PERGAMON

Automatica 35 (1999) 1427–1436

www.elsevier.com/locate/automatica

Brief Paper

Real-time estimation of process frequency response and step response from relay feedback experiments☆

L. Wang[1], M.L. Desarmo[2], W.R. Cluett*

Department of Chemical Engineering, University of Toronto, Toronto, Ontario, Canada M5S 3E5

Received 28 July 1997; revised 28 September 1998; received in final form 10 February 1999

Abstract

This paper describes two new methods for obtaining frequency response and step response models for processes operating under relay feedback control. Both methods are based on the frequency sampling filter (FSF) model structure and a recursive least-squares (RLS) estimator. With the first method, a standard relay feedback experiment is performed and the FSF/RLS algorithm is used along with the generated input–output data to provide real-time estimates of the process frequency response at the dominant harmonics of the limit cycle. In the second method, a novel relay experiment is proposed that involves switching the error signal back-and-forth between a standard relay element and an integrator in series with a relay. The FSF/RLS algorithm is then applied to the generated process input–output data to obtain a real time estimate of the process step response. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Process identification; Recursive estimation; Frequency responses; Step function responses; Relays; Frequency sampling filters

1. Introduction

The relay feedback experiment was made popular in the field of process control by Åström and Hägglund (1984). This experiment was suggested by the authors as a means to automate the Ziegler–Nichols scheme for determining ultimate gain and frequency information about a process. Their approach followed directly from a describing function approximation (DFA) to the non-linear relay element with their objective being to use the obtained process information for automatic tuning of PID controllers.

This work by Åström and Hägglund (1984) has prompted research in several different directions. One of these directions, and the focus of this paper, is in the area of process identification, where the objective is to develop a more complete and accurate model of the process from data generated under relay feedback. This identified model can be used for either obtaining better tuning parameters for a PID-type regulator, designing other, more "advanced" controllers, or initializing an adaptive controller. Luyben (1987) suggested combining the relay-generated ultimate gain and frequency information with prior information on the process steady-state gain and delay information obtained from the initial part of the time response curves to estimate a continuous transfer function model. Later on, Luyben and co-workers (Li et al., 1991) proposed the use of information generated from a second relay experiment, obtained by adding hysteresis to the relay element or additional delay in the feedback loop, to remove the need for prior knowledge of the steady-state gain. Li et al. (1991) also made the important observation that the DFA approach to obtaining the ultimate process information can give rise to significant errors that depend on certain process characteristics such as the order of the system and the delay to time constant ratio. Schei (1994) used the idea of two different relay experiments to generate limit cycle oscillations at approximately the crossover and critical frequencies of the loop transfer function. The relay in Schei's work was placed around an existing closed-loop feedback system and the objective was to estimate a process

* Corresponding author. Tel.: +1-416-978-6697; fax: +1-416-978-8605; e-mail: cluett@ecf.utoronto.ca.

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor H. Hjalmarsson under the direction of Editor T. Söderström.

[1] Current address: Department of Electrical and Computer Engineering, University of Newcastle, Callaghan NSW 2308, Australia

[2] Current address: Honeywell Ltd., Industrial Automation and Control, Toronto, Ontario, Canada M2J 1S1.

PII: S0005-1098(99)00053-9

model from the relay-generated data. The fact that Schei chose to estimate a discrete ARX transfer function model raised other issues such as choice of sampling frequency and prefilter design. Lundh and Åström (1994) suggested using data generated from a relay experiment to estimate a continuous transfer function model with the main purpose being to use this model to initialize a self-tuning controller. These authors also pointed out that classical frequency response analysis techniques based on discrete Fourier transforms (DFT) may be used to accurately estimate the values of the transfer function at the first and third harmonics during the stable limit cycle portion of the relay experiment. Wang et al. (1997b) derived exact expressions for the periods and amplitudes of limit cycles under relay feedback for first order plus delay systems and then used this information with data generated from a single relay test to fit a first order plus delay model to a wide range of stable processes. Wang et al. (1997a) proposed the use of the relay test to obtain an estimate of the process frequency response at multiple frequencies. The authors introduced an exponential decay data window to permit use of both the transient and periodic parts of the relay feedback generated data. In a related paper, Bi et al. (1997) combined a modified relay experiment with the FFT algorithm to obtain accurate frequency response estimates at multiple frequencies around the process critical frequency.

In this paper, we make two contributions to this field of research. The first deals with the estimation of frequency response information from process input–output data generated under relay feedback, where we propose to solve this problem using a novel approach based on the frequency sampling filter (FSF) model (Wang and Cluett, 1997). Several authors, e.g. Friman and Waller (1995) and Sung et al. (1995), have examined the procedure followed by Åström and Hägglund (1984) and have proposed modified relay experiments to improve upon the accuracy of the ultimate gain and frequency estimates. It will be shown in this paper that the FSF model structure can be used in conjunction with the process input–output data generated from a standard relay experiment to obtain accurate frequency response estimates at the dominant harmonics of the limit cycle. A recursive implementation of the least-squares algorithm is suggested as a means to produce estimates in real time.

The second contribution is a recursive method for estimating a process step response model from input–output data generated using a novel relay experiment. Most of the cited literature dealing with process identification from relay feedback data requires the determination of the "correct" process model order from the generated frequency response information, e.g. Li et al. (1991), Lundh and Åström (1994), or the process model order is imposed in advance by the user, e.g. Schei (1994), Wang et al. (1997b). The method proposed in this paper is based on the FSF model structure and therefore avoids the need for prior knowledge or estimation of the model order or time delay. In addition, the method works very well in a fast sampling environment and there is no need for any data re-sampling or data prefilter design. The only requirement is an estimate of the process settling time and it will be shown here how even this information can be estimated directly from the relay feedback data. The proposed relay experiment is conducted by switching the error signal back-and-forth between a standard relay element and an integrator in series with a standard relay element. As a result, the input signal is no longer periodic but instead is typically rich in the frequency range needed for accurate step response model identification. The obtained step response model can then be used directly in a model-based controller design, or for subsequent fitting of a lower order transfer function model.

2. Recursive estimation of frequency response using standard relay experiment

2.1. Frequency sampling filter model

The frequency sampling filter (FSF) model of a stable, linear, time-invariant process has a z-transfer function given by (Rabiner and Gold, 1975)

$$G(z) = \sum_{k=-(N-1)/2}^{(N-1)/2} G(e^{j2\pi k/N}) \frac{1}{N} \frac{1-z^{-N}}{1-e^{j2\pi k/N} z^{-1}}, \tag{1}$$

where $G(e^{j2\pi k/N})$ represents the discrete frequency response of the process at $w_k = 2\pi k/N$ radians. The frequency response is related to the impulse response coefficients h_i, $i = 0, 1, \dots, N-1$, by

$$h_i = \frac{1}{N} \sum_{k=-(N-1)/2}^{(N-1)/2} G(e^{j2\pi k/N})\, e^{j2\pi ki/N}. \tag{2}$$

The FSF model is obtained from a linear transformation of the finite impulse response (FIR) model. The connection is that an FIR model of order N can be completely described by the value of its frequency response at N equally spaced frequencies around the unit circle. The N frequency sampling filters themselves are described by

$$H^k(z) = \frac{1}{N} \frac{1-z^{-N}}{1-e^{j2\pi k/N} z^{-1}} \tag{3}$$

and these filters have the following properties:

Property A. The magnitude of the kth filter's frequency response $|H^k(e^{jw})| = 1$ for $w = 2\pi k/N$, and $|H^k(e^{jw})| = 0$ for $w = 2\pi m/N$, where m is an integer like k in the range $-(N-1)/2$ to $(N-1)/2$ and $m \neq k$. The magnitude of the filter's frequency response forms a narrow band centered at $w = 2\pi k/N$ with a bandwidth of $4\pi/N$.

Property B. For a discrete sinusoidal input signal $u(t) = A\cos(2\pi kt/N) = A(e^{j2\pi kt/N} + e^{-j2\pi kt/N})/2$, where t is an integer variable used to enumerate the sampling instants, the discrete output responses of the N frequency sampling filters, after a transient period of N sampling instants, are given by

$$f(t)^k = H^k(z)\,u(t) = \frac{A}{2}e^{j2\pi kt/N}, \tag{4}$$

$$f(t)^{-k} = H^{-k}(z)\,u(t) = \frac{A}{2}e^{-j2\pi kt/N}, \tag{5}$$

$$f(t)^m = H^m(z)\,u(t) = 0 \tag{6}$$

for all $m \neq \pm k$.

2.2. Formulation of the identification problem

It is well known that many processes when placed under relay feedback control will oscillate at their critical frequency corresponding to the point of intersection of the process Nyquist curve with the negative real axis. If hysteresis is added to the relay element, the oscillation will occur at some lower frequency, and if a linear dynamic element is placed in series with the relay, the closed-loop system will oscillate at some frequency either above or below the critical frequency. For example, if the dynamic element is an integrator, the frequency of oscillation will correspond to the point of intersection of the process Nyquist curve with the negative imaginary axis. Suppose that the process to be identified is placed under relay feedback control and oscillates with some period T. Using a sampling interval of Δt, the number of samples within a period is $N' = T/\Delta t$. The periodic square wave $u(t)$ generated by the relay output can be completely described over this period $[0, T]$ by its discrete Fourier expansion (Godfrey, 1993)

$$u(t) = \sum_{k=-(N'-1)/2}^{(N'-1)/2} A_k e^{j2\pi kt/N'} \tag{7}$$

for $t = 0, 1, \ldots, N'-1$, where

$$A_k = \frac{1}{N'}\sum_{t=0}^{N'-1} u(t)\,e^{-j2\pi kt/N'} \tag{8}$$

for $k = 0, \pm 1, \ldots, \pm(N'-1)/2$. If the input signal is also symmetric and the time origin is taken at one of the relay switches, $A_k = 0$ for $k = 0, \pm 2, \pm 4, \ldots$. The magnitudes of the nonzero values of A_k decrease with increasing $|k|$.

For this problem, we will set the parameter N in the FSF model to be equal to N', and describe the process output $y(t)$ by

$$y(t) = G(z)\,u(t) + v(t), \tag{9}$$

where $G(z)$ is defined by Eq. (1) and $v(t)$ is the disturbance term. We will now define the parameter vector to be estimated as

$$\theta = [G(e^{j0}) \quad G(e^{j2\pi/N}) \quad G(e^{-j2\pi/N}) \ \ldots\ G(e^{j(n-1)\pi/N}) \quad G(e^{-j(n-1)\pi/N})]^{T}$$

and its corresponding regressor vector as

$$\phi(t) = [f(t)^0 \quad f(t)^1 \quad f(t)^{-1} \ldots f(t)^{(n-1)/2} \quad f(t)^{-(n-1)/2}]^{T},$$

where

$$f(t)^r = \frac{1}{N}\frac{1 - z^{-N}}{1 - e^{j2\pi r/N}z^{-1}}\,u(t) \tag{10}$$

for $r = 0, \pm 1, \ldots, \pm(n-1)/2$. Therefore, the total number of frequencies included in this process output description is $(n-1)/2 + 1$. (The above process description consists of both complex parameters and complex regressors, except for the terms corresponding to $r = 0$. Bitmead and Anderson (1981) have shown that, by combining the complex conjugate pairs of terms in both the parameter and regressor vectors, the above FSF model description can be converted to a model representation with n real parameters. However, Wang and Cluett (1997) have shown that the model structure containing the complex terms provides more direct insight into the frequency content of the input signal.)

If $u(t)$ is a periodic and symmetric signal, then from Eq. (7) and Property B, the filter outputs for even values of r ($r = 0, \pm 2, \pm 4, \ldots$) are equal to zero after one complete period N. In addition, the magnitudes of the nonzero filter outputs corresponding to $r = \pm 1, \pm 3, \ldots$ decrease as $|r|$ increases. Therefore, in this situation, the only terms required to accurately describe the process output $y(t)$ in Eq. (9) for processes with a monotonically decreasing frequency response may be those with $r = \pm 1, \pm 3$ and ± 5. However, because output disturbances and measurement noise are encountered in most practical situations, $u(t)$ is seldom an ideal periodic and symmetric signal. In many cases though, $u(t)$ would be nearly periodic and the parameter N could be chosen based on an estimate of the average period. In order to avoid biased estimates of the frequency response parameters in such circumstances, we suggest that all of the terms corresponding to $r = 0, \pm 1, \ldots, \pm 5$ be included in the process description.

Given the process input–output data generated from the relay experiment, the parameter vector θ can be estimated in real time using a recursive algorithm. Here, we propose to use the recursive least squares algorithm (Goodwin and Sin, 1984) given by

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t-1)\phi(t)(y(t) - \phi(t)^{T}\hat{\theta}(t-1)) \tag{11}$$

and

$$P(t-1) = P(t-2) - \frac{P^*(t-2)\phi(t)\phi^*(t)P(t-2)}{1 + \phi^*(t)P(t-2)\phi(t)}, \tag{12}$$

where * denotes the complex-conjugate transpose. We also suggest using the process input–output data from the first complete period along with a standard batch least squares estimator to initialize $\hat{\theta}(0)$ and $P(-1)$.

Bitmead and Anderson (1981) proposed the use of a recursive least mean squares (LMS) algorithm to estimate frequency response coefficients using the FSF model structure. The authors treated the problem as a collection of independent, or decoupled, estimation problems by assuming that the outputs of the frequency sampling filters satisfy an approximate orthogonality property. This assumption allowed the authors to reduce the larger single estimation problem (Nth order) to a collection of several smaller (one or two parameter) estimation problems, with their objective being to avoid the dimensionality and ill-conditioning problems known to be associated with the larger problem. Our choice of the more rapidly converging recursive least-squares approach to solve the larger single estimation problem is justified by the fact that we only need to estimate n parameters. Because we can choose $n \ll N$, the dimensionality and ill-conditioning problems can be avoided. Also, the FSF outputs are only truly orthogonal if the input is either white noise or periodic (Goberdhansingh et al., 1992). Therefore, in practice, it is desirable to solve the larger single estimation problem to avoid biased parameter estimates.

2.3. Simulation example

The following transfer function is representative of the dynamics typically associated with paper machine basis weight (EnTech, 1993)

$$G(s) = \frac{e^{-100s}}{45s+1}. \tag{13}$$

The objective is to recursively estimate the process frequency response at the dominant harmonic frequencies generated under a standard relay experiment. A white noise disturbance sequence with unit variance has been added to the output. The relay amplitude has been set equal to 2 and the hysteresis level has been set equal to 3 (3 times the standard deviation of the output noise). The process is sampled with a time interval of 0.67 s. Note that, although the hysteresis level is larger than the noise-free process output response to an input change of magnitude 2, a limit cycle still occurs due to the presence of the noise.

Fig. 1 shows the process input-output data generated under a standard relay experiment. An estimate of $N'/2$ was taken as the number of samples between the second and third switches of the relay. Fig. 2 illustrates how the real and imaginary parts of the parameter estimates behave for the first harmonic ($r = 1$) and the third harmonic ($r = 3$) over the duration of the relay experiment. The estimates of these parameters up until the third switch in the relay output ($\approx$300 s) are constant and equal to $\hat{\theta}(0)$. Fig. 3 compares the two estimated frequency responses after 500 s with the true process frequency response. These latter two figures show that the parameter estimates have effectively converged after 500 s which compares very favorably with the total process settling time of approximately 325 s.

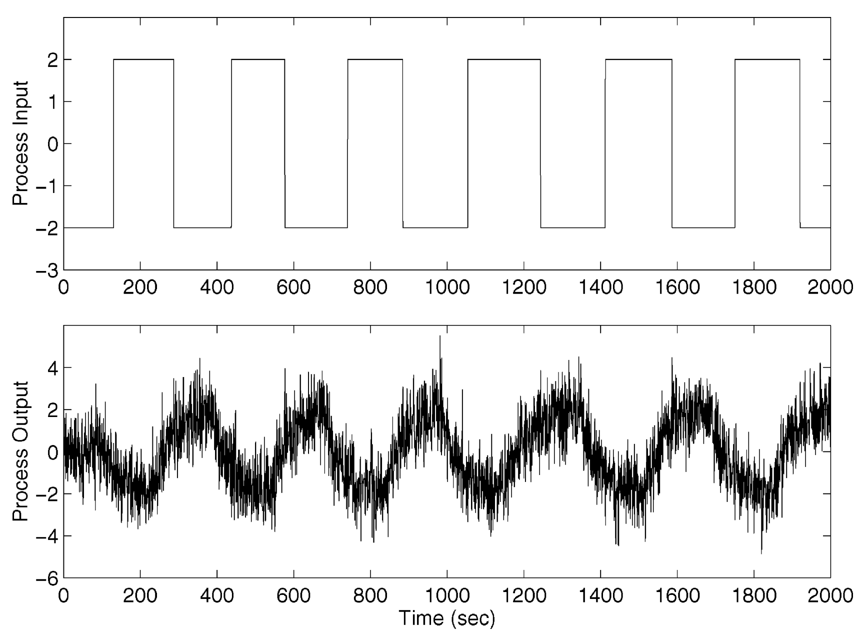

Fig. 1. Process input–output data for paper machine basis weight example.

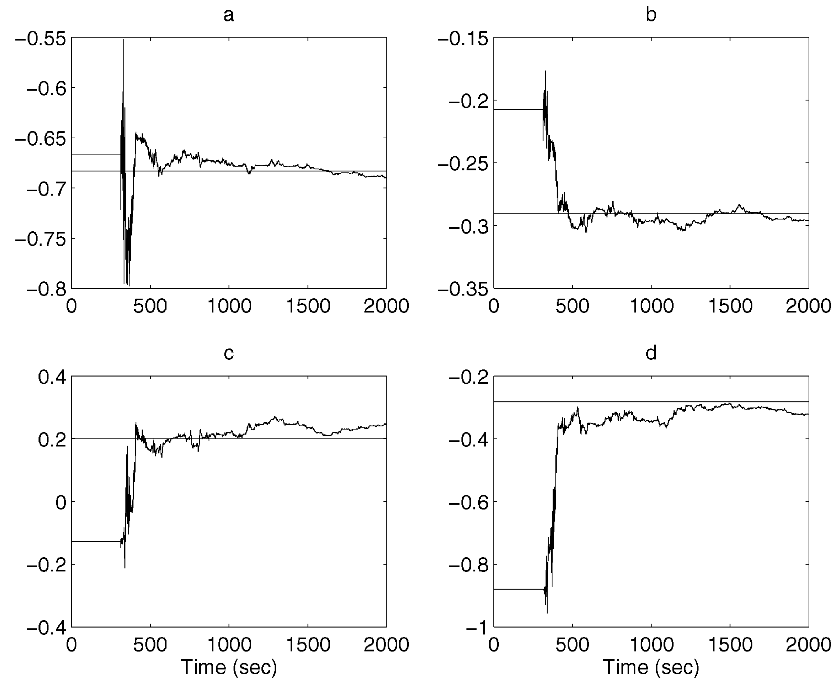

Fig. 2. Real (a,c) and imaginary (b,d) parts of the frequency response estimates corresponding to the first and third harmonics, respectively.

The example used here is a first order plus delay process with very little roll-off at higher frequencies. For processes with more roll-off at higher frequencies, it may not be realistic to expect accurate identification at frequencies beyond the first harmonic, depending on the signal-to-noise ratio at the higher harmonics.

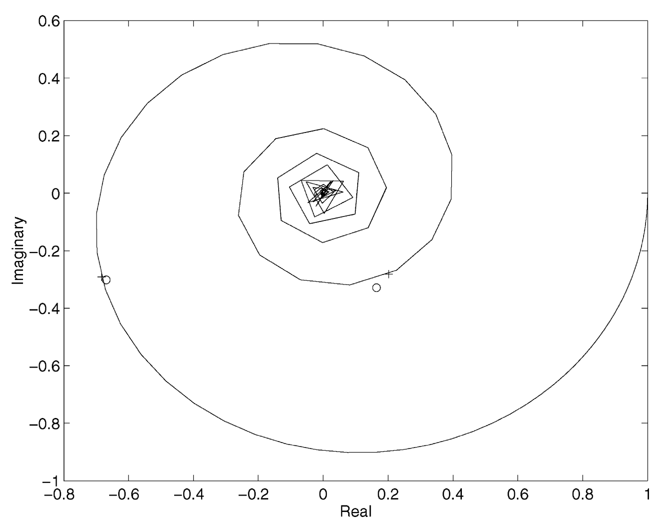

Fig. 3. Comparison of true ('+' with solid line) frequency responses with estimated values ('○') after 500 s.

3. Recursive estimation of step response using novel relay experiment

3.1. Method and apparatus

Wang and Cluett (1997) have proposed the use of the FSF model structure for the identification of process step response models. When Eq. (9) is used to estimate a step response model, the parameter N must be chosen according to the process settling time T_s, i.e. $N \approx T_s/\Delta t$. The step response coefficients g_m, $m = 0, 1, \ldots, N-1$ are obtained from the FSF model parameters using the following weighted sum relationship:

$$g_m = \sum_{k=-(n-1)/2}^{(n-1)/2} G(e^{j2\pi k/N}) \frac{1}{N} \frac{1 - e^{j(2\pi k/N)(m+1)}}{1 - e^{j2\pi k/N}}, \qquad (14)$$

where $(n-1)/2 + 1$ is the number of frequencies in the FSF model that we choose to estimate for the construction of the step response model. As pointed out in Wang and Cluett (1997), the FSF model in this case is just a linear reparameterization of an Nth-order FIR model when $n = N$.

Calculating a step response model from the FSF parameter estimates has several advantages over direct estimation of the step response or impulse response coefficients. First, the FSF parameters are well-behaved with respect to the sampling interval in the sense that, as the sampling interval $\Delta t \to 0$, the parameters in the FSF model converge to the continuous-time process frequency responses at $w = 0, 2\pi/T_s, 4\pi/T_s, \ldots, (n-1)\pi/T_s$ rad/s. Therefore, the number of significant parameters in the FSF model is independent of the choice of sampling interval. Second, because the frequency response of processes with strictly proper transfer functions, which covers most physical systems, approaches zero at high frequencies, the number of significant FSF model parameters (n) is small relative to N. These characteristics of the FSF model, combined with the fact that the step response coefficients in Eq. (14) are obtained from a weighted summation of the FSF model parameters that emphasizes primarily the parameters in the lower frequency region, make for an effective approach to estimating step response models.

As discussed in the previous section, a standard relay experiment produces in most cases a limit cycle dominated by a single frequency. However, this information is not sufficient for the estimation of an accurate process step response model. This raises the issue of how to generate the appropriate information. One approach is to inject a 'dither' signal while the process is under some sort of feedback control, either additively to the controller output or via the setpoint. However, this requires design of the dither signal, i.e. decisions must be made concerning its power spectrum. Another alternative is to make use of multiple relay experiments to generate frequency response information at several frequencies. Our objective was to develop a single relay experiment that automatically generates the desired information.

The proposed apparatus combines in parallel a relay element with an integrator in series with a relay element. Fig. 4 provides a block diagram of this apparatus. The experiment is performed by alternatively switching the error signal between the relay path and the integrator-relay path. The design of the experiment then reduces to the selection of this switching sequence. The input signal generated from this type of relay experiment will no longer be dominated by a single frequency but will instead contain frequencies over a range corresponding to the process phase shift of $-\pi/2$ and $-\pi$ in the noise-free case, and over a slightly lower frequency range when hysteresis is added to the relay elements.

The input–output data generated from this relay experiment can then be used along with Eqs. (9), (14), (11) and (12) to estimate in real time the process step response, where N is chosen according to the process settling time.

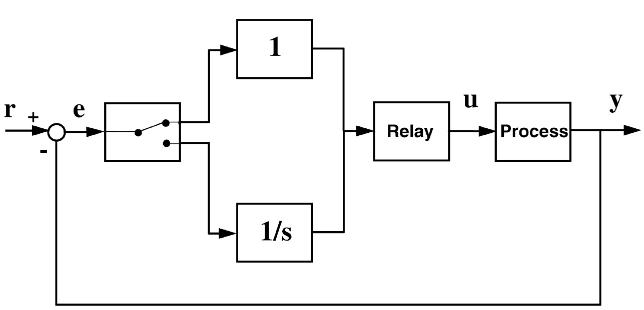

Fig. 4. Block diagram for proposed relay apparatus.

If prior information on the approximate settling time is available, then this may be used here to preselect a value for N. However, it would be desirable to find a simple way to estimate this parameter on-line during the relay experiment. When an integrator is placed in series with a relay with zero hysteresis, the closed-loop system will oscillate at a frequency corresponding to the point of intersection of the process Nyquist curve with the negative imaginary axis. From a large number of simulation studies, we have found that, for many processes, the frequency response at $w = 2\pi/N$ rad (or $w = 2\pi/T_s$ rad/s) is located in the vicinity of this point of intersection. Therefore, if the relay experiment begins with the error following the integrator-relay path for at least four successive switches in the relay output, an estimate of $N/2$ can be taken as the number of samples between the third and fourth switches of the relay. In addition, $\hat{\theta}(0)$ and $P(-1)$ can be estimated at the same time from the previous N sets of input–output data.

The only remaining parameter to be determined is n, the number of FSF model parameters to be estimated in Eq. (9). Using a batch, generalized least-squares algorithm, Patel et al. (1997) applied the PRESS statistic (Wang and Cluett, 1996) for selecting n. For the recursive approach proposed here, we suggest fixing the number of parameters to be estimated by preselecting $n = 11$ which we have found to be sufficient for a wide range of processes.

3.2. Simulation case studies

In this section, we will study the performance of the proposed method and apparatus by performing a set of five case studies, with four of these cases selected from the test batch presented in Åström and Hägglund (1995).

Process A:

$$G(s) = \frac{e^{-s}}{(s+1)^2}. \tag{15}$$

Process B:

$$G(s) = \frac{1}{(s+1)^8}. \tag{16}$$

Process C:

$$G(s) = \frac{1}{(1+s)(1+0.5s)(1+0.5^2 s)(1+0.5^3 s)}. \tag{17}$$

Process D:

$$G(s) = \frac{1-0.2s}{(s+1)^3}. \tag{18}$$

Process E:

$$G(s) = \frac{e^{-s}}{s^2 + 2\times 0.45s + 1}. \tag{19}$$

Process E is not found in Åström and Hägglund's test batch but we have added it to provide a case with underdamped dynamics.

The quality of the estimated model will be measured by the amount of departure from the true impulse and step response coefficients. More precisely, we will use the sum of the squared deviations from the true impulse response coefficients (E_{impulse}) and the sum of squared deviations from the true step response coefficients (E_{step}) to quantify the closeness of the model fit, i.e.

$$E_{\text{impulse}} = \sum_{i=0}^{N-1} (h_i - \hat{h}_i)^2, \tag{20}$$

$$E_{\text{step}} = \sum_{i=0}^{N-1} (g_i - \hat{g}_i)^2. \tag{21}$$

The impulse response error, E_{impulse}, measures the closeness of the estimated process dynamics to the true process dynamics with equal weighting at all frequencies, while the step response error, E_{step}, measures the accuracy of the estimated model with more emphasis in the low-frequency range (Dayal and MacGregor, 1996).

For the simulations, a white noise disturbance sequence with variance equal to 0.8^2 has been added to the process output. The relay amplitude has been set equal to 2 and the hysteresis level has been initially set equal to 0.08. The relay experiment is started with the error following the integrator-relay path. Because of the averaging effect of integration, a small hysteresis level can be tolerated and is used here initially to obtain an estimate of the frequency corresponding to the point of intersection of the process Nyquist curve with the negative imaginary axis, and in turn an estimate of N. The experiment is allowed to proceed until four switches in the relay output have occurred at which time the value of N is estimated. The hysteresis level is then fixed to a value of 2.5 (approximately 3 times the standard deviation of the output noise) and the error is switched to the relay path for one complete period (two switches). Then, the error is switched back-and-forth between the integrator-relay and relay paths, after 2–3 switches in the relay output have occurred along a given path. Total simulation time for each case is approximately four times the process settling time. The sampling rate is adjusted for each case so that 3000 sets of input–output data are collected within the total simulation time.

To illustrate the results, the estimated FSF model parameters have been converted into step response and impulse response coefficients at integer multiples of the process settling time. The values of E_{impulse} and E_{step} are summarized in Table 1, along with the estimated T_s values obtained from the relay experiment. These results show that the estimated models are all converging as the length of the relay experiment increases.

We have selected Process A for further study. Fig. 5 presents the process input–output data collected from the new relay experiment. The estimated step response models at different points in time over the course of the experiment are presented in Fig. 6 along with the true step response. This figure clearly shows that the estimated step response model converges to the true step response. In fact, the estimated step response model changes very little after only two settling times of data. Fig. 7 confirms that the estimated FSF model parameters after two settling times of data are very accurate in the low and medium frequency regions ($r = 0, \pm 1, \pm 2$) and the deviations in the higher frequency region ($r = \pm 3, \pm 4, \pm 5$) are modest but do not have a significant effect on the step response model accuracy.

With the FSF approach, the diagonal elements of the correlation matrix associated with the least-squares parameter estimates are proportional to the periodogram of the input signal in the vicinity of the FSF frequencies, and to the data length. Therefore, the accuracy of a particular FSF parameter is directly related to the magnitude of the corresponding diagonal element in the correlation matrix (Wang and Cluett, 1997). To illustrate the energy distribution of the input signal generated using the new relay experiment for Process A, Fig. 8 shows the magnitude plots of the diagonal elements of the correlation matrix corresponding to the FSF model with $n = 11$. After one settling time of data, the energy of the input signal is focused mainly at the first pair of frequencies ($r = \pm 1$). However, after just two settling times, energy is clearly present at both the zero frequency ($r = 0$) and at the second pair of frequencies ($r = \pm 2$), and to a lesser degree at the higher frequencies ($r = \pm 3, \pm 4, \pm 5$). This figures shows how the data generated from the new relay experiment with this example leads to accurate estimates of the frequency response in the low and medium frequencies, including the steady-state gain,

Table 1
Case study results

Process	After T_s	After $2T_s$	After $3T_s$	After $4T_s$
Process A ($T_s = 11$ s)				
E_{impulse}	9.11×10^{-4}	1.27×10^{-4}	8.66×10^{-5}	1.04×10^{-4}
E_{step}	26.83	0.66	0.43	0.11
Process B ($T_s = 31$ s)				
E_{impulse}	8.76×10^{-4}	2.62×10^{-5}	3.45×10^{-5}	3.28×10^{-5}
E_{step}	4.45	3.47	0.68	0.18
Process C ($T_s = 6.3$ s)				
E_{impulse}	3.57×10^{-4}	8.69×10^{-5}	4.66×10^{-5}	3.25×10^{-5}
E_{step}	2.35	0.69	0.43	0.15
Process D ($T_s = 19$ s)				
E_{impulse}	1.5×10^{-2}	9.5×10^{-4}	8.49×10^{-4}	8.73×10^{-4}
E_{step}	19.5	0.32	0.24	0.31
Process E ($T_s = 9.3$ s)				
E_{impulse}	4.31×10^{-4}	8.18×10^{-5}	4.88×10^{-5}	5.08×10^{-5}
E_{step}	0.8	0.32	0.25	0.09

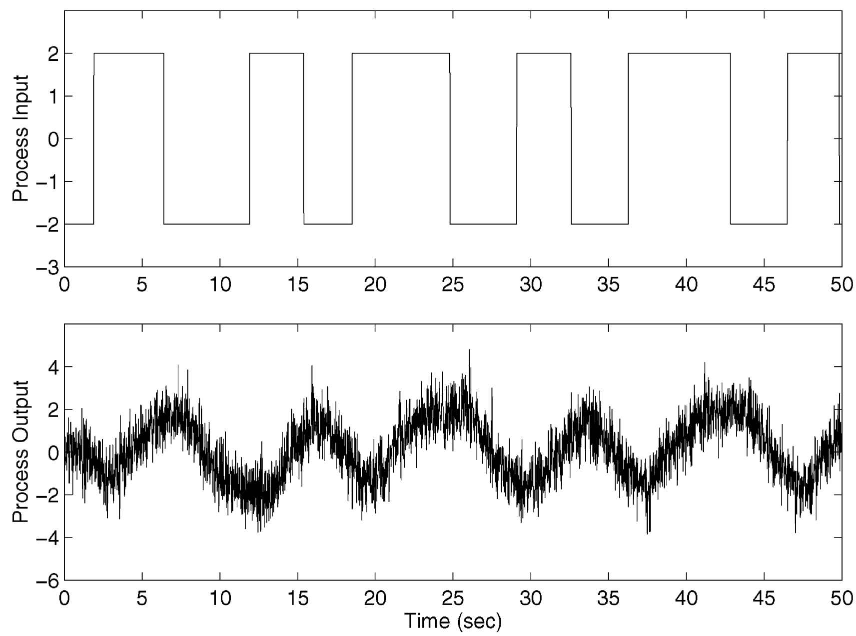

Fig. 5. Process input–output data for Process A.

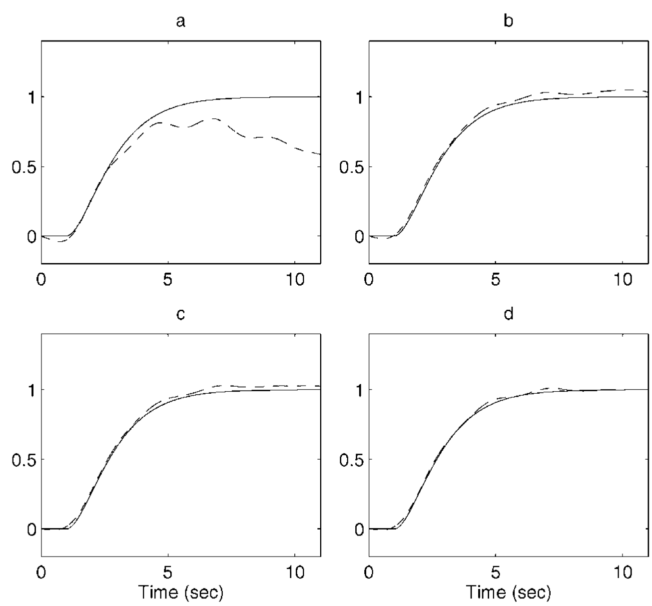

Fig. 6. Step response for Process A. Solid: true response; dashed: estimated response using FSF model after (a) T_s, (b) $2T_s$, (c) $3T_s$, and (d) $4T_s$.

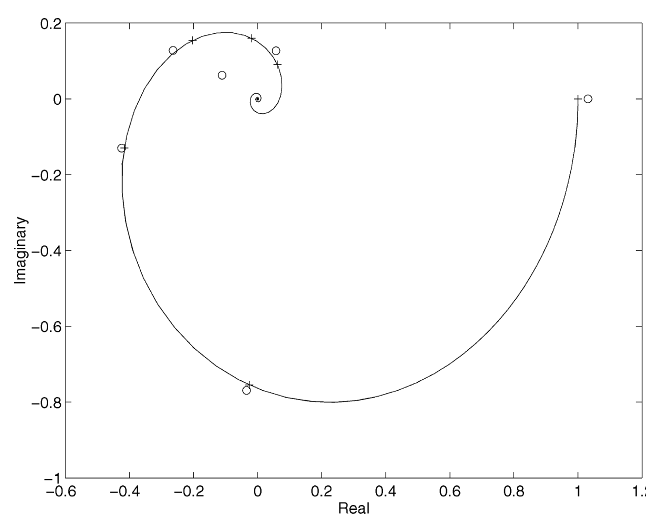

Fig. 7. Frequency response for Process A. Solid line: continuous-time frequency response; $+$: true FSF parameters; ○: estimated FSF parameters after $2T_s$.

and in turn an accurate estimate of the step response via the FSF model.

3.3. Automated design of an identification experiment

The proposed relay device and FSF algorithm could readily be bundled together as a stand-alone apparatus for real time process frequency/step response identification and subsequent controller design and tuning (see, for example, Hägglund and Åström, 1985). However, the proposed relay experiment on its own provides some interesting and new ideas about how to design input signals for process identification. One of the main benefits of the methodology described earlier in this section is that the design of an identification experiment suitable for obtaining an accurate step response model for stable processes has now been automated. A standard approach to the design of an identification experiment involves first the off-line design of the input signal followed by the

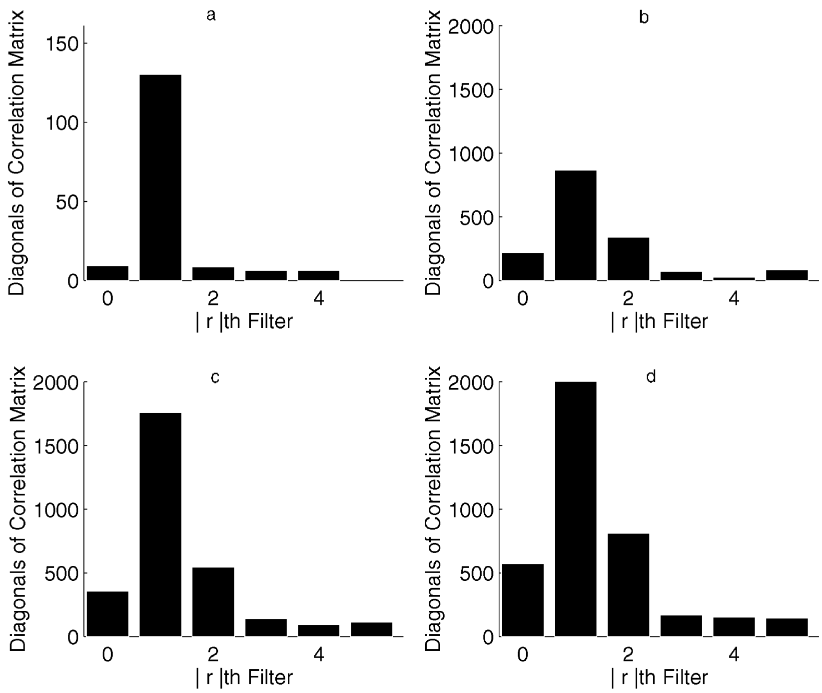

Fig. 8. Diagonal elements of correlation matrix for Process A after (a) T_s, (b) $2T_s$, (c) $3T_s$, and (d) $4T_s$.

identification experiment itself. This off-line design typically requires prior information on the dominant process dynamics and/or the desired power spectrum of the input signal. Under the proposed relay experiment, the relevant frequency response information is automatically generated without the requirement of any prior process information except the sign of the gain. Here, a general input signal design algorithm suitable for process identification is proposed which does not require an actual relay device for implementation.

Step 1: With the process initially at or near steady state, estimate the output noise level by calculating its standard deviation (σ) and set the hysteresis width $\varepsilon = 3\sigma$. Preselect the amplitude (a) of the input signal where the input signal will switch $\pm a$ around a nominal value $\bar{u}$. Designate the corresponding nominal value of the process output as the reference output value r.

Step 2: Set the input signal $u(t) = \bar{u} + a$ and start calculating the integrated error signal according to $e_I(t) = e(t)\Delta t + e_I(t-1)$, with $e(t) = r - y(t)$ and the initial value of $e_I = 0$.

Step 3: At each sampling instant, calculate the current value of the input signal $u(t)$ according to:

If

$$|e_I(t)| \geq \varepsilon \tag{22}$$

then (for positive process gain)

$$u(t) = \bar{u} + a \times sign(e_I(t)) \tag{23}$$

then (for negative process gain)

$$u(t) = \bar{u} - a \times sign(e_I(t)) \tag{24}$$

else

$$u(t) = u(t-1). \tag{25}$$

Step 4: Repeat Step 3 until three switches in the process input have occurred.

Step 5: After the final switch based on Step 3, calculate the current value of the input signal $u(t)$ according to:

If

$$|e(t)| \geq \varepsilon \tag{26}$$

then (for positive process gain)

$$u(t) = \bar{u} + a \times sign(e(t)) \tag{27}$$

then (for negative process gain)

$$u(t) = \bar{u} - a \times sign(e(t)) \tag{28}$$

else

$$u(t) = u(t-1) \tag{29}$$

Step 6: Repeat Step 5 until two switches in the process input have occurred.

Step 7: Alternate between Steps 3 and 5 after 2–3 switches in the process input have occurred within a given step. At the beginning of Step 3, always set the initial value of $e_I = 0$.

4. Conclusions

This paper has introduced two approaches, based on the frequency sampling filter (FSF) model and a recursive least-squares algorithm, for real-time estimation of process frequency response and step response models from relay feedback experiments. The approach to frequency response estimation uses data generated from a standard relay experiment. The method for step response model estimation involves a new relay experiment, which on its own provides some interesting and new ideas about how to design input signals for process identification experiments without prior knowledge of the process dynamics.

Acknowledgements

The authors would like to acknowledge financial support from the Government of Canada's Mechanical Wood-Pulps Network of Centres of Excellence.

References

Åström, K. J., & Hägglund, T. (1984). Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica*, *20*, 645–651.

Åström, K. J., & Hägglund, T. (1995). *PID controllers: theory, design, and tuning*. Research Triangle Park, NC: Instrument Society of America.

Bi, Q., Wang, Q.-G., & Hang, C.-C. (1997). Relay-based estimation of multiple points on process frequency response. *Automatica*, *33*, 1753–1757.

Bitmead, R. R., & Anderson, B. D. O. (1981). Adaptive frequency sampling filters. *IEEE Transactions on Circuits and Systems*, *28*, 524–533.

Dayal, B. S., & MacGregor, J. F. (1996). Identification of finite impulse response models: methods and robustness issues. *Industrial and Engineering Chemistry Research*, *35*, 4078–4090.

EnTech (1993). *Automatic controller dynamic specification*. Toronto, EnTech Control Engineering Inc.

Friman, M., & Waller, K. V. (1995). Closed-loop identification by use of single-valued nonlinearities. *Industrial and Engineering Chemistry Research*, *34*, 3052–3058.

Goberdhansingh, E., Wang, L., & Cluett, W. R. (1992). Robust frequency domain identification. *Chemical Engineering Science*, *47*, 1989–1999.

Godfrey, K. (1993). *Perturbation signals for system identification*. Englewood Cliffs, NJ: Prentice-Hall.

Goodwin, G. C., & Sin, K. S. (1984). *Adaptive filtering prediction and control*. Englewood Cliffs, NJ: Prentice-Hall.

Hägglund, T., & Åström, K. J. (1985). Method and an apparatus in tuning a PID-regulator. United States Patent, 4549123.

Li, W., Eskinat, E., & Luyben, W. L. (1991). An improved autotune identification method. *Industrial and Engineering Chemistry Research*, *30*, 1530–1541.

Lundh, M., & Åström, K. J. (1994). Automatic initialization of a robust self-tuning controller. *Automatica*, *30*, 1649–1662.

Luyben, W. L. (1987). Derivation of transfer functions for highly nonlinear distillation columns. *Industrial and Engineering Chemistry Research*, *26*, 2490–2495.

Patel, U. J., Wang, L., Foley, M., & Cluett, W. R. (1997). *Application of the frequency sampling filter model in process identification: an industrial case study*. (pp. 25–30) Preprints IFAC ADCHEM, Banff, Canada.

Rabiner, L. R., & Gold, B. (1975). *Theory and application of digital signal processing*. Englewood Cliffs, NJ: Prentice-Hall.

Schei, T. S. (1994). Automatic tuning of PID controllers based on transfer function estimation. *Automatica*, *30*, 1983–1989.

Sung, S. W., Park, J. H., & Lee, I.-B. (1995). Modified relay feedback method. *Industrial and Engineering Chemistry Research*, *34*, 4133–4135.

Wang, L., & Cluett, W. R. (1996). Use of PRESS residuals in dynamic system identification. *Automatica*, *32*, 781–784.

Wang, L., & Cluett, W. R. (1997). Frequency-sampling filters: an improved model structure for step-response identification. *Automatica*, *33*, 939–944.

Wang, Q.-G., Hang, C.-C., & Bi, Q. (1997a). Process frequency response estimation from relay feedback. *Control Engineering Practice*, *5*, 1293–1302.

Wang, Q.-G., Hang, C.-C., & Zou, B. (1997b). Low-order modeling from relay feedback. *Industrial and Engineering Chemistry Research*, *36*, 375–381.

Liuping Wang obtained a B.Sc. degree in Electrical Engineering from Chongqing University, China in 1982 and a Ph.D. degree in Control Engineering from the University of Sheffield, England in 1989. During her Ph.D. program, she studied in the Department of Mathematics at the University of Strathclyde, Scotland from 1985 to 1988. She also worked as a research assistant in the Department of Mechanical Engineering at the University of Strathclyde after completing her Ph.D. From the end of 1989 to the beginning of 1998, she worked in the Department of Chemical Engineering, University of Toronto, Canada, where she taught both undergraduate and graduate courses and co-supervised postgraduate students. She is currently a Senior Lecturer in the Centre for Integrated Dynamics and Control, Department of Electrical and Computer Engineering, University of Newcastle, Australia. Her current research interests include robust control, system identification and control applications to process industry.

Michelle Desarmo received her B.A.Sc. degree in Engineering Science from the University of Toronto, Canada in 1995. Following this, she received an M.A.Sc. degree in 1998 from the Department of Chemical Engineering at the University of Toronto, where her research focused on the area of process identification. She is currently working in the area of alarm management for process control at Honeywell Industrial Automation & Control.

William R. Cluett was born in London, Canada in 1959. He received a B.Sc. degree in Chemical Engineering from Queen's University, Canada in 1981 and a Ph.D. degree in Chemical Engineering from the University of Alberta, Canada in 1986. Since 1986, he has been with the Department of Chemical Engineering and Applied Chemistry at the University of Toronto where he is currently a Professor and is also serving as the Vice-Dean of the Faculty of Applied Science and Engineering. In 1997, he received the Syncrude Canada Innovation Award of the Canadian Society for Chemical Engineering and in 1998 he was elected as a Fellow of The Chemical Institute of Canada.