Université de Lyon
CNRS, Ecole Centrale Lyon, INSA Lyon, Université Claude Bernard Lyon 1

Laboratoire Ampère
Unité Mixte de Recherche du CNRS - UMR 5005
Génie Electrique, Automatique, Bio-ingénierie

Mémoire doctorant 1ère année
2017 -2018

Nom - Prénom	**Morelli Federico**
email	**federico.morelli@ec-lyon.fr**
Titre de la thèse	**Optimal experiment design for the identification of one module in the interconnection of locally controlled systems**
Directeur de thèse	**Bombois Xavier et Bako Laurent**
Co- encadrants	**Korniienko Anton**
Dpt. de rattachement	**MIS**
Date début des travaux	**01/11/2017**
Type de financement	**Contract de recherche ECL, dans le cadre du projet NEXT4MEMS financé par la BPI**

Laboratoire Ampère – Ecole Centrale de Lyon – 36, avenue Guy de Collongue - 69134 Ecully cedex – France
Tél : +33 (0) 4 72 18 60 99 Fax : +33 (0) 4 78 43 37 17 http://www.ampere-lab.fr

Abstract

This document reports the work done during the first year of the research activity. The concepts acquired regarding the System Identification are reported, together with their connection to the problem of designing the identification experiment. Then the literature regarding the identification of the interconnection of locally controlled systems is reviewed and the main problem subject of this activity is stated. Finally, it is proposed a solution that represents a novelty in the robust experiment design literature, obtained for a simple case, and are presented some simulative results.

Contents

List of Symbols

$E(x)$	Expected value of x
$\bar{E}(x)$	$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E(x(i))$
$A > 0$	Where A is matrix: A is positive definite
$\bar{\lambda}(A)$	Largest eigenvalue of A
$Re(x)$	Real part of x
$Im(x)$	Imaginary part of x
A^*	Transpose conjugate of A
$\|x\|$	Module of x: $\sqrt{Re(x)^2 + Im(x)^2}$
$\|v\|_2$	Norm 2 of v: $\sqrt{\sum_{k=1}^{n} \|v_k\|^2}$
$\chi^2(k)$	Chi-squared probability density function of k degrees of freedom
$diag(d_1, d_2, \dots d_n)$	Block diagonal matrix: $\begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}$

Introduction

This report regards the first year of activity of the doctoral program titled "Optimal experiment design for the identification of one module in the interconnection of locally controlled systems". This activity is one asset of the MIS (Méthodes pour l'ingénierie des systèmes) department of Laboratoire Ampère, in particular the axis "large-scale and networked systems" of the priority M1.

Nowadays, distributed systems are used in many fields, from power grids to monitoring a post-disaster scenario e.g. finding survivor after an avalanche [1]. Let us consider the latter example where a group of drones must coordinate themselves in order to optimally cover the area of interest and effectively help the rescuers. This type of task is usually faced as a problem of formation control: every drone has its own controller, but its desired position is obtained according to the position of the other agents. Every agent communicates its own position to the others according to the network topology, then the desired position is obtained using a consensus algorithm that takes also into account the shape of the desired formation e.g. a semi-circle. Due to the interconnected nature of the whole system, an undesired behaviour of one agent will affect also the rest of the network.

In an emergency environment the service continuity of the distributed system is really important. An agent may encounter a situation that leads to a loss of local performance, thus affecting the rest of the network and then the global performance of the distributed system. In this case, it is necessary to re-identify a model for this agent in order to restore the performance of the whole system.

The identification of a dynamical network is a really recent and challenging research topic [18],[8]. Recently, the conditions for obtaining a consistent estimate, when dealing with the identification of only one module, have been obtained. This is a non-trivial problem, since the consistency of the estimate depends on which signals are measured in the network, thus posing important issues in the experiment design [11]. The main goal of this research activity is to design the experimental conditions for the identification of only one module, under the requirement that the they should affect as least as possible the rest of the network. This can be seen as an extension of the least costly experiment design [3], indeed we want to identify the true system while the plant is operating i.e. without stopping the production. In fact, the identification experiment on one system of the network requires to perturb it in a controlled way, thus affecting its normal operation. Its perturbed output is then used by the

other systems in the network for their own control purpose, e.g. formation control. So the whole network will be perturbed by the identification experiment of only one module, compromising the normal operation of the whole network. Moreover, since this type of configuration is widely used for autonomous systems [9], we have to take into account that the energy used for the generating the excitation signal represents a cost. This type of problem falls under the optimal experiment design framework, since we want to minimize the cost of the identification experiment. Up to our knowledge, the only contributions on optimal experiment design for dynamical networks are [4] and [12]. In [4] have been proposed techniques to identify a whole network, so every system is excited, thus affecting the service continuity.

In order to limit the effect of the experiment of the rest of the network, we will have to deal with the output signals of the agents inside the network and define some constraints on them, thus needing the knowledge of the unknown true systems. This means that it is not possible to verify whether the constraints are satisfied or not. In literature this type of problem is tackled in the robust experiment design literature. These techniques are based on finding a tractable version of the constraint, robust w.r.t. the unknown true system, using the prior information about the system, in order to assure that the actual constraint will be satisfied. The constraint is then said to be robustified. However, the solutions proposed in the literature rely on an approximation of the constraint, and thus they do not fully guarantee the actual ones. The second goal of this research activity is to propose a robustification technique that does not approximate the constraint.

This report is structured as follows:

Section 1 The theoretical framework on optimal experiment is presented, together with the needed concept of System Identification and Prediction Error criterion.

Section 2 The structure for the interconnection of locally controlled systems is presented, together with the formulation of the identification procedure for this type of network. Then the main problem is presented and treated in details for a simple case. Afterwards, two solutions for effectively robustifying the constraint are presented, together with simulative results for comparing them;

Section 3 The next step of this research activity are here presented.

1 Experiment Design

Models are needed in every fields of science and many industrial applications. The purpose of the model can be, for example, control design, fault detection, filtering, etc. Building a model starting from physical laws may lead, sometimes, to a model that is too complex for the desired application, or which involves constants hard to estimate. For these reasons we are interested in building a parametric model relating only the input and output quantities of interest, directly from experimental data. Many techniques for estimating parametric models from experimental data were developed in the literature [15]. Here we will restrict our attention only to the estimation of linear transfer functions that relate the output signal $y(t)$ to the controlled input signal $u(t)$ and the disturbances $v(t)$. Let us consider the data generating system:

$$y(t) = G_0(z)u(t) + v(t), \tag{1}$$

where $G_0(z)$ is a linear transfer function and $v(t) = H_0(z)e(t)$, where $e(t)$ is a zero-mean white noise with variance σ_e^2 and $H_0(z)$ a linear transfer function. The quantity $u(t)$ represents the user-controlled input signal, while $v(t)$ represents all the unmodeled inputs and noises. Then the true system is denoted by $\mathcal{S} = \{G_0(z), H_0(z)\}$. The user must choose a parametric model structure

$\mathcal{M} = \{G(z,\theta), H(z,\theta)\}$, where $\theta \in \mathbb{R}^k$ is the vector containing the parameters of G and H (named the parameter vector), suitable for the application. If the model structure is able to fully represents the true system behaviour, then $\mathcal{S} \in \mathcal{M}$, i.e. $\exists\theta_0 \in \mathbb{R}^k$, named the true parameter vector, s.t. $G(z,\theta_0) = G_0(z), H(z,\theta_0) = H_0(z)$. The model structure is then said to be a *full-order model structure*. Once we have chosen $\mathcal{M}$ and we have collected experimental data, by measuring the input and output signal of $\mathcal{S}$ during an experiment, we have to choose an identification criterion for identifying a model in $\mathcal{M}$, i.e. identifying a parameter vector $\hat{\theta}$, that describes the observed experimental behaviour. To summarize, system identification is made of three main ingredients: experimental data, a model structure and an identification criterion.

The model has also to be a "good" model for the application i.e. must have a certain accuracy over the frequency band of interest. The accuracy of the identified model, $G(z,\hat{\theta})$ and $H(z,\hat{\theta})$, is related to the experimental conditions [15], so it is possible to design them in order to obtain a model appropriate for the intended application [13]. However, there may be some limitations in the experimental setting, such as a limitation on the maximum power for the excitation signal $u(t)$. The goal of the experiment design is to find the experimental conditions such that the identified model will satisfy a given quality requirement, while respecting given constraints on the experiment.

The experiment design framework for the Prediction Error identification criterion will be presented in the following parts of this section. The identification of networked systems will be presented in the remainder of this report.

1.1 Prediction Error

Let us consider the true system $\mathcal{S} = \{G_0(z), H_0(z)\}$, as in (1), represented in Figure 1 and assume $G_0(z)$ to be monic and stable, while $H_0(z)$ to be stable, monic and inversely stable. The parametric model structure is given by $\mathcal{M} = \{G(z,\theta), H(z,\theta)\}$, where $\theta \in \mathbb{R}^k$ is the parameter vector, and it is assumed that $\mathcal{S} \in \mathcal{M}$.

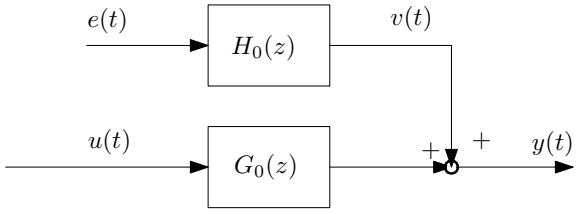


Figure 1: Schematic of the data generating system.

The system is excited with the input signal $u(t)$ and N input-output data points are collected in $Z^N = \{u(t), y(t) \quad t = 1..N\}$. For a given θ and Z^N it is possible to compute the prediction error, defined as follows:

$$\epsilon(t,\theta) := H^{-1}(z,\theta)(y(t) - G(z,\theta)u(t)) \tag{2}$$

This definition comes from the difference between the measured output of the system, $y(t)$, and its prediction made from the model [15]. For the dataset Z^N and for a given θ we can compute the following function of the prediction error:

$$V_N(Z^N,\theta) := \frac{1}{N}\sum_{t=1}^{N}\epsilon^2(t,\theta) \tag{3}$$

The PEM, *Prediction Error Method*, takes as identification criterion the minimization of $V_N(Z^N, \theta)$. So, the identified parameter vector $\hat{\theta}_N$ is obtained by solving the following minimization problem, over the data set Z^N:

$$\hat{\theta}_N = \arg\min_{\theta} V_N(Z^N, \theta) \tag{4}$$

If the input signal is informative enough and the model structure is identifiable [15], then $\hat{\theta}_N$ is unique for a given dataset Z^N and will tend to θ_0 with probability 1 as $N \to \infty$. However, due to the noise corrupting the data, $\hat{\theta}_N$ is a random variable and has the following statistical property:

$$\sqrt{N}\left(\hat{\theta}_N - \theta_0\right) \sim AsN(0, P_\theta) \tag{5}$$

This means that the difference $\sqrt{N}\left(\hat{\theta}_N - \theta_0\right)$ is asymptotically normally distributed with asymptotic covariance matrix P_θ, given by the following relation [15]:

$$P_\theta = \sigma_e^2 \left[\bar{E}\psi(t, \theta_0)\psi(t, \theta_0)^T\right]^{-1} \tag{6}$$

where $\psi(t, \theta) = -\frac{d\epsilon(t,\theta)}{d\theta}$. Starting from P_θ and using the asymptotic properties of $\hat{\theta}_N$, it is possible to define, for N sufficiently large, a confidence region U for $\hat{\theta}_N$, which contains θ_0 with a chosen probability level α [15]:

$$U := \left\{\theta | (\theta - \hat{\theta}_N)^T P_\theta^{-1} (\theta - \hat{\theta}_N) < \chi_\alpha\right\} \tag{7}$$

where χ_α is defined as $Pr(\chi^2(k) < \chi_\alpha) = \alpha$. The set U is an ellipsoid in the parameter space, shaped by P_θ^{-1}. The higher the eigenvalues of P_θ^{-1}, the smaller the set U. According to (6), P_θ depends on two unknown quantities: θ_0 and σ_e^2. However, it is possible to compute an approximation of P_θ^{-1}, by replacing θ_0 and σ_e with their estimates, respectively $\hat{\theta}_N$ and $\hat{\sigma}_e^2 = V_N(Z^N, \hat{\theta}_N)$.

Now, we want a relation between P_θ^{-1} and the experimental conditions, the experiment length N and the applied excitation signal $u(t)$, allowing us to shape the uncertainty region by modifying the experimental conditions. In [15] a relation which connects P_θ^{-1} with the frequency domain properties of the model and of the signals is presented:

$$P_\theta^{-1} = \frac{N}{\sigma_e^2 2\pi} \int_{-\pi}^{\pi} F_u(e^{j\omega}, \theta_0) F_u^*(e^{j\omega}, \theta_0) \Phi_u(\omega) d\omega + \frac{N}{2\pi} \int_{-\pi}^{\pi} F_e(e^{j\omega}, \theta_0) F_e^*(e^{j\omega}, \theta_0) d\omega \tag{8}$$

where $F_u(z, \theta) = H(z, \theta)^{-1} \frac{dG(z,\theta)}{d\theta}$, $F_e(z, \theta) = H(z, \theta)^{-1} \frac{dH(z,\theta)}{d\theta}$ and $\Phi_u(\omega)$ the power spectrum of $u(t)$.[1] The inverse of the covariance matrix represents the amount of information available, so, according to (8), increasing the experiment duration N will increase our knowledge of the system, thus reducing the dimension of the uncertainty set. Moreover, increasing the total power of the excitation signal will also increase the amount of information, as well as choosing an appropriate power spectrum $\Phi_u(\omega)$. Indeed, by knowing $F_u(z, \theta)$, it is possible to find at which frequencies the power spectrum will have a greater effect on P_θ^{-1}. Then, (8) is a fundamental equation when dealing with the experiment design problem.

Let us consider a generic industrial plant. Performing an identification using the procedure just presented requires to stop the production for performing the experiment, thus the experiment has a

[1] The power spectrum is a positive function $\Phi_u : [-\pi, \pi] \to \mathbb{R}_{\geq 0}$, defined as the Fourier Transform of $R_u(\tau) := \bar{E}(u(k)u(k-\tau))$ for $u(k)$ a *quasi-stationary* signal [15]

cost related to the lost in the productivity. Moreover, also the power of the excitation signal represents a cost, since energy is not free, in particular for autonomous systems, since they have a limited amount of energy available. It is then useful to define a cost of the identification experiment as a function of the experimental conditions $\mathcal{F}(N, \Phi_u(\omega))$. Moreover, we usually desire to identify a model accurate enough for the intended application [13]. Since the accuracy of the identified model depends on the set U, we may require it to be contained in a set U_{app} which is the admissible uncertainty region for the application [13]. It is then interesting to formulate the following optimal experiment design problem:

Problem 1. Let us consider a cost function of the identification experiment, $\mathcal{F}(N, \Phi_u(\omega))$, and an accuracy requirement $U \subseteq U_{app}$, where U is defined as in (7) and U_{app} is dependent on the application of the identified model [13]. Find the experimental conditions that lead to the least costly experiment, while satisfying the accuracy requirement for the to-be-identified model.

It is straightforward to rewrite this problem as the following minimization problem:

$$\min_{\Phi_u, N} \mathcal{F}(N, \Phi_u) \tag{9}$$

$$U \subseteq U_{app} \tag{10}$$

Since the experimental conditions obtained are optimal w.r.t. the experiment cost, this problem is called *least costly experiment design problem*. Note that there may be also additional constraints related to the limitations of the experimental apparatus [14]. Moreover, we want the optimization problem to be convex [13], so the constraints are usually relaxed to a convex tractable formulation.

1.2 Least Costly Experiment Design for closed loop systems

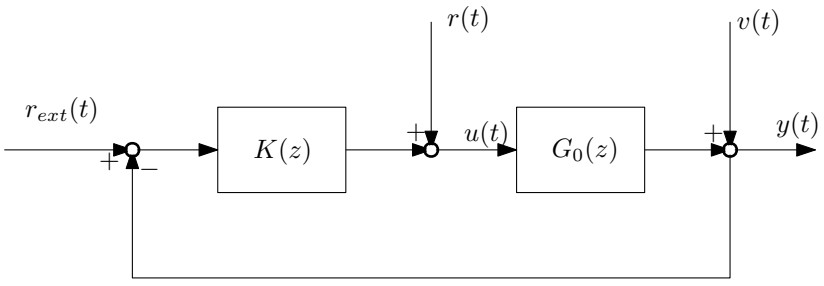


Figure 2: Closed loop setting

Let us consider that the system $\mathcal{S} = \{G_0(z)H_0(z)\}$ must be operated according to certain requirements e.g. tracking an external reference $r_{ext}(t)$ or rejecting the external disturbances $v(t)$. For this reason it is necessary to operate it in closed loop, i.e. with a control law that uses the output signal in order to adapt the input signal according to the requirements. Let us restrict our attention to control laws given by a transfer function $K(z)$, named the controller, as in Figure 2. Then the input $u(t)$ and the output $y(t)$ of the system are given by:

$$\begin{cases} y(t) = S(z)v(t) + G_0(z)S(z)r(t) \\ u(t) = K(z)S(z)r_{ext}(t) - K(z)S(z)v(t) + S(z)r(t) \\ \text{where } S(z) = \frac{1}{1+G_0(z)K(z)} \end{cases} \tag{11}$$

Suppose we want to identify the parameter vector $\hat{\theta}_N$ for the model structure $\mathcal{M}$ as in the previous section. Let us then set $r_{ext}(t)$ equal to 0, excite the system with $r(t)$ and collect N input-output data points in $Z^N = \{u(t), y(t) \quad t = 1..N\}$. Then it is possible to identify $\hat{\theta}_N$ using (4). If the signal $u(t)$

is informative enough, the model structure is of full-order and identifiable, then we have an unique $\hat{\theta}_N$ for a given dataset Z^N and we can define an uncertainty region U as in (7). Note that in this case the signal $u(t)$ is correlated to the disturbance $v(t)$, this fact is reflected on the technical conditions for having a consistent estimate [15]. The relation for computing the covariance matrix is slightly different from (8) and is given by:

$$P_\theta^{-1} = \frac{N}{\sigma_e^2 2\pi}\int_{-\pi}^{\pi} F_r(e^{j\omega},\theta_0)F_r^*(e^{j\omega},\theta_0)\Phi_u(\omega)d\omega + \frac{N}{2\pi}\int_{-\pi}^{\pi} F_v(e^{j\omega},\theta_0)F_v^*(e^{j\omega},\theta_0)d\omega \tag{12}$$

where $F_r(z,\theta) = S(z)H^{-1}(z,\theta)\Lambda_G(z,\theta)$, $F_v(z,\theta) = H^{-1}(z,\theta)\Lambda_H(z,\theta) - K(z)S(z)\Lambda_G(z,\theta)$, $\Lambda_G(z,\theta) = \frac{\partial G(z,\theta)}{\partial\theta}$, $\Lambda_H(z,\theta) = \frac{\partial H(z,\theta)}{\partial\theta}$. Nevertheless is still possible to use this relation for the experiment design problem, as in the previous section. The configuration in Figure 2 is typical of many industrial plants. Sometimes the disconnection of the system from the controller may require to stop the production for an unacceptable time. On the other hand, performing the identification during the normal operation of the system could spoil the production, since we add the signal $r(t)$ that is equal to 0 during the normal operation. So, it is really interesting to search the experimental conditions that affect as least as possible the normal operation of the system, while obtaining a model with the desired accuracy.

To this aim, in [3] the cost of the identification is determined by the signals $u_r(t)$ and $y_r(t)$, the effects of the excitation $r(t)$ on the signals $u(t)$ and $y(t)$ respectively. So, considering a fixed experiment length N, the experiment cost $\mathcal{F}(\Phi_r)$ is defined as a weighted sum of the power of $u_r(t)$ and $y_r(t)$:

$$\begin{aligned}\mathcal{F}(\Phi_r) &= \alpha_y \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{y_r}(\omega)d\omega + \alpha_u\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{u_r}(\omega)d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\alpha_y|G_0(e^{j\omega})S(e^{j\omega})|^2 + \alpha_u|S(e^{j\omega})|^2\right)\Phi_r(\omega)d\omega\end{aligned} \tag{13}$$

where $\Phi_{y_r}(\omega)$ and $\Phi_{u_r}(\omega)$ are the power spectrum of $u_r(t)$ and $y_r(t)$ respectively, while α_y and α_x are weight coefficients chosen by the user.

The accuracy requirement can also be obtained considering that the model will be used for designing an improved controller $\hat{K}(z)$. Then the constraint $U \subseteq U_{app}$ is formulated such that the new controller will satisfy certain performance requirement when connected to the true system. This formulation is obtained in [3] using the results of Robust Control theory, then connecting the performance requirement to the uncertainty region U. We can now formulate the following least costly experiment design problem:

Problem 2. Let us consider the cost function of the identification experiment, $\mathcal{F}(\Phi_r)$ defined in (13), and an accuracy requirement $U \subseteq U_{app}$, where U is defined as in (7) and U_{app} is dependent on the application of the identified model as in [3]. Find the power spectrum $\Phi_r(\omega)$ that leads to the least costly experiment, while satisfying the accuracy requirement for the to-be-identified model.

It is straightforward to rewrite this problem as the following minimization problem:

$$\min_{\Phi_r} \mathcal{F}(\Phi_r) \tag{14}$$

$$U \subseteq U_{app} \tag{15}$$

Note that, while the statement of the problem remains the same as Problem 1, the experimental configuration considered here allows us to perform the identification experiment without stopping the production. This fact has really important consequences, since it allows us to implement an algorithm such that the plant can (semi)autonomously perform the identification experiment whenever it is necessary e.g. change in the performance due to aging phenomena.

2 Optimal experiment design for the interconnection of locally controlled systems

Dynamic networks can be seen as complex structured systems, where the nodes are the internal signals and the connections between them are represented by dynamic systems. These interconnections may generate a complex closed loop configuration for each system. Starting from this observation, in recent literature the PEM has been extended to dynamic networks [18] using the known results for closed loop identification [15]. However, when we want to identify only one system the choice of the signals to measure affects the consistency of the estimate [11]. In fact, when dealing with a network, it may be impossible to measure every node inside it, so the signals to be measured must be chosen in order to guarantee the consistency of the identification. To this aim, in [8] are provided sufficient conditions on the set of signals to measure, in order to obtain a consistent estimate of one system. However, the interconnected nature presents some challenges due to the propagation of the signal through the network. In fact, the prediction error (2) is defined for one node that must be selected appropriately in order to identify the systems of interest [8],[11]. Moreover, in order to guarantee consistency, the formulation of the prediction error may require to add an excitation signal to more than one node, then requiring a parametric model for every system in the path from the excitation signal to the one for which we have defined the prediction error. Moreover, we have to take into account that the excitation signal will be propagated also towards other nodes in the network, thus affecting their operation. For this reason, it is interesting to consider least costly experiment design for reducing the effects of the excitation signal in the network. To our knowledge, optimal experiment design for a dynamic network has been treated only in [12] and [4].

2.1 System Identification for the Interconnection of Locally Controlled Systems

Let us consider the interconnection of N_m modules as in [4]. The i-th module is made of a system $\mathcal{S}_i = \{G_{i,0}(z), H_{i,0}(z)\}$, controlled in closed loop by its own controller $K_i(z)$ as in Figure 3.

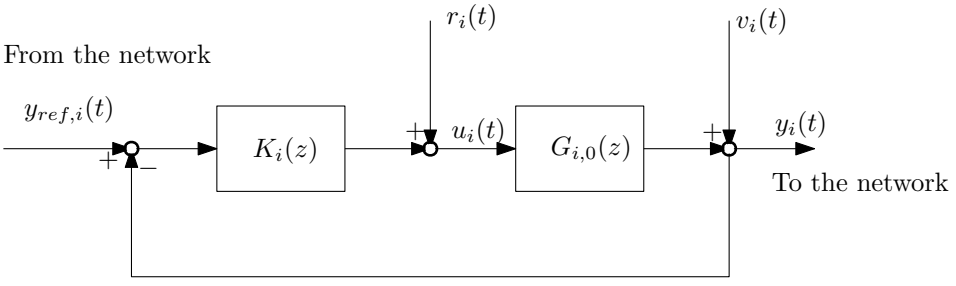


Figure 3: Schematic of one system in the network

Note that this configuration is identical to the one in Figure 2, with the only difference that the system receives the external reference, here $y_{ref,i}(t)$, from the network and sends its output $y_i(t)$ to

the network. The external reference $y_{ref,i}(t)$ is considered to be given by a consensus strategy that depends on the outputs of the other systems $y_i(t)$ and on the global external scalar reference $ref_{ext}(t)$ given to the network. The signal $u_i(t)$ is the input of $G_{i,0}(z)$, while $r_i(t)$ is the external excitation signal, used only during the identification experiment, of the i-th module. Then, the governing equations of the system $\mathcal{S}_i$ are defined as follows:

$$y_i(t) = G_{i,0}(z)u_i(t) + v_i(t) \tag{16}$$

$$u_i(t) = r_i(t) + K_i(z)(y_{ref,i} - y_i(t)) \tag{17}$$

$$\bar{y}_{ref}(t) = A_d\bar{y}(t) + B_d ref_{ext}(t) \tag{18}$$

where $\bar{y}_{ref}(t) = (y_{ref,1}(t), y_{ref,2}(t), \ldots, y_{ref,N_m}(t))$ and $\bar{y}(t) = (y_1(t), y_2(t), \ldots y_{N_m}(t))$. Every disturbance term $v_i(t)$ is given by $v_i(t) = H_{i,0}(z)e_i(t)$ where $H_{i,0}(z)$ is the noise model of the i-th module and $e_i(t)$ is a white noise process. The quantities A_d and B_d are, respectively, the normalized adjacency matrix of the network, which represents the consensus strategy implemented in the network, and the input vector for $ref_{ext}(t)$. In the network here considered, it is assumed that the i-th module computes $y_{ref,i}(t)$ as an average of the measures received from its neighbours without any delays. In order to describe this algorithm as in (18), we need to define the in-degree matrix D and the adjacency matrix A of the graph associated to the network.

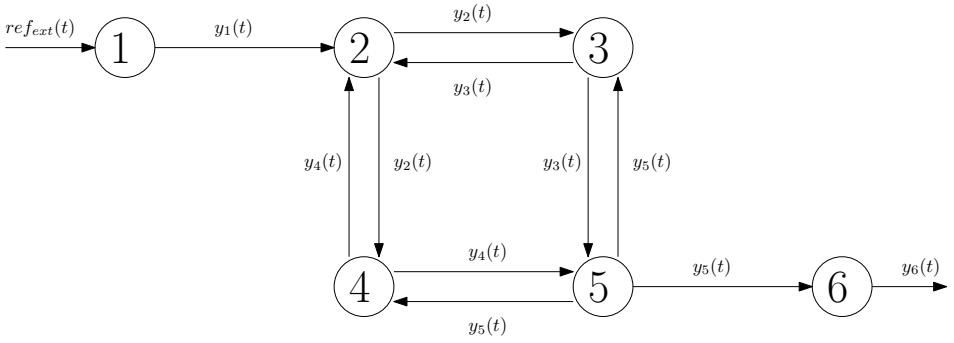


Figure 4: Example of graph associated to the network.

Let us consider the graph associated to the network as in Figure 4, where the nodes are the modules and the edges are the communication links between the modules. The adjacency matrix $A \in \mathbb{R}^{N_m \times N_m}$ is defined such that if there exists an edge from the node j to the node i, then the i,j-th entry of A is equal to 1. If for every edge connecting the node j to the node i, there exists also an edge connecting the node i to the node j, then the graph is said to be undirected and the matrix A is symmetric. In addition, let $D \in \mathbb{R}^{N_m \times N_m}$ be the in-degree matrix of the graph, defined as $D = diag\{d(1), d(2), \ldots, d(N_m)\}$, where $d(i)$ is the number of edges directed to the node i. The normalized adjacency matrix A_d used in (18) can be computed as $A_d = D^{-1}A$. Then, for the topology represented in Figure 4 we have:

$$Ad = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad Bd = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{19}$$

The whole network is represented in Figure 5 as a MIMO system (Multiple Input, Multiple Output) $\mathcal{S} = \{G_0(z), H_0(z)\}$ operated in closed loop, where $G_0(z) = diag(G_{1,0}(z), \dots G_{N_m,0}(z))$ and $H_0(z) = diag(H_{1,0}(z), \dots H_{N_m,0}(z))$.

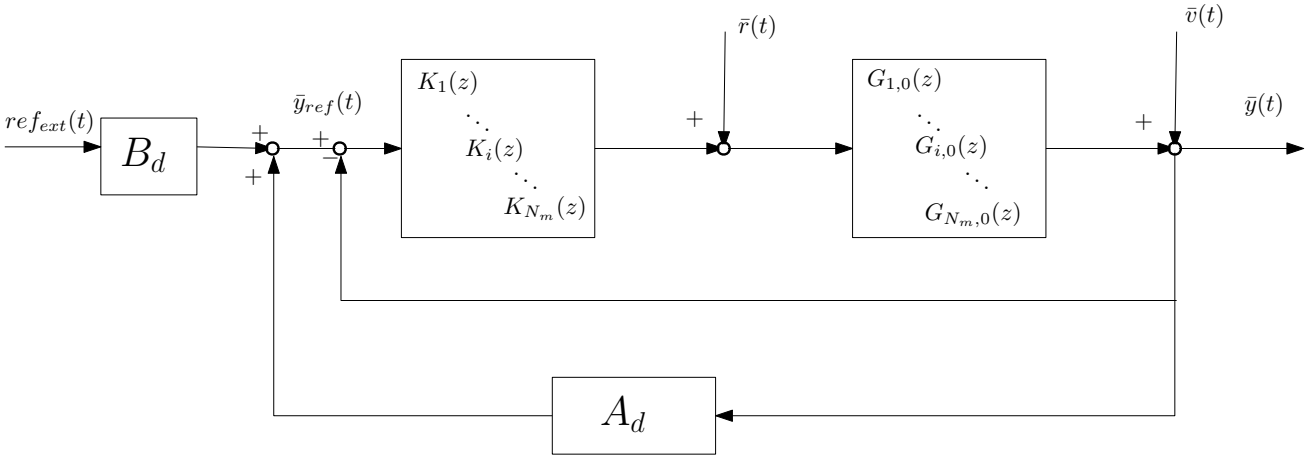


Figure 5: Schematic of the whole network

The covariance matrix of the vector $\bar{e}(t) = (e_1(t), e_2(t), \dots e_{N_m}(t))$, given by $E\bar{e}(t)\bar{e}(t)^T = \Lambda$, is positive definite and it is diagonal when signals $e_i(t)$ are mutually independents. Let us assume that we have access to the signals $u_i(t)$ and $y_i(t)$ for every module and let $ref_{ext}(t)$ be set to 0. Then we excite the network using the signals $r_i(t)$ and we collect N input output data points for each module $Z_i^N = \{u_i(t), y_i(t) \quad t = 1 \dots N\}$. Let $\mathcal{M}_i = \{G_i(z, \theta_i), H_i(z, \theta_i)\}$, where $\theta_i \in \mathbb{R}^{k_i}$, be a full-order parametric model structure for i-th module and $\theta_{i,0}$ the true parameter vector s.t. $G(z, \theta_{i,0}) = G_{i,0}(z), H(z, \theta_{i,0}) = H_{i,0}(z)$. It is possible to identify a global parameter vector $\hat{\theta} = \{\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_{N_m}\}$ through a PEM for MIMO systems. However, this procedure leads to a minimization problem involving a large number of parameters. Under certain conditions on the matrix Λ, provided in [4], the global parameter vector $\hat{\theta}$, identified with the MIMO formulation of PEM, is equal to define separately the prediction error for each module $\epsilon_i(\theta_i, t)$ and then identify separately each parameter vector $\hat{\theta}_i$ with the SISO formulation of PEM. This fact implies that, under the conditions given in [4], it is possible to identify only one module without performing the identification of the whole network. The signal $u_i(t)$ can be expressed as a function of the signal applied to the network during the identification experiment:

$$u_i(t) = \sum_{j=1}^{N_m} (R_{i,j}(z) r_j(t) + Q_{i,j}(z) e(t)) \tag{20}$$

for given transfer functions $R_{i,j}(z)$ and $Q_{i,j}(z)$ that can be obtained from the network [4]. The transfer functions $R_{i,j}(z)$ and $Q_{i,j}(z)$ are non zero if and only if there exists a path from the node j to the node i. This relation implies that every signal $r_j(t)$ that have a path to the node i are involved in expression of $\epsilon_i(\theta_i, t)$. So, assuming that the signals $r_i(t)$ are mutually independent, the inverse of the covariance matrix for $\hat{\theta}_i$ has the following form [4]:

$$P_{\theta_i}^{-1} = \frac{N}{2\pi\Lambda_{i,i}} \int_{-\pi}^{\pi} F_i(e^{j\omega}, \theta_{i,0}) F_i^*(e^{j\omega}, \theta_{i,0}) \left(\sum_{j=1}^{N_m} |R_{i,j}(e^{j\omega})|^2 \Phi_{r_i}(\omega) \right) d\omega + \frac{N}{2\pi\Lambda_{i,i}} \int_{-\pi}^{\pi} \mathcal{Z}_i(e^{j\omega}, \theta_{i,0}) \Lambda \mathcal{Z}_i^*(e^{j\omega}, \theta_{i,0}) d\omega \tag{21}$$

where $F_i(z,\theta_i) = H_i(z,\theta_i)^{-1}\frac{dG_i(z,\theta_i)}{d\theta_i}$, $\Lambda_{i,i}$ is the i,i-th entry of Λ and $\mathcal{Z}_i$ a transfer matrix of dimension $k_i \times N_m$. The i-th column of $\mathcal{Z}_i$ is equal to $L_i + F_i Q_{i,i}$, with $L_i = H_i(z,\theta_i)^{-1}\frac{dG_i(z,\theta_i)}{d\theta_i}$, and the j-th column, for $j \neq i$, is $F_i Q_{i,j}$. It is then possible to define an uncertainty region U_i for $\hat{\theta}_i$, that contains $\theta_{i,0}$ with a given probability level β, as in (7):

$$U_i := \left\{\theta_i | (\theta_i - \hat{\theta}_i)^T P_{\theta_i}^{-1} (\theta_i - \hat{\theta}_i) < \chi_\beta\right\} \tag{22}$$

where χ_β is defined as $Pr(\chi^2(n_\theta) < \chi_\beta) = \beta$ and n_θ is the dimension of $\hat{\theta}$ [4].

Instead of designing the excitation signals $r_i(t)$ for every system, it is possible to identify every module by applying the excitation only to one node, as long as the conditions for consistency in [4] are satisfied. Moreover, in (21), the excitation signal $r_j(t)$, $j \neq i$, is always present as long as there exists a path from the node j to the node i. This is a great advantage, since every excitation signal $r_j(t)$ contributes to the reduction of the size of the uncertainty region U_i. On the other hand, the covariance matrix P_{θ_i} depends also on the other true systems $G_{j,0}$. In addition, the signal propagation through the network is a problem when we want to identify only one module, while not affecting the normal operation of the rest of the network.

2.2 The problem

In recent literature there are only two contributions on optimal experiment design for dynamic networks, [4] and [12]. In [12] both the excitation signal and an ad-hoc open loop controller are designed for performing the identification experiment of one module. In [4] the network considered is the same as the one presented above, but the problem addressed is that of identifying the whole network. Both contributions solve the optimal experiment design problem in an application oriented framework [13], then giving an important contribution to the extension of this framework to dynamical networks. The problem considered in this thesis has a different goal: we want to identify only one module of the network, without affecting the behaviour of the other systems. The main philosophy is similar to Section 1.2, where the identification experiment was performed during the normal operation of the plant, so the applied excitation signal had to affect as least as possible the output of the plant. In this case, the interconnected nature represents an important challenge: every time a system $\mathcal{S}_i$ is excited with $r_i(t)$, also the rest of the network will be affected (20).

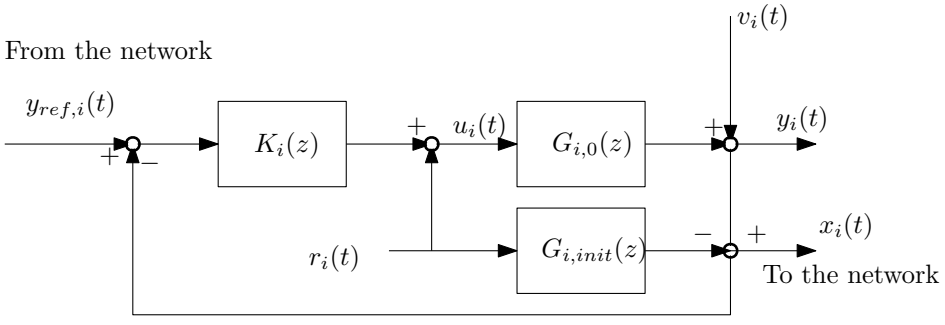


Figure 6: The proposed solution of modification of the modules.

Let us consider that we want to re-identify a module i and that the initial model $G_{i,init}(z) = G(\theta_{i,init}, z)$ for $G_{i,0}(z)$, used for designing the controller $K_i(z)$, is already available. Let us suppose that the new model is needed to design a new controller $\hat{K}_i(z)$, in order to restore the performance of the system. So we will excite the module with the signal $r_i(t)$, collect N measurements of the input-output data in

$Z_i^N = \{u_i(t), y_i(t) \quad t = 1..N\}$ and identify a new parameter vector $\hat{\theta}_i$ in a full order model structure. Then the uncertainty region U_i of $\hat{\theta}_i$ must respect an application dependent accuracy constraint $U_i \subseteq U_{app}$

Unfortunately, this procedure will affect the rest of the network, since every other system $j \neq i$ will be affected by the excitation signal $r_i(t)$ according to (20). Then, we decide to modify the module as in Figure 6, taking inspiration from the Stealth Identification proposed in [16]. Then, the module i will send to the network the signal $x_i(t) = y_i(t) - G_{i,init}(z)r_i(t)$ instead of $y_i(t)$. Note that in [16] it is assumed that $G_{i,init}(z) \approx G_{i,0}(z)$, thus making $\mathcal{P}_{x_j r_i}$, the power of $x_i(t)$ induced by $r_i(t)$, negligible. In our case it would be particularly dangerous, since we would not be able to control the effects of $x_i(t)$ and it could spoil the network behaviour in an unforeseeable way. Here we do not make this approximation, since we want to be sure that the normal operation will be preserved. So we must design $r_i(t)$ such that $\mathcal{P}_{x_j r_i}$ is small. However, this is still not enough to preserve the normal operation, because the actual output of every module will still be affected by the excitation signal. So we must take into account also $\mathcal{P}_{y_j r_i}$, the power of the output of the module j, $y_j(t)$, induced by the identification experiment, in order to assure the normal operation of the network:

$$\mathcal{P}_{y_j r_i} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |R_{j,i}(z) G_{j,0}(e^{j\omega})|^2 \Phi_{r_i}(\omega) d\omega \tag{23}$$

Another critical aspect is the total power of the excitation signal $\mathcal{P}_{r_i}$. In fact, the type of network here considered is typically used with autonomous system, so the energy used for the identification will come directly from the batteries of the module, then reducing its autonomy. For this reason we will consider $\mathcal{P}_{r_i}$ as the cost of the experiment. Now, we can present the problem considered in this thesis:

Problem 3. Let us consider the interconnection of N_m locally controlled modules $\mathcal{S}_i$, as defined in Section 2.1. Let $G_{i,init}(z)$ be an initial parametric model for the module i. Find the power spectrum of the excitation signal Φ_{r_i} s.t. the identification experiment performed as in Figure 6 on the module i while the network is operating, will lead to an identified model which satisfies a given application dependent accuracy requirement. Moreover, the experiment must perturb as least as possible the network, or, at least, have a limited effect.

Thank to the previous considerations, we can face this problem by solving the following minimization problem:

$$\min_{\Phi_{r_i}(\omega)} \mathcal{P}_{r_i} \tag{24}$$

$$\mathcal{P}_{y_j r_i}(\Phi_{r_i}(\omega)) < \gamma_{y,j} \quad \forall j = 1 \dots N_m \tag{25}$$

$$\mathcal{P}_{x_i r_i}(\Phi_{r_i}(\omega)) < \gamma_x \tag{26}$$

$$U_i \subseteq U_{app} \tag{27}$$

where the constants γ_x and $\gamma_{y,j} \quad \forall j = 1 \dots N_m$ are chosen in order to keep the perturbation acceptable and $\gamma_x \ll \gamma_{y,i}$. Note that the main philosophy is the same as [3], where the identification experiment must affect as least as possible the normal operation of the plant. The problem here addressed is actually similar, but presents the additional constraint (26), needed to reduce the effect of the experiment on the rest of the network. Moreover, this constraint represents a novelty w.r.t [16], since we are designing $r_i(t)$ in order to keep $\mathcal{P}_{x_i r_i}$ small, instead of assuming it to be negligible.

2.2.1 A simple case

In order to face this problem, it is convenient to start with a simple case. Let us consider a simple case of only one module, where the true system $G_0(z)$ is the module 1 of the network in Figure 4 and it is operated in open loop, as in Fig 7. Let us consider that an initial model $G_{init}(z) = G(\theta_{init}, z)$ for $G_0(z)$ and $H_{init}(z) = H(\theta_{init}, z)$ for $H_0(z)$, is available and was obtained under the same assumptions of Section 2.2. The linear transfer functions $G(z, \theta)$ and $H(z, \theta)$ have the following form:

$$G(z, \theta) = \frac{Z_1 \theta}{1 + Z_2 \theta} \quad H(z, \theta) = \frac{1 + Z_3 \theta}{1 + Z_4 \theta} \tag{28}$$

where Z_1, Z_2, Z_3 and Z_4 are appropriate vectors of zeros and delays.

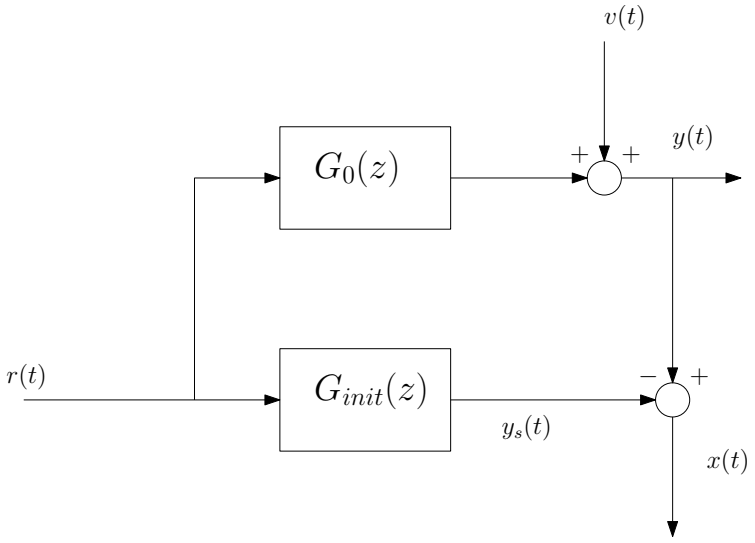


Figure 7: Schematic of the simple case considered.

Let $P_{\theta_{init}}$ be the covariance matrix of θ_{init} and U_{init} the uncertainty region that contains the true parameter vector θ_0 with a given probability level β:

$$U_{init} := \left\{ \theta | (\theta - \theta_{init})^T P_{\theta_{init}}^{-1} (\theta - \theta_{init}) < \chi_\beta \quad \theta \in \mathbb{R}^k \right\} \tag{29}$$

where χ_β is defined as $Pr(\chi^2(k) < \chi_\beta) = \beta$. Let us suppose that it is necessary to perform another identification experiment, while the system is operating. The system is then operated as in Figure 7, where $G_0(z)$ is controlled in open loop with the signal $r(t)$, the output $y(t)$ is measured for identification purpose and the signal sent to the network $x(t)$ is here given by:

$$x(t) = (G_0(z) - G_{init}(z)) r(t) + v(t) \tag{30}$$

The power of the signals $y(t)$ and $x(t)$, induced by the excitation signal $r(t)$, and the power of the excitation signal are here given by:

$$\mathcal{P}_{yr} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |G_0(e^{j\omega})|^2 \Phi_r(\omega) d\omega \tag{31}$$

$$\mathcal{P}_{xr} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |G_0(e^{j\omega}) - G_{init}(e^{j\omega})|^2 \Phi_r(\omega) d\omega \tag{32}$$

$$\mathcal{P}_{r} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_r(\omega) d\omega \tag{33}$$

where $\Phi_r(\omega)$ is the power spectrum of the excitation signal. The quantities $\mathcal{P}_{yr}$ and $\mathcal{P}_{xr}$ can be bounded by scalars γ_y and γ_x, respectively, in order to keep the effect of $r(t)$ bounded. The new experiment must have the lowest cost, here considered to be given by $\mathcal{P}_r$. The to-be-identified model will have covariance matrix P_θ, defined as in (8):

$$P_\theta^{-1} = \frac{N}{\sigma_e^2 2\pi}\int_{-\pi}^{\pi} F_r(e^{j\omega},\theta_0) F_r^*(e^{j\omega},\theta_0) \Phi_u(\omega) d\omega + \frac{N}{2\pi}\int_{-\pi}^{\pi} F_e(e^{j\omega},\theta_0) F_e^*(e^{j\omega},\theta_0) d\omega \tag{34}$$

where $F_r(z,\theta) = H(z,\theta)^{-1}\frac{dG(z,\theta)}{d\theta}$, $F_e(z,\theta) = H(z,\theta)^{-1}\frac{dH(z,\theta)}{d\theta}$. Then we can define the uncertainty region U as in (7) and express an application dependent accuracy requirement $U \subseteq U_{app}$. Finally, the experiment duration N is considered fixed. Then the experiment design problem considered in this thesis can be written, for this simple case, as the following minimization problem:

$$\min_{\Phi_r(\omega)} \mathcal{P}_r \tag{35}$$

$$\mathcal{P}_{yr}(\Phi_r(\omega)) < \gamma_y \tag{36}$$

$$\mathcal{P}_{xr}(\Phi_r(\omega)) < \gamma_x \tag{37}$$

$$U \subseteq U_{app} \tag{38}$$

This problem, as it is, is still not tractable since it suffers from the following issues:

- The power spectrum is a continuous function, making the problem of infinite dimension.
- We have not done any assumption on the accuracy constraint $U \subseteq U_{app}$.
- The quantities $\mathcal{P}_{xr}$, $\mathcal{P}_{yr}$ and (34) depend on the unknown true vector. How it is possible to deal with this dependency?

Parametrization of the power spectrum The first point is easily solved in the literature with the choice of the parametrization for $\Phi_r(\omega)$. It is desired that the expression of $\Phi_r(\omega)$ is linear in the parameters and it leads to a convex optimization problem. To this aim, in literature are considered two main type of signals: multisine or filtered white noise. The power spectrum of a signal is a positive function, so it may be necessary to add a constraint in order to guarantee that this property is satisfied by the parametrization. Here we decided to parametrize $\Phi_r(\omega)$ as the power spectrum of a multisine signal.

A multisine signal is given by a summation of a finite number n_ω of sine waves. There are three parameters for each harmonic i: the amplitude, A_i, the frequency, ω_i, and the phase ϕ_i. The signal and its power spectrum are given by:

$$r(t) = \sum_{i=1}^{n_\omega} A_i \sin(\omega_i t + \phi_i) \quad \Phi_r(\omega) = \sum_{i=1}^{n_\omega} \frac{A_i^2 \pi}{2} \left(\delta(\omega - \omega_i) + \delta(\omega + \omega_i)\right) \tag{39}$$

where $\delta(\omega)$ is the delta Dirac function. Note that the phase ϕ_i does not appear in the power spectrum, so it may be chosen in order to minimize the maximum amplitude of the signal or for creating a random phase multisine signal. The power spectrum is now a finitely parametrized function, giving the possibility to simplify the computation of P_θ^{-1}:

$$P_\theta^{-1} = \frac{N}{\sigma_e^2} \sum_{i=1}^{n_\omega} Re\left(F_r(e^{j\omega_i}, \theta_0) F_r^*(e^{j\omega_i}, \theta_0)\right) \frac{A_i^2}{2} + \frac{N}{2\pi} \int_{-\pi}^{\pi} F_e(e^{j\omega}, \theta_0) F_e^*(e^{j\omega}, \theta_0) d\omega \tag{40}$$

With this parametrization the total power of the signal simply becomes $\mathcal{P}_r = \sum_{i=1}^{n_\omega} \frac{A_i^2}{2}$, while the positivity of $\Phi_r(\omega)$ can be assured by simply imposing $A_i^2 > 0 \quad \forall i$. However, there is the challenge of choosing the decision variables. Indeed, using both ω_i and A_i^2 would make the problem non-linear. For this reason, only the amplitudes squared A_i^2 are taken as decision variables, while the ω_i are taken as grid points of the frequency axis. So P_θ^{-1} becomes a linear combination of the decision variables A_i^2. Nevertheless, this parametrization usually requires a fine frequency grid, since it is not available an a-priori knowledge about which are the most important ones.

Accuracy Constraint In recent literature are presented many approaches for expressing an accuracy constraint that represents an application dependent requirement. Many of them concern the design of a controller that is able to reach a certain performance level for every system inside the uncertainty region [5],[3],[2],[4]. Here we consider the approach in [5] based on a bound $r_{adm}(\omega)$ for the modelling error $|G_0(e^{j\omega}) - G(e^{j\omega}, \hat{\theta}_N)|$. The dependency on the covariance matrix is obtained by approximating the modelling error with a Taylor expansion truncated at the first order, leading to the following constraint:

$$\chi_\beta \bar{\lambda}\left(T(e^{j\omega}, \theta_{init}) P_\theta^{-1} T(e^{j\omega}, \theta_{init})^T\right) \leq r_{adm}^2(\omega) \quad \forall \omega \tag{41}$$

Where $T(z,\theta) = (Re(\Lambda_G^T(z,\theta)) \quad Im(\Lambda_G^T(z,\theta)))^T$, $\Lambda_G(z,\theta) = \frac{dG(z,\theta)}{d\theta}$. Moreover, since P_θ^{-1} depends on the true parameter vector, here it is computed by substituting θ_0 for θ_{init}. This constraint needs to be evaluated at every frequency ω, which is impossible in practice. For this reason it is usually verified over a fine frequency grid Ω_r of n_{ω_r} points. Moreover, we want this constraint to be an affine function of the decision variables A_i^2. Using the Schur's complement [7] we can rewrite the constraint as follows:

$$P_\theta^{-1} > R_{adm}(\omega) \quad \omega \in \Omega_r \tag{42}$$

where $R_{adm}(\omega) = \frac{\chi_\beta}{r_{adm}^2(\omega)} T^T(e^{j\omega}, \theta_{init}) T(e^{j\omega}, \theta_{init})$.

A tractable formulation for $\mathcal{P}_{yr}$ and $\mathcal{P}_{xr}$ The quantities $\mathcal{P}_{yr}$ and $\mathcal{P}_{xr}$ are dependent on the unknown true system $G_0(z)$. Substituting $G_{init}(z)$ for $G_0(z)$ in (32) would make $\mathcal{P}_{xr}$ identically equal to 0. This approach is the same as assuming $G_{i,init}(z) \approx G_{i,0}(z)$ and does not allow us to keep the power $x(t)$ small. Moreover, we want to be sure that effect of $r(t)$ on $y(t)$ will remain small, in order to avoid unpredicted behaviour during critical operation e.g. dangerous enviroment [1]. For these reasons, it is necessary to use a robust formulation of these constraints. Thanks to the fact that the model $G_{init}(z)$ has been obtained under the full-order model structure, we known that there exists

the true parameter vector θ_0 s.t. $G_0(z) = G(z, \theta_0)$ and that $\theta_0 \in U_{init}$. So it is possible to define a worst case version of (31) and (32):

$$\mathcal{P}_{yr}^{wc} := \max_{\theta \in U_{init}} \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{j\omega}, \theta)|^2 \Phi_r(\omega) d\omega \tag{43}$$

$$\mathcal{P}_{xr}^{wc} := \max_{\theta \in U_{init}} \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(e^{j\omega}, \theta) - G_{init}(e^{j\omega})|^2 \Phi_r(\omega) d\omega \tag{44}$$

Then, if the minimization problem is solved using (43) and (44), then it is assured that $\mathcal{P}_{yr}$ and $\mathcal{P}_{xr}$ will remain below the given bounds. So, the minimization problem can be rewritten as follows:

$$\min_{A_i^2} \sum_{i=1}^{n_\omega} \frac{A_i^2}{2} \tag{45}$$

$$\max_{\theta \in U_{init}} \sum_{i=1}^{n_\omega} \frac{A_i^2}{2} |G(e^{j\omega_i}, \theta)|^2 < \gamma_y \tag{46}$$

$$\max_{\theta \in U_{init}} \sum_{i=1}^{n_\omega} \frac{A_i^2}{2} |G(e^{j\omega_i}, \theta) - G_{init}(e^{j\omega_i})|^2 < \gamma_x \tag{47}$$

$$P_\theta^{-1} > R_{adm}(\omega) \quad \forall \omega \in \Omega_r \tag{48}$$

However, it still remains the problem of verifying whether (46) and (47) hold. Some possible solutions come from the literature in robust experiment design. The approach proposed in [3], equivalent to the one used in [17], consists in a discretization of U_{init} in a finite number of N_l points $\theta_l \in U_{init}$. Then (46) and (47) are approximated by the following set of constraints:

$$\sum_{i=1}^{n_\omega} \frac{A_i^2}{2} |G(\theta_l, e^{j\omega_i})|^2 < \gamma_y \quad \forall l = 1 \dots N_l \tag{49}$$

$$\sum_{i=1}^{n_\omega} \frac{A_i^2}{2} |G(\theta_l, e^{j\omega_i}) - G_{init}(e^{j\omega_i})|^2 < \gamma_x \quad \forall l = 1 \dots N_l \tag{50}$$

Another possibility comes from [14] and [10], where the approaches proposed rely on a first order or second order Taylor approximation of the constraints. However, the approaches just cited rely on an approximation of (46) and (47). In the problem here considered we are focused in assuring that these constraints will be satisfied. This concern comes from the fact that in the original problem we have to perform the identification experiment while the network is operating, maybe also in a dangerous context [1].

To this aim, in the sequel we will propose two approaches for assuring that (46) and (47) will be satisfied. These two approaches, named α-approach and SG-approach, represent the first contribution of this thesis. In the following they will be presented only for (47), since their application to (46) directly follows by simply imposing $G_{init}(z) = 0$ in (47).

2.2.2 α approach

Substituting an upper bound for (44) in (47) assures that $\mathcal{P}_{xr}^{wc}$ will respect the constraint during the experiment. Then, it is always possible to bound (44) using the following relation:

$$\max_{\theta\in U_{init}} \sum_{i=1}^{n_\omega} \frac{A_i^2}{2}|G(e^{j\omega_i},\theta) - G_{init}(e^{j\omega_i})|^2 \leq \sum_{i=1}^{n_\omega} \frac{A_i^2}{2} \max_{\theta\in U_{init}} |G(e^{j\omega_i},\theta) - G_{init}(e^{j\omega_i})|^2 \tag{51}$$

Note that now we do not have the problem of computing (44) in in (47) reduces to the computation of the following terms:

$$\alpha_x(\omega) := \max_{\theta\in U_{init}} |G(e^{j\omega},\theta) - G_{init}(e^{j\omega})|^2 \tag{52}$$

The frequencies ω_i are given, so it is possible to compute the terms $\alpha_x(\omega_i)$ before solving the main experiment design problem. The value of $\alpha_x(\omega)$ at a given frequency ω is obtained as $\alpha_x(\omega) = \alpha_{opt}$, where α_{opt} is the optimal value of α in the following optimization problem:

$$\min_{\alpha} \alpha \text{ subject to: } |G(e^{j\omega},\theta) - G_{init}(e^{j\omega})|^2 < \alpha \quad \forall\theta \in U_{init} \tag{53}$$

This problem can be easily solved by means of a convex LMI-based optimization problem [6]. Note that the upper bound (51) is an affine function of the decision variable. Then, the experiment design problem can be rewritten as the following LMI-based convex optimization problem:

$$\min_{A_i^2} \sum_{i=1}^{n_\omega} \frac{A_i^2}{2} \tag{54}$$

$$\sum_{i=1}^{n_\omega} \frac{A_i^2}{2}\alpha_x(\omega_i) < \gamma_x \tag{55}$$

$$P_\theta^{-1} > R_{adm}(\omega) \quad \forall\omega \in \Omega_r \tag{56}$$

where the satisfaction of (55) implies that (47) holds.

2.2.3 SG approach

The second approach proposed aims at finding a sufficient condition for (47) to hold, using the separation of graph theorem. Let's recall that $\mathcal{P}_{xr}^{wc} < \gamma_x$ is equivalent to:

$$\sum_{i=1}^{n_\omega} \frac{A_i^2}{2}|G(e^{j\omega_i},\theta) - G_{init}(e^{j\omega_i})|^2 < \gamma_x \quad \forall\theta \in U_{init} \tag{57}$$

Since $G(z,\theta)$ is a fractional function of θ, it is then possible to rewrite every term $h(e^{j\omega_i}) = G(e^{j\omega_i},\theta) - G_{init}(e^{j\omega_i})$ as an LFT in θ of the following system, represented in Figure 8:

$$\begin{pmatrix} p_i \\ h(e^{j\omega_i}) \end{pmatrix} = \begin{pmatrix} -Z_2(e^{j\omega_i}) & 1 \\ Z_1(e^{j\omega_i}) & -G_{init}(e^{j\omega_i}) \end{pmatrix} \begin{pmatrix} q_i \\ 1 \end{pmatrix} \quad q_i = \theta p_i \tag{58}$$

Then, it is possible to aggregate all the terms in only one system M, given by:

$$\begin{pmatrix} \bar{p} \\ \bar{h} \end{pmatrix} = M \begin{pmatrix} \bar{q} \\ 1 \end{pmatrix} = \begin{pmatrix} -\bar{Z}_2 & \bar{1} \\ \bar{Z}_1 & -\bar{G}_{init} \end{pmatrix} \begin{pmatrix} \bar{q} \\ 1 \end{pmatrix} \quad \bar{q} = \Theta\bar{p} \tag{59}$$

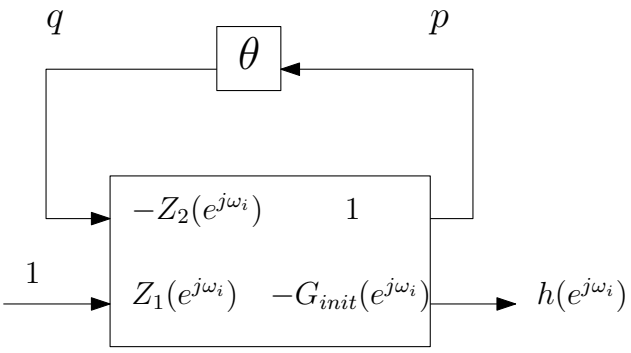


Figure 8: LFT representation of $h(e^{j\omega_i})$

where:

$$\Theta = I_{n_\omega} \otimes \theta \quad \bar{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{n_\omega} \end{pmatrix} \quad \bar{q} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_{n_\omega} \end{pmatrix} \quad \bar{G}_{init} = \begin{pmatrix} G_{init}(e^{j\omega_1}) \\ G_{init}(e^{j\omega_2}) \\ \vdots \\ G_{init}(e^{j\omega_{n_\omega}}) \end{pmatrix} \quad \bar{h} = \begin{pmatrix} h(e^{j\omega_i}) \\ h(e^{j\omega_i}) \\ \vdots \\ h(e^{j\omega_{n_\omega}}) \end{pmatrix} \tag{60}$$

$$\bar{Z}_1 = \begin{pmatrix} Z_1(e^{j\omega_1}) & 0 & \dots & 0 \\ 0 & Z_1(e^{j\omega_2}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Z_1(e^{j\omega_{n_\omega}}) \end{pmatrix} \quad \bar{Z}_2 = \begin{pmatrix} Z_2(e^{j\omega_1}) & 0 & \dots & 0 \\ 0 & Z_2(e^{j\omega_2}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Z_2(e^{j\omega_{n_\omega}}) \end{pmatrix} \tag{61}$$

where $\bar{1} = (1 \dots 1)^T \in \mathbb{R}^{kn_\omega}$ and $\Theta \in \Delta = \{\Theta | \Theta$ is as in (60) with $\theta \in U_{init}\}$. Then, the output $\bar{h}$ is given by:

$$\bar{h} = \bar{Z}_1 \Theta (I + \bar{Z}_2 \Theta)^{-1} \bar{1} - \bar{G}_{init} \tag{62}$$

Let us define the set of multiplier for the uncertainty set Δ as an affine parametrization of matrices $\Pi = \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12}^* & \Pi_{22} \end{pmatrix}$ having the following property:

$$\begin{pmatrix} I_{n_\omega} \\ \Theta \end{pmatrix}^* \Pi \begin{pmatrix} I_{n_\omega} \\ \Theta \end{pmatrix} > 0 \quad \forall \Theta \in \Delta \tag{63}$$

Such a parametrization can be easily found following the reasoning in [2]. It is then possible to verify the constraint $\mathcal{P}_{xr}^{wc} < \gamma_x$ using the following result:

Proposition 2.1. *Let us consider a set of multiplier Π for the uncertainty set Δ and the system M defined as in* (59)*. The existence of Π such that the following matrix inequality holds:*

$$\begin{pmatrix} M \\ I_{kn_\omega} \end{pmatrix}^* \left(\begin{array}{cc|cc} \Pi_{11} & 0 & \Pi_{12} & 0 \\ 0 & \mathcal{A} & 0 & 0 \\ \hline \Pi_{12}^* & 0 & \Pi_{22} & 0 \\ 0 & 0 & 0 & -2\gamma_x \end{array}\right) \begin{pmatrix} M \\ I_{kn_\omega} \end{pmatrix} < 0 \tag{64}$$

with $\mathcal{A} = diag(A_1^2, A_2^2, \dots A_{n_\omega}^2)$, is a sufficient condition for:

$$\sum_{i=1}^{n_\omega} \frac{A_i^2}{2} |h(e^{j\omega_i})|^2 < \gamma_x \quad \forall \theta \in U_{init} \tag{65}$$

with $h(e^{j\omega_i})$ *defined as* $h(e^{j\omega_i}) = G(e^{j\omega_i}, \theta) - G_{init}(e^{j\omega_i})$.

Proof. Let us first note that if exist Π such that (64) holds, this implies for the same Π:

$$\begin{pmatrix} -\bar{Z}_2 \\ I_{n_\omega} \end{pmatrix}^* \Pi \begin{pmatrix} -\bar{Z}_2 \\ I_{n_\omega} \end{pmatrix} < 0 \tag{66}$$

Using the separation of graph theorem, the existence of Π for the the uncertainty set Δ such that (66) holds, is a sufficient condition for $det(I + \bar{Z}_2\Theta) \neq 0$ to hold $\forall \Theta \in \Delta$, thus assuring the well-posedeness of (59) for every $\Theta \in \Delta$. This implies that the signals $\bar{q}$, $\bar{p}$ and $\bar{h}$ are well-defined.

Let us now pre and post multiply (64) by the input of (59) i.e. $\begin{pmatrix} \bar{q} \\ 1 \end{pmatrix}^*$ and $\begin{pmatrix} \bar{q} \\ 1 \end{pmatrix}$:

$$\begin{pmatrix} \bar{q} \\ 1 \end{pmatrix}^* \begin{pmatrix} M \\ I_{kn_\omega} \end{pmatrix}^* \left(\begin{array}{cc|cc} \Pi_{11} & 0 & \Pi_{12} & 0 \\ 0 & \mathcal{A} & 0 & 0 \\ \hline \Pi_{12}^* & 0 & \Pi_{22} & 0 \\ 0 & 0 & 0 & -2\gamma_x \end{array}\right) \begin{pmatrix} M \\ I_{kn_\omega} \end{pmatrix} \begin{pmatrix} \bar{q} \\ 1 \end{pmatrix} < 0 \tag{67}$$

we can expand this inequality using (59):

$$\begin{pmatrix} \bar{p} \\ \bar{q} \end{pmatrix}^* \Pi \begin{pmatrix} \bar{p} \\ \bar{q} \end{pmatrix} + \begin{pmatrix} \bar{h} \\ 1 \end{pmatrix}^* \begin{pmatrix} \mathcal{A} & 0 \\ 0 & -2\gamma_x \end{pmatrix} \begin{pmatrix} \bar{h} \\ 1 \end{pmatrix} < 0 \tag{68}$$

The first term on the left hand side is (63) pre and post multiplied by $\bar{p}^*$ and $\bar{p}$. Then it is greater than zero $\forall \Theta \in \Delta$, due to the definition of Π. Consequently, we have:

$$\begin{pmatrix} \bar{h} \\ 1 \end{pmatrix}^* \begin{pmatrix} \mathcal{A} & 0 \\ 0 & -2\gamma_x \end{pmatrix} \begin{pmatrix} \bar{h} \\ 1 \end{pmatrix} < 0 \quad \forall \Theta \in \Delta \tag{69}$$

which is equivalent to:

$$\sum_{i=1}^{n_\omega} \frac{A_i^2}{2} |h(e^{j\omega_i})|^2 < \gamma_x \quad \forall \theta \in U_{init} \tag{70}$$

thus concluding the proof. □

Then, the experiment design problem can be rewritten as the following LMI-based convex optimization problem:

$$\min_{A_i^2, \Pi} \sum_{i=1}^{n_\omega} \frac{A_i^2}{2} \tag{71}$$

$$\begin{pmatrix} M \\ I_{kn_\omega} \end{pmatrix}^* \left(\begin{array}{cc|cc} \Pi_{11} & 0 & \Pi_{12} & 0 \\ 0 & \mathcal{A} & 0 & 0 \\ \hline \Pi_{12}^* & 0 & \Pi_{22} & 0 \\ 0 & 0 & 0 & -2\gamma_x \end{array}\right) \begin{pmatrix} M \\ I_{kn_\omega} \end{pmatrix} < 0 \tag{72}$$

$$P_\theta^{-1} > R_{adm}(\omega) \quad \forall \omega \in \Omega_r \tag{73}$$

Note that (64) is just a sufficient condition for (47). This implies that the constraint (64) is conservative with respect to the original one (47).

2.2.4 Simulative Results

In order to compare the α and SG approach we made some simulative tests. We created the following true system $\mathcal{S}$ in Matlab:

$$G_0(z) = \frac{0.10276z^{-3} + 0.18123z^{-4}}{1 - 1.99185z^{-1} + 2.20265z^{-2} - 1.84083z^{-3} + 0.89413z^{-4}} \tag{74}$$

$$H_0(z) = \frac{1}{1 - 1.99185z^{-1} + 2.20265z^{-2} - 1.84083z^{-3} + 0.89413z^{-4}} \tag{75}$$

Then an initial model G_{init} has been obtained by simulating $\mathcal{S}$ excited by a white noise of unit variance, then collecting 1000 input-output data points and identifying G_{init} with a PEM in a full-order model structure. The two approaches have been implemented in Matlab, using LMItoolbox for solving the LMI-based optimization problem.

	$\mathcal{P}_r$	$\mathcal{P}_{xr}$
without constraint on $\mathcal{P}_{xr}$	1.1861	1.5
with constraint on $\mathcal{P}_{xr}$	2.6541	0.015

Table 1: Comparison of the results with the constraint active or not. In the test with constraint active, γ_x was set equal to 0.015

A first test involved the solution of the optimization problem, obtained from the α-approach, performed over a fine frequency grid. In Figure 9 it is possible to see how changes the power spectrum obtained when the constraint on $\mathcal{P}_{xr}^{wc}$ is active or not. The total power of the experiments designed in both cases are in reported in Tab 1. Note that with the constraint active the total power of the excitation signal doubles, while the $\mathcal{P}_{xr}^{wc}$ is reduced by a factor 100.

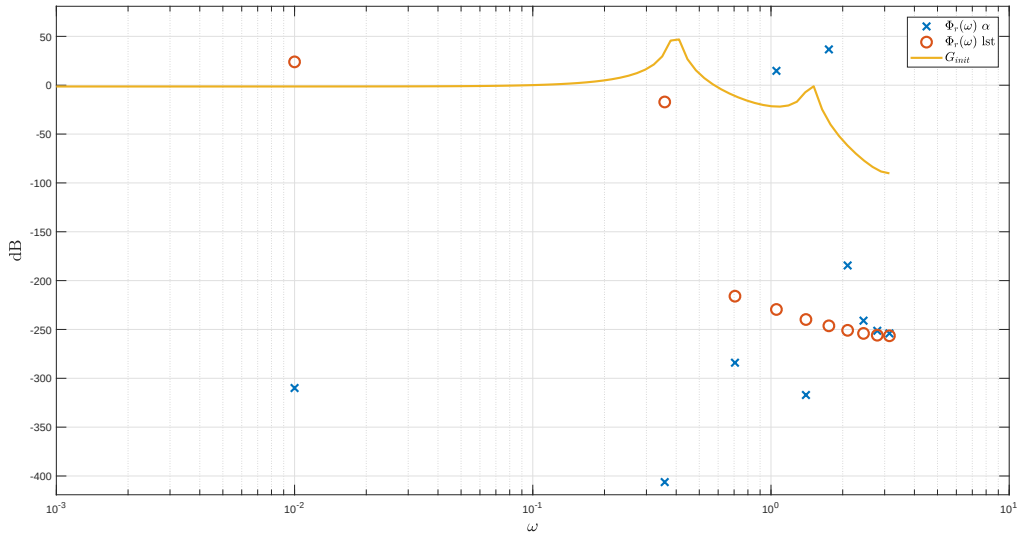


Figure 9: Comparison of the results with the constraint active or not. G_{init} is the module of the frequency resposne of the initial model G_{init}, $\Phi_r(\omega)$ α is the power spectrum obtained with the constraint on $\mathcal{P}_{xr}^{wc}$ active, $\Phi_r(\omega)$ lst is the power spectrum obtained without the constraint on $\mathcal{P}_{xr}^{wc}$

Then, from the results obtained when the constraint was active, we selected the two most important frequencies: [1.0539 1.7498]. Afterwards, using these two frequencies, the experiment design problems, obtained with the α and the SG approach, were solved for different values of γ_x. In Figure 10 we can observe how $\mathcal{P}_r$ varies with different values of γ_x. It can be observed that the difference between the two approaches is more significant for lower values of γ_x, while seems to be almost negligible for relatively high values.

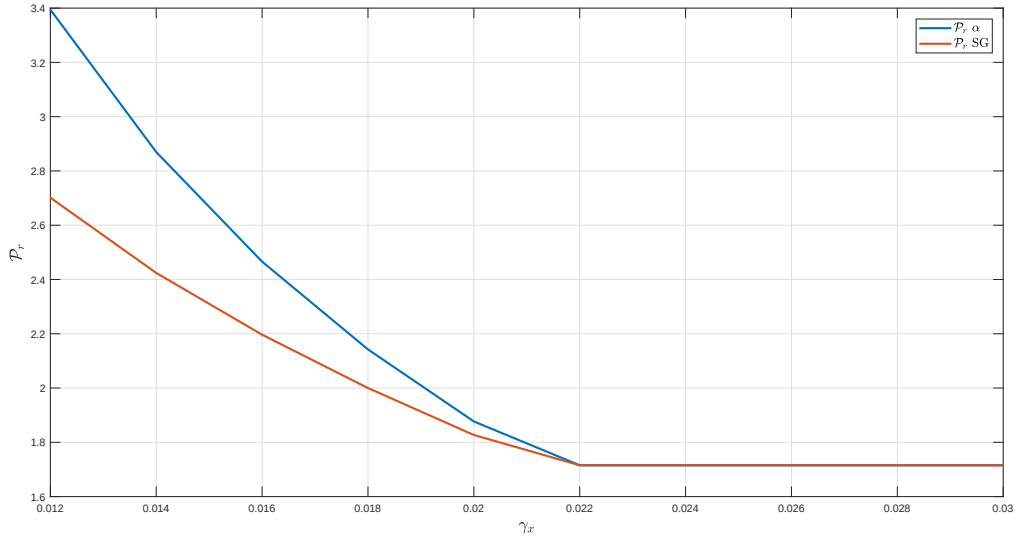


Figure 10: Comparing $\mathcal{P}_r$ obtained with different values of γ_x with both approaches. $\mathcal{P}_r$ α is the power obtained using the α approach, while $\mathcal{P}_r$ SG is the power obtained using the SG approach.

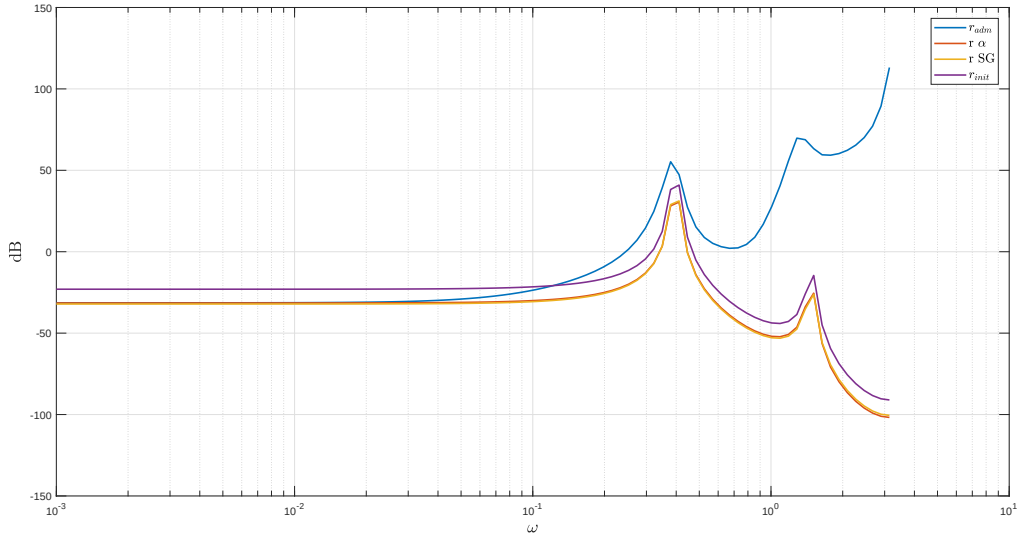


Figure 11: Modelling error obtained from the new identification experiment. r_{adm} is the accuracy requirement on the modelling error as in (41), r_{init} the initial modelling error, r α the modelling error coming from the experiment designed with the α approach, r SG the modelling error coming from the experiment designed with the SG approach

Finally, the new identification experiment was performed using the power spectra designed with $\gamma_x = 0.012$ from both approaches. In Figure 11 we can see the obtained modelling error from the experiment designed with both approaches. Note that both lead to a model which satisfies the accuracy requirement, but from Figure 10 we see that the power spectrum obtained thanks to the SG approach requires less power.

Conclusion and Roadmap

2.3 Roadmap

During this first year, the main focus was to acquire the necessary knowledge about System Identification, in particular the Prediction Error Method, experiment design and the application of robust control theory in experiment design. Particular attention was posed on some basic concepts of this topic, that were unrelated to my initial background. In the remainder of this thesis two paths will be followed: the former regards my formation, aimed at improving my knowledge in robust control and system identification, the latter regards the research activity according to the following roadmap:

- The results will be extended to the complete interconnected situation;
- The methodology here proposed relies on the fact that we have access to the measured output $y(t)$ and we can modify its value before sending this information. However, this may not be possible in practice. Another step will be to consider this case. A possible solution could be to act directly on the output of the system with an actuator;
- Here we limited our attention to the interconnection of locally controlled systems. It would be interesting to extend the formulation also to the general dynamical network configuration [8].
- Finally, another interesting step could be the robustification of the accuracy constraint. Indeed, the computation of the covariance matrix is usually performed by replacing the true parameter vector θ_0 with its estimate. Up to our knowledge, the only attempt to robustify it is in [17], where actually a scalar measure of the Fisher information matrix is robustified and then its computation is performed using an approximation.

It must be reported that this thesis started within the scope of the NEXT4MEMS project, financed by BPI France. However, during this year the activities in the project did not require a strong involvement, so we started this new activity. During the next year I will be working, on a parallel path, on the recursive identification of the time-varying resonance frequency of the MEMS gyroscope considered in the project.

2.4 Conclusion

This report presented the activity of this first year. The study of the main subjects related to the problem allowed to obtain a better understanding of it. Moreover, there were presented the first results of this research activity, that represent the first contribution to the research in this field. Indeed, even if the problem considered until now is much more simpler than the main one, the two approaches here proposed are actually a novelty in the robust experiment design literature, since they allow to

effectively guarantee that the constraints will be satisfied for every system inside the uncertainty region.

These results represent the first important step towards the main goal of this thesis: the optimal costly experiment design of one module in the interconnection of locally controlled system.

References

[1] Marco Andreetto et al. "A cooperative monitoring technique using visually servoed drones". In: *Environmental, Energy and Structural Monitoring Systems (EESMS), 2015 IEEE Workshop on*. IEEE. 2015, pp. 244–249.

[2] Märta Barenthin et al. "Identification for control of multivariable systems: Controller validation and experiment design via LMIs". In: *Automatica* 44.12 (2008), pp. 3070–3078.

[3] Xavier Bombois et al. "Least costly identification experiment for control". In: *Automatica* 42.10 (2006), pp. 1651–1662.

[4] Xavier Bombois et al. "Optimal identification experiment design for the interconnection of locally controlled systems". In: *Automatica* 89 (2018), pp. 169–179.

[5] X Bombois et al. "Cheapest open-loop identification for control". In: *Decision and Control, 2004. CDC. 43rd IEEE Conference on*. Vol. 1. IEEE. 2004, pp. 382–387.

[6] X Bombois et al. "Least costly identification experiment for control. A solution based on a high-order model approximation". In: *American Control Conference, 2004. Proceedings of the 2004*. Vol. 3. IEEE. 2004, pp. 2818–2823.

[7] Stephen Boyd et al. *Linear matrix inequalities in system and control theory*. Vol. 15. Siam, 1994.

[8] Arne Dankers et al. "Identification of dynamic models in complex networks with prediction error methods: Predictor input selection". In: *IEEE Transactions on Automatic Control* 61.4 (2016), pp. 937–952.

[9] J Alexander Fax and Richard M Murray. "Information flow and cooperative control of vehicle formations". In: *IEEE transactions on automatic control* 49.9 (2004), pp. 1465–1476.

[10] Marco Forgione, Xavier Bombois, and Paul MJ Van den Hof. "Data-driven model improvement for model-based control". In: *Automatica* 52 (2015), pp. 118–124.

[11] Michel Gevers and Alexandre S Bazanella. "Identification in dynamic networks: identifiability and experiment design issues". In: *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on*. IEEE. 2015, pp. 4005–4010.

[12] Per Hägg and Bo Wahlberg. "On optimal input design for networked systems". In: *Automatica* 53 (2015), pp. 275–281.

[13] Håkan Hjalmarsson. "System identification of complex and structured systems". In: *Control Conference (ECC), 2009 European*. IEEE. 2009, pp. 3424–3452.

[14] Abhishankar Kumar and Sridharakumar Narasimhan. "Robust Plant Friendly Optimal Input Design". In: *IFAC Proceedings Volumes* 46.32 (2013), pp. 553–558.

[15] L Ljung. *System Identification-Theory for the User 2nd edition PTR Prentice-Hall*. Upper Saddle River, USA, 1999.

[16] Marcus Gerardus Potters, PMJ van den Hof, and XJA Bombois. "Experiment design for identification of structured linear systems". PhD thesis. Delft University of Technology, 2016.

[17] Cristian R Rojas et al. “Robust optimal experiment design for system identification”. In: *Automatica* 43.6 (2007), pp. 993–1008.

[18] Paul MJ Van den Hof et al. “Identification of dynamic models in complex networks with prediction error methods—Basic methods for consistent module estimates”. In: *Automatica* 49.10 (2013), pp. 2994–3006.

Université de Lyon
CNRS, Ecole Centrale Lyon, INSA Lyon, Université Claude Bernard Lyon 1

Laboratoire Ampère
Unité Mixte de Recherche du CNRS - UMR 5005
Génie Electrique, Automatique, Bio-ingénierie

Mémoire doctorant 1ère année
2017 -2018

Nom - Prénom	PEÑA, Robert Alfie S.
email	robert.pena@univ-lyon1.fr
Titre de la thèse	Amélioration de la durée de vie des systèmes de stockage d'énergie électrique par optimisation des convertisseurs d'énergie modulaires associés
Directeur de thèse	Pascal VENET
Co- encadrants	Alaa HIJAZI
Dpt. de rattachement	Méthodes pour l'Ingénierie des Systèmes (MIS)
Date début des travaux	4 septembre 2017
Type de financement	Financement étranger

Amélioration de la durée de vie des systèmes de stockage d'énergie électrique par optimisation des convertisseurs d'énergie modulaires associés

Robert Alfie S. Peña*

2 July 2018

1 Introduction

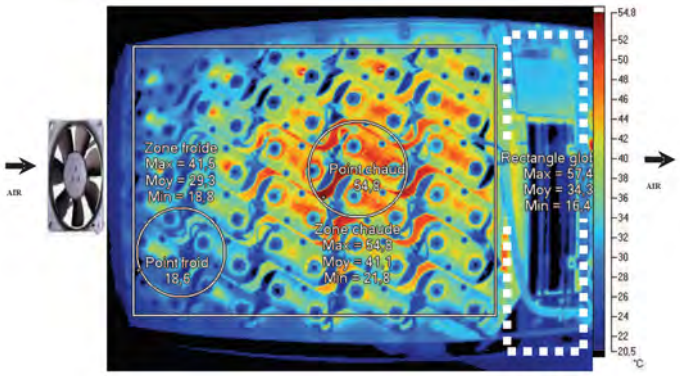
There is currently a lot of opportunities in energy storage. Recently, Renault-Nissan-Mitsubishi, currently the world's largest automotive alliance will invest as much as $1 billion to fund startups working on vehicle electrification, autonomy, connectivity, and artificial intelligence (AI) over the next five years [1]. It follows similar ventures by competitors and their goal is to cut battery costs. On the other hand, General Electric is betting on large-scale applications with their recently released 1.2-MW containerized energy storage product [2].

So indeed, energy storage will play a huge role in a future where energy is more integrated, decentralized, and sustainable. In general, we use devices with high energy density like batteries for bulk energy storage and devices with high power density like supercapacitors as energy buffer in high power transient situations. Energy density tells us how much energy these devices can store per unit mass while power density tells us how much power a device can output per unit mass.

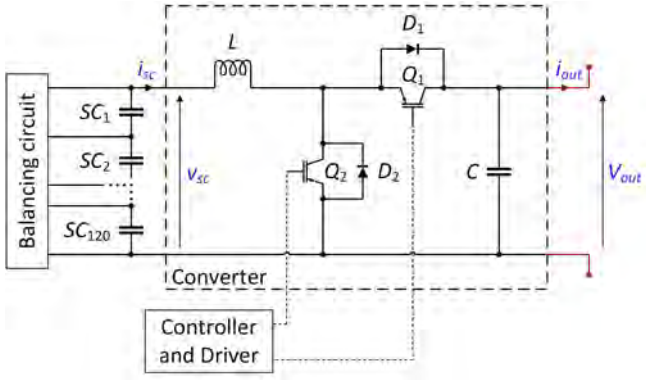
Supercapacitors, otherwise called electrical double-layer capacitors (EDLC), have much higher charge storage capability compared to capacitors. Unlike usual capacitors, there is no dielectric separating the electrodes of supercapacitors. An electrolyte fills the space between the electrodes instead. The electrical double layer (EDL) refers to the two charged layers formed at the electrode-electrolyte interfaces. The high charge storage capability, manifested in the capacitance, is due to the large surface area of activated carbon used in the electrodes and the extremely thin distance between the double layers [3]–[5]. Compared to batteries, much higher charge-discharge cycles and therefore, longer lifetimes are possible with supercapacitors [5].

When they are made, there is a natural tolerance to the parameters of the supercapacitors because manufacturing processes themselves are not uniform [6]. This is also true of other electrical energy storage devices like batteries. The initial imbalance is exacerbated during operation because a cell's

* PhD Student at Laboratoire Ampère – CNRS UMR5005, Université Claude Bernard Lyon 1, Université de Lyon, 69622 Villeurbanne Cedex (e-mail: robert.pena@univ-lyon1.fr).



(a)



(b)

Fig. 1. (a) Thermal image of a 120-supercapacitor stack and (b) the stack connected to a bidirectional buck-boost converter.

initial characteristics affect its performance. In the Hybus project, which was a past research project at Laboratoire Ampère aimed at the recovery of the braking energy of trolleybuses in a stack of supercapacitors, the difference between the maximum and minimum temperatures in a supercapacitor stack could be as high as 20 °C, as can be seen in Fig. 1(a) [7], [8]. Such a high dispersion in operating conditions leads to the uneven aging of supercapacitor cells. An increase of 10 °C reduces supercapacitor lifetime by a factor of two [9]. There is thus a need for a balancing system that can minimize the imbalance in aging by biasing the supercapacitors according to their states of health (SoH).

1.1 Description of the problem

The system in the Hybus project serves as the example and basis for the system which we consider in this thesis. The project aimed to have supercapacitors as auxiliary supply during power cuts in the operation of trolleybuses. Supercapacitors were the energy storage device of choice because of the power requirement of the application. They supply energy during the power cuts, and recharge and recover energy when the trolleybus brakes. The study involved dimensioning the supercapacitors and designing the power electronics considering the power requirements and the conditions under which the system were to operate [7], [8], [10].

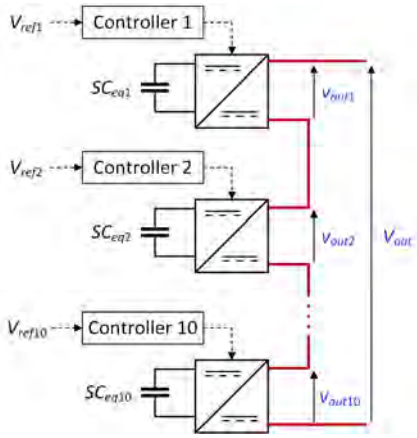


Fig. 2. 120 supercapacitors grouped by ten modular buck-boost converters.

In the Hybus project, 120 supercapacitors with a total voltage varying from 120 to 300 V depending on the charge stored in the stack, were connected to a bidirectional buck-boost converter whose output was regulated at 350 V. The system is shown in Fig. 1(b). In light of minimizing the imbalances among the cells through balancing circuits to extend the system life, the interesting approach of partitioning the big buck-boost converter into smaller modular ones arises. Fig. 2 illustrates such partitioning. The 120 supercapacitors were grouped by ten and the groups are connected to smaller converters whose outputs are connected in series.

Where does balancing to maximize the lifetime come into the picture? As in the one-converter system, the output voltage V_{out} of the system in Fig. 2 is regulated. While evenly dividing the load of the big converter into ten achieves modularity and has reliability benefits in itself, we can further take advantage of the new architecture by considering the parameters crucial to the aging of the supercapacitors, namely voltage, current, and temperature. These, on the other hand, are related to the output voltages and currents of the converters according to the operational characteristics of the converters. Since the converter outputs are connected in series, the same current flows through them. Thus, we can control the amount of power that goes into and out of the supercapacitor groups based on the output voltages of the converters. The research problem, then is looking for the optimum values of converter output voltages that maintain the regulation of the total output voltage and that maximize system lifetime.

1.2 Objectives of the thesis

In this thesis, we will be exploring the latest approach to balancing circuits using both modular power converters and SoH balancing. The objectives of the thesis are:

1. Optimize energy storage performance and lifetime by controlling the connected modular converters according to the states of health (SoH) of the supercapacitor cells.

2. Compare the new strategy to the classical balancing technique in terms of performance and economic benefits.

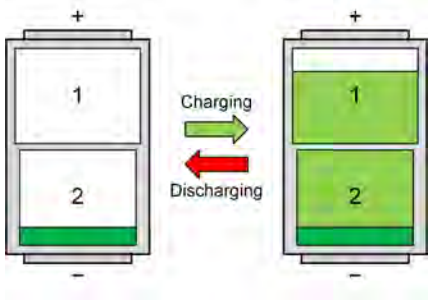


Fig. 3. Energy storage module with cells of different capacities during charging and discharging.

In industrial and commercial applications, passive balancing methods (see subsection 2.2) are still the preferred choice because of the cheaper and easier implementation. If we are able to obtain substantially longer lives and adequate modularity, the interesting question is whether those are enough to compensate for the more involved and expensive implementation.

2 Review of the literature

2.1 Balancing cells

While supercapacitors have high power densities, their individual cell operating voltages are relatively limited for use in most applications. They are thus connected in series to obtain suitable higher voltages. Parallel connection of series strings could also be done to augment current if the application so requires.

As discussed in the previous section, there is a natural dispersion to the characteristics of supercapacitors, which in turn, manifest during operation. A balancing system reduces this dispersion. To illustrate this principle, we have in Fig. 3 an energy storage module with two cells of different capacities connected in series. The capacities are represented by the different sizes of the cells. In addition, cell 2 has an initial charge represented by the area in dark green. Since the cells have the same current, we can only charge up to the capacity of cell 2. Otherwise, cell 2 will become overcharged. Cell 2 thus limits the module resulting in cell 1 being undercharged. The same principle operates during discharge.

This simple example shows the need for a system to balance the states of charge (SoC) of the cells in a module. SoC indicates the remaining electricity available from a storage capacity. Real systems are far more complex and other factors like temperature and aging are important as well during operation [5]. These different parameters interact in complex ways that ultimately result in nonuniform aging [6]. Balancing is thus needed for the safe operation of the cells and for the reduction of aging. Energy storage management systems are therefore generally required. Aside from balancing, these management systems take on a variety of functions such as: monitoring of voltages, currents, and temperature and interfacing with other devices [5].

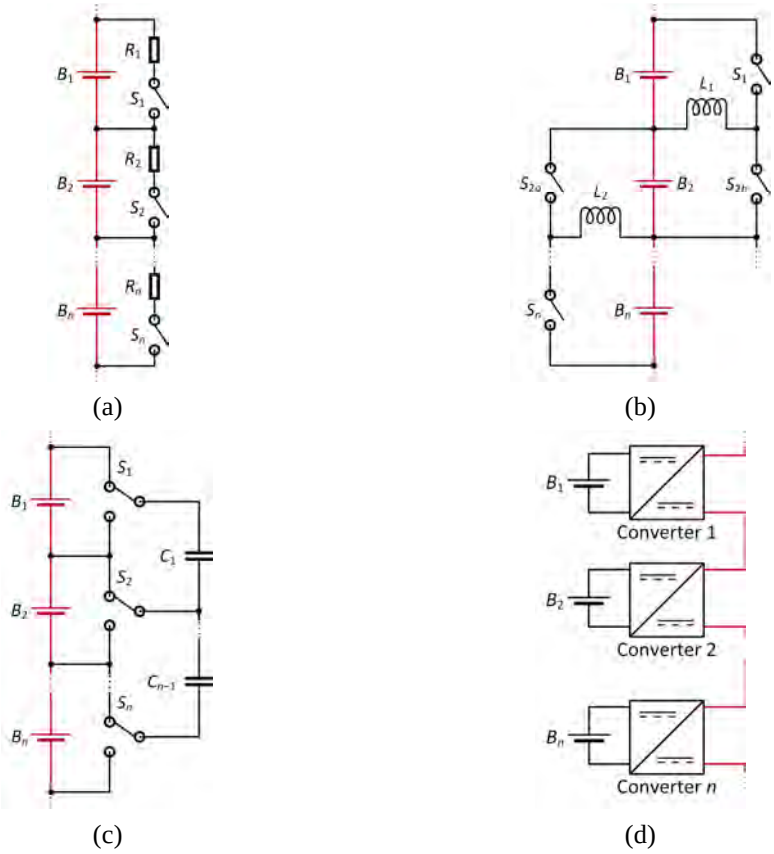


Fig. 4. (a) Passive balancing circuit; (b) switched-mode converter active balancing circuit; (c) switched-capacitor active balancing circuit; and (d) modular power converter balancing circuit.

2.2 Balancing circuits

In the literature, there are three types of balancing circuits: passive, active, and modular power converter balancing circuits [11]. These three have fundamental differences in terms of circuits, control, and operation. Each has its own advantages and disadvantages as well.

Fig. 4 shows the different types of circuits used in balancing cells in energy storage systems such as batteries or supercapacitors. In passive balancing circuits, shown in Fig. 4(a), the excess charge in the cells is dissipated through resistors. This is the simplest and cheapest strategy to implement. On the other hand, active balancing circuits, seen in Fig. 4(b)–(c), redistribute charge from overcharged to undercharged cells. We can see that active balancing circuits are more complicated and thus, more expensive to implement. The advantage, however, is in their better efficiencies.

The latest approach in balancing is to use modular power converters as shown in Fig. 4(d). Its architecture is fundamentally different. Unlike in the previous circuits, power to the rest of the system is drawn from the output of the converter. The output voltage of converters can be controlled, and by doing that, we also control the power we draw from the individual cells [12]–[14].

2.3 Power converters in balancing

Power converters of the type in Fig. 4(d) have a long history of being used in interfacing energy storage devices. Their flexibility in terms of architecture and control lend themselves for use in electric vehicles and grid applications [13], [15], [16]. Their use for balancing energy storage devices was born out of the need to integrate more storage cells for more power.

Aside from the task of voltage regulation, the converters have to be controlled in such a way that the other task of balancing is achieved. Most of the algorithms do these two by balancing the cells' SoCs. The total energy is shared among the converters such that, when charging for example, cells with higher SoCs get less charged compared to cells with lower SoCs [14]. For incorporating second-life batteries that have varying SoHs into a system, SoC balancing can still be done but system stability becomes a crucial factor. Thus, more robust control is needed [13], [17]. One study looked into adding the third task of thermal balancing to ensure the uniform aging of battery cells [18]. In the case of supercapacitors, converters were designed for reduced inductor size and weight, and controlled for voltage balancing of the supercapacitors [19].

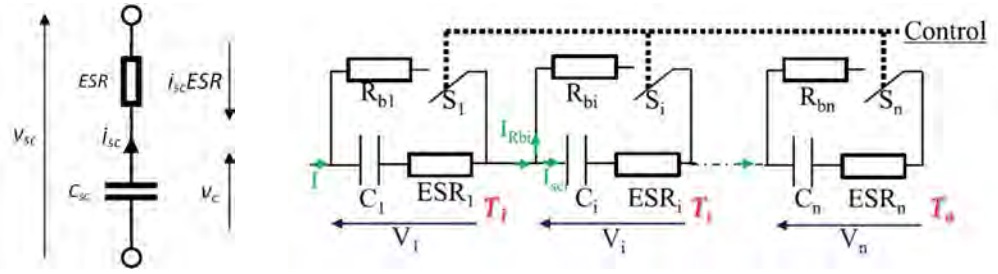
2.4 Balancing to maximize lifetime

The conventional goal of balancing circuits is to balance the SoCs of individual cells, or in the case of supercapacitors, to balance voltages [5]. These balancing circuits only approximately minimize aging. An altogether different approach was started in the laboratory by the work of Shili [20], [21]. She directly used a model for the lifetime of each cell to control a passive balancing circuit with the express goal of minimizing the aging imbalance and prolonging the lifetime of the whole system. This is the first time that instead of balancing SoCs, we are balancing SoHs. SoH is the parameter that indicates storage degradation in terms of capacity and power fade [22]. It tells us how much useful life is left.

Shili's work uses a lifetime model for supercapacitors which was also developed in the laboratory [23]:

$$\tau_d(t) = \frac{\tau_0}{\frac{1}{T}\int_{t-T}^{t} \exp\left(\frac{v(t)}{V_0} + \frac{\theta(t)}{\theta_0} + \frac{i_{rms0}(t)}{I_{rms0}}\right)dt}, \tag{1}$$

where τ_d is the dynamic lifetime (s); $v(t)$ is the dynamic value of the voltage (V) across the supercapacitor; $\theta(t)$ is its dynamic temperature (°C); and $i_{rms}(t)$ is the dynamic rms current (A). V_0, θ_0, and I_{rms0} are constants that relate the latter three quantities to their effect on lifetime, while τ_0 is the theoretical lifetime for a supercapacitor at 0 V and 0 °C [10], [20].



(a) (b)

Fig. 5. (a) A simple model of a supercapacitor and (b) the model used in a passive balancing circuit.

In Fig. 5(a), we have the basic supercapacitor model, which is that of a capacitor in series with the supercapacitor's equivalent series resistance (ESR). As a supercapacitor is used, its capacitance decreases and the *ESR* increases until its end of life. The usual end-of-life criteria are when a reduction of 20% is reached for a supercapacitor's capacitance or its *ESR* increases by 100% [24].

This model was used to propose a method of controlling the passive balancing circuit to prolong system life. Dissipation of the shunt resistors in the circuit in Fig. 5(b) was controlled through finite control set model predictive control with the goal not to balance voltages or SoCs, but to balance supercapacitors' SoHs instead. The novel method was compared to the classical voltage balancing strategy and was found to have extended lifetime by 20%. For both strategies, the stack's lifetime was set by the cell that reached its end of life first.

3 Balancing through modular power converters

3.1 Control allocation based on reliability indicators

As mentioned, a modular system has reliability benefits. This has mainly to do with the maintenance and replaceability of components. Consequently, another advantage it brings for the system in Fig. 2 lies in redundancy. For example, when there is a fault in one of the converters, the rest can still function and shoulder the load. Controlling such a system with the express goals of fault-tolerance and reliability is a difficult task.

For a general linearized system such as:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_u u(t) \\ y(t) &= Cx(t) \end{aligned} \tag{2}$$

where x, u, and y are the state, control, and output vectors; A, B_u, and C are the state, control, and output matrices; and where additionally B_u can be factorized as $B_u = B_v B$, $v(t)$ is the virtual control input and represents the total control effort by the actuators. The control allocation problem can be expressed as the constrained linear mapping problem:

$$v(t) = Bu(t), \tag{3}$$

$$u_{min} \leq u \leq u_{max}, \tag{4}$$

where u_{min} and u_{max} in (4) are the saturation points of the actuators. Furthermore, in managing the redundant actuators, we use a weighting matrix W_u, which is based on the health or reliability of the actuators. It is defined as:

$$W_u = \begin{bmatrix} \frac{r_1}{r_{max}} & 0 & \cdots & 0 \\ 0 & \frac{r_2}{r_{max}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{r_n}{r_{max}} \end{bmatrix} \tag{5}$$

where r_i is the reliability indicator taken as the baseline or nominal failure rate (λ_i in 1/s) of the ith actuator and r_{max} is the maximum among all the reliability indicators [25], [26].

In control allocation, the weighting matrix is used to allot the control input into the different actuators. Maximizing the reliability of a system based on it is thus an optimization problem. Optimal control-based solutions are used to solve the control allocation problem in (3). They aim to look for the optimal solution, the best one if there many solutions, or the most feasible one if no exact solution exists. For example, the optimal solution can be solved using a two-step sequential quadratic programming:

$$\psi = \underset{u_{min} \le u \le u_{max}}{\arg\min} \; \|Bu - v\|_2, \tag{6}$$

$$u = \arg\min_{u \in \psi} \|W_u(u - u_d)\|_2, \tag{7}$$

where u_d is the desired control input. If the constraint in (4) is not considered, the quadratic programming problem can be minimized into:

$$\begin{aligned} &\min_u J = \|W_u u\| \\ &\text{s.t. } Bu(t) = v(t), \end{aligned} \tag{8}$$

where an explicit analytic expression for the optimal solution can be obtained. This solution is based on the pseudoinverse method for control allocation and is given by:

$$u^*(t) = W_u^{-1}\left(BW_u^{-1}\right)^+ v(t), \tag{9}$$

where the "+" sign in superscript is the pseudoinverse operator. The analytic nature of the optimal solution in the pseudoinverse method makes its application in many systems easy. While there is no guarantee that the solution is within the saturation points because the constraint was disregarded, there are methods to take care of it. In the redistributed pseudoinverse method, for example, saturated actuators are removed from the optimization problem. The pseudoinverse formula is then just solved for the remaining unsaturated actuators [25], [26].

To achieve the objectives in subsection 1.2 according to the formulation of the problem in subsection 1.1, we employ a control allocation module that is separate from the task of regulation. Its implementation is easy with the formula for the optimal solution in (9). The previous discussion of the technique demonstrates that it considers the health of the components in the system and aims to improve overall system reliability.

Adapting the control allocation method using the pseudoinverse operator for our system in Fig. 2, the control input is: $u^*(t) = [V_{ref1} \quad V_{ref2} \quad \cdots \quad V_{ref10}]^{\mathrm{T}}$ and $B = [1 \quad 1 \quad \cdots \quad 1]$, such that:

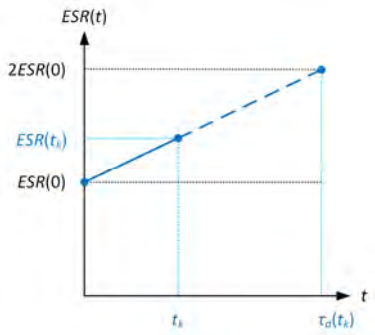


Fig. 6. Supercapacitor *ESR*(*t*)

$$Bu^*(t) = v(t) = \sum_{i=1}^{10} V_{refi} = V_{out}, \tag{10}$$

where V_{refi} is the reference voltage (V) for the *i*th converter and V_{out} is the total output or bus voltage (V). Using the pseudoinverse method in (9), should result in supercapacitors with better health being loaded more and supercapacitors with less health being loaded less. Thus, to use this method to obtain the reference voltages needed to control the modular power converters, the weighting matrix based on the model for supercapacitor lifetime must be designed.

3.4 Supercapacitor remaining life as the reliability indicator

In the general system considered in the previous subsection, an actuator's failure rate served as the reliability indicator for the control allocation. There are many definitions for failure rate and specific ones depend highly on the components, systems, and conditions under consideration [25], [26]. However, it is generally understood as the frequency of failure of a component or a system and is expressed as the number of failures per unit time.

In the system we are considering, we take the inverse of the remaining life of supercapacitors as the reliability indicator:

$$r_i(t) = \frac{1}{\tau_{di}(t) - t}, \tag{11}$$

where $r_i(t)$ is the reliability indicator of the *i*th supercapacitor group and $\tau_{di}(t)$ is its dynamic lifetime (s), because a supercapacitor is deemed failed when it reaches the end of its lifetime and its remaining life is the amount of time left until failure.

To obtain an expression for the remaining life, we consider Fig. 6. This initial design for the supercapacitor remaining life as the reliability indicator assumes that from a macro perspective, *ESR* evolves linearly. It is thus easy to obtain an expression for the remaining life:

$$\tau_{di}(t) - t = \frac{t[2ESR_{0i} - ESR_i(t)]}{ESR_i(t) - ESR_{0i}}, \tag{12}$$

where ESR_{0i} and $2ESR_{0i}$ are the initial and end-of-life *ESR*s (Ω) of the *i*th supercapacitor group. However, this initial design for the reliability indicator disregards the dynamic evolution of *ESR* and has to be modified. A better expression must depend not on the initial value of *ESR*, but on the value from a previous time step. Work is currently underway to accomplish this. In the meantime, (12) was used for the initial results in the succeeding section.

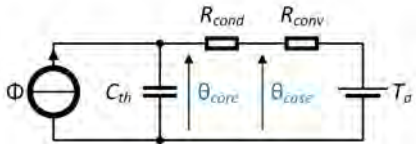


Fig. 7. A thermal model for supercapacitors.

4 Implementation and initial results

4.1 Supercapacitor modelling

The model for a supercapacitor in Fig. 5(a) was implemented using Simscape, an environment within Simulink for multidomain physical modelling and simulation. A variable resistor block for ***ESR*** and a variable capacitor block for C_{sc} represented the supercapacitor model. Considering the end-of-life criteria for a supercapacitor, we use the dynamic lifetime model in (1) to update the value of $ESR_i(t)$ every time:

$$ESR_i(t) = \int_0^t \frac{ESR_{0i}}{\tau_{di}(t)} dt + ESR_{0i}. \tag{13}$$

$ESR_i(t)$ here represents the equivalent ***ESR*** of the ***i***th supercapacitor group. Each group is composed of 12 supercapacitors. It is assumed for simulation purposes that supercapacitors within a group have uniform characteristics. The same principles hold for the capacitance, but because the end-of-life criterion for this parameter is when it reaches 80% of its original value, the equivalent initial capacitance has to be multiplied with −0.2:

$$C_{sci}(t) = \int_0^t \frac{-C_{sc0i}}{\tau_{di}(t)} dt + C_{sc0i}, \tag{14}$$

where C_{sc0i} is the initial capacitance (F) of the ***i***th supercapacitor group. Because of (13) and (14), both the ***ESR*** and C_{sc} change at the same rate and reach the end-of-life criteria at the same time. Since it does not matter which of the two parameters to track, we express supercapacitor ***SoH*** in terms of ***ESR***:

$$SoH_i(t) = 100 \cdot \frac{2ESR_{0i} - ESR_i(t)}{ESR_{0i}}. \tag{15}$$

The reliability indicator in (11) and (12) is also based on ***ESR***(***t***).

Another important parameter in the aging of supercapacitors is temperature. A circuit model for the thermal behavior was used and is shown in Fig. 7 [8], [21]. In this model, Φ is the heat generation from Joule heating (W); C_{th} is the thermal capacity (J/K) of the supercapacitor; R_{cond} is the thermal resistance (K/W) to conduction and models heat transfer from the core of the supercapacitor to its case; R_{conv} is the thermal resistance (K/W) to convection and models heat transfer between the case and the air; and T_a is the ambient temperature (K). θ_{core} represents the temperature (K) of the supercapacitor core and θ_{case} represents the case temperature (K).

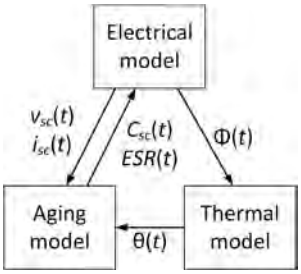


Fig. 8. The supercapacitor multiphysical model and the interactions among the models

The supercapacitor core temperature (in °C) is fed into the dynamic lifetime model in (1). To complete the parameters in the lifetime model, voltage- and current-sensing blocks were employed in the Simscape model to measure these parameters. An additional rms block was needed for the measured current before being fed into the lifetime model. With the aging, thermal, and electrical models, and the interdependency among them as can be seen in Fig. 8, we now have a suitable multiphysical model we can use for the supercapacitors.

4.2 Modelling the converter

To be able to use our supercapacitor model in simulation with modular power converters, we must have a model for the converter as well. A nonisolated bidirectional buck-boost dc-dc converter is used in this case and is shown in Fig. 9. The original converter for 120 supercapacitors in the Hybus project was sized to have an inductor with L_{120sc} =160 mH and a capacitor with C_{120sc} = 1,600 mF for a maximum voltage ripple of 2% and a maximum current ripple of 12% [7]. A pulse-width modulated (PWM) signal with a frequency of 10 kHz serves as the control signal [7]. For the modular converter approach, the 120 supercapacitors were grouped by ten. Each group has 12 supercapacitors, which now shoulders a tenth of the original load. Based on this and on the formulas for the ripple voltage and current [27], the new inductor size L_{12sc} = 16 mH should be ten times smaller than the original and the new capacitor C_{12sc} = 16 mF ten times larger to keep the same ripple values.

The PWM signal controlling the two insulated-gate bipolar transistor (IGBT) switches are complementary. Its duty cycle is represented as d. While it is possible to implement the switched model in Simscape, we will have faster simulation if we used the averaged model instead. The goal is to simulate the whole system lifetime, so it is imperative that we do it fast, and derive and use the averaged model of the converter.

We consider the two states of the converter. The first state is: switch Q_1 is off and Q_2 is on; and the second state is: switch Q_1 is on and Q_2 is off. We then look for each state's system equation of the form $\dot{x} = Ax + Bu$. We then solve for the average values of A and B over the switching period: $\bar{A} =$

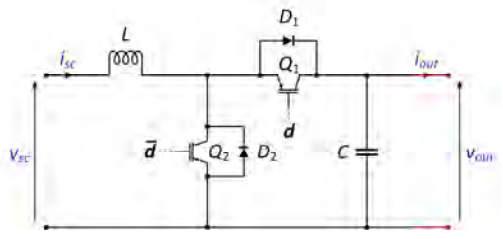


Fig. 9. A nonisolated bidirectional buck-boost dc-dc converter.

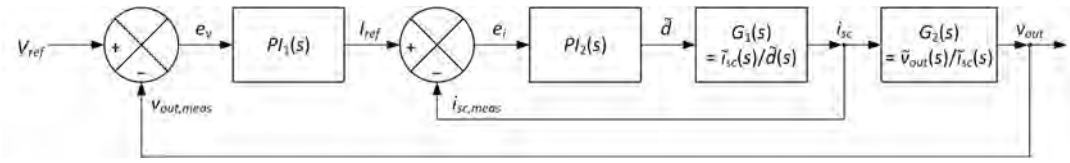


Fig. 10. A two-loop control system for controlling the converter.

$dA_1 + (1-d)A_2$ and $\bar{B} = dB_1 + (1-d)B_2$ to find the averaged model. Doing that, we arrive at our goal, the averaged model of the converter:

$$\begin{bmatrix} \frac{di_{sc}}{dt} \\ \frac{dv_{out}}{dt} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1-d}{L} \\ \frac{1-d}{C} & 0 \end{bmatrix} \begin{bmatrix} i_{sc} \\ v_{out} \end{bmatrix} + \begin{bmatrix} \frac{1}{L} & 0 \\ 0 & -\frac{1}{C} \end{bmatrix} \begin{bmatrix} v_{sc} \\ i_{out} \end{bmatrix}. \tag{16}$$

We then use this averaged model in (16) to build the Simscape model of the nonisolated bidirectional buck-boost dc-dc converter in Fig. 9 that we then connect to our supercapacitor model.

4.3 Controlling the converter

For controlling the converter's voltage output, the two-loop control system in Fig. 10 was used. The slower outer loop is for controlling the voltage while the faster inner one is for controlling the current. We use two cascaded control loops because they respond to disturbances better and correct nonlinearities better than a single-loop system. The reference voltage from the control allocation module serves as input to the control system. The first proportional-integral (PI) controller outputs the current reference and the second one outputs the duty cycle of the control signal for the converter.

In Fig. 10, we see the two plant models $G_1(s)$ and $G_2(s)$. They describe the behavior of the converter. We can obtain them by using the averaged model of the converter in (16) and doing a small-signal analysis of the system. In small-signal analysis, we assume that each signal has a nominal value and a small perturbation about that value. It is basically a method of linearizing an otherwise nonlinear system. For example, our system in (16) is nonlinear because the control variable d has to be multiplied with the state variables i_{sc} and v_{out}. We are able to linearize nonlinear systems because we can disregard multiplication of two perturbed variables [28]. Without going into the details of the derivation, the small-signal model based on the averaged model of the converter is:

$$\begin{bmatrix} \frac{d\tilde{\imath}_{sc}}{dt} \\ \frac{d\tilde{v}_{out}}{dt} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1-D}{L} \\ \frac{1-D}{C} & 0 \end{bmatrix} \begin{bmatrix} \tilde{\imath}_{sc} \\ \tilde{v}_{out} \end{bmatrix} + \begin{bmatrix} \frac{1}{L} & 0 & V_{out} \\ 0 & -\frac{1}{C} & -\frac{I_{sc}}{C} \end{bmatrix} \begin{bmatrix} \tilde{v}_{sc} \\ \tilde{\imath}_{out} \\ \tilde{d} \end{bmatrix}, \tag{17}$$

where the variables in capital letters are the nominal values of the signals and the variables with the tilde on top are the perturbed variables. We see that $\tilde{d}$ is now an element of the control vector. The small-signal model in (17) was solved in the s domain for the transfer functions of the two plants:

$$G_1(s) = \frac{\tilde{\imath}_{sc}(s)}{\tilde{d}(s)} = \frac{V_{out}RCs + 2V_{out}}{RLCs^2 + Ls + R(1-D)^2}, \tag{18}$$

$$G_2(s) = \frac{\tilde{v}_{out}(s)}{\tilde{\imath}_{sc}(s)} = \frac{V_{out}R(1-D) - I_{sc}RLs}{V_{out}RCs + 2V_{out}}, \tag{19}$$

The component sizes and the nominal values at the input and output of the converter were substituted into (18) and (19) to obtain the specific transfer functions. These were then used in tuning

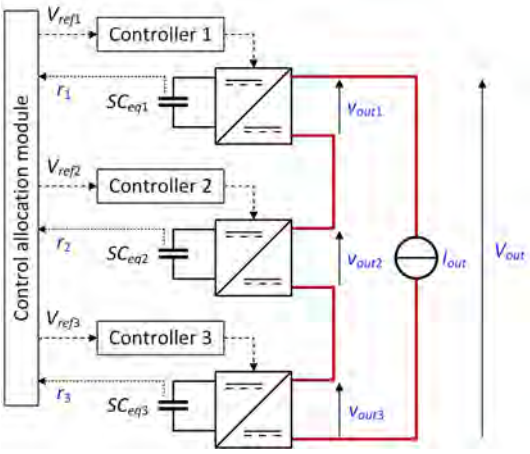


Fig. 11. The three-converter system used for simulation.

the two PI controllers in the MATLAB Control System Designer tool. The response time for the inner loop was set at 1 ms and 5 ms for the outer loop. A faster inner loop is required for the effective operation of the control system in Fig. 10. Phase margin for the two loops are at least 60°. The tuned transfer functions for the two controllers are:

$$PI_1(s) = K_{p1} + \frac{K_{i1}}{s} = 13.181 + \frac{12{,}792.161}{s}, \tag{20}$$

$$PI_2(s) = K_{p2} + \frac{K_{i2}}{s} = 0.00052096 + \frac{0.20307}{s}, \tag{21}$$

where K_p is the proportional gain and K_i is the integral gain of the controller. The control system was implemented using Simulink blocks, while measurement was done using Simscape blocks. At this point, it is now possible to simulate the complete system with the supercapacitor, converter, control allocation, and control system modules.

4.4 Simulation

We simulate a smaller system of three converters for the initial simulation results here. The schematic diagram in Fig. 11 shows the system under consideration. The total voltage output V_{out} is regulated at 105 V. An ideal current source I_{out} sinks and sources current at ±50 A. It roughly simulates what happens in an electric vehicle for example: the supercapacitors supplying power to the motor (or discharging) and the supercapacitors drawing in power from regenerative braking (or charging). The period for this discharge-charge cycle was set at 40 s. The regular discharging and charging makes it easy as well to compare the modular converter approach with the classical balancing one. On average, each converter processes 5.25 kW of power.

Table 1 shows the parameters used in the simulation. The control allocation module provides the reference voltages, which are held for a period of 10 s. The values for the three groups, each connected to a converter, are the same except for the initial capacitance C_{sc0} and the thermal resistance to convection R_{conv}. The variation in capacitances was set not only to simulate real conditions, but to induce a difference in the voltages of the supercapacitor groups at the end of the discharge-charge cycling. This is required for the comparison with the classical balancing later. In future simulations, such a variation in capacitances must be accompanied by a variation in the initial ESR_0 as well.

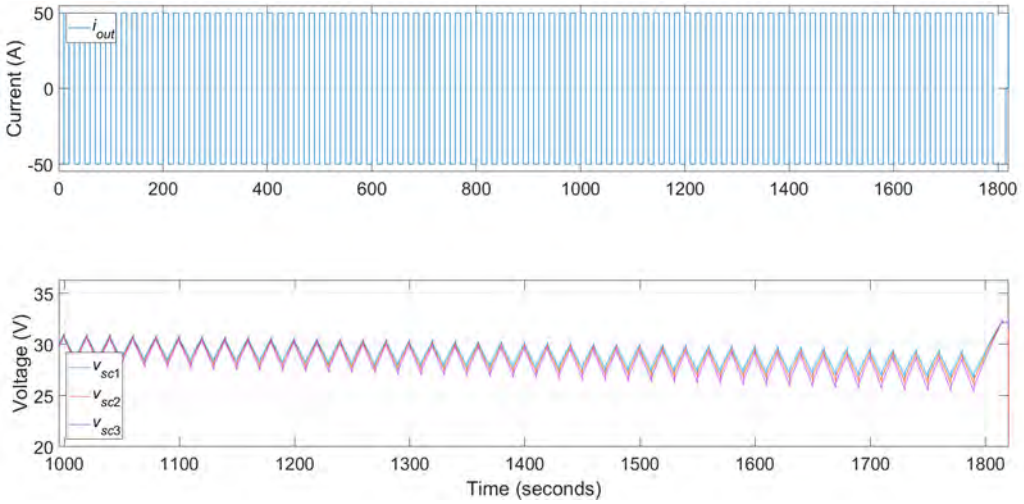


Fig. 12. The output current of the converters and the voltages of the supercapacitors.

Meanwhile, the variation in the thermal resistance to convection values was set to simulate the different positions of the supercapacitor groups in a 120-cell stack. A supercapacitor group's heat generation affects the thermal behavior of its neighboring groups.

The normal discharge-charge cycling of supercapacitors happens for 30 min. Fig. 12 shows the 1,000–1,800-s portion of the converter output current and the voltages of the three supercapacitor groups during the first discharge-charge cycling. The decreasing supercapacitor voltages are mainly due to the efficiencies of the supercapacitors, which are around 95%. A portion of the power that goes through a supercapacitor is lost as heat in its *ESR*. The uneven dynamics of the converters for their buck and boost modes also contribute to this loss but is negligible. Since the converters are ideal and process a constant amount of power at steady state, the decrease in the voltages is compensated by the increase in supercapacitor currents.

As can be seen in Fig. 12, starting at 1,790 s, a different phase operates for around 15 s. This is the charging phase and during this period, a voltage-balancing technique operates. Once the supercapacitor cells reach 2.7 V, charging stops and the system enters a rest period of around 5 s until the normal discharge-charge cycling operates again at 1,820 s. Instead of getting the reference

Table 1. Simulation parameters.

Parameter	**SC group 1**	**SC group 2**	**SC group 3**
V_{ref} (V)	Control allocation	Control allocation	Control allocation
V_{ref} period (s)	10	10	10
C_{sc0} (F)	3,150 ÷ 12	3,000 ÷ 12	2,850 ÷ 12
ESR_0 (mΩ)	12 × 0.27	12 × 0.27	12 × 0.27
V_0 (V)	12 × 2.7	12 × 2.7	12 × 2.7
T_a (°C)	25	25	25
C_{th} (J/K)	700	700	700
R_{conv} (K/W)	4.9	5.9	6.9
R_{cond} (K/W)	0.627	0.627	0.627
Acceleration factor	10	10	10

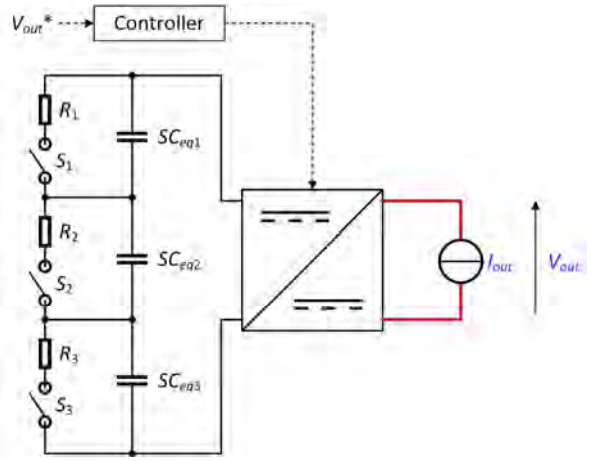


Fig. 13. The three supercapacitor groups using the classical balancing technique.

voltages from the control allocation module, we get them from the voltage-balancing module governed by the equation:

$$V_{crefi}(t) = V_{out} \cdot \frac{\omega_i \upsilon_{sci}(t)}{\sum_{m=1}^{n} \omega_m \upsilon_{scm}(t)}, \tag{22}$$

$$\omega_i(t) = \frac{C_{sci}(t)[V_{max} - \upsilon_{sci}(t)]}{\sum_{m=1}^{n} \upsilon_{scm}(t) C_{scm}(t)[V_{max} - \upsilon_{scm}(t)]}, \tag{23}$$

where $V_{crefi}(t)$ is the reference voltage for the ith supercapacitor group during the charging phase, $\omega_i(t)$ is its weighting based on the charge capacities of all the n converters, and V_{max} = 12 × 2.7 V. The same controllers were used. The disadvantage with this voltage-balancing technique is that $C_{sci}(t)$ must be always known. For simulation throughout the system lifetime, the discharge-charge cycling and the charging phase will be repeated until one of the supercapacitor groups reaches its end of life.

An earlier technique used during the charging phase involved the use of bypassing switches at the output of the converters. For this technique, the reference voltages still came from the control allocation module. When the supercapacitor cells in a group reach 2.7 V, that group is bypassed. This happens until all three groups reach V_{max}. The supercapacitors thus start at the same maximum voltage for the next discharge-charge cycling. The problem with using this technique was the voltages do not reach the maximum possible at the same time. It was a crude implementation, so it was discontinued in favor of the voltage-balancing technique in (22) and (23). The bypassing switches could still be useful when the supercapacitors or the converter in a group has a fault [13], [17]. The switches could bypass the faulty group, so the rest of the system can shoulder its load.

4.5 Comparison with classical balancing

We now compare the new strategy towards lifetime balancing, the modular converter approach, with the classical balancing strategy. For the classical balancing strategy, the switched resistors (SR) circuit, as shown in Fig. 13, was used. Supercapacitor voltages are measured every 0.1 s and if a group has a voltage that is greater than 1% of the minimum among all the groups, that group is shunted. Each shunt resistor was designed to have R = 120 Ω. This was done to ensure that the current through the shunt resistor has a value of 270 mA when the supercapacitor group has a voltage of V_{max}. Power dissipation in the shunt resistors was also limited to 8.75 W because of this.

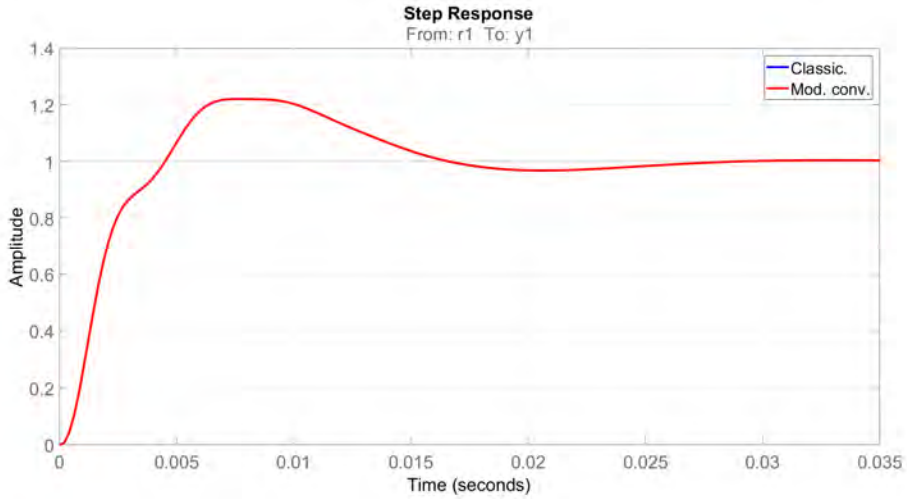
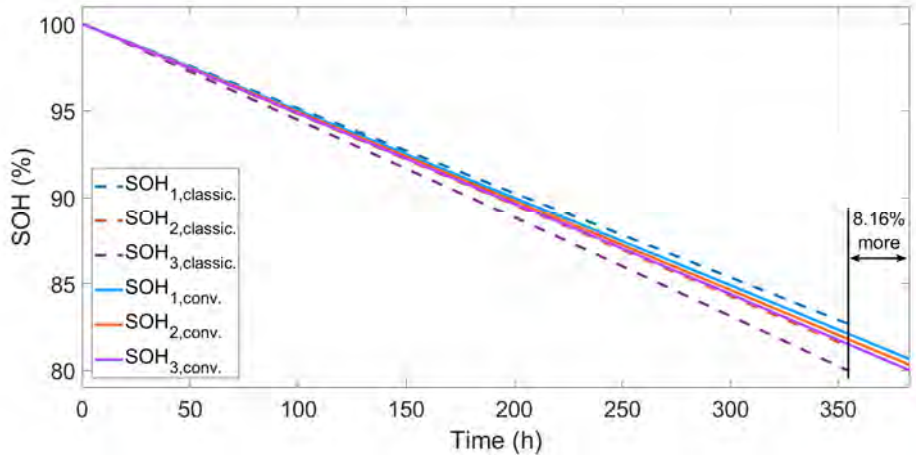


Fig. 14. Comparison of the step responses in the classical and modular converter approaches.

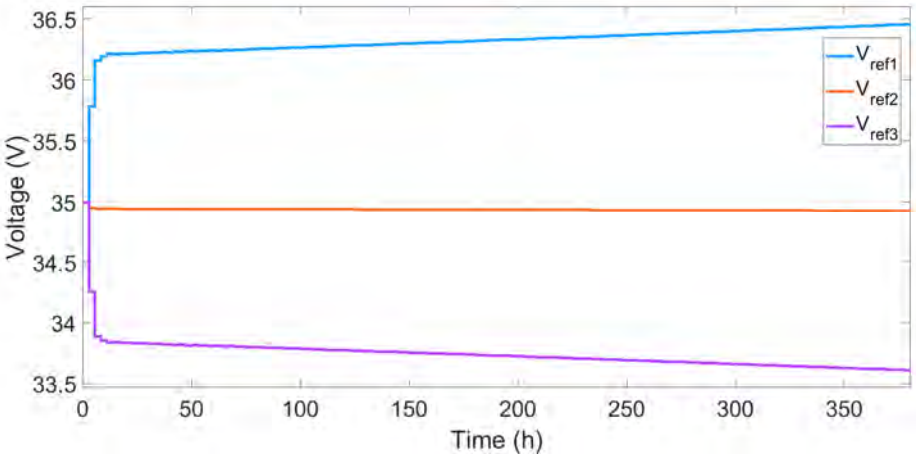
The same procedures for modelling and controlling the converter in the SR circuit were followed as in subsections 4.2 and 4.3 but adapted for the sizing in Fig. 13. Additionally, the dynamic performance of the converter for the classical balancing strategy and its control was compared with that of the modular converter approach to ensure that the comparison is valid. A comparison of their step responses is shown in Fig. 14. There is a very small difference in terms of the performance of their dynamics, so in this aspect, it was taken to mean that the control systems for the two approaches are comparable.

Fig. 15(a) shows the comparison of the simulation results for the two approaches. It shows that the modular converter approach can prolong the system lifetime by 8.16%. The *SoHs* of the supercapacitors in the modular converter approach are much closer to one another, compared to those of the classical balancing approach, which are more dispersed. In both cases, supercapacitor group 1 has a higher *SoH* compared to the others. Group 2 follows it, then group 3. The modular converter approach is able to extend system lifetime because supercapacitor group 1 is loaded more, compared to the other two. The load that group 1 takes is freed up for the other two groups. Because of this, supercapacitor groups 2 and 3 are loaded less. However, the effect is most noticeable for group 3, which has the worst *SoH*. Fig. 15(b) illustrates this point. It shows the reference voltages for the three converters, which are also the output voltages of the converters at steady state. The output current is the same for all three converters, so the reference voltages basically show how the supercapacitor groups are loaded. Remember that the reference voltages are computed through the pseudoinverse method in (9) based on the supercapacitor remaining life as the reliability indicator. The extension of the system lifetime would not have been possible if the three converters were loaded equally with a reference voltage of $V_{out} \div 3 = 35$ V.

While this initial result looks promising, some factors still have to be modified to make the comparison valid. One of those factors is the algorithm for the SR circuit. In the current setup, the resistors act as protective shunts that turn on when V_{max} is reached. In fact, the shunt resistors did not turn on at all because the 1% condition was not reached yet. An actual voltage-balancing strategy for the SR circuit must be implemented instead. Another factor is that for the simulation of the modular converter approach, the voltage-balancing strategy in (22) and (23) was not yet implemented. Future



(a)



(b)

Fig. 15. (a) Comparison of SoHs for the classical and modular converter approaches and (b) the reference voltages.

simulations will incorporate the algorithm for this voltage-balancing strategy. The earlier technique of using bypassing switches was used for the initial results in Fig. 15.

5 Conclusion and future work

We have shown how important it is to solve the problem of uneven aging among supercapacitor cells. The reliability of an energy storage system using supercapacitors depends on it. Solving the problem also has the additional benefit of extending the system lifetime, which then results in reduced costs and savings. We have proposed in this thesis a novel method of balancing that uses modular power converters to control the loading of supercapacitor cells based on their states health (SoHs). Conventional balancing methods are based on cells' states of charge (SoC) and do not directly optimize system lifetime. We demonstrated how the new method is possible through modelling and simulation. Control allocation based on the supercapacitor remaining life as the reliability indicator

Table 2. The Gantt chart for the thesis.

Activity	2017–2018				2018–2019				2019–2020			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Review the literature and the state of the art.												
Perform simulation studies.												
Pseudoinverse method of control allocation.												
Remaining life reliability indicator.												
Compare with classical balancing .												
Control allocation of saturated converters.												
Design control algorithms.												
Run fast multiphysical simulation.												
Develop test bench.												
Prepare the thesis manuscript.												

is at its core. While there remains work to improve the modelling and simulation, initial results look promising. They show that the use modular power converters can balance the SoHs of the supercapacitors and extend overall system lifetime.

Table 2 shows the Gantt chart for the required steps to achieve the objectives and finish the thesis per quarter in three academic years. The green color indicates full or partial completion and the yellow color indicates that the task is currently underway. We are currently doing simulation studies and literature review at the same time. Issues and possible solutions to what have been done so far especially with the remaining life as the reliability indicator, were identified and discussed in the previous sections. Fine-tuning, improvements, or changes needed will be made. They are on track to be completed before the first year ends. Once these are done, we will proceed to the crucial step of implementing the redistributed pseudoinverse method for the control allocation of saturated converters. For the next academic year, the next steps are designing the control algorithms, developing the test bench, and implementing the algorithms to gather experimental data. In parallel, improvements to the multiphysical simulation will be done to make it more functional. And lastly, of course, the thesis manuscript will be written.

Along the course, we will also submit articles to journals and conferences. The following list the planned articles for submission to journals and conferences:

1. The effectiveness of the voltage-balancing strategy in (22) and (23) considering the voltage-dependence of the charge storage capacity of supercapacitors.

2. Cyber-physical cooperative control of modular power converters for the voltage-balancing of the connected supercapacitors.

3. Comparison of the modular power converter approach implementing the control allocation module with the classical balancing approach in terms of lifetime optimization.

4. Accelerated aging experiments comparing the modular power converter (with a control allocation module) and the classical balancing approaches in terms of lifetime optimization.

Acknowledgment

R. A. S. Peña would like to thank the Commission on Higher Education (CHED) of the Philippines and the Embassy of France to the Philippines for granting him a CHED-PhilFrance Scholarship, which supports his doctoral studies.

References

[1] "Nissan-Renault Plans $1 Billion Fund for Auto Tech Startups," *Bloomberg.com*, 09-Jan-2018.

[2] J. Spector, "GE Says Its New Battery Product Can Cut Grid-Scale Storage Installation Times in Half," 14-Mar-2018.

[3] "Murata Supercapacitor Technical Note," Murata Manufacturing Co., Ltd., Technical Note.

[4] S. Zhang and N. Pan, "Supercapacitors Performance Evaluation," *Adv. Energy Mater.*, vol. 5, no. 6, p. 1401401, Mar. 2015.

[5] A. Rufer, *Energy Storage: Systems and Components*. Boca Raton, FL, USA: CRC Press, Taylor & Francis Group, 2018.

[6] M. M. Hoque, M. A. Hannan, A. Mohamed, and A. Ayob, "Battery charge equalization controller in electric vehicle applications: A review," *Renew. Sustain. Energy Rev.*, vol. 75, pp. 1363–1385, Aug. 2017.

[7] A. Hijazi, "Modélisation électrothermique, commande et dimensionnement d'un système de stockage d'énergie par supercondensateurs avec prise en compte de son vieillissement: application à la récupération de l'énergie de freinage d'un trolleybus," Université Claude Bernard Lyon 1, Université de Lyon, Villeurbanne, France, 2010.

[8] A. Hijazi, P. Kreczanik, E. Bideaux, P. Venet, G. Clerc, and M. Di Loreto, "Thermal Network Model of Supercapacitors Stack," *IEEE Trans. Ind. Electron.*, vol. 59, no. 2, pp. 979–987, Feb. 2012.

[9] D. Linzen, S. Buller, E. Karden, and R. W. DeDoncker, "Analysis and Evaluation of Charge-Balancing Circuits on Performance, Reliability, and Lifetime of Supercapacitor Systems," *IEEE Trans. Ind. Appl.*, vol. 41, no. 5, pp. 1135–1141, Sep. 2005.

[10] P. Kreczanik, "Étude de la fiabilité et du vieillissement d'un système de stockage par supercondensateurs pour l'alimentation partielle et ponctuelle d'un trolleybus grâce à la récupération de l'énergie de freinage: approche du composant au système de stockage," Université Claude Bernard Lyon 1, Université de Lyon, Villeurbanne, France, 2011.

[11] F. Altaf, "On modeling and optimal control of modular batteries: thermal and state-of-charge balancing," PhD thesis, Chalmers University of Technology, Göteborg, Sweden, 2016.

[12] F. Altaf, L. Johannesson, and B. Egardt, "On Thermal and State-of-Charge Balancing using Cascaded Multi-level Converters," *J. Power Electron.*, vol. 13, no. 4, pp. 569–583, Jul. 2013.

[13] N. Mukherjee and D. Strickland, "Control of Cascaded DC–DC Converter-Based Hybrid Battery Energy Storage Systems—Part I: Stability Issue," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2340–2349, 2016.

[14] W. Huang and J. A. Abu Qahouq, “Energy Sharing Control Scheme for State-of-Charge Balancing of Distributed Battery Energy Storage System,” *IEEE Trans. Ind. Electron.*, vol. 62, no. 5, pp. 2764–2776, May 2015.

[15] L. Solero, A. Lidozzi, and J. A. Pomilio, “Design of Multiple-Input Power Converter for Hybrid Vehicles,” *IEEE Trans. Power Electron.*, vol. 20, no. 5, pp. 1007–1016, Sep. 2005.

[16] A. S. Samosir and A. H. M. Yatim, “Implementation of Dynamic Evolution Control of Bidirectional DC–DC Converter for Interfacing Ultracapacitor Energy Storage to Fuel-Cell System,” *IEEE Trans. Ind. Electron.*, vol. 57, no. 10, pp. 3468–3473, Oct. 2010.

[17] N. Mukherjee and D. Strickland, “Control of Cascaded DC–DC Converter-Based Hybrid Battery Energy Storage Systems—Part II: Lyapunov Approach,” *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3050–3059, 2016.

[18] F. Altaf, B. Egardt, and L. Johannesson Mardh, “Load Management of Modular Battery Using Model Predictive Control: Thermal and State-of-Charge Balancing,” *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 1, pp. 47–62, Jan. 2017.

[19] D. Montesinos-Miracle, M. Massot-Campos, J. Bergas-Jane, S. Galceran-Arellano, and A. Rufer, “Design and Control of a Modular Multilevel DC/DC Converter for Regenerative Applications,” *IEEE Trans. Power Electron.*, vol. 28, no. 8, pp. 3970–3979, Aug. 2013.

[20] S. Shili, “Contrôle des circuits d’équilibrage des systèmes de stockage d’énergie (supercondensateurs) en vue d’estimer et d’améliorer leur durée de vie,” PhD thesis, Université Claude Bernard Lyon 1, Université de Lyon, Villeurbanne, France, 2016.

[21] S. Shili, A. Hijazi, A. Sari, X. Lin-Shi, and P. Venet, “Balancing Circuit New Control for Supercapacitor Storage System Lifetime Maximization,” *IEEE Trans. Power Electron.*, vol. 32, no. 6, pp. 4939–4948, Jun. 2017.

[22] S. J. Moura, “Techniques for battery health conscious power management via electrochemical modeling and optimal control,” PhD dissertation, The University of Michigan, Ann Arbor, MI, USA, 2011.

[23] P. Kreczanik, P. Venet, A. Hijazi, and G. Clerc, “Study of Supercapacitor Aging and Lifetime Estimation According to Voltage, Temperature, and RMS Current,” *IEEE Trans. Ind. Electron.*, vol. 61, no. 9, pp. 4895–4902, Sep. 2014.

[24] “Life Duration Estimation,” Maxwell Technologies, Inc., Application Note.

[25] A. Khelassi, D. Theilliol, and P. Weber, “Control design for over-actuated systems based on reliability indicators,” 2010, pp. 536–541.

[26] A. Khelassi, J. Jiang, D. Theilliol, P. Weber, and Y. M. Zhang, “Reconfiguration of Control Inputs for overactuated Systems based on Actuators health,” *IFAC Proc. Vol.*, vol. 44, no. 1, pp. 13729–13734, Jan. 2011.

[27] M. H. Rashid, Ed., *Power electronics handbook: devices, circuits, and applications handbook*, 3rd ed. Burlington, MA, USA: Elsevier, 2011.

[28] R. W. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*. New York, USA: Springer, 2001.

Université de Lyon
CNRS, Ecole Centrale Lyon, INSA Lyon, Université Claude Bernard Lyon 1

Laboratoire Ampère
Unité Mixte de Recherche du CNRS - UMR 5005
Génie Electrique, Automatique, Bio-ingénierie

Mémoire doctorant 1ère année
2017 -2018

Nom - Prénom	**SAGGIN Fabrício**
email	**fabricio.saggin@ec-lyon.fr**
Titre de la thèse	**Robust control for MEMS inertial sensors**
Directeur de thèse	**BOMBOIS Xavier et SCORLETTI Gérard**
Co- encadrants	**KORNIIENKO Anton**
Dpt. de rattachement	**MIS**
Date début des travaux	**01/10/2017**
Type de financement	**Contrat de recherche Ecole Centrale de Lyon (projet NEXT4MEMS – BPI France)**

Laboratoire Ampère – Ecole Centrale de Lyon – 36, avenue Guy de Collongue - 69134 Ecully cedex – France
Tél : +33 (0) 4 72 18 60 99 Fax : +33 (0) 4 78 43 37 17 http://www.ampere-lab.fr

Robust Control of MEMS Inertial Sensors

Fabrício Saggin

July 2, 2018

Abstract

Micro-electro-mechanical systems (MEMS) inertial sensors are micromachined devices that have been widely employed to measure motion and rotation in numerous applications. This kind of devices has the advantage to be easily integrated into electronic devices, cheap and low power consuming. However, compared to traditional mechanical inertial sensors, MEMS sensors have a degraded precision. This issue is mainly due to fabrication imperfections and sensitivity to parameters and environmental variations. In this context, this PhD project strives to contribute to the development of MEMS inertial sensors performance by the use of robust control techniques. Through this document, a literature review on the inertial sensors model(s) and the modes of operation are presented, discussing the advantages, inconvenient and possible improvements. Then, some approaches based on the celebrated $\mathcal{H}_\infty$ synthesis are proposed and preliminary practical results are presented. In addition, an amplitude envelope modeling is proposed. Perspectives of the future work are drawn at the end of this document.

Keywords: *MEMS inertial sensors, gyroscope, robust control, $\mathcal{H}_\infty$ synthesis, amplitude envelope modeling.*

Contents

1 Introduction

This report describes the work realised since October 2017 on the PhD thesis entitled "Robust Control of MEMS Inertial Sensors". This thesis is advised by Xavier Bombois, Gérard Scorletti and Anton Korniienko, and it is realised within MIS (*Méthodes pour l'Ingénierie des Systèmes*) department of Laboratoire Ampère, at Ecole Centrale de Lyon.

This research work is part of the project NEXT4MEMS, which main objective is to develop a new generation of high-performance MEMS inertial sensors. The consortium is composed by French leaders of the inertial sensors industry, and French laboratories. For further details about the project, refer to [Lab18]. The funding is granted by BPI France through *Projets de Recherche et Développement Structurants pour la Compétitivité* (PSCS).

Inertial sensors make use of the effects of the inertial forces over an object to determine its dynamic comportment. This characteristic can be directly exploited when the short-term dynamic changes of an object are of interest, as in crash-detection systems or vibrational measurements, for instance. Nonetheless, there are numerous applications, such as navigation, where inertial sensors are employed to determine the position or the orientation of an object. In these cases, the measured inertial forces have to be integrated to estimate the position/orientation of the object, what points to the necessity of high-accuracy sensors [Kem11].

Conventional high-performance inertial sensors often are too large, too heavy and/or too expensive for use in emerging applications. However, advances in the fabrication of micro-electro-mechanical systems (MEMS) allowed the miniaturization of inertial sensors. These miniaturized sensors have the advantage to be easily integrated into electronic devices, cheap and low power consuming. Nevertheless, due to fabrication imperfections, MEMS sensors have a degraded accuracy, compared to the conventional ones. Nowadays, a lot of effort has been dispensed to increase the performance of the micromachined ones, allowing them to be applied where high accuracy is required.

Inertial sensors can be classified as accelerometers if they are sensitive to linear forces, or gyroscopes if they are sensitive to rotational forces. Despite their differences, MEMS accelerometers and MEMS gyroscopes have similar working principles. Furthermore, since gyroscopes have a more challenging architecture and integrate most of the particularities of accelerometers, this document will be focused on the gyroscopes.

A MEMS gyroscope is composed of a proof mass that is suspended over a substrate by silicon springs, which allows the mass to move along two perpendicular axes: the primary (or x-) axis, and the secondary (or y-) axis. Electrostatic actuators are used to drive oscillations along the x-axis. If the device is submitted to an angular speed Ω_z perpendicular to the previous axes, a Coriolis force[1] appears along the secondary axis. This force is proportional to Ω_z and to the speed of the proof mass along the primary axis. Then, if the position along the x-axis can be described by a sinusoidal function over time, its derivative is also sinusoidal. Thus, the Coriolis force can be seen as the modulation of the angular rate with the derivative of the x-position. Hence, by estimating and demodulating the Coriolis force, one is able to recover the angular rate Ω_z. Nevertheless, MEMS devices are highly sensitive to fabrication imperfections and to environmental changes, what can quickly degrade their performance. To cope with these issue, control loops are implemented, considerably increasing the performance of the device [Sau08, OAL+05, You11].

In the literature, several methods are proposed to increase the performance of MEMS gyroscopes. However, these methods normally present at least one of the three limitations below:

[1] The principle that underlies the working principle of the MEMS gyroscope is the same that underlies the working principle of the celebrated Foucault pendulum: the Coriolis force.

Insufficient performance: since some methods are based on ideal models, the performances are easily degraded in real conditions;

Lack of a rigorous formulation: even presenting good performance, some methods do not have rigorous formulation. So, no formal guaranties of stability or performance can be given;

Implementability: finally, some methods present a high potential to overcome the previous issues. Nonetheless, their practical implementation is not proven.

In order to achieve high performance, it is mandatory to have a precise model of the device. Then, with an adequate model, a high-performance controller can be designed. Finally, a fine performance analysis shall be realized to verify if the required performance specifications are achieved. Therefore, in the scope of the project NEXT4MEMS, other thesis are being developed concomitantly:

- "Joint Identification and Control of MEMS Sensors" by Kévin Colin [Col18], where identification methods for control are being investigated; and
- "Performance Validation of MEMS Sensors Using Nonlinear Uncertain Models" by Jorge Ivan Ayala Cuevas [Cue18], who aims to develop analysis methods to verify the performance of the whole system in real conditions.

The main goal of this thesis is to develop a control design method to be integrated into the industrial MEMS sensors design flow. It has to ensure a desired performance level considering not only the nonidealities of the MEMS sensors, but also the electronic circuitry, parameter/environmental variations, technological dispersion, etc. Furthermore, it shall be an easy-to-use tool for engineers with no advanced knowledge on control theory.

Since multiple control loops compose the system, classic methods fail to ensure performance and stability of the closed-loop system. Then, robust control is considered, allowing one to take into account the stability and performance specifications of the system since the beginning of the controller design.

This document is structured as follows. Section 2, presents the research problem that guided the work realized during this first year. Section 3 presents a literature review on the gyroscope model, as well as on different approaches for the device operation, leading to the proposition of the main development lines. Section 4 presents the first line of research. A controller design based on the specifications of the sensor is described, and practical results are presented and discussed. Section 5 presents the second development line, introducing the envelope modeling. Section 6 draws the perspectives and the expected contributions of this thesis. Finally, Section 7 concludes this document.

In addition, details about the specifications of the gyroscope are provided in Appendix.

2 Research Problem

The main objective of this research work is to provide a computational tool to, based on specifications, systematically design high performance controllers for MEMS inertial sensors. It shall present the following features:

Ease of use: the proposed method should be accessible for an engineer with no strong background in advanced control tools;

Generality: the method must be valid for different devices (accelerometers and gyroscopes) and for different desired performances.

Calculability: the calculation time shall remain reasonable even for high performance specifications;

Implementability: the controller issued from the proposed method has to be implementable in the embedded processor of the sensor.

Besides the main objective, the particular control objectives with respect to the MEMS gyroscopes have to be considered. The principal goal of the primary (or drive) mode is to keep the proof mass oscillating, so in the presence of an angular rate, a Coriolis force will be created in the secondary mode. Furthermore, if one can provoke oscillations exactly at the resonance frequency, the best trade-off between the dispensed power and the vibrations amplitude is achieved. However, the performance of the system can be sensibly degraded due to temperature fluctuations, coupling effects, environmental variations or aging. So, the problem of the **primary mode** can be stated as

To implement a mechanism capable of oscillating the proof mass at the resonance frequency in a reasonable time with a controlled amplitude despite the variations of the environment or the device parameters.

The secondary (or sense) mode is the element that is sensitive to the Coriolis force, allowing one to estimate the angular rate. Similarly to the primary mode, imperfections are presented and can also degrade the performance of the sensor. Hence, the problem of the **secondary mode** can be formulated as

To implement a mechanism capable of detecting the Coriolis force acting on the y-axis despite the parasitic coupling forces, the variations of the environment or the device parameters.

Evidently, the implementation of the aforementioned mechanisms includes the choice of a control architecture, the design of a controller that ensures some performance specifications and its implementation in the embedded processor. It is important to highlight that the design of the inertial sensors and of the electronic circuitry is not in the scope of this thesis.

After having defined the research problems, it is necessary to define what are the performance specifications of the devices, as well as the experimental setup where the controllers shall be implemented.

2.1 Performance Specifications

For the production of a MEMS gyroscope, an exhaustive list of performance specifications has to be verified. However, only some of them concern the control design. These are as follows:

Start-up time: is the time the sensor gets to be operative. This characteristic is mainly related to the settling time (of the oscillations) of the primary mode;

Scale factor: is the ratio between the real angular rate and its measure;

Bandwidth: from the measurement point of view, a gyroscope works as a low-pass filter, filtering the real angular rate Ω_z and producing the measure $\tilde{\Omega}_z$. Then, the sensors bandwidth is maximum frequency for which the attenuation of the measure is not bigger than 3 dB; and

Latency: is the time-delay between the stimulation (applied angular rate) and the measure of this stimulus;

The bandwidth and the scale factor can be directly related to the control loops performance specifications, as detailed in Appendix A. The latency and the bandwidth are also related to the low-pass filter present in the synchronous demodulation.

2.2 Experimental Setup

The experimental tests shall be realized on the AS3125-SDK (Sensor Development Kit) development board, conceived by ASYGN. This development kit is mainly composed by:

SoC-class FPGA: field-programmable gate arrays (FPGA) are integrated circuits (IC) whose hardware is programmable. However, in addition to this feature, System-on-Chip-class (SoC-class) FPGAs also contain an integrated processor. Hence, both hardware and software are programmable, giving a high flexibility to the development board. The control loops are implemented via a code in C++ language;

AS3125 (Idyle): this ASIC (application-specific integrated circuit), is dedicated to the actuation and the measuring of MEMS sensors, making the interface between the microdevices and the FPGA. It has different mixers (modulators) that allow to modulate the actuation signal as well as the measurement signal;

Peripheral devices: as power supplies, numerous analog inputs and outputs, USB interface, auxiliary analog-to-digital converters (ADC), etc; and

MEMS Sensors: the electronic board has connectors where different MEMS sensors can be connected, as gyroscopes and accelerometers.

The interface with a computer is realized via USB, and dedicated Python library allows one to communicate with the development kit and to configure the Idyle IC. Thus, for practical experiments, a C++ code with the control loops has to be written, and a Python script has to be composed.

We emphasize that the Idyle IC allows to implement two different strategies: one with modulators for actuation and measurement, and another without modulators (a direct actuation and sensing). Both possibilities will be explored.

3 MEMS Gyroscopes: Model and Operation

A MEMS gyroscope is composed of a proof mass attached to a silicium structure by silicon springs. A simplified representation of its mechanical structure is presented in Figure 1. Electrostatic actuators are used to make the proof mass oscillate along the x- and y-directions [Sau08, Kem11]. Therefore,

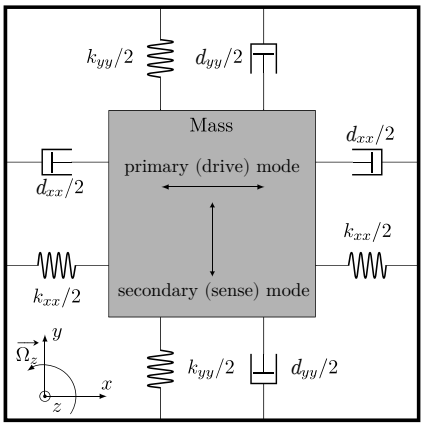


Figure 1: Simplified mechanical representation of a MEMS gyroscope.

MEMS gyroscopes are lightly damped mass-spring systems, whose dynamic behavior can be described by the differential equation below:

$$\begin{bmatrix} m_x & 0 \\ 0 & m_y \end{bmatrix} \begin{bmatrix} \ddot{x}(t) \\ \ddot{y}(t) \end{bmatrix} + \begin{bmatrix} d_{xx} & 0 \\ 0 & d_{yy} \end{bmatrix} \begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \end{bmatrix} + \begin{bmatrix} k_{xx} & 0 \\ 0 & k_{yy} \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} F_x(t) \\ F_y(t) \end{bmatrix}, \tag{1}$$

where

x and y are the drive (primary) mode and the sense (secondary) mode **positions**, respectively;

m_x and m_y are the drive mode and the sense mode **masses**, respectively [2];

F_x and F_y are the drive mode and the sense mode **applied forces**, respectively;

d_{xx} and d_{yy} are the **damping** in the x and the y directions, respectively; and

k_{xx} and k_{yy} are the respective **stiffness** coefficients.

The quality factors Q and the resonance frequencies ω_0 of each mode, identified by the indices x and y, can be cast as

$$\omega_{0,i} = \sqrt{\frac{k_{ii}}{m_i}} \quad \text{and} \quad Q_i = \frac{\sqrt{k_{ii} m_i}}{d_{ii}},$$

with $i = \{x, y\}$. This notation is very useful when dealing with the transfer functions G_i of each mode (x and y), that can be given by

$$G_i(s) = \frac{i(s)}{F_i(s)} = \frac{1/m_i}{s^2 + \omega_{0,i}/Q_i \cdot s + \omega_{0,i}^2}, \tag{2}$$

where s stands for the Laplace variable.

In the presence of an **angular rate** $\Omega_z(t)$ (in $\mathrm{rad\,s^{-1}}$) perpendicular to the xy-plane, a Coriolis force appears. This force is given by

$$\mathbf{F_{Cor}}(t) = \begin{bmatrix} F_{Cor,x}(t) \\ F_{Cor,y}(t) \end{bmatrix} = \Omega_z(t) \begin{bmatrix} 0 & 2m_y \\ -2m_x & 0 \end{bmatrix} \begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \end{bmatrix}. \tag{3}$$

[2] For the sake of completeness, two different masses are considered, as in [OAL+05]. This allows to reduce the influence of the sense mode over the drive mode as well as the mechanical coupling between them [Kem11].

Then, by including the Coriolis force into (1), the following expression is obtained, dropping the time dependence:

$$\begin{bmatrix} m_x & 0 \\ 0 & m_y \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} + \begin{bmatrix} d_{xx} & 0 \\ 0 & d_{yy} \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} + \begin{bmatrix} k_{xx} & 0 \\ 0 & k_{yy} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \end{bmatrix} + \Omega_z \begin{bmatrix} 0 & 2m_y \\ -2m_x & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}. \tag{4}$$

or, in a compact way,

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{D}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{F} + \mathbf{F_{cor}}, \tag{5}$$

with $\mathbf{M}$, $\mathbf{D}$ and $\mathbf{K}$ being respectively the masses, the damping and the stiffness matrices, and $\mathbf{q} = [x\ y]^T$.

The operating principle of a MEMS vibratory gyroscope can be understood from the previous equation. Let us consider a force $F_x(t)$ that provokes oscillations in the drive mass such that

$$x(t) = A_x \sin(\omega_{exc} t), \tag{6}$$

where A_x is the amplitude of the oscillations and ω_{exc} is the frequency of the excitation (in $\mathrm{rad\,s^{-1}}$). Since the matrices $\mathbf{M}$, $\mathbf{D}$ and $\mathbf{K}$ are diagonal, if $\Omega_z(t) \equiv 0$ and $F_y(t) \equiv 0$, no oscillations will be produced in the y-direction. However, if $\Omega_z(t) \neq 0$, a Coriolis force ($F_{cor,y} = -2m_x\Omega_z\dot{x}$) is applied in the y-direction. This force can be given by

$$F_{cor,y}(t) = -2A_x m_x \omega_{exc} \Omega_z(t) \cos(\omega_{exc} t). \tag{7}$$

Then, the position $y(t)$ can be described by

$$y(t) = -A_y(t) \cos(\omega_{exc} t + \varphi_y), \tag{8}$$

with amplitude $A_y(t) = 2m_x A_x \omega_{exc} \Omega_z(t) |G_y(j\omega_{exc})|$ and phase-delay $\varphi_y = \angle G_y(j\omega_{exc})$. Hence, by demodulating $y(t)$, $\Omega_z(t)$ can be estimated.

However, fabrication imperfections give origin to anisoelasticity and nonproportional damping effects, mechanically coupling the two modes of the gyroscope [Sau08, You11]. This mechanical coupling appears as nondiagonal terms in the matrices $\mathbf{D}$ and $\mathbf{K}$. Thus, the dynamic behavior of the device can then be described by the equation below:

$$\begin{bmatrix} m_x & 0 \\ 0 & m_y \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} + \begin{bmatrix} d_{xx} & d_{xy} \\ d_{yx} & d_{yy} \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} + \begin{bmatrix} k_{xx} & k_{xy} \\ k_{yx} & k_{yy} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \end{bmatrix} + \Omega_z \begin{bmatrix} 0 & 2m_y \\ -2m_x & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}. \tag{9}$$

In this new framework, the position $y(t)$ is rewritten as

$$y(t) = y_c(t) \cos(\omega_{exc} t + \varphi_y) + y_s(t) \sin(\omega_{exc} t + \varphi_y), \tag{10}$$

with

$$y_c(t) = -A_x \omega_{exc} |G_y(j\omega_{exc})| (2m_x \Omega_z + d_{yx}) \tag{11}$$

$$y_s(t) = -A_x |G_y(j\omega_{exc})| k_{yx}. \tag{12}$$

In equation (10), we can notice that y_c and y_s are modulated and in quadrature (one is multiplied by a sinus and the other one by a cosinus of same argument). Then, a synchronous demodulation can be used to recover y_c and y_s [PPK+04, OAL+05, Sau08]. A block diagram of the synchronous demodulation is presented in Figure 2, where the blocks $F_{LP}(s)$ represent low-pass filters, and $\tilde{y_c}$ and

$\tilde{y_s}$ are the estimates of y_c and y_s, respectively. Basically, by multiplying $y(t)$ by $\cos(\omega_{exc}t + \varphi_y)$ [3], a constant term proportional to y_c and terms at $2\omega_{exc}$ appear. The low-pass filter attenuate this terms at $2\omega_{exc}$, giving an estimate of y_c. The same reasoning is valid for y_s, but with a sinus instead a cosinus function.

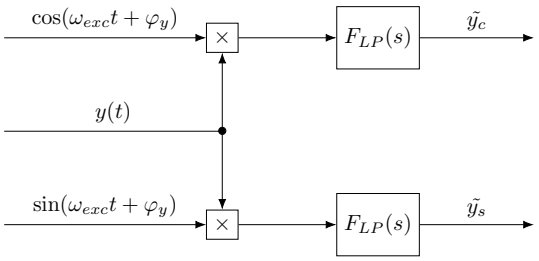


Figure 2: Block diagram of a synchronous demodulation.

In general, for the operation of a MEMS gyroscope, three main elements can be distinguished: the drive mode, the sense mode, and the synchronous demodulation or estimation of the angular rate. In the sequel of this section, a literature review on the different operation modes of each element is presented.

3.1 Control of the Primary Mode

We recall that the drive mode has to oscillate with controlled amplitude with a frequency close to the resonance frequency. Thus, two main control objectives can be cast: the amplitude control and the resonance frequency tracking. In the literature, different strategies are found to cope with this problem. The most common approaches are presented in the sequel, their advantages and drawbacks are also discussed.

Self-oscillating Automatic Gain Control loop: [MVG01] proposes a simple self-oscillating automatic gain control (AGC) loop, as shown in Figure 3, where the block AGC represents an envelope detection (of the velocity) and a PI controller. The output of this block is then multiplied by the derivative of the position x. In steady-state, the AGC block would produce a constant signal, such that the closed-loop system would behave as a pure oscillator. This approach is considered by numerous authors, as in [CMTB05, SHLK07, CYC+14]. In addition to be a relatively simple structure, this approach would ensure at the same time the amplitude control and the resonance frequency tracking, in the case of parameters or environmental variations. However, the nonlinearities present in the loop make it difficult to rigorously analyse the performance of the system. [MVG01] makes use of averaging methods to analyse the equilibrium points and their stability, but the regions of attraction and the effects of noise are not discussed. [SHLK07] proposes an envelope model of the gyroscope as an attempt to evaluate the performance of the closed-loop system. However, no information about the frequency is carried during the analysis.

Envelope-based Control: Since MEMS gyroscopes are resonator with a high quality-factor, even for resonance frequencies from the order of 10 kHz, they can take up to 10 s to achieve the sinusoidal steady-state. Thus, some authors, as [BCRM02, BRCD05, EMK12], propose a model for the evolution of the amplitude of x (A_x) instead of the evolution of x. Based on this envelope

[3] For the sake of simplicity, we consider φ_y known. However, if this phase-delay is unknown, y_c and y_s will be a combination of equations (11) and (12). For further details, refer to [Sau08].

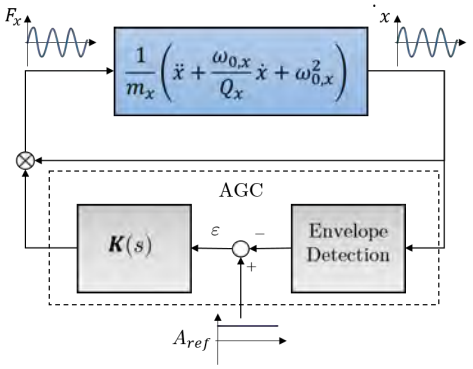


Figure 3: AGC loop proposed by [MVG01].

model, a PI controller can be tuned to ensure the amplitude control. However, in order to make available the measure of A_x instead of x, nonlinear functions (multiplication by a sinus and envelope detection) are included into the loop, as shown in Figure 4.

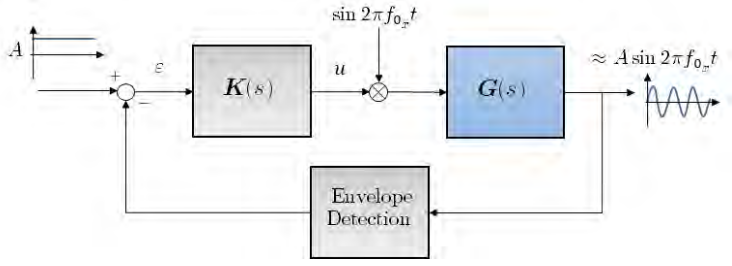


Figure 4: Block diagram of the envelope-based approach.

Resonance frequency tracking with a PLL: Some authors, as [KCV08] and [RCRW09], propose the use of a phase-locked loop (PLL) to track the resonance frequency. Since an ideal second-order resonator has a $-90°$ phase-delay at the resonance frequency, by imposing this phase-delay with a PLL, one would be able to track the resonance frequency of a gyroscope. Here, the issue resides in the fact that the electronic circuits add a phase-delay. So, the phase-delay at the resonance frequency depends also on the phase-delay included by electronic circuits. The problem is that the electronic phase-delay is unknown.

In order to control the amplitude, an additional control loop is required. This additional loop is also based on the envelope dynamics.

Instantaneous position (or direct) approach: It has to be stressed that in the previous approaches, the controller does not work with the instantaneous position (or speed) of the proof mass. The presented strategies use demodulation or envelope detection of the measured signals. From the point of view of the control theory, it is not intuitive to add nonlinear functions (such as the modulation/demodulation or the AGC loop) into the control loop to achieve a desired behaviour. The major issue is that the inclusion of these nonlinearities makes harder to formally ensure stability or performance of the closed-loop system. Thus, some authors have been proposing the use of classical control strategies applied to gyroscopes. Here, the instantaneous position is considered, as presented in Figure 5. For instance, sliding mode control (SMC) is used in [BSK06, Fei10], model predictive control (MPC) was applied by [PK14], and adaptive control was presented in [SHS+99, PH04, PHHN07].

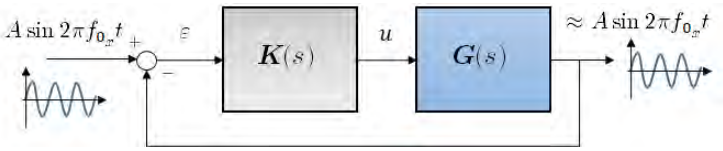


Figure 5: Block diagram of the instantaneous position (or direct) approach.

3.2 Control of the Sense Mode

Here, two approaches are very popular [OAL+05]: the open-loop and the closed-loop operating modes. The former one consists on the description given in the beginning of this section, where the oscillations in y are measured and demodulated, then the angular rate can be determined [SHS+99, LR02, SAH07, LLC+11]. This strategy is the simplest one, since it is based only on the measurements of the position y and posterior demodulation. However, the dynamics of the sense mode depends exclusively on its mechanical characteristics, which are often very slow due to its high-quality factor. This operating mode is suited for low-cost and low-performance applications.

In the other side, the closed-loop operating mode, also known as force-to-rebalance loop, consists on applying a force F_y in order to keep $y(t) \equiv 0$. In this case, it can be shown that

$$F_y(t) = F_{cor,y}(t) - d_{yx}\dot{x}(t) - k_{yx}x(t). \tag{13}$$

Thus, by demodulating F_y, the angular rate can be calculated [LR02, Sau08]. From the control theory point of view, this problem can be seen as a classical disturbance rejection problem, as illustrated in Figure 6. In comparison to the open-loop operating mode, in this case, electrostatic actuators are required to impose F_y. Nonetheless, since the closed-loop dynamics depends on the dynamics of the sense mode and on the dynamics of the controller, this approach allows one to have a larger bandwidth. Besides the dynamics improvements, since there is no movements on the sense mode ($y(t) \approx 0$, in the real case), there exist no coupling from the sense mode to the drive mode.

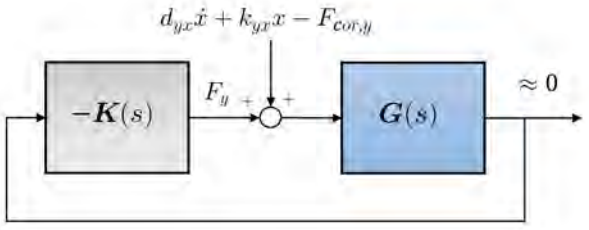


Figure 6: Block diagram of the sense mode control loop.

To cope with the sense mode problem, different methods are proposed in the literature. The first (and natural) candidate to ensure the disturbance rejection is the proportional–integral–derivative (PID) controller [PPK+04, SKL08], which is an easy to implement solution. However, its performance is not optimal. Then, some authors propose a linear quadratic Gaussian (LQG) [LWS10] or an $\mathcal{H}_\infty$ controller design [SLSK04, HZH+15], that can optimize the performance (and stability robustness) of the closed-loop sense mode. We stress that these strategies are similar to the direct approach used for the drive mode.

3.3 Estimation of the Angular Rate

Most of the commercially produced gyroscopes estimate the angular rate by means of the synchronous demodulation (of y in an open-loop operation, or of F_y in a closed-loop operation), as described earlier. Nevertheless, an increasing interest on the development of alternative solutions is found in the literature. These new solutions are mostly based on adaptive control, filtering and estimation [PH04, ZDG07, FY11, KS17]. Although these approaches appear to be promising, only simulation results are presented.

3.4 Approaches Considered for This Thesis

In this work, different approaches shall be considered and evaluated. They can be grouped into two main development lines:

Line 1: This main axis relies on the direct approach for drive and sense modes. A series of $\mathcal{H}_\infty$-based control methods will be investigated, starting from the simplest architecture, composed by two independent SISO controllers (one for each mode), passing by a MIMO framework, where both modes are controlled simultaneously by the same multivariable controller, and leading to an LPV approach, where the controller parameters can be adapted according to the gyroscope variations; and

Line 2: This alternative axis resides on the envelope-based approaches. A better description and a better understanding of the assumptions behind the envelope models can allow us to exploit more possibilities of this celebrated approach.

4 Direct Control Approach for MEMS Gyroscopes

This section presents the first research line of this thesis. Here, we justify the use of advanced control techniques, and we propose some strategies for the direct control of primary and secondary modes of a MEMS gyroscope.

Since the control problem of the gyroscope is similar to the vibration control problems (generation and attenuation of oscillations), classical frequency-domain design methods would be the first candidates to be considered. However, since multiple control loops compose the system, classic methods fail to ensure performance and stability of the closed-loop system. Hence, advanced frequency-domain controller design methods shall be considered, allowing one to take into account the performance of the system since the beginning of the controller design stage.

The $\mathcal{H}_\infty$ synthesis allows one to pose a controller design problem as a convex optimization problem subject to mathematical constraints. These constraints can be used to impose performance criteria as well as stability margins, for example. Then, due to these features, the $\mathcal{H}_\infty$ synthesis is a natural choice to the control design. Another advantage of the $\mathcal{H}_\infty$ synthesis is the flexibility. The same method used in a single-input single-output (SISO) framework can be extended to a multiple-input multiple-output (MIMO) one, and latter, to a linear parameter-varying (LPV) framework.

Firstly, let us consider a SISO framework with the closed-loop system of Figure 7, where $K(s)$ represents a controller, $G(s)$ stands for a plant to be controlled, r is the reference signal, ε is the error between the reference and the output signals y, u is the control signal, and d is an input disturbance[4].

[4]Without loss of generality, since an output disturbance d_o can be transposed to the input as $d = G^{-1}d_o$.

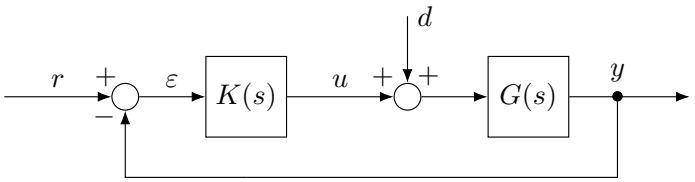


Figure 7: Block diagram of a closed-loop system.

The transfer functions of this closed-loop system are:

$$T_{r\to\varepsilon}(s) = \frac{1}{1+K(s)G(s)} = S(s) \tag{14}$$

$$T_{r\to u}(s) = \frac{K(s)}{1+K(s)G(s)} = K(s)S(s) \tag{15}$$

$$T_{d\to\varepsilon}(s) = -\frac{G(s)}{1+K(s)G(s)} = -G(s)S(s) \tag{16}$$

$$T_{d\to u}(s) = -\frac{K(s)G(s)}{1+K(s)G(s)} = -1 + S(s) = -T(s), \tag{17}$$

where $S(s)$ is the sensitivity function, and $T(s)$ is the complementary sensitivity function.

It required that this system shall satisfy some requirement specifications, such as stability, reasonable response time, a maximum allowed steady-state error, etc. These requirements can also be given as constraints on the closed-loop frequency responses, as follows:

Stability/Robustness: Further than the stability, it is desired to ensure a convenient stability margin, ensuring the robust stability of the system. A typical choice is to choose a modulus margin equal to 6 dB. The modulus margin is equivalent to $\|S(s)\|_\infty$. So, by restricting the maximal value of the sensitivity function, a minimal modulus margin can be imposed;

Reference Tracking This is a steady-state characteristic related to the error signal ε, or to the sensitivity function S. If the sensitivity function presents as zeros the modes (poles) of the reference signal, the error goes to zero at steady-state;

Response Time Complementary to the reference tracking, the response time is a transient characteristic linked to the error signal ε. It is related to the bandwidth of T or S;

Disturbance Rejection This characteristic is related to the capability of the controller to compensate exogenous disturbance. Similarly to the reference tracking, the transfer function $T_{d\to\varepsilon}$ must present as zeros the modes (poles) of the disturbance signal to ensure the disturbance rejection;

Control Effort It is desired the control effort to be as small as possible. Then, the transfer functions $T_{r\to u}$ and $T_{d\to u}$ shall not present high gains. Moreover, by shaping the gains of $T_{r\to u}$, some uncertainties of the plant can be taken into account; and

Noise Amplification The measurement noise is not explicitly presented in Figure 7. However, it can be shown that the transfer function from measurement noise to the error or to the control signal are equal (in modulus) to those from the reference.

Some of the previous constraints are contradictory. For instance, if we want to reduce the control effort, we would constraint the $T_{r\to u}$ and $T_{d\to u}$ transfer functions to be as small as possible. However,

$T_{d\to u} = T$ and $T = 1 - S$. By doing so, a constraint on S (and related specifications, as reference tracking) is indirectly imposed. Then, a trade-off between the desired contradictory specifications has to be found.

In order to shape the frequency response of the closed-loop system, filters can be used to weight the performance criteria. These filters are called weighting functions or weighting filters. Figure 8 presents the celebrated four-block criterion[5], where the signals w_1 and w_2, and z_1 and z_2 represent exogenous inputs and regulated outputs, respectively. These signals are not present in the real system, they are used only to design the controller.

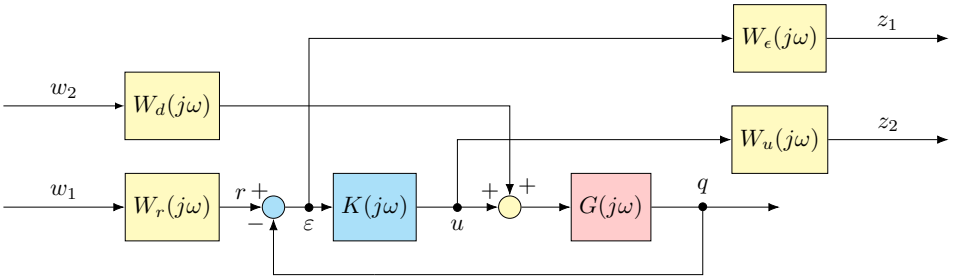


Figure 8: Four-block criterion.

The $\mathcal{H}_\infty$-problem can therefore be stated as to find $K(s)$ such that

$$\|\mathbf{T}_{\mathbf{w}\to\mathbf{z}}(s)\|_\infty = \left\| \begin{matrix} W_\epsilon(s)S(s)W_r(s) & W_\epsilon(s)G(s)S(s)W_d(s) \\ W_u(s)K(s)S(s)W_r(s) & W_u(s)T(s)W_d(s) \end{matrix} \right\|_\infty < \gamma. \tag{18}$$

Moreover, if $\gamma \leq 1$, then $\forall\omega \in \mathbb{R}$,

$$|S(j\omega)| < \frac{1}{|W_\epsilon(j\omega)W_r(j\omega)|}, \quad |G(j\omega)S(j\omega)| < \frac{1}{|W_\epsilon(j\omega)W_d(j\omega)|},$$

$$|K(j\omega)S(j\omega)| < \frac{1}{|W_u(j\omega)W_r(j\omega)|}, \quad \text{and} \quad |T(j\omega)| < \frac{1}{|W_u(j\omega)W_d(j\omega)|}.$$

In other words, the correct choice of the weighting functions allows one to shape the closed-loop transfer functions such that the requirement specifications are achieved.

It is important to highlight that the order of the controller is equal to the order of the system $G(s)$ added by the order of all the weighting functions. So, to obtain a controller with a reasonable order, the weighting functions have to be as simple (low order) as possible. Even though, if the resulting controller has a high order, model reduction can be performed [SP01, SF09].

Next, we will use the $\mathcal{H}_\infty$ synthesis to design the controller for the gyroscope. We will consider the direct approach for the drive mode, and the close-loop operation for the sense mode. In a first moment, the two modes of the gyroscope are considered to be uncoupled, and two SISO controllers are designed. Latter, extensions to the MIMO case are explored.

[5] Here, we consider the four-block criterion, that is rich enough to take into account the specifications. However, the flexibility of the $\mathcal{H}_\infty$ synthesis allows one to increase the richness of the criterion, as we be presented in the sequel.

4.1 Primary Mode - Reference Tracking

As discussed in Section 3.1 the control problem of the primary mode can be tackled as a reference tracking. Here, we consider a direct approach (see Figure 5), where the specifications for the controller design can be cast as follows:

1. Track a sinusoidal reference signal with frequency close to $\omega_{0,x}$;
2. Ensure the robust stability; and
3. Reduction of the control energy.

Considering the four-block criterion (Figure 8), the first specification can be achieved by constraining the sensitivity function to be small close to the resonance frequency. This restriction is made via the product $W_\epsilon W_r$, that has to be big at the resonance frequency. Moreover, this product can also be used to impose a modulus margin to the closed-loop system. The product $W_\epsilon W_r$ can be then written in the form

$$W_\epsilon(s)W_r(s) = k_e k_r \frac{s^2 + \alpha s + \omega_{min}\omega_{max}}{s^2 + \alpha\varepsilon_{max} s + \omega_{min}\omega_{max}}, \tag{19}$$

where

$$\alpha = \frac{\omega_{max} - \omega_{min}}{\overline{\varepsilon}} \sqrt{\frac{1 - \overline{\varepsilon}^2}{1 - \varepsilon_{max}^2}}. \tag{20}$$

For a fixed referente amplitude, this constraint imposes the reference tracking error to be lower than $\overline{\varepsilon}/(k_r k_e)$ for $\omega_{min} < \omega < \omega_{max}$, and lower than $\varepsilon_{max}/(k_r k_e)$ for $\omega = \sqrt{\omega_{min}\omega_{max}}$. The product $k_r \cdot k_e$ is used to impose the required modulus margin $M = 1/(k_r k_e)$ [SF09].

In order to reduce the control effort and to attenuate the measurement noise, the product $W_u(s)W_r(s)$ has to be small close to the resonance frequency, allowing $K(s)S(s)$ to be big, and big for low and high frequencies. So, the referred product has the form

$$W_u(s)W_r(s) = k_u k_r \frac{s^2 + \alpha u_{max} s + \omega_{min}\omega_{max}}{s^2 + \alpha s + \omega_{min}\omega_{max}}, \tag{21}$$

where

$$\alpha = \frac{\omega_{max} - \omega_{min}}{\overline{u}} \sqrt{\frac{1 - \overline{u}^2}{1 - u_{max}^2}}. \tag{22}$$

This constraint imposes the control signal to be lower than $\overline{u}/(k_u k_r)$ for $\omega_{min} < \omega < \omega_{max}$[6], and lower than $u_{max}/(k_u k_r)$ for $\omega = \sqrt{\omega_{min}\omega_{max}}$. The product $k_u \cdot k_r$ fixes the maximum gain in low and high frequencies. Further than the reduction of the control effort and the attenuation of the noise, this choice also improves the stability robustness of the closed-loop system to additive uncertanties of the model [SP01].

In addition to the design of the previous weighting functions, it is possible to avoid pole-zero compensation by constraining the $T_{d\to\epsilon}$ frequency response [SF09]. This can be done with the product $W_\epsilon(s)W_d(s)$. We choose here, $W_d(s) = k_d$.

A numerical example and practical implementation related to these choices are presented at the end of this section.

[6]The choice of ω_{min} and ω_{max} for $W_u(s)W_r(s)$ can be different from those of $W_\epsilon(s)W_r(s)$.

4.2 Secondary Mode - Disturbance Rejection

As discussed in Section 3.2, the control problem of the secondary mode can be tackled as an input disturbance rejection (see Figure 6). The specifications for the controller design are cast as follows:

1. Reject sinusoidal disturbances with frequency close to $\omega_{0,x}$;
2. Ensure the robust stability; and
3. Reduction of the noise.

Similarly to the primary mode, we also consider the four-block criterion (Figure 8). To ensure disturbance rejection of sinusoidal signals, the $T_{d\to\epsilon}$ transfer function has to be small close to the resonance frequency. This restriction is made via the product $W_\epsilon W_d$, that has to be big at the resonance frequency. Therefore, the product $W_\epsilon W_d$ can be given by

$$W_\epsilon(s)W_d(s) = k_e k_d \frac{s^2 + \alpha s + \omega_{min}\omega_{max}}{s^2 + \alpha\varepsilon_{max} s + \omega_{min}\omega_{max}}, \tag{23}$$

with α of equation (20). Moreover, as mentioned before, this product can also be used to avoid pole-zero compensation strategies.

To ensure a modulus margin M to the closed-loop system. The minimum value of the product $W_\epsilon W_r$ has to be bigger than $1/M$. Moreover, as in the drive mode case, the product $W_u(s)W_r(s)$ can be used to improve stability robustness.

Finally, the effects of the measurement noise can be attenuated be constraining the $T_{r\to u}$ transfer function at low and high frequencies, where the controller does not have to actuate.

4.3 Extensions to the MIMO case

In the literature, drive and sense modes normally are treated separately. Nonetheless, with the $\mathcal{H}_\infty$-synthesis, it is possible to design a single multivariable controller that controls at the same time the primary and secondary modes of the gyroscope while ensuring the robust stability and the nominal performance specifications. The specifications for the controller are:

1. the drive mode must track a sinusoidal reference with frequency close to ω_x;
2. the sense mode must keep $y(t) \simeq 0$;
3. reduction of the control energy and noise;
4. ensure the stability robustness.

In order to design the controller, the block diagram of Figure 9 is considered, where the plant $\mathbf{G}$ represents both primary and secondary modes of the gyroscope, and has two inputs (F_x and F_y) and two outputs (x and y). The weighting functions are :

$$\mathbf{W_r}(s) = \mathrm{diag}\big(W_{r1}(s), W_{r2}(s)\big),$$

$$\mathbf{W_d}(s) = \mathrm{diag}\big(W_{d1}(s), W_{d2}(s)\big),$$

$$\mathbf{W}_\epsilon(s) = \mathrm{diag}\big(W_{\epsilon 1}(s), W_{\epsilon 2}(s)\big),$$

$$\mathbf{W_u}(s) = \mathrm{diag}\big(W_{u1}(s), W_{u2}(s)\big)$$

where the indices 1 and 2 refer to the primary and secondary modes, respectively.

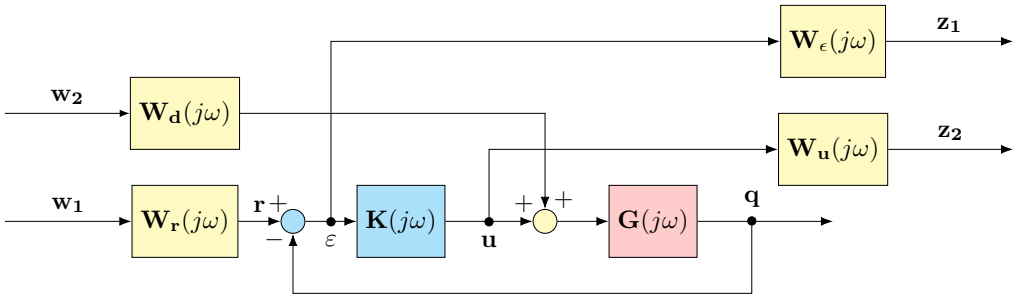


Figure 9: Four-block criterion for the MIMO framework.

The $\mathcal{H}_\infty$ problem can then be stated as to find $\mathbf{K}(s) \in \mathbb{C}^{2\times 2}$ such that[7]

$$\|\mathbf{T_{w\to z}}\|_\infty = \left\| \begin{matrix} \mathbf{W_e S W_r} & \mathbf{W_e S G W_d} \\ \mathbf{W_u K S W_r} & \mathbf{W_u T W_d} \end{matrix} \right\|_\infty < \gamma, \tag{24}$$

where $\mathbf{S} = (\mathbf{I}_2 + \mathbf{GK})^{-1}$ and $\mathbf{T} = \mathbf{I}_2 - \mathbf{S}$. Moreover, if $\gamma \leq 1$, then $\forall \omega \in \mathbb{R}$,

$$\overline{\sigma}(\mathbf{S}) < \frac{1}{\min\{|W_{\epsilon 1} W_{r1}|, |W_{\epsilon 2} W_{r2}|\}}, \quad \overline{\sigma}(\mathbf{GS}) < \frac{1}{\min\{|W_{\epsilon 1} W_{d1}|, |W_{\epsilon 2} W_{d2}|\}}$$

$$\overline{\sigma}(\mathbf{KS}) < \frac{1}{\min\{|W_{u1} W_{r1}|, |W_{u2} W_{r2}|\}}, \quad \overline{\sigma}(\mathbf{T}) < \frac{1}{\min\{|W_{u1} W_{d1}|, |W_{u2} W_{d2}|\}},$$

where $\overline{\sigma}(\mathbf{M})$ is the maximum singular value of $\mathbf{M}$.

The main difference in the MIMO case, with respect to the SISO framework, is that instead of regarding the frequency responses of the closed-loop system, now the singular values are taken into account. For instance, in order to impose a criterion similar to the modulus margin in the MIMO case, we could bound the sensitivity function $\mathbf{S}$, or its $\mathcal{H}_\infty$-norm, that is defined as

$$\|\mathbf{S}\|_\infty = \sup_{\omega \in \mathbb{R}} \overline{\sigma}(\mathbf{S}(j\omega)). \tag{25}$$

This can be done by the adequate choice of W_{r1}, W_{r2}, $W_{\epsilon 1}$ and $W_{\epsilon 2}$.

In the other hand, specific performance criteria (evaluated from a single input to a single output), as the reference tracking of the drive mode or the disturbance rejection of the sense mode, can be imposed by the choice of the individual combination of weighting functions. For example, to ensure the reference tracking, the product $W_{\epsilon 1} W_{r1}$ has to present the form of the equation (19), other weighting function are not directly related to this criterion. Similarly to this example, the choice of the other weighting functions to impose the other specifications is inspired on the SISO case.

4.4 Extensions to Joint Control and Estimation

The main objective of the sense mode is to estimate the Coriolis force. Normally, this is made by imposing $y(t) \simeq 0$, and therefore the equation (13) is verified. However, due to the flexibility of the $\mathcal{H}_\infty$ synthesis, we propose some modifications to directly estimate the Coriolis force, $F_{Cor,y}$. The new block-diagram is presented in Figure 10. For the synthesis, the model of the gyroscope $\mathbf{G}$ has

[7] For the sake of simplicity, the dependency on s is omitted.

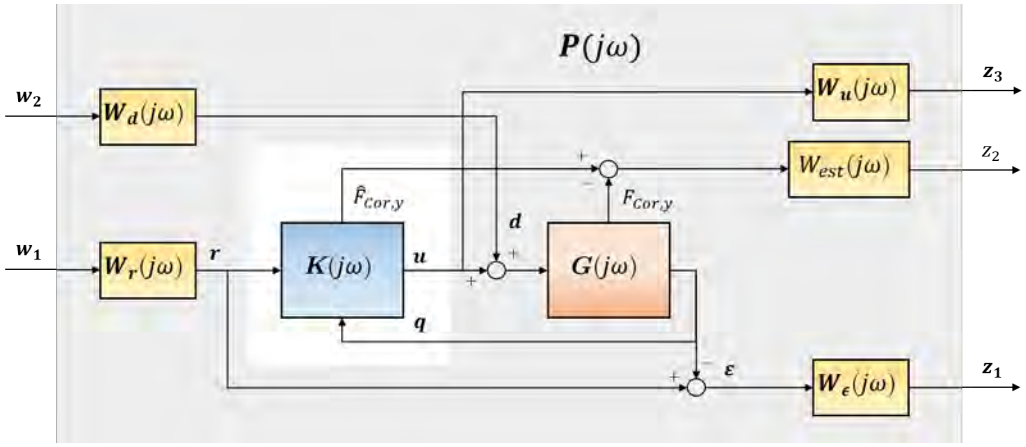


Figure 10: Four-block criterion for the MIMO framework.

an additional output, the Coriolis force $F_{Cor,y}$. The controller receives the reference signals and the measures of the drive and the sense modes, and calculates the control signal $\mathbf{u} = [F_x \ F_y]^T$ and an estimate of the Coriolis force, $\hat{F}_{Cor,y}$. The weighting function, $W_{est}(j\omega)$, is used to minimize the error between $F_{Cor,y}$ and $\hat{F}_{Cor,y}$, that can be used to calculate the angular rate of the gyroscope.

Similarly to the previous cases, the $\mathcal{H}_\infty$-problem can be stated as to find $\mathbf{K}(s)$ such that

$$\|\mathbf{T}_{\mathbf{w}\to\mathbf{z}}\|_\infty < \gamma, \tag{26}$$

with $\mathbf{w} = [\mathbf{w_1}^T \ \ \mathbf{w_2}^T]^T$ and $\mathbf{z} = [\mathbf{z_1}^T \ \ z_2 \ \ \mathbf{z_3}^T]^T$.

Variations on the angular rate can be modeled as an input disturbance. Then, in order to have an accurate estimation of the Coriolis force, the product $W_{est}W_{d2}$ has to present a high-gain at least close to the resonance frequency. So, it can be chosen with the same form presented in equation (21).

4.5 First Practical Results - Drive Mode

Before to run the first practical experiments, some modifications on the C++ code of the development kit were required in order to work with arbitrary actuation signals. Validations tests were performed to ensure the proper operation of the whole system. Then, the identification of the gyroscope was possible. For further details about the identification, refer to [Col18]. Figure 11 presents the Bode diagram of the drive mode of the gyroscope (and electronic circuits for actuation and measurements), where a high quality-factor resonator is observed ($Q_x \approx 100\,\text{dB}$). In addition, it was observed the existence of an electrical coupling between the actuation and the measurement of the gyroscope. This coupling is modeled as a linear system in parallel with the gyroscope and it is also presented in Figure 11.

During the study of the control design, the electrical coupling effect was ignored. Nonetheless, to compensate this effect, an additional block is considered in the closed-loop, as presented in Figure 12, where $E(s)$ represents the real electrical coupling and $\hat{E}(s)$ is the estimate of $E(s)$. Similarly to the controller, the block $\hat{E}(s)$ is also implemented in the embedded processor.

Finally, since the nonlinearity due to the actuation mechanism can be neglected, and the electrical coupling can be compensated, we face a structure similar to the one presented in Figure 7. Then, the $\mathcal{H}_\infty$ synthesis can be performed. We consider here the following specifications and respective constraints:

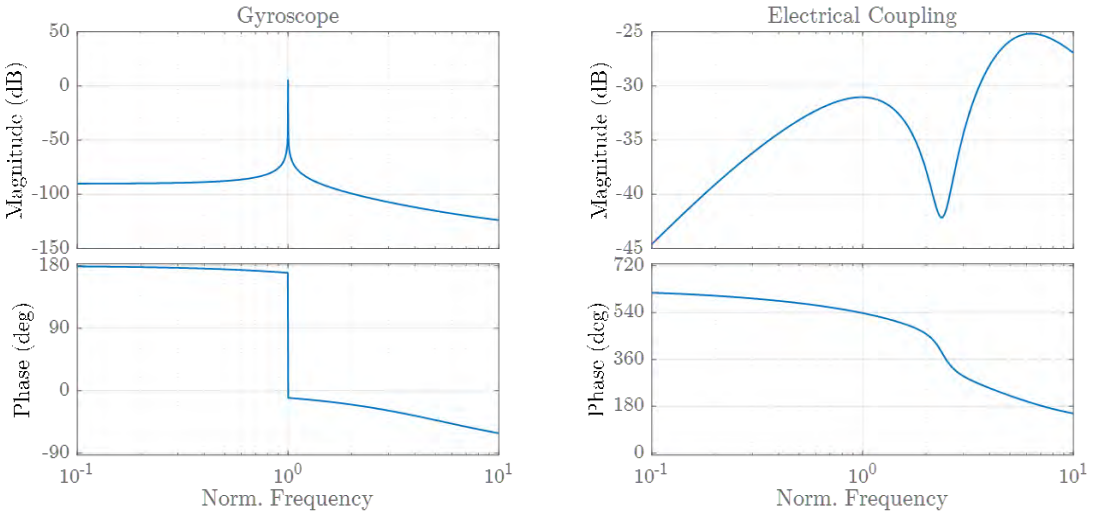


Figure 11: Bode diagrams of the drive mode of the gyroscope (left) and the electrical coupling (right).

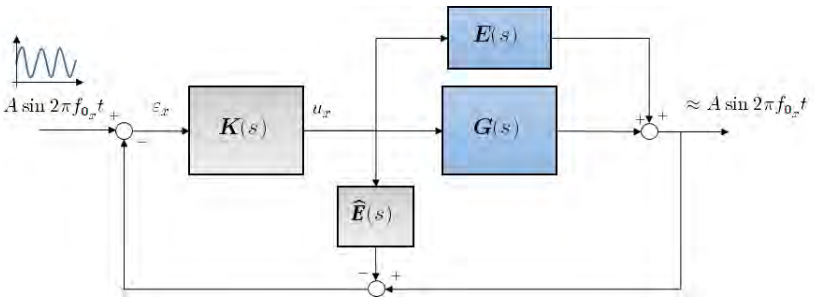


Figure 12: Block diagram for the electrical coupling compensation.

Sinusoidal Reference Tracking: tracking error not superior to 0.5% close (±20 ppm) to the resonance frequency. This is imposed by making $|W_r W_\epsilon| \geq 200$ (or 46 dB) close to the resonance frequency;

Robust Stability: modulus margin superior to 6 dB. This specification can be ensured by making $|W_r W_\epsilon| \geq 0.5 \ \forall \omega \in \mathbb{R}$; and

Reduction of the control energy: reduce as much as it is possible the control energy, mainly at low and high frequencies, where no control effort is required and noise is present. This constraint is ensured by making $|W_u W_r|$ as bigger as possible for low and high frequencies.

By solving the $\mathcal{H}_\infty$ problem with the weighting functions presented in Figure 13, we obtain $\gamma = 1.42$, the closed-loop frequency responses presented in the same figure, and the controller whose bode diagram is presented Figure 14.

The controller was discretised and implemented in C++ language in the embedded processor. The results of the first practical implementations are presented in the following figures. The reference, given by

$$ref(t) = A(t)\sin(\omega_{exc} t),$$

and the measured output are shown in Figure 15. The discontinuous aspect of the measurements are justified by the ratio of sampling frequency over the resonance frequency, approximately equal to 6. In other words, in a period, only 6 points of the output sinus are measured.

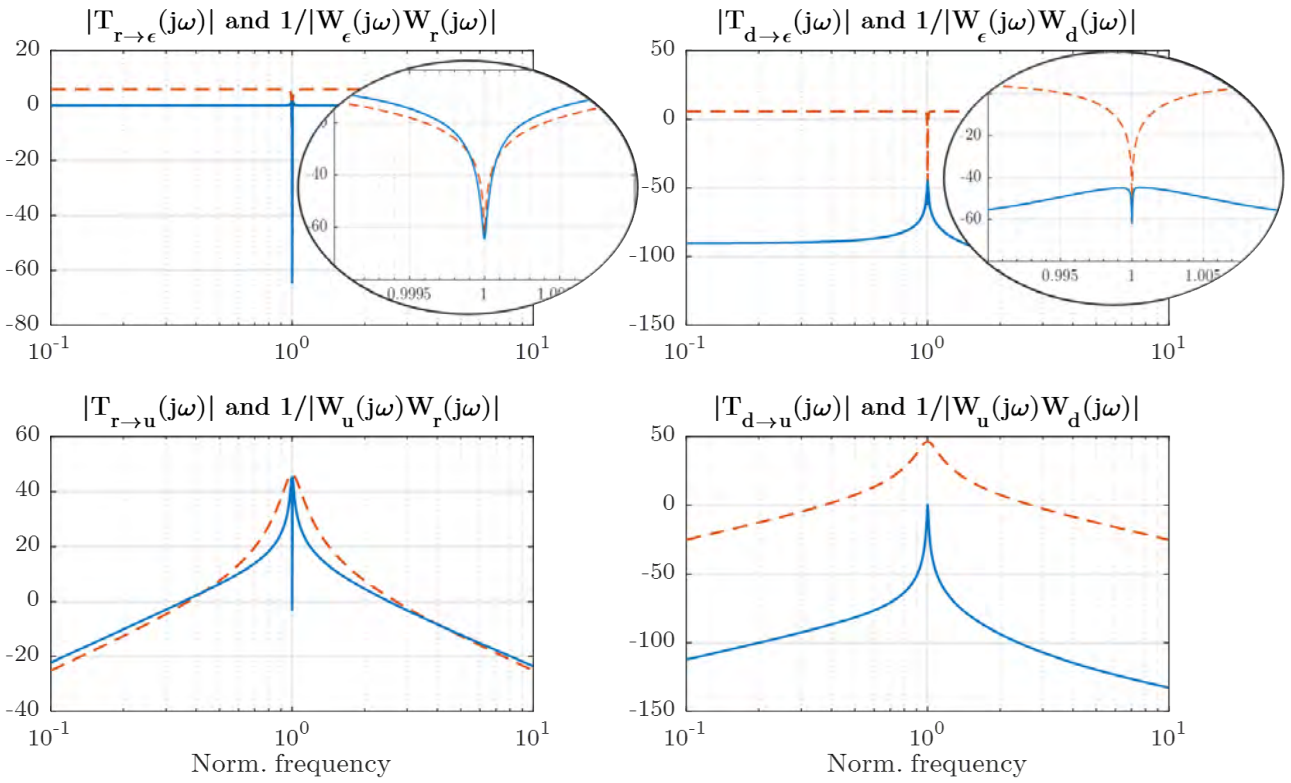


Figure 13: Constraints (combination of weighting functions - dashed lines) and closed-loop frequency responses (solid lines). The gains are in dB.

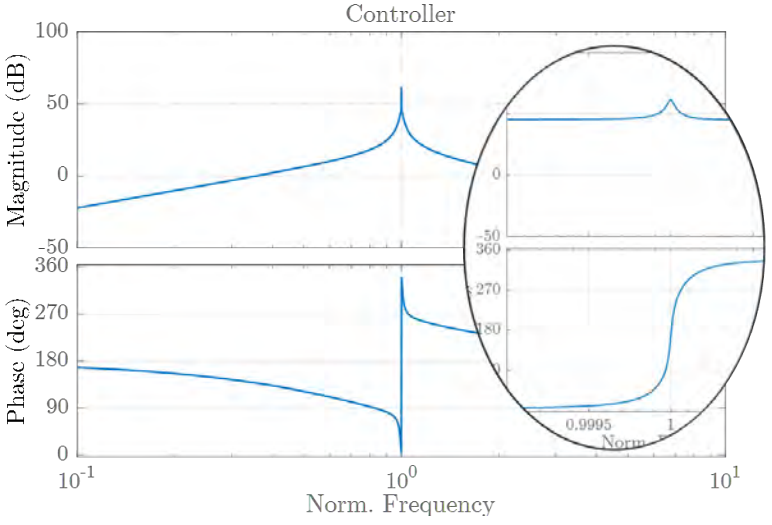


Figure 14: Bode diagram and zoom around the resonance frequency of the controller.

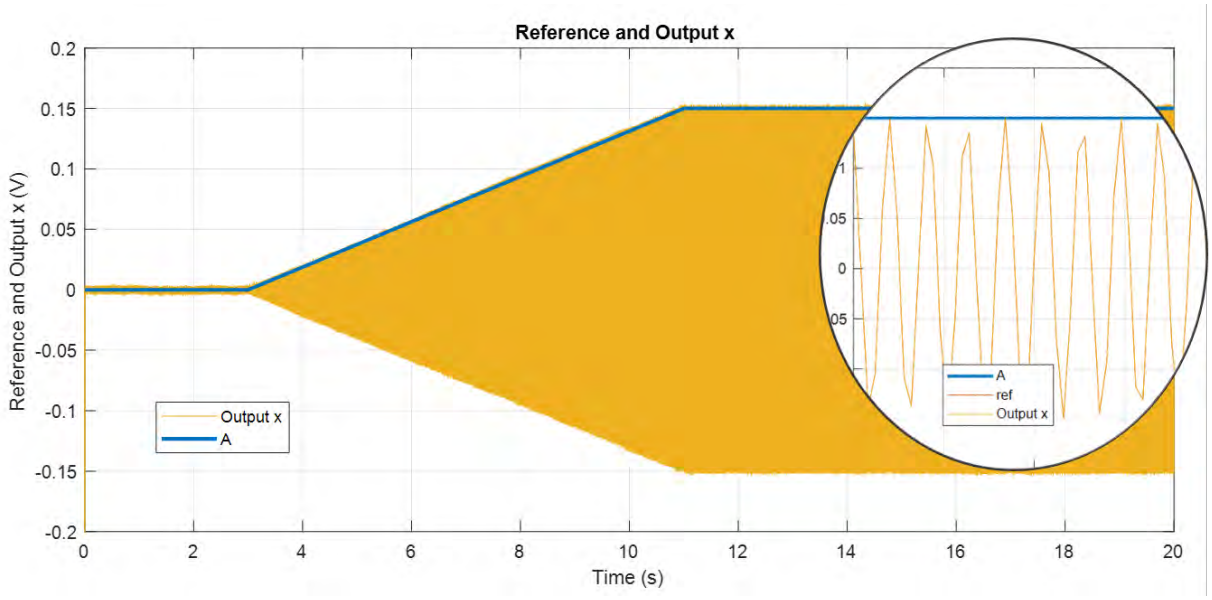


Figure 15: Results of practical implementation - reference tracking.

Since the error signal is contaminated by the measurement noise, it was evaluated by considering the power spectrum density (PSD) of the error, $\mathcal{S}_\varepsilon(j\omega)$, over the PSD of the reference signal $\mathcal{S}_r(j\omega)$. This ratio is presented Figure 16. We can notice that in the practical implementation, a tracking of a sinusoidal reference is obtained with an error close to the specifications ($-46\,\text{dB} \approx 0.5\%$). Nonetheless, since the gyroscope is very sensitive to temperature variations and has a very high quality factor, if the resonance frequency is not perfectly known, the gain of the gyroscope can be highly degraded, and in consequence, the performance of the whole system is also degraded.

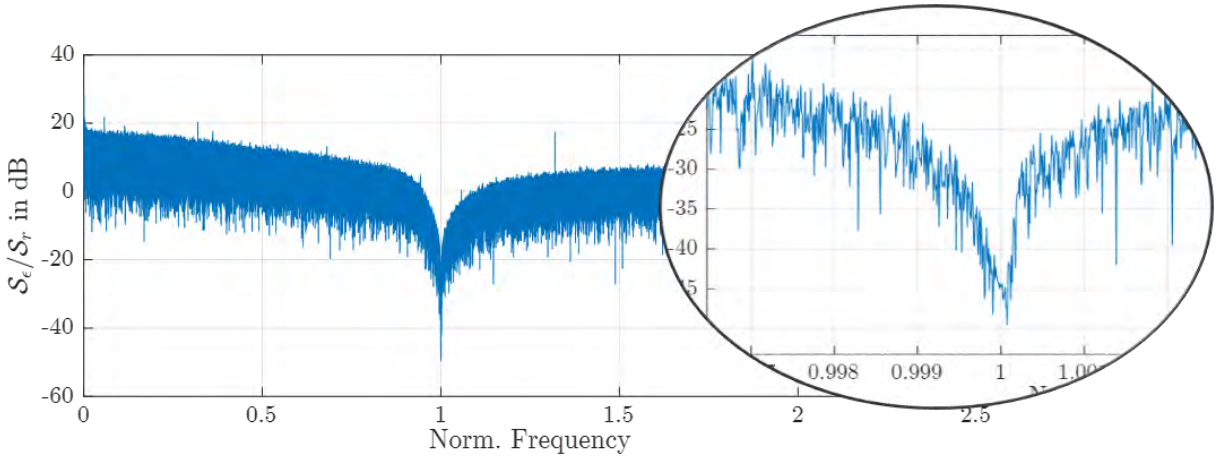


Figure 16: Results of practical implementation - error evaluation.

5 Envelope Dynamics Modeling

This section presents the work concerning the second research line, where we are interested on the modeling of the amplitude envelope of the signals. The technique presented in this section can be applied to a system if its input signal is modulated (multiplied by a sinus and/or a cosinus) and its output is demodulated. This structure is the same as the one presented in Figure 4, for control methods based on the envelope dynamics. This approach is inspired on the works [EK10, EMK12, BCRM02, BRCD05], and gives a richer model compared to the models used for envelope-based control techniques, as in [Sau08] for instance.

Let us consider an LTI system given by

$$\begin{cases} d\mathbf{x}(t)/dt &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases}, \tag{27}$$

where $t \in \mathbb{R}$ is the time variable, $\mathbf{x} \in \mathbb{R}^n$ is the state vector, $\mathbf{u} \in \mathbb{R}^{n_u}$ is the input vector, $\mathbf{y} \in \mathbb{R}^{n_y}$ is the output vector, $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the state matrix, $\mathbf{B} \in \mathbb{R}^{n \times n_u}$ is the input matrix, $\mathbf{C} \in \mathbb{R}^{n_y \times n}$ is the output matrix, and $\mathbf{D} \in \mathbb{R}^{n_y \times n_u}$ is the direct transfer matrix.

Let us also consider a modulated signal $\mathbf{s} \in \mathbb{R}^p$ in the form

$$\mathbf{s}(t) = \Re\{\overline{\mathbf{s}}(t) \cdot e^{j\omega_{exc}t}\}, \tag{28}$$

where $\overline{\mathbf{s}} \in \mathbb{C}^p$ is the "dynamic phasor"[8] of the signal $\mathbf{s}$ at the frequency ω_{exc} (in $\mathrm{rad\,s^{-1}}$), and $j^2 = -1$. Since $\overline{\mathbf{s}}$ is a complex vector, it can be decomposed in real and imaginary parts

$$\overline{\mathbf{s}}(t) = \mathbf{s_R}(t) + j\mathbf{s_I}(t) \tag{29}$$

Now, if an input $\mathbf{u}$ has the same form as $\mathbf{s}$ (i.e., a modulated input signal), the equation (27) can be rewritten as

$$\begin{cases} d\overline{\mathbf{x}}(t)/dt &= (\mathbf{A} - j\omega_{exc}\mathbf{I}_n)\overline{\mathbf{x}}(t) + \mathbf{B}\overline{\mathbf{u}}(t) \\ \overline{\mathbf{y}}(t) &= \mathbf{C}\overline{\mathbf{x}}(t) + \mathbf{D}\overline{\mathbf{u}}(t) \end{cases}, \tag{30}$$

where $\overline{\mathbf{x}}$, $\overline{\mathbf{y}}$ and $\overline{\mathbf{u}}$ are the dynamic phasors of $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{u}$, respectively. The transfer function of the system can be cast as

$$\overline{\mathbf{G}}(s) = \frac{\overline{\mathbf{x}}(s)}{\overline{\mathbf{u}}(s)} = \mathbf{C}(s\mathbf{I}_n - \mathbf{A} + j\omega_{exc}\mathbf{I}_n)^{-1}\mathbf{B} + \mathbf{D}. \tag{31}$$

Since

$$\Re\{\lambda(\mathbf{A})\} = \Re\{\lambda(\mathbf{A} - j\omega_{exc}\mathbf{I}_n)\}, \tag{32}$$

where $\lambda(\mathbf{A})$ stands for the eigenvalues of the matrix $\mathbf{A}$, the stability properties of the original system are conserved. Moreover, since the $\mathcal{H}_\infty$-norm of a proper stable system $\mathbf{G}(s)$ is defined as

$$\|\mathbf{G}\|_\infty = \sup_{\omega \in \mathbb{R}} \overline{\sigma}(\mathbf{G}(j\omega)) \tag{33}$$

and

$$\|\mathbf{G}\|_\infty = \sup_{\omega \in \mathbb{R}} \overline{\sigma}(\mathbf{G}(j\omega)) = \sup_{\omega \in \mathbb{R}} \overline{\sigma}(\overline{\mathbf{G}}(j\omega - j\omega_{exc})) = \left\|\overline{\mathbf{G}}\right\|_\infty, \tag{34}$$

the $\mathcal{H}_\infty$-norm is also conserved.

[8] If $\overline{\mathbf{s}}$ is constant through time, its definition coincides with the classical definition of phasors. Since $\overline{\mathbf{s}}$ can evolve over time, some authors call it dynamic phasors.

Instead of considering a system with complex-valued parameters and signals, (30) can be equivalently rewritten in an augmented form, where all matrix and signals are real-valued:

$$\begin{cases} \frac{d}{dt}\begin{bmatrix}\mathbf{x_R}(t)\\ \mathbf{x_I}(t)\end{bmatrix} &= \begin{bmatrix}\mathbf{A} & \omega_{exc}\mathbf{I}_n\\ -\omega_{exc}\mathbf{I}_n & \mathbf{A}\end{bmatrix}\begin{bmatrix}\mathbf{x_R}(t)\\ \mathbf{x_I}(t)\end{bmatrix} + \begin{bmatrix}\mathbf{B} & \mathbf{0}_{n\times n_u}\\ \mathbf{0}_{n\times n_u} & \mathbf{B}\end{bmatrix}\begin{bmatrix}\mathbf{u_R}(t)\\ \mathbf{u_I}(t)\end{bmatrix}\\ \begin{bmatrix}\mathbf{y_R}(t)\\ \mathbf{y_I}(t)\end{bmatrix} &= \begin{bmatrix}\mathbf{C} & \mathbf{0}_{n_y\times n}\\ \mathbf{0}_{n_y\times n} & \mathbf{C}\end{bmatrix}\begin{bmatrix}\mathbf{x_R}(t)\\ \mathbf{x_I}(t)\end{bmatrix} + \begin{bmatrix}\mathbf{D} & \mathbf{0}_{n_y\times n_u}\\ \mathbf{0}_{n_y\times n_u} & \mathbf{D}\end{bmatrix}\begin{bmatrix}\mathbf{u_R}(t)\\ \mathbf{u_I}(t)\end{bmatrix} \end{cases}.$$

The previous equation can also be presented in a compact way as the LPV system below:

$$\begin{cases} \frac{d}{dt}\mathbf{x_{dp}}(t) &= \mathbf{A_{dp}}(\omega_{exc})\mathbf{x_{dp}}(t) + \mathbf{B_{dp}}\mathbf{u_{dp}}(t)\\ \mathbf{y_{dp}}(t) &= \mathbf{C_{dp}}\mathbf{x_{dp}}(t) + \mathbf{D_{dp}}\mathbf{u_{dp}}(t) \end{cases} \tag{35}$$

Two possibilities to evaluate a dynamic phasor model were presented: with complex-valued parameters and signals, but with the same order of the original system; or with real-valued parameters and signals, but with a doubled order. The former one will be referred to as the $\mathbb{C}^n$-approach, and the latter one will be referred to as the $\mathbb{R}^{2n}$-approach.

The $\mathbb{C}^n$-approach has the advantage to conserve the order and the number of inputs and output of the original system. In addition, in a SISO framework, classical analysis tools, as Bode diagram and Nyquist plot can be used. However, since the parameters are complex-valued, the symmetry properties are lost, therefore the analysis shall cover positive and negative frequencies. The most important drawback of this approach is that synthesis tools that treat complex-valued parameters are rarely available.

In the other hand, even an original SISO system becomes a MIMO one in the $\mathbb{R}^{2n}$-approach. In this framework, Bode diagrams and Nyquist plots are harder to be interpreted. Nonetheless, the biggest advantage of this approach is that there exist synthesis tools adapted to MIMO systems.

We highlight that if the original system is linear, the relationships between the real and imaginary parts of the signals remain linear. Nonetheless, the amplitude $\mathbf{S}(t)$ (or the envelope) and the phase $\varphi_{\mathbf{s}}(t)$ are nonlinear with respect to the real and imaginary parts, as follows:

$$\mathbf{S}(t) = \sqrt{\mathbf{s_R}^2(t) + \mathbf{s_I}^2(t)} \quad \text{and} \quad \varphi_{\mathbf{s}}(t) = \arctan\frac{\mathbf{s_I}(t)}{\mathbf{s_R}(t)}$$

Numerical Example

In order to illustrate the dynamic phasor modeling, let us consider a resonant system (similar to the drive mode of a gyroscope)

$$G(s) = \frac{\omega_0^2}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2}, \tag{36}$$

with $\omega_0 = 2\pi 11500\ \mathrm{rad\,s^{-1}}$ and $Q = 50$. We consider that this system will be excited with $\omega_{exc} = 2\pi 10000\ \mathrm{rad\,s^{-1}}$. In order to evaluate the transient and the steady-state behavior of the systems, a simulation is realized and presented in Figure 17. For the phasor-space system (referred to as DPM, from dynamic phasor model), steps are applied to the real and to the imaginary parts, $u_R(t)$ and $u_I(t)$, respectively. The input of the direct system ($G(s)$) is given by

$$u(t) = u_R(t)\cos(\omega_{exc}t) - u_I(t)\sin(\omega_{exc}t), \tag{37}$$

in the same form as equation (28). DPM represents the instantaneous amplitude of y, $Y(t) = \sqrt{y_{Re}^2(t) + y_{Im}^2(t)}$. The simulation shows the exact correspondence between the output of $G(s)$ and the envelope model (DPM). Since we are interested on modeling the low-frequency dynamics of the envelope, a model reduction could be performed to consider only the low-frequency dynamics (eliminating the small oscillations of the envelope).

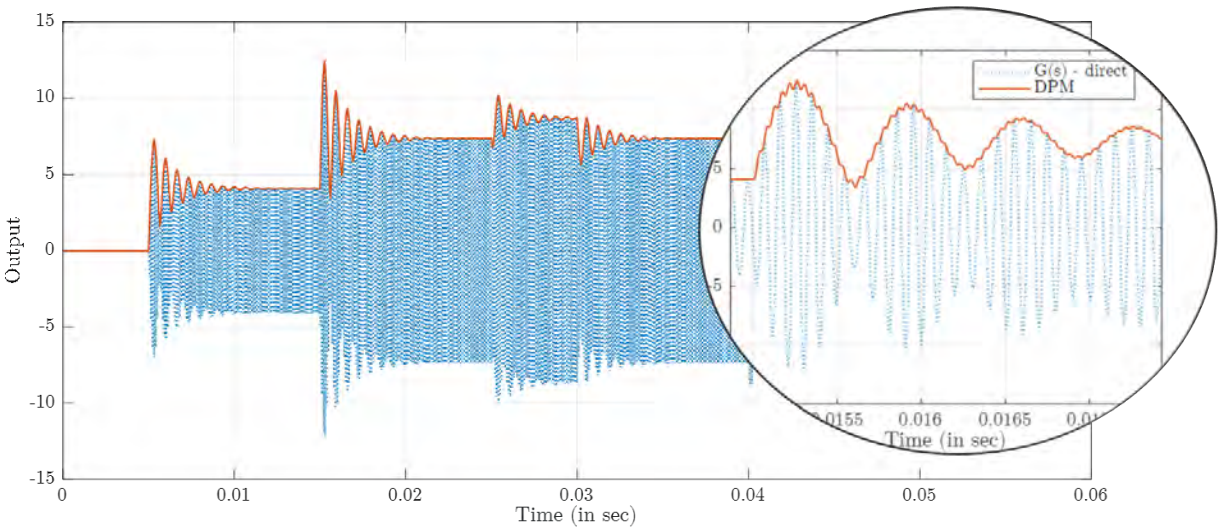


Figure 17: Numerical example 1.

In order to compare the real and imaginary parts of DPM and the real and imaginary parts of the output of $G(s)$, a synchronous demodulation is considered for the latter one. Real and imaginary terms of $G(s)$ are directly related to the direct and quadrature terms ($\tilde{y}_c$ and $\tilde{y}_s$, respectively) of the synchronous demodulation (SD). The results are presented in Figure 18, where we can notice the very similar comportment of the systems. The main difference that can be noticed is the delay present in $G(s)$ - SD (solid lines). This delay is due to the filtering stage of the synchronous demodulation. This non-ideality was also modeled and allows one to take it into account.

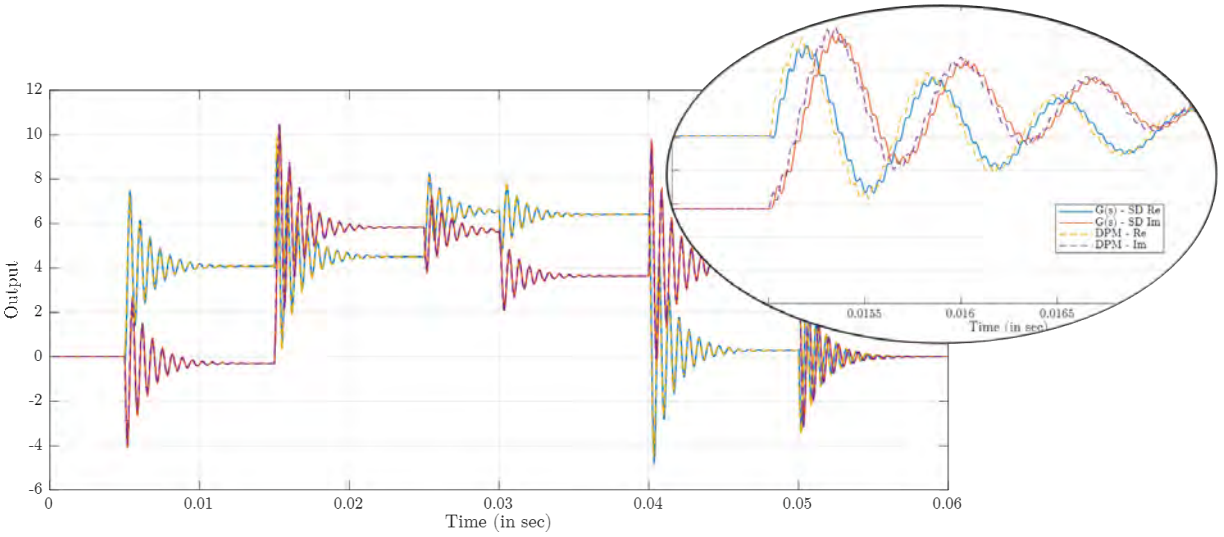


Figure 18: Numerical example 2.

6 Perspectives and Time Schedule

Regarding the development lines presented in Section 3.4 and the work executed until the present moment, the perspectives of future work are as follows :

Line 1: The next steps consist on the implementation of the SISO controller for the sense mode, and of the multivariable architectures. In parallel, the evaluation of their robustness (for stability and performance) has been realized by Jorge Ivan Ayala Cuevas (see report "Performance validation of MEMS sensor using nonlinear uncertain models").

The electrical coupling was not taken into account for the controller design. However, some improvements can be achieved if this effect is considered for the controller synthesis.

MEMS inertial sensors are very sensitive to temperature variations, and their performance can be quickly deteriorated. Thus, if one would be able to detect the temperature or the resonance frequency in real-time, an LPV controller can be considered;

Line 2: The proposed envelope modeling allows to describe the gyroscope behavior when considering modulation and synchronous demodulation inside the control loop. This allows to couple excitation frequency (or excitation instantaneous phase) with the instantaneous amplitude. Considering this framework for the drive mode, it would be possible to regulate the amplitude of the oscillations and to track the resonance frequency at the same time with a multivariable controller;

6.1 Provisional Time Schedule

In order to achieve the expected objectives, the following time schedule is proposed (Table 1). This table presents the remaining tasks and the month(s) of execution, counting from September 2018 (month 1) and previewing to finish by September 2020 (month 25). The description of the tasks is described below:

Table 1: Time Schedule

Task\Month	1-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16	17-18	19-20	21-22	23-24	25
A	X	X											
B	X	X											
C			X										
D				X									
E					X	X							
F							X	X					
G				X	X			X					
H			X			X			X				
I									X	X	X	X	
J													X

A Implementation of $\mathcal{H}_\infty$ MIMO controllers without and with direct estimation of the Coriolis force;

B Simulations and implementation of $\mathcal{H}_\infty$ envelope-based controllers;

C Study on the resonance frequency tracking;

D Implementation of resonance frequency tracking strategies;

E Study on the LPV controllers;

F Development and implementation of LPV controllers (based on the resonance frequency or temperature);

G Adaptation and implementation of the control design methods to accelerometers (depends on the availability of the device);

H Paper writing. Three papers can be considered: one about the multivariable controller, one about the envelope modeling, and one about the LPV approach;

I Thesis writing;

J Thesis defense.

7 Conclusions

This document summarizes the work realized since October 2017. The bibliographic review about the MEMS inertial sensors (mainly the gyroscopes) and related control techniques allowed to better understand the application as well as to define two main lines of research: one based on the direct approach and on different $\mathcal{H}_\infty$-based controllers, and the second one which is focused on the envelope modeling and related control techniques.

First practical results showed good performance for the direct approach. However, there is place for further improvements. The implementation of more complex structures, as the MIMO extensions, have to be realized. Moreover, the inclusion of the electrical coupling for the controller synthesis and an LPV approach will be considered.

Besides the direct approach, an approach for the modeling of the envelope dynamics was proposed. The modeling is rich and accurate, what would allow us to perform a model reduction and take into account only the low-frequency dynamics of the system while mastering the errors that are introduced. This approach is suited for implementations where the modulation and demodulation stages are already implemented in the actuation and sensing chain of the system.

Finally, by exploiting the different approaches presented in this document, we intend to be able to propose a general method to systematically design controllers for MEMS inertial sensor while ensuring a required performance.

References

[BCRM02] B. Boivin, P. Coirault, L. Rambault, and N. Maamri. Modelling and H_∞ Controller Applied to a Gyrometer. In *IEEE Int. Conf. Syst. Man Cybern.*, volume vol.4, page 6. IEEE, 2002. doi:10.1109/ICSMC.2002.1173294.

[BRCD05] B. Boivin, L. Rambault, P. Coirault, and C. Dewez. Multivariable Output Feedback Stabilisation with Structure Constraint: Application to a Gyrometer. In *Proc. 2005, Am. Control Conf. 2005.*, pages 1148–1153. IEEE, 2005. doi:10.1109/ACC.2005.1470116.

[BSK06] C. Batur, T. Sreeramreddy, and Q. Khasawneh. Sliding Mode Control of a Simulated MEMS Gyroscope. *ISA Trans.*, 45(1):99–108, 2006. doi:10.1016/S0019-0578(07)60069-X.

[CMTB05] Y. C. Chen, R. T. M'Closkey, T. A. Tran, and B. Blaes. A Control and Signal Processing Integrated Circuit for the JPL-Boeing Micromachined Gyroscopes. *IEEE Trans. Control Syst. Technol.*, 13(2):286–300, March 2005. doi:10.1109/TCST.2004.839558.

[Col18] Kévin Colin. Joint Identification and Control of MEMS Sensors. Technical report, Laboratoire Ampère, 2018.

[Cue18] Jorge Ivan Ayala Cuevas. Performance Validation of MEMS Sensors Using Nonlinear Uncertain Models. Technical report, Laboratoire Ampère, 2018.

[CYC+14] F. Chen, W. Yuan, H. Chang, G. Yuan, J. Xie, and M. Kraft. Design and Implementation of an Optimized Double Closed-Loop Control System for MEMS Vibratory Gyroscope. *IEEE Sens. J.*, 14(1):184–196, January 2014. doi:10.1109/JSEN.2013.2271586.

[EK10] M. Egretzberger and A. Kugi. A Dynamical Envelope Model for Vibratory Gyroscopes. *Microsyst. Technol.*, 16(5):777–786, 2010. doi:10.1007/s00542-009-0979-y.

[EMK12] M. Egretzberger, F. Mair, and A. Kugi. Model-Based Control Concepts for Vibratory MEMS Gyroscopes. *Mechatronics*, 22(3):241–250, April 2012. doi:10.1016/j.mechatronics.2011.06.003.

[Fei10] J. Fei. Robust Adaptive Vibration Tracking Control for a MEMS Vibratory Gyroscope with Bound Estimation. *IET Control Theory Appl.*, 4(6):1019–1026, 2010. doi:10.1049/iet-cta.2008.0199.

[FY11] J. Fei and Y. Yang. System Identification of MEMS Vibratory Gyroscope Sensor. *Math. Probl. Eng.*, pages 1–12, 2011. doi:10.1155/2011/829432.

[HZH+15] C. He, Q. Zhao, Q. Huang, D. Liu, Z. Yang, D. Zhang, and G. Yan. A MEMS Vibratory Gyroscope with Real-Time Mode-Matching and Robust Control for the Sense Mode. *IEEE Sens. J.*, 15(4):2069–2077, April 2015. doi:10.1109/JSEN.2014.2371456.

[KCV08] C. Kharrat, E. Colinet, and A. Voda. $\mathcal{H}_\infty$ Loop Shaping Control for PLL-Based Mechanical Resonance Tracking in NEMS Resonant Mass Sensors. In *2008 IEEE Sensors*, pages 1135–1138. IEEE, 2008. doi:10.1109/ICSENS.2008.4716641.

[Kem11] V. Kempe. *Inertial MEMS - Principles and Practice*. Cambridge University Press, Cambridge, 2011.

[KS17] L. Král and O. Straka. Nonlinear Estimator Design for MEMS Gyroscope with Time-varying Angular Rate. *IFAC-PapersOnLine*, 50(1):3195–3201, July 2017. doi:10.1016/j.ifacol.2017.08.433.

[Lab18] Laboratoire Ampère (institution). NEXT4MEMS (2017-2020) : Capteurs Inertiels MEMS de Haute Performance, 2018. URL: http://www.ampere-lab.fr/spip.php?article885.

[LLC+11] L. Lin, D. Liu, J. Cui, Z. Guo, Z. Yang, and G. Yan. Digital Closed-Loop Controller Design of a Micromachined Gyroscope Based on Auto Frequency Swept. In *2011 6th IEEE Int. Conf. Nano/Micro Eng. Mol. Syst.*, number 1, pages 654–657. IEEE, February 2011. doi:10.1109/NEMS.2011.6017440.

[LR02] P. W. Loveday and C. A. Rogers. The Influence of Control System Design on the Performance of Vibratory Gyroscopes. *J. Sound and Vibration*, 255(3):417–432, 2002. doi:10.1006/jsvi.2001.4163.

[LWS10] Y. Liu, L. Wang, and D. Sun. The LQG Controller Design for Micromachined Tunneling Gyroscope. In *2010 IEEE 5th Int. Conf. Nano/Micro Eng. Mol. Syst.*, pages 348–351. IEEE, January 2010. doi:10.1109/NEMS.2010.5592236.

[MVG01] R. T. M'Closkey, A. Vakakis, and R. Gutierrez. Mode Localization Induced by a Nonlinear Control Loop. *Nonlinear Dyn.*, 25(1/3):221–236, 2001. doi:10.1023/A:1012934112137.

[OAL+05] R. Oboe, R. Antonello, E. Lasalandra, G. Spinola Durante, and L. Prandi. Control of Z-Axis MEMS Vibrational Gyroscope. *IEEE/ASME Trans. Mechatronics*, 10(4):364–370, August 2005. doi:10.1109/TMECH.2005.852437.

[PH04] S. Park and R. Horowitz. New Adaptive Mode of Operation for MEMS Gyroscopes. *J. Dyn. Syst. Meas. Control*, 126(4):800, 2004. doi:10.1115/1.1849252.

[PHHN07] S. Park, R. Horowitz, S. K. Hong, and Y. Nam. Trajectory-Switching Algorithm for a MEMS Gyroscope. *IEEE Trans. Instrum. Meas.*, 56(6):2561–2569, 2007. doi:10.1109/TIM.2007.908597.

[PK14] M. H. Pishrobat and J. Keighobadi. Model Predictive Control of MEMS Vibratory Gyroscope. *IFAC Proc. Vol.*, 47(3):7278–7283, 2014. doi:10.3182/20140824-6-ZA-1003.02322.

[PPK+04] Y. Park, S. Park, D. Kwak, H. Ko, T. Song, D. Cho, K. Huh, and J. H. Park. Feedback Control of MEMS Gyroscope to Achieve the Tactical-Grade Specifications. *IFAC Proc. Vol.*, 37(6):671–676, June 2004. doi:10.1016/S1474-6670(17)32253-X.

[RCRW09] J. Raman, E. Cretu, P. Rombouts, and L. Weyten. A Closed-Loop Digitally Controlled MEMS Gyroscope with Unconstrained Sigma-Delta Force-Feedback. *IEEE Sens. J.*, 9(3):297–305, March 2009. doi:10.1109/JSEN.2008.2012237.

[SAH07] M. Saukoski, L. Aaltonen, and K. A. I. Halonen. Zero-Rate Output and Quadrature Compensation in Vibratory MEMS Gyroscopes. *IEEE Sens. J.*, 7(12):1639–1652, December 2007. doi:10.1109/JSEN.2007.908921.

[Sau08] M. Saukoski. *System and Circuit Design for a Capacitive MEMS Gyroscope*. PhD thesis, 2008.

[SF09] G. Scorletti and V. Fromion. *Automatique Fréquentielle Avancée*. 2009. URL: https://cel.archives-ouvertes.fr/cel-00423848v2.

[SHLK07] W. T. Sung, C. Hyun, J. G. Lee, and T. Kang. Design of an AGC Driving Loop in MEMS Gyroscopes. *IFAC Proc. Vol.*, 40(7):335–340, 2007. doi:10.3182/20070625-5-FR-2916.00058.

[SHS+99] A. M. Shkel, R. Horowitz, A. A. Seshia, S. Park, and R. T. Howe. Dynamics and Control of Micromachined Gyroscopes. In *Proc. 1999 Am. Control Conf. (Cat. No. 99CH36251)*, volume 3, pages 2119–2124. IEEE, 1999. doi:10.1109/ACC.1999.786303.

[SKL08] W. T. Sung, T. Kang, and J. G. Lee. Controller Design of a MEMS Gyro-Accelerometer with a Single Proof Mass. *Int. J. Control Autom. Syst.*, 6(6):873–883, 2008.

[SLSK04] W. T. Sung, J. G. Lee, J. W. Song, and T. Kang. H_∞ Controller Design of MEMS Gyroscope and its Performance Test. In *PLANS 2004. Position Locat. Navig. Symp. (IEEE Cat. No.04CH37556)*, pages 63–69. IEEE, 2004. doi:10.1109/PLANS.2004.1308975.

[SP01] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control - Analisys and Design.* John Wiley & Sons, second edition, 2001.

[You11] M. I. Younis. *MEMS Linear and Nonlinear Statics and Dynamics*, volume 20 of *Microsystems.* Springer US, 2011. doi:10.1007/978-1-4419-6020-7.

[ZDG07] Q. Zheng, L. Dong, and Z. Gao. Control and Rotation Rate Estimation of Vibrational MEMS Gyroscopes. In *2007 IEEE Int. Conf. Control Appl.*, number May, pages 118–123. IEEE, 2007. doi:10.1109/CCA.2007.4389216.

APPENDIX

A Requirements/Specifications

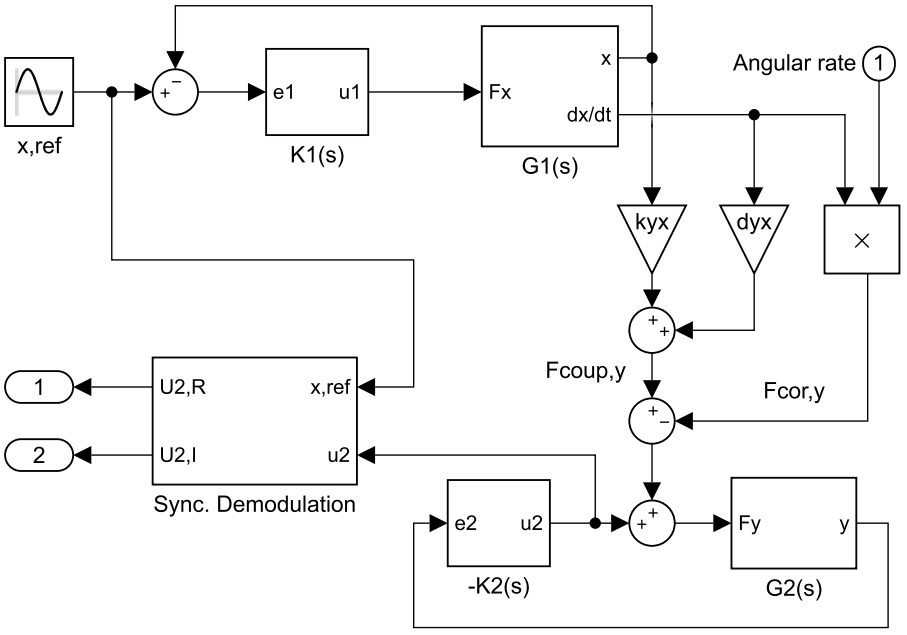


Figure 19: Block diagram illustrating the considered control strategy.

As the principal objective of the gyroscope is to estimate the angular rate, firstly we determine the influence of each control loop on this estimation. Then, by having these relationships and the required estimation performance, one would be able to establish the specifications for each control loop.

Primary mode: The objective of the primary mode is to make the proof mass oscillate at (or close to) the resonance frequency. Nonetheless, the control loop is not ideal, and the output of the drive mode is contaminated by an error. So, in sinusoidal steady-state (at the frequency ω), the output can be defined by

$$\overline{x} = T_x(j\omega)\overline{x_{ref}} \tag{38}$$

where x_{ref} is the reference (or desired output) signal and T_x is the frequency response from x_{ref} to x, given by

$$T_x(j\omega) = \frac{K_1(j\omega)G_x(j\omega)}{1 + K_1(j\omega)G_x(j\omega)}. \tag{39}$$

The symbol $\overline{(\cdot)}$ is used to represent the phasorial notation of $(\cdot)$.

Coriolis force: The Coriolis force acting on the sense mode can be given by

$$\overline{F_{Cor,y}} = -j\omega\tilde{\Omega}\overline{x} \tag{40}$$

$$\overline{F_{Cor,y}} = -j\omega\tilde{\Omega}T_x(j\omega)\overline{x_{ref}} \tag{41}$$

where $\tilde{\Omega} = 2m_x\Omega_z$.

Coupling force: The coupling force acting on the sense mode can be given by

$$\overline{F_{coup,y}} = (k_{yx} + j\omega d_{yx})T_x(j\omega)\overline{x_{ref}} \tag{42}$$

Secondary mode: Often, the control loop of the sense mode is used to reject the "input disturbances" provoked by the Coriolis and the coupling forces. Thus, one can affirm that in steady state, the transfer function from the disturbance to the control signal u_2, $-T_y$, must be close to -1, at least close to the working frequency. This transfer is defined as

$$T_y(j\omega) = \frac{K_2(j\omega)G_y(j\omega)}{1 + K_2(j\omega)G_y(j\omega)}. \tag{43}$$

Then, the control signal u_2 can be given by

$$\overline{u_2} = T_y(j\omega)(-\overline{F_{Cor,y}} + \overline{F_{coup,y}}) \tag{44}$$

$$\overline{u_2} = T_y(j\omega)(j\omega\tilde{\Omega} + k_{yx} + j\omega d_{yx})T_x(j\omega)\overline{x_{ref}}. \tag{45}$$

Ideally (i.e., if $T_x(j\omega_x) = T_y(j\omega_x) = 1$), the control signal can be given by

$$\overline{u_2}^{\star} = (j\omega_x\tilde{\Omega} + k_{yx} + j\omega_x d_{yx})\overline{x_{ref}}, \tag{46}$$

where $(\cdot)^{\star}$ represents the ideal case. Hence, an estimate of the error, ε_{u2}, can be cast as

$$\varepsilon_{u2} = \frac{|\overline{u_2} - \overline{u_2}^{\star}|}{|\overline{u_2}^{\star}|} \tag{47}$$

$$\varepsilon_{u2} = |(1 - T_y(j\omega_x)T_x(j\omega_x))|. \tag{48}$$

Since the transfer functions T_x and T_y can be also expressed in terms of the sensibility functions ($S_x = 1 - T_x$ and $S_y = 1 - T_y$), the equation (48) can be approximated by

$$\varepsilon_{u2} \approx |S_x(j\omega_x)| + |S_y(j\omega_x)|. \tag{49}$$

Synchronous demodulation: Here, we consider that the synchronous demodulation just separates the real from the imaginary part of $\overline{u_2}$. Since $\overline{u_2} \approx \overline{u_2}^{\star}$, we can consider that Ω_z can be deduced by the imaginary part of $\overline{u_2}$ (if d_{yx} is known). Here, it is considered that the synchronous demodulation does not include any errors to the estimation.

It is important to notice that this approach can somehow be optimist since no error during the synchronous demodulation is considered. However, it demonstrates that the error of the estimation is highly and equally sensitive to the error generated in the drive and the sense modes.