf{Y}}(\mathbf{y})$ yields

$$\begin{aligned}
&f_{\mathbf{Y}}(\mathbf{y}) \\
&= \sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} 1 d\mathbf{x} + \sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_1}(\mathbf{x}) d\mathbf{x} \\
&\quad + \sum_{j=0}^{2^k-1} \int_{\mathbf{u}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_2}(\mathbf{u}) d\mathbf{u} \\
&\quad + \sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x}) \cdot g_{\mathbf{X}_1}(\mathbf{x}) d\mathbf{x}.
\end{aligned}$$

Since

$$\sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} 1 d\mathbf{x} = \int_{\mathbf{x}\in[0,1]^k} 1 d\mathbf{x} = 1,$$

and for $i = 1, 2$,

$$\begin{aligned}
&\sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_i}(\mathbf{x}) d\mathbf{x} \\
&= \int_{\mathbf{x}\in[0,1]^k} f_{\mathbf{X}_i}(\mathbf{x}) d\mathbf{x} - 1 = 0,
\end{aligned}$$

we have

$$f_{\mathbf{Y}}(\mathbf{y}) = 1 + \sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} g_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x}) \cdot g_{\mathbf{X}_1}(\mathbf{x}) d\mathbf{x}.$$

Theorem 1 now follows easily because

$$\begin{aligned} & |f_{\mathbf{Y}}(\mathbf{y}) - 1| \\ & \leq \sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} |g_{\mathbf{X}_2}(\delta_j + \mathbf{y} - \mathbf{x})| \cdot |g_{\mathbf{X}_1}(\mathbf{x})| d\mathbf{x} \\ & \leq \sum_{j=0}^{2^k-1} \int_{\mathbf{x}\in A_{j,\mathbf{y}}} \epsilon_1 \cdot \epsilon_2 d\mathbf{x} \\ & = \epsilon_1 \epsilon_2. \square \end{aligned}$$

Theorem 1 is a k-dimensional extension of a result proved in Deng and George (1990). Essentially, they proved the combination generator will improve the "uniformity" of the generator and we show that it will also improve the "independence" when considering the joint distribution of any k-dimensional random vectors.

Corollary 1. *Let* $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ *be* n *independent random vectors over* $[0,1]^k$, *with the p.d.f.* $f_{\mathbf{X}_i}(\mathbf{x})$, *for* $i = 1, 2, ...n$. *Let* $\mathbf{Y} = \sum_{i=1}^n \mathbf{X}_i \ mod\ 1$, *and* $f_{\mathbf{Y}}(\mathbf{y})$ *be the p.d.f. of* $\mathbf{Y}$.

1. *If* $|f_{\mathbf{X}_i}(\mathbf{x}) - 1| \leq \epsilon_i, i = 1, 2, ...n$, *then* $|f_{\mathbf{Y}}(\mathbf{y}) - 1| \leq \prod_{i=1}^n \epsilon_i$.
2. *If* $\prod_{i=1}^n \epsilon_i \longrightarrow 0$, *then* $\mathbf{Y}$ *converges to uniform distribution over* $[0,1]^k$.
3. *In particular, if one of the* $\mathbf{X}_i$ *is uniformly distributed over* $[0,1]^k$, *then* $\mathbf{Y}$ *is uniformly distributed over* $[0,1]^k$.

Proof. Corollary 1 follows easily from Theorem 1 and mathematical induction. □

According to Theorem 1 and its Corollary, one can produce a random vector whose distribution is closer to $U[0,1]^k$ by taking the fractional part of the sum of several random vectors.

3 CONCLUDING REMARKS

Brown and Solomon (1979) showed that the combination generator is at least as uniformly distributed as the individual generators, using the techniques on majorization. Their result holds for an arbitrary symmetric norm (a symmetric norm is invariant under permutation) over k-space. Marsaglia (1985) proved, with a very special case, that the combination of two RNG's produces not only "more uniformly" but also "more independent" sequence. On page 5 of his paper, Marsaglia (1985) introduced the operation $x \cdot y$ which can be easily seen that

$$x \cdot y = ((x + y) \bmod m) + 1$$

with $m = 3$. The distance measure used is the usual Euclidean L_2 norm between two probability distributions. Under the L_2 norm, he showed that the distribution corresponding to the combination generator is at least as uniform and also as independent as the individual generators. Clearly, the result obtained by Brown and Solomon (1979) is stronger than that of Marsaglia (1985). However, neither results demonstrates that the combination generator actually *improves* over its component RNG's. Furthermore, no information about the "degree of improvement" is provided. We show that combining two nearly uniformly distributed RNG's will indeed yield a RNG with a smaller maximum deviation. An upper bound of the maximum deviation of the probability distribution of the combination generator is given in Theorem 1.

We have provided some theoretical justification for the practice of combining RNG's by taking the fractional parts of sums of RNG's. It is shown to improve not only the uniformity but also the independence of the generator. The reader should be reminded that all of the above mentioned justifications *assumed* the independence between the two component generators. In reality, however, any sequence generated by a RNG is a deterministic one. No independence in either *between RNG's* or *within RNG* can be safely assumed. Deng, George and Chu (1989, 1991) considered combining generators that are not necessarily independent and/or uniformly distributed. Their empirical study suggested that the combination generator will indeed improve the uniformity and independence of RNG's, even if the component RNG's are highly correlated.

Recently, several authors have studied two extensions of Lehmer's classical random number generator, namely (a) the multiple recursive generator (MRG) and (b) the matrix congruential generator (MCG). An excellent review on these and other generators is given in L'Ecuyer (1989, 1990). Both MRG and MCG have a long maximum period $p^k - 1$, where k is the dimension of the matrix for MCG and the degree of primitive polynomial for MRG and p is a prime number. Deng and Rousseau (1991) proposed efficient algorithms for finding MCGs and MRGs with the maximum period. The MRG can be viewed as a combination generator and its statistical justification is given here. Brown and Solomon (1979), Marsaglia (1985), L'Ecuyer (1988) and Deng, George and Chu (1989, 1991) also gave some justification from a sta-

tistical theory viewpoint. Deng (1991) has developed a general theory concerning the asymptotic uniformity and asymptotic independence for both the MRG and the MCG.

REFERENCES

Anderson, S. L. 1990. Random number generators on vector supercomputers and other advanced architectures, *SIAM Review*, **32**, 221-251.

Brown, M. and H. Solomon. 1979. On Combining Pseudorandom Number Generators, *Annals of Statistics*, **3**, 691-695.

Collings, B. J. 1987. Compound random number generators, *Journal of the American Statistical Association*, **82**, 525-527.

Deng, L. Y. 1991. Multivariate congruential generators,Submitted for publication.

Deng, L. Y. and E. O. George. 1990. Generation of Uniform Variates from Several Nearly Uniformly Distributed Variables, *Communications In Statistics*, **B19, No 1**, 145-154.

Deng, L. Y., E. O. George and Y. C. Chu. 1989. On Improving "Bad" Pseudo-Random Number Generators, In *Proceedings of the 21st Symposium on the Interface*, April 9-12, 284-286.

Deng, L. Y., E. O. George and Y. C. Chu. 1991. On Improving Pseudo-Random Number Generators, In *Proceedings of the 1991 Winter Simulation Conference.*

Deng, L. Y. and C. Rousseau. 1991. Recent development in random number generation, In *Proceedings of the 29th Annual ACM Southeast Regional Conference, Auburn Alabama, April 10-12*, 89-94.

Horton, H. B. 1948. A Method for Obtaining Random Numbers, *Annals of Mathematical Statistics*, **19**, 81-85.

Horton, H. B., and R. T. Smith III. 1949. A Direct Method for Producing Random Digits in Any Number System, *Annals of Mathematical Statistics*, **20**, 82-90.

L'Ecuyer, P. 1988. Efficient and Portable Combined Random Number Generators, *Communications of the ACM*, **31**, 742-748,774.

L'Ecuyer, P. 1989. A tutorial on uniform variate generation, In *Proceedings of the 1989 Winter Simulation Conference*, 40-49.

L'Ecuyer, P. 1990. Random Numbers For Simulation, *Communications of the ACM*, **33**, 85-97.

Lehmer, D. H. 1951. Mathematical Methods in Large-scale Computing Units, In *Proceedings of the Second Symposium on Large Scale Digital Computing Machinery: Harvard University Press*, Cambridge, 141-146.

Marsaglia, G. 1985. A Current View of Random Number Generators, In *Proceedings of the 16th Symposium on the Interface*, editor L. Billard, 3-10, North-Holland: Elsevier Science Publishers.

Tausworthe, R. C. 1965. Random numbers generated by linear recurrence modulo two, *Mathematics of Computation*, **19**, 201-209.

Wichmann, B. A. and I. D. Hill. 1982. An Efficient and Portable Pseudo-Random Number Generator, *Applied Statistics*, **31**, 188-190.

AUTHOR BIOGRAPHIES

LIH-YUAN DENG is a visiting associate professor in the Department of Mathematical Sciences at the University of Houston–Clear Lake. He is on leave from the Department of Mathematical Sciences at Memphis State University. He received the B.S. and M.S. degree in Mathematics in 1975 and 1977, both from the National Taiwan University. He then received the M.S. degree in Computer Science in 1982 and Ph. D. degree in Statistics in 1984, both from the University of Wisconsin–Madison. His research interests are in random number generation, survey sampling design and analysis, variance estimation, and statistical computing. He is a member of ACM and ASA.

YU-CHAO CHU is a research associate in the Department of Preventive Medicine at the University of Tennessee - Memphis. She received the M.S. and Ph. D. degree in Mathematics (with concentration in Statistics) in 1985 and 1990, both from Memphis State University. Her research interests include random number generation, statistical computing, and biostatistics.