Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Process Control 15 (2005) 729–740
JOURNAL OF PROCESS CONTROL
www.elsevier.com/locate/jprocont

Closed-loop identification with a quantizer

M. Wang [a], N.F. Thornhill [b,*], B. Huang [c]

[a] Centre for Process Systems Engineering, Imperial College London, Prince Consort Road, London SW7 2BY
[b] Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE
[c] Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 2G6

Received 15 September 2003; received in revised form 7 June 2004; accepted 4 October 2004

Abstract

This paper proposes a new closed-loop identification scheme for a single-input single-output (SISO) control loop based upon a quantizer inserted into the feedback path. The quantizer can be used to generate an equivalent persistently exciting signal to which the well known two-stage method of closed-loop identification can be applied. The paper examines the performance and behaviour of the quantizer-based closed-loop identification and gives suggestions for the choice of quantizer interval. Simulation and experimental examples are used to illustrate the proposed new CLID scheme.

Keywords: Chemical process; Closed-loop identification; Frequency response; Quantization; Persistently exciting; PID control; Power spectrum; Process control; Process monitoring

1. Introduction

A process model is useful for tuning of classical PID controllers and necessary for model-based process control. The purpose of closed-loop identification (CLID) is to identify a process model while the process is still under feedback control [1]. Closed-loop identification might be necessary because the system is unstable in open loop or the system contains inherent feedback mechanisms [2], but in most cases CLID is preferred for economic reasons since keeping the controller in the loop during system identification means less disruption to the normal manufacturing process. The advantage of CLID is that the process will not drift from the nominal operating point.

The motivation of this paper is to find a novel way to do CLID by inserting a quantizer in the feedback path. The quantizer is used to generate a persistently exciting signal so that the two-stage method by Van den Hof and Schrama [3] can be used directly for identification.

A benefit of the proposed CLID scheme is that there is no requirement for external excitation, an activity that requires design and generation of a persistently exciting signal or of harmonic signals at selected frequencies. Rather, the quantizer interval is adjusted on the basis of routine measurements of the signals within the closed loop so that the disturbance caused by quantization becomes persistently exciting. The proposed method requires the setting of just one parameter, h, the quantizer interval. The quantizer can be implemented in software and when identification is complete the quantizer interval is made negligibly small and the controller stays in place during the identification. The method determines the model order and identifies the complete frequency response.

The paper presents the techniques of CLID with a quantizer, gives recommendations for the choice of quantizer interval and shows results from simulation and experiment that demonstrate the method. It is laid out as follows. Section 2 reviews the background of

* Corresponding author. Tel.: +44 (0)20 7679 3983; fax: +44 (0)20 7388 9325.
E-mail address: n.thornhill@ee.ucl.ac.uk (N.F. Thornhill).

0959-1524/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2004.10.007

the work and other relevant approaches. An analysis of the quantization error is given in Section 3 to provide a basis for closed-loop identification with a quantizer. Section 4 presents the procedures of CLID with a quantizer. Simulation and experimental results are presented in Sections 5 and 6 where a comparison is also made with an alternative published method devised by Welsh and Goodwin [4]. The paper ends with a conclusion section.

2. Background

2.1. CLID with external excitation

Achieving identification during normal operation without any external excitation or disruption would be an ideal target. Theoretical studies by Gustavsson et al. [5], MacGregor and Fogal [1] and an industrial case study by Bartee and McFarlane [6] showed, however, that the identification of process models using routine operating data (or archived data) has very limited potential because in CLID some measure has to be adopted to break the dependency between the process input signal and the process disturbance [2]. The conclusion from those studies was that external excitation is required. The external excitation should be persistently exciting across the frequency range of interest [7], and must be independent of the process disturbance.

External excitation is a dither signal injected into the original closed-loop system. It could be introduced from the set-point, the controller output or the feedback path as shown in Fig. 1. A single-input single-output (SISO) closed-loop system adopted from [1] illustrates the concept. In Fig. 1, $G_o(q)$ represents the true process, and the disturbance $v(t) = H_o(q)a(t)$ represents the effect of all unmeasured process disturbances on the measured output $y(t)$, q is the forward shift operator. $C(q)$ is the feedback controller. The set-point $y_{sp}(t)$ and the dither signal $d(t)$ are input signals that may be injected to aid the identification. In this paper, only $y_{sp}(t) = 0$ is considered, that is to say $d(t)$ is the only input signal and it is injected in the feedback path. The identification task is to use $y(t)$, $u(t)$ and $d(t)$ to identify the process $G_o(q)$.

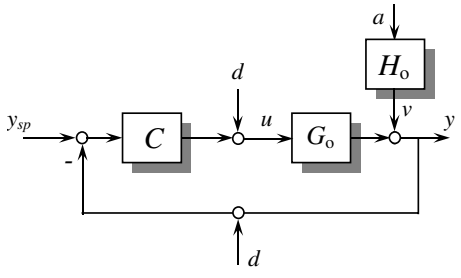

Fig. 1. Closed loop system with external excitation. Excitation may be inserted in the controller output, the feedback path or added to the set point. For CLID with a quantizer, it is inserted in the feedback path.

CLID approaches have been classified as: Direct Approach, Indirect Approach and Joint Input–Output Approach [5,8,9]. In the joint input–output approach as described in [8] the input $u(t)$ and the output $y(t)$ are jointly viewed as the output from a system driven by the external excitation signal $d(t)$ and noise $a(t)$ and the open-loop parameters are estimated from this augmented system. Three branches exist within the joint input–output approach. They are the coprime factor identification scheme, the two-stage method [3] or the two-step method [10] and the projection method [8]. Esmaili et al. [11] discussed the asymptotic and finite data behavior of some closed-loop identification methods while Gevers et al. [12] derived the asymptotic variance expressions for models that are identified on the basis of closed-loop data, showing that direct and indirect methods lead to the same asymptotic variance.

2.2. CLID with harmonic excitation

Fox and Godfrey [13] proposed a nonparametric closed-loop identification method with multi-harmonic perturbation excitation. The excitation signal must be carefully designed to excite the system response at the frequencies of interest which are often those around the system critical frequency. A phase-locked loop framework has also been used as a means for extracting a frequency response [14,15].

2.3. Use of relay

The classical relay identification method of Åström and Hägglund [16] determined gain margin and the cross-over frequency ω_{cg}. Many extensions have been reviewed in [17]. Wang et al. [18] introduced an exponential decay into the standard relay output and the process output so that with one relay test multiple points of the frequency response can be obtained by Fourier transform. De Arruda and Barros [19] used a relay-based procedure for finding the gain, phase and frequency of points on the frequency response curve. They used a relay in series with an integrator in a special feedback superstructure around the closed loop system to force a limit cycle at the frequency where the loop transfer function has a specified magnitude, r.

2.4. CLID with a quantizer

The publications closest to this paper are those by Welsh and Goodwin [4] and Goodwin and Welsh [20]. Welsh and Goodwin [4] proposed a novel autotuning method based on quantization as an extension to relay tuning. A quantizer at the controller output was used for identification instead of a relay. The main advantage

over relay-based identification is that the controller is kept in the control loop during the identification phase thus avoiding problems with other methods where the replacement of the PID controller by the relay disrupts the operation of the control loop. A second innovation in [4] was to retain the feedback blocking property of the quantizer while injecting other desirable (harmonic) test signals. Goodwin and Welsh [20] extended [4] into multivariable autotuning.

The main advance proposed in this paper is that there is no external signal injected into the loop, while in [4] and [20] harmonic excitation using an external signal is necessary. Instead, a quantizer is inserted into the feedback path in series with the A/D function of the sensor and the quantization error is used as the excitation.

3. CLID with a quantizer

3.1. Quantizer definition

The quantizer is shown in Fig. 2. This is the quantizer defined in MATLAB/Simulink (The MathWorks; Natick, MA). In Fig. 2, each horizontal level is called the quantization level and the gap between two subsequent quantization levels is called quantizer interval. It is a uniform quantizer because it has the same quantizer interval throughout the whole measurement range.

3.2. The proposed closed-loop identification scheme

The proposed closed-loop identification scheme is to insert a quantizer in the feedback path as shown in Fig. 3. As will be demonstrated below, the quantizer interval may be adjusted to generate a persistently exciting signal that satisfies the conditions for system identification.

The quantization error is defined as:

$$d = y_q - y \tag{1}$$

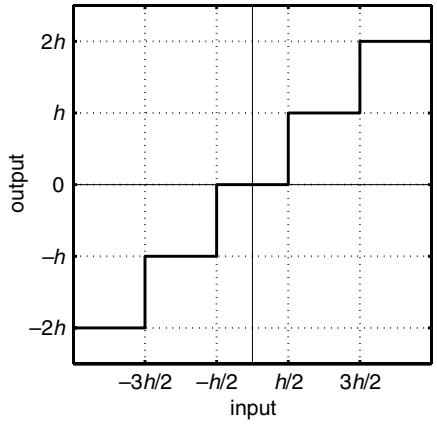

Fig. 2. Input–output characteristic of a quantizer.

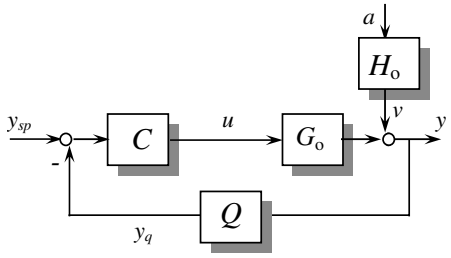

Fig. 3. Scheme for closed-loop identification with a quantizer.

where y_q is the quantizer output and y is the quantizer input. The characteristics of the quantization error signal vary as the quantizer interval, h, is adjusted.

3.3. Key issues for CLID with a quantizer

The statistical theory of quantization has been extensively studied by Gray [21] and Widrow et al. [22] while [23] discussed dithered quantizers in which a stochastic disturbance is added at the quantizer input by the designer. CLID does rely upon y having a stochastic component but the quantizer in CLID is not a dithered quantizer in the sense of [23] because the stochastic component v is not separable from y.

Both [21] and [22] point out that the quantizer output y_q always has a deterministic relationship with the input y, however the quantization error d can be modelled by an additive noise that is uniformly distributed, has a white spectrum and is uncorrelated with the input signal if the quantizer interval is small enough. If the input to the quantizer has a Gaussian distribution then [22] showed the additive noise assumption is very strong if $h \leqslant \sigma_y$.

In a closed loop system subject to a stochastic disturbance it is true that $\sigma_y \geqslant \sigma_a$ provided there is a sample delay in the loop, the equality being achieved with a minimum variance controller. The recommendation made from successful empirical tests in Section 5.2 will be to choose h in the range $1.3\sigma_{\mathrm{mv}}$–$2.0\sigma_{\mathrm{mv}}$ where σ_{mv} is an estimate of σ_a derived from closed loop measurements. Since $\sigma_y \geqslant \sigma_a$ the recommendation is close to the Widrow condition but would exceed it in cases when the controller is close to minimum variance.

For closed-loop identification with a quantizer the quantizer interval has to be small enough that the quantization error excitation will be white and persistently exciting and uncorrelated with the process disturbance. However, the signal-to-noise ratio may not be large enough to do CLID. If the quantizer interval becomes larger then the quantization error excitation will be large enough, but the correlation between the quantization error excitation and the process disturbance will be larger. Such correlation will also endanger the CLID and may lead to a model which is unreliable in some frequency ranges. Therefore, there is the need for a

trade-off between the signal-to-noise ratio and the cross-correlation between the signal and noise in determining the appropriate range of quantizer intervals.

An example of the relationship between the signal-to-noise ratio, the correlation and the quantizer interval is shown in Fig. 4. Fig. 4 was generated from a simulation scheme in which the noise sequence a was white noise with variance 1.00 and the quantizer interval ranged from 0.1 to 3.0 units. The simulation used the models for G_o and H_o which are described in Section 5.1 with a unit feedback control law ($C = 1$). It is necessary to consider the correlation between the quantization error and the disturbance v because of the conditions stated in [7]. The *signal* refers to the quantization error d, the *noise* refers to the process disturbance v and the correlation coefficient is:

$$r_{v,d} = E(v' \times d') \tag{2}$$

where E is the expectation operator and v' and d' are v and d after mean centring and scaling to unit standard deviation. It is noted that v is a coloured noise sequence because it has been through the filter $H_o(q)$ but its distribution is still Gaussian. Fig. 4 shows that the correlation starts to become significant in this simulation experiment when h exceeds a value of about 2.5 (i.e. $h > 2.5\sigma_a$).

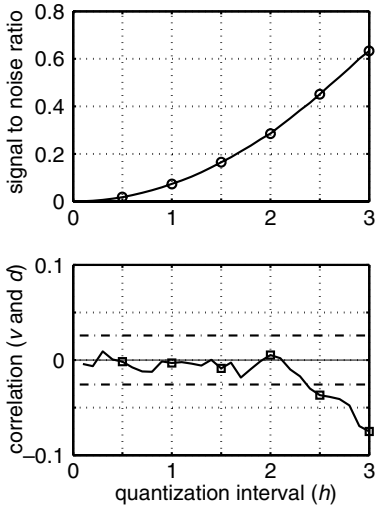

Fig. 4. Effect of quantizer interval in simulation. Upper panel: Signal-to-noise ratio, lower panel: Correlation between disturbance and quantizer error. The dashed lines are ±99% confidence limits beyond which correlation is significant.

3.4. The characteristics of the quantization error

Fig. 5 shows an example of the unquantized signal, quantized signal and quantization error. The noise sequence a was white noise with variance 1.00 and the quantizer interval was 2.0 units. The top panels show the input and output of the quantizer and the difference between these two signals (the quantizer error). The lower panels show the power spectrum of the quantizer error on a normalised frequency axis, its autocovariance

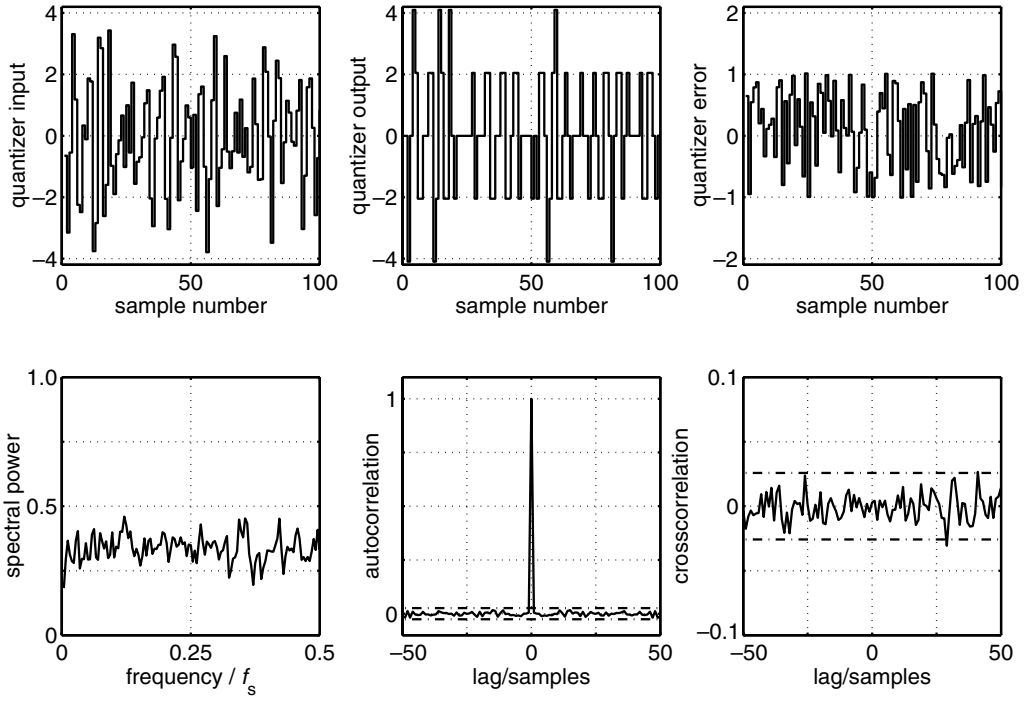

Fig. 5. Simulation with white noise sequence with variance 1.00 and quantizer interval of 2.0 units. Upper panels: (left) unquantized signal, (middle) quantized signal and (right) quantization error. Lower panels: (left) power spectrum of the quantization error, (middle) autocorrelation of the quantization error and (right) cross correlation between quantization error and process disturbance.

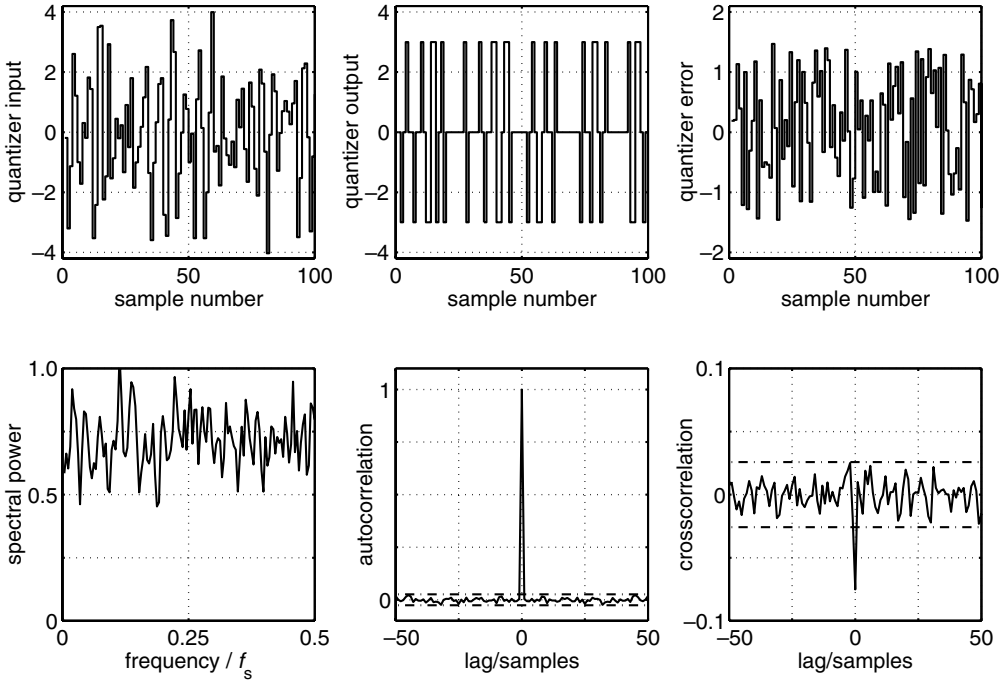

Fig. 6. Simulation with white noise sequence with variance 1.00 and quantizer interval of 3.0 units, showing the same sequence of plots as Fig. 5. The spectral power is the higher than in Fig. 5 but the quantization error and process disturbance have become correlated.

function and also the cross-correlation between the quantization error and the process disturbance, v from the filter $H_o(q)$. The three lower panels of the figure show the quantization error is persistently exciting because the quantization error has a nearly flat power spectrum for all frequencies up to one half the sampling frequency. The lower right panel shows that the correlation between the quantization error and the process disturbance is not significant. Therefore the quantizer provides an independent persistently exciting signal which guarantees the closed loop identifiability.

Fig. 6 was generated under the same condition as Fig. 5 except that the quantizer interval is 3.0 units. From the left bottom and middle bottom panels, the quantization is still persistently exciting. However, from the right bottom panel, there is now correlation between the quantization error and the process disturbance v. The benefit of the larger quantizer interval, however, is that the spectral power of the quantization error, and hence the signal-to-noise ratio, is higher.

A close inspection of Fig. 6 shows why a correlation starts to grow between the quantization error excitation and the process disturbance as the quantizer interval becomes larger. In Fig. 6 there are many instances when the quantizer output is zero because the magnitude of the process output is less than one half of the quantizer interval. The quantizer error is therefore equal to $-y$ and when the loop is operating at a steady set point it becomes correlated with the process disturbance v since y is dominated by v in steady state.

3.5. Use of minimum variance

As was shown in [24,25], y includes an irreducible minimum variance component which remains consistent for different controllers. The standard deviation of the minimum variance component of the process output (σ_{mv}) can be used to give an estimate of σ_a for specifying a suitable quantizer interval for CLID. Following Harris [24] and Desborough and Harris [25], the process output y in Fig. 3 can be expressed as:

$$y = \hat{y} + w \tag{3}$$

where $\hat{y}$ is a forward prediction of y and w is the residual. The minimum variance is the variance of the prediction error sequence, w.

4. Method

4.1. Closed-loop identification with external excitation

A system with the external excitation signal in the feedback path is considered as shown in Fig. 1. The following can be derived from Fig. 1 when d is in the feedback path:

$$u(t) = \frac{-C(q)}{1 + C(q)G_o(q)}d(t) - \frac{C(q)}{1 + C(q)G_o(q)}v(t) \tag{4}$$

$$y(t) = \frac{-C(q)G_o(q)}{1 + C(q)G_o(q)}d(t) + \frac{1}{1 + C(q)G_o(q)}v(t) \tag{5}$$

From (4) and (5)

$$u(t) = -C(q)S_o(q)d(t) - C(q)S_o(q)H_o(q)a(t) \quad (6)$$

$$y(t) = -C(q)G_o(q)S_o(q)d(t) + S_o(q)H_o(q)a(t) \quad (7)$$

where S_o is the true sensitivity function, equal to $1/(1 + G_oC)$.

Defining a noise-free input u^d:

$$u^d(t) = -C(q)S_o(q)d(t) \quad (8)$$

gives:

$$y(t) = G_o(q)u^d(t) + S_o(q)H_o(q)a(t) \quad (9)$$

The two-stage method [3] was chosen because it was characterized in [8] as robust and simple to use. It may be adapted for use with the above expressions using the following procedure:

- The quantity $C(q)S_o(q)$ is obtained from the excitation d and the process input u. Then, a noise-free process input u^d is simulated from excitation d and the $C(q)S_o(q)$ using (8);
- The estimate of the process G_o is calculated from the noise-free process input u^d and the process output y.

The method changes the closed-loop identification into two open-loop identifications and gives unbiased results when excitation $d(t)$ is persistently exciting and uncorrelated with the process disturbance v.

The prediction error method was used for identification with a Box–Jenkins model [2] and was implemented using the System Identification Toolbox (Mathworks, Natick, USA) of MATLAB. Caution should be taken to guarantee the model validation is passed for each open-loop identification. Model validation is achieved by inspection of the residuals between the fitted model and the measurements used to create the model. To pass model validation the residual should be white noise with zero mean and uncorrelated with the input.

4.2. The proposed CLID procedure

A quantizer is inserted in the feedback path to generate quantization errors equivalent to the dither signal d. The adapted methods are given by expressions (6), (8) and (9). The procedures for the proposed CLID scheme are as follows:

- The process is run as normal with quantizer interval very small, for example 0.005 mA, equivalent to 12 bit A/D on a 4–20 mA scale.
- The standard deviation of the minimum variance for the process output (σ_{mv}) is calculated from [25].
- A suitable quantizer interval is chosen (see Section 5.2). The process is run with this quantizer interval. The quantizer error excitation $d = y_q - y$ is used to do closed-loop identification.

4.3. Model accuracy measure

For assessment of the accuracy of the estimate, the identification error is measured by the maximum absolute error expressed as a percentage:

$$\left(\max_{\omega_i} \mid \widehat{G}(j\omega_i) - G(j\omega_i) \mid\right) \times \frac{100}{|G(j\omega_{i(\max)})|} \quad (10)$$

where $G(j\omega_i)$ and $\widehat{G}(j\omega_i)$ are the actual and estimated process frequency response respectively and $\omega_{i(\max)}$ is the angular frequency at which the absolute error is largest. The absolute error is the length of a vector joining the actual and estimated points on a traditional Nyquist plot, it takes account of phase errors as well as magnitude errors.

A second measure is the RMS relative error in the frequency domain, expressed as a percentage.

$$100 \times \sqrt{\text{mean}\left(\left|\frac{\widehat{G}(j\omega_i) - G(j\omega_i)}{G(j\omega_i)}\right|^2\right)} \quad (11)$$

Only the Nyquist curve from zero frequency to the crossover frequency ω_{cg} where the argument is $-180°$ were considered since this part is the most important for system identification and controller design.

4.4. Comparison of the proposed method with [4]

The method of Welsh and Goodwin [4] was implemented for comparison. The frequency response points were obtained from injected test signals at three frequencies ω with the quantizer used to block the feedback path. Frequencies near the cross-over frequency were used as recommended by Welsh and Goodwin. The magnitude and phase angle of the frequency response of the plant were derived from the harmonic signal $A\sin(\omega t + \phi)$ that gave the best fit to the output at the sampling instants $t = nT$, where T is the sampling interval. Following [4] a second-order time delay model was fitted in the least-squares sense to the measured frequency points using a Nelder-Mead simplex search.

5. Simulation examples

5.1. ARMAX model simulation

In order to demonstrate the methods discussed above, the second order ARMAX model presented in [10] was chosen as a simulation example:

$$y(t) = G_o(q)u(t-1) + H_o(q)a(t) \quad (12)$$

where

$$G_o(q) = \frac{0.3403 + 0.2417q^{-1}}{1 - 0.7859q^{-1} + 0.3679q^{-2}}$$

and

$$H_o(q) = \frac{1 - 0.8q^{-1} + 0.12q^{-2}}{1 - 0.7859q^{-1} + 0.3679q^{-2}}$$

A proportional feedback control law was implemented in this simulation. The white noise $a(t)$ was applied with $\sigma_a = 0.1$ or $\sigma_a = 1$ and the number of sampled data points in the simulation was 10 000.

The example presents a challenge for CLID. The cross-over frequency of G_o at $\omega T = 1.62$ is not far from the Nyquist frequency where $\omega T = \pi$. This feature means the excitation created by the quantizer has to contain frequency content all the way up to the Nyquist frequency. Moreover, the open loop system G_o and the noise model H_o both have resonant responses and the resonant frequencies are similar. CLID with a quantizer must accurately determine the frequency and magnitude of the resonance in G_o in the presence of the disturbance from H_o.

5.2. Choice of quantizer interval

Simulations were conducted for different proportional-only controllers under a range of quantizer intervals. A white noise sequence with $\sigma_a = 1$ was applied and fixed throughout different tests.

Initially, the quantizer interval was set to a very small value in order to determine the true variability of the output y and to get an estimate of the minimum variance. Then, CLID with a quantizer was attempted for a range of quantizer intervals summarized in Table 1. The fourth row shows the range of quantizer intervals which allowed successful closed-loop identification.

In Table 1, the standard deviation of the process output σ_y is calculated directly from the measurements while the value of σ_{mv} was calculated using the algorithm given in [25] using a prediction horizon of 1. The results highlight the benefit of the choice of σ_{mv} as a reference for choice of quantizer interval because the estimated σ_{mv} was similar for all three trials, whereas σ_y varied greatly with the largest σ_y being almost double the smallest. If it is based on σ_{mv} then the quantizer interval does not change if the controller gain changes.

The range of quantizer intervals leading to successful CLID shows the procedure was robust in this simulated example and suggests h in the range $1.3\sigma_{mv}$–$2.0\sigma_{mv}$ is a good value to choose because it encompasses all the successful results in Table 1. Case 2 shows successful CLID with h up to $2.37\sigma_{mv}$. The reason for cautiously selecting $2.0\sigma_{mv}$ as the upper limit rather than $2.37\sigma_{mv}$ is to be sure the recommendation does not get too far away from the Widrow criterion $h \leqslant \sigma_y$. With the $h = 1.3\sigma_{mv}$ to $2.0\sigma_{mv}$ recommendation the worst case will be $h = 2\sigma_y$ which will occur when the loop has a minimum variance controller.

Table 1
CLID at various controller gains and quantizer intervals

	Case 1	Case 2	Case 3
Controller gain	1.0	0.5	2.0
σ_{mv}	1.01	1.01	1.01
σ_y	1.37	1.18	2.32
Quantizer interval, h	1.00–2.80	1.30–2.40	0.60–2.90
h/σ_{mv}	0.99–2.76	1.28–2.37	0.59–2.86

5.3. CLID with a quantizer

In order to explore the proposed method under different disturbance and different quantizer intervals, the three cases in Table 2 were designed and tested in simulation using quantizer intervals h within the $1.3\sigma_{mv}$–$2.0\sigma_{mv}$ recommendation established above. The controller gain was fixed at 1.

Results (Bode frequency response plots and step response plots) identified from simulations using CLID with a quantizer are shown in Figs. 7 and 8. Fig. 7 shows frequency response plots and Fig. 8 shows the sampled data step responses. The top panels labeled (a) correspond to Cases 4 and 6 listed in Table 2 which gave identical results because the h/σ_{mv} ratio was the same. The middle panels labeled (b) are from Case 5 and the bottom panels (c) show results from the Welsh and Goodwin method and will be discussed in the next subsection. The identified model accuracy can be seen qualitatively from the figures and the model accuracy measures are listed in Table 3. The most prominent features of the results are:

- The step response comprising an overshoot and damped oscillation has been captured correctly. For instance, the peak and valley of the step response align correctly in Fig. 8;
- The resonant frequency and magnitude in the frequency response plots has been captured correctly in Fig. 7. There are errors at low frequency in the frequency response plot for Cases 4 and 6 which can also be seen in the steady state following the step response in Fig. 8.

Cases 4 and 5 have the same disturbance, $\sigma_a = 0.1$. The only difference is in the quantizer interval applied.

Table 2
CLID with a quantizer simulation conditions

	Case 4	Case 5	Case 6
σ_a	0.10	0.10	1.00
σ_{mv}	0.101	0.101	1.01
Quantizer interval, h	0.137	0.2025	1.373
h/σ_{mv}	1.35	2.0	1.35
Quantizer interval, h equivalent to A/D bit	7	6	3 or 4

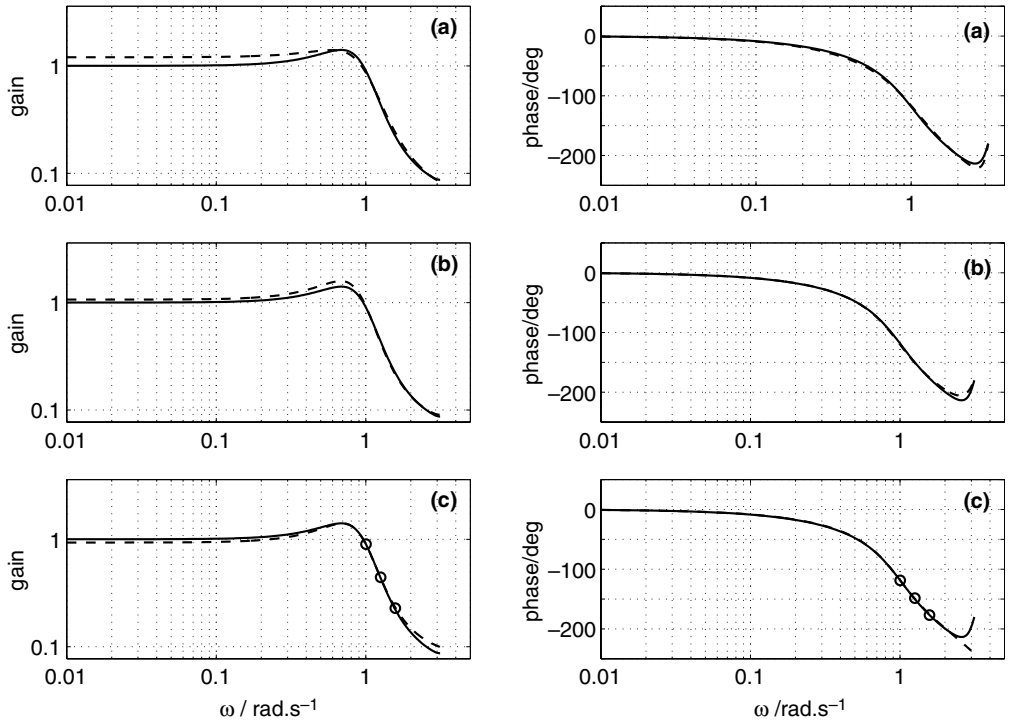

Fig. 7. Frequency response results for simulation: (a) cases 4 and 6, (b) case 5 and (c) using the method of Welsh and Goodwin [4]. (—) true process, (---) identified response and (○) injected frequencies.

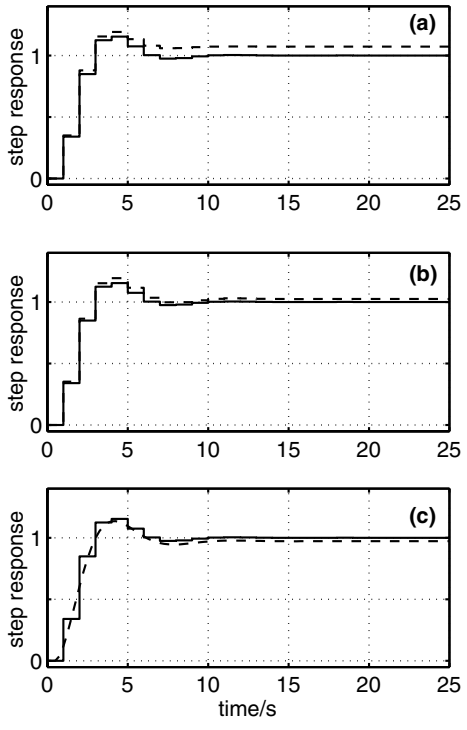

Fig. 8. Step response results for simulation: (a) cases 4 and 6, (b) case 5 and (c) using the method of Welsh and Goodwin [4]. (—) true process and (---) identified response.

The (b) panels of Figs. 7 and 8 compared with the panels labeled (a) show improved accuracy of the identified model. This is because the signal-to-noise ratio was larger in Case 5 due to the increased quantizer interval.

Table 3
Errors and disruption measures for the identified models

	Max	RMS		
	Abs error	$\omega_{i(\max)}$	Rel error	h/σ_{mv}
Case 4	7.2%	0.01	6.1%	1.35
Case 5	4.4%	0.17	3.4%	2.0
Case 6	7.2%	0.01	6.1%	1.35
Welsh and Goodwin [4]	2.7%	0.01	2.5%	7.92

Cases 4 and 6 have different values for σ_a, the standard deviation of the noise sequence. However, in both cases, the ratio between the quantizer interval and the standard deviation σ_{mv} is the same and they achieved the same identification model accuracy. This indicates the proposed scheme is suitable for both small and large disturbances.

5.4. Comparison of the proposed method with Welsh and Goodwin

The bottom panels labelled (c) in Figs. 7 and 8 show the frequency and step responses determined from the method of Welsh and Goodwin [4]. The quantizer is used to block the feedback path while external signals of various frequencies are injected at the input of the system G_o (q) to achieve open-loop identification. The

number of sampled data points in the simulation was 10000. The injected sine wave had a peak-to-peak value of 0.2025 and thus the Welsh and Goodwin test is comparable to Case 5 in terms of its signal-to-noise ratio. The blocking quantizer required an interval larger than the peak-to-peak value of the injected wave because of the presence of noise as was noted in [4]. The input to the quantizer was a sum of the sinusoidal output of $G_o(q)$ and the Gaussian noise $v(t)$. Out of 10000 sampled data points in $v(t)$ about 10 are expected to be further than three standard deviations from the mean. A quantizer with an interval of 0.8 was needed to fully block the feedback path against such extreme values. The disruption to the process is thus greater in the Welsh and Goodwin test because the blocking quantizer interval is 0.8 (h/σ_{mv} ratio of 7.92) rather than 0.2025 in Case 5.

The continuous-time second order with time delay model identified from the frequency response points gave a good fit for the frequency response curve. The model had a delay of 0.5 s which is one half the sampling interval of the sampled data system. It also matches well to the sampled data response although the identified response is continuous rather than sampled. Panels (b) and (c) in Figs. 7 and 8 show that the results achieved by Case 5 of CLID with a quantizer are comparable with the fit from Welsh and Goodwin [4]. The benefits of CLID with a quantizer, however, are that the disruption to the process using CLID with a quantizer is less by a factor of about 4 and there is no need for injection of external signals.

6. Experimental demonstration

6.1. Experimental method

An experimental apparatus (Fig. 9) was used for a practical validation of the proposed scheme for CLID with a quantizer. It is a doubled-walled glass tank. The tank has a height of 50 cm and an inside diameter of 14.5 cm. This system is located in the computer process control laboratory at University of Alberta.

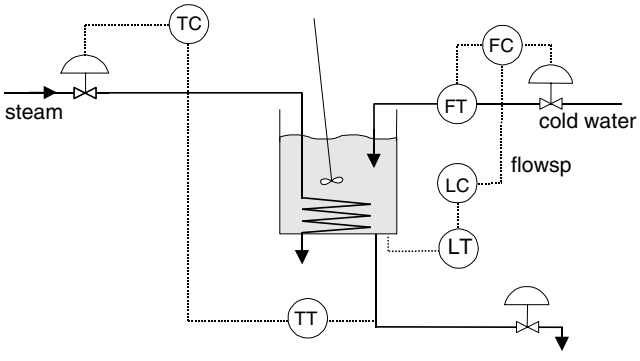

Fig. 9. Diagram of the experimental apparatus, courtesy of the University of Alberta.

6.2. Open loop identification

Open loop identification of the dynamics of the temperature system was carried out in the pilot plant to give a reference against which the results of CLID with a quantizer could be compared. A random binary (RBS) test signal was applied at the steam valve and the output was measured from the temperature sensor in the outlet pipe of the tank. The time delay for the temperature control loop is known to be about 8 s, and the time constant is about 35–40 s so a sampling interval of 4 s was used as a convenient sub-multiple of the delay time giving an adequate number of samples per time constant. The RBS frequency band was selected to identify the frequency response in the range $0–\pi/4\ \text{rad}\cdot\text{s}^{-1}$ (i.e. up to the Nyquist frequency of the signal with 4 s samples). It deviated by −0.6 to 0.6 mA on a 4–20 mA scale around a steam valve set point of 12.57 mA and was generated by the System Identification Toolbox in MATLAB. The temperature system model identified with the RBS test was taken as the benchmark because the RBS test was a direct measurement of the open loop and used a large amplitude excitation relative to that of the CLID test.

The open loop identification results are shown in Fig. 10. The noisy trace in the upper panel is a directly-measured open loop unit amplitude step test while the heavy dashed line is the step response identified from the RBS test. The lower panels of Fig. 10 show the gain and phase of the frequency response identified by the RBS test (dashed lines). The heavy solid lines in Fig. 10 will be discussed shortly.

6.3. Closed loop tests

The controlled variable was the temperature of the water leaving the tank measured by the temperature transmitter TT.

A well-tuned PI controller with continuous transfer function $C(s) = \frac{3s+0.1}{s}$ was used and the manipulated variable was the steam flow. It is noted that the closed loop time constant is smaller than the open loop time constant and the sampling rate required to maintain good closed loop control was 1s.

The setpoint for the temperature loop is 10.5 mA (equivalent to 41.5 °C). The level of the tank was maintained at 12 mA on a 4–20 mA scale (in the middle of the tank) during the experiment. The cold water inflow and the hot water outflow were always balanced when steady-state was reached in the level controller.

There is a natural disturbance to the process in the form of bubbles. Compressed air flows through a pipe, which was submerged into the bottom of the tank

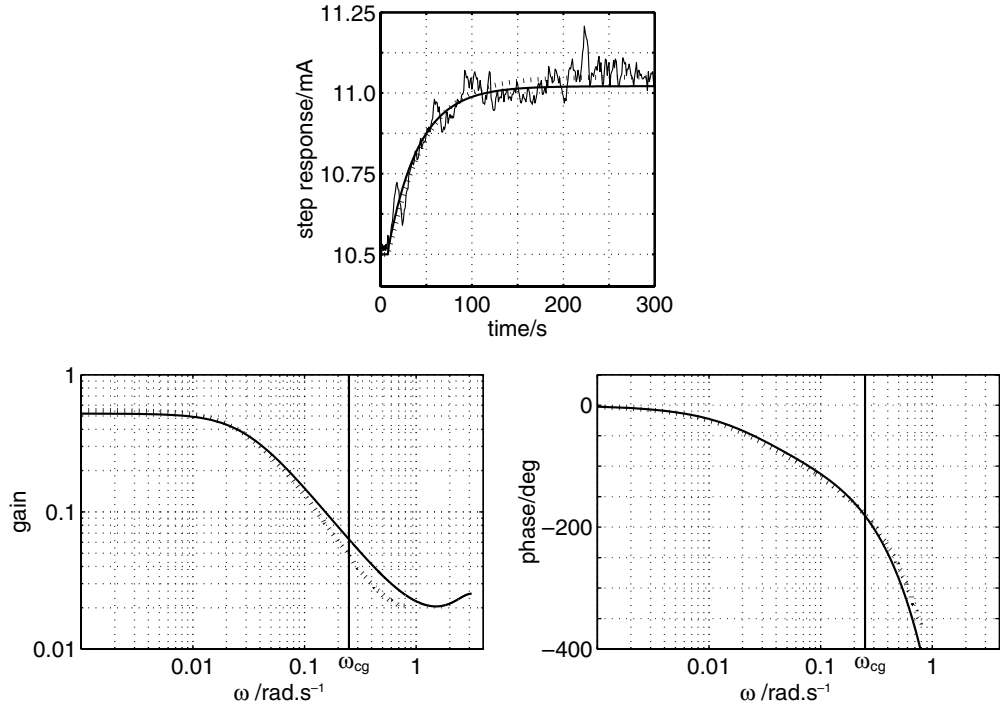

Fig. 10. Identification of the temperature loop. (Upper panel) identified and experimental step responses and (lower panels) identified frequency response function. Results are from open-loop RBS identification (---) and CLID with a quantizer (—).

leading to observable disturbance to both the level and temperature measurements. The experimental procedure is as follows:

- The process is run with the bubbles described above with a very small quantizer interval. The value of σ_{mv} from the minimum variance calculation was $\sigma_{mv} = 0.064$.
- A quantizer interval of $h = 0.105$ mA was chosen, equal to $h = 1.64\sigma_{mv}$. This is within the recommended band of $1.3\sigma_{mv} - 2.0\sigma_{mv}$ from Section 5.2. The quantizer interval in CLID is therefore about 11 times smaller than the peak-to-peak magnitude of the open loop RBS test signal. The process was run with this quantizer interval and the quantizer error excitation $d = y_q - y$ was used for CLID with a quantizer.

6.4. Pre-filtering and sub-sampling

The accuracy of closed-loop identification using experimental data is dependent on the sampling interval. Ljung [2] discusses the effect of sampling interval in Section 13.7 of his book pointing out that experimental system identification finds a set of parameters that minimizes the mismatch between the frequency response of the true system and model integrated over the whole frequency band $-\pi < \omega\ T < \pi$, where T is the sampling interval. The frequency band becomes larger as the sampling interval decreases. Thus, in general, a small sampling interval leads to a model estimate fitted over a wide frequency range while a larger sampling interval will be better for estimation of low frequencies.

The approach taken in this work was to apply the two-stage method [3] for system identification as follows:

- The 1 s data from the CLID with a quantizer experiment were used to identify the high frequency behaviour, i.e. the time delay, the time constant, the corner frequency, the cross-over frequency and gain margin;
- The 1 s data were subsampled by selecting every fourth data point and the resulting 4 s data used to identify the low frequency gain.

6.5. CLID with a quantizer

The results of CLID with a quantizer are the heavy solid lines shown in Fig. 10. Table 4 compares the parameters of the models identified by the benchmark open-loop RBS test and closed-loop identification with the quantizer. The steady-state gain is the low frequency value on the gain frequency response plot, τ is the time taken to accomplish a fraction $1 - e^{-1}$ of the step response, ω_c is the frequency where the gain has reduced by 3 dB, ω_{cg} is the cross-over frequency where the phase is $-180°$ and the gain margin is the reciprocal of the gain at the cross-over frequency. Overall the CLID with a quantizer gives results close to those of the open loop

Table 4
System parameters from open loop identification and CLID with a quantizer

Parameter	Open loop RBS test CLID with quantizer	
Disturbance	±0.6 mA	0.105 mA
Steady state gain	0.550	0.521
Time delay/s	8	7
Time constant, τ/s	40	35
ω_c/rad·s^{-1}	0.027	0.030
ω_{cg}/rad·s^{-1}	0.250	0.247
Gain margin	21.1	15.6

RBS test but with much less disturbance and with no need to take the loop out of service to perform open loop tests.

The top panel in Fig. 10 shows that the shape of the step response from CLID with a quantizer matches well with the open loop RBS test and to the experimental open loop step response. The steady state gain identified using CLID with a quantizer matches the gain identified by the open loop RBS test to within 5%.

The lower panels show the frequency responses. The frequency response identified using CLID with a quantizer extends to higher frequency than the RBS test because the sampling interval was smaller (1 s rather than 4 s). The phase response (lower right panel in Fig. 10) of the frequency response and the cross-over frequency ω_{cg} match well. The overall shape of the gain versus frequency plots (lower left panel) also match well in the region of the corner frequency at about 0.027–0.03 rad·s^{-1}. The gain of the response identified using CLID with a quantizer (the heavy solid line) is overestimated in the frequency range above $\omega = 0.1$, however, and also shows a small increase at very high frequency. As a result, the estimated gain margin is too small by about 25%.

The reason for the mismatch in gain margin is likely to be because there was no anti-aliasing filter in the closed loop control system. Some higher frequencies present in the noisy temperature measurement are aliasing (folding back) into the identified frequency band and boosting signal power.

7. Conclusions

In this paper, a new scheme has been described for closed-loop identification (CLID) with a quantizer inserted in the feedback path. The one parameter to be set is the quantizer interval and a means of choosing the quantizer interval was proposed in order to generate a persistently exciting signal suitable for closed-loop identification. The design requires only that the minimum variance component of the controlled variable be determined using measurements from routine operation. The quantizer interval is made negligibly small once CLID is complete. Simulation examples and experimental evaluation indicate that the strategy is successful. The results compared well with those from the method of Welsh and Goodwin [4] and from open loop identification using a random binary signal.

Acknowledgement

Meihong Wang gratefully acknowledges the financial support of the Centre for Process Systems Engineering, Imperial College London. The UK authors thank the CPC group of the Department of Chemical and Materials Engineering at the University of Alberta for use of pilot plant and facilities.

References

[1] J.F. MacGregor, D.T. Fogal, Closed-loop identification: the role of the noise model and pre-filters, Journal of Process Control 5 (1995) 163–171.
[2] L. Ljung, System Identification: Theory for the User, Prentice-Hall, Englewood Cliffs, NJ, 1999.
[3] P.M.J. Van den Hof, R. Schrama, An indirect method for transfer function estimation from closed loop data, Automatica 29 (1993) 1523–1527.
[4] J.S. Welsh, G.C. Goodwin, A novel mechanism for autotuning based on quantization, in: Proceedings IFAC 14th World Congress, Beijing China, 1999.
[5] I. Gustavsson, L. Ljung, T. Soderstrom, Identification of processes in closed loop—identifiability and accuracy aspects, Automatica 13 (1977) 59–75.
[6] J.F. Bartee, R.C. McFarlane, Identification of linear systems operating in closed loop, Symposium on Advances in Process Control 5, Swansea, 1998, pp. 111–120.
[7] K.R. Godfrey, Perturbation Signals for System Identification, Prentice Hall, Herts, UK, 1993.
[8] U. Forssell, L. Ljung, Closed-loop identification revisited, Automatica 35 (1999) 215–1241.
[9] S.B. Jorgensen, J.H. Lee, Recent advances and challenges in process identification, AIChE Symposium Series 326 (98) (2002) 55–72.
[10] B. Huang, S.L. Shah, Closed-loop identification: a two step approach, Journal of Process Control 7 (1997) 425–438.
[11] A. Esmaili, J.F. MacGregor, P.A. Taylor, Direct and two-step methods for closed-loop identification: a comparison of asymptotic and finite data set performance, Journal of Process Control 10 (2000) 525–537.
[12] M. Gevers, L. Ljung, P. Van den Hof, Asymptotic variance expressions for closed-loop identification, Automatica 37 (2001) 781–786.
[13] P.D. Fox, K.R. Godfrey, Multiharmonic perturbations for nonparametric autotuning, IEE Proceedings–Control Theory and Applications 146 (1999) 1–8.
[14] J. Crowe, M.A. Johnson, Process identifier and its application to industrial control, IEE Proceedings–Control Theory and Applications 147 (2000) 196–204.
[15] D.W. Clarke, J.W. Park, Phase-locked loops for plant tuning and monitoring, IEE Proceedings–Control Theory and Applications 150 (2003) 155–169.
[16] K.J. Åström, T. Hägglund, Automatic tuning of simple regulators with specifications on phase and amplitude margins, Automatica 20 (1984) 645–651.

[17] C.C. Yu, Autotuning of PID Controllers, Springer-Verlag, London, UK, 1999.
[18] Q.G. Wang, C.C. Hang, Q. Bi, Process frequency response estimation from relay feedback, Control Engineering Practice 5 (1997) 1293–1302.
[19] G.H.M. De Arruda, P.R. Barros, Relay-based closed loop transfer function frequency points estimation, Automatica 39 (2003) 309–315.
[20] G.C. Goodwin, J.S. Welsh, Analysis of a novel method of autotuning a multivariable plant based on quantization, in: Proceedings of American Control Conference, San Diego, California, 1999, pp. 347–3351.
[21] R.M. Gray, Quantization noise spectra, IEEE Transactions on Information Theory 36 (1990) 1220–1244.
[22] B. Widrow, I. Kollar, M.C. Liu, Statistical theory of quantization, IEEE Transactions on Instrumentation and Measurement 45 (1996) 353–361.
[23] R.M. Gray, T.G. Stockham, Dithered quantizers, IEEE Transactions on Information Theory 39 (1993) 805–812.
[24] T.J. Harris, Assessment of control loop performance, Canadian Journal of Chemical Engineering 67 (1989) 856–861.
[25] L. Desborough, T.J. Harris, Performance assessment measures for univariate feedback control, Canadian Journal of Chemical Engineering 70 (1992) 1186–1197.