Computers and Chemical Engineering 106 (2017) 572–581

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

# On anti-aliasing filtering and over-sampling scheme in system identification☆

Mengyuan Fang[a], Yucai Zhu[a,*], Håkan Hjalmarsson[b]

[a] State Key Lab of Industry Control Technology, College of Control Science & Engineering, Zhejiang University, Hangzhou, China
[b] Department of Automatic Control, School of Electrical Engineering, KTH Royal Institute of Technology Sweden, Stockholm, Sweden

ARTICLE INFO

*Article history:*
Received 1 January 2017
Received in revised form 3 June 2017
Accepted 28 July 2017
Available online 12 August 2017

*Keywords:*
System identification
Sampling frequency
Anti-aliasing filtering
Variance
Over-sampling scheme

ABSTRACT

In this work, two fundamental issues in system identification, sampling frequency and anti-aliasing filtering are revisited, questions like, "Can higher sampling frequency help increase model accuracy?", and, "Can we do better than anti-aliasing filtering?" will be answered. First it is shown that, the traditional anti-aliasing filtering procedure does not give a consistent estimate of the desired model. Then an identification method is proposed based on the over-sampling scheme where a high frequency model is first identified and then converted to the (lower) control sampling frequency. It is shown that when the output noise contains energy beyond the bandwidth of the plant, the proposed method performs anti-aliasing in both open-loop and closed-loop identification; and in closed-loop identification, extra excitation is achieved from the high frequency noise. Simulation studies are used to illustrate the findings. The result has important implications in control applications.



## 1. Introduction

The basic requirement for system identification, especially in large-scale control applications, e.g. chemical process industries, is to obtain good quality models at a low cost. Noise aliasing is an issue that an identification experiment may have and will reduce the model quality. It happens when sampling the continuous-time process output that contains high frequency noise outside the frequency band of the plant (1/2 of the sampling frequency), according to the well-known Nyquist-Shannon sampling theorem. The noise aliasing involved in the output will then decrease the signal-to-noise ratio. Therefore, anti-aliasing measure is needed to avoid this noise aliasing.

The generally known way to avoid the noise aliasing is the anti-aliasing filtering procedure, see Ljung (1999). In this procedure, first the input and output are sampled at a high enough frequency to ensure the Nyquist sampling condition (the output noise contains no energy outside 1/2 of this sampling frequency) and give no chance to the noise aliasing; then the sampled signals are low-pass filtered to remove the frequency content beyond the frequency band of the plant; finally the filtered data are down-sampled to the needed sampling frequency for model estimation. Beside the anti-aliasing filtering procedure, the anti-aliasing problem is seldom discussed in the system identification literature. And, the anti-aliasing filtering procedure is lack of analysis, so we question that if this procedure is feasible in system identification.

On the other hand, to solve the anti-aliasing problem, we are inspired by the over-sampling scheme proposed in Sun et al. (1997), see also Sun and Sano (2009). The over-sampling scheme also begins with sampling the process data at a high frequency. The difference is that, then they identify a high frequency model from the high frequency sampled data without filtering and down-sampling the data; and they get the desired model by converting the high frequency model to the original low frequency. Wang et al. (2004b) gives solutions to the identification of multirate sampled-data system where the input and output are sampled at different sampling rates, this can be viewed as extension of the over-sampling problem. We are curious about the over-sampling scheme, can it achieve anti-aliasing by also using high-frequency sampling, or even do better than the anti-aliasing filtering procedure? And also, we would like to raise a fundamental question, can higher sampling frequency help increase model quality?

In Sun and co-workers' work (Sun and Sano, 2009; Sun and Zhu, 2012; Wang et al., 2004a), the over-sampling scheme is used in closed-loop system identification with no external excitation applied. They have proved, in different ways, that the

☆ This work is supported by the National Science Foundation of China (No. 61673343).

* Corresponding author.
*E-mail addresses:* myfang@iipc.zju.edu.cn (M. Fang), yczhu@iipc.zju.edu.cn (Y. Zhu), hakan.hjalmarsson@ee.kth.se (H. Hjalmarsson).

http://dx.doi.org/10.1016/j.compchemeng.2017.07.010
0098-1354/

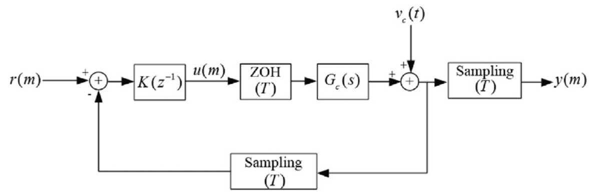


**Fig. 1.** A typical closed-loop system.

over-sampling scheme can achieve informative closed-loop identification tests even when there is no external excitation. However, they have not given the exact conditions on which the over-sampling scheme can work, while Wang et al. (2004a) points out that the performance of the over-sampling scheme is sensitive to many factors and sometimes the estimated model is of poor quality. Actually, external test signals are permitted in industrial applications, although they should be kept small in amplitudes. Therefore, in order to achieve high model quality and low disturbing test, we propose to use the over-sampling scheme with external excitation. We will further analyze the over-sampling scheme by using the asymptotic theory (Ljung, 1985) and propose an identification method based on the result. It will be surprisingly found that, the proposed method is a way that can both achieve anti-aliasing and make use of the high frequency noise as a part of excitation. We will also give the exact condition on which the good performance can be expected. And we point out it is also useful to apply the over-sampling scheme in open-loop identification.

Moreover, we would like to mention that the over-sampling operation is practically costless. In industrial MPC control systems, the controller sampling frequency is much lower than the sampling frequency of the lower level DCS (distributed control system) or PLC (programmable logic controller) systems. For example, in the refining/petrochemical industry, the sampling period of DCS systems is 1 s or smaller, while the sampling period of a typical MPC control system is 60 s. This means that in MPC identification, the sampling frequency can be made 60 times as high if necessary. The situations in other industries are similar. The redundancy of control systems sampling power is the foundation of the over-sampling scheme, but to the best of our knowledge, it has rarely been explored in the identification community.

The outline is as follows. Section 2 gives the problem statement. The anti-aliasing filtering procedure is described and analyzed in Section 3. In Section 4, the over-sampling scheme is introduced and further analyzed in frequency domain. The identification method based on the over-sampling scheme is proposed in Section 5. In Section 6, we compare the proposed method with the conventional identification. Some numerical examples are given in Section 7 to illustrate the results. Section 8 contains the conclusion.

## 2. Problem statement

In this work, we consider the identification of a closed-loop system depicted in Fig. 1. This system contains a linear continuous-time plant $G_c(s)$ and a linear digital controller $K(z^{-1})$. The control period is $T$, so is the sampling period in the feedback loop. The discrete-time control input $\{u(m)\}_{m=0,1,2,\ldots}$ is zero-order-held with holding time $T$ before applying to the continuous-time plant. $\{r(m)\}_{m=0,1,2,\ldots}$ is the reference signal with sample time $T$. $\{v_c(t)\}_{t\geq 0}$ is the continuous-time output noise.

The identification problem is to estimate the discretized model of $G_c(s)$ with respect to the sampling time $T$, denoted as $G_T(z^{-1})$. This model will be used in the following model based controller design. Although the problem statement is given with a closed-loop system, open-loop identification is also covered in our framework.

The model structure we consider is

$$G_T(z^{-1}) = \frac{B_T(z^{-1})}{A_T(z^{-1})} = \frac{b_{\tau_b}z^{-\tau_b} + \cdots + b_{n_b}z^{-n_b}}{1 + a_1z^{-1} + \cdots + a_{n_a}z^{-n_a}} \tag{1}$$

where $n_a$ and $n_b$ are the orders of $A_T(z^{-1})$ and $B_T(z^{-1})$, $\tau_b$ is the delay.

The main assumptions through this paper are listed here.

**Assumption 1** *(True process).* The true model is contained in the model set defined by (1). The order of the true model and its delay are known as a priori. The delay $\tau_b \geq 1$.

**Assumption 2** *(Noise).* The output noise $v_c(t)$ is a stationary stochastic signal and the sampled signals from it can be described by stable and inversely stable discrete-time transfer functions.

**Assumption 3** *(Input).* In closed-loop identification, the reference signal is independent of the output noise; in open-loop case, the input signal is independent of the output noise.

## 3. Anti-aliasing filtering

### 3.1. General

To do identification in the closed-loop system in Fig. 1, first an identification experiment is conducted: the external excitation is applied (generally on the reference signal $r(m)$); the input and output are observed for the model estimation. Conventionally, the input and output are sampled at sampling period $T$ because we need to estimate the model $G_T(z^{-1})$. In Fig. 1, the control input $u(m)$ can be used as the $T$-sampled input. The $T$-sampled output is denoted as $\{y(m)\}_{m=0,1,2,\ldots}$. $u(m)$ and $y(m)$ are then used to estimate the model $G_T(z^{-1})$:

$$y(m) = G_T(z^{-1})u(m) + v(m), \tag{2}$$

where $v(m)$ is the $T$-sampled signal from the continuous-time output noise $v_c(t)$.

However, there is a problem when sampling the data in this way: when the output noise $v_c(t)$ contains high frequency term outside the frequency band $[-\pi/T, \pi/T]$, the $T$-sampled output will contain aliasing from the high frequency noise, following from the Nyquist-Shannon sampling theorem. The noise aliasing will increase the model error in identification.

The anti-aliasing filtering procedure is aimed to avoid this kind of high frequency noise aliasing. When the following assumption holds,

**Assumption 4.** The spectrum of the output noise $v_c(t)$ is band limited and there exists a positive integer $p_0$ that the output noise $v_c(t)$ contains no frequency term outside the frequency band $[-p_0\pi/T, p_0\pi/T]$.

The procedure goes in three steps:

- Sample the input and output at a short enough sampling period $\Delta = T/p$, i.e. $p$ is a positive integer and $p > p_0$. Denote the $\Delta$-sampled input and output by $\{u_\Delta(k)\}_{k=0,1,2,\ldots}$ and $\{y_\Delta(k)\}_{k=0,1,2,\ldots}$, then $y_\Delta(k)$ will not have high frequency noise aliasing according to Assumption 4.
- Filter the $\Delta$-sampled data using a low-pass filter $F(z^{-1})$, of which the pass band is $[-\pi/p, \pi/p]$. Denote the filtered data by $\{u^f_\Delta(k)\}_{k=0,1,2,\ldots}$ and $\{y^f_\Delta(k)\}_{k=0,1,2,\ldots}$:

$$u^f_\Delta(k) = F(z^{-1})u_\Delta(k), \quad y^f_\Delta(k) = F(z^{-1})y_\Delta(k); \tag{3}$$

- Down-sample the filtered $\Delta$-sampled data at interval $p$. Since $T = p\Delta$, the sample time of the data generated in this step is $T$. Denote these data by $\{u^f(m)\}_{m=0,1,2,\ldots}$ and $\{y^f(m)\}_{m=0,1,2,\ldots}$: