Optimizing Stability Bounds of Finite-Precision Controller Structures for Sampled-Data Systems in the delta Operator domain

S. Chen[†0], *J. Wu*[‡], *R.H. Istepanian*[¶], *J. Chu*[‡] *and J.F. Whidborne*[§]

[†] Department of Electronics and Computer Science
University of Southampton, Highfield
Southampton SO17 1BJ, U.K.

[‡] National Key Laboratory of Industrial Control Technology
Institute of Industrial Process Control
Zhejiang University, Hangzhou, 310027, P. R. China

[¶] Department of Electrical and Computer Engineering
Ryerson Polytechnic University, 350 Victoria Street
Toronto, Ontario, Canada, M5B 2K3

[§] Department of Mechanical Engineering
King's College London, Strand
London WC2R 2LS, U.K.

Abstract

The paper derives a tractable closed-loop stability related measure for controller structures, realized using the δ operator and digitally implemented with finite-word-length (FWL). The optimal realizations of the general finite-precision controller are defined as those that maximize this measure and are shown to be the solutions of a constrained nonlinear optimization problem. For the special case of digital PID controllers, the constrained problem can be decoupled into two simpler unconstrained optimization problems. A global optimization strategy based on the adaptive simulated annealing (ASA) is adopted to provide an efficient method for solving this complex optimal realization problem. Two numerical examples are presented to illustrate the design procedure, and

[0]Contact author. Tel./Fax: +44 (0)23 8095 6660/4508; E-mail: sqc@ecs.soton.ac.uk.

the simulation results confirm that the optimal FWL realizations of the δ-operator based controller have better closed-loop stability margins than those of the usual shift-operator based controller, especially under fast sampling conditions.

Index Terms: Sampled-data system, shift and δ operators, finite word length, stability.

1 Introduction

Although the number of controller implementations using floating-point processors is increasing due to their reduced price, for reasons of cost, simplicity, speed, memory space and ease-of-programming, the use of fixed-point processors is more desired for many industrial and consumer applications, particularly for mass market applications in the automotive and consumer electronics sectors. Furthermore, due to their reliability and well-understood properties, fixed-point processors predominate in critical-safety systems. It is well known that a designed stable closed-loop system may become unstable when the "infinite-precision" controller is implemented using a fixed-point processor due to the FWL effect. The "robustness" of closed-loop stability under controller parameter perturbations is therefore a critical issue in fixed-point implementations.

Many studies have addressed the problem of digital controller realizations with finite-precision considerations [1]–[6]. The first FWL stability measure, denoted as μ_0, was proposed in 1994 [3], but computing explicitly this measure is still an unsolved open problem. Recently, two tractable FWL stability related measures, connected to μ_0, have been derived and the design procedures for searching for optimal FWL controller realizations have been developed [5],[6]. In the original works [5],[6], the term lower bound stability measure was used. We prefer to use the term stability related measure, as it is a lower bound of μ_0 only under some restricted conditions. It can be shown that the measure of [6] is "closer" to μ_0 than that of [5], and the investigations on FWL controller realizations have mainly been based on this stability related measure [7]–[13].

In all the above-mentioned studies, digital controller structures were described and realized with the usual shift operator z. A discrete-time system can also be described and realized with a different operator, called the delta operator δ [14]. Two major advantages are claimed for the use of δ operator realization: a theoretically unified formulation of continuous-time and discrete-time systems; and better numerical properties in FWL implementations [2]. For signal processing applications, it is well known that the digital filter realized with the δ operator has lower round-off noise, lower coefficient error and lower output error variance than the shift operator realization [15],[16]. The use of the δ operator as opposed to the shift operator in control applications has also been promoted [17],[18]. However, no study to date addresses the closed-loop stability issues of FWL controller structures using the δ operator formulation.

This paper analyzes the closed-loop stability of sampled-data control systems in the δ domain with FWL considerations. We derive a new measure quantifying FWL effects on the closed-loop stability. For the computational purpose, a tractable stability related measure is given. The optimal digital controller realization in the δ domain can be obtained by maximizing this measure. As the optimization criterion is non-smooth and non-convex, an efficient non-gradient based global optimization method, known as the ASA [19]–[23], is employed to search for the true optimal controller realization. It turns out that the approach of analyzing FWL digital controllers in the z domain [3],[5],[6] can be extended to study δ-based FWL digital controllers, and the δ-based controller realization has better closed-loop stability robustness to FWL effects over the z-domain approach.

The paper is organized as follows. Section 2 summarizes the definition of δ operator and the equivalent relationship between representing a discrete-time system in the z and δ domains. Section 3 formulates the problem and establishes necessary notations and definitions. A closed-loop stability related measure that can be computed easily for a given controller realization is given in Section 4. The optimal controller realization problem is also defined in this section, which is to find a realization that maximizes the proposed measure. Section 5 presents the optimization framework for obtaining the general optimal FWL controller realization. The problem is formulated as a constrained optimization problem. Section 6 specifically studies the optimal FWL PID controller realization. It is shown that the constrained optimization problem can be decoupled into two unconstrained ones. In Section 7, the effectiveness of the proposed optimization strategy is illustrated by two numerical examples. The paper concludes in Section 8.

2 The δ operator

From a continuous-time transfer function $G(s)$, as the result of discretization procedure with a sampling period h, a discrete-time transfer function $G_z(z)$ based on the shift operator z can be obtained. Define the δ operator as [14]:

$$\delta = \frac{z-1}{h}\,. \tag{1}$$

The transfer function $G_z(z)$ can be re-expressed in δ form: $G_\delta(\delta) = G_z(z)$. Obviously $G_z(z)$ and $G_\delta(\delta)$ are two different but equivalent parameterizations representing the same input-output relationship. The state-space models in the z and δ domains are [24],[25]

$$\left.\begin{array}{l} z\,\mathbf{x}_z(k) = \mathbf{A}_z\mathbf{x}_z(k) + \mathbf{B}_z\mathbf{u}(k) \\ \mathbf{y}(k) = \mathbf{C}_z\mathbf{x}_z(k) + \mathbf{D}_z\mathbf{u}(k) \end{array}\right\} \tag{2}$$

and

$$\left.\begin{array}{l} \delta\,\mathbf{x}_\delta(k) = \mathbf{A}_\delta\mathbf{x}_\delta(k) + \mathbf{B}_\delta\mathbf{u}(k) \\ \mathbf{y}(k) = \mathbf{C}_\delta\mathbf{x}_\delta(k) + \mathbf{D}_\delta\mathbf{u}(k) \end{array}\right\} \tag{3}$$

respectively, where the various vectors and matrices are understood to have appropriate dimensions, and

$$G_z(z) = \mathbf{C}_z(z\,\mathbf{I} - \mathbf{A}_z)^{-1}\mathbf{B}_z + \mathbf{D}_z = G_\delta(\delta) = \mathbf{C}_\delta(\delta\,\mathbf{I} - \mathbf{A}_\delta)^{-1}\mathbf{B}_\delta + \mathbf{D}_\delta\,. \tag{4}$$

It follows from (4) that the relationships relate the two state-space representations are

$$\mathbf{A}_z = h\,\mathbf{A}_\delta + \mathbf{I},\ \ \mathbf{B}_z = h\,\mathbf{B}_\delta,\ \ \mathbf{C}_z = \mathbf{C}_\delta,\ \ \mathbf{D}_z = \mathbf{D}_\delta\,. \tag{5}$$

It is well known that the state-space realization of an input-output transfer function is not unique. Define a generalized operator ρ, where $\rho = z$ or δ, and

$$\mathcal{S}_\rho \triangleq \{(\mathbf{A}_\rho, \mathbf{B}_\rho, \mathbf{C}_\rho, \mathbf{D}_\rho) :\ G_\rho(\rho) = \mathbf{C}_\rho(\rho\,\mathbf{I} - \mathbf{A}_\rho)^{-1}\mathbf{B}_\rho + \mathbf{D}_\rho\}\,. \tag{6}$$

Then if $(\mathbf{A}_\rho, \mathbf{B}_\rho, \mathbf{C}_\rho, \mathbf{D}_\rho) \in \mathcal{S}_\rho$, $(\mathbf{T}^{-1}\mathbf{A}_\rho\mathbf{T}, \mathbf{T}^{-1}\mathbf{B}_\rho, \mathbf{C}_\rho\mathbf{T}, \mathbf{D}_\rho) \in \mathcal{S}_\rho$ for any nonsingular $\mathbf{T}$. Let $\{\lambda_{\rho,i}\}$ be the eigenvalues of $\mathbf{A}_\rho$. The following lemma relates $\{\lambda_{\delta,i}\}$ to $\{\lambda_{z,i}\}$.

Lemma 1 $\lambda_{z,i} = 1 + h\,\lambda_{\delta,i},\ \forall i.$

The proof of lemma 1 is straightforward based on the definition of eigenvalue and the relationship $\mathbf{A}_z = h\,\mathbf{A}_\delta + \mathbf{I}$. It is well known that the discrete-time system $(\mathbf{A}_z, \mathbf{B}_z, \mathbf{C}_z, \mathbf{D}_z)$ is stable if and only if all the eigenvalues $|\lambda_{z,i}| < 1$. From lemma 1, we have the condition of the stability for the discrete-time system described with the δ operator.

Lemma 2 The discrete-time system $(\mathbf{A}_\delta, \mathbf{B}_\delta, \mathbf{C}_\delta, \mathbf{D}_\delta)$ is stable if and only if

$$\left|\lambda_{\delta,i} + \frac{1}{h}\right| < \frac{1}{h}, \ \forall i \,. \tag{7}$$

3 FWL stability issue in the δ domain

Consider the sampled-data system depicted in Fig. 1, where $P(s)$ is strictly proper. The discrete-time plant $P(\delta) = S_h P(s) H_h$ has a state-space realization $(\mathbf{A}_p, \mathbf{B}_p, \mathbf{C}_p, \mathbf{0})$ in the δ domain, where $\mathbf{A}_p \in \mathcal{R}^{m\times m}$, $\mathbf{B}_p \in \mathcal{R}^{m\times l}$ and $\mathbf{C}_p \in \mathcal{R}^{q\times m}$. The controller $C(\delta)$ has a state-space realization $(\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c, \mathbf{D}_c)$ with $\mathbf{A}_c \in \mathcal{R}^{n\times n}$, $\mathbf{B}_c \in \mathcal{R}^{n\times q}$, $\mathbf{C}_c \in \mathcal{R}^{l\times n}$ and $\mathbf{D}_c \in \mathcal{R}^{l\times q}$. The corresponding realization $(\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{C}}, \bar{\mathbf{D}})$ of the closed-loop system is:

$$\begin{aligned}\bar{\mathbf{A}} &= \begin{bmatrix} \mathbf{A}_p + \mathbf{B}_p\mathbf{D}_c\mathbf{C}_p & \mathbf{B}_p\mathbf{C}_c \\ \mathbf{B}_c\mathbf{C}_p & \mathbf{A}_c \end{bmatrix} = \begin{bmatrix} \mathbf{A}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{B}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{D}_c & \mathbf{C}_c \\ \mathbf{B}_c & \mathbf{A}_c \end{bmatrix} \begin{bmatrix} \mathbf{C}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix} \\ &= \mathbf{M}_0 + \mathbf{M}_1\mathbf{X}_\delta\mathbf{M}_2 = \bar{\mathbf{A}}(\mathbf{X}_\delta)\,, \end{aligned} \tag{8}$$

$$\bar{\mathbf{B}} = \begin{bmatrix} \mathbf{B}_p \\ \mathbf{0} \end{bmatrix}, \ \bar{\mathbf{C}} = [\mathbf{C}_p \ \mathbf{0}], \ \bar{\mathbf{D}} = \mathbf{0}\,, \tag{9}$$

where all $\mathbf{0}$'s are zero matrices of proper dimensions, $\mathbf{I}_n$ is the $n \times n$ identity matrix, and

$$\mathbf{X}_\delta = \begin{bmatrix} \mathbf{D}_c & \mathbf{C}_c \\ \mathbf{B}_c & \mathbf{A}_c \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & \cdots & p_{q+n} \\ p_{q+n+1} & p_{q+n+2} & \cdots & p_{2(q+n)} \\ \vdots & \vdots & \cdots & \vdots \\ p_{(l+n-1)(q+n)+1} & p_{(l+n-1)(q+n)+2} & \cdots & p_{(l+n)(q+n)} \end{bmatrix} \tag{10}$$

is the controller matrix. Let $C(\delta)$ be chosen to make the feedback system stable. Then all the eigenvalues of $\bar{\mathbf{A}}(\mathbf{X}_\delta)$, denoted by $\{\lambda_i, 1 \le i \le m+n\}$, satisfy $\left|\lambda_i + \frac{1}{h}\right| < \frac{1}{h}$, $\forall i$.

When the realization $(\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c, \mathbf{D}_c)$ is implemented with a fixed-point digital control processor, $\mathbf{X}_\delta$ is perturbed into $\mathbf{X}_\delta + \Delta\mathbf{X}_\delta$ due to the FWL effects, where

$$\Delta\mathbf{X}_\delta = \begin{bmatrix} \Delta p_1 & \Delta p_2 & \cdots & \Delta p_{q+n} \\ \Delta p_{q+n+1} & \Delta p_{q+n+2} & \cdots & \Delta p_{2(q+n)} \\ \vdots & \vdots & \cdots & \vdots \\ \Delta p_{(l+n-1)(q+n)+1} & \Delta p_{(l+n-1)(q+n)+2} & \cdots & \Delta p_N \end{bmatrix} \tag{11}$$

and $N = (l+n)(q+n)$. Each element of $\Delta\mathbf{X}_\delta$ is bounded by $\frac{\epsilon}{2}$, that is,

$$\mu(\Delta\mathbf{X}_\delta) \triangleq \max_{1\le i\le N} |\Delta p_i| \le \frac{\epsilon}{2}\,. \tag{12}$$

For a fixed-point processor of B_s bits

$$\epsilon = 2^{-(B_s - B_X)}, \tag{13}$$

where 2^{B_X} is a normalization factor such that the absolute value of each element of $2^{-B_X}\mathbf{X}_\delta$ is not larger than 1. With the perturbation $\Delta\mathbf{X}_\delta$, λ_i is moved to $\tilde{\lambda}_i$. The sampled-data system will be unstable if and only if there exists $i \in \{1, \cdots, m+n\}$ such that $\left|\tilde{\lambda}_i + \frac{1}{h}\right| \geq \frac{1}{h}$.

To see when the round off error will cause the closed-loop system to become unstable, let us introduce the following stability measure:

$$\mu_{\delta 0}(\mathbf{X}_\delta) \stackrel{\triangle}{=} \inf\{\mu(\Delta\mathbf{X}_\delta) : \ \bar{\mathbf{A}}(\mathbf{X}_\delta + \Delta\mathbf{X}_\delta) \text{ is unstable}\}. \tag{14}$$

From this definition, it is obvious that:

Proposition 1 $\bar{\mathbf{A}}(\mathbf{X}_\delta + \Delta\mathbf{X}_\delta)$ is stable if $\mu(\Delta\mathbf{X}_\delta) < \mu_{\delta 0}(\mathbf{X}_\delta)$.

The larger $\mu_{\delta 0}(\mathbf{X}_\delta)$ is, the bigger FWL error the closed-loop stability can tolerate. Let $B_s^{\min}$ be the smallest word length that, when used to implement $\mathbf{X}_\delta$, can guarantee the closed-loop stability. It would be highly desirable to know $B_s^{\min}$. However, except in simulation, it is impossible to test the closed-loop system by reducing B_s until it becomes unstable. Assuming that h is realized exactly, an estimate of $B_s^{\min}$ can be provided by

$$\hat{B}_{s0}^{\min} = \mathrm{Int}[-\log_2(\mu_{\delta 0}(\mathbf{X}_\delta))] - 1 + B_X, \tag{15}$$

where $\mathrm{Int}[x]$ rounds x to the nearest integer and $\mathrm{Int}[x] \geq x$. From (12) to (15), it can be seen that the closed-loop system is stable when $\mathbf{X}_\delta$ is implemented with a fixed-point processor of at least $\hat{B}_{s0}^{\min}$ bits.

It is worth emphasizing an often overlooked constraint on the FWL implementation of δ-based controllers. The state-space equation of the δ-based controller, $\delta\,\mathbf{x}(k) = \mathbf{A}_c\mathbf{x}(k) + \mathbf{B}_c\mathbf{u}(k)$, is realized using: $\mathbf{x}(k+1) = \mathbf{x}(k) + h\,(\mathbf{A}_c\mathbf{x}(k) + \mathbf{B}_c\mathbf{u}(k))$. The sampling period h should be realized exactly without FWL errors. Otherwise, analysis based on $\mathbf{X}_\delta$ may not be valid. Specifically, assume that h can be realized exactly by B_h bits with the integer part of h requiring B_{hI} bits and the fractional part of h requiring B_{hF} bits. A modified estimate of the minimum bit length that can guarantee the closed-loop stability is

$$\hat{B}_{s0h}^{\min} = \max\{B_{hI}, B_X\} + \max\{B_{hF}, \hat{B}_{s0}^{\min} - B_X\}. \tag{16}$$

Notice that the δ-domain stability measure $\mu_{\delta 0}$, defined in (14), is similar to the stability measure μ_0 for z-operator based controller realizations given in [3]. Like μ_0, how to compute explicitly the value of $\mu_{\delta 0}$ for a given realization $\mathbf{X}_\delta$ is also an unsolved open problem. Thus, the stability measure $\mu_{\delta 0}(\mathbf{X}_\delta)$ has very limited practical value. Alternative measure that can not only quantify the FWL effects on stability robustness but also be computed easily must be sought.

4 FWL stability related measure in the δ domain

Roughly speaking, how easily the FWL error $\Delta\mathbf{X}_\delta$ can cause a stable control system to become unstable is strongly determined by how close λ_i are to the unstable boundary and how sensitive they are to the controller parameter perturbations. We propose the following stability related measure:

$$\mu_{\delta 1}(\mathbf{X}_\delta) \triangleq \min_{1\le i\le m+n} \frac{\frac{1}{h} - \left|\lambda_i + \frac{1}{h}\right|}{\sum_{j=1}^{N}\left|\left.\frac{\partial \lambda_i}{\partial p_j}\right|_{\mathbf{X}_\delta}\right|}. \tag{17}$$

Let λ_l and $\tilde{\lambda}_l$ be the lth eigenvalues of $\bar{\mathbf{A}}(\mathbf{X}_\delta)$ and $\bar{\mathbf{A}}(\mathbf{X}_\delta + \Delta\mathbf{X}_\delta)$, respectively, and define

$$\mathcal{P}(\mathbf{X}_\delta) \triangleq \left\{ \Delta\mathbf{X}_\delta : \left|\tilde{\lambda}_l + \frac{1}{h}\right| - \left|\lambda_l + \frac{1}{h}\right| \le \mu(\Delta\mathbf{X}_\delta)\sum_{j=1}^{N}\left|\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{X}_\delta}\right|, \forall l \right\}. \tag{18}$$

We have the following proposition:

Proposition 2 $\bar{\mathbf{A}}(\mathbf{X}_\delta + \Delta\mathbf{X}_\delta)$ is stable if $\Delta\mathbf{X}_\delta \in \mathcal{P}(\mathbf{X}_\delta)$ and $\mu_{\delta 1}(\mathbf{X}_\delta) > \mu(\Delta\mathbf{X}_\delta)$.

Proof: For $\Delta\mathbf{X}_\delta \in \mathcal{P}(\mathbf{X}_\delta)$,

$$\left|\tilde{\lambda}_l + \frac{1}{h}\right| - \left|\lambda_l + \frac{1}{h}\right| \le \mu(\Delta\mathbf{X}_\delta)\sum_{j=1}^{N}\left|\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{X}_\delta}\right|, \forall l\,. \tag{19}$$

It follows from $\mu_{\delta 1}(\mathbf{X}_\delta) > \mu(\Delta\mathbf{X}_\delta)$ that

$$\begin{aligned} \left|\tilde{\lambda}_l + \frac{1}{h}\right| &\le \left|\lambda_l + \frac{1}{h}\right| + \mu(\Delta\mathbf{X}_\delta)\sum_{j=1}^{N}\left|\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{X}_\delta}\right| \\ &< \left|\lambda_l + \frac{1}{h}\right| + \frac{\frac{1}{h} - \left|\lambda_l + \frac{1}{h}\right|}{\sum_{j=1}^{N}\left|\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{X}_\delta}\right|}\sum_{j=1}^{N}\left|\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{X}_\delta}\right| = \frac{1}{h},\ \forall l\,, \end{aligned} \tag{20}$$

which means that $\bar{\mathbf{A}}(\mathbf{X}_\delta + \Delta\mathbf{X}_\delta)$ is stable.

Remarks: The requirement $\Delta\mathbf{X}_\delta \in \mathcal{P}(\mathbf{X}_\delta)$ is not over restricted. In practice, we are only interested in those $\Delta\mathbf{X}_\delta$ which lie in a bounded region including $\Delta\mathbf{X}_\delta = 0$. More plainly, we are only interested for those $\Delta\mathbf{X}_\delta$ lying in (see proposition 1):

$$\mathcal{Q}(\mathbf{X}_\delta) \triangleq \{\Delta\mathbf{X}_\delta : \mu(\Delta\mathbf{X}_\delta) < \mu_{\delta 0}(\mathbf{X}_\delta)\} \ . \tag{21}$$

Since $\frac{\partial \lambda_l}{\partial p_j}$ is continuous,

$$\tilde{\lambda}_l + \frac{1}{h} = \lambda_l + \frac{1}{h} + \sum_{j=1}^{N} \int_{\mathcal{C}} \frac{\partial \lambda_l}{\partial p_j}\, dp_j = \lambda_l + \frac{1}{h} + \sum_{j=1}^{N} \left(\mathrm{Re}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{a}_j} \right] + i\, \mathrm{Im}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{b}_j} \right] \right) \Delta p_j \ , \tag{22}$$

where $\mathcal{C}$ is the oriented segment from $\mathbf{X}_\delta$ to $\mathbf{X}_\delta + \Delta\mathbf{X}_\delta$, $\mathbf{a}_j$ and $\mathbf{b}_j$ are some points on $\mathcal{C}$, $\mathrm{Re}[x]$ and $\mathrm{Im}[x]$ are the real and imaginary parts of the complex number x, respectively, and $i = \sqrt{-1}$. Hence

$$\left| \tilde{\lambda}_l + \frac{1}{h} \right| - \left| \lambda_l + \frac{1}{h} \right| \le \left| \sum_{j=1}^{N} \left(\mathrm{Re}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{a}_j} \right] + i\, \mathrm{Im}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{b}_j} \right] \right) \Delta p_j \right| \ . \tag{23}$$

Now let us compare

$$\left| \sum_{j=1}^{N} \left(\mathrm{Re}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{a}_j} \right] + i\, \mathrm{Im}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{b}_j} \right] \right) \Delta p_j \right| \tag{24}$$

with

$$\sum_{j=1}^{N} \left| \left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{X}_\delta} \right| \mu(\Delta\mathbf{X}_\delta) \ . \tag{25}$$

Notice that all the N real-valued items $\left| \left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{X}_\delta} \right|$ are in alignment; while the N complex-valued items $\left(\mathrm{Re}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{a}_j} \right] + i\, \mathrm{Im}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{b}_j} \right] \right)$ are generally out of alignment. Moreover, $|\Delta p_j| \le \mu(\Delta\mathbf{X}_\delta)$, $\mathrm{Re}\left[\frac{\partial \lambda_l}{\partial p_j}\right]$ and $\mathrm{Im}\left[\frac{\partial \lambda_l}{\partial p_j}\right]$ are differentiable. Thus, there exists a rather large κ such that $\forall \Delta\mathbf{X}_\delta \in \{\Delta\mathbf{X}_\delta : \mu(\Delta\mathbf{X}_\delta) \le \kappa\}$,

$$\left| \sum_{j=1}^{N} \left(\mathrm{Re}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{a}_j} \right] + i\, \mathrm{Im}\left[\left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{b}_j} \right] \right) \Delta p_j \right| \le \sum_{j=1}^{N} \left| \left.\frac{\partial \lambda_l}{\partial p_j}\right|_{\mathbf{X}_\delta} \right| \mu(\Delta\mathbf{X}_\delta) \ . \tag{26}$$

The above analysis shows that $\mathcal{P}(\mathbf{X}_\delta)$ exists and a rather large part of $\mathcal{Q}(\mathbf{X}_\delta)$ can be covered by $\mathcal{P}(\mathbf{X}_\delta)$, that is, the condition $\Delta\mathbf{X}_\delta \in \mathcal{P}(\mathbf{X}_\delta)$ is not too restricted.

Notice that, although $\mu_{\delta 1}(\mathbf{X}_\delta)$ can be used to describe the FWL stability characteristics, it is not generally true that "$\bar{A}(\mathbf{X}_\delta + \Delta\mathbf{X}_\delta)$ is stable if $\mu(\Delta\mathbf{X}_\delta) < \mu_{\delta 1}(\mathbf{X}_\delta)$". This is in contrast to $\mu_{\delta 0}(\mathbf{X}_\delta)$. For this reason, we prefer to call $\mu_{\delta 1}(\mathbf{X}_\delta)$ a stability related measure. Also, generally speaking, there is no rigor relationship between $\mu_{\delta 0}(\mathbf{X}_\delta)$ and $\mu_{\delta 1}(\mathbf{X}_\delta)$, but $\mu_{\delta 1}(\mathbf{X}_\delta)$ is connected with a lower bound of $\mu_{\delta 0}(\mathbf{X}_\delta)$ in some manners, as shown in the following corollary. First, define

$$\rho\left(\mathcal{P}(\mathbf{X}_\delta)\right) \stackrel{\triangle}{=} \inf_{\Delta \mathbf{X}_\delta \notin \mathcal{P}(\mathbf{X}_\delta)} \mu(\Delta \mathbf{X}_\delta)\,. \tag{27}$$

Corollary 1 $\mu_{\delta 1}(\mathbf{X}_\delta) \le \mu_{\delta 0}(\mathbf{X}_\delta)$ if $\rho\left(\mathcal{P}(\mathbf{X}_\delta)\right) > \mu_{\delta 0}(\mathbf{X}_\delta)$.

From corollary 1, it can be seen that $\mu_{\delta 1}(\mathbf{X}_\delta)$ can be considered as a lower bound of $\mu_{\delta 0}(\mathbf{X}_\delta)$, provided that $\mu_{\delta 0}(\mathbf{X}_\delta)$ is small enough. The assumption of small $\mu_{\delta 0}(\mathbf{X}_\delta)$ is not over restricted, as it does not make much sense to study the FWL effects on the closed-loop stability for those situations where the closed-loop systems have a very large stability robustness. It should be pointed out that most of digital control systems do have a small stability robustness, which is especially true when fast sampling is applied.

To compute $\mu_{\delta 1}(\mathbf{X}_\delta)$, $\{\frac{\partial \lambda_i}{\partial p_j}\}$ are needed. The following theorem shows that these eigenvalue sensitivities can easily be calculated.

Theorem 1 Let $\mathbf{A} = \mathbf{M}_0 + \mathbf{M}_1 \mathbf{X} \mathbf{M}_2 \in \mathcal{R}^{v \times v}$ be diagonalisable and denote $\{\lambda_i\}$ its eigenvalues, where $\mathbf{X} \in \mathcal{R}^{l \times r}$, and $\mathbf{M}_0$, $\mathbf{M}_1$ and $\mathbf{M}_2$ are independent of $\mathbf{X}$ with proper dimensions. Let $\mathbf{x}_i$ be a right eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ_i. Denote $\mathbf{M}_\mathbf{x} = [\,\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_v\,]$ and $\mathbf{M}_\mathbf{y} = [\,\mathbf{y}_1 \quad \mathbf{y}_2 \quad \cdots \quad \mathbf{y}_v\,] = \mathbf{M}_\mathbf{x}^{-\mathcal{H}}$, where $\mathcal{H}$ denotes the transpose and conjugate operation and $\mathbf{y}_i$ is called the reciprocal left eigenvector corresponding to λ_i. Then

$$\frac{\partial \lambda_i}{\partial \mathbf{X}} = \begin{bmatrix} \frac{\partial \lambda_i}{\partial x_{11}} & \frac{\partial \lambda_i}{\partial x_{12}} & \cdots & \frac{\partial \lambda_i}{\partial x_{1r}} \\ \frac{\partial \lambda_i}{\partial x_{21}} & \frac{\partial \lambda_i}{\partial x_{22}} & \cdots & \frac{\partial \lambda_i}{\partial x_{2r}} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \lambda_i}{\partial x_{l1}} & \frac{\partial \lambda_i}{\partial x_{l2}} & \cdots & \frac{\partial \lambda_i}{\partial x_{lr}} \end{bmatrix} = \mathbf{M}_1^T \mathbf{y}_i^* \mathbf{x}_i^T \mathbf{M}_2^T\,, \tag{28}$$

where x_{kj} is the (k, j)th element of $\mathbf{X}$, and the superscript $*$ denotes the conjugate operation.

Proof: Let α be a variable independent of $\mathbf{M}_1$ and $\mathbf{M}_2$. It follows from $\mathbf{y}_i^{\mathcal{H}} \mathbf{x}_i = 1$ that

$$\frac{\partial \mathbf{y}_i^{\mathcal{H}}}{\partial \alpha} \mathbf{x}_i + \mathbf{y}_i^{\mathcal{H}} \frac{\partial \mathbf{x}_i}{\partial \alpha} = 0\,. \tag{29}$$

Notice that $\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i$ and $\lambda_i = \mathbf{y}_i^{\mathcal{H}} \mathbf{A}\mathbf{x}_i$. Hence

$$\frac{\partial \lambda_i}{\partial \alpha} = \frac{\partial \mathbf{y}_i^{\mathcal{H}}}{\partial \alpha} \mathbf{A}\mathbf{x}_i + \mathbf{y}_i^{\mathcal{H}} \frac{\partial \mathbf{A}}{\partial \alpha} \mathbf{x}_i + \mathbf{y}_i^{\mathcal{H}} \mathbf{A} \frac{\partial \mathbf{x}_i}{\partial \alpha} . \tag{30}$$

It follows from (29) and $\mathbf{y}_i^{\mathcal{H}} \mathbf{A} = \lambda_i \mathbf{y}_i^{\mathcal{H}}$ that

$$\frac{\partial \lambda_i}{\partial \alpha} = \left(\frac{\partial \mathbf{y}_i^{\mathcal{H}}}{\partial \alpha} \lambda_i \mathbf{x}_i + \lambda_i \mathbf{y}_i^{\mathcal{H}} \frac{\partial \mathbf{x}_i}{\partial \alpha} \right) + \mathbf{y}_i^{\mathcal{H}} \frac{\partial \mathbf{A}}{\partial \alpha} \mathbf{x}_i = \mathbf{y}_i^{\mathcal{H}} \frac{\partial \mathbf{A}}{\partial \alpha} \mathbf{x}_i = \mathbf{y}_i^{\mathcal{H}} \mathbf{M}_1 \frac{\partial \mathbf{X}}{\partial \alpha} \mathbf{M}_2 \mathbf{x}_i . \tag{31}$$

Let $\alpha = x_{kj}$. Then

$$\frac{\partial \lambda_i}{\partial \alpha} = (\mathbf{y}_i^{\mathcal{H}} \mathbf{M}_1)_k (\mathbf{M}_2 \mathbf{x}_i)_j , \tag{32}$$

where $(\mathbf{y}_i^{\mathcal{H}} \mathbf{M}_1)_k$ and $(\mathbf{M}_2 \mathbf{x}_i)_j$ are the kth and jth elements of $\mathbf{y}_i^{\mathcal{H}} \mathbf{M}_1$ and $\mathbf{M}_2 \mathbf{x}_i$, respectively. This leads to (28).

Since $\mu_{\delta 1}(\mathbf{X}_\delta)$ is computationally tractable, for a given controller realization $\mathbf{X}_\delta$, we can estimate the smallest word length $B_s^{\min}$ based on $\mu_{\delta 1}(\mathbf{X}_\delta)$ using the following

$$\hat{B}_{s1}^{\min} = \mathrm{Int}[-\log_2(\mu_{\delta 1}(\mathbf{X}_\delta))] - 1 + B_X . \tag{33}$$

When the requirement for implementing h exactly is taken into account, the estimated smallest bit length should be modified to

$$\hat{B}_{s1h}^{\min} = \max\{B_{hI}, B_X\} + \max\{B_{hF}, \hat{B}_{s1}^{\min} - B_X\} . \tag{34}$$

It should be pointed out that although $\mu_{\delta 1}(\mathbf{X}_\delta)$ can be used to estimate $B_s^{\min}$, its importance lies in the fact that it can be used as the optimization criterion to search for an optimal controller realization, defined as:

$$\mathbf{X}_{\delta \mathrm{opt}} \triangleq \arg \max_{\mathbf{X}_\delta \in \mathcal{S}(\mathbf{X}_\delta)} \mu_{\delta 1}(\mathbf{X}_\delta) , \tag{35}$$

where

$$\mathcal{S}(\mathbf{X}_\delta) \triangleq \{\mathbf{X}_\delta : C(\delta) = \mathbf{C}_c(\delta \mathbf{I} - \mathbf{A}_c)^{-1} \mathbf{B}_c + \mathbf{D}_c\} . \tag{36}$$

is the set of all the realizations of the controller $C(\delta)$. The realization $\mathbf{X}_{\delta \mathrm{opt}}$ is optimal in the sense that it has maximum stability robustness to FWL effects. The digital controller implemented with an optimal realization means that the stability of the closed-loop system is guaranteed with a minimum hardware requirement in terms of word length. The detailed design procedure for finding an optimal controller realization will be discussed in the next section.

5 Optimal realization of FWL controller structures in the δ domain

To begin with the optimal design procedure, assume that an initial controller realization $\mathbf{X}_{\delta 0}$ is given to be

$$\mathbf{X}_{\delta 0} = \begin{bmatrix} \mathbf{D}_c^0 & \mathbf{C}_c^0 \\ \mathbf{B}_c^0 & \mathbf{A}_c^0 \end{bmatrix}, \tag{37}$$

where $C(\delta) = \mathbf{C}_c^0(\delta\,\mathbf{I} - \mathbf{A}_c^0)^{-1}\mathbf{B}_c^0 + \mathbf{D}_c^0$. Any realization of $C(\delta)$ can be expressed as:

$$\mathbf{X}_{\delta\mathbf{T}} \triangleq \begin{bmatrix} \mathbf{I}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-1} \end{bmatrix} \mathbf{X}_{\delta 0} \begin{bmatrix} \mathbf{I_q} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{bmatrix}, \tag{38}$$

where $\mathbf{T} \in \mathcal{R}^{n\times n}$ and $\det(\mathbf{T}) \neq 0$. From (8), it can be shown that the transition matrix of the closed-loop system is

$$\bar{\mathbf{A}}(\mathbf{X}_{\delta\mathbf{T}}) = \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-1} \end{bmatrix} \bar{\mathbf{A}}(\mathbf{X}_{\delta 0}) \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{bmatrix}. \tag{39}$$

From (39), applying theorem 1 results in

$$\left.\frac{\partial \lambda_i}{\partial \mathbf{X}_\delta}\right|_{\mathbf{X}_\delta = \mathbf{X}_{\delta\mathbf{T}}} = \begin{bmatrix} \mathbf{I}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{T} \end{bmatrix} \left.\frac{\partial \lambda_i}{\partial \mathbf{X}_\delta}\right|_{\mathbf{X}_\delta = \mathbf{X}_{\delta 0}} \begin{bmatrix} \mathbf{I}_q & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-T} \end{bmatrix}, \tag{40}$$

where $\left.\frac{\partial \lambda_i}{\partial \mathbf{X}_\delta}\right|_{\mathbf{X}_\delta = \mathbf{X}_{\delta 0}}$ is readily computed using theorem 1.

Let λ_i^0 be the ith eigenvalue of $\bar{\mathbf{A}}(\mathbf{X}_{\delta 0})$. Obviously, $\bar{\mathbf{A}}(\mathbf{X}_{\delta 0})$ and $\bar{\mathbf{A}}(\mathbf{X}_{\delta\mathbf{T}})$ have the identical eigenvalues, and the optimization problem (35) can be expressed as:

$$\mathbf{T}_{\text{opt}} \triangleq \arg \max_{\substack{\mathbf{T}\in\mathcal{R}^{n\times n} \\ \det(\mathbf{T})\neq 0}} \left(\min_{1\le i\le m+n} \frac{\frac{1}{h} - \left|\lambda_i^0 + \frac{1}{h}\right|}{\sum_{j=1}^{N} \left| \left.\frac{\partial \lambda_i}{\partial p_j}\right|_{\mathbf{X}_\delta = \mathbf{X}_{\delta\mathbf{T}}} \right|} \right). \tag{41}$$

Given $\mathbf{T}_{\text{opt}}$, the optimal controller realization $\mathbf{X}_{\delta\text{opt}}$ is readily computed using (38). For the complex-valued matrix $\mathbf{M} \in \mathcal{C}^{(l+n)\times(q+n)}$ with elements m_{ij}, define the matrix norm

$$\|M\|_F \triangleq \sum_{i=1}^{l+n}\sum_{j=1}^{q+n} |m_{ij}|. \tag{42}$$

The maximisation problem (41) is equivalent to the minimization problem

$$\begin{aligned} \mathbf{T}_{\text{opt}} &\triangleq \arg \min_{\substack{\mathbf{T}\in\mathcal{R}^{n\times n} \\ \det(\mathbf{T})\neq 0}} \left(\max_{1\le i\le m+n} \frac{\| \left.\frac{\partial \lambda_i}{\partial \mathbf{X}_\delta}\right|_{\mathbf{X}_\delta = \mathbf{X}_{\delta\mathbf{T}}} \|_F}{\frac{1}{h} - \left|\lambda_i^0 + \frac{1}{h}\right|} \right) \\ &= \arg \min_{\substack{\mathbf{T}\in\mathcal{R}^{n\times n} \\ \det(\mathbf{T})\neq 0}} \left(\max_{1\le i\le m+n} \left\| \begin{bmatrix} \mathbf{I}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{T} \end{bmatrix} \Phi_i \begin{bmatrix} \mathbf{I}_q & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-T} \end{bmatrix} \right\|_F \right), \end{aligned} \tag{43}$$

where

$$\Phi_i = \frac{\left.\frac{\partial \lambda_i}{\partial \mathbf{X}_\delta}\right|_{\mathbf{X}_\delta = \mathbf{X}_{\delta 0}}}{\frac{1}{h} - \left|\lambda_i^0 + \frac{1}{h}\right|} . \tag{44}$$

Thus finding an optimal controller realization is equivalent to obtaining a similarity transformation that is a solution of the following constrained nonlinear optimization problem

$$\mathbf{T}_{\text{opt}} = \arg \min_{\substack{\mathbf{T} \in \mathcal{R}^{n \times n} \\ \det(\mathbf{T}) \neq 0}} f_\delta(\mathbf{T}) \tag{45}$$

with the cost function

$$f_\delta(\mathbf{T}) = \max_{1 \le i \le m+n} \left\| \begin{bmatrix} \mathbf{I}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^T \end{bmatrix} \Phi_i \begin{bmatrix} \mathbf{I}_q & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-T} \end{bmatrix} \right\|_F . \tag{46}$$

Because the cost function $f_\delta(\mathbf{T})$ is non-smooth and non-convex, optimization must be based on a direct search without the aid of cost function derivatives. The conventional optimization methods for this kind of problem, such as Rosenbrock and Simplex algorithms [26]–[28], generally can only find a local minimum. Notice that, although the choice of initial realization will not affect the closed-loop eigenvalues, the eigenvalue sensitivities $\frac{\partial \lambda_i}{\partial \mathbf{X}_\delta}$ depend on the chosen initial realization. Thus for different $\mathbf{X}_{\delta 0}$ the shape of the cost function $f_\delta(\mathbf{T})$ will change, giving rise to different degree of difficulty in the optimization procedure. It is therefore important to use an efficient and preferably global optimization method. We adopt a global optimization strategy based on the ASA [19]–[23] to search for a true global optimum $\mathbf{T}_{\text{opt}}$. The detailed ASA optimizer applied to optimize the stability related measure in the z domain can be found in [12], where it is also shown how the constraint $\det(\mathbf{T}) \neq 0$ is dealt with during the optimization iteration.

6 Optimal realization of FWL PID controllers in the δ domain

In this section, we specifically discuss the optimal realization problem of FWL δ-based PID controllers. It is well-known that a constrained nonlinear optimization problem is generally much more difficult to solve than an unconstrained one. For the FWL z-based controller realization problem, the previous works have shown that the constrained optimization problem can be decoupled into two simpler unconstrained ones [11],[13].

This result can readily be extended to the case of FWL δ-based PID controller structures. A digital PID controller is an order $n = 2$ system. For notational simplicity, we will also restrict to the single-input and single-output controller, that is, $l = q = 1$. Let an initial realization for such a digital PID controller $C(\delta)$ be $(\mathbf{A}_c^0 \in \mathcal{R}^{2\times 2}, \mathbf{B}_c^0 \in \mathcal{R}^{2\times 1}, \mathbf{C}_c^0 \in \mathcal{R}^{1\times 2}, \mathbf{D}_c^0 \in \mathcal{R})$. From (43), the optimal PID controller realization problem is defined as the optimization problem:

$$\nu_\delta \triangleq \min_{\substack{\mathbf{T}\in\mathcal{R}^{2\times 2}\\ \det(\mathbf{T})\neq 0}} \left(\max_{1\le i\le m+2} \left\| \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^T \end{bmatrix} \Phi_i \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-T} \end{bmatrix} \right\|_F \right) . \tag{47}$$

The aim is to avoid handling the constraint $\det(\mathbf{T}) \neq 0$ directly in optimization. The following theorem shows that the optimization problem (47) can be solved by solving for the two simpler unconstrained problems. First define the two cost functions

$$f_{\delta 1}(x,y,w) = \max_{1\le i\le m+2} \left\| \begin{bmatrix} w & 0 & 0 \\ 0 & x & 0 \\ 0 & y & 1/x \end{bmatrix} \Phi_i \begin{bmatrix} 1/w & 0 & 0 \\ 0 & 1/x & 0 \\ 0 & -y & x \end{bmatrix} \right\|_F \tag{48}$$

and

$$f_{\delta 2}(x,y,u,w) = \max_{1\le i\le m+2} \left\| \begin{bmatrix} w & 0 & 0 \\ 0 & x & u \\ 0 & (xy-1)/u & y \end{bmatrix} \Phi_i \begin{bmatrix} 1/w & 0 & 0 \\ 0 & y & -u \\ 0 & (1-xy)/u & x \end{bmatrix} \right\|_F . \tag{49}$$

Theorem 2 Let

$$\nu_{\delta 1} = \min_{\substack{x\in(0,+\infty)\\ y\in(-\infty,+\infty)\\ w\in(0,+\infty)}} f_{\delta 1}(x,y,w) \tag{50}$$

and

$$\nu_{\delta 2} = \min_{\substack{x\in(-\infty,+\infty)\\ y\in(-\infty,+\infty)\\ u\in(0,+\infty)\\ w\in(0,+\infty)}} f_{\delta 2}(x,y,u,w) . \tag{51}$$

Then

$$\nu_\delta = \min\{\nu_{\delta 1}, \nu_{\delta 2}\} . \tag{52}$$

Moreover, if $\nu_\delta = \nu_{\delta 1}$ and $(x_{\text{opt1}}, y_{\text{opt1}}, w_{\text{opt1}})$ is the optimal solution of the problem (50), the optimal solution of the problem (47) is given as:

$$\mathbf{T}_{\text{opt}} = \frac{1}{w_{\text{opt1}}} \begin{bmatrix} x_{\text{opt1}} & y_{\text{opt1}} \\ 0 & 1/x_{\text{opt1}} \end{bmatrix} ; \tag{53}$$

if $\nu_\delta = \nu_{\delta 2}$ and $(x_{\text{opt2}}, y_{\text{opt2}}, u_{\text{opt2}}, w_{\text{opt2}})$ is the optimal solution of the problem (51), the optimal solution of the problem (47) is given as:

$$\mathbf{T}_{\text{opt}} = \frac{1}{w_{\text{opt2}}} \begin{bmatrix} x_{\text{opt2}} & (x_{\text{opt2}} y_{\text{opt2}} - 1)/u_{\text{opt2}} \\ u_{\text{opt2}} & y_{\text{opt2}} \end{bmatrix}. \tag{54}$$

The proof of theorem 2 is given in Appendix. Because $f_{\delta 1}(x, y, w)$ and $f_{\delta 2}(x, y, u, w)$ are still non-smooth and non-convex functions, an efficient global optimization method is preferred and we will adopt the ASA optimizer to solve for these two unconstrained nonlinear optimization problems.

7 Application examples

Two numerical examples were used to show how the optimization approach presented earlier can be used efficiently for designing optimal FWL δ-based controller structures. For the comparison purpose, both the z and δ based controllers were investigated in the simulation. The optimal realization problem of FWL z-based controller structures with the stability related measure $\mu_{z\,1}(\mathbf{X}_z)$ was defined in the previous works [6]–[13].

Example 1: We consider the following IFAC93 benchmark PID control system [29]. The continuous-time plant model is

$$P(s) = \frac{25(-0.4\,s + 1)}{(s^2 + 3\,s + 25)(5\,s + 1)} \tag{55}$$

and the designed PID controller is

$$C(s) = 1.311 + \frac{0.431}{s} + \frac{1.048\,s}{1 + 12.92\,s}. \tag{56}$$

The sampled-data system with the infinite-precision digital controller in z-domain is stable when the sampling period $h \le 2^3$. The range of the sampling period tested in the simulation was 2^3 to 2^{-12}, to cover the slow to very fast sampling conditions.

Given a sampling rate, the discrete-time plant model $P(\delta)$ and the digital controller $C(\delta)$ with the δ operator were obtained using the discretizing routines in MATLAB. The discretization procedure was based on the bilinear (Tustin) transformation

$$s = \frac{2}{h} \frac{1 - z^{-1}}{1 + z^{-1}} \tag{57}$$

with $\delta = \frac{z-1}{h}$. The initial realization $\mathbf{X}_{\delta 0}$ was chosen to be the "controllable" canonical form. When $\mathbf{X}_{\delta 0}$ was provided, the eigenvalues $\{\lambda_i^0\}$ of the ideal closed-loop system without FWL effects and the eigenvalue sensitivity matrices $\{\Phi_i\}$ were computed. The ASA was then used to search for an optimal transform matrix $\mathbf{T}_{\rm opt}$ by solving for the minimization problem (47) using theorem 2. This produced a corresponding optimal controller realization $\mathbf{X}_{\delta \rm opt}$ that maximizes the stability related measure $\mu_{\delta 1}(\mathbf{X}_\delta)$. The entire process was repeated with the z operator parameterization to obtain the optimal z-based realization $\mathbf{X}_{z\,\rm opt}$ that maximizes the stability related measure $\mu_{z\,1}(\mathbf{X}_z)$.

Fig. 2 shows the values of the FWL stability related measure $\mu_{\delta 1}$ given different sampling rates for the initial and optimal δ-based controller realizations $\mathbf{X}_{\delta 0}$ and $\mathbf{X}_{\delta \rm opt}$, respectively. It can be seen that, for this example, optimization achieved an improvement by more than an order of magnitude on the stability related measure. Fig. 3 (a) and (b) depict the estimated minimum bit lengths, $\hat{B}_{s1}^{\rm min}$, based only on the values of the stability related measure for $\mathbf{X}_{\delta 0}$ and $\mathbf{X}_{\delta \rm opt}$, respectively. As mentioned previously, for the δ operator parameterization, the sampling period h should be implemented exactly without FWL errors. Taking this into account, the modified estimate of the minimum bit length for the optimal realization $\mathbf{X}_{\delta \rm opt}$ is given in Fig. 3 (c).

Fig. 4 compares the FWL stability related measure for $\mathbf{X}_{\delta \rm opt}$ with that of the optimal z-operator controller realization $\mathbf{X}_{z\,\rm opt}$. It is seen that the optimal δ-based controller realization has much larger FWL closed-loop stability margin than its z-based counterpart. Furthermore, as the sampling rate is increasing, the stability related measure for $\mathbf{X}_{\delta \rm opt}$ is improving slightly and eventually leveling out while the stability related measure for $\mathbf{X}_{z\,\rm opt}$ is decreasing exponentially. This confirms with a well-known fact that the δ parameterization has significant advantages over the usual z parameterization, especially under fast sampling conditions. Fig. 5 gives the estimated minimal bit length for the optimal z-operator controller realization. Notice that it does not need to consider h separately in the z-operator parameterization, as the effect of h has already been included in the controller realization $\mathbf{X}_z$. Comparing Fig. 5 with Fig. 3 (c), even taking into account the requirement of implementing h exactly, the optimal δ-based realization requires a smaller bit length in FWL implementation than the optimal z-based realization.

Example 2: This example is the linearized model of a CH-47 tandem-rotor helicopter in horizontal motion about a nominal airspeed [30]. The continuous-time plant model $P(s)$ given by [30] is in the state-space form $(\mathbf{A}_s, \mathbf{B}_s, \mathbf{C}_s, \mathbf{D}_s)$ with

$$\mathbf{A}_s = \begin{bmatrix} -0.02 & 0.05 & 2.4 & -32 \\ -0.14 & 0.44 & -1.3 & -30 \\ 0 & 0.018 & -1.6 & 1.2 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{B}_s = \begin{bmatrix} 0.14 & -0.12 \\ 0.36 & -8.6 \\ 0.35 & 0.009 \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{C}_s = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 57.3 \end{bmatrix}, \quad \mathbf{D}_s = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{58}$$

A stabilizing continuous-time controller $C(s)$ was designed using the LQG method [6] and the controller $C(s)$ is given in the state-space form $(\mathbf{A}_t, \mathbf{B}_t, \mathbf{C}_t, \mathbf{D}_t)$ with

$$\mathbf{A}_t = \begin{bmatrix} -0.0175 & -0.1436 & 0.3852 & -26.3518 \\ 0.0084 & -17.6863 & -4.0536 & -13.9065 \\ 0.001 & 0.0018 & -6.7274 & -33.2584 \\ 0 & 0.0031 & 1 & -5.1191 \end{bmatrix}, \quad \mathbf{B}_t = \begin{bmatrix} 0.0158 & -0.2405 \\ 9.0660 & -0.1761 \\ 0.0091 & 0.2289 \\ -0.0031 & 0.0893 \end{bmatrix},$$

$$\mathbf{C}_t = \begin{bmatrix} 0.0033 & -0.0472 & -14.6421 & -60.8894 \\ -0.0171 & 1.0515 & -0.2927 & -3.2469 \end{bmatrix}, \quad \mathbf{D}_t = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{59}$$

The range of sampling rate used in the simulation was 2^2 to 2^{14}.

Using the generalized operator ρ to represent δ or z, depending on which operator is actually employed, the discrete-time plant model $P(\rho)$ and the discrete-time controller $C(\rho)$ were obtained for each given sampling rate using the discretizing routines in MATLAB. Because the version of MATLAB, which we have, does not have the discretizing routine that can provide the canonical state-space model for multi-input multi-output transfer functions, the initial controller realization $\mathbf{X}_{\rho 0}$ was chosen to be the non-canonical form as the result of a direct discretizing the state-space model of $C(s)$ given in (59). The ASA was used to find the optimal $\mathbf{T}_{\rm opt}$ and hence the optimal $\mathbf{X}_{\rho\, \rm opt}$ that maximizes the stability related measure $\mu_{\rho 1}(\mathbf{X}_\rho)$ for both $\rho = \delta$ and $\rho = z$.

Fig. 6 plots the FWL stability related measures as function of sampling rate for the initial and optimal δ-operator controller realizations $\mathbf{X}_{\delta 0}$ and $\mathbf{X}_{\delta \rm opt}$, where it can be seen that the optimization very effectively improves the FWL closed-loop stability robustness. Fig. 7 compares the FWL stability related measure for the δ-based optimal realization $\mathbf{X}_{\delta \rm opt}$ with that of the z-based optimal realization $\mathbf{X}_{z\, \rm opt}$. Again, as the sampling rate increases, the stability related measure for $\mathbf{X}_{z\, \rm opt}$ decreases exponentially while the stability related measure for $\mathbf{X}_{\delta\text{opt}}$ does not reduce, and the optimal δ-based controller realization has much better FWL closed-loop stability robustness than its z-based counterpart. Fig. 8 depicts the estimated minimum bit length, $\hat{B}_{s1}^{\min}$, based only on the value of $\mu_{\delta 1}$ for the optimal δ-operator realization $\mathbf{X}_{\delta\text{opt}}$ and the modified estimate of the minimum bit length, $\hat{B}_{s1h}^{\min}$, taking into account the sampling period h. The estimated minimum bit length for the z-operator realization $\mathbf{X}_{z\,\text{opt}}$ is given in Fig. 9. Again, $\mathbf{X}_{\delta\text{opt}}$ requires a smaller bit length to implement than $\mathbf{X}_{z\,\text{opt}}$.

8 Conclusions

The paper addresses the problem of digital controller structures realized using the δ operator and the relevant issues of closed-loop stability subject to FWL implementation. A tractable stability related measure, quantifying the robustness of closed-loop stability to the FWL effects in the δ domain, has been derived. It has been shown that the optimal realization problem of finite-precision δ-based digital controllers can be interpreted as a constrained nonlinear optimization problem. In particular, for δ-based PID controller realizations, the optimization can be decoupled into two unconstrained optimization problems. An efficient global optimization strategy based on the ASA has been adopted to solve for this FWL optimal controller realization problem in the δ domain.

Two numerical examples have been used to illustrate the optimal design procedure. The results obtained also demonstrate that the digital controllers described with the δ operator has much better FWL closed-loop stability robustness in fast sampling conditions, compared with the digital controllers described with the usual shift operator. In this work, the main emphasis has been focused on the important FWL closed-loop stability issues of sampled-data control systems. Ongoing work will explore the integration of the proposed optimization procedure with the closed-loop controller performance and the sparseness consideration of optimal controller realizations. This will provide a multi-objective framework to develop the optimal finite-precision controller realization that possesses the optimal trade off between minimum computational requirements, improved performance and stability robustness.

References

[1] P. Moroney, A.S. Willsky and P.K. Houpt, "The digital implementation of control compensators: the coefficient wordlength issue," *IEEE Trans. Automatic Control*, Vol.25, No.8, pp.621–630, 1980.

[2] M. Gevers and G. Li, *Parameterizations in Control, Estimation and Filtering Problems: Accuracy Aspects*. London: Springer Verlag, 1993.

[3] I.J. Fialho and T.T. Georgiou, "On stability and performance of sampled data systems subject to word length constraint," *IEEE Trans. Automatic Control*, Vol.39, No.12, pp.2476–2481, 1994.

[4] A.G. Madievski, B.D.O. Anderson and M. Gevers, "Optimum realizations of sampled-data controllers for FWL sensitivity minimization," *Automatica*, Vol.31, pp.367–379, 1995.

[5] G. Li, "On the structure of digital controllers with finite word length consideration," *IEEE Trans. Automatic Control*, Vol.43, pp.689–693, 1998.

[6] R.H. Istepanian, G. Li, J. Wu and J. Chu, "Analysis of sensitivity measures of finite-precision digital controller structures with closed-loop stability bounds," *IEE Proc. Control Theory and Applications*, Vol.145, No.5, pp.472–478, 1998.

[7] R.H. Istepanian, I. Pratt, R. Goodall and S. Jones, "Effect of fixed point parameterization on the performance of active suspension control systems," in *Proc. 13th IFAC World Congress* (San Francisco, USA), 1996, pp.291–295.

[8] R.H. Istepanian, J. Wu, J. Chu and J.F. Whidborne, "Maximizing lower bound stability measure of finite precision PID controller realizations by nonlinear programming," in *Proc. 1998 American Control Conf.* (Philadelphia, USA), 1998, pp.2596–2600.

[9] R.H. Istepanian, J. Wu, J.F. Whidborne, J. Yan and S.E. Salcudean, "Finite-word-length stability issues of teleoperation motion-scaling control system," in *Proc. UKACC Control'98* (Swansea, UK), 1998, pp.1676–1681.

[10] S. Chen, R.H. Istepanian and J. Wu, "Optimizing stability bounds of finite-precision PID controllers using adaptive simulated annealing," in *Proc. 1999 American Control Conf.* (San Diego, CA), 1999, pp.4316–4320.

[11] J. Wu, R.H. Istepanian and S. Chen, "Stability issues of finite precision controller structures for sampled data systems," *Int. J. Control*, Vol.72, No.15, pp.1331–1342, 1999.

[12] R.H. Istepanian, S. Chen, J. Wu and J.F. Whidborne, "Optimal finite-precision controller realization of sample data systems," *Int. J. Systems Science*, to appear, 1999.

[13] S. Chen, J. Wu, R.H. Istepanian and J. Chu, "Optimizing stability bounds of finite-precision PID controller structures," *IEEE Trans. Automatic Control*, Vol.44, No.11, pp.2149–2153, 1999.

[14] R.H. Middleton and G.C. Goodwin, *Digital Control and Estimation: A Unified Approach*. Englewood Cliffs, NJ: Prentice Hall, 1990.

[15] R.M. Goodall and B.J. Donoghue, "Very high sample rate digital filters using the δ operator," *IEE Proc. G*, Vol.140, pp.199–206,1993.

[16] Y. Sung and M. Kung, "Lower finite word-length effect on state space digital filter by δ operator realization," *Int. J. Electronics*, Vol.75, No.6, pp.1135–1141, 1993.

[17] G. Li and M. Gevers, "Comparative study of finite wordlength effects in shift and delta operator parameterizations," *IEEE Trans. Automatic Control*, Vol.38, No.5, pp.803–807, 1993.

[18] R.H. Istepanian, "Implementational issues for discrete PID algorithms using shift and delta operators parameterizations," in *Proc. 4th IFAC Workshop on Algorithms and Architectures for Real-Time Control* (Vilamoura, Portugal), 1997, pp.117–122.

[19] L. Ingber, "Simulated annealing: practice versus theory," *Mathematical and Computer Modelling*, Vol.18, No.11, pp.29–57, 1993.

[20] L. Ingber, "Adaptive simulated annealing (ASA): lessons learned," *J. Control and Cybernetics*, Vol.25, No.1, pp.33-54, 1996.

[21] B.E. Rosen, "Rotationally parameterized very fast simulated reannealing," submitted to *IEEE Trans. Neural Networks*, 1997.

[22] S. Chen, B.L. Luk and Y. Liu, "Application of adaptive simulated annealing to blind channel identification with HOC fitting," *Electronics Letters*, Vol.34, No.3, pp.234–235, 1998.

[23] S. Chen and B.L. Luk, "Adaptive simulated annealing for optimization in signal processing applications," *Signal Processing*, Vol.79, pp.117–128, 1999.

[24] C.P. Neuman, "Transformations between delta and forward shift operator transfer function models," *IEEE Trans. System Man and Cybernetics*, Vol.23, pp.295–296, 1993.

[25] C.P. Neuman, "Properties of the delta operator model of dynamic physical systems," *IEEE Trans. System Man and Cybernetics*, Vol.23, pp.296–301, 1993.

[26] J. Kowalik and M.R. Osborne, *Methods for Unconstrained Optimization Problems.* New York: Elsevier, 1968.

[27] G.S.G. Beveridge and R.S. Schechter, *Optimization: Theory and Practice.* McGraw-Hill, 1970.

[28] L.C.W. Dixon, *Nonlinear Optimisation.* London: The English Universities Press Ltd, 1972.

[29] J.F. Whidborne, G. Murad, D.W. Gu and I. Postlethwaite, "Robust control of an unknown plant – the IFAC93 benchmark," *Int. J. Control*, Vol.61, No.3, pp.589–640, 1995.

[30] P. Dorato, C. Abdullah and V. Cerone, *Linear-Quadratic Control: An Introduction.* New York: Prentice-Hall, 1995.

Appendix Proof of theorem 2

Define the diagonal matrix set: $\mathbf{U}_{\mathrm{diag}}(n) \triangleq \{\mathrm{diag}(u_1, u_2, \cdots, u_n) : u_i \in \{-1, 1\}, 1 \leq i \leq n\}$. From the definition (42),

Lemma 3 $\forall \mathbf{M} \in \mathcal{C}^{m\times n}$, $\mathbf{U}_1 \in \mathbf{U}_{\text{diag}}(m)$ and $\mathbf{U}_2 \in \mathbf{U}_{\text{diag}}(n)$,

$$\|\mathbf{U}_1\mathbf{M}\|_F = \|\mathbf{M}\|_F \text{ and } \|\mathbf{M}\mathbf{U}_2\|_F = \|\mathbf{M}\|_F \,.$$

Define the sets

$$\begin{aligned}
\mathcal{T}_0 &\triangleq \left\{\mathbf{T} = \begin{bmatrix} t_1 & t_2 \\ t_3 & t_4 \end{bmatrix} :\ t_1 \in \mathcal{R}, t_2 \in \mathcal{R}, t_3 \in \mathcal{R}, t_4 \in \mathcal{R}, t_1t_4 - t_2t_3 \neq 0\right\}, \\
\mathcal{T}_1 &\triangleq \left\{\mathbf{T} = \begin{bmatrix} t_1 & t_2 \\ 0 & t_4 \end{bmatrix} :\ t_1 \in \mathcal{R}, t_2 \in \mathcal{R}, t_4 \in \mathcal{R}, t_1t_4 \neq 0\right\}, \qquad (60) \\
\mathcal{T}_2 &\triangleq \left\{\mathbf{T} = \begin{bmatrix} t_1 & t_2 \\ t_3 & t_4 \end{bmatrix} :\ t_1 \in \mathcal{R}, t_2 \in \mathcal{R}, t_3 \in \mathcal{R}, t_4 \in \mathcal{R},\ t_3 \neq 0, t_1t_4 - t_2t_3 \neq 0\right\}.
\end{aligned}$$

Construct the optimization problems

$$\nu_{\delta 1} \triangleq \min_{\mathbf{T}\in\mathcal{T}_1} \max_{1\le i\le m+2} \left\| \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^T \end{bmatrix} \Phi_i \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-T} \end{bmatrix} \right\|_F \qquad (61)$$

and

$$\nu_{\delta 2} \triangleq \min_{\mathbf{T}\in\mathcal{T}_2} \max_{1\le i\le m+2} \left\| \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^T \end{bmatrix} \Phi_i \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-T} \end{bmatrix} \right\|_F . \qquad (62)$$

Obviously $\mathcal{T}_0 = \mathcal{T}_1 \cup \mathcal{T}_2$ and, therefore, $\nu_\delta = \min\{\nu_{\delta 1}, \nu_{\delta 2}\}$. Define the function

$$\text{sgn}(x) = \begin{cases} 1, & x \ge 0\,, \\ 0, & x < 0\,. \end{cases} \qquad (63)$$

Consider the optimization problem (61). $\forall \mathbf{T} \in \mathcal{T}_1$ and $\forall i \in \{1, \cdots, m+2\}$, utilizing lemma 3, we have:

$$\left\| \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^T \end{bmatrix} \Phi_i \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-T} \end{bmatrix} \right\|_F = \left\| \begin{bmatrix} 1 & 0 & 0 \\ 0 & t_1 & 0 \\ 0 & t_2 & t_4 \end{bmatrix} \Phi_i \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/t_1 & 0 \\ 0 & -t_2/(t_1t_4) & 1/t_4 \end{bmatrix} \right\|_F =$$

$$\left\| \begin{bmatrix} 1/\sqrt{|t_1t_4|} & 0 & 0 \\ 0 & \sqrt{|t_1/t_4|} & 0 \\ 0 & \text{sgn}(t_4)t_2/\sqrt{|t_1t_4|} & \sqrt{|t_4/t_1|} \end{bmatrix} \Phi_i \begin{bmatrix} \sqrt{|t_1t_4|} & 0 & 0 \\ 0 & \sqrt{|t_4/t_1|} & 0 \\ 0 & -\text{sgn}(t_4)t_2/\sqrt{|t_1t_4|} & \sqrt{|t_1/t_4|} \end{bmatrix} \right\|_F . \qquad (64)$$

Define

$$\begin{aligned}
x &= \sqrt{\frac{|t_1|}{|t_4|}} \in (0, +\infty)\,, \\
y &= \text{sgn}(t_4)\frac{t_2}{\sqrt{|t_1t_4|}} \in (-\infty, +\infty)\,, \qquad (65) \\
w &= \frac{1}{\sqrt{|t_1t_4|}} \in (0, +\infty)\,.
\end{aligned}$$

Then

$$f_{\delta 1}(x,y,w) \triangleq \max_{1\le i\le m+2}\left\|\begin{bmatrix} w & 0 & 0\\ 0 & x & 0\\ 0 & y & 1/x\end{bmatrix}\Phi_i\begin{bmatrix} 1/w & 0 & 0\\ 0 & 1/x & 0\\ 0 & -y & x\end{bmatrix}\right\|_F$$
$$= \max_{1\le i\le m+2}\left\|\begin{bmatrix} 1 & \mathbf{0}\\ \mathbf{0} & \mathbf{T}^T\end{bmatrix}\Phi_i\begin{bmatrix} 1 & \mathbf{0}\\ \mathbf{0} & \mathbf{T}^{-T}\end{bmatrix}\right\|_F, \tag{66}$$

and

$$\nu_{\delta 1} \triangleq \min_{\mathbf{T}\in\mathcal{T}_1}\max_{1\le i\le m+2}\left\|\begin{bmatrix} 1 & \mathbf{0}\\ \mathbf{0} & \mathbf{T}^T\end{bmatrix}\Phi_i\begin{bmatrix} 1 & \mathbf{0}\\ \mathbf{0} & \mathbf{T}^{-T}\end{bmatrix}\right\|_F = \min_{\substack{x\in(0,+\infty)\\ y\in(-\infty,+\infty)\\ w\in(0,+\infty)}} f_{\delta 1}(x,y,w)\,. \tag{67}$$

If $\nu_\delta = \nu_{\delta 1}$ and $(x_{\text{opt1}}, y_{\text{opt1}}, w_{\text{opt1}})$ is the solution of the optimization problem (67),

$$\nu_\delta = \nu_{\delta 1} = \max_{1\le i\le m+2}\left\|\begin{bmatrix} w_{\text{opt1}} & 0 & 0\\ 0 & x_{\text{opt1}} & 0\\ 0 & y_{\text{opt1}} & 1/x_{\text{opt1}}\end{bmatrix}\Phi_i\begin{bmatrix} 1/w_{\text{opt1}} & 0 & 0\\ 0 & 1/x_{\text{opt1}} & 0\\ 0 & -y_{\text{opt1}} & x_{\text{opt1}}\end{bmatrix}\right\|_F$$
$$= \max_{1\le i\le m+2}\left\|\frac{1}{w_{\text{opt1}}}\begin{bmatrix} w_{\text{opt1}} & 0 & 0\\ 0 & x_{\text{opt1}} & 0\\ 0 & y_{\text{opt1}} & 1/x_{\text{opt1}}\end{bmatrix}\Phi_i\begin{bmatrix} 1/w_{\text{opt1}} & 0 & 0\\ 0 & 1/x_{\text{opt1}} & 0\\ 0 & -y_{\text{opt1}} & x_{\text{opt1}}\end{bmatrix}w_{\text{opt1}}\right\|_F, \tag{68}$$

which means that

$$\mathbf{T}_{\text{opt}} = \frac{1}{w_{\text{opt1}}}\begin{bmatrix} x_{\text{opt1}} & y_{\text{opt1}}\\ 0 & 1/x_{\text{opt1}}\end{bmatrix} \tag{69}$$

is the optimal solution of the problem (47).

By considering (62) in a similar way, we can prove the rest of theorem 2.

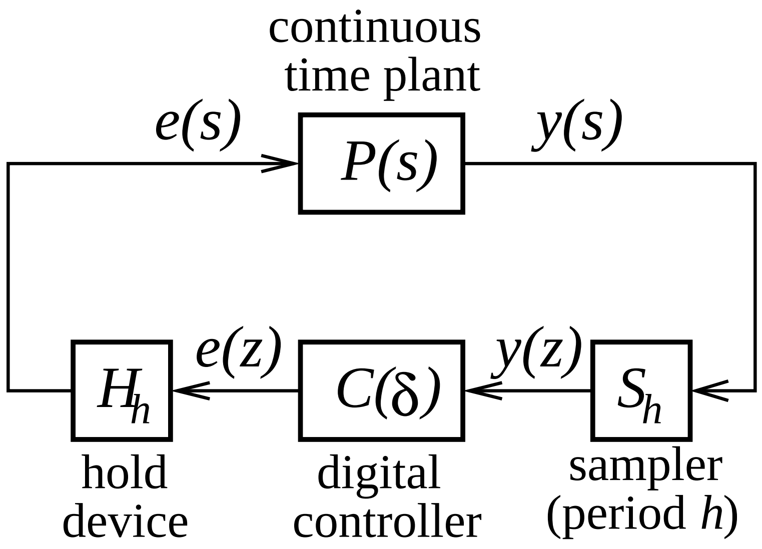


Figure 1: Sampled-data system with digital controller realization.

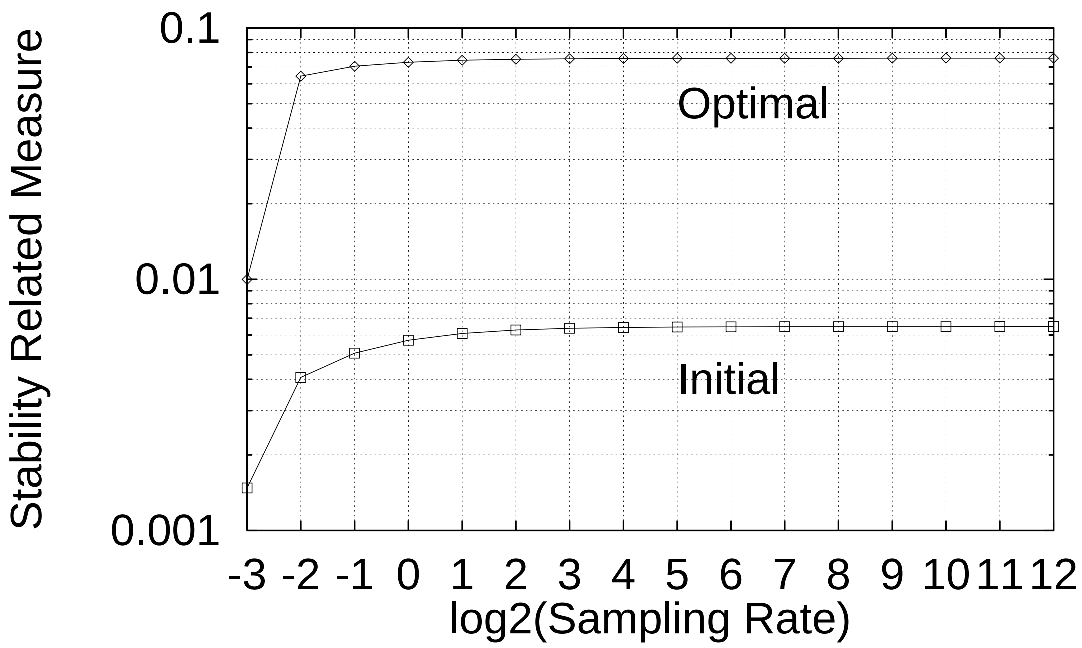
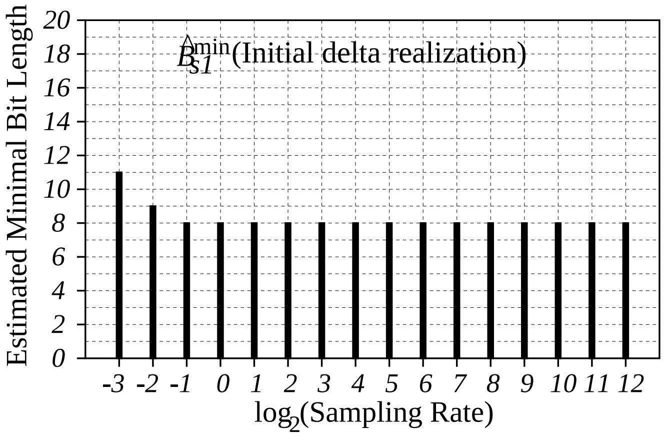
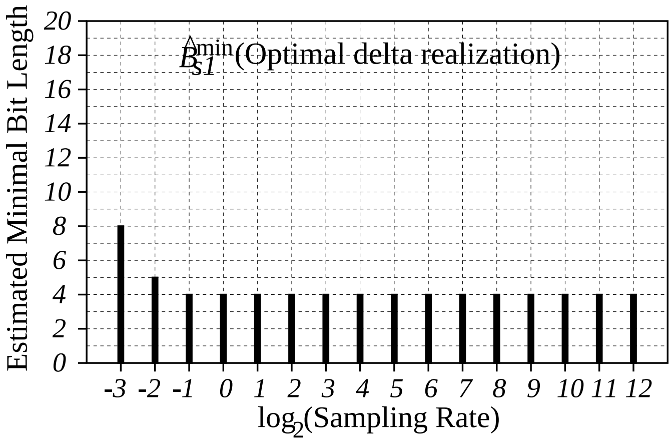


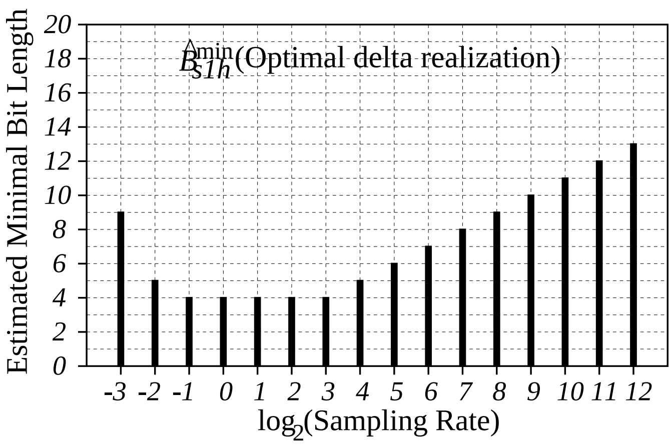
Figure 2: FWL stability related measure $\mu_{\delta 1}$ as a function of sampling rate for two different δ-based controller realizations. IFAC93 benchmark PID control system.



(a) $\hat{B}_{s1}^{\min}$ based on $\mu_{\delta 1}$ only for the initial realization $\mathbf{X}_{\delta 0}$



(b) $\hat{B}_{s1}^{\min}$ based on $\mu_{\delta 1}$ only for the optimal realization $\mathbf{X}_{\delta \mathrm{opt}}$



(c) $\hat{B}_{s1h}^{\min}$ based on $\mu_{\delta 1}$ and h for the optimal realization $\mathbf{X}_{\delta \mathrm{opt}}$

Figure 3: Estimated minimum bit lengths as a function of sampling rate for two different δ-based controller realizations. IFAC93 benchmark PID control system.

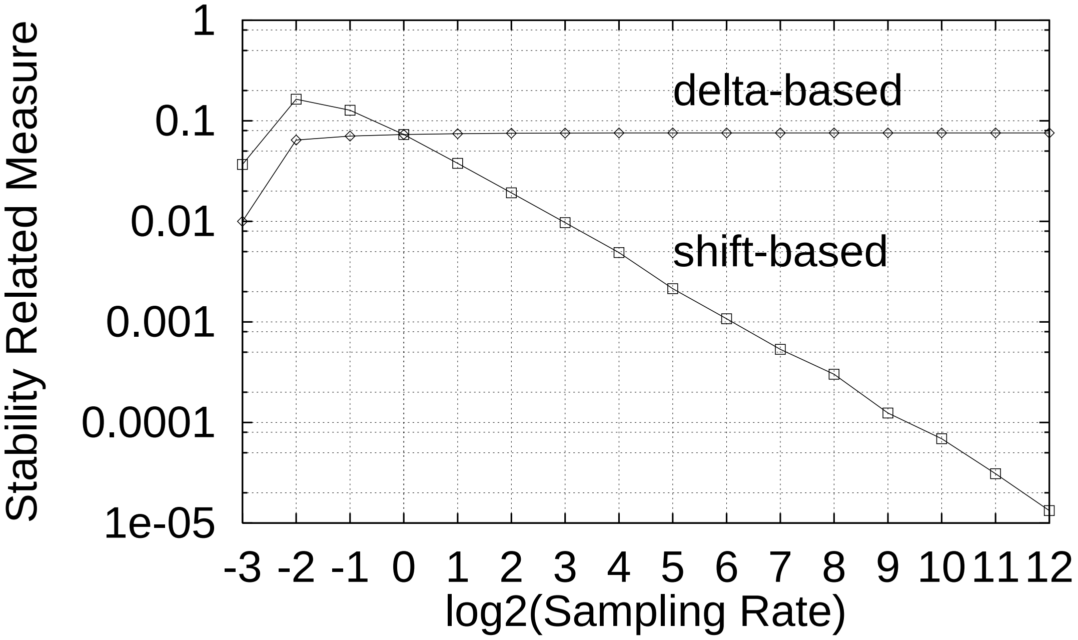


Figure 4: Comparison of FWL stability related measures for the optimal z-based and δ-based controller realizations. IFAC93 benchmark PID control system.

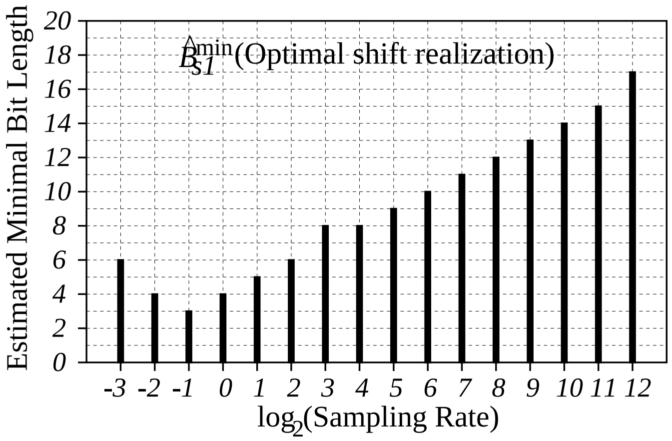


Figure 5: Estimated minimum bit length as a function of sampling rate for the optimal z-based controller realization. IFAC93 benchmark PID control system.

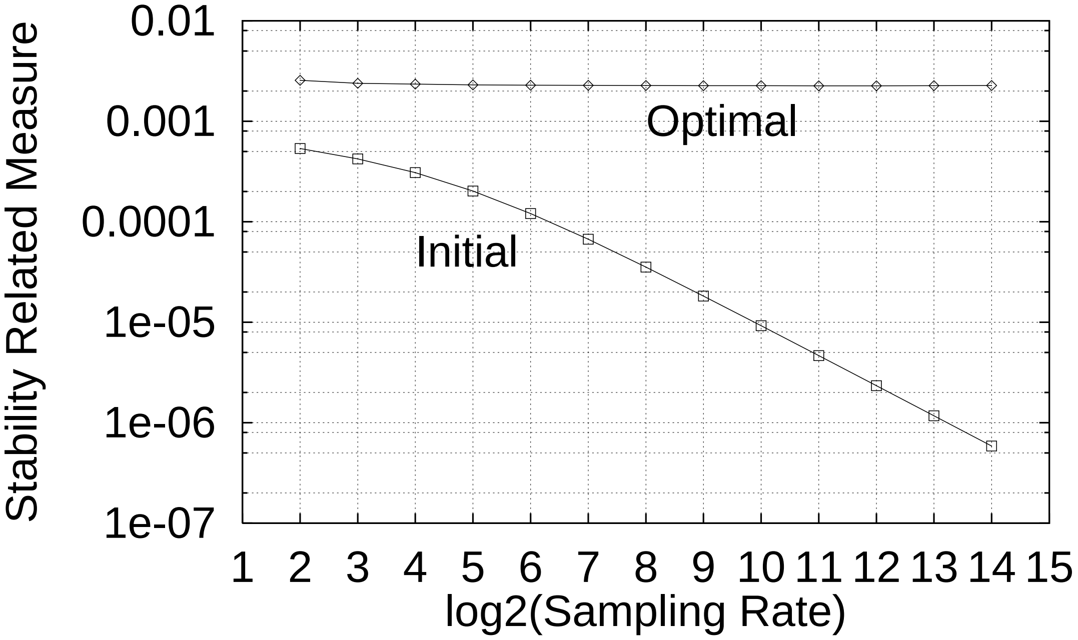


Figure 6: FWL stability related measure $\mu_{\delta 1}$ as a function of sampling rate for two different δ-based controller realizations. Helicopter control system example.

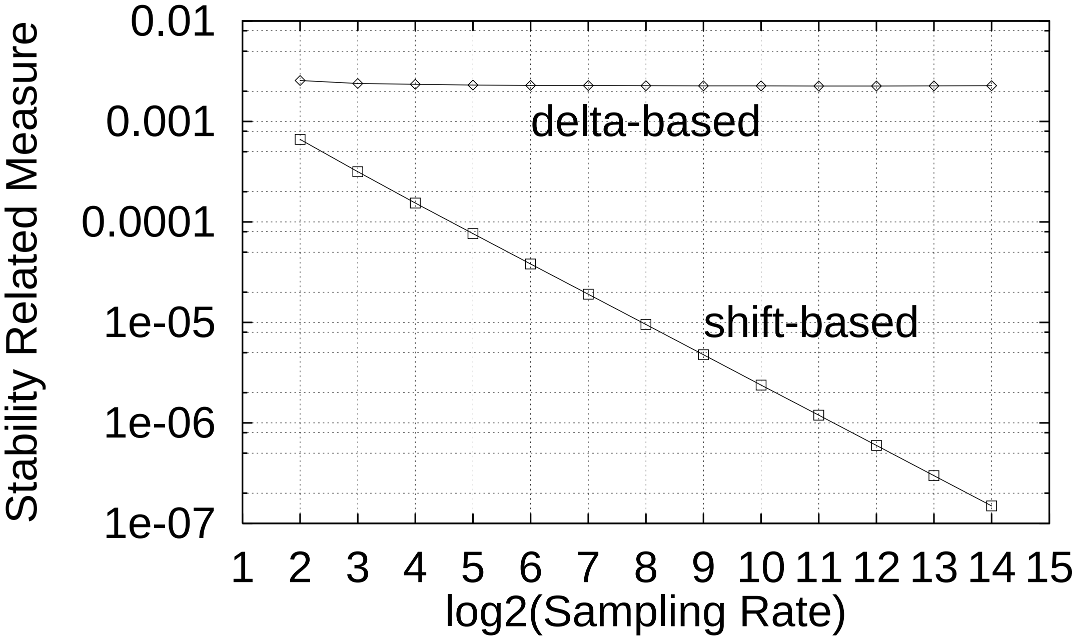
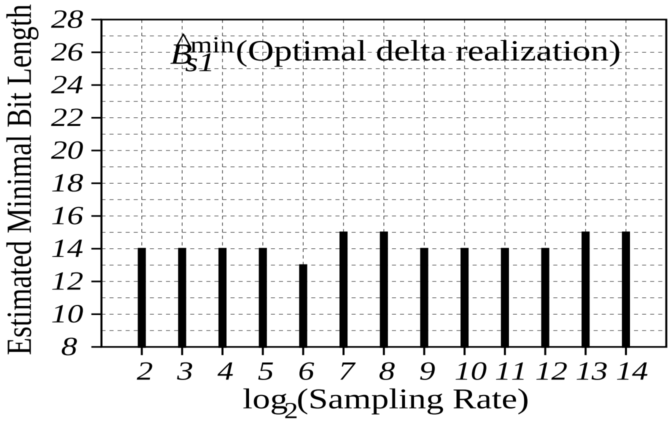
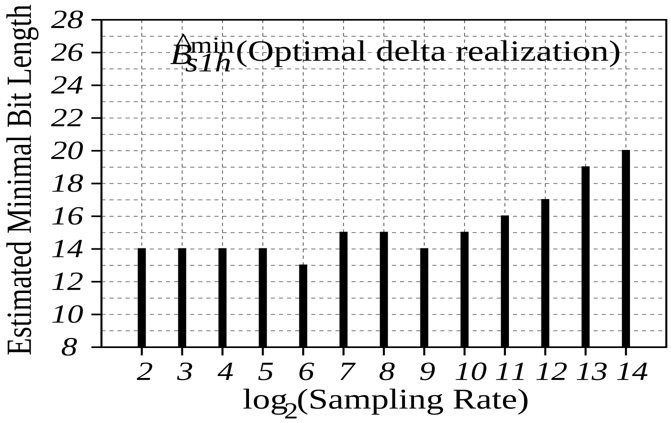


Figure 7: Comparison of FWL stability related measures for the optimal z-based and δ-based controller realizations. Helicopter control system example.



(a) $\hat{B}_{s1}^{\min}$ based on $\mu_{\delta 1}$ only for the optimal realization $\mathbf{X}_{\delta\mathrm{opt}}$



(b) $\hat{B}_{s1h}^{\min}$ based on $\mu_{\delta 1}$ and h for the optimal realization $\mathbf{X}_{\delta\mathrm{opt}}$

Figure 8: Estimated minimum bit lengths as a function of sampling rate for the optimal δ-based controller realization. Helicopter control system example.

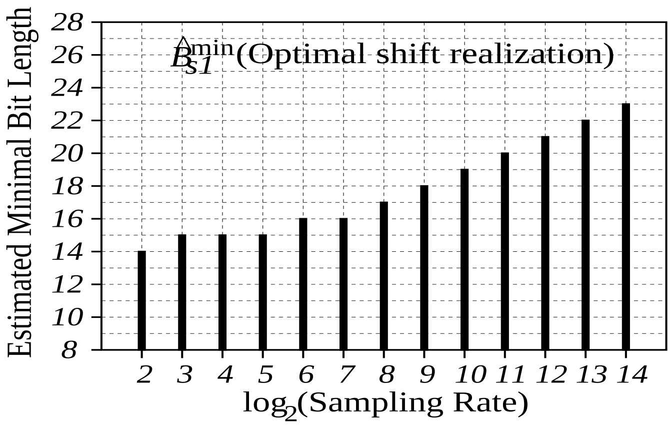


Figure 9: Estimated minimum bit length as a function of sampling rate for the optimal z-based controller realization. Helicopter control system example.