

# Continuous-Time Model Identification and State Estimation Using Non-Uniformly Sampled Data

Rolf Johansson

Lund University, Dept Automatic Control,  
PO Box 118, SE 22100 Lund, Sweden,  
Tel: +46 462228791; Fax: +46 46138118;  
Email Rolf.Johansson@control.lth.se

**Abstract**—This contribution reviews theory, algorithms, and validation results for system identification of continuous-time state-space models from finite input-output sequences. The algorithms developed are autoregressive methods, methods of subspace-based model identification and stochastic realization adapted to the continuous-time context. The resulting model can be decomposed into an input-output model and a stochastic innovations model. Using the Riccati equation, we have designed a procedure to provide a reduced-order stochastic model that is minimal with respect to system order as well as the number of stochastic inputs, thereby avoiding several problems appearing in standard application of stochastic realization to the model validation problem. Next, theory, algorithms and validation results are presented for system identification of continuous-time state-space models from finite non-uniformly sampled input-output sequences. The algorithms developed are methods of model identification and stochastic realization adapted to the continuous-time model context using non-uniformly sampled input-output data. The resulting model can be decomposed into an input-output model and a stochastic innovations model. For state estimation dynamics and Kalman filters, we have designed a procedure to provide separate continuous-time temporal update and error feedback update based on non-uniformly sampled input-output data.

## I. INTRODUCTION

The accurate knowledge of a continuous-time transfer function is a prerequisite to many methods in physical modeling and control system design. System identification, however, is often made by means of time-series analysis applied to discrete-time transfer function models. As yet, there is no undisputed algorithm for parameter translation from discrete-time parameters to a continuous-time description. Problems in this context are associated with translation of the system zeros from the discrete-time model to the continuous-time model whereas the system poles are mapped by means of complex exponentials. As a result, a poor parameter translation tends to affect both the frequency response such as the Bode diagram and the transient response such as the impulse response. One source of error in many existing algorithms is that computation of the system zeros is affected by the assumed and actual inter-sample behavior of the control variables. Early contributions on continuous-time identification can be found in [29], [2], [30], [31], [8], [22], [23], [21], [6].

There are two circumstances, however, that favor the traditional indirect approach via discrete-time identification: Firstly, data are in general available as discrete measurements. Another problem is the mathematical difficulty to treat continuous-time random processes. In the context of discrete-time measurements, however, it is in many cases sufficient to model disturbances as a noise sequence of finite spectral range. A relevant question is, of course, why there is no analogue to ARMAX models for continuous-time systems. One reason is that polynomials in the differential operator can not be used for identification immediately due to the implementation problems associated with differentiation. The successful ARMAX-models correspond to transfer function polynomials in the  $z$ -transform variable  $z$  or  $z^{-1}$ —i.e., the forward or the backward shift operators, with advantages for modeling and signal processing, respectively, and translation between these two representations is not difficult. A related problem is how to identify accurate continuous-time transfer functions from data and, in particular, how to obtain good estimates of the zeros of a continuous-time transfer function. The difficulties to convert a discrete-time transfer function to continuous-time transfer function are well known and related to the mapping  $f(z) = (\log z)/h$ —for non-uniform sampling [3], [19].

We derive an algorithm that fits continuous-time transfer function models to discrete-time non-uniformly sampled data and we adopt a hybrid modeling approach by means of a discrete-time disturbance model and a continuous-time transfer function.

## II. A MODEL TRANSFORMATION

This algorithm introduces an algebraic reformulation of transfer function models. In addition, we introduce discrete-time noise models in order to model disturbances. The idea is to find a causal, stable, realizable linear operator that may replace the differential operator while keeping an exact transfer function. This shall be done in such a way that we obtain a linear model for estimation of the original transfer function parameters  $a_i, b_i$ . We will consider cases where we obtain a linear model in all-pass or low-pass filter operators. Actually, there is always a linear one-to-one transformation which relates the continuous-time

parameters and the convergence points for each choice of operator [10].

Then follows investigations on the state space properties of the introduced filters and the original model. The convergence rate of the parameter estimates is then considered. Finally, there are two examples with applications to time-invariant and time-varying systems, respectively. Consider a linear  $n$ th order transfer operator formulated with a differential operator  $p = d/dt$  and unknown coefficients  $a_i, b_i$ .

$$G_0(p) = \frac{b_1 p^{n-1} + \dots + b_n}{p^n + a_1 p^{n-1} + \dots + a_n} = \frac{B(p)}{A(p)} \quad (1)$$

where it is assumed that  $A$  and  $B$  are coprime. It is supposed that the usual isomorphism between transfer operators and transfer functions, *i.e.*, the corresponding functions of a complex variable  $s$ , is valid. Because of this isomorphism,  $G_0$  will sometimes be regarded as a transfer function and sometimes as a transfer operator. A notational difference will be made with  $p$  denoting the differential operator and  $s$  denoting the complex frequency variable of the Laplace transform.

On any transfer function describing a physically realizable continuous-time system, it is a necessary requirement that because pure derivatives of the input cannot be implemented. This requirement is fulfilled as  $\lim_{s \rightarrow \infty} G_0(s)$  is finite, *i.e.*,  $G_0(s)$  has no poles at infinity. An algebraic approach to system analysis may be suggested. Let  $a$  be point on the positive real axis and define the mapping

$$f(s) = \frac{a}{s+a}, \quad s \in \mathbb{C}$$

Let  $\bar{\mathbb{C}} = \mathbb{C} \cup \infty$  be the complex plane extended with the ‘infinity point’. Then  $f$  is a bijective mapping from  $\bar{\mathbb{C}}$  to  $\bar{\mathbb{C}}$  and it maps the ‘infinity point’ to the origin and  $-a$  to the ‘infinity point’. The unstable region—*i.e.*, the right half plane ( $\text{Re } s > 0$ )—is mapped onto a region which does not contain the ‘infinity point’. Introduction of the operator

$$\lambda = f(p) = \frac{a}{p+a} = \frac{1}{1+p\tau}, \quad \tau = 1/a \quad (2)$$

This allows us to make the following transformation

$$G_0(p) = \frac{B(p)}{A(p)} = \frac{B^*(\lambda)}{A^*(\lambda)} = G_0^*(\lambda)$$

with

$$A^*(\lambda) = 1 + \alpha_1 \lambda + \alpha_2 \lambda^2 + \dots + \alpha_n \lambda^n \quad (3)$$

$$B^*(\lambda) = \beta_1 \lambda + \beta_2 \lambda^2 + \dots + \beta_n \lambda^n \quad (4)$$

An input-output model is easily formulated as

$$A^*(\lambda)y(t) = B^*(\lambda)u(t) \quad (5)$$

or on regression form

$$y(t) = -\alpha_1[\lambda y](t) - \dots - \alpha_n[\lambda^n y](t) + \beta_1[\lambda u](t) + \dots + \beta_n[\lambda^n u](t) \quad (6)$$

This is now a linear model of a dynamical system at all points of time. Notice that  $[\lambda u], [\lambda y]$  etc. denote filtered inputs and outputs. The parameters  $\alpha_i, \beta_i$  may now be estimated by any suitable method for estimation of parameters of a linear model. A reformulation of the model (6) to a linear regression form is

$$\begin{aligned} y(t) &= \varphi_\tau^T(t)\theta_\tau, \\ \theta_\tau &= (\alpha_1 \quad \alpha_2 \dots \alpha_n \quad \beta_1 \quad \beta_2 \dots \beta_n)^T \\ \varphi_\tau(t) &= (-[\lambda y](t), \dots, -[\lambda^n y](t), \\ &\quad [\lambda u](t), \dots, [\lambda^n u](t))^T \end{aligned} \quad (7)$$

with parameter vector  $\theta_\tau$  and the regressor vector  $\varphi_\tau$ . We may now have the following continuous-time input-output relations:

$$\begin{aligned} y(t) &= G_0(p)u(t) = G_0^*(\lambda)u(t), \\ Y(s) &= \mathcal{L}\{y(t)\} = G_0^*(\lambda(s))U(s) \end{aligned} \quad (8)$$

$$\begin{aligned} y(t) &= \varphi_\tau^T(t)\theta_\tau \\ Y(s) &= \Phi_\tau^T(s)\theta_\tau \quad \text{where } \Phi_\tau(s) = \mathcal{L}\{\varphi_\tau(t)\}(s) \end{aligned} \quad (9)$$

where  $\mathcal{L}$  denotes a Laplace transform. As a consequence of the linearity of the Laplace transform, one can conclude that the same linear relation holds in both the time domain and the frequency domain. Notice that this property holds without any approximation or any selection of data.

#### A. Example—Estimation of two constant parameters

Consider the system with input  $u$ , output  $y$ , and the transfer operator  $G_0$

$$y(t) = G_0(p)u(t) = \frac{b_1}{p+a_1}u(t) \quad (10)$$

Use the operator transformation  $\lambda$  of (11) Use the operator transformation  $\lambda$  of (11)

$$\lambda = \frac{1}{1+p\tau} \quad (11)$$

This gives the transformed model

$$G_0^*(\lambda) = \frac{b_1\tau\lambda}{1+(a_1\tau-1)\lambda} = \frac{\beta_1\lambda}{1+\alpha_1\lambda} \quad (12)$$

A linear estimation model of the type (6) is given by

$$y(t) = -\alpha_1[\lambda y](t) + \beta_1[\lambda u](t) = \varphi_\tau^T(t)\theta_\tau(t) \quad (13)$$

with regressor  $\varphi_\tau(t)$  and the parameter vector  $\theta_\tau$  and

$$\varphi_\tau(t) = \begin{pmatrix} -[\lambda y](t) \\ [\lambda u](t) \end{pmatrix}, \quad \theta_\tau = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \quad (14)$$

The original parameters are found via the relations

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\tau}(\alpha_1 + 1) \\ \frac{1}{\tau}\beta_1 \end{pmatrix} \quad (15)$$

and their estimates from

$$\begin{pmatrix} \hat{a}_1 \\ \hat{b}_1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\tau}(\hat{\alpha}_1 + 1) \\ \frac{1}{\tau}\hat{\beta}_1 \end{pmatrix} \quad (16)$$

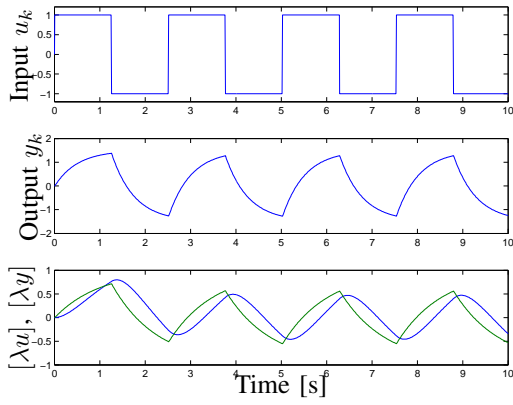


Fig. 1. Non-uniformly sampled data from simulation of continuous-time system of Example 1 with  $a_1 = 2$ ,  $b_1 = 3$  and continuous-time regressors for identification: Input  $u$  (upper), disturbance-free output  $y$  (middle), regressors  $[\lambda u]$ ,  $[\lambda y]$  (lower) for operator  $\lambda$  with  $\tau = 1$ .

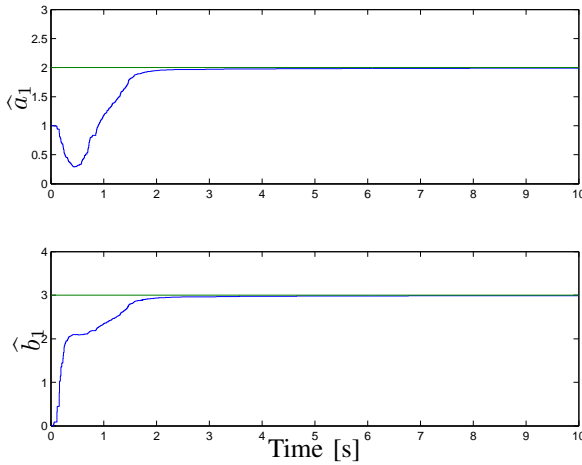


Fig. 2. Continuous-time model identification of Example 1 with  $a_1 = 2$ ,  $b_1 = 3$  and recursive least-squares identification using continuous-time regressors for input  $u$  and disturbance-free output  $y$ . The estimates  $\hat{a}_1$ ,  $\hat{b}_1$  converge to the correct values  $a_1 = 2$ ,  $b_1 = 3$  for  $N = 1000$  samples.

Sampling—*e.g.*, non-uniform sampling—of all variables in Eq. (13) and application of the recursive least-squares estimation algorithm is obviously possible.

The filter constant  $a$  (or  $\tau$ ) of the operator  $\lambda$  should be regarded as a design parameter to be chosen appropriately. As the components of the regressor vector  $\varphi_\tau$  tend to become small for high frequency input one should match the filter constant with respect to the dynamics of the system investigated.

*Remark—Operator Representation Singularities:* A relevant question is, of course, how general is the choice  $\lambda$  and if it can, for instance, be replaced by some other bijective mapping

$$\mu = \frac{bs + a}{s + a}, \quad b \in \mathbb{R}, \quad a \in \mathbb{R}^+, \quad \text{and} \quad s = \frac{\mu a - a}{b - \mu} \quad (17)$$

One can treat this problem by considering the example

$$G_0(s) = \frac{1}{s + a/b + \epsilon} \quad \text{where} \quad \epsilon \in \mathbb{R} \quad \text{is small}$$

Application of the operator translation  $\mu$  gives

$$\begin{aligned} G_0(s) &= \frac{1}{s + a/b + \epsilon} & (18) \\ &= \frac{1}{\mu - b - \epsilon b + (a(\frac{1}{b} - 1) + \epsilon)\mu} = G_0^*(\mu) & (19) \end{aligned}$$

Obviously, the zero-order denominator polynomial coefficient will vanish for  $\epsilon = 0$  so that  $G_0^*(\mu)$  exhibits a pole at  $z = 0$ . The estimation model would be

$$y = \alpha[\mu y] + \beta_1[\mu u] + \beta_0[u] \quad (20)$$

$$= \left( \frac{1}{\epsilon} \frac{a}{b} \left( \frac{1}{b} - 1 \right) + \frac{1}{b} \right) [\mu y] - \frac{1}{\epsilon b} [\mu u] + \frac{1}{\epsilon} [u] \quad (21)$$

which exhibits coefficients of very large magnitudes for small  $\epsilon$ . This would constitute a serious sensitivity problem—at least for  $b > 0$  for which  $G_0(s)$  is stable. An operator  $\mu$  with  $b < 0$  according to Eq. (17) would give rise to large coefficients of the transformed model only for unstable systems which might be more ‘affordable’. By comparison, a model transformation using  $\lambda$  would not exhibit any such singularities. Hence, use of the operator  $\mu$  should for sensitivity reasons be restricted to cases with  $b = 0$  (or  $b_{min} < b \leq 0$  for some number  $b_{min}$  chosen according to some *a priori* information about the system dynamics). Note that the set of polynomials associated with  $b < 0$  is related to the orthogonal Laguerre polynomials.

### B. Parameter transformations

Before proceeding, we should make clear the relationship between the parameters  $\alpha_i, \beta_i$  of (4) and the original parameters  $a_i, b_i$  of the transfer function (1). Let the vector of original parameters be denoted by

$$\theta = (-a_1 \quad -a_2 \quad \dots \quad -a_n \quad b_1 \quad \dots \quad b_n)^T \quad (22)$$

Using the definition of  $\lambda$  (11) and (11) it is straightforward to show that the relationship between operator-transformed parameters (7) and original parameters (22) is

$$\theta_\tau = F_\tau \theta + G_\tau \quad (23)$$

where the  $2n \times 2n$ -matrix  $F_\tau$  is

$$F_\tau = \begin{pmatrix} M_\tau & 0_{n \times n} \\ 0_{n \times n} & M_\tau \end{pmatrix} \quad (24)$$

and where  $M_\tau$  is the Pascal matrix

$$M_\tau = \begin{pmatrix} m_{11} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ m_{n1} & \dots & m_{nn} & \end{pmatrix}, \quad (25)$$

$$m_{ij} = (-1)^{i-j} \binom{n-j}{i-j} \tau^j \quad (26)$$

Furthermore, the  $2n \times 1$ -vector  $G_\tau$  are given by

$$G_\tau = (g_1 \dots g_n \quad 0 \dots 0)^T; \quad g_i = \binom{n}{i} (-1)^i \quad (27)$$

The matrix  $F_\tau$  is invertible when  $M_\tau$  is invertible, *i.e.* for all  $\tau > 0$ . The parameter transformation is then one-to-one and

$$\theta = F_\tau^{-1}(\theta_\tau - G_\tau) \quad (28)$$

We may then conclude that the parameters  $a_i, b_i$  of the continuous-time transfer function  $G_0$  may be reconstructed from the parameters  $\alpha_i, \beta_i$  of  $\theta_\tau$  by means of basic matrix calculations. As an alternative we may estimate the original parameters  $a_i, b_i$  of  $\theta$  from the linear relation

$$y(t) = \theta_\tau^T \varphi_\tau(t) = (F_\tau \theta + G_\tau)^T \varphi_\tau(t) \quad (29)$$

where  $F_\tau$  and  $G_\tau$  are known matrices for each  $\tau$ . Furthermore, elaborated identification algorithms adapted for numerical purposes sometimes contain some weighting or orthogonal linear combination of the regressor vector components by means of some linear transformation matrix  $T$ . Thus, one can modify (29) to

$$y(t) = (T\varphi_\tau(t))^T T^{-T} F_\tau \theta + (T\varphi_\tau(t))^T T^{-T} G_\tau$$

Hence, the parameter vectors  $\theta_\tau$  and  $\theta$  are related via known and simple linear relationships so that translation between the two parameter vectors can be made without any problem arising. Moreover, identification can be made with respect to either  $\theta$  or  $\theta_\tau$ .

### C. Orthogonalization and Numerics

In some cases, independent regressor variables are chosen at the onset of identification so as to be orthogonal in order to save computational effort. The case of polynomial regression provides an example of the use of orthogonalized independent variables. Orthogonal expansions are suitable and advantageous for of MA-type (FIR) and AR-type models but cannot be fully exploited in the case of regression vectors containing both input and output data, *e.g.* ARX and ARMAX models. A suitable approach to orthogonalization is to consider an asymptotically stable state space realization  $(A, B, C)$  by means of computation of the correlation between the impulse responses of the different state vector components  $x_i$ . As these impulse responses can be written  $x(t) = e^{At}B$  one finds that

$$P_c = \int_0^\infty x(t)x^T(t)dt = \int_0^\infty e^{At}BB^T e^{A^T t} dt \quad (30)$$

It is well known that  $P_c$  of (30) fulfills the Lyapunov equation

$$AP_c + P_c A^T = -BB^T \quad (31)$$

with a unique positive semidefinite solution  $P_c$ . If  $P_c$  were diagonal, one could conclude from Eq. (31) that the components of  $x(t)$  be orthogonal. To the purpose of orthogonalization, let  $P_c$  be factorized according to the Cholesky factorization

$$P_c = R_1^T R_1 \quad (32)$$

and choose  $T = R_1^{-T}$ . Now introduce the state-space transformation

$$z = Tx \quad (33)$$

with the dynamics

$$\dot{z} = A_1 z + B_1, \quad \text{where} \quad \begin{cases} A_1 = TAT^{-1} \\ B_1 = TB \end{cases} \quad (34)$$

By means of (31) it can be verified that

$$A_1 + A_1^T = -B_1 B_1^T, \quad P_z = \int_0^\infty z(t)z^T(t)dt = I$$

which implies that the components of  $z = Tx$  are mutually orthogonal over the interval  $[0, \infty)$ . As  $P_z$  is diagonal, it is clear that the components of  $z(t)$  be orthogonal.

Thus, the  $\lambda$ -operator is still effective for higher-order systems, the upper order limit of application being determined by the quality of the numerical solvers of Lyapunov equations and numerical Cholesky factorization.

*Example—Orthogonalization:* Let  $\lambda(s) = 1/(s+1)$  and consider the third order state-space realization

$$\begin{aligned} X(s) &= \begin{pmatrix} \lambda(s)U(s) \\ \lambda^2(s)U(s) \\ \lambda^3(s)U(s) \end{pmatrix} \\ &= \begin{pmatrix} s+1 & 0 & 0 \\ -1 & s+1 & 0 \\ 0 & -1 & s+1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} U(s) \end{aligned} \quad (35)$$

The solution of the Lyapunov equation (31) gives

$$P_c = \begin{pmatrix} 0.5000 & 0.2500 & 0.1250 \\ 0.2500 & 0.2500 & 0.1875 \\ 0.1250 & 0.1875 & 0.1875 \end{pmatrix} > 0 \quad (36)$$

so that

$$T = \begin{pmatrix} 1.0000 & 0 & 0 \\ -1.0000 & 2.0000 & 0 \\ 1.0000 & -4.0000 & 4.0000 \end{pmatrix} \quad (37)$$

As a result of orthogonality there is no need to solve a set of linear equation for MA-models with white-noise inputs. This might be advantageous for special purpose hardware implementation or when the linear equations become nearly singular or when the model order increases.

### D. Non-uniform Sampling

Assume that data acquisition has provided finite sequences of non-uniformly sampled input-output data  $\{y(t_k)\}_0^N, \{u(t_k)\}_0^N$  at sample times  $\{t_k\}_0^N$ , where  $t_{k+1} > t_k$  for all  $k$ .

As the regression model of Eq. (6) is valid for all times, it is also a valid regression model at sample times  $\{t_k\}_0^N$

$$\begin{aligned} y(t_k) &= -\alpha_1[\lambda y](t_k) - \dots - \alpha_n[\lambda^n y](t_k) \\ &\quad + \beta_1[\lambda u](t_k) + \dots + \beta_n[\lambda^n u](t_k) \end{aligned} \quad (38)$$

Introduce the following brief notation for non-uniformly sampled filtered data

$$[\lambda^j u]_k = [\lambda^j u](t_k), \quad 0 \leq j \leq n, \quad 0 \leq k \leq N \quad (39)$$

$$[\lambda^j y]_k = [\lambda^j y](t_k) \quad (40)$$

so that

$$y_k = -\alpha_1 [\lambda y]_k - \dots - \alpha_n [\lambda^n y]_k + \beta_1 [\lambda u]_k + \dots + \beta_n [\lambda^n u]_k \quad (41)$$

Introduce the regressor-state dynamics

$$x_u = \begin{pmatrix} [\lambda^1 u] \\ [\lambda^2 u] \\ \vdots \\ [\lambda^n u] \end{pmatrix}, \quad x_y = \begin{pmatrix} [\lambda^1 y] \\ [\lambda^2 y] \\ \vdots \\ [\lambda^n y] \end{pmatrix} \quad (42)$$

with dynamics

$$\frac{1}{\tau} \dot{x}_u = \begin{pmatrix} -1 & 0 & 0 & \dots & 0 \\ 1 & -1 & \ddots & \ddots & 0 \\ 0 & 1 & -1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -1 \end{pmatrix} x_u + \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} u \quad (43)$$

$$\frac{1}{\tau} \dot{x}_y = \begin{pmatrix} -1 & 0 & 0 & \dots & 0 \\ 1 & -1 & \ddots & \ddots & 0 \\ 0 & 1 & -1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -1 \end{pmatrix} x_y + \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} y \quad (44)$$

or

$$\frac{1}{\tau} \dot{x}_u = A_\lambda x_u + B_\lambda u, \quad \frac{1}{\tau} \dot{x}_y = A_\lambda x_y + B_\lambda y \quad (45)$$

Adopting a zero-order-hold (ZOH) approximation, the non-uniformly sampled discretized model will be

$$x_u(t_{k+1}) = A_k x_u(t_k) + B_k u(t_k) \quad (46)$$

$$x_y(t_{k+1}) = A_k x_y(t_k) + B_k y(t_k) \quad (47)$$

where

$$A_k = e^{A_\lambda(t_{k+1}-t_k)/\tau} \quad (48)$$

$$B_k = \int_0^{(t_{k+1}-t_k)/\tau} e^{A_\lambda s} B_\lambda ds \quad (49)$$

Summarizing the regressor model of Eq. (41) including the regressor filtering, we have

$$\phi(t_k) = \begin{pmatrix} x_y(t_k) \\ x_u(t_k) \end{pmatrix} \quad (50)$$

$$\phi(t_{k+1}) = \begin{pmatrix} A_k & 0 \\ 0 & A_k \end{pmatrix} \phi(t_k) \quad (51)$$

$$+ \begin{pmatrix} -B_k & 0 \\ 0 & B_k \end{pmatrix} \begin{pmatrix} y(t_k) \\ u(t_k) \end{pmatrix} \quad (52)$$

$$\theta = (\alpha_1 \ \dots \ \alpha_n \ \beta_1 \ \dots \ \beta_n)^T \quad (53)$$

$$y(t_k) = \phi(t_k) \theta + w(t_k) \quad (54)$$

where  $\{w(t_k)\}$  represents an uncorrelated non-uniformly sampled noise sequence.

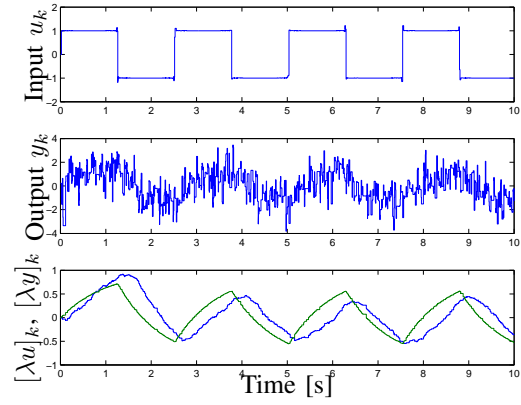


Fig. 3. Non-uniformly sampled data used for continuous-time model identification: Input  $\{u_k\}$  (upper), output  $\{y_k\}$  with stochastic disturbance (middle), regressors  $\{[\lambda u]_k\}$ ,  $\{[\lambda y]_k\}$  (lower).

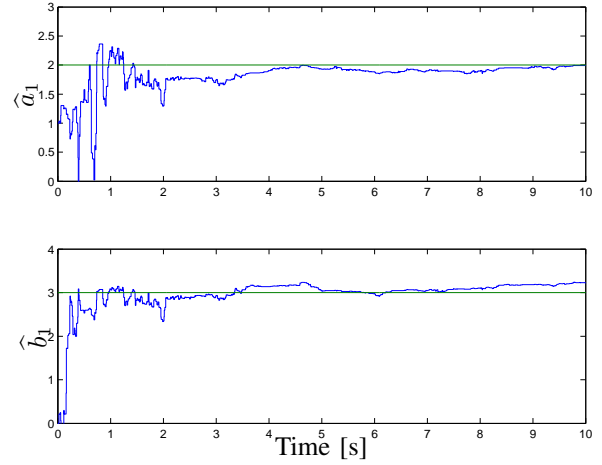


Fig. 4. Continuous-time model identification of Example 1 with  $a_1 = 2$ ,  $b_1 = 3$  and recursive least-squares identification using non-uniformly sampled input  $u$  and disturbance-contaminated output  $y$ . The estimates  $\hat{a}_1$ ,  $\hat{b}_1$  converge towards the correct values  $a_1 = 2$ ,  $b_1 = 3$  ( $N = 1000$ ).

### E. Example (cont'd)—A first-order system

A simulated example of Ex. II-A is shown in Figs. 1-4 for parameters  $a_1 = 2$ ,  $b_1 = 3$  and with operator time constant  $\tau = 1$ . A histogram of the sampling intervals is shown in Fig. 5. Whereas a least-squares estimate based on the  $N = 1000$  deterministic data of Figs. 1-2 reproduced the exact parameters  $a_1 = 2$ ,  $b_1 = 3$ , the estimates  $\hat{a}_1 = 1.988$ ,  $\hat{b}_1 = 3.168$  were obtained for a signal-to-noise ratio equal to one of inputs (input  $u$  and noise  $w$ ) in Figs. 3-4.

### F. Innovations Model and Prediction

Adopting a standard continuous-time innovations model to complement the system model of Eq. (1), we have

$$\dot{x} = Ax + Bu + Kw, \quad G_0(s) = C(sI - A)^{-1}B \quad (55)$$

$$y = Cx + w \quad (56)$$

with the innovations model inverse (or Kalman filter)

$$\hat{x} = (A - KC)\hat{x} + Bu + Ky \quad (57)$$

$$\hat{w} = y - C\hat{x} \quad (58)$$

Updates for non-uniformly sampled input-output data can be made as the non-uniformly sampled discrete-time system

$$\hat{x}(t_{k+1}) = F_k \hat{x}(t_k) + \begin{pmatrix} G_k & H_k \end{pmatrix} \begin{pmatrix} u(t_k) \\ y(t_k) \end{pmatrix} \quad (59)$$

$$\hat{w}(t_k) = y(t_k) - C\hat{x}(t_k) \quad (60)$$

for

$$F_k = e^{(A-KC)(t_{k+1}-t_k)} \quad (61)$$

$$G_k = \int_0^{t_{k+1}-t_k} e^{(A-KC)s} B ds \quad (62)$$

$$H_k = \int_0^{t_{k+1}-t_k} e^{(A-KC)s} K ds \quad (63)$$

which permits state estimation and standard residual analysis for purposes of validation [9, Ch.9]. In prediction-correction format, separate time update and error update can be made as follows

$$\hat{x}(t_{k|k}) = \hat{x}(t_{k|k-1}) + \kappa_k (y(t_k) - C\hat{x}(t_{k|k-1})) \quad (64)$$

$$\hat{x}(t_{k+1|k}) = \Phi_k \hat{x}(t_{k|k}) + \Gamma_k u(t_k) \quad (65)$$

$$\hat{w}(t_{k|k}) = y(t_k) - C\hat{x}(t_{k|k}) \quad (66)$$

for  $\{\Phi_k, \Gamma_k, \kappa_k\}_{k=0}^N$  obtained from non-uniform discretization of Eqs. (55-56).

### III. STATE-SPACE MODEL IDENTIFICATION

Consider a continuous-time time-invariant system  $\Sigma_n(A, B, C, D)$  with the state-space equations

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + v(t) \\ y(t) &= Cx(t) + Du(t) + e(t) \end{aligned} \quad (67)$$

with input  $u \in \mathbb{R}^m$ , output  $y \in \mathbb{R}^p$ , state vector  $x \in \mathbb{R}^n$  and zero-mean disturbance stochastic processes  $v \in \mathbb{R}^n$ ,  $e \in \mathbb{R}^p$  acting on the state dynamics and the output, respectively. The continuous-time system identification problem is to find estimates of system matrices  $A, B, C, D$  from finite sequences  $\{u_k\}_{k=0}^N$  and  $\{y_k\}_{k=0}^N$  of input-output data. The underlying discretized state sequence  $\{x_k\}_{k=0}^N$  and discrete-time stochastic processes  $\{v_k\}_{k=0}^N, \{e_k\}_{k=0}^N$  correspond to disturbance processes  $v$  and  $e$  which can be represented by the components

$$v_k = \int_{t_{k-1}}^{t_k} e^{A(t_k-s)} v(s) ds, \quad k = 1, 2, \dots, N \quad (68)$$

$$e_k = e(t_k) \quad (69)$$

with the covariance  $Q \geq 0$ ,  $q = \text{rank}(Q)$

$$\mathcal{E} \left\{ \begin{bmatrix} v_i \\ e_j \end{bmatrix} \begin{bmatrix} v_j \\ e_j \end{bmatrix}^T \right\} = Q \delta_{ij} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} \delta_{ij}, \quad (70)$$

Consider a discrete-time time-invariant system  $\Sigma_n(A, B, C, D)$  with the state-space equations with input  $u_k \in \mathbb{R}^m$ , output  $y_k \in \mathbb{R}^p$ , state vector  $x_k \in \mathbb{R}^n$

and noise sequences  $v_k \in \mathbb{R}^n$ ,  $e_k \in \mathbb{R}^p$  acting on the state dynamics and the output, respectively.

**Remark:** As computation and statistical tests deal with discrete-time data, we assume the original sampled stochastic disturbance sequences to be uncorrelated with a uniform spectrum up to the Nyquist frequency, thereby avoiding the mathematical problems associated with Brownian motion [10].

#### Continuous-Time State-Space Linear System

From the set of first-order linear differential equations of Eq. (67) one finds the Laplace transform

$$\begin{aligned} sX &= AX + BU + V + sx_0, \quad x_0 = x(t_0) \\ Y &= CX + DU + E \end{aligned} \quad (71)$$

Introduction of the complex variable transform

$$\lambda(s) = \frac{1}{1 + s\tau} \quad (72)$$

corresponding to a stable, causal operator permits an algebraic transformation of the model

$$\begin{aligned} X &= (I + \tau A)[\lambda X] + \tau B[\lambda U] + \tau[\lambda V] + (1 - \lambda)x_0 \\ Y &= CX + DU + E \end{aligned} \quad (73)$$

Reformulation while ignoring the initial conditions to linear system equations gives

$$\begin{aligned} \begin{bmatrix} \xi \\ y \end{bmatrix} &= \begin{bmatrix} I + \tau A & \tau B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} \tau v \\ e \end{bmatrix}, \quad x(t) = [\lambda \xi](t) \\ &= \begin{bmatrix} A_\lambda & B_\lambda \\ C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} \tau v \\ e \end{bmatrix}, \quad \begin{cases} A_\lambda = I + \tau A \\ B_\lambda = \tau B \end{cases} \end{aligned} \quad (74)$$

the mapping between  $(A, B)$  and  $(A_\lambda, B_\lambda)$  being bijective. Provided that a standard positive semi-definiteness condition of  $Q$  is fulfilled so that the Riccati equation has a solution, it is possible to replace the linear model of Eq. (74) by the innovations model

$$\begin{bmatrix} \xi \\ y \end{bmatrix} = \begin{bmatrix} A_\lambda & B_\lambda \\ C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} K_\lambda \\ I \end{bmatrix} w, \quad K_\lambda = \tau K \quad (75)$$

By recursion it is found that

$$\begin{aligned} y &= Cx + Du + w \quad (76) \\ &= CA_\lambda[\lambda x] + CB_\lambda[\lambda u] + Du + CK_\lambda[\lambda w] + w \\ &\vdots \\ &= CA_\lambda^k[\lambda^k x] + \sum_{j=1}^k CA_\lambda^{k-j} B_\lambda[\lambda^{k-j} u] + Du \\ &\quad + \sum_{j=1}^k CA_\lambda^{k-j} K_\lambda[\lambda^{k-j} w] + w \end{aligned} \quad (77)$$

To the purpose of subspace model identification, it is straightforward to formulate extended linear models for the original models and its innovations form

$$\mathcal{Y} = \Gamma_x \mathcal{X} + \Gamma_u \mathcal{U} + \Gamma_v \mathcal{V} + \mathcal{E} \quad (78)$$

$$\mathcal{Y} = \Gamma_x \mathcal{X} + \Gamma_u \mathcal{U} + \Gamma_w \mathcal{W} \quad (79)$$

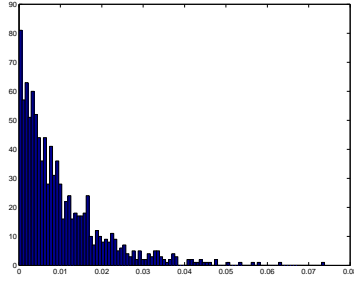


Fig. 5. Histogram of sampling intervals

with state variables  $\mathcal{X} = [\lambda^{i-1}x]$  and input-output variables

$$\mathcal{Y} = \begin{bmatrix} [\lambda^{i-1}y] \\ [\lambda^{i-2}y] \\ \vdots \\ [\lambda^1 y] \\ y(t) \end{bmatrix}, \mathcal{U} = \begin{bmatrix} [\lambda^{i-1}u] \\ [\lambda^{i-2}u] \\ \vdots \\ [\lambda^1 u] \\ u(t) \end{bmatrix}, \quad (80)$$

and stochastic processes of disturbance

$$\mathcal{V} = \begin{bmatrix} [\lambda^{i-1}v] \\ [\lambda^{i-2}v] \\ \vdots \\ [\lambda^1 v] \\ v(t) \end{bmatrix}, \mathcal{E} = \begin{bmatrix} [\lambda^{i-1}e] \\ [\lambda^{i-2}e] \\ \vdots \\ [\lambda^1 e] \\ e(t) \end{bmatrix}, \mathcal{W} = \begin{bmatrix} [\lambda^{i-1}w] \\ [\lambda^{i-2}w] \\ \vdots \\ [\lambda^1 w] \\ w(t) \end{bmatrix} \quad (81)$$

and parameter matrices of state variables and input-output behavior

$$\Gamma_x = \begin{bmatrix} C \\ CA_\lambda \\ \vdots \\ CA_\lambda^{i-1} \end{bmatrix} \in \mathbb{R}^{ip \times n} \quad (82)$$

$$\Gamma_u = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB_\lambda & D & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ CA_\lambda^{i-2}B_\lambda & CA_\lambda^{i-3}B_\lambda & \cdots & D \end{bmatrix} \in \mathbb{R}^{ip \times im}$$

and for stochastic input-output behavior

$$\Gamma_v = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \tau C & 0 & \cdots & 0 & 0 \\ \tau CA_\lambda & \tau C & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ \tau CA_\lambda^{i-2} & \tau CA_\lambda^{i-3} & \cdots & \tau C & 0 \end{bmatrix} \in \mathbb{R}^{ip \times im}$$

and

$$\Gamma_w = \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ CK_\lambda & I & \ddots & \vdots & \vdots \\ \vdots & CK_\lambda & \ddots & 0 & 0 \\ CA_\lambda^{i-3}K_\lambda & \vdots & \ddots & I & 0 \\ CA_\lambda^{i-2}K_\lambda & CA_\lambda^{i-3}K_\lambda & \cdots & CK_\lambda & I \end{bmatrix} \quad (83)$$

It is clear that  $\Gamma_x$  of Eq. (82) represents the extended observability matrix as known from linear system theory [26], [25], [24].

### System Identification Algorithms

The theory provided permits formulation of a variety of algorithms with the same algebraic properties as the original discrete-time version though with application to continuous-time modeling and identification. Below is presented one realization-based algorithm. Subspace-based algorithms and theoretical justification is to be found in [13].

*Algorithm 1 (System realization [7], [13]):*

- 1) Use least-squares identification to find a multi-variable transfer function

$$G(\lambda(s)) = D_L^{-1}(\lambda)N_L(\lambda) = \sum_{k=0}^{\infty} G_k \lambda^k \quad (84)$$

where  $D_L(\lambda)$ ,  $N_L(\lambda)$  are polynomial matrices obtained by means of some identification method such as linear regression with

$$\varepsilon(t, \theta) = D_L(\lambda)y(t) - N_L(\lambda)u(t) \quad (85)$$

$$G(\lambda) = D_L^{-1}(\lambda)N_L(\lambda) \quad (86)$$

$$D_L(\lambda) = I + D_1\lambda + \cdots + D_n\lambda^n \quad (87)$$

$$N_L(\lambda) = N_0 + N_1\lambda + \cdots + N_n\lambda^n \quad (88)$$

- 2) Solve for the transformed Markov parameters

$$G_k = N_k - \sum_{j=1}^k D_j G_{k-j}, \quad k = 0, \dots, n \quad (89)$$

$$G_k = -\sum_{j=1}^n D_j G_{k-j}, \quad k = n+1, \dots, N \quad (90)$$

- 3) For suitable numbers  $q, r, s$  such that  $r+s \leq N$  arrange the Markov parameters in the Hankel matrix

$$G_{r,s}^{(q)} = \begin{bmatrix} G_{q+1} & G_{q+2} & \cdots & G_{q+s} \\ G_{q+2} & G_{q+3} & \cdots & G_{q+s+1} \\ \vdots & \vdots & \ddots & \vdots \\ G_{q+r} & G_{q+r+1} & \cdots & G_{q+r+s-1} \end{bmatrix} \quad (91)$$

- 4) Determine rank  $n$  and resultant system matrices

$$G_{r,s}^{(0)} = U\Sigma V^T \quad (\text{SVD}) \quad (92)$$

$$E_y^T = [I_{p \times p} \quad 0_{p \times (r-1)p}] \quad (93)$$

$$E_u^T = [I_{m \times m} \quad 0_{m \times (s-1)m}] \quad (94)$$

$$\Sigma_n = \text{diag} \{ \sigma_1, \sigma_2, \dots, \sigma_n \} \quad (95)$$

$$U_n = \text{matrix of first } n \text{ columns of } U \quad (96)$$

$$V_n = \text{matrix of first } n \text{ columns of } V \quad (97)$$

Finally, calculate the state-space matrices

$$A_n = \Sigma_n^{-1/2} U_n^T G_{r,s}^{(1)} V_n \Sigma_n^{-1/2}, \hat{A} = \frac{1}{\tau} (A_n - I) \quad (98)$$

$$B_n = \Sigma_n^{1/2} V_n^T E_u, \hat{B} = \frac{1}{\tau} B_n \quad (99)$$

$$C_n = E_y^T U_n \Sigma_n^{1/2}, \hat{C} = C_n \quad (100)$$

$$D_n = G_0, \hat{D} = D_n \quad (101)$$

which yields the  $n$ th-order state-space realization

$$\begin{aligned} \dot{x}(t) &= \hat{A}x(t) + \hat{B}u(t) \\ y(t) &= \hat{C}x(t) + \hat{D}u(t) \end{aligned} \quad (102)$$

## IV. DISCUSSION AND CONCLUSIONS

We have formulated an identification method for continuous-time transfer function models and equivalent to ARMAX and state-space models for discrete-time systems [9]. The transformation by means of  $\lambda$  allows an *exact* reparametrization of a continuous-time transfer function. High-frequency dynamics and low-frequency dynamics thus appear without distortion in the mapping from input to output. The low-pass filters implemented for the estimation model have a filtering effect in producing regressor variables for identification. Also orthogonal regressor variables can be used in this context. Both the operator translation and filtering approaches such as the Poisson moment functional (PMF) or the Laguerre polynomials give rise to similar estimation models for the deterministic case [22], [21], [6], [30], [31], [4]. Implementation of the operator  $\lambda$  may be done as continuous-time filters, discrete-time filters or by means of numerical integration methods [1]. As elaborated by Middleton and Goodwin [20], it may be valuable to replace  $z$  by the  $\delta$ -operator in order to improve on numerical accuracy in discrete-time implementation. Whereas ZOH only was studied here, inter-sample behavior is significant for approximation properties.

The main differences between this method and previous approaches to continuous-time model identification consist of a different estimation model and a new parametrization of the continuous-time transfer function whereas the parameter estimation method—*i.e.*, least-squares estimation or maximum-likelihood estimation, is the same as that used in ARMAX-model identification [10]. The hybrid approach involves a discrete-time model of the stochastic disturbances with little specification of the continuous-time noise, the properties of which are not known or measured in detail. With reference to the central limit theorem, this approach appears to be an appropriate assumption in physical modeling contexts where small continuous-time disturbances add up to a normally distributed disturbance on the sampled signals. Thus, the resulting model has the same modeling power of the stochastic environment as that of an ARMAX model [10]. Analysis of convergence and statistical consistency was presented in [10].

## REFERENCES

- [1] C. T. Chou, M. Verhaegen, and R. Johansson. Continuous-time identification of SISO systems using Laguerre functions. *IEEE Trans. Signal Processing*, 47:349–362, 1999.
- [2] A. B. Clymer. Direct system synthesis by means of computers, Part I. *Trans. AIEE*, 77:798–806, 1959.
- [3] F. Eng and F. Gustafsson. Identification with stochastic sampling time jitter. *Automatica*, 44:637–646, 2008.
- [4] H. Garnier, M. Gilson, and E. Huselstein. Developments for the Matlab CONTSID Toolbox. In *13th IFAC Symp. System Identification (SYSID'2003)*, Rotterdam, The Netherlands, 2003.
- [5] H. Garnier, M. Mensler, and A. Richard. Continuous-time model identification from sampled data: Implementation issues and performance evaluation. *Int. J. Control*, 76(13):1337–1357, 2003.
- [6] R. Haber and H. Unbehauen. Structure identification of nonlinear dynamic systems—A survey on input/output approaches. *Automatica*, 26:651–677, 1990.
- [7] B. L. Ho and R. E. Kalman. Effective construction of linear state-variable models from input/output functions. *Regelungstechn.*, 14:545–548, 1966.
- [8] R. Johansson. Identification of continuous-time dynamic systems. In *Proc. 25th IEEE Conf. Decision and Control*, pages 1653–1658, Athens, Greece, 1986.
- [9] R. Johansson. *System Modeling and Identification*. Prentice Hall, Englewood Cliffs, NJ, 1993.
- [10] R. Johansson. Identification of continuous-time models. *IEEE Transactions on Signal Processing*, 4:887–897, 1994.
- [11] R. Johansson. *Identification of Continuous-time Models from Sampled Data (H. Garnier, L. Wang, Eds.)*, chapter Subspace-based Continuous-time Identification, pages 291–309. Springer-Verlag, London, 2008.
- [12] R. Johansson. *Identification of Continuous-time Models from Sampled Data (H. Garnier, L. Wang, Eds.)*, chapter Continuous-time Model Identification Using Spectrum Analysis with Passivity-preserving Model Reduction, pages 393–406. Springer-Verlag, London, 2008.
- [13] R. Johansson, M. Verhaegen, and C. T. Chou. Stochastic theory of continuous-time state-space identification. *IEEE Trans. Signal Processing*, 47:41–51, January 1999.
- [14] R. Johansson, M. Verhaegen, and C. T. Chou. Stochastic theory of continuous-time state-space identification. *IEEE Transactions on Signal Processing*, 47:41–51, January 1999.
- [15] R. Johansson, M. Verhaegen, C. T. Chou, and A. Robertsson. Residual models and stochastic realization in state-space identification. *Int. J. Control*, 74(10):988–995, 2001.
- [16] R. Johansson, M. Verhaegen, C. T. Chou, and A. Robertsson. Residual models and stochastic realization in state-space system identification. *Int. J. Control*, 74:988–995, 2001.
- [17] T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation*. Prentice Hall, Upper Saddle River, NJ, 2000.
- [18] K. Mahata and H. Garnier. Identification of continuous-time errors-in-variables models. *Automatica*, 42(9):1470–1490, 2006.
- [19] F. Marvasti, editor. *Nonuniform sampling: Theory and Practice*. Kluwer, Dordrecht, NL, 2001.
- [20] R. H. Middleton and G. C. Goodwin. *Digital Control and Estimation: A Unified Approach*. Prentice Hall, Englewood Cliffs, NJ, 1990.
- [21] D. C. Saha and G. P. Rao. *Identification of continuous dynamical systems—The Poisson moment functional (PMF) approach*, volume 56 of *LNCIS*. Springer-Verlag, Berlin, 1983.
- [22] H. Unbehauen and G. P. Rao. *Identification of continuous-time systems*. North-Holland, Amsterdam, 1987.
- [23] H. Unbehauen and G. P. Rao. Continuous-time approaches to system identification—A survey. *Automatica*, 26:23–35, 1990.
- [24] P. van Overschee and B. de Moor. *Subspace Identification for Linear Systems—Theory, Implementation, Applications*. Kluwer Academic Publishers, Boston-London-Dordrecht, 1996.
- [25] M. Verhaegen and P. Dewilde. Subspace model identification—Analysis of the elementary output-error state-space model identification algorithm. *Int. J. Control*, 56:1211–1241, 1992.
- [26] M. Verhaegen and P. Dewilde. Subspace model identification—The output-error state-space model identification class of algorithms. *Int. J. Control*, 56:1187–1210, 1992.
- [27] L. Wang. Continuous-time model predictive control design using orthonormal functions. *Int. J. Control*, 74:1588–1600, 2001.
- [28] L. Wang. Discrete model predictive controller design using Laguerre function. *J. Process Control*, 14:131–142, 2004.
- [29] N. Wiener. *Nonlinear Problems in Random Theory*. The MIT Press, Cambridge, MA, 1958.
- [30] P. C. Young. Applying parameter estimation to dynamic systems. *Control engineering*, 16:Oct 119–125, Nov 118–124, 1969.
- [31] P. C. Young. Parameter estimation for continuous-time models—A survey. *Automatica*, 17:23–29, 1981.