A SIMULATION STUDY OF IDENTIFICATION OF THE MODEL OF THE FLOW IN CONTINUOUS STIRRED TANK REACTOR BY PRBS

A. HAVLÍČEK and J. ČERMÁK

Institute of Chemical Process Fundamentals, Czechoslovak Academy of Sciences, 165 02 Prague 6 - Suchdol

Received February 25th, 1976

Numerical simulation has been used to study the effect of certain factors on the accuracy of determination of the parameter of a first-order system — a continuous perfectly stirred tank reactor. The employed technique utilizes the pseudo-random binary signal (PRBS). In comparison with other perturbation signals that have been applied to the model of this system to date, PRBS offers the advantage of a shorter duration of experiment as well as a shorter interval Δt leading to the same accuracy. The most serious error is that incurred while generating the sequence of PRBS, this error having a greater weight when occurring toward the end rather than at the beginning of the period. In the presence of noise the output information/noise ratio is favourably affected by extension of the interval Δt as well as by augmenting the period N of PRBS. This, however, is accompanied by the adverse influence on the evaluation of the impulse characteristic by the method of moments. Optimum design of the parameters of the signal must therefore compromise between the two effects. The results of simulation studies confirm that the length of a single period $N\,\Delta t$ should not exceed the time required by the system to come to steady state, *i.e.* about 2—3 times the system time constant T_c.

The PRBS method is rather sensitive to monotonous trends. Filtration of the output signal by the "floating average" method destroys the information contained in the cross-correlation function. As more suitable appears the method of filtration due to Nikiforuk in the modification for a non-zero mean value.

The results of this simulation study have been used to model a real system.

Models of the flow of a fluid in apparatuses of chemical technology are commonly set-up using techniques based on the measurement of the response of the system in the effluent stream to a known concentration change of a tracer in the input stream. Most commonly used input signals are generalized deterministic functions such as *e.g.* pulse or a unit step change. The evaluation of the corresponding dynamic responses, which have become familiar as C or F-curves[1], may yield parameters of the model of a chosen structure.

Application of this already classic technique of identification of linear systems with a single input and output from the impulse or transient characteristic must necessarily suffer from the deficiencies of the physical realization of the given input signal — primarily the δ-function. The other drawback is that all output information is concentrated into a rather narrow space of time. Particularly important, however, is the fact that this method of identification in its most extensively used form does not account for the effect of various disturbances, noise of all kinds, *i.e.* external noise, noise generated by the system, high frequency noise, drift *etc.* In order to preserve the accuracy of estimation of a given parameter we have to resort to higher amplitudes of the input signals which may introduce the effects of nonlinearity or call for unnecessarily large perturbations of the quasi-stationary operation. The experiments have to be repeated, the data averaged with the adverse effect of nonstationarity.

The effect of certain disturbances may be apparent even from a simple experiment such as the identification of the model of the flow in a single-phase CSTR (ref.[2]). So much the more is this true in a two-phase gas–liquid reactor (*e.g.* aerated tanks), in a pilot-plant or a full scale equipment with poorly defined external conditions. Thus it is understandable that the possibilities offered by statistical identification methods were explored very early. Unlike the previous types of signals these methods utilize random or pseudo-random signals of small amplitude as inputs while preserving the quasi-stationary conditions. With the aid of the correlation and spectral analysis the sought output information is then obtained in the form of an estimate of the impulse characteristics or the transfer function of the system. The correlation method is the optimal method of evaluation in the sense of the mean square estimate.

Statistical identification enables the effect of the noise to be eliminated as well as other above mentioned deficiencies. On the other hand, its use brings along a substantially higher requirements as far as the computational effort is concerned which, until recently, has hampered its broader application.

Earlier works devoted to the identification of the flow models in CSTR's by statistical methods analyzed mostly the application of the true random signals.

Lapidus and Angus[3] have used true random sequence of binary (two-value) signals (*e.g.* ± 1) while the switching of the two values took place in fixed time intervals. Injection of diluted acid at the frequency of this signal into liquid entering a laboratory CSTR served to evaluate the time constant from pH response of the effluent liquid using spectral analysis. With the aid of the interval of switching, Δt, equal 1/6 of the time constant and a sequence of 240 intervals, the time constant, T_c, was found with an error of 11%.

Homan and Tierney[4] in a numerical simulation study analyzed identification of the same system using the inlet signal in the form of the Gaussian white noise generated by means of the tabulated normal distribution N(0, 1). The results of the simulation lead the authors to the conclusion that the accuracy of the estimate depends primarily on the ratios of the time interval, Δt, the length of the sequence, T_N, and the maximum lag of the correlation functions, τ_{max}, to the time constant of the system. In their simulation runs these ratios ranged between

$$\Delta t/T_c = 0{\cdot}04 - 0{\cdot}16\,, \quad T_N/T_c \approx 70 - 350\,, \quad \tau_{max}/T_c = 2{\cdot}7 - 21{\cdot}6.$$

Based on these results the authors recommend for the design of an identification experiment the following values: $\Delta t/T_c = 0{\cdot}16$, $T_N/T_c > 100$, $\tau_{max}/T_c = 10$, breaking eventually the sequence of the length T_N into parts of the length $T_p/T_c > 35$ and averaging the estimates of the time constant obtained from individual parts.

Goodman and coworkers[5] have tested the so-called thermal noise, *i.e.* the signal whose amplitude assumes values given by the normal distribution of constant mean and standard deviation. The appropriate values are switched at the time instants given by the Poisson distribution with the parameter $1/\Delta t$. Determination of the time constant to the accuracy of 10% for a laboratory CSTR and the amplitude of the perturbations of the temperature of the entering liquid equalling 1/4 of the steady state value required, according to the theoretical analysis, these values: $\Delta t/T_c = 0{\cdot}22$, $T_n/T_c = 65$, $\tau_{max}/T_c = 0{\cdot}12$. A suitable form of the input signals for estimation of the parameters of the models of real systems appears to be pseudo-random binary signals[6].

The principal advantage of the application of pseudo-random binary signals rests in the relative ease of their physical realization and the well-defined autocorrelation function. The form of the autocorrelation function permits direct determination of the impulse characteristic from the cross correlation function of the input and output. From this information we can determine the parameters of the model in the time domain

$$h(k\,\Delta t) = (1/\Delta t a^2)\,(N/N+1)\,[R_{xy}(k\,\Delta t) + \sum_{i=0}^{N-1} R_{xy}(i\,\Delta t)]\,. \qquad (1)$$

The pseudo-random binary signal (PRBS) in the form of a "linear recurrence sequence of maximum length" is realized by an equally spaced time sequence of two values "$+a$", "$-a$" with an interval Δt and a period N, where $N = 2^n - 1$, generated in a prescribed order following the general recurrence formula

$$\eta_k = C_1\eta_{k-1}^{(+)}C_2\eta_{k-2}^{(+)} \ldots C_{n-1}\eta(k-n-1)\,(+)\,C_n\eta(k-n) \qquad (2)$$

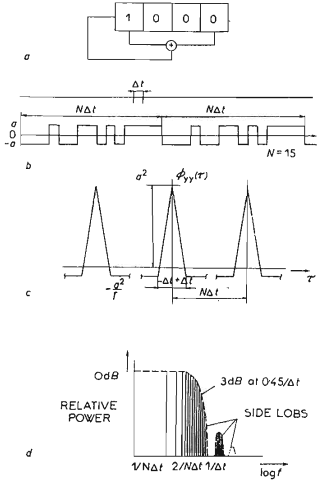

FIG. 1

Course of PRBS

a) Shift register generating PRBS for $N = 15$; *b*) PRBS for $N = 15$; *c*) autocorrelation function of PRBS; *d*) power spectral density of PRBS.

C_i are parameters given by the used period N, $(+)$ is the symbol of summation of the modulus 2. The process of generating the sequence for a period N may be represented by a shift register of length n whose first term initially equals unity and the remaining ones are zeroes.

The autocorrelation function of PRBS is given by

$$R_{\eta\eta}(\tau) = \begin{cases} a^2 \left[\dfrac{N+1}{N}\left(1 - \left|\dfrac{\tau}{\Delta t}\right|\right) - \dfrac{1}{N}\right. & |\tau| < \Delta t \\ -\dfrac{a^2}{N} & \Delta t \leqq |\tau| < (N-1)\,\Delta t \end{cases} \tag{3}$$

For $N = 2^4 - 1 = 15$ the shift register and the course of PRBS are shown in Fig. 1. The figure also shows the course of the autocorrelation function and the power spectral density for a general period N.

Prior to the application of PRBS to modelling the flow in a definite real equipment proper we focused our attention on the assessment of the effect of characteristic parameters and perturbations of PRBS on the accuracy of determination of the parameter of the CSTR through a simulation study. At the same time we were able to test the reliability of the data logging and processing system. The results of the analog study have been presented earlier[7].

A solution of the model in the form:

$$T_c(\mathrm{d}C_2/\mathrm{d}t) = C_1(t) - C_2(t)\,, \tag{4}$$

where $T_c = V/Q$ is the mean residence time of liquid in the reactor, C_2 is the outlet concentration of a tracer, C_1 is the inlet concentration of the tracer in the form of PRBS, has been obtained by integrating Eq. (4) using the Runge–Kutta fourth-order method in Gill's modification. The integration was performed on the Hewlet--Packard 9821 calculator for the forcing function $C_1(t)$ generated by the recurrence formula (2), where for the period $N = 31$ we took $C_1 = 2$ and $C_2 = 5$ while for the period $N = 63$ $C_1 = 1$ and $C_2 = 6$. Duration of a single pulse of PRBS, Δt, was taken identical with the integration step. In the discrete points Δt apart we calculated the impulse characteristic from the cross correlation function given in Eq. (1).

The simulation enabled the effect of the following factors to be studied independently: the effect of the initial condition of the model, the effect of the choice of the origin of PRBS, the effect of $\Delta t/T_c$ ratio, the effect of the length of the period, N, the effect of error in the generated sequence of PRBS, the effect of noise at the output of the system, the effect of drift at the output of the system.

A block diagram of the program for the simulation is shown in Fig. 2.

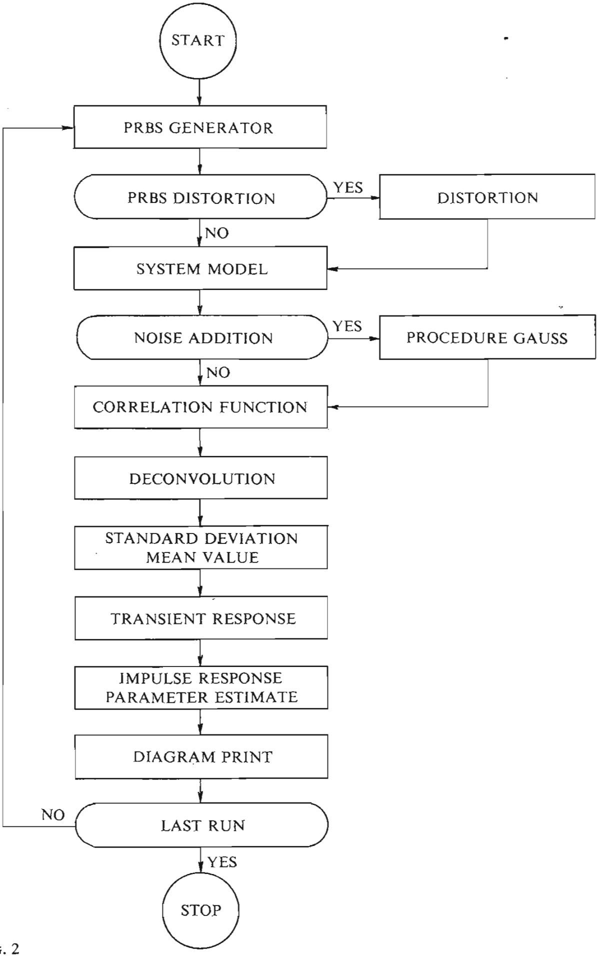

FIG. 2

Block Diagram of the Program "SIMUL"

RESULTS

Initial Condition of the Solution

The choice of the initial condition for the solution of the model (*i.e.* the initial concentration of the tracer in the system at the physical realization) is of considerable importance for the accuracy of determination of the time constant. As has been noted earlier, PRBS is a periodic deterministic signal and hence also the response of the system under ideal conditions is also periodic. The operator of the transfer of the model (*4*) may be defined in the Laplace domain by:

$$C_2(p) = F(p)\, C_1(p) + Q(p)\,, \tag{5}$$

where for the operator of the initial conditions we may write:

$$Q(p) = C_2(0)/(p + k)\,. \tag{6a}$$

In the time domain this operator transforms into the response due to the initial condition

$$Q(t) = C_2(0) \exp(-kt)\,. \tag{6b}$$

Thus it may seem that a zero initial condition (in the physical interpretation zero concentration of the tracer in the effluent stream from CSTR) should eliminate this term from the transfer characteristic. It turns out, however, that the operator $C_1(p)$ introduces into the time response another transient term due to the lack of the symmetry of PRBS about the mean (defined as the arithmetic mean of both values of the signal) in time. The transients of PRBS due to this deviation from the mean must die out first before the outlet concentration, $C_2(t)$, may fluctuate in the rythm of the period of PRBS. The effect of incorrectly adjusted initial condition is usually eliminated by measuring the response of several periods of PRBS and the first one carrying the error is then dropped from processing.

The final value of the first, dropped period represents then the effective initial value for the next period. On the contrary, if we started from this initial value already at the first period we might use already the results from the first period of PRBS.

In practice this conclusion can be interpreted as follows: The determination of the final value of the first period of PRBS, *e.g.* by a simulation calculation using the parameter of the real system as a first estimate could provide an initial value for experiment. The physical realization of the experiment with this initial value thus saves the experimental effort associated with the first period which otherwise would have been fruitless.

This conclusion, which have been utilized in our experiments, may thus considerably cut down the duration of experiments, particularly those involving large time constant. This is usually the case in large apparatuses of chemical technology.

The effect of a well and ill estimated initial condition for typical values of the parameters of PRBS and the model is illustrated in Fig. 3, 4.

The Effect of the $\Delta t/T_c$ *Ratio on the Accuracy of Determination of the Time Constant* T_c

In an application of PRBS to a system with an unknown time constant we have to estimate the time interval Δt so as to have $N \,.\, \Delta t$, *i.e.* the time of duration of a single period of PRBS, longer than the time the system takes to steady down. In such case we have to make a preliminary estimate prior to the experiment proper. This estimate can be made either from a theoretical analysis or from the transient response to a single step change. The study of the effect of the time interval Δt on the accuracy of the estimate of the time constant T_c was carried out for two periods of PRBS, namely N equal 31 and 63. These periods were then also used in the experimental work proper.

First we carried out a preliminary simulation calculation in order to determine the time constant for the given model of the first-order system on the TESLA 200 computer for the periods N equalling successively 31, 63, 127 and 255. Because the estimates of the time constants differed only very little, in subsequent more detailed

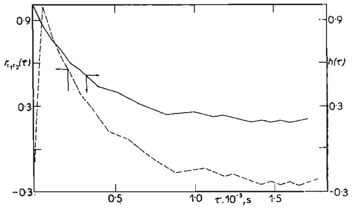

FIG. 3

The Effect of Incorrectly Selected Initial Condition

----- Normalized cross correlation function; —— impulse characteristic.

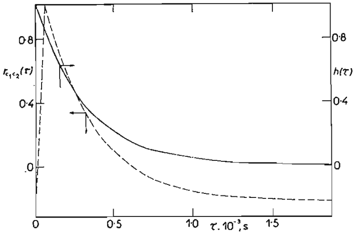

FIG. 4

The Effect of Correctly Selected Initial Condition

----- Normalized cross correlation function; —— impulse characteristic.

studies we used the shorter period $N = 31$ and $N = 63$ as being more convenient from the point of view of experimental effort. The time of duration of the experiment can be proportionately cut down for longer periods by decreasing the time interval Δt. In this respect, however, we were constrained by the frequency properties of the system as well as the probes and measuring instruments.

The simulation was carried out for the model (*4*) with the parameters corresponding to the real experimental set-up, correctly adjusted initial condition, the period of PRBS $N = 31$ and 63, and a number of values of the ratio $\Delta t/T_c$.

The values of the time constant $\tilde{T}_c$ obtained from the evaluation of the impulse responses by the method of moments exhibit an interesting dependence on the value of the ratio $\Delta t/T_c$. These results yield different conclusions than the so far published recommendations for the choice of the repetition frequency of PRBS in relation to the magnitude of the dominant time constant of the identified system.

This dependence is plotted in Fig. 5 for the periods $N = 31$ and $N = 63$.

For the period $N = 31$ it was found that the experimentally determined value of the time constant of the modelled system well agrees with the theoretical one in the interval $\Delta t/T_c \in (0{\cdot}15, 1{\cdot}20)$.

In this interval the maximum error of determination of the time constant was 7·5%, the optimum value of the ratio $\Delta t/T_c$ was 0·45. Outside this interval the found values significantly differ from the theoretical value.

The corresponding interval of $\Delta t/T_c$ for the period $N = 63$ is somewhat broader, namely $(0{\cdot}10, 1{\cdot}20)$. This conclusion is at odds with that found in the literature[5] recommending a substantially narrower interval $(0{\cdot}10; 0{\cdot}15)$. The found dependences may be explained by the following reasoning pertaining the frequency domain. The density of spectral lines in the spectrum of PRBS shown in Fig. 1*d* is given by the product $N\,\Delta t$; their number is given by the value of N. The frequency characteristic of the system is then covered by a certain number of the spectral lines the number of which depends more on the ratio of Δt and T_c than N alone. Accordingly, the curves of the above plot for $N = 31$ and $N = 63$ in this region are approximately identical.

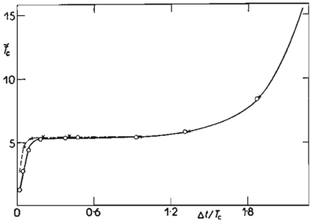

FIG. 5

Dependence of the Estimate of T_c on the Ratio $\Delta t/T_c$ for the Theoretical Time Constant $T_c = 319{\cdot}2$ s

——— Curve for $N = 31$, - - - - - curve for $N = 63$.

At low values of $\Delta t/T_c$ the system at the time $N\ \Delta t$ has not yet steadied (in the extreme case $N\ \Delta t < T_c$), which explains the rather large error of the estimate of T_c for the values of $\Delta t/T_c$ below the lower limit of the corresponding interval. The spectral lines of the power spectral density of PRBS are then crowded into the region of higher frequencies and cover therefore only a part of the frequency characteristic. The error for large values of the $\Delta t/T_c$ ratio can be explained by the fact that the spectral lines in the power spectral density spectrum of PRBS are, on the contrary, crowded into the region of low frequencies. This effect, however, does not show so markedly as in the opposite case.

The Effect of the Length of the Period

From the above we may conclude that over a wide range of the values $\Delta t/T_c$ there is no substantial difference in the accuracy of determination of the constant T_c for the periods $N = 31$ and $N = 63$.

Under such conditions the period $N = 31$ thus should be satisfactory. The effect of the length of the period for selected conditions of the simulation is illustrated in Fig. 6.

The Effect of Error in the Sequence of PRBS

Technological pilot-plant and large scale units respond usually slowly to a disturbance; the dominant time constants range mostly between 10^1 and 10^4 s.

The generators of PRBS working on any principle provide without problem signals in higher frequency region. However, at low frequencies, *i.e.* at a large time interval, the probability of malfunction due to *e.g.* power service interruption, electromagnetic influences, *etc.*, increases. As a result, the sequence of the two-value signal is either incomplete or imperfect. Numerical values evaluated from the response of the system to such a signal necessarily introduces numerous errors and inaccuracies into the solution.

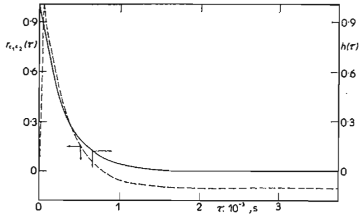

FIG. 6

Typical Identification of a First-Order System for $N = 63$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s

----- Normalized cross correlation function; ——— impulse characteristic.

The autocorrelation function of PRBS then does not correspond to the definition (Fig. 1) and/or to the assumptions implied in the determination of the impulse characteristic of the system.

Continuous test of the function of the generator poses problem. In the work proper we tested only the number of impulses in a single period by a counter. Unfortunately, this test does not indicate the type of the error in the sequence.

In the simulation experiments we tested the effect of various defects in the sequence of PRBS on the quality of the impulse characteristics of the given model system.

The results of malfunction of the generator of the sequence of PRBS of $N = 31$ in the 5-th and the 26-th point of the sequence are shown in Figs 7, 8.

Expressing the difference between the theoretical impulse characteristic and the characteristic obtained from the erroneous PRBS for $N = 31$ and the conditions: $T_c = 319{\cdot}8$ s, $\Delta t/T_c = 0{\cdot}188$ as

$$\sigma_h = [(1/N - 1) \sum_{i=1}^{N} (h_i \exp - h_{ir})^2]^{1/2} , \qquad (7)$$

where $h_{i\exp}$ are found ordinates of the impulse characteristic, h_{ir} are theoretical ordinates of the impulse characteristic, then for the first case (the error in the initial

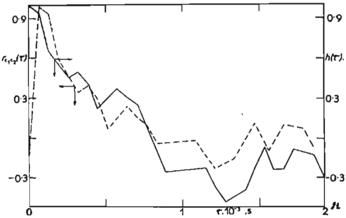

FIG. 7

The Effect of Defect in the 5-th Point of the Sequence for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s

- - - - - Normalized cross correlation function; ——— impulse characteristic.

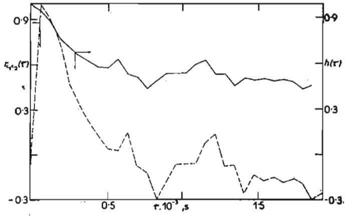

FIG. 8

The Effect of Defect in the 26-th Point of the Sequence for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s

- - - - - Normalized cross correlation function; ——— impulse characteristic.

part of PRBS) we obtain $\sigma_h = 0{\cdot}0124$; for the second case (the error near the end of PRBS) we obtain $\sigma_h = 0{\cdot}1725$.

Evaluation of the time constant by the method of moments yields in the first case $\tilde{T}_c = 178{\cdot}7$ s, in the second case $\tilde{T}_c = 801{\cdot}6$ s. The eror near the end of the sequence thus has a much greater weight ruining the experiment and putting the result beyond acceptable limits.

The Effect of Noise in the Response of the System

The usual simplified approach to the evaluation of the effect of external noise assumes that noise of any kind superimposes on the response of the system. As a noise we regard disturbances of higher frequency in the order of magnitude than the frequency region of the system proper. Lower-order noise is usually referred to as drift or trend. Statistical methods display relatively low sensitivity to external noise provided the cross correlation between the input signal, *i.e.* PRBS, and the noise is low or none. The error due to the noise may be expressed in terms of the cross correlation of the input and the noise, $R_{xn}(t)$, *i.e.* as a function of time. For a compact expression over the period we can use standard deviation of this function.
Thus:

$$\sigma_R^2 = E\{R_{xn}^2(t)\} - (E\{R_{xn}(t)\})^2 . \tag{8}$$

In simplification for the white noise Briggs[9] published the final relationship between the mean of the cross correlation function of the input and output in the form:

$$\frac{\bar{R}_{xy}}{\sigma_R} = \left[\frac{aG}{\sigma_n}\left(\frac{\Delta t}{T_s}\right)^{1/2}\right]\left(\frac{N\,\Delta t}{T_s}\right), \tag{9}$$

where G is static amplification of the system, a is the amplitude of PRBS, σ_n is the standard deviation of the noise, T_s is the time for steadying the cross correlation.

For the purpose of analyzing the effect of the noise in real and simulated systems we express the ratio of the information contained in the input/output cross correlation function to the standard deviation of the input/noise cross correlation function. From Eq. (9) it further follows that increased interval Δt, amplitude and the length of the period, N, favourably affect this ratio. The choice of the parameters Δt and N, however, is constrained by additional aspects as show the results of this simulation study as well as of real experiments. Duration of a single period of PRBS should not exceed too much the time T_s for two reasons.

First of all, the bulk information about the dynamic behaviour of the studied system is obtained from that section of the cross correlation function before the instant T_s. The rest of the curve unnecessarily prolongs the experiment. In addition, if we use for evaluation of the impulse response the method of moments (obtained from the cross correlation function) the impulse response after the time T_s oscillates only due to the presence of noise. The effect of this portion of the impulse characteristic becomes effective in the evaluation of the r-th moment proportionately to the r-th power of time. The adverse effect of the prolongation then outweights the improvement of the ratio of meaningfull information to noise. For a quantitative expression of the signal/noise ratio let us define further the standard deviation of the output signal of the system as:

$$\sigma_{c_2} = \{\sum_{i=1}^{N} (c_2(i\,\Delta t) - \bar{c}_2)^2/(N-1)\}^{1/2}\,, \tag{10}$$

where σ_{c_2} is an estimate of the standard deviation of the output signal from N samples and $\bar{c}_2$ is the corresponding mean.

The noise was generated by the generator of random numbers with the normal distribution[10] $N(0, 1)$.

The random number at the end of each integration interval Δt was added to the output of the model. The sample mean and sample variance for a sample of 1000 random numbers was 0·0294 and 1·0692. The autocorrelation function of noise was approximated by the expression $(\exp -0{\cdot}0425t)$ for the time increment $\Delta t = 60$ s. The estimate of the autocorrelation function was calculated for 250 random numbers and a maximum delay of 24 intervals. Eq. (*8*) was verified for the period $N = 31$,

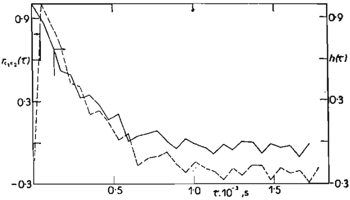

FIG. 9

The Effect of Noise Superimposed on the System Output for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s; $\sigma_{c_2}/\sigma_n = 1{\cdot}2$
------ Normalized cross correlation function; ——— impulse characteristic.

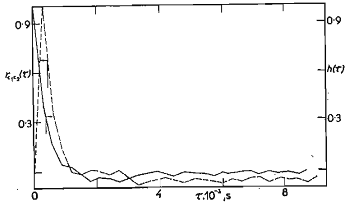

FIG. 10

The Effect of Noise Superimposed on the System Output for $N = 31$, $\Delta t = 300$ s, $T_c = 319{\cdot}2$ s; $\sigma_{c_2}/\sigma_n = 1{\cdot}2$
------ Normalized cross correlation function; ——— impulse characteristic.

Δt = 60 s, the time constant T_c = 319·2 s and the ratio σ_{c_2}/σ_n = 1·2 with the result shown in Fig. 9. For the same parameters, except that Δt = 300 s, the result is shown in Fig. 10. The evaluated time constant T_c for the whole period of PRBS in the first case carried an error of 106%, in the second case 31%. Increased effect of noise for Δt = 60 s from σ_{c_2}/σ_n = 1·2 to 0·58, *i.e.* approximately twice, changes the error of the time constant T_c to only about 120%. This evidences the damping effect of the correlation method on the effect of noise.

If the impulse characteristic is evaluated only from the time T_s, then for Δt = 60 s the error of determination of the time constant T_c diminishes for σ_{c_2}/σ_n = 1·2 to 6·3% and for σ_{c_2}/σ_n = 0·58 to 6·9%. However, if we change the period N from 31 to 63 keeping the time interval Δt = 60 s and σ_{c_2}/σ_n = 1·2 (Fig. 11) the error of the time constant evaluated from the impulse characteristic over the whole period of PRBS increases from 106% to 208%. This exemplifies the adverse effect of prolonged experimental time on the calculation of the time constant by the method of moments.

It is apparent that the optimum choice of the parameters N and Δt must be made keeping Eq. (*9*) as well as the accuracy of the method of moments in mind.

The Effect of Drift

The effect of drift becomes particularly annoying in real long-term measurements when the logged data may be affected by fluctuations of the voltage in the power network, ambient temperature, *etc.*

In the simplest case the effect of the trend may be diminished by fitting a straight line to it, or we may use digital filters of various types and function. Prior to examining the effect of some typical trends we would like to point out that the estimates of time constant T_c obtained by PRBS method with moment evaluation are rather insensitive to the additions of harmonic components at the output of system. A sinus trend for the period of which the group $N\,\Delta t/T_{trend}$ is an integer plays no role. If this ratio is greater than 10 then even if the above group is not an integer the effect of the harmonic component remains small.

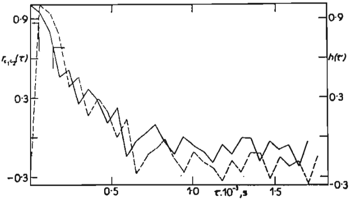

FIG. 11

The Effect of Noise Superimposed on the System Output for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s, $\sigma_{c_2}/\sigma_n = 0{\cdot}58$
----- Normalized cross correlation function; ——— impulse characteristic.

Fig. 12 shows the response of the studied model to PRBS of $N = 31$ and $\Delta t = 60$ s to which we added a sinus component with the period equalling 1/3 of the period of PRBS. This means that the ratio of both periods is an integer. The amplitude of the sinus component was equal 0·1 of the amplitude of PRBS. The evaluated time constant T_c carried an error of 2·1%. Fig. 13 corresponds to the same conditions except that the amplitude of the sinus component equals unity. The error of determination of the time constant of the system was 2·2%. These results are due to the specific features of the moment method. Had we used any other method of parameter estimation from impulse characteristics in Figs 12, 13 (*e.g.* direct search) we would have obtained quite misleading results, notwithstanding the fact that the presumed struc-

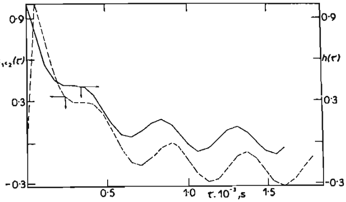

FIG. 12

The Effect of a Sinus Trend on the Cross Correlation Function and Impulse Characteristic

Amplitude of the trend is 0·1, $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s; ----- normalized cross correlation function; ——— impulse characteristic.

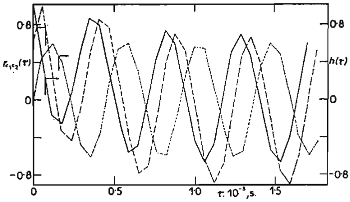

FIG. 13

The Effect of a Sinus Trend on the Cross Correlation Function and Impulse Characteristic. Amplitude of the Trend is 0·1, $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s

----- Normalized cross correlation function; ——— impulse characteristic.

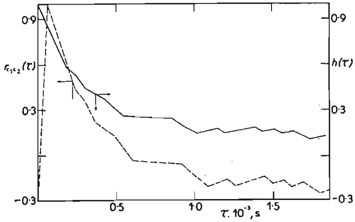

FIG. 14

The Effect of a Linear Trend $1 + i/31$ for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s

----- Normalized cross correlation function; ——— impulse characteristic.

ture of a linear, first-order model is no longer justified. Fig. 14 shows the results for the run when the input of the system was mixed with a linear drift $(1 + (i\,\Delta t)/31)$. The determination of the time constant had in this case an error of 49%. The effect of the trend of this kind becomes the more manifest the longer the experiment.

Elimination of the effect of the trend was firstly studied on a "floating average"[11] digital filter which may be described by the equation:

$$\bar{c}_{2i} = c_{2i} - (1/2M) \sum_{k=-M}^{+M} c_{2(i-h)}, \quad i = 1, \ldots, N. \tag{11}$$

The application of this filter in the time domain corresponds in the Fourier domain to the relation:

$$F\{\bar{c}_{2i}\} = F\{c_{2i}\}\left[1 - (\sin \omega T/\omega T)\right], \tag{12}$$

where $T = M\,\Delta t$.

This filter preserves frequencies higher than $1/2T$ and supresses lower frequencies.

Illustration of the function of this filter is shown in Fig. 15, where the trend has parabolic form $1 + (i\,\Delta t/10)^2$, $i \in \langle 1, 31 \rangle$. Superimposed to this trend is random noise with the Gauss distribution $N(0, 1)$.

The trend was filtered according to Eq. (11) where M was put equal 3. The filtered curve is also shown in Fig. 15*a*. Fig. 16 represents the same filter applied to the curve described by the equation $\{1 + (i\,\Delta t/10) + 5 \sin[(3\pi/31)\,.\,i\,\Delta t]\}$ mixed with the same noise. The plotted curve after filtration preserves higher frequencies; the lower frequencies, *i.e.* the equation proper have been smoothed. The curvature of the

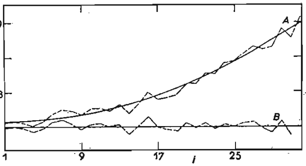

FIG. 15

Application of a Floating Average Filter to a Mixture of Gaussian Noise and a Parabola $1 + (i/10)^2$

Curve *A* is the mixture; after filtration we obtain curve *B* without parabolic trend.

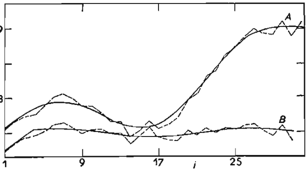

FIG. 16

Application of a Floating Average Filter to a Mixture of Gaussian Noise and a Trend $1 + (1/10)^2 + 5 \sin [3\pi/31\ \Delta i]$

Curve *A* is before, curve *B* after filtration.

filtered curve for the first three points corresponds to summation between $M = 3$ in Eq. (*11*). Up to this bound the filter is inoperational.

Further we analyzed the effect of this filter on the cross correlation function serving for evaluation of the response to PRBS. For the cross correlation function of PRBS and the filtered output signal $\bar{c}_2$ one can derive:

$$R_{c_1\bar{c}_2}(j) = R_{c_1c_2}(j) - (1/2M) \sum_{k=-M}^{+M} R_{c_1c_2}(j - k)\,. \qquad (13)$$

This equation points at the limitations of the application of the "floating average filtration" to the response to PRBS. It may be interpreted as follows: Filtration of the response of a system to PRBS means at the same time filtration of the calculated cross correlation function with simultaneous destruction of the contained information. This fact is illustrated by the results of simulation in Fig. 17 where the response of the model was filtered according to Eq. (*11*) for $M = 3$.

A similar effect for the case of PRBS as the input signal would exhibit also some other general digital filters.

For this reason we sought a special method of filtering respecting the specific features of PRBS.

Of several possible alternatives worked out for this particular case in the literature[12-14] we took the one due to Nikiforuk[15]. This author tested the method exclusively on simulated data and zero-mean PRBS.

The values of the impulse characteristic are solved in this case from the following convolution equation between the cross correlation function and the autocorrelation function of the filtered signals C_1 and C_2:

$$R_{S_Q(C_1)D_Q(C_2)}(k) = \sum_{j=0}^{N-1} h_j(-1)^Q R_{C_1C_1}(k - Q - j)\,, \qquad (14a)$$

where the operation on the signal C_1 *i.e.* $S_Q(C_1)$ may be defined by the recurrence

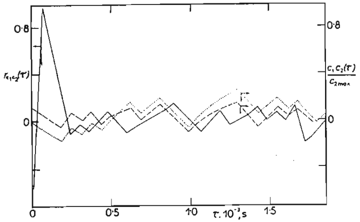

FIG. 17

Destruction of the Normalized Cross Correlation Function by the Floating Average Filter for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s ············ Response of system before filtration; ----- response after filtration; ——— normalized cross correlation function.

formula as:

$$S_{Q(C_{1i})} = -(1/N) \sum_{j=r-N+1}^{r} (r - j + 1) S_{Q-1}(C_{1j}) + \sum_{j=r-N+1}^{i} S_{Q-1}(C_{1j}) \qquad (14b)$$

while $S_0(C_{1i}) = C_{1i}$.

The operation on the signal C_2, *i.e.* $D_Q(C_2)$ is defined by

$$D_Q(C_{2i}) = \sum_{q=0}^{Q} (-1)^q \cdot \frac{Q!}{q!(Q-q)!} \cdot C_{2(i-q)}, \qquad (15)$$

where r is a current index, Q is the order of the filter.

The case of a non-zero mean may be regarded as a constant drift and increased order of the difference eliminates this non-zero value too.

On applying the filter described by Eqs (*14a,b*) (*15*) for $Q = 2$ to eliminate the drift $(1 + i\,\Delta t/31)$ whose effect has been shown in Fig. 14 (the error of T_c is 49%) for the same model we obtain the result plotted in Fig. 18. The error of the time constant evaluated by the method of moments is then 3·7%. The application of the filter with the third difference leads to additional improvement; the error of the time constant is then only 0·3% – see Fig. 19.

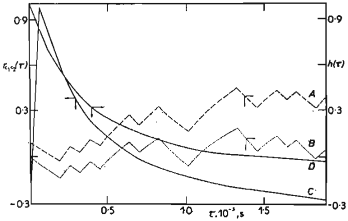

FIG. 18

Nikiforuk's II-nd Order Method of Filtration of a Linear Trend $1 + \Delta ti/31$ for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s

Curve *A* is the response of the system before filtration, curve *B* is the response of the system after filtration, curve *C* is the normalized cross correlation function, curve *D* is the impulse characteristic.

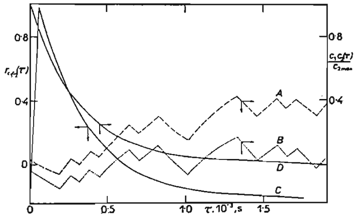

FIG. 19

Nikiforuk's III-rd Order Method of Filtration of a Linear Trend $1 + \Delta ti/31$ for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s

Curve *A* is the response of the system before filtration, curve *B* is the response after filtration, curve *C* is the normalized cross correlation function, curve *D* is the impulse characteristic.

It must be stressed, however, that in these simulated runs we did not allow for the presence of high-frequency noise as is the case of real systems.

The Nikiforuk's method somewhat amplifies the effect of noise. The simulation of the model with noise in the output signal is shown graphically in Fig. 9. In this case the deviation of the theoretical and evaluated impulse characteristic, expressed in terms of the characteristics σ_h (Eq. (7)), was $3{\cdot}73 \,.\, 10^{-3}$.

After filtration according to Eq. (*14a,b*), with the second difference, this characteristic increases to $5{\cdot}16 \,.\, 10^{-2}$.

Such filters may be applied for the purpose of removing the effect of improperly set initial conditions as has been already discussed in one of the previous paragraphs.

In case of an incorrectly selected initial condition illustrated in Fig. 4 the time constant was evaluated with an error of 48%, after filtration with the above II-nd order filter the error dropped to only 2·6%.

This filter was applied also to eliminate the effect of initial conditions for real experimental data. To our knowledge, this is the first application of this filter to real data as well as a new method of evaluating data from the very first period of PRBS. Application of the filter with the third difference to a parabolic drift in the form $[0{\cdot}1 + (i\,\Delta t/31) + 0{\cdot}5(i\,\Delta t/31)^2]$ is shown graphically in Fig. 20; the error of T_c was 0·3%. Fig. 21 shows the unfiltered result where the error of determination amounted to 161%.

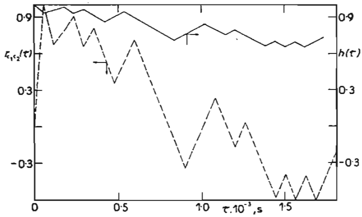

FIG. 20

Nikiforuk's III-rd Order Method of Filtration of a Trend $0{\cdot}1 + (\Delta ti/31) + 0{\cdot}5(\Delta ti/31)^2$ for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s

Curve *A* is the response of the system before filtration, curve *B* is the response after filtration, curve *C* is the normalized cross correlation function, curve *D* is the impulse characteristic.

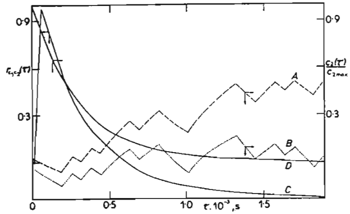

FIG. 21

Distortion of the Normalized Cross Correlation Function by the Trend $0{\cdot}1 + \Delta ti/30 + 0{\cdot}5(i\,\Delta t/31)^2$ for $N = 31$, $\Delta t = 60$ s, $T_c = 319{\cdot}2$ s

------ cross correlation function;

——— impulse characteristic.

CONCLUSIONS

The results of this numerical simulation of statistical identification of the model of the flow in a continuous perfectly stirred tank reactor using PRBS indicate that in a comparison with the previously published results this method of estimation yields results of the same or better accuracy for the model in question while calling for shorter experimental time and fewer samples. For instance, the results of the simulation study by PRBS with the period $N = 31$ suggest that it can be worked in the interval of parameters $T_N/T_c \in (4{\cdot}5 - 37{\cdot}2)$ and $\Delta t/T_c \in (0{\cdot}10 - 1{\cdot}20)$. The recommended values for other statistical signals leading to a comparable accuracy ranged as follows: $T_N/T_c \in (40; 100)$ and $\Delta t/T_c \in (0{\cdot}04 - 0{\cdot}22)$.

It turns out that the most serious source of error for this method are the defects in the generated sequence of signals. The weight of this error is greater toward the end of the period in contrast to that in the initial part of the period.

In the presence of noise our result confirms the relationship derived by Briggs, according to which the accuracy of the method and the ratio of the meaningful information to noise is favourably affected by increased principal interval Δt as well as the length of the period N. On the other hand, the evaluation by the method of moments causes that an increase of these parameters is undesirable. Optimum choice of the parameters of PRBS should be made keeping both aspects in mind. The results of this simulation indicate that the duration of a single period $N\ \Delta t$ should not exceed the time the system takes to steady down: $T_s = (2-3)\ T_c$. At the same time, the evaluation of the moments in the presence of noise should utilize only that part of the impulse response corresponding to the time prior to steadying.

The method of PRBS is sensitive to monotonous trends. Although filtration of input data by the floating average method does eliminate the influence of these trends, at the same time it distorts the information contained in the cross correlation function. The more suitable is method of filtration developed by Nikiforuk, modified for a non-zero mean of PRBS, though increases the influence of noise. It has been successfully applied to eliminate the effect of an incorrectly adjusted initial condition.

The results of the simulation study were used to design experiments and methods of data evaluation for real systems.

It turns out that additional problems arising in connection with the simulation studies are associated with the effect of internal noise of the system and distortion of input signals in transducers and active elements.

LIST OF SYMBOLS

a	amplitude of PRBS
c_1	inlet concentration
c_2	outlet concentration
C	constants in recurrence formula (*2*)
E	operator of expected value
f	frequency
h	impulse characteristics
i	sequence index in PRBS
k	constant
M	constant of digital filter according to Eq. (*11*)
n	exponent in formula for the length of period PRBS
N	length of period of PRBS
p	Laplace variable
Q	volume flow rate in CSTR
$r(\tau)$	normalized correlation function
$R(\tau)$	correlation function
t	time
Δt	time interval of the forcing signal
T_c	time constant
T_N	time of realization of experiment
T_p	length of period of realization
T_s	time for steadying
V	volume of CSTR
η	variable in Eq. (*2*)
σ	standard deviation
τ	lag of correlation function
F	Fourier transform operator
—	mean value
$\sim$	experimental value

REFERENCES

1. Seinfeld J. H., Lapidus L.: *Process Modelling, Estimation and Identification.* Prentice-Hall 1974.
2. Václavek V.: Chem. Eng. Sci. *22*, 1209 (1967).
3. Angus M., Lapidus L.: AIChE J. *9*, 810 (1963).
4. Homan C. J., Tierney W. J.: Chem. Eng. Sci. *12*, 153 (1960).
5. Goodman N. R., Katz S., Ramer B. H., Kuo T.: Technometrics *3*, 245 (1961).
6. Briggs P. A. N., Hammond P. H., Hughes N. T. D., Plumb G. O.: Proc. Conf. on Advances in Aut. Contr. Paper 7, Nottingham 1965.
7. Havlíček A., Čermák J.: Presented at the Meeting of Czechoslovak Chemists, Brno, August 1974.
8. Pekárek O.: *Generátor* PSG-2C. Technický manuál. UTIA-Czechoslovak Academy of Sciences, Prague 1974.
9. Briggs P. A. N., Godfrey K. R., Hamond P. H.: Proc. 1st. IFAC Symp. Identification in Automatic Control Systems, Prague 1967.

10. Skáva L., Slavík J., Vavruška K.: *Malá výpočetní technika pro řešení dopravních a spojových úloh.* ALFA, Bratislava 1974.
11. Jenkins M. G., Watts G. D.: *Spectral Analysis and its Applications.* Holden-Day, San Francisco 1968.
12. Douce J. L., Ngk. C.: Proc. IFAC Symp. Self-Adaptive Control Systems 1965.
13. Douce J. L., Ngk. C.: Electr. Lett. *2*, 243 (1966).
14. Macleod C. J.: Int. J. Control. *14*, 97 (1971).
15. Nikiforuk P. N., Gupta M. M., Hofman R.: Automatica *7*, 249 (1971).

Translated by V. Staněk.