

OPTIMAL LOCAL LPV IDENTIFICATION EXPERIMENT DESIGN

D. Ghosh⁽¹⁾, X. Bombois⁽¹⁾, J. Huillery⁽¹⁾, G. Scorletti⁽¹⁾ et
G. Mercère⁽²⁾

1. Laboratoire Ampère UMR CNRS 5005
2. LIAS, Université de Poitiers

ERNSI workshop - 28 September 2016

Introduction

An LPV system is a system whose parameters depend on an exogenous (scheduling) variable $p(t)$

If $p(t)$ is kept constant, the LPV system is an LTI system

The dynamics of this LTI system depend on the value of the constant p

We have a collection of LTI dynamics at different operating points

Such a representation can be used to deal with non-linear systems (gain scheduling)

Introduction

An LPV system is a system whose parameters depend on an exogenous (scheduling) variable $p(t)$

If $p(t)$ is kept constant, the LPV system is an LTI system

The dynamics of this LTI system depend on the value of the constant p

We have a collection of LTI dynamics at different operating points

Such a representation can be used to deal with non-linear systems (gain scheduling)

Introduction

An LPV system is a system whose parameters depend on an exogenous (scheduling) variable $p(t)$

If $p(t)$ is kept constant, the LPV system is an LTI system

The dynamics of this LTI system depend on the value of the constant p

We have a collection of LTI dynamics at different operating points

Such a representation can be used to deal with non-linear systems (gain scheduling)

Introduction

Local LPV identification approach: $p(t)$ is kept constant at successive operating points and local LTI identification experiments are performed

We determine those operating points and the local LTI identification experiments to guarantee a certain model accuracy with the least input energy

Related work on the selection of the scheduling sequence: *Khalate et al: 2009*, *Vizer et al: 2015*

Introduction

Local LPV identification approach: $p(t)$ is kept constant at successive operating points and local LTI identification experiments are performed

We determine those operating points and the local LTI identification experiments to guarantee **a certain model accuracy with the least input energy**

Related work on the selection of the scheduling sequence: *Khalate et al: 2009*, *Vizer et al: 2015*

Introduction

Local LPV identification approach: $p(t)$ is kept constant at successive operating points and local LTI identification experiments are performed

We determine those operating points and the local LTI identification experiments to guarantee **a certain model accuracy with the least input energy**

Related work on the selection of the scheduling sequence: *Khalate et al: 2009, Vizer et al: 2015*

Description of the LPV system

We consider the following LPV-OE system for simplicity:

$$y_{nf}(t) = - \sum_{i=1}^{n_a} a_i^0(p(t)) y(t-i) + \sum_{i=1}^{n_b} b_i^0(p(t)) u(t-i)$$

$$y(t) = y_{nf}(t) + e(t)$$

The parameter vector $\xi^0(p(t)) = (a_1^0(p(t)), \dots, b_{n_b}(p(t)))^T$ depends on the **time-varying scheduling variable** $p(t)$

$$a_i^0(p(t)) = a_{i,0}^0 + \sum_{j=1}^{n_p} a_{i,j}^0 p^j(t)$$

$$b_i^0(p(t)) = b_{i,0}^0 + \sum_{j=1}^{n_p} b_{i,j}^0 p^j(t)$$

Description of the LPV system

$$a_i^0(p(t)) = a_{i,0}^0 + \sum_{j=1}^{n_p} a_{i,j}^0 p^j(t) \quad i = 1 \dots n_a$$

$$b_i^0(p(t)) = b_{i,0}^0 + \sum_{j=1}^{n_p} b_{i,j}^0 p^j(t) \quad i = 1 \dots n_b$$

This defines a mapping $T(p(t))$ between the global parameter vector θ^0 and the time-varying parameter vector $\xi^0(p(t))$

$$\xi^0(p(t)) = T(p(t)) \theta^0$$

$$\xi^0(p(t)) = (a_1^0(p(t)), \dots, b_{n_b}^0(p(t)))^T \quad \theta^0 = (a_{1,0}^0, \dots, b_{n_b, n_p}^0)^T$$

Identification objective

$$\xi^0(p(t)) = T(p(t)) \theta^0$$

The parameter vector θ^0 entirely determines the LPV system

Objective. Determine with the least powerful excitation an estimate $\hat{\theta}$ of θ^0 having a given accuracy:

$$P_{\hat{\theta}}^{-1} > R_{adm}$$

Identification objective

$$\xi^0(p(t)) = T(p(t)) \theta^0$$

The parameter vector θ^0 entirely determines the LPV system

Objective. Determine with the least powerful excitation an estimate $\hat{\theta}$ of θ^0 having a given accuracy:

$$P_{\theta}^{-1} > R_{adm}$$

Identification of an LPV system: local approach

Suppose $p(t)$ is kept constant to an operating point \mathbf{p}_m

$$p(t) = \mathbf{p}_m \quad \forall t$$

The LPV system then reduces to an LTI system described by a time-invariant parameter vector $\xi^0(\mathbf{p}_m)$

$$y(t) = G(z, \xi^0(\mathbf{p}_m))u(t) + e(t)$$

$$\xi^0(\mathbf{p}_m) = T(\mathbf{p}_m) \theta^0$$

This LTI system can of course then be identified using LTI prediction error identification

Identification of an LPV system: local approach

Suppose $p(t)$ is kept constant to an operating point \mathbf{p}_m

$$p(t) = \mathbf{p}_m \quad \forall t$$

The LPV system then reduces to an LTI system described by a time-invariant parameter vector $\xi^0(\mathbf{p}_m)$

$$y(t) = G(z, \xi^0(\mathbf{p}_m))u(t) + e(t)$$

$$\xi^0(\mathbf{p}_m) = T(\mathbf{p}_m) \theta^0$$

This LTI system can of course then be identified using LTI prediction error identification

LTI identification at an operating point \mathbf{p}_m

If we apply an input signal u_m of spectrum Φ_{u_m} to

$$y_m(t) = G(z, \xi^0(\mathbf{p}_m))u_m(t) + e_m(t),$$

we can collect a data set $Z_m^N = \{y_m(t), u_m(t) \mid t = 1 \dots N\}$ and identify an estimate $\hat{\xi}_m$ of $\xi^0(\mathbf{p}_m)$ using:

$$\hat{\xi}_m = \arg \min_{\xi} \frac{1}{N} \sum_{t=1}^N (y_m(t) - G(z, \xi)u_m(t))^2$$

This estimate is (asymptotically) such that $\hat{\xi}_m \sim \mathcal{N}(\xi^0(\mathbf{p}_m), P_{\hat{\xi}_m})$

LTI identification at an operating point \mathbf{p}_m

The estimate $\hat{\xi}_m$ is (asymptotically) such that $\hat{\xi}_m \sim \mathcal{N}(\xi^0(\mathbf{p}_m), P_{\hat{\xi}_m})$

The covariance matrix $P_{\hat{\xi}_m}$ depends on $\xi^0(\mathbf{p}_m)$ and Φ_{u_m} :

$$P_{\hat{\xi}_m}^{-1} = \frac{N}{\sigma_e^2} \frac{1}{2\pi} \int_{-\pi}^{\pi} F(e^{j\omega}, \xi^0(\mathbf{p}_m)) F(e^{j\omega}, \xi^0(\mathbf{p}_m))^* \Phi_{u_m}(\omega) d\omega$$

$$F(z, \xi^0(\mathbf{p}_m)) = \left. \frac{dG(z, \xi)}{d\xi} \right|_{\xi^0(\mathbf{p}_m)}$$

This operation has to be repeated at different \mathbf{p}_m to deduce an estimate of θ^0 since $\dim(\xi^0(\mathbf{p}_m)) < \dim(\theta^0)$

LTI identification at an operating point \mathbf{p}_m

The estimate $\hat{\xi}_m$ is (asymptotically) such that $\hat{\xi}_m \sim \mathcal{N}(\xi^0(\mathbf{p}_m), P_{\hat{\xi}_m})$

The covariance matrix $P_{\hat{\xi}_m}$ depends on $\xi^0(\mathbf{p}_m)$ and Φ_{u_m} :

$$P_{\hat{\xi}_m}^{-1} = \frac{N}{\sigma_e^2} \frac{1}{2\pi} \int_{-\pi}^{\pi} F(e^{j\omega}, \xi^0(\mathbf{p}_m)) F(e^{j\omega}, \xi^0(\mathbf{p}_m))^* \Phi_{u_m}(\omega) d\omega$$

$$F(z, \xi^0(\mathbf{p}_m)) = \left. \frac{dG(z, \xi)}{d\xi} \right|_{\xi^0(\mathbf{p}_m)}$$

This operation has to be repeated at different \mathbf{p}_m to deduce an estimate of θ^0 since $\dim(\xi^0(\mathbf{p}_m)) < \dim(\theta^0)$

We obtain M estimates $\hat{\xi}_m$ of $\xi^0(\mathbf{p}_m) = T(\mathbf{p}_m) \theta^0$:

$$\hat{\xi}_m = T(\mathbf{p}_m) \theta^0 + \delta_m \quad \delta_m \sim \mathcal{N}(0, P_{\hat{\xi}_m})$$

The estimate $\hat{\theta}$ of θ^0 is classically determined using ordinary least squares based on the observations $\hat{\xi}_m$ and the regressor $T(\mathbf{p}_m)$

This is however **not the minimum variance estimator** since the respective variances of $\hat{\xi}_m$ are neglected

⇒ use of **weighted least squares**:

$$\hat{\theta} = \arg \min_{\theta} \sum_{m=1}^M \left(\hat{\xi}_m - T(\mathbf{p}_m) \theta \right)^T P_{\hat{\xi}_m}^{-1} \left(\hat{\xi}_m - T(\mathbf{p}_m) \theta \right)$$

We obtain M estimates $\hat{\xi}_m$ of $\xi^0(\mathbf{p}_m) = T(\mathbf{p}_m) \theta^0$:

$$\hat{\xi}_m = T(\mathbf{p}_m) \theta^0 + \delta_m \quad \delta_m \sim \mathcal{N}(0, P_{\hat{\xi}_m})$$

The estimate $\hat{\theta}$ of θ^0 is classically determined using ordinary least squares based on the observations $\hat{\xi}_m$ and the regressor $T(\mathbf{p}_m)$

This is however **not the minimum variance estimator** since the respective variances of $\hat{\xi}_m$ are neglected

⇒ use of **weighted least squares**:

$$\hat{\theta} = \arg \min_{\theta} \sum_{m=1}^M \left(\hat{\xi}_m - T(\mathbf{p}_m)\theta \right)^T P_{\hat{\xi}_m}^{-1} \left(\hat{\xi}_m - T(\mathbf{p}_m)\theta \right)$$

We obtain M estimates $\hat{\xi}_m$ of $\xi^0(\mathbf{p}_m) = T(\mathbf{p}_m) \theta^0$:

$$\hat{\xi}_m = T(\mathbf{p}_m) \theta^0 + \delta_m \quad \delta_m \sim \mathcal{N}(0, P_{\hat{\xi}_m})$$

The estimate $\hat{\theta}$ of θ^0 is classically determined using ordinary least squares based on the observations $\hat{\xi}_m$ and the regressor $T(\mathbf{p}_m)$

This is however **not the minimum variance estimator** since the respective variances of $\hat{\xi}_m$ are neglected

⇒ use of **weighted least squares**:

$$\hat{\theta} = \arg \min_{\theta} \sum_{m=1}^M \left(\hat{\xi}_m - T(\mathbf{p}_m) \theta \right)^T P_{\hat{\xi}_m}^{-1} \left(\hat{\xi}_m - T(\mathbf{p}_m) \theta \right)$$

$$\hat{\theta} = \arg \min_{\theta} \sum_{m=1}^M \left(\hat{\xi}_m - T(\mathbf{p}_m)\theta \right)^T P_{\hat{\xi}_m}^{-1} \left(\hat{\xi}_m - T(\mathbf{p}_m)\theta \right)$$

The estimate $\hat{\theta}$ is such that $\hat{\theta} \sim \mathcal{N}(\theta^0, P_{\theta})$

$$P_{\theta}^{-1} = \sum_{m=1}^M T^T(\mathbf{p}_m) P_{\hat{\xi}_m}^{-1} T(\mathbf{p}_m)$$

with $P_{\hat{\xi}_m}^{-1}$ linear in Φ_{u_m}

P_{θ}^{-1} is the sum of the contribution of each local experiments !!

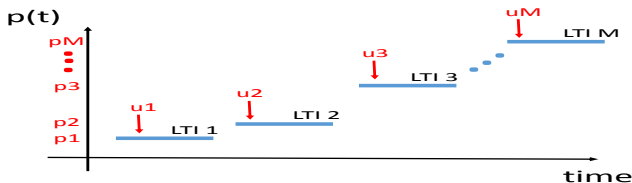
Optimal experimental design

To-be-optimized variables:

- the number M of local identification experiments M ,
- the operating points \mathbf{p}_m ($m = 1 \dots M$)
- the spectra Φ_{u_m} of the input signal u_m ($m = 1 \dots M$) used in the local identification experiments

To-be-minimized cost: $\mathcal{J} = N \sum_{m=1}^M \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$

Accuracy constraint: $P_{\theta}^{-1} > R_{adm}$



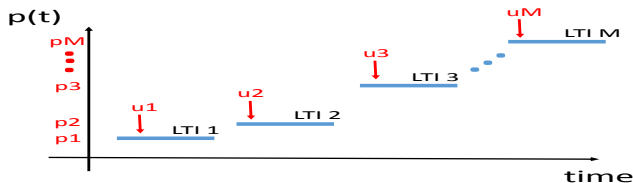
Optimal experimental design

To-be-optimized variables:

- the number M of local identification experiments M ,
- the operating points \mathbf{p}_m ($m = 1 \dots M$)
- the spectra Φ_{u_m} of the input signal u_m ($m = 1 \dots M$) used in the local identification experiments

To-be-minimized cost:
$$\mathcal{J} = N \sum_{m=1}^M \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$$

Accuracy constraint: $P_{\theta}^{-1} > R_{adm}$



Convex optimization for the design of the spectra

Suppose that we have a-priori chosen M and \mathbf{p}_m ($m = 1 \dots M$)

The design of Φ_{u_m} ($m = 1 \dots M$) is then a **convex optimization problem**

$$\Phi_{u_m} \min_{(m=1 \dots M)} N \sum_{m=1}^M \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$$

$$\sum_{m=1}^M T^T(\mathbf{p}_m) P_{\hat{\xi}_m}^{-1}(\Phi_{u_m}) T(\mathbf{p}_m) > R_{adm}$$

Convex optimization for the design of the spectra

Suppose that we have a-priori chosen M and \mathbf{p}_m ($m = 1 \dots M$)

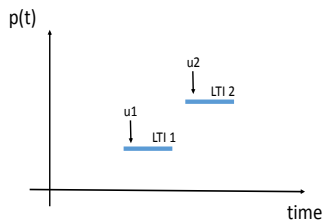
The design of Φ_{u_m} ($m = 1 \dots M$) is then a **convex optimization problem**

$$\min_{\Phi_{u_m} \ (m=1 \dots M)} N \sum_{m=1}^M \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$$

$$\sum_{m=1}^M T^T(\mathbf{p}_m) P_{\hat{\xi}_m}^{-1}(\Phi_{u_m}) T(\mathbf{p}_m) > R_{adm}$$

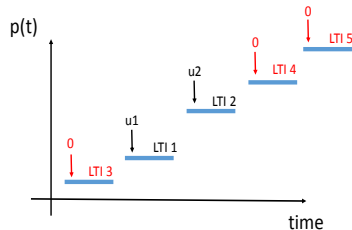
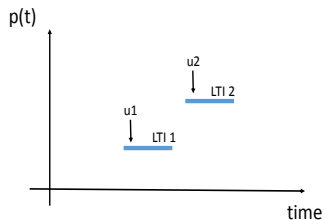
How to perform the selection of the operating points \mathbf{p}_m ?

$$\min_{\Phi_{u_m}} N \sum_{m=1}^M \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$$
$$\sum_{m=1}^M T^T(\mathbf{p}_m) P_{\hat{\xi}_m}^{-1}(\Phi_{u_m}) T(\mathbf{p}_m) > R_{adm}$$



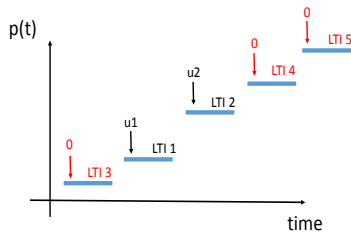
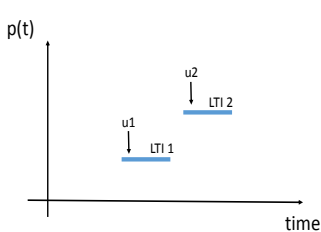
How to perform the selection of the operating points \mathbf{p}_m ?

$$\min_{\Phi_{u_m}} N \sum_{m=1}^M \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$$
$$\sum_{m=1}^M T^T(\mathbf{p}_m) P_{\hat{\xi}_m}^{-1}(\Phi_{u_m}) T(\mathbf{p}_m) > R_{adm}$$



How to perform the selection of the operating points \mathbf{p}_m ?

$$\min_{\Phi_{u_m}} N \sum_{m=1}^M \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$$
$$\sum_{m=1}^M T^T(\mathbf{p}_m) P_{\hat{\xi}_m}^{-1}(\Phi_{u_m}) T(\mathbf{p}_m) > R_{adm}$$



These experiments are equivalent from a mathematical point of view since they lead to the same cost \mathcal{J} and the same P_{θ}^{-1} !!

Convex formulation of the experiment design problem

Consider a fine grid $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{M_{grid}}\}$ of the scheduling space

We will determine a spectrum Φ_{u_m} for all \mathbf{p}_m in this fine grid

The optimal experiment design problem can thus be formulated as:

$$\min_{\Phi_{u_m} \ (m=1 \dots M_{grid})} N \sum_{m=1}^{M_{grid}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$$
$$\sum_{m=1}^{M_{grid}} T^T(\mathbf{p}_m) P_{\hat{\xi}_m}^{-1}(\Phi_{u_m}) T(\mathbf{p}_m) > R_{adm}$$

The local experiments will of course only be performed at the operating points \mathbf{p}_m for which $\Phi_{u_m}^{opt} \neq 0$

Convex formulation of the experiment design problem

Consider a fine grid $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{M_{grid}}\}$ of the scheduling space

We will determine a spectrum Φ_{u_m} for all \mathbf{p}_m in this fine grid

The optimal experiment design problem can thus be formulated as:

$$\min_{\Phi_{u_m} \ (m=1 \dots M_{grid})} N \sum_{m=1}^{M_{grid}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$$
$$\sum_{m=1}^{M_{grid}} T^T(\mathbf{p}_m) P_{\hat{\xi}_m}^{-1}(\Phi_{u_m}) T(\mathbf{p}_m) > R_{adm}$$

The local experiments will of course only be performed at the operating points \mathbf{p}_m for which $\Phi_{u_m}^{opt} \neq 0$

Convex formulation of the experiment design problem

Consider a fine grid $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{M_{grid}}\}$ of the scheduling space

We will determine a spectrum Φ_{u_m} for all \mathbf{p}_m in this fine grid

The optimal experiment design problem can thus be formulated as:

$$\min_{\Phi_{u_m} \ (m=1 \dots M_{grid})} N \sum_{m=1}^{M_{grid}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{u_m}(\omega) d\omega$$
$$\sum_{m=1}^{M_{grid}} T^T(\mathbf{p}_m) P_{\hat{\xi}_m}^{-1}(\Phi_{u_m}) T(\mathbf{p}_m) > R_{adm}$$

The local experiments will of course only be performed at **the operating points \mathbf{p}_m for which $\Phi_{u_m}^{opt} \neq 0$**

Chicken-and-egg problem

The covariance matrix P_θ depends on θ^0

We can determine a first estimate θ_{init} of θ^0 using an initial local LPV identification

The optimal experiment design problem will then be used to complement the information delivered by this initial experiment

In this optimal experiment design problem, θ^0 will be replaced by θ_{init}

Numerical illustration

Consider the following LPV-OE system: $y(t) = y_{nf}(t) + e(t)$

$$y_{nf}(t) = -a_1^0(p(t)) y(t-1) + b_1^0(p(t)) u(t-1)$$

$$a_1^0(p(t)) = -0.9 + 0.1 p(t) \quad b_1^0(p(t)) = 10 - 1 p(t)$$

$$\underbrace{\begin{pmatrix} a_1^0(p(t)) \\ b_1^0(p(t)) \end{pmatrix}}_{=\xi^0(p(t))} = \underbrace{\begin{pmatrix} 1 & p(t) & 0 & 0 \\ 0 & 0 & 1 & p(t) \end{pmatrix}}_{=T(p(t))} \underbrace{\begin{pmatrix} -0.9 \\ 0.1 \\ 10 \\ -1 \end{pmatrix}}_{=\theta^0}$$

$p(t)$ can take values in the scheduling space $[0 \ 8]$

Numerical illustration

Consider the following LPV-OE system: $y(t) = y_{nf}(t) + e(t)$

$$y_{nf}(t) = -a_1^0(\rho(t)) y(t-1) + b_1^0(\rho(t)) u(t-1)$$

$$a_1^0(\rho(t)) = -0.9 + 0.1 \rho(t) \quad b_1^0(\rho(t)) = 10 - 1 \rho(t)$$

$$\underbrace{\begin{pmatrix} a_1^0(\rho(t)) \\ b_1^0(\rho(t)) \end{pmatrix}}_{=\xi^0(\rho(t))} = \underbrace{\begin{pmatrix} 1 & \rho(t) & 0 & 0 \\ 0 & 0 & 1 & \rho(t) \end{pmatrix}}_{=T(\rho(t))} \underbrace{\begin{pmatrix} -0.9 \\ 0.1 \\ 10 \\ -1 \end{pmatrix}}_{=\theta^0}$$

$\rho(t)$ can take values in the scheduling space $[0 \ 8]$

Numerical illustration

Consider the following LPV-OE system: $y(t) = y_{nf}(t) + e(t)$

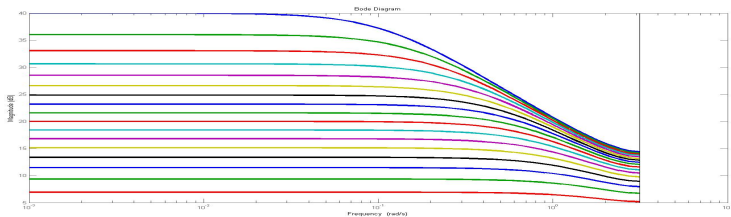
$$y_{nf}(t) = -a_1^0(p(t)) y(t-1) + b_1^0(p(t)) u(t-1)$$

$$a_1^0(p(t)) = -0.9 + 0.1 p(t) \quad b_1^0(p(t)) = 10 - 1 p(t)$$

$$\underbrace{\begin{pmatrix} a_1^0(p(t)) \\ b_1^0(p(t)) \end{pmatrix}}_{=\xi^0(p(t))} = \underbrace{\begin{pmatrix} 1 & p(t) & 0 & 0 \\ 0 & 0 & 1 & p(t) \end{pmatrix}}_{=T(p(t))} \underbrace{\begin{pmatrix} -0.9 \\ 0.1 \\ 10 \\ -1 \end{pmatrix}}_{=\theta^0}$$

$p(t)$ can take values in the scheduling space $[0 \ 8]$

Frequency responses of the corresponding $G(z, \xi^0(\mathbf{p}_m))$



We choose: $N = 1000$, $\sigma_e^2 = 0.5$ and R_{adm} enforces a standard deviation of 0.3% on each parameter of θ^0

Optimization problem based on the $M_{grid} = 17$ operating points

$$\mathbf{p}_m = 0, 0.5, 1, 1.5, \dots, 8$$

\implies only three nonzero Φ_{u_m} at $\mathbf{p}_m = 0, 1$ and 8

Corresponding $G(z, \xi^0(\mathbf{p}_m))$ and Φ_{u_m}

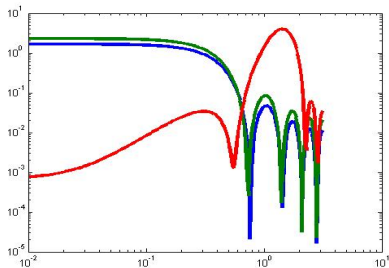
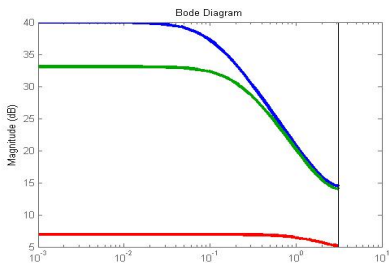
We choose: $N = 1000$, $\sigma_e^2 = 0.5$ and R_{adm} enforces a standard deviation of 0.3% on each parameter of θ^0

Optimization problem based on the $M_{grid} = 17$ operating points

$$\mathbf{p}_m = 0, 0.5, 1, 1.5, \dots, 8$$

\Rightarrow only three nonzero Φ_{u_m} at $\mathbf{p}_m = 0, 1$ and 8

Corresponding $G(z, \xi^0(\mathbf{p}_m))$ and Φ_{u_m}



Let us compare the required input energy \mathcal{J} to obtain $P_\theta^{-1} > R_{adm}$ for different choices of \mathbf{p}_m

\mathbf{p}_m	required input energy \mathcal{J}
$\mathbf{p}_m = 0, 1, 8$	1380
$\mathbf{p}_m = 0, 4, 8$	2320
$\mathbf{p}_m = 0, 1$	23000
$\mathbf{p}_m = 0, 8$	16000
$\mathbf{p}_m = 1, 8$	23000

Conclusions

First attempt to tackle the optimal experiment design problem for LPV systems

Local approach: $p(t)$ follows a staircase shape

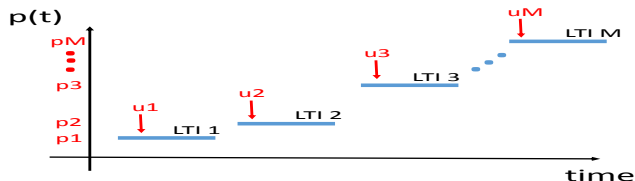
A staircase $p(t)$ is certainly not (fully) optimal

Future work will consider other shapes of $p(t)$ (global LPV identification)

Conclusions

First attempt to tackle the optimal experiment design problem for LPV systems

Local approach: $p(t)$ follows a **staircase shape**



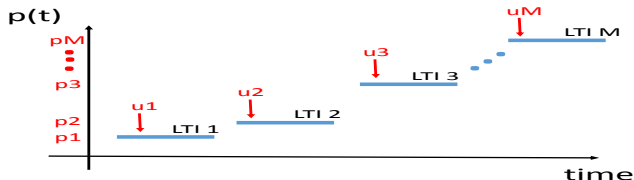
A staircase $p(t)$ is certainly not (fully) optimal

Future work will consider other shapes of $p(t)$ (global LPV identification)

Conclusions

First attempt to tackle the optimal experiment design problem for LPV systems

Local approach: $p(t)$ follows a staircase shape



A staircase $p(t)$ is certainly not (fully) optimal

Future work will consider other shapes of $p(t)$ (global LPV identification)