Is Random Random Enough?

T. Ala-Nissila[a,c], K. Kankaala[a,b], and I. Vattulainen[a,c]

[a]Department of Electrical Engineering
Tampere University of Technology
P.O. Box 692, FIN - 33101 Tampere
Finland

[b]Center for Scientific Computing
P.O. Box 405, FIN - 02101 Espoo
Finland

[c]Research Institute for Theoretical Physics
P.O. Box 9 (Siltavuorenpenger 20 C)
FIN - 00014 University of Helsinki
Finland

September 8, 1994

"Random numbers are too important
to be left to chance."
- R. Coveyou

When something is said to be random, most of us have a rather good intuitive idea what is being meant: random is something that we cannot predict, something opposite to order and regularity, or something that the casinos in Monte Carlo and Las Vegas rely on to make profit of (assuming that they play fair, of course). However, it is not easy to come up with a rigorous, formal definition of randomness; a fact which in the course of time has caused a lot of worry and despair among mathematicians. Nature, on the other hand, does not worry about abstract definitions but generates random behavior as it is usually understood: through radioactive decay, noise in semiconductor devices and thermal fluctuations, for example. In fact, through physical systems it is relatively easy to grasp an idea what is actually *meant* by randomness.

In modelling and simulations of natural phenomena, randomness is needed to mimic the unpredictable elements of nature. In these cases, random events are generated with the help of random numbers. These numbers could be extracted directly from natural sources such as white noise with suitable analog-to-digital converters, but due to its impracticality this method is seldom used. Therefore, random numbers are often produced by computer algorithms known as (pseudo)random number generators (PRNG's). Since these algorithms are deterministic by construction, the best they can do is to imitate random behavior. Since there is no single, unique all-encompassing definition of randomness we can rather freely define a set of requirements that these algorithms need to fulfill in order to qualify for a given application.

Traditionally, the PRNG's are scrutinized by constructing a variety of different tests to probe certain properties of random number sequences. These tests can be divided into different categories. *Theoretical tests* probe the algorithm itself using number theoretic means whereas *empirical tests* probe the properties of the pseudorandom number sequences as produced by the algorithm in situ. Roughly speaking, a very practical criterion for a good generator is then easy to formulate: the more tests the generator passes, the better it is. Surprisingly enough, this almost trivial definition is the best available to date, if an operational approach for randomness is needed. We are thus faced with a practical task of developing tests to detect various kinds

of correlations in pseudorandom number sequences.

A number of such tests have been proposed over the years. The most well known set is included in the colossal work of Knuth [1]. These tests probe the properties of the output of PRNG's in many different ways (pair correlations, deviations from the desired distribution etc.). Over the years, this set of tests has become a benchmark that any generator should pass. However, as PRNG's with increasingly good properties have been devised there is a need for more efficient tests. In part this need has been filled by Marsaglia's test bench DIEHARD [2], a set of tests that he claims to be "more stringent" than the tests proposed before. However, recently problems have been reported [3] in the physics literature when high-precision Monte Carlo simulations have been performed with certain PRNG's, which perform very well in the existing tests [4]. Thus new and better tests are still needed!

A novel approach to testing random numbers is that based on physical systems. We have undertaken such an approach [5,6] to develop new tests, which have direct connections to the properties of certain models within statistical mechanics, most notably the Ising model and random walks. We call these tests *application specific*, since they mimic or directly use physical applications. We shall summarize some of our recent results below and show that these new tests constitute a very powerful tool in detecting subtle correlations not tractable by any other existing tests.

Our new tests can be divided into two categories. The first two, the *cluster test* and the *autocorrelation test* utilize known results of the two-dimensional Ising model, while the other two are based on random walks. In the cluster test [5], each bit of the output of a PRNG is transformed into up or down Ising spins, each spin sitting on a square lattice of size $L \times L$. When the whole lattice is filled the resulting configuration should have the same properties as the two-dimensional Ising model at an infinite temperature. Based on this we study the distribution of connected clusters that are formed by spins of the same sign which sit on nearest neighbor sites on the lattice. We then compare the observed distribution with the theoretical prediction. For some commonly used generators the results of this test are given in Table 1.

Another way of utilizing the Ising model is to study its properties at the critical temperature through Monte Carlo simulations [3]. This is in fact how problems with some PRNG's were recently discovered, most notably with a generator called R250. We have extended [6] this work by concentrating on a novel quantity, called the integrated autocorrelation time which turns out

to be a particularly sensitive measure of correlations present in computer simulations. Using this *autocorrelation test*, we have been able to show that the errors observed in simulations are due to very small correlations, whose effect can in some cases accumulate and cause erroneous results to appear.

Despite its physical appeal, the information given by the autocorrelation test in the Ising model is somewhat indirect. To be able to quantitatively measure the range of correlations we have devised two tests: the *random walk* and *n-block tests* [6]. In the random walk test we study the distribution of the final position of a two-dimensional random walk after n steps. In the n-block test, we basically utilize the same idea in one dimension. In both tests, if the distribution after n steps deviates from the expected significantly, the generator fails the test. Thus, by varying n we can find its smallest value where correlations start to appear. Some results for a group of generators with the random walk test are presented in Table 2 with $n = 1000$, where the failing of *e.g.* R250 tells us that the correlations are within less than 1000 successive numbers apart. Results of the n-block test with varying n are shown in Fig. 1, where the χ^2 values for R250 as a function of n are given. A dramatic increase close to $n = 250$ is observed, which implies that the range of correlations in R250 is approximately equal to the longer lag parameter $p = 250$ in its algorithm. Our results provide the first quantitative proof of this fact, which has been conjectured in the literature [7] to be the source of problems for R250.

To summarize, we have shown that application specific tests are needed to evaluate the performance of modern PRNG's. We have shown that some of these tests are capable of detecting correlations not unravelled by any other existing tests. However, it should be stressed that no test or no set of tests can guarantee that a given PRNG is good enough for all applications; careful tests mimicking the nature of the intended application should be conducted before any PRNG is employed!

[1] D. E. Knuth, *The Art of Computer Programming: Seminumerical Algorithms*, Vol. 2, 2nd Edition (Addison-Wesley, Reading, 1981).

[2] G. Marsaglia, in *Computer Science and Statistics: The Interface*, edited by L. Billard (Elsevier, Amsterdam, 1985), p. 3.

[3] A. M. Ferrenberg, D. P. Landau, and Y. J. Wong, Phys. Rev. Lett. **69**, 3382 (1992); P. D. Coddington, `cond-mat@babbage.sissa.it`

No. 9309017 (unpublished); W. Selke, A. L. Talapov, and L. N. Shchur, JETP Lett. **58**, 665 (1993); P. Grassberger, Phys. Lett. A **181**, 43 (1993); P. Grassberger, J. Phys. A **26**, 2769 (1993).

[4] I. Vattulainen, K. Kankaala, J. Saarinen, and T. Ala-Nissila, cond-mat@babbage.sissa.it No. 9304008 (unpublished).

[5] K. Kankaala, T. Ala-Nissila, and I. Vattulainen, Phys. Rev. E **48**, 4211 (1993).

[6] I. Vattulainen, T. Ala-Nissila, and K. Kankaala, cond-mat@babbage.sissa.it No. 9406054 (unpublished).

[7] P. Grassberger, Phys. Lett. A **181**, 43 (1993); A. Compagner and A. Hoogland, J. Computat. Phys. **71**, 391 (1987).

Random number generator	Failing bits	
	Cluster test	Equidistribution of bits
GGL	none	none
R250	none	none
R1279	none	none
RANMAR	$25-31$	$25-31$
RAN3	$1-4, 25-30$	$1-11, 24-30$
RAND	$7-31$	$22-31$

Table 1: Some results of the cluster test where bit number one denotes the most significant bit [5]. The last column denotes bits which fail in testing the equidistribution of ones. For a description of the PRNG's, see Ref. [4].

RNG	Result
R250	FAIL
RAN3	FAIL
R1279	PASS
RAND	PASS
GGL	PASS
RANMAR	PASS

Table 2: Some results of the random walk test with $N = 10^6$ samples [6].

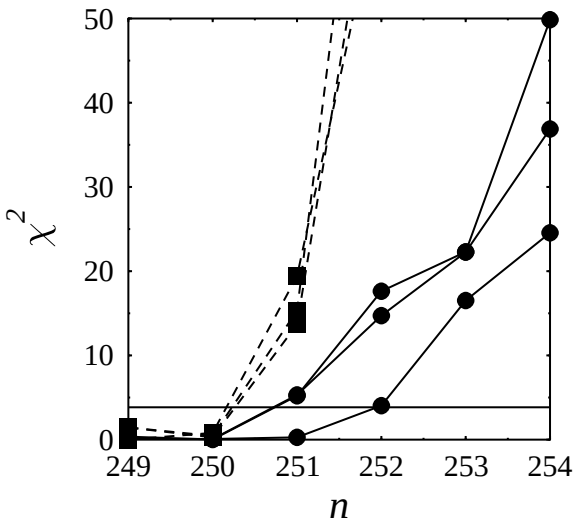

Figure 1: The χ^2 values for R250 in the n-block test [6]. Curves with circles and squares correspond to $N = 10^8$ and $N = 10^9$ samples, respectively. In both cases three independent runs have been performed. The points which lie above the horizontal line indicate failures in the test.