POLITECNICO DI MILANO

Continuous-time subspace identification in closed-loop

Marco Bergamasco and Marco Lovera

Dipartimento di Elettronica e Informazione, Politecnico di Milano

Index

- Introduction
- Problem statement and preliminaries
- From continuous-time to discrete-time using Laguerre filters
- Continuous-time predictor-based subspace model identification (PBSID)
- Simulation examples

Introduction

- Subspace Model Identification (SMI) methods:
  originally developed for discrete-time models
- MOESP, PO-MOESP, N4SID, … : SMI in open-loop
- PBSID, SIMPCA, IEM, SSARX, ... : SMI in closed-loop
- The PBSID algorithm is the present state-of-the-art for closed-loop SMI
- Continuous-time subspace identification: mainly studied in open-loop setting; limited understanding of the closed-loop case

Introduction

- Main issue of continuous-time SMI:
  the need of computing high-order derivatives of input-output data

- This problem is faced in the literature using a number of different approaches:
  - (Bastogne et al. 2001): Poisson Moment Functionals
  - (Haverkamp 2001): Laguerre filtering
  - (Ohta and Kawai 2005): orthonormal basis projections (e.g., Laguerre)
  - (Mohd-Moktar and Wang 2008): Laguerre filtering+EIV

- We will rely on results first presented in
  - Y. Ohta. Realization of input-output maps using generalized orthonormal basis functions. *Systems & Control Letters*, 22(6):437–444, 2005
  - Y. Ohta and T. Kawai. Continuous-time subspace system identification using generalized orthonormal basis functions. In *16th International Symposium on Mathematical Theory of Networks and Systems, Leuven, Belgium*, 2004

  which allow to obtain a discrete-time equivalent model using a system transformation

Problem statement: model class

- Consider the LTI continuous-time stochastic system

$$
\begin{aligned}
dx(t) &= Ax(t)dt + Bu(t)dt + dw(t), x(0) = x_0 \\
dz(t) &= Cx(t)dt + Du(t)dt + dv(t) \\
y(t)dt &= dz(t)
\end{aligned}
$$

where

$$x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^p, w \in \mathbb{R}^n \text{ and } v \in \mathbb{R}^p$$

$$E\left\{ \begin{bmatrix} dw(t) \\ dv(t) \end{bmatrix} \begin{bmatrix} dw(t) \\ dv(t) \end{bmatrix}^T \right\} = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} dt$$

Some assumptions

- (A, B, C, D) such that
  - (A, C) observable
  - $(A, [B, Q^{1/2}])$ controllable
  - A asymptotically stable

Problem statement: the Laguerre basis

Some definitions

- A transfer function $w(s)$ is called inner (all-pass) if $w(j\omega) \in H_\infty$ such that $\left| w(j\omega) \right| = 1$ almost everywhere on the imaginary axis.

- Define the multiplication operator $\Lambda_w, L_2(0,\infty) \mapsto L_2(0,\infty)$ as

$$\Lambda_w u(t) = F^{-1}[wF[u(t)]]$$

F: Fourier Transformation

- Consider the first order inner function

$$w(s) = \frac{s-a}{s+a}$$

Problem statement: the Laguerre basis

$$w(s) = \frac{s-a}{s+a}$$

- It can be shown that
  - $w(s)H_2$ is a proper closed subspace of H_2, its orthogonal complement is denoted as $S = H_2 \ominus w(s)H_2$
  - a basis of the (one-dimensional) subspace S is
  $$L_0(s) = \frac{\sqrt{2a}}{s+a}$$
  - the set
  $$\left\{ L_0(s), w(s)L_0(s), \ldots, w^k(s)L_0(s), \ldots \right\}$$
  is an orthonormal basis of $H_2 = \bigoplus_{k=0}^{\infty} w^k S$

  Laguerre basis in H_2
  - Letting $l_0(t) = F^{-1}[L_0(s)]$, the set
  $$\left\{ l_0(t), \Lambda_w l_0(t), \ldots, \Lambda_w^k l_0(t), \ldots \right\}$$
  is an orthonormal basis of $L_2(0,\infty) = \bigoplus_{k=0}^{\infty} \Lambda_w^k S$

  Laguerre basis in L_2

From continuous-time to discrete-time using Laguerre filters

$$\begin{aligned} dx(t) &= Ax(t)dt + Bu(t) + dw(t) \\ dz(t) &= Cx(t) + Du(t) + dv(t) \\ y(t)dt &= dz(t) \end{aligned}$$

$$\Lambda_w^k l_0(t) = F^{-1}[w^k(s)L_0(s)]$$

State space matrices transformation

$$\begin{aligned} A_o &= (A - aI)^{-1}(A + aI) \\ B_o &= \sqrt{2a}(A - aI)^{-1}B \\ C_o &= -\sqrt{2a}C(A - aI)^{-1} \\ D_o &= D - C(A - aI)^{-1}B \end{aligned}$$

Signals transformation

$$\begin{aligned} \tilde{u}(k) &= \int_0^\infty \left(\Lambda_w^k l_0(t)\right) u(t)dt \\ \tilde{y}(k) &= \int_0^\infty \left(\Lambda_w^k l_0(t)\right) y(t)dt \\ \tilde{w}(k) &= \int_0^\infty \left(\Lambda_w^k l_0(t)\right) dw(t) \\ \tilde{v}(k) &= \int_0^\infty \left(\Lambda_w^k l_0(t)\right) dv(t) \end{aligned}$$

$$\begin{aligned} \xi(k+1) &= A_o\xi(k) + B_o\tilde{u}(k) + B_{ow}\tilde{w}(k) \\ \tilde{y}(k) &= C_o\xi(k) + D_o\tilde{u}(k) + D_{ow}\tilde{w}(k) + D_{ov}\tilde{v}(k) \end{aligned}$$

$$\begin{aligned} \xi(k+1) &= A_o\xi(k) + B_o\tilde{u}(k) + K_o\tilde{e}(k) \\ \tilde{y}(k) &= C_o\xi(k) + D_o\tilde{u}(k) + \tilde{e}(k) \end{aligned}$$

k: basis order

Discrete-time model in prediction form

- The transformed system has the state space representation

$$\begin{aligned} \xi_i(k+1) &= A_o \xi_i(k) + B_o \tilde{u}_i(k) + K_o \tilde{e}_i(k),\ \xi_i(0) = x(t_i) \\ \tilde{y}_i(k) &= C_o \xi_i(k) + D_o \tilde{u}_i(k) + \tilde{e}_i(k) \end{aligned}$$

Innovation Form

Closed-loop predictor matrices

$$\begin{aligned} \bar{A}_o &= A_o - K_o C_o \\ \bar{B}_o &= B_o - K_o D_o \end{aligned}$$

$$\begin{aligned} \tilde{z}_i(k) &= \begin{bmatrix} \tilde{u}_i^T(k) & \tilde{y}_i^T(k) \end{bmatrix}^T \\ \tilde{B}_o &= \begin{bmatrix} \bar{B}_o & K_o \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \xi_i(k+1) &= \bar{A}_o \xi_i(k) + \tilde{B}_o \tilde{z}_i(k),\ \xi_i(0) = x(t_i) \\ \tilde{y}_i(k) &= C_o \xi_i(k) + D_o \tilde{u}_i(k) + \tilde{e}_i(k), \end{aligned}$$

Prediction Form

- Iterating *p-1* times the state equation

$$
\begin{aligned}
\xi_i(k+2) &= \bar{A}_o^2 \xi_i(k) + \begin{bmatrix} \bar{A}_o \tilde{B}_o & \tilde{B}_o \end{bmatrix} \begin{bmatrix} \tilde{z}_i(k) \\ \tilde{z}_i(k+1) \end{bmatrix} \\
&\vdots \\
\xi_i(k+p) &= \bar{A}_o^p \xi_i(k) + \mathcal{K}^p Z_i^{0,p-1}
\end{aligned}
$$

where

$$
\mathcal{K}^p = \begin{bmatrix} \bar{A}_o^{p-1} \tilde{B}_0 & \dots & \tilde{B}_o \end{bmatrix}
$$

Extended controllability matrix

and

$$
Z_i^{0,p-1} = \begin{bmatrix} \tilde{z}_i(k) \\ \vdots \\ \tilde{z}_i(k+p-1) \end{bmatrix}
$$

Input-output "past" data

Continuous-time PBSID: the data equation

- $\bar{A}_o$ has all eigenvalues inside the open unit circle, so $\bar{A}_o^p \xi_i(k) \simeq 0$ for sufficiently large values of *p*

$$\xi_i(k+p) \simeq \mathcal{K}^p Z_i^{0,p-1}$$

p: past window length

- Then, the input-output behaviour of the system is

$$\begin{aligned}
\tilde{y}_i(k+p) &\simeq C_o \mathcal{K}^p Z_i^{0,p-1} + D_o \tilde{u}_i(k+p) + \tilde{e}_i(k+p) \\
&\vdots \\
\tilde{y}_i(k+p+f) &\simeq C_o \mathcal{K}^p Z_i^{f,p+f-1} + D_o \tilde{u}_i(k+p+f) + \tilde{e}_i(k+p+f)
\end{aligned}$$

- Introducing the vector notation

f: future window length

$$\begin{aligned}
Y_i^{p,f} &= \begin{bmatrix} \tilde{y}_i(k+p) & \tilde{y}_i(k+p+1) & \ldots & \tilde{y}_i(k+p+f) \end{bmatrix} \\
U_i^{p,f} &= \begin{bmatrix} \tilde{u}_i(k+p) & \tilde{u}_i(k+p+1) & \ldots & \tilde{u}_i(k+p+f) \end{bmatrix} \\
E_i^{p,f} &= \begin{bmatrix} \tilde{e}_i(k+p) & \tilde{e}_i(k+p+1) & \ldots & \tilde{e}_i(k+p+f) \end{bmatrix} \\
\Xi_i^{p,f} &= \begin{bmatrix} \xi_i(k+p) & \xi_i(k+p+1) & \ldots & \xi_i(k+p+f) \end{bmatrix} \\
\bar{Z}_i^{p,f} &= \begin{bmatrix} Z_i^{0,p-1} & Z_i^{1,p} & \ldots & Z_i^{f,p+f-1} \end{bmatrix}
\end{aligned}$$

Continuous-time PBSID: the data equation

- The system can be rewritten

$$
\begin{aligned}
\Xi_i^{p,f} &\simeq \mathcal{K}^p \bar{Z}_i^{p,f} \\
Y_i^{p,f} &\simeq C_o \mathcal{K}^p \bar{Z}_i^{p,f} + D_o U_i^{p,f} + E_i^{p,f}.
\end{aligned}
$$

- Considering the sequence of sampling instants t_i, $i=1,\dots, N$

$$
\begin{aligned}
\tilde{u}_i(k) &= \int_0^\infty \left(\Lambda_w^k l_0(\tau)\right) u(t_i + \tau) d\tau \\
\tilde{e}_i(k) &= \int_0^\infty \left(\Lambda_w^k l_0(\tau)\right) de(t_i + \tau) \\
\tilde{y}_i(k) &= \int_0^\infty \left(\Lambda_w^k l_0(\tau)\right) y(t_i + \tau) d\tau
\end{aligned}
$$

$$
Y^{p,f} = \begin{bmatrix} \tilde{y}_1(k+p) & \cdots & \tilde{y}_N(k+p) & \cdots & \tilde{y}_1(k+p+f) & \cdots & \tilde{y}_N(k+p+f) \end{bmatrix}
$$

and similarly for $U_i^{p,f}, E_i^{p,f}, \Xi_i^{p,f}$ and $\bar{Z}_i^{p,f}$

- The data equation is given by

$$
\begin{aligned}
\Xi^{p,f} &\simeq \mathcal{K}^p \bar{Z}^{p,f} \\
Y^{p,f} &\simeq C_o \mathcal{K}^p \bar{Z}^{p,f} + D_o U^{p,f} + E^{p,f}
\end{aligned}
$$

Data Equation

Continuous-time PBSID: first estimation step

- Considering $f = p$ and solving the least-squares problem

$$\min_{C_o\mathcal{K}^p, D_o} \left\| Y^{p,p} - C_o\mathcal{K}^p \bar{Z}^{p,p} - D_o U^{p,p} \right\|_F \Rightarrow \widehat{C_o\mathcal{K}^p}, \hat{D}_o$$

- Defining

$$\Gamma^p = \begin{bmatrix} C_o \\ C_o\bar{A}_o \\ \vdots \\ C_o\bar{A}_o^{p-1} \end{bmatrix}$$

Extended observability matrix

- Noting that the product of Γ^p and $\mathcal{K}^p = \begin{bmatrix} \bar{A}_o^{p-1}\tilde{B}_0 & \dots & \tilde{B}_o \end{bmatrix}$

$$\boxed{\bar{A}_o^p \simeq 0}$$

$$\Gamma^p\mathcal{K}^p = \begin{bmatrix} C_o\bar{A}_o^{p-1}\tilde{B}_o & \dots & C_o\tilde{B}_o \\ C_o\bar{A}_o^{p}\tilde{B}_o & \dots & C_o\bar{A}_o\tilde{B}_o \\ \vdots & & \\ C_o\bar{A}_o^{2p-2}\tilde{B}_o & \dots & C_o\bar{A}_o^{p-1}\tilde{B}_o \end{bmatrix} \simeq \begin{bmatrix} C_o\bar{A}_o^{p-1}\tilde{B}_o & \dots & C_o\tilde{B}_o \\ 0 & \dots & C_o\bar{A}_o\tilde{B}_o \\ \vdots & & \\ 0 & \dots & C_o\bar{A}_o^{p-1}\tilde{B}_o \end{bmatrix}$$

- An estimate of $\Gamma^p\mathcal{K}^p$ is obtained using $\widehat{C_o\mathcal{K}^p}$

Continuous-time PBSID: estimation of the state sequence

- Recalling that

$$\Xi^{p,p} \simeq \mathcal{K}^p \bar{Z}^{p,p}$$

- It holds that

$$\Gamma^p \Xi^{p,p} \simeq \Gamma^p \mathcal{K}^p \bar{Z}^{p,p}$$

- Computing the SVD of

$$\widehat{\Gamma^p \mathcal{K}^p} \bar{Z}^{p,p} = U \Sigma V^T$$

- an estimate of the state sequence can be obtained

$$\hat{\Xi}^{p,p} = \Sigma_n V_n^T$$

Continuous-time PBSID: second estimation step

- Using the estimated state sequence, the state space system matrices can be computed

$$\min_{C_o} \left\| Y^{p,p} - \hat{D}_o U^{p,p} - C_o \hat{\Xi}^{p,p} \right\|_F \quad \Rightarrow \quad \hat{C}_o$$

- An estimate of the innovation data matrix can be obtained as

$$E^{p,p} = Y^{p,p} - \hat{C}_o \hat{\Xi}^{p,p} - \hat{D}_o U^{p,p}$$

- And solving the last least squares problems

$$\min_{A_o, B_o, K_o} \left\| \hat{\Xi}^{p+1,p} - A_o \hat{\Xi}^{p,p-1} - B_o U^{p,p-1} - K_o E^{p,p-1} \right\|_F \quad \Rightarrow \quad \hat{A}_o, \hat{B}_o, \hat{K}_o$$

- Finally, the state space continuous-time system matrices are

$$\begin{bmatrix} \hat{A}_o & \hat{B}_o \\ \hat{C}_o & \hat{D}_o \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix}$$

Simulation examples

- Considered algorithms
  - PO-MOESP$_w$

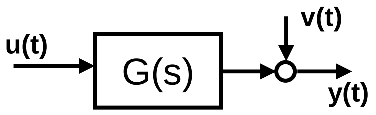
    B. R. J. Haverkamp. *State space identification: theory and practice*. PhD thesis, Delft University of Technology, 2001
  - PO-MOESP$_o$

    Y. Ohta and T. Kawai. Continuous-time subspace system identification using generalized orthonormal basis functions. In *16th International Symposium on Mathematical Theory of Networks and Systems, Leuven, Belgium*, 2004
  - PBSID$_o$
- The input is a sequence of filtered white Gaussian noise
- Implementation approximation

$$
\begin{aligned}
\tilde{u}_i(k) &= \int_0^{\infty} \left(\Lambda_w^k l_0(\tau)\right) u(t_i + \tau) d\tau = \\
&= \int_{t_i}^{\infty} \left(\Lambda_w^k l_0(\tau - t_i)\right) u(\tau) d\tau = \\
&\simeq \int_{t_i}^{t_{N/2}+t_i} \left(\Lambda_w^k l_0(\tau - t_i)\right) u(\tau) d\tau, \, i = 1, \ldots, \frac{N}{2}
\end{aligned}
$$

Simulation examples

Open-loop



$$G(s) = \frac{32}{(s+8)(s+2)}$$

- Monte Carlo study (400 runs)
- $p = f = 10$
- $a = 20$ $\quad \boxed{w(s) = \frac{s-a}{s+a}}$
- $T_{Sim} = 10s$

$\Delta t = 0.005s$ $\qquad e = \lambda - \hat{\lambda}_{mean}$

σ_v^2/σ_y^2	PO-MOESP$_w$	PO-MOESP$_o$	PBSID$_o$
0.01	−0.17±(0.50) 0.02±(0.06)	0.35±(0.49) −0.03±(0.05)	0.05±(0.24) −0.00±(0.03)
0.05	−0.96+i0.03±(1.27+i0.20) 0.13−i0.03±(0.21+i0.20)	0.32±(0.75) −0.02±(0.09)	0.07±(0.55) −0.00±(0.07)
0.1	−1.92+i0.15±(1.83+i0.43) 0.23−i0.15±(0.70+i0.43)	0.19±(0.91) −0.00±(0.12)	−0.01±(0.78) 0.01±(0.10)

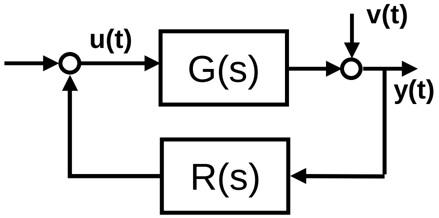
$\Delta t = 0.01s$

σ_v^2/σ_y^2	PO-MOESP$_w$	PO-MOESP$_o$	PBSID$_o$
0.01	−0.35±(0.69) 0.04±(0.09)	43.31±(80.17) −0.48±(0.27)	0.48±(0.53) −0.04±(0.05)
0.05	−1.78+i0.11±(1.63+i0.36) 0.28−i0.11±(0.35+i0.36)	54.07±(121.65) −2.09±(32.42)	0.62±(1.09) −0.05±(0.10)
0.1	−3.38+i0.51±(2.07+i0.71) 0.31−i0.51±(1.01+i0.71)	39.91±(60.04) −15.33±(296.98)	0.76±(2.28) −0.03±(0.16)

- PO-MOESP$_w$ leads to complex estimates of the real eigenvalues
- PO-MOESP$_o$ performance is not satisfactory for large sampling intervals

Simulation examples

Closed-loop



$$G(s) = \frac{32}{(s+8)(s+2)} \qquad R(s) = 2$$

- Monte Carlo study (400 runs)
- $p = f = 10$
- $a = 20$ $\quad \boxed{w(s) = \frac{s-a}{s+a}}$
- $T_{Sim} = 10s$

$\Delta t = 0.005s$ $\qquad e = \lambda - \hat{\lambda}_{mean}$

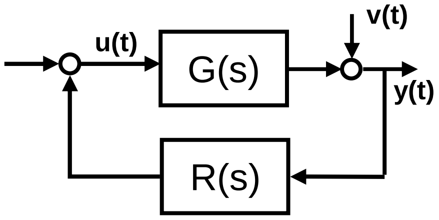
σ_v^2/σ_y^2	PO-MOESP$_w$	PO-MOESP$_o$	PBSID$_o$
0.01	−0.08±(0.37)	0.06±(0.21)	0.01±(0.15)
	0.01±(0.07)	−0.01±(0.04)	−0.00±(0.04)
0.05	−0.31±(0.77)	0.06±(0.37)	0.03±(0.36)
	0.06±(0.17)	0.00±(0.08)	0.00±(0.08)
0.1	−0.79+i0.02±(1.19+i0.15)	0.03±(0.53)	−0.01±(0.50)
	0.18−i0.02±(0.33+i0.15)	0.01±(0.12)	0.01±(0.11)

$\Delta t = 0.01s$

σ_v^2/σ_y^2	PO-MOESP$_w$	PO-MOESP$_o$	PBSID$_o$
0.01	−0.12±(0.43)	1.46±(2.02)	0.05±(0.22)
	0.02±(0.09)	−0.11±(0.20)	−0.01±(0.05)
0.05	−0.67+i0.012±(1.13+i0.10)	1.42±(1.72)	0.05±(0.48)
	0.16−i0.01±(0.32+i0.10)	−0.10±(0.22)	0.00±(0.12)
0.1	−1.40+i0.11±(1.54+i0.40)	1.36±(1.92)	0.06±(0.65)
	0.32−i0.11±(0.47+i0.40)	−0.10±(0.24)	−0.00±(0.15)

- Algorithms based on Laguerre projections provide superior performance
- PBSID$_o$ leads to better results with respect to PO-MOESP$_o$ with larger sampling intervals

Closed-loop Unstable System



$$G(s) = \frac{8}{(s+4)(s-2)} \qquad R(s) = 2$$

- Monte Carlo study (400 runs)
- $p = f = 10$
- $a = 20$ $\quad \boxed{w(s) = \frac{s-a}{s+a}}$
- $T_{Sim} = 10s$

$\Delta t = 0.005s$ $\qquad e = \lambda - \hat{\lambda}_{mean}$

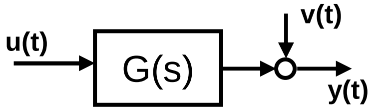
σ_v^2/σ_y^2	PO-MOESP$_w$	PO-MOESP$_o$	PBSID$_o$
0.01	−0.02±(0.13) 0.02±(0.02)	0.06±(0.10) 0.01±(0.05)	0.03±(0.07) 0.02±(0.03)
0.05	−0.10±(0.31) 0.02±(0.05)	0.06±(0.18) 0.02±(0.08)	0.04±(0.16) 0.04±(0.06)
0.1	−0.16±(0.46) 0.02±(0.07)	0.04±(0.25) 0.02±(0.11)	0.06±(0.23) 0.05±(0.08)

$\Delta t = 0.01s$

σ_v^2/σ_y^2	PO-MOESP$_w$	PO-MOESP$_o$	PBSID$_o$
0.01	0.01±(0.19) 0.03±(0.03)	0.75±(1.04) 0.13±(0.24)	0.06±(0.11) 0.06±(0.04)
0.05	−0.11±(0.44) 0.03±(0.07)	0.70±(1.04) 0.17±(0.34)	0.07±(0.24) 0.09±(0.09)
0.1	−0.28±(0.65) 0.04±(0.12)	0.82+0.00i±(1.29+i0.03) 0.27−0.00i±(0.51+i0.03)	0.10±(0.32) 0.12±(0.13)

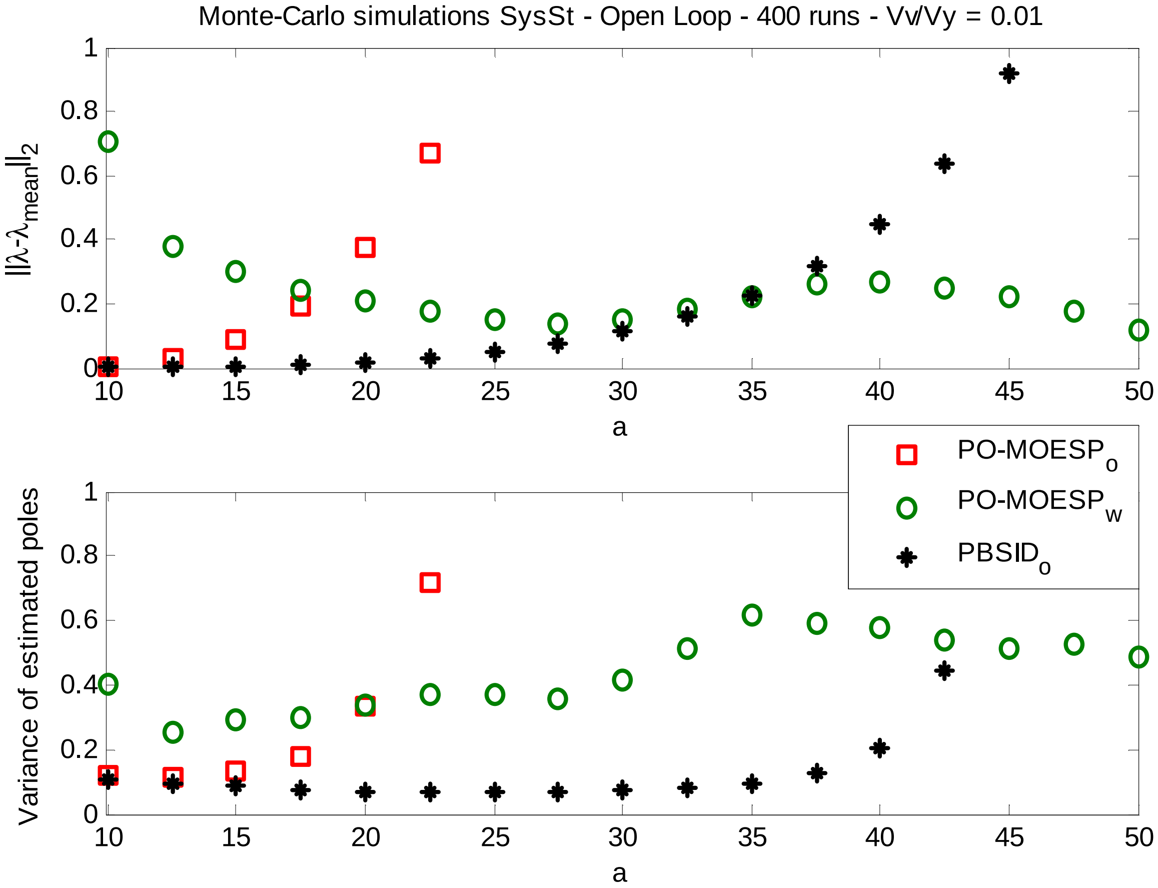
Simulation examples: choice of *a*

Open-loop



$$G(s) = \frac{32}{(s+8)(s+2)}$$

- Monte Carlo study (400 runs)
- $p = f = 10$
- $T_{Sim} = 10s$
- $\Delta t = 0.005s$
- $\sigma_v^2 \; / \; \sigma_y^2 = 0.01$

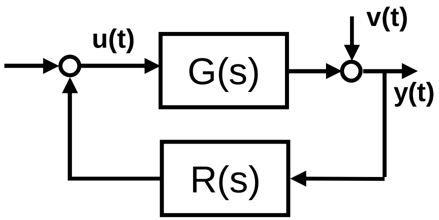


- PO-MOESP$_w$ results are irregular for increasing of *a*
- Algorithms based on projections give poor performance for large values of *a*. For PBSID$_o$, the variance is smaller over a wide range of values of *a*.

Conclusions

- The problem of continuous-time subspace model identification has been studied
- An algorithm combining Laguerre projections and predictor-based SMI has been proposed
- The new approach leads to more accurate and reliable results than the comparing algorithms
- It has been shown that approaches based on Laguerre projection provide better performance than the ones based on Laguerre filtering

Closed-loop



$$G(s) = \frac{32}{(s+8)(s+2)} \qquad R(s) = 2$$

- Monte Carlo study (400 runs)
- $p = f = 10$
- $T_{Sim} = 10s$
- $\Delta t = 0.005s$
- $\sigma_v^2 \ / \ \sigma_y^2 = 0.01$