# Effective Control of Simulation Runs by a New Evaluation Algorithm for Correlated Random Sequences*

by Friedrich Schreiber**

Report from the Lehrstuhl für Allgemeine Elektrotechnik und Datenfernverarbeitung
Aachen University of Technology

A new algorithm is presented which allows the analysis of correlated random sequences in order to gain the stationary distribution function $F_n(x)$. This algorithm being based on recent statistical investigations of Markov chains can be used to control systematically the required number of trials $n$ of a computer simulation run by a formula which depends on the desired minimum value $F_{\min}$ of $F_n(x)$, on the prescribed upper limit $d_{\max}$ of relative error and also on the measured mean value of the correlation coefficient $\bar{\varrho}(x)$. As shown by selected examples the algorithm can process any correlated $x$-sequence of discrete and/or continuous type, detects rare event details of $F_n(x)$ and may replace therefore the conventional batch means evaluation method and other methods.

**Effektive Steuerung von Simulationsläufen durch einen neuen Auswertealgorithmus für korrelierte Zufallssequenzen**

Es wird ein neuer Algorithmus vorgestellt, der die Analyse von korrelierten Zufallssequenzen zwecks Gewinnung der stationären Verteilungsfunktion $F_n(x)$ durchführt. Dieser Algorithmus beruht auf neueren statistischen Untersuchungen an Markov-Ketten und eignet sich zur systematischen Kontrolle der bei einer Rechnersimulation erforderlichen Versuchszahl $n$ mittels einer Formel, die von dem gewünschten Minimalwert $F_{\min}$ der Funktion $F_n(x)$, von der oberen Grenze $d_{\max}$ des relativen Fehlers und auch von dem gemessenen mittleren Korrelationskoeffizienten $\bar{\varrho}(x)$ abhängt. Anhand ausgesuchter Beispiele wird gezeigt, daß dieser Algorithmus beliebige korrelierte $x$-Sequenzen vom diskreten und/oder kontinuierlichen Typ verarbeiten kann, seltene Ereignisse von $F_n(x)$ aufdeckt und daher die konventionelle Batch-Means-Methode und andere Methoden ersetzen kann.

## 1. Introduction

Since the work of Kosten [4] discrete event simulation on large computers has become the major tool for analyzing the performance of complex teletraffic and data processing installations. Essentially a simulation system deals with three tasks [3]:

1) Generation of random numbers of several prescribed distribution types.
2) Organization of task scheduling, list processing etc. representing the main body of the system.
3) Statistical evaluation of the output random sequence yielding the desired information about the performance of the simulated object: e.g. the mean and variance or if possible the complete stationary distribution function of a delay time.

Task (3) is often carried out by the *batch means* method [3], [9]; several other methods are known but it can be said that the statistical evaluation is up to now one of the weak points of simulation, see discussion on page 101 in [1]. This is an unsatisfactory situation because the evaluation results should be reliable in a well specified manner such that the engineer may take responsibility for any decisions based on the simulated results. For this purpose a trustworthy, always applicable evaluation method is necessary in order to control the length and also the cost of a simulation run which may become excessive considering typical run times on large computers of several hours. The new method to be described here is an extension of the Limited Relative Error (LRE) algorithm for *independent* $x$-sequences [5] to the correlated case and is based on the fundament of the Bayes-Laplace-statistics [6].

## 2. Statistic for Markovian Random Sequences

The essential result of recent statistical investigations of Markov chains can be summarized as follows. Consider a chronological $x$-sequence represented by the vector $\boldsymbol{x}_t = (x_1, x_2, \ldots, x_{n+1})$ with unknown correlation between neighbours[1] and with an unknown

---

* Revised and augmented version of a paper presented at the 12th International Teletraffic Congress (ITC), Torino (Italy), June 1–8, 1988.

** Prof. Dr.-Ing. F. Schreiber, Lehrstuhl für Allgemeine Elektrotechnik und Datenfernverarbeitung, Rheinisch-Westfälische Technische Hochschule Aachen, Templergraben 55, D-5100 Aachen.

[1] In the present context it is not necessary to use the term "autocorrelation".

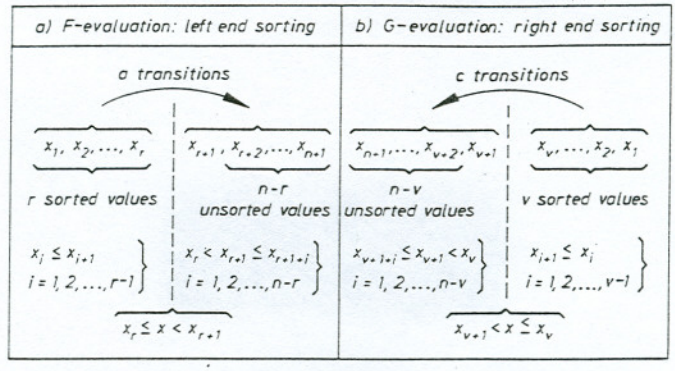


Fig. 1. Two modes for partial sorting of $n+1$ measured $x$-values.

stationary d.f. $F(x)$ resp. complementary distribution function (c.d.f.) $G(x) = 1 - F(x)$. This vector can be sorted as shown in Fig. 1 in two different ways: for gaining the empirical stationary d.f. $F_n(x)$ $r$ values form the ordered vector $\mathbf{x}_r = (x_1, x_2, \ldots, x_r)$, $x_i \leqq x_{i+1}$ at the *left* end resp. for gaining the c.d.f. $G_n(x)$ $v$ values form the ordered vector $\mathbf{x}_v = (x_v, \ldots, x_2, x_1)$, $x_{i+1} \leqq x_i$ at the *right* end of the observed $x$-range. The rest of measured $x$-values beyond each ordered vector forms a "global range" $x > x_r$ resp. $x < x_v$ which for the present we may consider to remain unordered. The essential point of the new statistical method is that not only the number of ordered $x$-values $r$ resp. $v$ but also the number of transitions $a$ resp. $c$ from each ordered range to the range $x > x_r$ resp. $x < x_v$ will be measured and evaluated in order to deal with correlation.

Then on basis of the BL-statistics for the 2-node Markov chain [7], [8] and observing the *large sample conditions*

$$n \geqq 10^3; \quad (r, v) \geqq 10^2; \quad (a, r-a, c, v-c) \geqq 10 \qquad (1)$$

the following posterior statements can be made [9]:

## 3. The Empirical Stationary Distribution Function

### 3.1. Evaluation method with prescribed F-levels

In many conventional methods the measured $x$-range must be divided into a number of $x$-intervals having a constant width $\Delta x$. This is a problematical task because only *a posteriori* i.e. after the bulk of measured $x$-values is known an adequate choise of $\Delta x$ can be attempted. Looking out for another solution we find that it is principally better to gain intervals by dividing the vertical $F$-axis with its known range $0 \leqq F \leqq 1$ as shown in Fig. 2.

Here we are able to prescribe *a priori* the desired range of investigation $F_{\max} \leqq F \leqq F_{\min}$ and also the desired number of intervals $k$ which determines the local resolution. Assuming as usually a logarithmical ordinate[2] we may compute a multiplier $m$ and the $F$-levels $F_j$

$$\left.\begin{aligned} m &= (F_{\min}/F_{\max})^{1/k}; \quad m < 1; \\ F_{j+1} &= m F_j = m^j F_1; \quad j = 1, 2, \ldots, k; \\ F_1 &= F_{\max}; \quad F_{k+1} = F_{\min}. \end{aligned}\right\} \qquad (5)$$

$F$-evaluation, $x_r \leqq x < x_{r+1}$ | $G$-evaluation, $x_{v+1} < x \leqq x_v$

*empirical stationary distribution function*

$$F_n(x) = r/n; \qquad (2\,a) \qquad | \qquad G_n(x) = v/n; \qquad (2\,b)$$

*mean correlation coefficient* $(-1 < \bar{\varrho} < 1)$

$$\bar{\varrho}_F(x) = 1 - \frac{a/r}{1 - r/n} \qquad (3\,a) \qquad \left| \qquad \bar{\varrho}_G(x) = 1 - \frac{c/v}{1 - v/n} \right. \qquad (3\,b)$$

*relative error concerning the statements eq. (2 a, b)*

$$d_F(x) = \left[\frac{1 - r/n}{r} \cdot \frac{1 + \bar{\varrho}_F(x)}{1 - \bar{\varrho}_F(x)}\right]^{1/2} \qquad (4\,a) \qquad \left| \qquad d_G(x) = \left[\frac{1 - v/n}{v} \cdot \frac{1 + \bar{\varrho}_G(x)}{1 - \bar{\varrho}_G(x)}\right]^{1/2} \right. \qquad (4\,b)$$

Eq. (2 a, b) represents simply the measured relative frequency as it should be for the large sample case. The control of the evaluation procedure will be performed by the error formula eq. (4 a, b) which takes into account the measured correlation effect. All further deductions will be restricted to the case "$F$-evaluation" since the treatment of the case "$G$-evaluation" is completely analogue, see also Section 4.1.

Fig. 2 shows that this procedure leads to $x$-intervals $I_j$ in the range[3] $X_{j-1} < x \leqq X_j$ whose variable widths

[2] For the case "linear ordinate" a linear incrementation of the type $F_{j+1} = F_j - \Delta F = F_1 - j\,\Delta F$ would have to be implemented instead of eq. (5).

[3] For avoiding an additional index the interval endpoints are expressed by *capital* letters: $X_j$ etc.

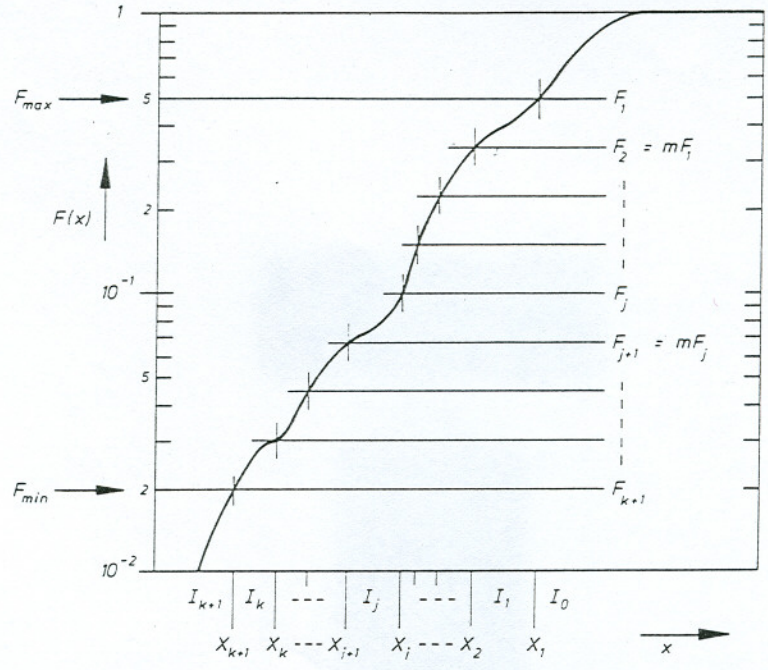


Fig. 2. "$F$-level" method: prescribed levels $F_j$ on the ordinate lead to variable $x$-intervals $I_j$.

$\Delta x_j = X_j - X_{j+1}, j = 1, 2, \ldots, k$ are adapted to the local shape of the $F$-curve. The $F$-level method is also well suited to detect *discrete points*, see Section 3.2.2.

### 3.2. The LRE-algorithm II

The LRE-algorithm II extracts automatically from a given stationary random $x$-sequence with unknown properties the empirical stationary d.f. $F_n(x)$ whereby the relative error $d_F(x)$ eq. (4a) is always kept smaller than the prescribed error limit $d_{max}$. This algorithm is typically performed by a statistical software module as part of a simulation system. A first fully operable version labeled "EDF 14b" has been written in PASCAL. In the following only the principal procedure steps are explained thereby omitting the *initialization phase*.

3.2.1. *The main iteration cycle.* In Fig. 3 it is assumed that the algorithm is in processing state $j$ i.e. that $x$-interval $I_j$ Fig. 2 has to be established next. In this state the limits $X_j, X_{j-1}, \ldots, X_2, X_1$ of the already established $x$-intervals $I_i$, $i = j-1, \ldots, 2, 1$ together with the measured values $l_i$, $r_i$ and $a_i$ are stored in the *F-result memory* FR, where $r_i$ and $a_i$ are defined by Fig. 1 and $l_i$ is the number how often the $x$-value $X_i$ has occurred: $l_i > 1$ resp. $l_i = 1$ means that $X_i$ is a *discrete* point resp. a *continuous* point belonging to a continuous part of the $F$-curve.

The *F-sorting memory* FS contains in state $j$ the left end ordered vector $(x_1, x_2, \ldots, x_{r_j})$ with $x_{r_j} = X_j$ being the right limit of interval $I_j$ and contains also to each sorted value $x_i$ its *chronological follower* $x_i^*$.

During the main iteration cycle the next $x$-value $x_t$ is sorted into FS provided that $x_t \leqq X_j$; if not then $x_t$ belongs to an interval $X_{h+1} < x_t \leqq X_h$ somewhere in memory FR. This means that we can state always a right *insertion limit* $X_h$ of $x_t$, $h = j, j-1, \ldots, 0$ with $X_0 = \infty$ including the case $h = j$, when $x_t$ is sorted into FS. Let $X_p$ be the insertion limit of the *predecessor* $x_{t-1}$; then in case $x_t > X_j$ i.e. $h < j$ the procedure "increment FR" in Fig. 3 works as follows:

a) increment the $r_i$-counters:
FOR $i = 1, 2, \ldots, h$ DO $(r_i := r_i + 1)$;
b) increment the $a_i$-counters:
IF $X_p < X_h$ THEN FOR $i = p, p-1, \ldots, h-1$ DO $(a_i := a_i + 1)$;
c) increment the $l_h$-counter:
IF $x_t = X_h$ THEN $l_h := l_h + 1$.

The cycle with either "sort into FS" or "increment FR"[4] is repeated until the relative error check $d_j \leqq d_{max}$ is fulfilled with $d_j = d_F(x)$, $X_{j+1} < x \leqq X_j$ eq. (4a). Then the values $l_j$, $r_j$ and $a_j$ are stored in FR whereby $a_j$ is computed depending on the locations of the followers $x_i^*$ in FS and $l_i$ is computed as the length of the $x_{r_j}$-section in FS:

d) compute $a_j$: $a_j = 0$; FOR $i = 1, 2, \ldots, r_j$
DO (IF $x_i^* > X_j$ THEN $a_j := a_j + 1$)
e) compute $l_j$: $l_j = 1$; $h = r_j$;
WHILE $x_{h-1} = x_{r_j}$ DO $(l_j := l_j + 1;\ h := h - 1)$.

3.2.2. *Finding the next sorting limit; detection of a discrete point.* Since the stationary d.f. $F(x)$ of a gener-

[4] The feature "increment FR" which has not been used in the former LRE-algorithm [5] improves a vast range of the $F_n(x)$-curve by decreasing the relative error $d_F(x)$ eq. (4a) far below its limit $d_{max}$ at the expense of a somewhat higher evaluation CPU-time.

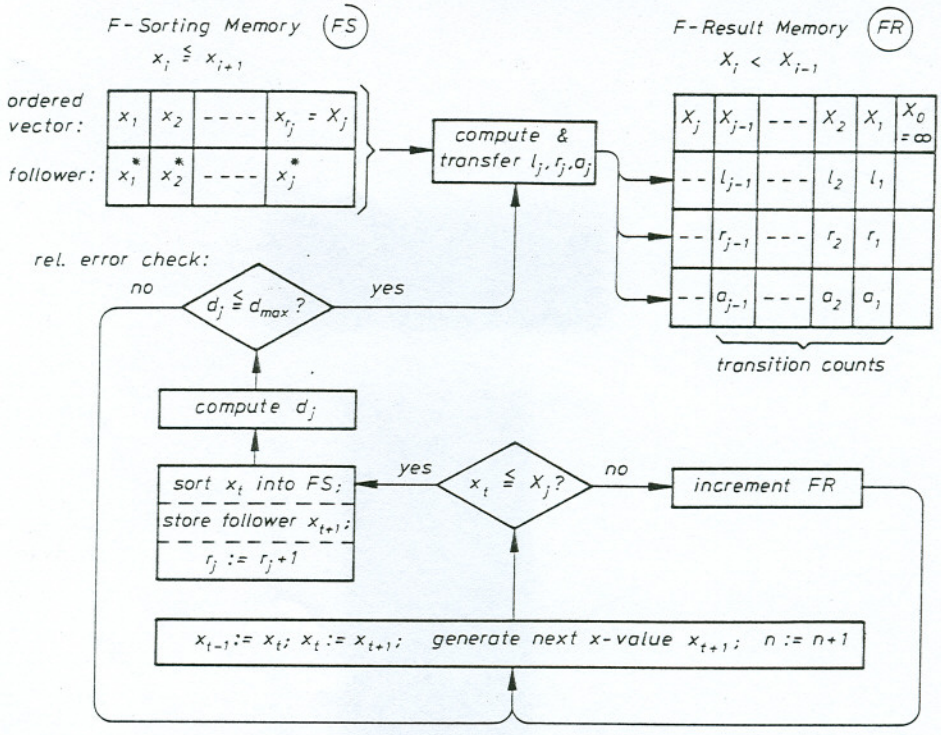


Fig. 3. LRE-algorithm II: Structure of $F$-memories and of the "main iteration cycle".

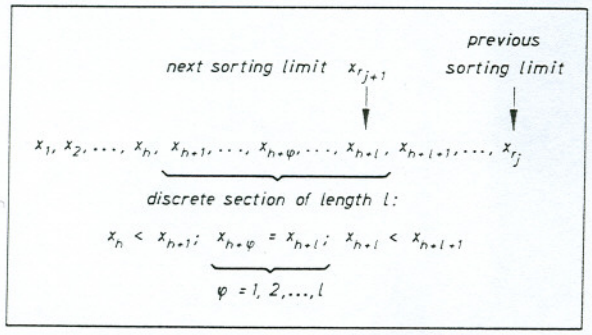


Fig. 4. LRE-algorithm II: Sorted $x$-values in sorting memory FS in case that the next sorting limit $x_{r_{j+1}}$ belongs to a discrete section.

al random process could have discrete points (this possibility is not shown in Fig. 2) the ordered vector in FS may contain "discrete sections" being characterized by a length $l > 1$ of *equal* $x$-values. When computing the next sorting limit $x_{r_{j+1}}$ the algorithm detects also such a discrete section and deals with it as follows (Fig. 4):

f) check for "end of evaluation": IF $(r_j/n) < F_{\min}$ GOTO Section 3.2.3 ELSE GOTO g);
g) find the *left* end of $x_{r_j}$-section in FS:
$r_j := r_j - (l_j - 1)$;
h) find $x_{r_{j+1}}$ in FS by computing:
$r_{j+1} = \text{INTEGER}\,(m\, r_j)$;[5] store $X_{j-1} = x_{r_{j-1}}$ in FR;
i) find the *right* end of $x_{r_{j-1}}$-section in FS:
$h - \varphi = r_{j+1}$;
WHILE $x_{h+\varphi+1} = x_{r_{j-1}}$ DO $(\varphi := \varphi + 1)$;
$r_{j+1} = h + \varphi$.

Then the algorithm returns to the main iteration cycle Section 3.2.1 with sorting limit $x_{r_{j-1}} = X_{j+1}$.

3.2.3. *Plot of $F_n(x)$-curve.* After the stop-condition $(r_j/n) \leqq F_{\min}$ has been met the evaluation results stored in result memory FR can be used to plot the desired $F_n(x)$-curve from right to left. For any interval $I_j$ with the range $X_{j+1} < x \leqq X_j$ we can compute the upper and lower level $F_j = r_j/n$ and $F_{j+1} = r_{j+1}/n$, $F_{j+1} < F_j$, see eq. (2a). The interval $I_j$ contains a discrete point at $X_j$ if $l_j > 1$, see Section 3.2.2. Then applying *linear* interpolation between continuous points of the interval it can be plotted as follows:

j) *pure continuous* interval due to $l_j = 1$: connect the coordinates $(X_j, F_j)$ and $(X_{j+1}, F_{j+1})$ by a straight line, see Fig. 5a;
k) *discrete* point a $X_j$ due to $l_j > 1$: draw at $X_j$ a vertical line between $F_j$ and $F_j^* = (r_j - l_i)/n$; connect the coordinates $(X_j, F_j^*)$ and $(X_{j-1}, F_{j-1})$ by a straight line.
This construction leads either to a partial discrete and partial continuous representation of the interval if $l_j < (r_j - r_{j+1})$, see Fig. 5b, or in case $l_j = (r_j - r_{j+1})$ i.e. $F_j^* = F_{j+1}$ to a pure discrete representation, see Fig. 5c.

Since the number of intervals $k$ can be chosen relatively high compared to the batch means method (e.g.

[5] This is due to eq. (5) with $F_j \approx r_j/n$ and $F_{j-1} \approx r_{j-1}/n$.

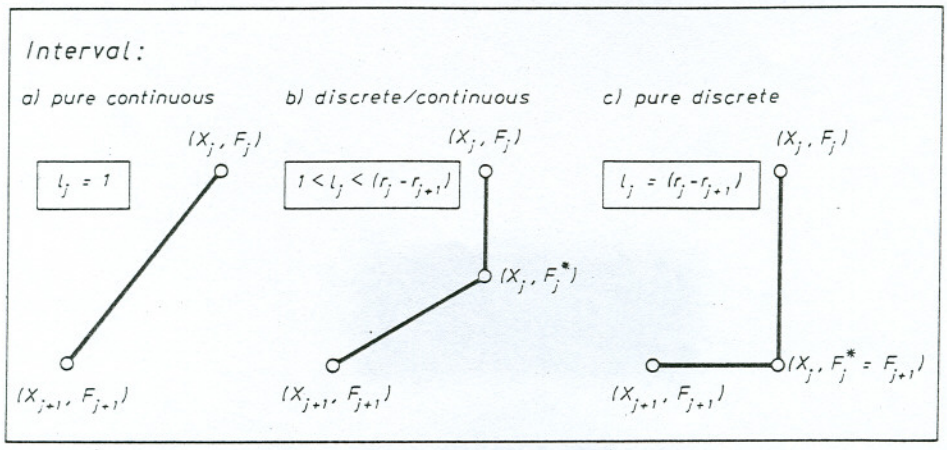


Fig. 5. LRE-algorithm II: Plot of interval $I_j$ of $F_n(x)$-curve depending on the number $l_j$ of *equal* $X_j$-values.

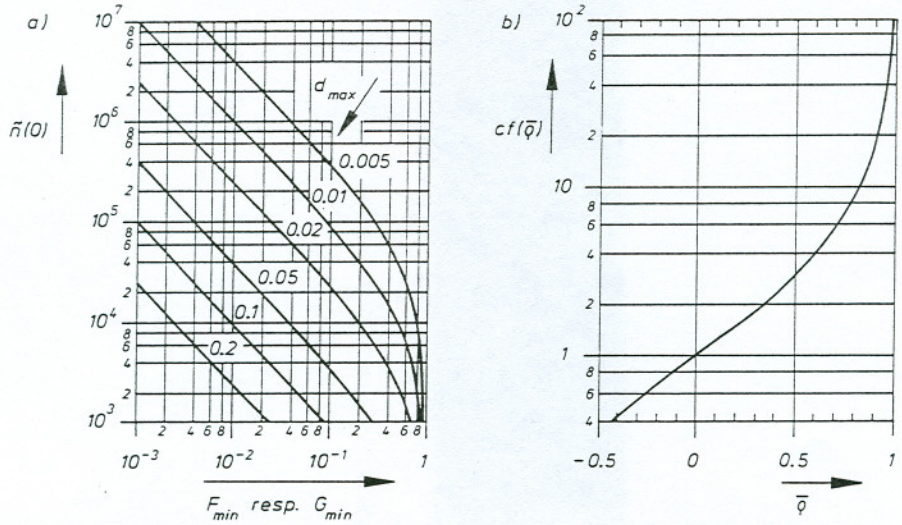


Fig. 6. (a) Required number of trials $\tilde{n}(0)$ in the uncorrelated case and (b) correlation factor $\mathrm{cf}(\bar{\varrho})$, see eq. (6).

$k = 140$ in Fig. 8) and since the relative error limit $d_{max}$ might be prescribed as low as it seems desirable the algorithm described here may trace the empirical $F_n(x)$-curve in a very fine manner according to the measured data.

In much the same way the curves for $\bar{\varrho}_F(x)$ eq. (3a) and $d_F(x)$ eq. (4a) can be plotted and the whole procedure described here for the case "$F$-evaluation" can be easily transcribed to the case "$G$-evaluation" which has been used for the experimental results in Section 3.4.

### 3.3. Control of simulation run length

From eq. (4a, b) and eq. (2a, b) we obtain the *required number of trials* $\tilde{n}(\bar{\varrho})$ for a given simulation problem

$$\tilde{n}(\bar{\varrho}) = \tilde{n}(0) \cdot \mathrm{cf}(\bar{\varrho}); \quad \mathrm{cf}(\bar{\varrho}) = \frac{1+\bar{\varrho}}{1-\bar{\varrho}};$$

$$\tilde{n}(0) = \frac{1-F_{min}}{F_{min}\, d_{max}^2} \quad \text{resp.} \quad \frac{1-G_{min}}{G_{min}\, d_{max}^2}, \qquad (6)$$

see Fig. 6. In this product formula the term $\tilde{n}(0)$ expresses the *uncorrelated* case $\bar{\varrho} = 0$ in accordance with eq. (5a, b) [5] which depends on prescribed parameters only, namely $F_{min}$ resp. $G_{min}$ and $d_{max}$; the term $\tilde{n}(0)$ is multiplied by the *correlation factor* $\mathrm{cf}(\bar{\varrho})$ where $\bar{\varrho}$ is the mean correlation coefficient eq. (3a, b) at level $F_n(x) = F_{min}$ resp. $G_n(x) = G_{min}$.[6] The simulation run length under control of the LRE-algorithm II follows exactly eq. (6), see the examples in Section 3.4. Under the real conditions of practical simulation $\bar{\varrho}$ is always positive and may well be in the range $0.8 < \bar{\varrho} < 1$. Therefore the correlation factor $\mathrm{cf}(\bar{\varrho})$ may attain e.g. value of $10 \ldots 20$ which means that the simulation run time will be substantially increased by this factor compared to the case $\bar{\varrho} = 0$. Since a simulation expert knows very soon the approximate value of $\bar{\varrho}$ resp. $\mathrm{cf}(\bar{\varrho})$ for a certain field of investigation he may forcaste realistically the length and cost of further simu-

[6] Eq. (6) can be generalized by omitting the fixation to the lowest level $F_{min}$ resp. $G_{min}$.

lation runs by means of eq. (6). All calculations with this formula confirm the experience that indeed only a powerful computer is adequate for simulation in the teletraffic field and in other fields like e.g. physics of elementary particles.

### 3.4 Experimental evidence

Several simulation experiments with various random processes have shown that the LRE-algorithm II can be universally applied. Its adaptability and performance will be demonstrated by evaluating the correlated $x$-sequences of a continuous, a discrete and of a mixed continuous/discrete random process.

- The cases of a *continuous* and of a *discrete* process are represented by the normalized delay time $x = \mu\tau_D$ resp. the occupancy $x$ of queueing system $M/M/1$ whose stationary c.d.f. $G(x)$ and correlation coefficient $\varrho(x)$ [2] are given by the equations:

*delay time*

$$G(x) = e^{-(1-A)x}; \tag{7}$$

$$\varrho(x) = 1 - \frac{(1-A)(1-e^{-(1+A)x})}{(1+A)(1-e^{-(1-A)x})}; \quad \varrho(\infty) = \frac{2A}{1+A}.$$

*occupancy*

$$G(x) = A^i; \quad (i-1) < x \leqq i; \quad i = 1, 2, \ldots;$$

$$\varrho(x) = 1 - \frac{1-A}{(1+A)(1-A^i)}; \quad \varrho(\infty) = \frac{2A}{1+A}. \tag{8}$$

A verification of the simulation result curves $F_n(x)$ and $\bar{\varrho}(x)$ Fig. (7 a, b) by means of eq. (7) resp. eq. (8) makes clear that these empirical curves have been established by the LRE-algorithm II very closely to the theoretical functions. The same is true for all other result curves traced by the algorithm up to now, though in principle a certain amount of statistical fluctuation may always occur especially in the vicinity of $F_{\min}$ resp. $G_{\min}$ where the relative error $d_F(x)$ resp. $d_G(x)$ comes close to its limit $d_{\max}$.

In both simulation experiments Fig. 7 a, b the prescribed number of desired intervals has been $k = 40$ but in Fig. 7 b due to the discrete nature of the occupancy process only 12 intervals have been actually established as it must be under the circumstances given here. The ability of the algorithm to adapt itself to discrete points rests on the $F$-level method Section 3.1 and on the procedure steps Section 3.2.2.

- The example of a mixed continuous/discrete random process is described by a *constant* correlation coefficient [7] $\varrho(x) = 0.5$ and by a stationary density $f(x)$ with three continuous sections and two discrete points as defined in Figs. 8 a. An interesting feature of this density is the "weak peak" with weight $p_3 = 0.05$ at the very right in Fig. 8 a which must be detected and traced by the evaluation algorithm.
An inspection of the simulation result curves Fig. 8 b shows that here too the empirical stationary c.d.f. $G_n(x)$ including its discrete points and the "rare event detail" and also the associated mean correlation coefficient $\bar{\varrho}(x)$ have been established very closely to the theoretical functions.

The numbers of trials $n$ needed for the experiments Fig. 7 and Fig. 8 confirm the formula for $\tilde{n}(\bar{\varrho})$ eq. (6). The number of *sorted events* $g$ in sorting memory GS (corresponding to memory FS in Fig. 3) is found to be much smaller than $n$ which means that the CPU-time for sorting is greatly reduced compared to the case that *all* $n$ measured values of the $x$-sequence would have to be sorted, see Section 5 [5].

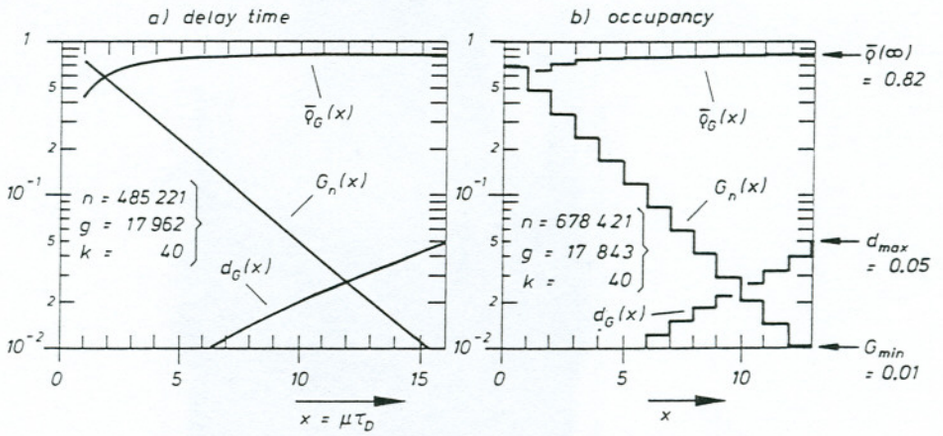


Fig. 7. Elementary queueing system $M/M/1$, load $A = 0.7$: empirical stationary c.d.f. $G_n(x)$ and associated empirical functions for (a) delay time and (b) occupancy.

Due to the feature "increment FR" explained in Section 3.2.1 the relative error $d_G(x)$ eq. (4 b) is far below its prescribed limit $d_{\max}$ in a vast range of the

[7] According to a proposal by W. Ding, Aachen the here applied new principle for generating correlated random sequences with the property "$\varrho(x) = \text{const.}$" allows any prescribed stationary density $f(x)$.

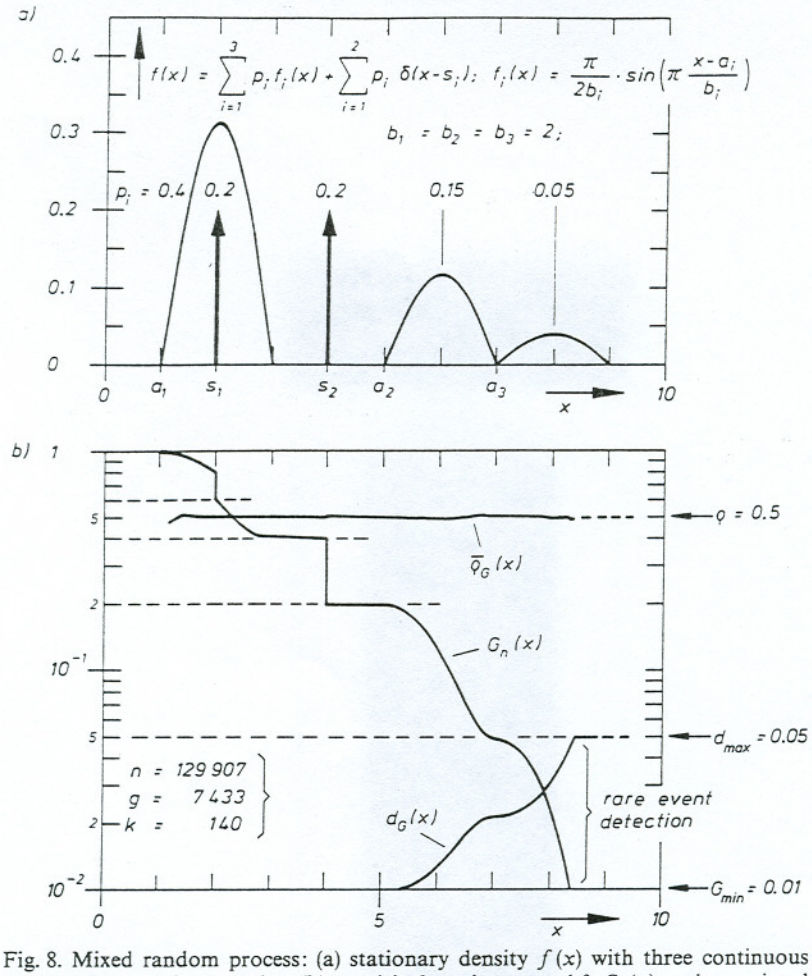


Fig. 8. Mixed random process: (a) stationary density $f(x)$ with three continuous sinus-sections and two poles; (b) empirical stationary c.d.f. $G_n(x)$ and associated functions.

$G_n(x)$-curve and only toward its right tail at $G$-levels close to $G_{min}$ it will approach the limit $d_{max} = 0.05$.

## 4. Further Remarks

1) In the teletraffic and queueing network field it is often necessary to investigate primarily the *right tail* of the distributions involved which means that the $G$-evaluation eq. (2b)–(4b) has to be applied. In other cases of course the $F$-evaluation eq. (2a)–(4a) for investigating the *left tail* might be preferred. Therefore in a general version of the LRE-Algorithm II the processing mode can be switched to either the $F$- or $G$-evaluation alone or to both evaluation types at the same time. In the latter case we may set $F_{max} = G_{max} = 1/2$ and the memories FS and FR in Fig. 3 are operated simultaneously with the corresponding memories GS and GR: now each generated value $x_t$ will be handled according to the alternative "sort into FS" of "increment FR" as described by Fig. 3 and also independently from this according to the alternative "sort into GS" or "increment GR".

2) The LRE-algorithm II can be developed further to include a modified version of the *window mechanism* Section 6.2 [5]: given the parameters $F_{max}$, $F_{min}$, $d_{max}$ and $k$ we might be interested to observe the details of a certain segment $F_a$ to $F_b$ ($F_{min} \leqq F_a < F_b \leqq F_{max}$) of the $F_n(x)$-curve with higher local resolution i.e. with a reduced relative error limit $d^*_{max} < d_{max}$ and with $k^*$ $F$-levels in this segment. To achieve this the window is opened at $F_b$ by switching $d_{max}$ to $d^*_{max}$ and the multiplier $m$ eq. (5) to $m^* = (F_a/F_b)^{1/k^*}$ and is closed at $F_a$ by switching back to $d_{max}$ and $m$. This concept can be generalized to allow for several windows along $F_n(x)$.

3) When simulating the number of trials i.e. of available $x$-values $n$ can be usually raised under control of the algorithm and is limited only by the duration resp. cost of the computer experiment. In other fields of statistical investigations the number $n$ might be limited from the beginning. Then the evaluation parameters has to be adapted to the given facts, and especially the relative error limit $d_{max}$ must be raised appropriately.

In cases where the large sample conditions eq. (1) can not be met i.e. where a *small sample* evaluation is necessary it is recommended to sort *all* measured $x$-values and to apply then the exact formula for the moments $M_l\{Q\}$ eq. (10) [8] with $F_n(x) = M_1$ and $d_F(x) = (M_2/M_1^2 - 1)^{1/2}$ and also for $\varrho(x)$ the exact formula eq. (17a) [7], $x_r \leqq x < x_{r+1}$. Here the relative er-

ror $d_F(x)$ could turn out to be relatively high but this would be in full accordance with the objectiveness of the posterior statements provided by the Bayes-Laplace statistics [6].

4) The $x$-values to be evaluated by the algorithm can be selected in a certain distance $\varkappa = 1, 2, \ldots$ within the given chronological sequence. For increasing values of $\varkappa$ this leads normally to reduced correlation and might therefore lead to a reduced resp. minimized simulation run time, see Section 6.2 [9]. A thorough discussion of this interesting point must be postponed.

5) The correlation coefficient $\varrho(x)$ of the "$F(x)$-equivalent" 2-node Markov chain as defined in [9] is not a mere matter of statistics but is an important, well defined property of such correlated random variables like occupancy, waiting and delay time etc. which occur in queueing systems, but are usually described only by their stationary properties. A first investigation has yielded among others the formulae eqs. (7), (8) of the $M/M/1$ queue [2], but it seems desirable to devote more work to the $\varrho(x)$-function.

*Acknowledgement*

The author thanks W. Ding, Aachen, for his contributions concerning the $\varrho(x)$-function and new random generators.

(Received June 24, 1988.)

**References**

[1] Bratley, P.; Fox, B. L.; Schrage, L. E.: A guide to simulation. New York: Springer, 1983.

[2] Ding, W.; Schreiber, F.: Correlation properties of random sequences generated by queueing system $M/M/1$ (in preparation).

[3] Fishman, G. S.: Principles of discrete event simulation. New York: J. Wiley & Sons, 1978.

[4] Kosten, L.: Simulation in traffic theory. Proc. 6th ITC, Munich 1970, 411/1–8.

[5] Schreiber, F.: Time efficient simulation: the LRE-algorithm for producing empirical distribution functions with limited relative error. AEÜ **38** (1984), 93–98.

[6] Schreiber, F.: The extended Bayes-postulate, its potential effect on statistical methods and some historical aspects. Proc. Int. Symposium on Probability and Bayesian Statistics, Innsbruck (Austria) 1986, 423–430. (R. Viertl ed.), New York: Plenum Press, 1987.

[7] Schreiber, F.: Measurement of the correlation coefficient of the 2-node Markov chain. AEÜ **40** (1986), 402–405.

[8] Schreiber, F.: Measurement of the stationary state probabilities of the 2-node Markov chain. AEÜ **41** (1987), 117–122.

[9] Schreiber, F.: The empirical stationary distribution function of Markovian correlated random sequences. AEÜ **41** (1987), 257–263.

NOVA ACTA LEOPOLDINA

Abhandlungen der Deutschen Akademie der Naturforscher Leopoldina. Im Auftrag des Präsidiums herausgegeben von Prof. Dr. Dr. Dr. h. c. J.-H. Scharf — Director Ephemeridum der Akademie.

Jetzt lieferbar:

**Raum und Zeit**

Vorträge anläßlich der Jahresversammlung vom 9. bis 12. April 1980 zu Halle (Saale)
Herausgegeben von J.-H. Scharf.
158 Abb., 5 Tab., Geb. DM 149,—
(Nr. 244)

WVDG **Wissenschaftliche Verlagsgesellschaft mbH**
**Postfach 10 53 39**
**D-7000 Stuttgart 10**

**WANTED!**
**WANTED!!**
**WANTED!!!**

Video & Video cassettes, office materials, Jewelleries, Watches & Giftwares, Lighters, Toys, Electrical & Electronics appliances, Utensils, Telephone, accessories, Hardwares, Airconditioners, Photographic equipments, Compugraphic machines & equipments, Motor – genuine spare parts & accessories. Also electrical trades Directory wanted.

Airmail all catalogues with Export prices to:

**The Manager,**
**ETS AYOOLA S.A.,**
**P.O. Box 2374,**
**Yaba, Nigeria**