



**HAL**  
open science

# The Effect of Prior Knowledge on the Least Costly Identification Experiment

Georgios Birpoutsoukis, Xavier Bombois

► **To cite this version:**

Georgios Birpoutsoukis, Xavier Bombois. The Effect of Prior Knowledge on the Least Costly Identification Experiment. 18th IFAC Symposium on System Identification, Jul 2018, Stockholm, Sweden. pp.209-214, 10.1016/j.ifacol.2018.09.131 . hal-01997195

**HAL Id: hal-01997195**

**<https://hal.science/hal-01997195>**

Submitted on 25 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Effect of Prior Knowledge on the Least Costly Identification Experiment <sup>★</sup>

Georgios Birpoutsoukis<sup>\*</sup> Xavier Bombois<sup>\*\*</sup>

<sup>\*</sup> *Department of ELEC, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium, (e-mail: georgios.birpoutsoukis@vub.ac.be).*

<sup>\*\*</sup> *Laboratoire Ampère UMR CNRS 5005, Ecole Centrale de Lyon, 69134 Ecully, France, (e-mail: xavier.bombois@ec-lyon.fr)*

---

**Abstract:** Design of optimal input excitations is one of the most challenging problems in the field of system identification. The main difficulty lies in the fact that the optimization problem cannot always be formulated to be convex, therefore a globally optimal excitation for the dynamic system of interest cannot be guaranteed. In this paper, optimal input design (OID) for linear systems in the presence of prior knowledge is studied. Information related to exponential decay and smoothness is incorporated in the optimal input design problem by making use of the Bayes rule of information. Three different cases of modeling the linear dynamics are considered, namely Finite Impulse Response (FIR) model with and without prior knowledge, as well as the rational transfer function case. It is shown that the prior information affects the spectrum of the minimum power optimal input. The input with the least power is always obtained for the transfer function model case.

*Keywords:* System identification, Optimal experiment design, Regularization

---

## 1. INTRODUCTION

The purpose of optimal input design, as a subfield of optimal experiment design, is to determine the optimal excitation signal for a dynamic system. Two basic frameworks of optimal design problems can be found in the literature, with respect to the cost function, as well as the constraints, that determine the optimality of the signal.

The first framework considers the optimal input signal as the one that maximizes a measure of information in the data given constraints on the signal characteristics. In case of an asymptotically efficient estimator, i.e. one that achieves the Cramér–Rao bound, the covariance matrix of the parameter vector converges to the inverse of the Fisher information matrix for increasing number of samples. Different measures of the covariance matrix have been extensively used to define the optimizing measure of information in the data (Wong, W. K. (1994)).

The second framework, known as the Least Costly identification framework, considers the optimal input signal as the one that minimizes the cost of the experiment given constraints on the estimated model performance, Bombois, X., Scorletti, G., Gevers, M., Hildebrand, R., & Van den Hof, P. (2004), Bombois, X., Scorletti, G., Van den Hof, P., & Gevers, M. (2004), Bombois, X., Scorletti, G., Gevers, M., Van den Hof, P. M. J., & Hildebrand, R. (2006), Bombois, X., & Scorletti, G. (2012), Forgiione, M., Bombois, X., & Van den Hof, P. M. J. (2013). This is

also the one used in this work, to determine the optimal excitation in the presence of prior knowledge.

In particular, we consider the optimal input design problem when the impulse response coefficients of a system have to be identified. In this particular situation, in order to reduce the variance of the estimate, prior information on the decay rate and the smoothness of these impulse response coefficients is generally added. It is to be noted that, in recent years, this type of identification has known a regained interest for various reasons Pillonetto, G., & De Nicolao, G. (2010), Chen, T., Ohlsson, H., & Ljung, L. (2012), Pillonetto, G., Dinuzzo, F., Chen, T., De Nicolao, G., & Ljung, L. (2014), Lataire, J., & Chen, T. (2016). The main objective of the present paper is to analyze the role of such prior information on the optimal input signal.

To investigate the role of such a prior information, we will compare the optimal input signal for three different cases, namely: i) FIR modeling of a linear system without any prior knowledge about the underlying dynamics, ii) FIR modeling with prior knowledge about exponential decay and smoothness of the impulse response and, iii) identification of the rational transfer function of the linear system, which corresponds to an infinite impulse response (IIR) representation, without any prior knowledge considered.

A least costly identification framework is adopted to investigate and compare the aforementioned cases. As such, the minimum power input excitation is computed for all the aforementioned cases given a minimum requirement on the information contained in the measured data of the identification experiment. This minimum requirement is related to bounds on the modulus of the error between the estimated and the true underlying transfer function.

---

<sup>★</sup> This work was supported in part by the Fund for Scientific Research (FWO-Vlaanderen), by the Flemish Government (Methusalem), the Belgian Government through the Inter university Poles of Attraction (IAP VII) Program, and by the ERC advanced grant SNLSID, under contract 320378.

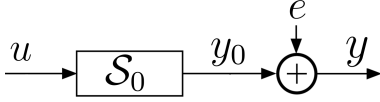


Fig. 1. Block scheme of the system structure considered in this work.

The dependence of the optimal input spectrum on the prior information about the underlying impulse response becomes clear through a numerical example. In this example, it will be observed that, without too much surprise, the least powerful input signal is obtained in the case where a rational model structure is used. It will also be observed that the presence of prior information on decay rate and smoothness when identifying the impulse response coefficients allows to reduce the required power spectrum especially in high frequencies (meaning that this type of prior information mainly pertains to these higher frequencies).

It is to be noted that a similar input design problem has been considered in Fujimoto, Y., & Sugie, T. (2016), Mu, B., & Chen, T. (2017). The main contribution of the present paper with respect to these earlier contributions is in the analysis of the influence of the prior information on the optimal spectrum.

The paper is structured as follows. In Section 2 necessary definitions, assumptions and problem formulation are given. In Section 3 the amount of information available at the end of the identification experiment is defined for all the modeling cases considered. In Section 4 it is shown how requirements on the accuracy of the estimated model are translated to minimum requirements on the information extracted from the identification experiment. The least costly OID problem is also formulated in the same section. In Section 5, the least costly OID method is demonstrated on a numerical example. Finally, some conclusions are given in Section 6.

## 2. PROBLEM DEFINITION AND ASSUMPTIONS

*Definition 1.* (System Class) The unknown data-generating system  $\mathcal{S}_0$  is stable and Linear Time-Invariant (LTI), and an Output Error (OE) framework is considered (Fig. 1):

$$\mathcal{S}_0 : y(k) = y_0(k) + e(k) = G_0(z)u(k) + e(k) \quad (1)$$

where  $G_0(z)$  is in general a rational transfer function (Infinite Impulse Response (IIR) system),  $y_0$  is the true noiseless system output as a reaction to the input excitation  $u$ ,  $e$  represents the identically and independently normally distributed measurement noise of finite variance ( $e \sim \mathcal{N}(0, \sigma_e^2)$ ) and  $y$  is the measured noisy system output.

To identify the underlying dynamics based on  $N$  input  $u$  - output  $y$  measured samples, a model structure must be defined.

*Definition 2.* (Model classes) The true system dynamics  $\mathcal{S}_0$  are modeled with the following OE structure as:

$$y(k) = y_m(k) + e(k) = G(z, \theta)u(k) + e(k) \quad (2)$$

where  $G(z, \theta)$  is the model structure parameterized with the vector  $\theta \in \mathbb{R}^{n_\theta}$ , to be estimated from measured input-output data. In this work, two different cases are considered for the model structure candidate to describe the true system dynamics. The first one corresponds to a finite impulse response (FIR) representation:

$$G(z, \theta_{np}) = \sum_{n_p=1}^{n_{\theta_{np}}} g(n_p)z^{-n_p} \quad (3)$$

with  $\theta_{np} = [g(1) \ g(2) \ \dots \ g(n_{\theta_{np}})] \in \mathbb{R}^{n_{\theta_{np}}}$  the parameter vector to estimate in this nonparametric FIR model case and  $g(n_p), n_p = 1, \dots, n_{\theta_{np}}$  denote the impulse response coefficients of the LTI system  $G(z)$ . The second model structure considered is a rational transfer function:

$$G(z, \theta_p) = \frac{\sum_{k=0}^{n_\beta} \beta_k z^k}{z^{n_\alpha} + \sum_{k=0}^{n_\alpha-1} \alpha_k z^k} \quad (4)$$

with  $\theta_p = [\beta_0 \ \beta_1 \ \dots \ \beta_{n_\beta} \ \alpha_0 \ \alpha_1 \ \dots \ \alpha_{n_\alpha-1}] \in \mathbb{R}^{n_{\theta_p}}$  the parameter vector to estimate in this parametric model case and  $n_\alpha > n_\beta$  (causality condition).

*Assumption 3.* The FIR model structure (eq. 3) is considered to be long enough such that model errors due to truncation of the impulse response are negligible.

*Definition 4.* (Input class) The signal used for excitation is a multisine defined as follows:

$$u(k) = \sum_{n=1}^F A_n \cos(\omega_n k + \phi_n) \quad (5)$$

where  $\omega_n, n = 1, \dots, F$  are the frequencies contained in the multisine,  $A_n, n = 1, \dots, F$  denote the amplitude for each frequency contained in the multisine and  $\phi_n, n = 1, \dots, F$  is the phase of each cosine term.

The power spectrum  $\Phi_u$  of the input signal  $u$  is the design variable of the optimal design problem. The discrete power spectrum of this signal (power/frequency) is given by:

$$\Phi_u(\omega_n) = \Phi_u(-\omega_n) = \frac{A_n^2}{4}, \quad n = 1, \dots, F \quad (6)$$

The power of the input signal  $u$  is defined as:

$$p_u = 2 \sum_{n=1}^F \Phi_u(\omega_n) = \sum_{n=1}^F \frac{A_n^2}{2} \quad (7)$$

## 3. INFORMATION IN THE DATA AND PRIOR KNOWLEDGE

### 3.1 Case of no prior knowledge available

Consider the case when the system is modeled with the OE model structure in eq. 2.

*Assumption 5.* In the case when no prior knowledge about the system dynamics is available, an asymptotically efficient estimator is used to estimate the parameter vector  $\theta$  for the model  $G(z, \theta)$  in eq. 2. As such, the covariance matrix of the parameter vector converges to the inverse of the Fisher information matrix for increasing number of samples.

An example of an asymptotically efficient estimator is the Maximum Likelihood (ML) estimator, to obtain an estimate of the parameter vector  $\theta$ , which for the OE case would boil down to:

$$\hat{\theta}_N^{OE} = \arg \min_{\theta} \sum_{k=0}^{N-1} |y(k) - G(z, \theta)u(k)|^2 \quad (8)$$

*Definition 6.* The  $(i, j)$ -element of the Fisher information matrix, denoted by  $M(\theta) \in \mathbb{R}^{n_{\theta} \times n_{\theta}}$ , is given by:

$$M(\theta)_{i,j} = \mathbb{E}\left[\left(\frac{\partial \ln(f_y(Y_N, \theta))}{\partial \theta(i)}\right)\left(\frac{\partial \ln(f_y(Y_N, \theta))}{\partial \theta(j)}\right)\right] \quad (9)$$

where  $Y_N \in \mathbb{R}^N$  is a vector which contains the measured output samples and  $f_y$  corresponds to the probability density function of  $Y_N$ . The mathematical expectation operator is taken with respect to  $Y_N$ .

In case of no prior information available about the underlying dynamics, the total information available at the end of the experiment is represented by:

$$M_{\text{total}} = M(\theta)$$

where  $M(\theta)$  is the Fisher matrix defined in eq. 9. In Ljung, L. (1999) it is shown that the Fisher information matrix can be expressed as a function of the input power spectrum, which for the case of discrete power spectrum and an OE framework considered here, is given by:

$$M = \frac{N}{\sigma_e^2} \sum_{n=-F, n \neq 0}^F \Lambda_G(e^{j\omega_n}, \theta_0) \Lambda_G(e^{j\omega_n}, \theta_0)^* \Phi_u(\omega_n) \quad (10)$$

with  $\Lambda_G(z, \theta) = \frac{\partial G(z, \theta)}{\partial \theta}$ ,  $\theta_0$  denoting the true parameter vector to estimate and  $\Phi_u(\omega)$  the spectrum of the input signal. Eq. 10 shows clearly the linear dependence of the information matrix on the input power spectrum  $\Phi_u$ .

*Assumption 7.* The output of the excited LTI system is transient-free. As a result, no transient response is estimated and the information matrix related to the identification experiment can be constructed based on the assumption of a periodic excitation.

### 3.2 Case of prior knowledge available

Consider the case when the system in eq. 1 is modeled with the FIR model structure in eq. 3. If there is prior knowledge about exponential decay (stability) and smoothness of the underlying impulse response, a regularized FIR model can be estimated using a Bayesian framework as described in Chen, T., Ohlsson, H., & Ljung, L. (2012), Pilonetto, G., Dinuzzo, F., Chen, T., De Nicolao, G., & Ljung, L. (2014). Assuming that the parameter vector  $\theta$  is a Gaussian random variable itself with zero mean and prior covariance matrix  $P \in \mathbb{R}^{n_{\theta} \times n_{\theta}}$ , i.e.  $\theta \sim \mathcal{N}(0, P)$ , and taking into account that:

$$Y_N = K_N \theta + E, \quad e(k) \sim \mathcal{N}(0, \sigma^2)$$

with

$$\begin{aligned} Y_N &= [y(0) \ y(1) \ \dots \ y(N-1)]^T \\ \phi_1(k) &= [u(k-1) \ u(k-2) \ \dots \ u(k-n+1)]^T \\ K_N &= [\phi_1(0) \ \phi_1(1) \ \dots \ \phi_1(N-1)]^T \\ E &= [e(0) \ e(1) \ \dots \ e(N-1)]^T \end{aligned}$$

then  $\theta$  and  $Y_N$  are jointly Gaussian variables:

$$\begin{bmatrix} \theta \\ Y_N \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} P & PK_N^T \\ K_N P & K_N P K_N^T + \sigma^2 I \end{bmatrix}\right)$$

where  $I$  denotes the identity matrix. The posterior distribution of  $\theta$  conditioned on the observed data  $Y_N$  is given by:

$$\theta | Y_N \sim \mathcal{N}(\hat{\theta}_N^{\text{apost}}, P^{\text{apost}})$$

$$\hat{\theta}_N^{\text{apost}} = (K_N^T K_N + \sigma^2 P^{-1})^{-1} K_N^T Y_N \quad (11)$$

$$P^{\text{apost}} = ((\sigma^2 (K_N^T K_N)^{-1})^{-1} + P^{-1})^{-1} \quad (12)$$

The last equation 12 represents the Bayes rule of information. In this work, for the case of FIR estimation in the presence of prior knowledge, the total information available at the end of the experiment is defined as:

$$M_{\text{total}} = (P^{\text{apost}})^{-1} = \frac{1}{\sigma^2} (K_N^T K_N) + P^{-1} = M + P^{-1} \quad (13)$$

where  $M$  denotes the Fisher matrix representing the information linked to the new experiment and  $P^{-1}$  represents the prior information about the identified parameter vector  $\theta$ .

*Remark 8.* In the case when the prior information about the estimated FIR model is related to exponential decay and smoothness of the underlying response, the Diagonal/Correlated (DC) structure for the prior covariance matrix  $P$  can be used Chen, T., Ohlsson, H., & Ljung, L. (2012), Pilonetto, G., Dinuzzo, F., Chen, T., De Nicolao, G., & Ljung, L. (2014):

$$P_{\text{DC}}(i, j) = c \cdot e^{-\alpha|i-j|} e^{-\beta \frac{(i+j)}{2}} \quad (14)$$

where  $0 \leq \alpha, \beta \leq \infty$ ,  $c > 0$ .

*Assumption 9.* In this work the values of the hyperparameters  $c, \alpha, \beta$  are assumed to be known from a prior experiment. As such, the prior information matrix  $P^{-1}$  is considered to be known.

### 3.3 Least Costly OID framework

Given the aforementioned definitions and assumptions, a least costly OID framework is considered as defined below.

*Definition 10.* The optimal input design problem is defined as:

$$\begin{aligned} u_{\text{opt}} &= \arg \min_u p_u \\ \text{s.t. } M_{\text{total}} &> R_{\text{adm}} \end{aligned} \quad (15)$$

where  $p_u$  is the power of the input signal as defined in eq. 7,  $M_{\text{total}}$  represents the total information matrix which corresponds to the information obtained at the end of the identification experiment and  $R_{\text{adm}}$  is a lower bound on  $M_{\text{total}}$ .

The information matrix  $M_{\text{total}}$  can be given either by eq. 10 or eq. 13, depending on the absence or presence of prior knowledge. It is shown in Bombois, X., & Scorletti, G. (2012) that accuracy constraints on the identified model can be transformed into a constraint in the form  $M_{\text{total}} > R_{\text{adm}}$  where  $R_{\text{adm}}$  is in our case a frequency-dependent constraint on the model error achieved at transfer function level of the identified model. This constraint is analyzed in Section 4.

#### 4. BOUND ON MODEL ACCURACY AND THE OID PROBLEM

As mentioned in Section 2, three different cases of model structures are investigated, a FIR model (eq. 3) with and without prior knowledge and the rational transfer function (eq. 4). Note that the number of parameters of the parametric model in eq. 4, namely  $n_{\theta_p}$ , is different than the one for the nonparametric model in eq. 3,  $n_{\theta_{np}}$  (it is considered that  $n_{\theta_p} \ll n_{\theta_{np}}$ ). Therefore the information matrices for the two different model structures will have different dimensions and comparison of the two cases with respect to information cannot be done fairly on a parameter level.

Under this condition, the constraint on the model accuracy used in the optimal input design problem (eq. 15) should be formulated such that the different model structures can be equally compared. The latter can be done by translating the bound on the information matrix ( $M_{\text{total}} > R_{\text{adm}}$ ) to a bound on the model error of the estimated transfer function. As such, the optimal input can be derived such that the same bound on the model accuracy is imposed for different structures.

To formulate the bound on the model accuracy at a transfer function level, the method found in Bombois, X., & Scorletti, G. (2012) (also related to Jansson, H., & Hjalmarsson, H. (2005)) is used and presented briefly in this section. It is shown that, given a constant  $\alpha_p$  related to a chosen probability level, one can define a frequency-by-frequency constraint on the information matrix:

$$M > R_{\text{adm}}(\omega), \quad \forall \omega \quad (16)$$

with  $R_{\text{adm}}(\omega) = \frac{\alpha_p^2}{r_{\text{adm}}^2(\omega)} T^T(e^{j\omega}, \theta_0) T(e^{j\omega}, \theta_0)$ ,

$$T(e^{j\omega}, \theta_0) = \begin{pmatrix} \text{Re}(\Lambda_G^T(e^{j\omega}, \theta_0)) \\ \text{Im}(\Lambda_G^T(e^{j\omega}, \theta_0)) \end{pmatrix} \in \mathbb{R}^{2 \times n_{\theta}} \text{ and } r_{\text{adm}}(\omega)$$

represents the desired bound on the modeling error at each frequency. More precisely, if eq. 16 holds, then it is guaranteed that  $|G_0(e^{j\omega}) - G(e^{j\omega}, \hat{\theta}_N)| < r_{\text{adm}}(\omega), \forall \omega$ , with a certain probability. It can be observed that the higher the value of  $r_{\text{adm}}(\omega)$ , the higher the allowable modeling error resulting in a lower minimum requirement  $R_{\text{adm}}(\omega)$  for the information matrix. The constraint eq. 16 has to be verified for all frequencies. However, in practice, this constraint will only be posed for a finite amount of frequencies covering the interval  $[0 \pi]$ .

The OID problem is defined as:

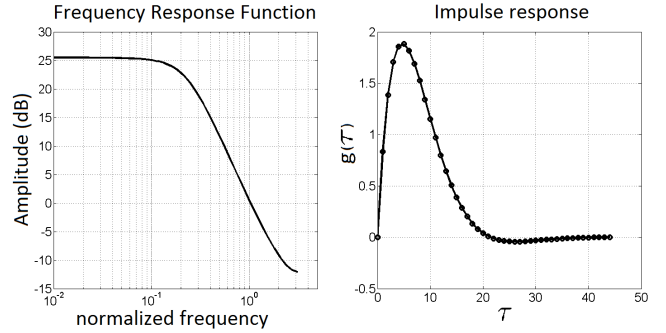


Fig. 2. The linear system used for illustration of the least costly input design framework. Left: The FRF of the linear system  $G(z)$  (frequency normalized to  $\pi$ ). Right: The impulse response  $g(\tau)$  of the linear system  $G(z)$ .

$$\begin{aligned} \Phi_{u,opt}(\omega_n) &= \arg \min_{\Phi_u(\omega_n)} p_u \quad (= \arg \min_{A_1, \dots, A_F} \sum_{n=1}^F A_n^2) \\ \text{s.t. } M_{\text{total}} &> R_{\text{adm}}(\omega), \quad \forall \omega \\ \Phi_u(\omega_n) &\geq 0 \quad \forall \omega_n \end{aligned} \quad (17)$$

which is affine in the input power spectrum  $\Phi_u(\omega_n)$ . The second constraint (non-negativity of the spectrum) is necessary for the signal to be realizable. The optimal design problem (17) is convex in the design variables, therefore there is no risk of resulting in a local minimum. In case of prior information available, the total information matrix is given by  $M_{\text{total}} = M + P^{-1}$ , as defined already in Section 3.2.

*Remark 11.* Prior information about smoothness of the impulse response is directly related to a certain behavior of the modulus of corresponding Frequency Response Function (FRF) (amplitude of transfer function). A smooth impulse response corresponds to a FRF which has low power in the high frequency region. It is shown in Section 5 that this fact is important when bounds related to the amplitude of the transfer function are considered in the OID problem.

*Remark 12.* In eq. 16, the lower bound on the information matrix depends on the true parameter vector  $\theta_0$ . Even though the true vector is used for the numerical illustration in Section 5, in practice a prior estimate of the parameter vector can be used to solve the OID problem and bypass the well-known “chicken-and-egg” problem.

#### 5. NUMERICAL EXAMPLE

The optimal input design problem (eq.17) is applied to the linear system  $G_0(z)$  given by:

$$G_0(z) = \frac{0.8347z}{z^2 - 1.657z + 0.7017}$$

whose FRF  $G_0(e^{j\omega})$  (amplitude) and impulse response  $g(\tau)$  are given in Fig. 2. The input signal is a multisine (eq. 5) and a grid of 40 frequencies, covering the full frequency band  $[0 \pi]$ , are considered to be excited by the input signal. The noise in the measured output is Gaussian i.i.d with variance  $\sigma^2 = 0.25$ . For the bound on the modulus of the estimated transfer function in eq. 16,

the same discrete set of 40 frequencies is chosen on which the constraint will be imposed. The value of  $\alpha_p$  in eq. 16 is chosen, for the three considered situations (see below), in such a way that the modeling error is guaranteed to be smaller than  $r_{adm}(\omega)$ ,  $\forall \omega$ , with a probability of 0.95 (Bombois, X., & Scorletti, G. (2012), page 4).

The parameter  $r_{adm}(\omega)$  is kept constant across the chosen set of frequencies and it can take three values, 0.1, 0.5 and 0.9, resulting in an increasing admissible error in the transfer function modulus, therefore in a smaller and smaller bound on the information matrix. For the case of the FIR model structure and prior knowledge available, the DC covariance structure in eq. 14 is considered for the prior covariance matrix  $P$ , with  $c = 5.6$ ,  $\alpha = 0.033$  and  $\beta = 0.232$ .

The optimal input design problem eq. 17 is considered for three different model structures:

- (1) The system is modeled with a finite impulse response of  $n_{\theta_{np}} = 45$  lags, long enough such that the response decays properly to zero.
- (2) The system is modeled with the same FIR model but prior information about smoothness and decay is available.
- (3) The system is modeled with a rational transfer function complex enough to capture the underlying dynamics of the linear system (three parameters to be estimated as it is the case for the true underlying system) without any more prior information about the underlying dynamics.

The results of the optimal input design problem eq. 17 for all the aforementioned cases are depicted in Fig. 3. It can be observed that:

- (1) The spectrum of the optimal input for the case of FIR modeling approaches the one of a white noise signal, as expected.
- (2) The spectrum of the optimal signal for the case of modeling the linear system with the rational transfer function (IIR) converges to two dominant frequencies which should be optimally excited, as expected for the case of a second order linear system.
- (3) For both aforementioned cases, the constraint on the transfer function of the estimated model affects only the power of the input signal and not the spectrum.
- (4) This is not the case when FIR modeling in the presence of prior knowledge is considered. The spectrum of the input signal is directly related to the constraint imposed on the information matrix. Since the prior information is related to smoothness of the impulse response, it will suppress the power of the estimated transfer function in the high frequency region. Under this condition, the more relaxed the constraint on the modulus of the transfer function, the more is the prior information able to deliver a model inside the allowable model error bounds. Therefore, the power of the optimal input signal decreases in the high frequency region as the constraint on the transfer function modulus relaxes.
- (5) In the right column of Fig. 3, we observe the value of the objective function, namely the total power of the optimal signal for each case of the OID problem. It is clear that modeling the system with a nonparametric

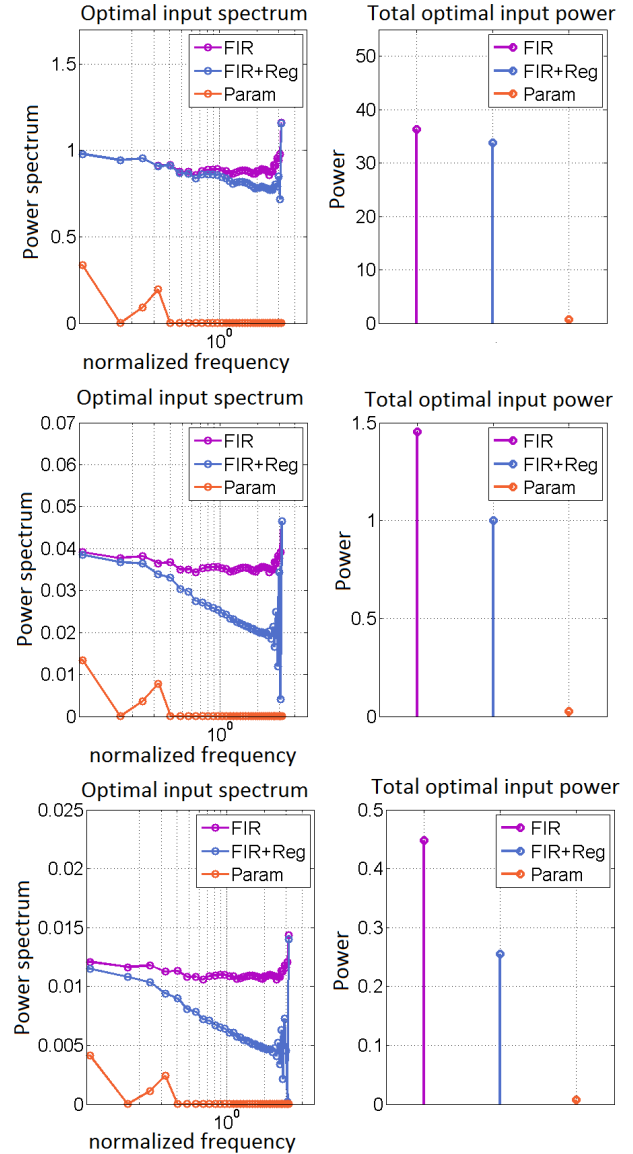


Fig. 3. The optimal input for a linear system in three different cases of model structures and bound on the model accuracy. The admissible modeling error  $r_{adm}$  increases from the upper to the lower row. Left column: Optimized input spectra for the different model structures. Right: Total power of the optimal input signals (the lowest frequency of the 40-frequency grid chosen has been omitted for visualization reasons).

FIR structure under no prior knowledge leads to the largest total power for the optimal signal. Moreover, when prior knowledge is used, the minimum power needed to satisfy the estimated transfer function constraints is reduced. Finally, for the case of parametric modeling and the rational transfer function model, the power of the optimal signal is in every case the minimum one. This result can be intuitively explained by the fact that the transfer function model, complex enough to capture the dynamics of the linear system, contains information about the structure of the system. As such, it constitutes a more compact and informative representation of the linear system

## ACKNOWLEDGEMENTS

This work was supported in part by the Fund for Scientific Research (FWO-Vlaanderen), by the Flemish Government (Methusalem), the Belgian Government through the Inter university Poles of Attraction (IAP VII) Program, and by the ERC advanced grant SNLSID, under contract 320378.

## REFERENCES

- Bombois, X., & Scorletti, G. (2012). Design of least costly identification experiments: The main philosophy accompanied by illustrative examples. *Journal Européen des Systèmes Automatisés (JESA)*, 46(6-7), 587–610.
- Bombois, X., Scorletti, G., Gevers, M., Hildebrand, R., & Van den Hof, P. (2004). Cheapest open-loop identification for control. In *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, volume 1, 382–387. IEEE.
- Bombois, X., Scorletti, G., Gevers, M., Van den Hof, P. M. J., & Hildebrand, R. (2006). Least costly identification experiment for control. *Automatica*, 42(10), 1651–1662.
- Bombois, X., Scorletti, G., Van den Hof, P., & Gevers, M. (2004). Least costly identification experiment for control. a solution based on a high-order model approximation. In *American Control Conference, 2004. Proceedings of the 2004*, volume 3, 2818–2823. IEEE.
- Chen, T., Ohlsson, H., & Ljung, L. (2012). On the estimation of transfer functions, regularizations and gaussian processes - revisited. *Automatica*, 48(8), 1525–1535.
- Forgione, M., Bombois, X., & Van den Hof, P. M. J. (2013). Experiment design for batch-to-batch model-based learning control. In *American Control Conference (ACC), 2013*, 3912–3917. IEEE.
- Fujimoto, Y., & Sugie, T. (2016). Informative input design for kernel-based system identification. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, 4636–4639. IEEE.
- Jansson, H., & Hjalmarsson, H. (2005). Input design via lmis admitting frequency-wise model specifications in confidence regions. *IEEE transactions on Automatic Control*, 50(10), 1534–1549.
- Lataire, J., & Chen, T. (2016). Transfer function and transient estimation by gaussian process regression in the frequency domain. *Automatica*, 72, 217–229.
- Ljung, L. (1999). *System Identification: Theory for the User*, PTR Prentice Hall Information and System Sciences Series.
- Mu, B., & Chen, T. (2017). On input design for regularized lti system identification: Power-constrained input. *arXiv preprint arXiv:1708.05539*.
- Pillonetto, G., & De Nicolao, G. (2010). A new kernel-based approach for linear system identification. *Automatica*, 46(1), 81–93.
- Pillonetto, G., Dinuzzo, F., Chen, T., De Nicolao, G., & Ljung, L. (2014). Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, 50(3), 657–682.
- Wong, W. K. (1994). Comparing robust properties of a, d, e and g-optimal designs. *Computational statistics & data analysis*, 18(4), 441–448.

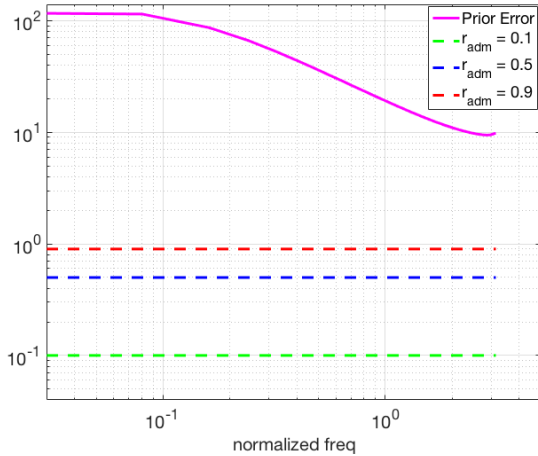


Fig. 4. The maximal modeling error on the modulus of the transfer function, that would be obtained if only the prior information was used, is depicted (solid magenta line). The different levels of admissible error considered in this numerical example are also plotted and compared with the maximal error  $r_{\text{adm}}$ .

and can be considered as extra prior information introduced in the optimal input design problem.

The optimal spectrum obtained for the case when prior information about smoothness is available, and shown in Fig. 3, left column, in blue, can be explained by Fig. 4. In the latter plot, the maximal modeling error on the modulus of the transfer function, that would be obtained if only the prior information was used, is depicted (solid magenta line). This error corresponds to  $\alpha_p \sqrt{\lambda_{\max}(T(e^{j\omega}, \theta_0) P T^T(e^{j\omega}, \theta_0))}$ , where  $\lambda_{\max}(A)$  corresponds to the largest eigenvalue of a matrix  $A$  and  $P$  is the prior covariance matrix. It can be observed that in the high frequency region, the gap that has to be filled by the optimal input is smaller than for the lower frequencies. Consequently, the identification experiment will essentially have to bring information for low frequencies (resulting in a spectrum which is essentially located in these lower frequencies). This phenomenon will be increasingly important for increasing values of  $r_{\text{adm}}$ . This is precisely what is observed for the blue curves in the left part of Fig. 3.

## 6. CONCLUSION

The effect of prior knowledge, about the system dynamics, on the minimum power excitation has been investigated in this work. The optimal input design has been formulated in a way that prior knowledge can be easily incorporated in the design problem. The input excitation belongs to the class of multisine signals and the spectrum of the signal is the optimizing design variable. The constraints of the OID problem are related to the bounds on the error of the transfer function modulus between the true system and the estimated model. It has been shown that prior knowledge related to smoothness of the impulse response affects the spectrum of the optimal input excitation by decreasing the power of the signal in the high frequency region.