UNIVERSITÀ DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
"TECNOLOGIE DELL'INFORMAZIONE"

CICLO XXXIV

Big Data-based Identification and Fault Detection for Industrial Monitoring with Application to a Pharmaceutical Freeze-Dryer

Coordinatore:
Chiar.mo Prof. Marco Locatelli

Tutore:
Chiar.mo Prof. Luca Consolini
Chiar.mo Prof. Gianluigi Ferrari

Dottorando: Gabriele Calzavara

Anni 2018/2021

To my family

Contents

List of Figures

Introduction

Maintenance operations on industrial machinery have traditionally been of two types: corrective or preventive. In the first case, repairs are carried out after the breakdown has occurred: this leads to inevitable and unexpected machine failures that cause delays in the manufacturing process. In the second case, on the other hand, it is established a priori when the equipment will be replaced, regardless of the actual deterioration state of the equipment that compose it: the risk here is that of not exploiting components and machinery until the end of their life cycle and therefore of replacing an asset that is still effectively functioning. These are clearly two ineffective and unnecessarily expensive approaches that do not provide any added value from a 4.0 industrial competitiveness perspective but, despite this, they are still the two paradigms most used by most companies today. In order to favor the widespread adoption of smart maintenance, it is crucial to provide systems that can facilitate the monitoring of the current state of machines and enabling a deepen interaction within human and machines. In order to bridge the gap between machines and humans, digital transformation can provide useful services to final users as a consequence of the processing work on the huge amount of data collected by machine sensors [1, 2].

Given such a complex environment there is a need to develop a smarter supply chain and, in this sense, the Pharma sector is one of the most fertile grounds for digital innovation [3, 4]. In fact, in the Pharmaceutical industries, there is a need to contain costs and improve business results while maintaining high quality standards. The traditional supply chain approach based on requirements planning is no longer effective. The solution is an Industry 4.0 that can effectively anticipate failures, i.e.,

an intelligent supply chain based on digital twin and predictive analytics [5, 6].

As a result, Big Data is becoming an increasingly decisive factor in enabling companies to advance in terms of efficiency, productivity and quality. Big Data Machine Analytics is a concept that defines the accumulation of data through the use of sensors placed within an industrial facility [7]. These large amounts of data are accumulated in efficient databases and then analyzed and interpreted for future purposes. However, despite widespread recognition of the usefulness of insights gained from IoT data, often almost all of the information assets collected by companies are not actually used. In fact, data management remains a critical issue for many companies, primarily due to two problems. First, there is often a lack of competent figures with respect to information management. The second problem is due to the fact that most of the data acquired comes from signals aimed at monitoring manufacturing processes and not the actual state of components deterioration of the plant itself [8]. This implies that most of these data are not useful to explain the machine health state and it is very complex to identify the hidden relations necessary to create a diagnostic model. To deal with these problems, companies need domain experts, who are familiar with processes and equipment in the industry, and specialized in computer science and mathematical modeling, in order to manage and process the acquired data [9]. The most wanted figure possessing these skills is the Data Scientist [10]. These figures are able to perform Data Mining, they extract from Big Data the information necessary to create diagnostic models and simulate the machine status. Predictive maintenance is one of the most active fields of study for Industry 4.0, as it is expected to significantly decrease the maintenance costs of the equipment. The big challenge then is to be able to create concrete solutions that are easy to evaluate.

Study Overview

Two of the main goals of the digital transformation are (i) the development of a process simulator that represents a "digital twin" of a machine and (ii) the development of diagnostic and health management system for the machines. The first objective allows to develop a tool for simulation-based preliminary tests of changes to the ma-

chine control system and for the study of process improvements. Thanks to the tools developed to reach the second goal it is possible to detect and predict the degradation of machine equipment and to schedule maintenance interventions more efficiently.

With the intention of addressing the first objective, in Chapter 2, we begin with a theoretical study regarding the heat exchanges that can occur within machines of this type. We exploit the equivalence that exists between thermal circuits and resistive-capacitive (RC) networks. We consider the problem of recovering the network connection structure from measured input-output data. We address this problem as a structured identification one, that is, we assume to have a state-space model of the system (identified with standard techniques, such as subspace methods) and find a coordinate transformation that puts the identified system in a form that reveals the nodes connection structure. We characterize the solution set, that is, the set of all possible RC-networks that can be associated with the input-output data. We present a possible solution algorithm and show some computational experiments.

To achieve the second goal and create solutions applicable to pharmaceutical freeze-dryers, in Chapter 3, we focus on freeze-dryers and the study of their sensor signals. It is often impossible to accurately predict the deterioration of a component, as the reliability of predictive models strongly depends on the available sensory data and on the specific characteristics of the monitored component. So we present a time-aware clustering-based approach to the analysis of sensory data with the aim of predicting the temporal evolution of the health status of a machine component in a pharmaceutical plant. The developed strategy allows to obtain a time segmentation of the component's operational points, which are then clustered using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN). In particular, this approach has the advantages of being general and making use of a limited amount of features extracted from a single sensor signal. The proposed approach becomes attractive when the data collected from a single sensor are not sufficient to build a physical model capable of identifying changes in the system status.

In some cases, however, having a simple physical model is the key to obtaining a reliable diagnostic system. In Chapter 4, we therefore focused on studying the Leak Test process, which verifies the sealing of the machine. In this process, the product

has to be brought to a very low pressure and the lyophilization chamber has to be perfectly sealed. Even small external leaks can contaminate the entire drug batch. Since a single batch may contain thousands of product vials, freeze-dryer leakages are one of the most critical problems of the entire production chain of lyophilized drugs. We describe a simple mathematical model for lyophilizer leaks and address the problem of identifying and separating internal and external leaks. We propose a leak identification method based on the use of multiple leak detection tests. By using the real data of a pharmaceutical lyophilizer, we show that the proposed method allows to identify internal and external leaks and to estimate their evolution in time.

Thesis Structure

Chapter 1 provides the background of the freeze-dryer machine and lyophilization processes in a pharmaceutical plant. In Chapter 2 is faced the problem of recovering the network connection structure from measured input-output data of a Resisitive-capacitive (RC) circuit. Chapter 3 presents a general method to predict the deterioration of machine components thanks to a time-aware clustering-based approach. In Chapter 4, a mathematical model to identify leaks in a pharmaceutical freeze-dryer is proposed.

Chapter 1

The Freeze-Dryer Case Study

GlaxoSmithKline plc (GSK) is a British multinational pharmaceutical company headquartered in London, England. Established in 2000 by a merger of Glaxo Wellcome and SmithKline Beecham, GSK is one of the world's largest pharmaceutical company. It is a science-led global healthcare company with a special purpose: to help people do more, feel better, live longer. The company has three global businesses that research, develop and manufacture innovative pharmaceutical medicines, vaccines and consumer healthcare products. Legacy products developed at GSK include several listed in the World Health Organization's List of Essential Medicines, some of which, are freeze-dried products. This means that the supply chain for these products must always be efficient and high-performing. In addition, freeze-dried biopharm drugs have an extremely high development and manufacturing cost, therefore the freeze-drying process is critical and must be constantly monitored and kept under control.

GSK is also present in Italy with six sites. Thanks to its expertise and ability to evolve as a manufacturing facility, Parma is a center of excellence in GSK's industrial network that distributes its products worldwide. This PhD study is framed within a project created in collaboration with the digital department of GSK Parma, with the aim of building monitoring systems for industrial freeze-dryers present within the site.

1.1 Freeze-Drying

Freeze-drying is a standard procedure in pharmaceutical industry, used to stabilize, store and increase the shelf life of drug products.

The freeze-drying (or lyophilization) process is a method to dehydrate a product and it has important applications in the pharmaceutical and biotechnology industries. Pharmaceutical freeze-drying is now a standard process used to stabilise, store or increase the shelf life of drug products. For a detailed description of freeze-drying, see for instance [11, 12].

Freeze-drying works by freezing the material, then reducing the pressure and increasing the temperature to allow the frozen water in the material to sublimate. The cycle consists of three processes: freezing, primary drying, secondary drying.

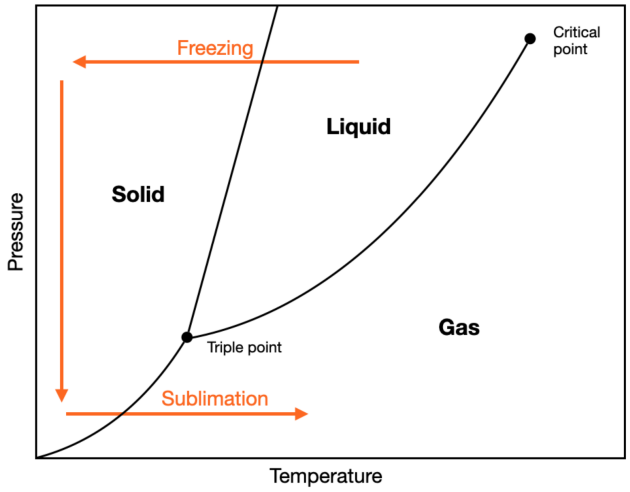


Figure 1.1: Water state diagram: the blue arrows indicate the three steps performed during freeze-drying.

- *Freezing*
  During the freezing step the entire product is brought to temperatures below the

triple point of the product solvent (that is the point at which the solid, liquid and gas phases can coexist). Most of the solvent contained in the product solidifies and forms ice crystals. The operating conditions at which freezing takes place have effects on the final product. More precisely, these conditions determine the structure of the ice crystals that form, which can have consequences on the characteristics of the final product.

The solidification is the result of two consecutive phenomena: the nucleation of the ice crystal and its growth. If solidification occurs in a time prolonged, growth is faster than nucleation: crystals form with larger sizes, creating a structure with wider pores during the next drying and reducing the resistance to transport of steam during sublimation. Moreover, large crystal sizes (due to long solidification) can irreversibly damage the crystalline structure of the solid product. Therefore fast freezing is recommended in order to get small ice crystals, which limited impact on the quality of the product.

- *Primary drying (sublimation)*
  The pressure in the chamber is reduced until the partial pressure of the water vapor is lower than the value of the triple point shown in Fig. 1.1, and the energy (heat) is supplied to the product for the frozen water to sublimate. This step is necessary because the sublimation process is endothermic and, therefore, requires energy. During sublimation a dry layer is formed in the material whose thickness increases over time: the process is thus characterized by a mobile sublimation interface. While the ice sublimates, the generated steam flows through the dry material and creates empty spaces in the product, leading to a porous structure which facilitates rehydration. The driving force of the process is the difference in partial pressure of the water vapor between the sublimation interface and the freeze-dryer chamber in which the product is contained. The partial pressure of the steam in the chamber is controlled through a condenser, which involves the continuous removal of the water eliminated from the product by sublimation. At the end of primary drying the product can still contain a low quantity of water: this water is not removed by sublimation during primary drying, but is desorbed during the secondary drying phase.

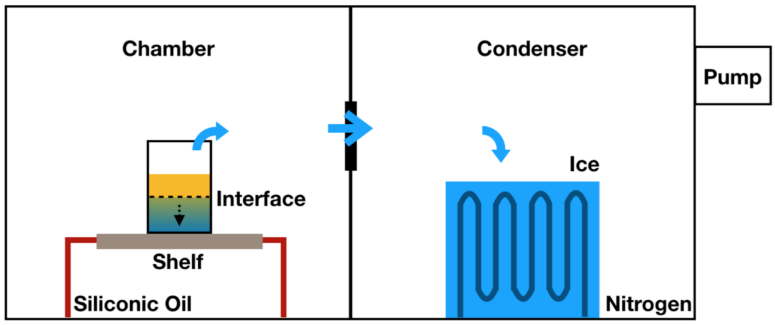


Figure 1.2: Primary drying process.

- *Secondary drying (adsorption)*
  The remaining water, still adsorbed on the porous surface from the weak bonds within the dry material, is desorbed into conditions of low pressure and temperature higher than the primary drying phase. This process allows the content to reach of desired humidity level.

Freezing and secondary drying phases are faster than the primary drying phase. From the point of view of the quality of the final product, freeze-drying is considered the best drying method thanks to the low temperatures at which it is carried out.

1.2 The Freeze-Dryer Machine

The freeze-dryer is a steel machine, consists of two chambers: the main chamber, where the drug bottles are inserted, and the condensation chamber, which is necessary to trap the sublimated water during primary and secondary drying. The main chamber is made of multiple shelves above which the vials containing the product to be lyophilized are placed. Inside these shelves silicon oil flows and is heated or cooled in relation to the phase of the process that is in progress.

Nitrogen flows in the gaseous state in the condensing chamber coils and after having absorbed heat from the chamber, it comes out in the liquid state. The freeze-

Figure 1.3: The freeze-dryer machine.

dryer is subject to strong mechanical stresses due to high variations in pressure and temperature: this requires a continuous maintenance to keep the machine operating correctly. As the high cost of the freeze dried product, it is necessary to carry out studies to prevent machine malfunctions, in order to make maintenance operations more efficient, improve process planning and avoid failure that would have a greater economic impact. In addition to freeze-drying, other cycles are performed inside the machine, in order to wash (Cleaning In Place - CIP), sterilize (Steam In Place - SIP) and test the sealing of the freeze-dryer (Leak Test - LT).

1.3 Cleaning In Place (CIP)

The Cleaning In Place (CIP) process is used to wash the freeze-dryer. During the process, the machine is cleaned with purified water: chamber walls, shelves, condenser walls, and condenser plates are sprayed by nozzles. The water temperature is around $350°K$.

The sequence of phases of the CIP relies on the following mechanical elements:

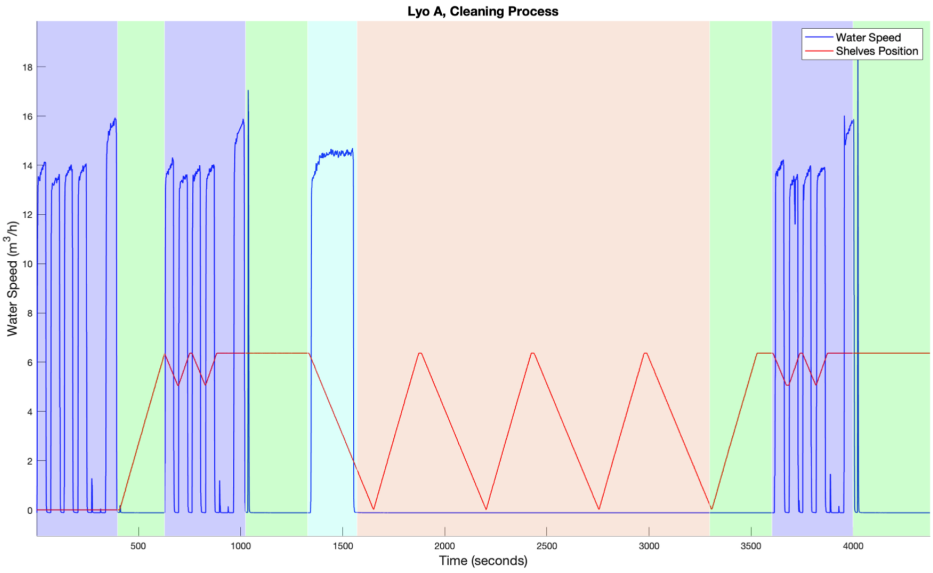


Figure 1.4: The CIP phases.

- the valves that allow the water to be poured into the machine: they are four for the chamber and one for the condenser;
- the valves that allow water to be drained: they are two for the chamber and one for the condenser;
- the piston that, thanks to the oil under pressure, moves the shelves in order to rinse the machine better.

Below we describe the phases of the CIP:

1. Before the cleaning process starts, the chamber and condenser must be vented to atmospheric pressure.
2. Condenser and chamber drain valve are opened.

3. Water is poured into the chamber: the four sprays inside the chamber are activated one by one. Then, a single spray pours water into the condenser. All sprays have an activation duration of 50 seconds.

4. When the condenser spray finishes pouring water, there is a waiting period to allow the water to flow out of the drains.

5. The activation sequence of the sprays is repeated a second time and during this phase machine's shelves are moved. The hydraulic pump is activated and hanks to a directly operated proportional valve, the oil flow is directed to move the piston alternately towards the up or down.

6. There is a waiting time for water drainage: during this time, the plates remain still.

7. Then the water drain valves are closed and the first chamber spray is activated: after a set time the spray is turned off and, at this point, the plates are immersed in water.

8. The shelves are moved several times up and down to allow the water to wash them.

9. Subsequently, the following phases take place: a new drainage, another phase of spraying during which the plates are moved and, finally, there a last drainage.

1.4 Leak Test (LT)

Due to large variations in temperature and pressure, freeze-dryer machines are subject to considerable mechanical stresses. In particular, cyclic large temperature variations cause thermal fatigue, due to repeated expansions and contractions of the material constituting the lyophilization chamber. This may cause microscopic cracks, especially in tubes and support structures. These cracks may cause a *leak*, that is an influx of gas into the drying chamber. Even very small leaks can contaminate the sterile product environment inside the chamber. In the event of a contamination, the entire

drug batch must be rejected. Since a single batch may contain thousands of product vials, freeze-dryer leakages are one of the most critical problems of the entire production chain of lyophilized drugs.

In order to monitor possible gas infiltrations, a common practice is to run a *leak detection test* before the beginning of the lyophization process. In this test, the lyophilization chamber is put under vacuum and, then, the trend of growing pressure over time is monitored. If the rise of pressure with respect to time is above a given threshold, it is concluded that the lyophization chamber is not correctly sealed and the machine is not used until the problem is solved.

A typical leak test measures the rise of pressure with time in the lyophilization chamber, after the air has been removed by a vacuum pump. More precisely, it consists of the following three phases:

1. *evacuation*: the freeze-dryer is evacuated by vacuum pumps until the chamber pressure is below 10 μbar;

2. *leak test preparation*: for a given time interval, the pressure is kept in the interval between 8 and 10 μbar. In order to achieve this behaviour, the vacuum pumps are switched on (if the pressure exceeds 10 μbar) or off (if it falls below 8 μbar). Note that, in this phase, the pressure rises mainly due to internal leaks;

3. *leak test*: the vacuum pumps remain switched off, and the freeze dryer is left in this state for a fixed amount of time (typically between 30 min and 1 hour).

In common practice, in order to quantitatively assess leaks, one measures the chamber pressure difference Δp (dimension: $[\mu\text{bar}]$) between the end and the beginning of the leak test. If V is the volume of the lyophilization chamber (dimension: $[l]$) and Δt is the time difference between the ending and initial instants of the leak test (dimension: $[s]$) , the overall *leak rate* can be estimated as (see [13])

$$Q_L = \frac{\Delta p V}{\Delta t}. \tag{1.1}$$

The unit of measure of the leak rate is typically $[\text{mbar l/s}]$. Namely, in an enclosed, evacuated vessel with a volume equal to 1 l, there is a leak rate of $Q_L =$

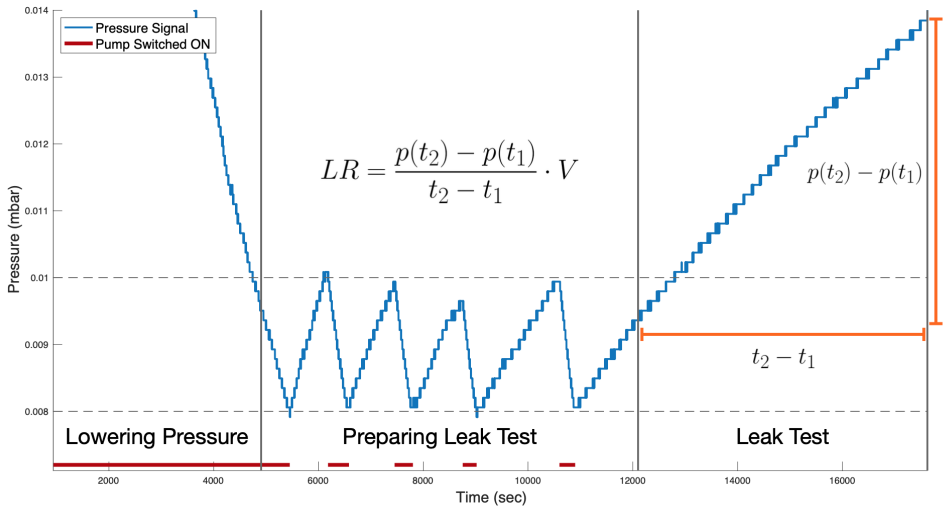


Figure 1.5: The three phases of Leak Test Process.

1 mbar l/s when the pressure rises by 1 mbar per second. In standard practice, if the measured leak rate is above a given threshold, then the machine is considered faulty and is stopped for diagnosis and repair. Even though Food and Drug Administration (FDA) does not provide clear guidelines for the choice of the leak rate threshold, common practice in the pharmaceutical industry is to use a value equal to 2×10^{-2} mbar l/s [14].

1.5 Data Transfer

These studies rely on a large amount of process data that have been historicized over the years by GSK. Thanks to FDA regulation, pharmaceutical companies are required to record all production data, so GSK has been historicizing process signals for years. This is instrumental to train algorithms that describe the behavior of the machine even without knowing a classical physical model. The tool used to analyze data and create forecasting models is MATLAB, a versatile and relatively simple tool to use both to run simulations and to create statistical models. Process data are stored in a Histo-

rian database. There is no method to directly transfer data from Historian to Matlab, so that it was necessary to use SQL Server Management Studio as communication bridge between the two softwares and compress the data.

Figure 1.6: Data transfer sequence.

The method to query Historian with SQL Server is to write a text string that is passed to Historian via OPENQUERY. The advantage of OPENQUERY is that it passes the supplied SQL statement as its second argument to the linked server you supply as its first argument, directly without modification. Once process data are inserted into a table, it was necessary to create a new table in SQL Server Management Studio in which the same but compressed data were inserted. Furthermore, our strategy was to convert the Tag-name related to the signals from character to whole numbers. Finally, the timestamp was converted to a numerical value: it was possible to do so because the data are logged once a second, so that it was sufficient to calculate the difference in seconds between the recording time of any value and the process start time. After the application of this method the data was transferred to Matlab thanks to the Database Toolbox functionalities. The data has been inserted in a three-dimensional matrix, whose coordinates are:

- the process cycle number;
- the number corresponding to the tag, which refers to a single signal sensor;
- the instant signal integer indicating the time, second by second.

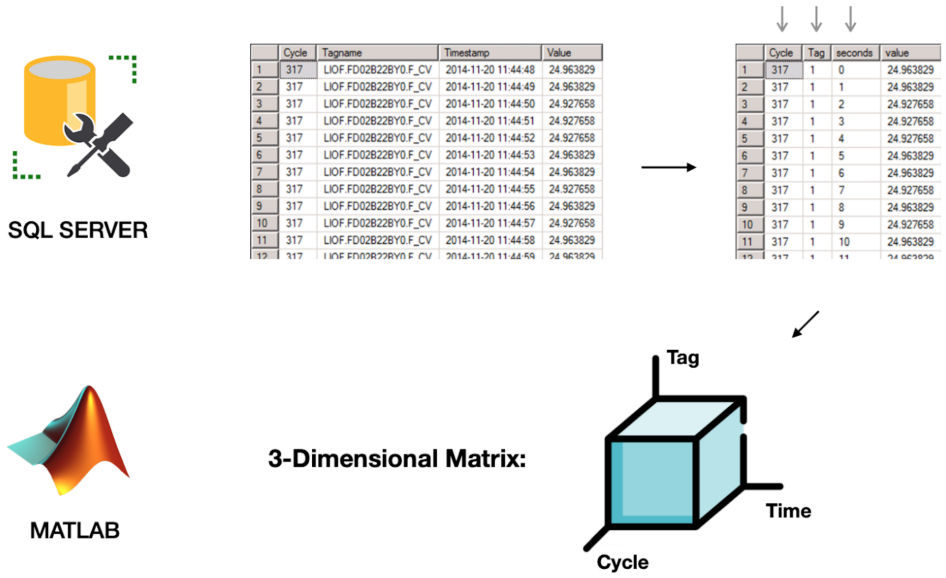
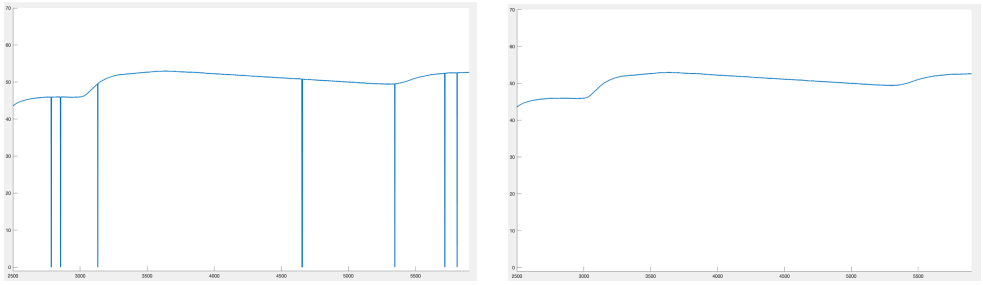


Figure 1.7: Strategy to transfer data from SQL Server to Matlab.

Only the values of the measured signals are inserted into the matrix. In a first analysis of the data, it was observed that in some moments all the signals were worth 0: this happened because Historian had a data loss during the acquisition. Even though this loss was very marginal (a cumulative maximum of 20 seconds per hour), it was necessary to eliminate those spikes. In order to do this, an algorithm was therefore written to automatically identify such data holes and replace them with linearly interpolated values.

To occupy less memory space there are various methodologies that eliminate repetitive data at successive instants, an example of compression strategies can be found in [15].



(a) (b)

Figure 1.8: Example of signal (a) before and (b) after the replace holes algorithm.

Chapter 2

Circuit Identification

2.1 Introduction

When it is required to simulate the behavior of a machine, one of the most common strategy is to use finite element methods. However, the lyophilizer is a very complex machine, and its behavior varies over time in relation to small modifications carried out on the machine (weldings, repairs, etc.), difficult to describe in a three-dimensional model. Various dynamical models of processes in engineering, physics or biology have the following form

$$\begin{aligned} G\dot{x}(t) &= Sx(t) + Bu(t) \\ y(t) &= Cx(t)\,, \end{aligned} \tag{2.1}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the input, $y(t) \in \mathbb{R}^p$ is the output. We assume that $S \in \mathbb{R}^{n\times n}$ is symmetric, $G \in \mathbb{R}^{n\times n}$ is diagonal and positive definite, while B and C have no special structure.

For instance, model (2.1) may represent a generic RC (resistive and capacitive) network, in which the components of $x \in \mathbb{R}^n$ are the node potentials, S is the admittance matrix and the diagonal elements of G are the nodes capacitances. As an example, consider the RC circuit represented in Fig. 2.1 and assume that the output is the potential of node 1. The corresponding model has form (2.1):

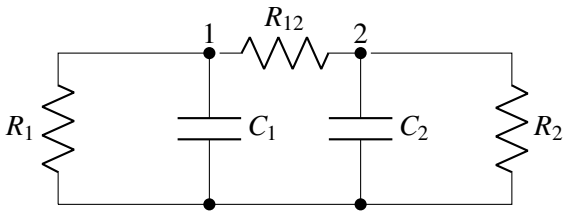


Figure 2.1: Example of an RC-circuit.

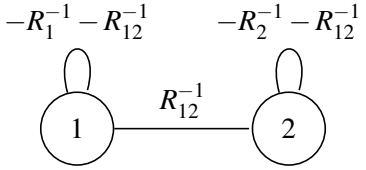


Figure 2.2: Graph representation of model associated with the circuit in Fig. 2.1.

$$\begin{aligned} \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix} \begin{pmatrix} \dot{v}_1(t) \\ \dot{v}_2(t) \end{pmatrix} \\ = \begin{pmatrix} -R_1^{-1}-R_{12}^{-1} & R_{12}^{-1} \\ R_{12}^{-1} & -R_2^{-1}-R_{12}^{-1} \end{pmatrix} \begin{pmatrix} v_1(t) \\ v_2(t) \end{pmatrix}, \\ y(t) = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} v_1(t) \\ v_2(t) \end{pmatrix}. \end{aligned} \tag{2.2}$$

Here, v_1 and v_2 represent the potentials of nodes 1 and 2. We can associate a weighted undirected graph to matrix S in (2.1) by considering S a weighted adjacency matrix. Namely, we define a node for each component of vector x and we define an edge between node i and node j if the entry of row i and column j of S is nonzero. The numerical value of the entry represents the edge weight. For instance, Fig. 2.2 represents the graph associated with the model of the circuit in Fig. 2.1. For simplicity, in the rest of the chapter we will omit self loops when representing the graph associated to the S matrix of a system in form (2.1).

Form (2.1) also models those systems that have the same mathematical representation of RC-circuits, such as thermal systems, physical network systems ([16]), dendritic structures ([17]), mammillary systems ([18]), or, more generally, diagonally symmetrizable compartmental systems (see [19]). Therefore, the RC-circuit model could also be used to simulate heat exchanges and pressure variations that occur inside a freeze-dryer. Note that also other linear circuit models (for instance, lithium battery models) can be put in form (2.1), for instance by passing to the Cauer form.

Consider a second system of form

$$\begin{aligned} \dot{z}(t) &= \hat{A}z(t) + \hat{B}u(t) \\ y(t) &= \hat{C}z(t)\,, \end{aligned} \tag{2.3}$$

in which $\hat{A}$, $\hat{B}$, $\hat{C}$ have the same dimensions of S, B, C. In this chapter we consider the following algebraic problem.

Problem 1. *Consider systems (2.1), (2.3), where matrices $\hat{A},\hat{B},\hat{C},B,C$ are given. Find, if possible, an invertible matrix T, a symmetric matrix S and a strictly positive diagonal matrix G such that*

$$\begin{cases} T^{-1}\hat{A}T = G^{-1}S \\ \hat{C}T = C \\ T^{-1}\hat{B} = G^{-1}B\,. \end{cases} \tag{2.4}$$

In other words, we are looking for a state-space transformation of form $z = Tx$ and suitable matrices S and G such that (2.3) takes on form (2.1).

Problem 1 can be interpreted as a structured identification one. Namely, system (2.3) represents an identified black-box model, obtained from experimental data with standard techniques, for instance state space identification methods. Our aim is to check if this system, after state transformation $z = Tx$, can be given the form of model (2.1), and, if this is possible, find such a transformation. The structural requirements imposed in Problem 1 are the following ones

- The transformed system matrix $T^{-1}\hat{A}T$ must be the product of a positive diagonal and a symmetric matrix.

- The input and output matrices are assigned.

The first requirement ensures that the transformed system matrix can be written as the product of a diagonal matrix (containing the inverse capacities, in the case of RC-circuits) and a symmetric matrix (the admittance matrix in the case of RC-circuits). The reduction to the RC-model form (2.1), gives a meaningful description of the system. In particular, matrix S can be directly associated with an undirected graph, which represents the interactions among the states.

The second requirement is due to the fact that, in some cases, the input matrix B or the output matrix C may be known from structural properties of the system at hand. For instance, in an RC-circuit we may be able to measure the potentials of the first n_1 nodes, while the potentials of the remaining ones are not accessible. In this case, C would correspond to the projection matrix on the first n_1 components. Note that we may not have this requirement or know only one among B and C. In this case, the second, the third condition in (2.4), or both could be omitted.

Furthermore, we consider a more restrictive version of Problem 1, based on the observation that, in various dynamical models, matrix S in (2.1) is Metzler, that is, all its off-diagonal entries are non-negative. For instance, in RC-networks, off-diagonal entries correspond to the values of the resistances connecting the network nodes (see (2.2)). This suggests the following formulation.

Problem 2. *Solve Problem 1 with the additional requirement that S is Metzler.*

Further, if Problem 2 has multiple solutions, one may minimize the number of nonzero components of S, that is, minimize $\|S\|_0$, the so-called zero-norm of S. This follows the principle of parsimony of finding the simplest model of form 1 that fits the data. This leads to the following additional problem.

Problem 3. *Find the solutions of Problem 2 in which* $\|S\|_0$ *is minimum.*

For instance, in an RC-circuit, the nonzero elements of S represent the resistive connections between the nodes. Hence, minimizing $\|S\|_0$ corresponds to reducing the overall number of resistive components.

2.1.1 Statement of contribution

Regarding Problem 1, we will show that it is convenient to parameterize the set of solutions as $T = PQ\sqrt{G}$, where P, Q is the polar decomposition of $T\sqrt{G}^{-1}$. In particular,

- Proposition 4 shows that P and G are the solution of a problem with a reduced number of unknowns. Essentially, this result leverages the symmetry of S. Furthermore, if either the second or the third condition in (2.4) is missing, P and G are the solution of a convex problem. We will consider this last case in more detail. We will mention that, in many cases, the solution P, G is unique up to a scaling factor.
- Proposition 7 parameterizes all solutions of Q corresponding to each solution P, G.

These results can also be used to reduce the number of unknowns in Problems 2 and 3. Anyway, these last two problems are more difficult than Problem 1. To solve them, we will resort to general local search algorithms.

2.1.2 Comparison with literature

A problem similar in structure to Problem 1, but more general, consists in solving the following system with respect to unknown parameter vector θ

$$\begin{cases} T^{-1}\hat{A}T = A(\theta) \\ \hat{C}T = C(\theta) \\ T^{-1}\hat{B} = B(\theta)\,, \end{cases} \tag{2.5}$$

in which matrices A, B, C depend on θ. Problem 1 may be considered a special case of (2.5), in which C and B do not depend on θ and the only constraint on A is symmetry.

Problem (2.5) has been extensively studied in recent literature. A common approach consists in two phases. First, one finds a black-box model (this is in general an easy one, for instance, resorting to subspace-based methods). Second, one finds a

coordinate transformation T and a parameter vector θ that satisfy (2.5). In the general case, this second step is not an easy task. In fact, even assuming, as commonly done, that A, B, C are a linear function of θ, Problem (2.5) is bilinear and, thus, nonconvex.

This approach has been introduced in [20] and studied in various subsequent works. For instance, [21] studied the problem of parameter initialization.
Works [22], [23], [24], [25] present different numerical approaches for the solution.

With respect to these general approaches, this work leverages the special structure imposed by Problem (2.1), in particular the symmetry of S, to obtain specific properties (see Propositions 4 and 7) that do not apply to the more general case (2.5). As far as we know, the results presented in these two propositions are new, perhaps also due to the specificity of the problem discussed in this chapter.

The literature on network reconstruction has mainly been concerned on determining conditions ensuring network identifiability. Loosely speaking, this problem consists in assessing whether the network structure can be uniquely identified from input/output measurements. Depending on the network model and the assumed prior knowledge, different formalizations of this problem can arise. For instance, in [26], the authors consider systems in state-space form, where the network structure is given by the system matrix A, for which no prior information is available. Reference [27] shows that, if no prior information is known on matrices A and B, then the presence of unmeasured nodes is a sufficient condition for network unidentifiability. On the other hand, it is shown in [28] that, even if there are unmeasured nodes, by assuming that the boolean structure of the network is known a priori, it is still possible to guarantee identifiability by appropriately selecting the measured nodes. In [29], identifiability is discussed for autonomous dynamical networks where all nodes are measured and matrix A is constrained to be a symmetric, an adjacency or a Laplacian matrix. For each of these cases different identifiability conditions are derived.

The results on identifiability of [29, 28] are the one that most naturally apply to our case. However, we take a different approach. Rather than studying the conditions which guarantee network identifiability for RC-networks, we take for granted that unidentifiability might arise (especially in the recurrent situation in which the number of measured nodes is not sufficiently high) and present a complete characterization

of all possible RC-circuits compatible to the available I/O measurements.

The problem of designing a network reconstruction algorithm is different from the one of network identifiability. In this work, the network reconstruction problem is casted into a structured identification one, for which a solution algorithm is proposed. Other approaches of network reconstruction are available in literature. For instance, in [30] an adaptive control strategy allows, as a side result, to recover the network structure for controllable systems. For systems whose state-matrix is constrained to be a Laplacian matrix (for example, this is the case for consensus networks or even RC-circuits having $G = I$ and no grounding resistances), the so-called node knockout procedure for network identification is presented in [31]. In [29], a network identification algorithm based on Lyapunov equations is presented for autonomous dynamical networks without unmeasured nodes.

With respect to these approaches, our work does not require controllability which is instead required in [30] and might also be needed in [31] to achieve the node knockout. With respect to [29], nor we constrain our discussion to autonomous systems and neither we assume to measure all nodes of the network. Nevertheless, we discuss in detail the autonomous case in our examples. In this sense the research in [29] is the most similar to this work.

Notation: Matrix $A \in \mathbb{R}^{n\times n}$ is diagonalizable if there exist a nonsingular matrix V and a diagonal matrix Λ such that $AV = V\Lambda$. The columns of V are the right eigenvectors of A. Set $W = V^{-1}$, then we have that $WA = V^{-1}A = \Lambda V^{-1} = \Lambda W$, which shows that the rows of W are left eigenvectors of A. We also write that (V, Λ, W) is a diagonalization of A. We denote the orthogonal group over $\mathbb{R}$ by

$$O(n) = \{M \in \mathbb{R}^{n\times n} \text{ such that } M \text{ is invertible and } M^T M = I\},$$

that is the set of real orthonormal matrices of dimension n. Given a subspace $V \subset \mathbb{R}^n$, we will denote by $V^\perp$ its orthogonal subspace.

2.2 Discussion of Problem 1

The following proposition presents a necessary condition for the feasibility of Problem 1.

Proposition 1. *Problem 1 has a solution only if $\hat{A}$ is diagonalizable and has real eigenvalues.*

Proof. Assume that Problem 1 has a solution. By the first of (2.4) it follows that $G^{1/2}T^{-1}\hat{A}TG^{-1/2} = G^{-1/2}SG^{-1/2}$. Note that this last matrix is symmetric, hence $\hat{A}$ is diagonalizable and has real eigenvalues, being similar to a symmetric matrix. □

Due to the previous proposition, we will make this assumption throughout the chapter.

Assumption 1. *$\hat{A}$ is diagonalizable and has real eigenvalues.*

Remark 1. *By Proposition 1, if Assumption 1 does not hold for $\hat{A}$, then Problem 1 does not have a solution. This means that the identified system has not the structure of an RC-network.*

We will parameterize the set of possible solutions T of Problem 1 as

$$T = PQ\sqrt{G}, \tag{2.6}$$

where P is a symmetric and positive definite matrix and $Q \in O(n)$. Note that PQ corresponds to the left polar decomposition of $T\sqrt{G}^{-1}$, which is unique, being $T\sqrt{G}^{-1}$ invertible. In particular, P corresponds to a scaling and Q to a rotation or reflection, further $P = \sqrt{TG^{-1}T^T}$. In the following, we will show that parameterization (2.6) is convenient since couple P, G can be found separately from Q. As a first step, the following proposition shows that the feasibility of Problem 1 is equivalent to the existence of a solution of an equation independent of Q. The proof is presented in the Appendix.

Proposition 2. *Problem 1 has a solution T,S,G if and only if there exist positive definite matrices M and G, with G diagonal, such that*

$$\begin{cases} \hat{A}M = M\hat{A}^T \\ \hat{C}M\hat{C}^T = CG^{-1}C^T \\ \hat{B}^T M^{-1}\hat{B} = B^T G^{-1}B \\ \hat{C}\hat{B} = CG^{-1}B. \end{cases} \tag{2.7}$$

Moreover, $M = P^2$, where P is defined as in (2.6).

Note that, with respect to form (2.6), equation (2.7) contains variables $M = P^2$, G, but does not contain Q. The structure of (2.7) can be simplified by diagonalizing $\hat{A}$. In particular, if (V,Λ,W) is a diagonalization of $\hat{A}$ (i.e., $\hat{A} = V\Lambda W$), the following proposition shows that the first equation in (2.7) can be substituted with $M = VDV^T$, where D is a matrix that commutes with Λ (i.e., $D\Lambda = \Lambda D$).

Proposition 3. *Let (V,Λ,W) be a diagonalization of $\hat{A}$ and let $M \in \mathbb{R}^{n\times n}$, then the following statements are equivalent*

i) $\hat{A}M = M\hat{A}^T$

ii) there exists a matrix D, that commutes with Λ, such that $M = VDV^T$.

Proof. i) $\Rightarrow$ ii) Substituting $\hat{A} = V\Lambda V^{-1}$ in i) we obtain $V\Lambda V^{-1}M = MV^{-T}\Lambda V^T$, which implies $\Lambda V^{-1}MV^{-T} = V^{-1}MV^{-T}\Lambda$. Set $D = V^{-1}MV^{-T}$, then $\Lambda D = D\Lambda$ and $M = VDV^T$.

ii) $\Rightarrow$ i) Let D be any matrix such that $\Lambda D = D\Lambda$ and set $M = VDV^T$. Then $\hat{A}M = \hat{A}VDV^T = V\Lambda V^{-1}VDV^T = V\Lambda DV^T = VD\Lambda V^T = VDV^TV^{-T}\Lambda V^T = M\hat{A}^T$.

□

Remark 2. *The requirement that D commutes with Λ reduces the actual number of unknown entries of D. In fact, setting $\Lambda = diag\,\{\lambda_1,\ldots,\lambda_n\}$, $D = (d_{ij})$ commutes with Λ if and only if*

$$d_{ij} = 0, \text{ for all } i,j \text{ such that } \lambda_i \neq \lambda_j.$$

For instance, if all eigenvalues of $\hat{A}$ are distinct, D must be diagonal. In the general case, D has a block-diagonal structure.

Combining the results of Propositions 2 and 3, we derive the following result.

Proposition 4. *Problem 1 has a solution T, S, G if and only if there exist a symmetric matrix D and a diagonal matrix G such that*

$$\begin{cases} D > 0 \\ \Lambda D = D\Lambda \\ G > 0 \\ \hat{C}VDV^T\hat{C}^T = CG^{-1}C^T \\ \hat{B}^T W^T D^{-1} W\hat{B} = B^T G^{-1} B \\ \hat{C}\hat{B} = CG^{-1}B. \end{cases} \tag{2.8}$$

Moreover, $P = \sqrt{VDV^T}$, where P is defined in (2.6).

In Problem (2.8) the optimization variables are D and G. This problem is non-convex, since variable D appears in it together with its inverse.

Remark 3. *If the third condition is not present in (2.4), the third and fourth conditions disappear from* (2.7) *and, setting $H = G^{-1}$, Problem* (2.8) *reduces to the convex one:*

$$\begin{cases} D > 0 \\ \Lambda D = D\Lambda \\ H > 0 \\ \hat{C}VDV^T\hat{C}^T = CHC^T. \end{cases} \tag{2.9}$$

If the third condition is not present in (2.4), we do not impose any structural requirement on matrix B. Note that the solution of (2.9) *is not unique. In fact, if D, H is a solution of* (2.9)*, any scaling αD, αH, with $\alpha > 0$ is still a solution. In particular, if $\hat{A}$ has distinct eigenvalues, D must be diagonal and Problem* (2.9) *reduces to finding positive diagonal matrices D, H such that*

$$\hat{C}VDV^T\hat{C}^T = CHC^T. \tag{2.10}$$

The set of all solutions of (2.9) *corresponds to a polyhedral cone and can be expressed as a conical combination of a finite set of vertices (by Weyl-Minkowski Theorem), that is, we can find vectors $v_1, \ldots, v_l$ (called generators) such that the set of all*

solutions of (2.9) *is*

$$\{\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_l v_l, \quad \alpha_1, \alpha_2, \ldots, \alpha_l > 0\}. \tag{2.11}$$

These considerations also hold if the second condition of (2.4) is not present.

Remark 4. *We present an intuitive discussion on the number of distinct solutions of* (2.9)*. Assuming $\hat{A}$ has distinct eigenvalues, so that D is diagonal, Problem* (2.9) *reduces to finding positive diagonal matrices D, H such that* (2.10) *holds. The solutions of* (2.10) *are represented by vector $x = [diag\, D, diag\, H]$, that contains the elements on the diagonal of the two matrices D and H.*

Note that, since the left and right-hand sides of (2.10) *are symmetric $p \times p$ matrices, equation* (2.10) *corresponds to a set of $n_e = p(p+1)/2$ equations. We have $n_u = 2n$ unknown terms (the diagonal elements of D and H). Hence, for a generic choice of problem data (i.e., matrices $\hat{A}$, C, $\hat{C}$ are randomly selected) we have a solution consisting of a unique ray (that is, unique up to scaling) if $n_e \geq n_u - 1$, that is $p \geq (\sqrt{16n-7}-1)/2$, where the -1 term is due to the fact that a ray has dimension one. However, if the problem data are not generic, we may have multiple solutions even if this condition is satisfied. For instance, if C is a projection on the first p components, then term CHC^T does not contain the last $n-p$ elements of the diagonal of H. Hence, these are left undetermined and can be chosen as arbitrary positive values. In this case, the number of remaining unknowns is $n_u = n + p$, so that, if remaining parameters $\hat{A}$, $\hat{C}$ are generic, we have only one solution for D and for the first p elements of the diagonal of H (up to a scaling factor) if $n_e \geq n_u - 1$, that is*

$$p \geq \frac{\sqrt{8n-7}+1}{2}. \tag{2.12}$$

These considerations intuitively justify the fact that, in generic cases, if p is sufficiently high, equation (2.9) *has only one solution (up to a scaling factor). Our numerical experiments confirm this fact, however, we do not have a formal proof.*

Example 1. *Consider the RC circuit depicted in Fig. 2.3. If $x(t) \in \mathbb{R}^{4\times 1}$ represents the node potentials, $C_1 = C_2 = C_3 = C_4 = 1$, $R_1 = R_4 = R_2 = 1$, $R_3 = 2$, the model of*

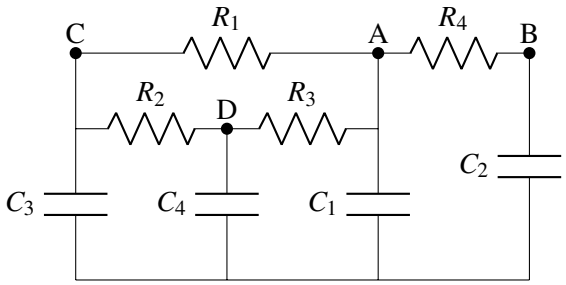


Figure 2.3: Circuit used in Example 1.

the system corresponds to (2.1) with $G = I$ and

$$S = \begin{pmatrix} -4 & 1 & 1 & 2 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -2 & 1 \\ 2 & 0 & 1 & -3 \end{pmatrix}. \tag{2.13}$$

We assume that the system is autonomous (i.e., B is not present) and the outputs are the potentials of the first three nodes, that is

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Suppose that we do not known matrices G and S, but we do know matrix C, since our output consists in the potentials of the first three nodes. Assume also that, by using standard identification techniques (for instance subspace-based methods), we are able to identify a state-space model in form

$$\begin{aligned} \dot{z}(t) &= \hat{A}z(t) \\ y(t) &= \hat{C}z(t), \end{aligned}$$

with

$$\hat{A} = \begin{pmatrix} -10 & -4 & -23 & 4 \\ 1 & -1 & 3 & -1 \\ 3 & 1 & 7 & -2 \\ 1 & -1 & 3 & -6 \end{pmatrix}, \hat{C} = \begin{pmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix}.$$

Then, solving Problem 1 consists in finding a coordinate transformation $z = Tx$ and matrices G and S such that (2.4) *holds. In particular, matrix S is a key piece of information since it allows to reconstruct the network structure. Note that Assumption 1 is satisfied. We use parameterization* (2.6) *and apply Proposition 4 to find matrices P and G. Since we do not have any requirement on input matrix B, we have to solve the convex Problem* (2.9)*. Furthermore, since $\hat{A}$ has distinct eigenvalues, by Remark 2, D must be diagonal, so that Problem* (2.9) *reduces to solving* (2.10)*, in which the variables are the diagonal matrices D and H. As previously noted, the solution to this problem is not unique, since, if D and H are a solution of* (2.10) *for any $\alpha \in \mathbb{R}$, with $\alpha > 0$, also $\alpha D, \alpha H$ is a solution of* (2.10)*. The solution set has form* (2.11)*, in particular there are two generators, so that the set of all solutions is given by*

$$x = \{v_1\alpha_1, v_2\alpha_2\}, \tag{2.14}$$

with $v_1 = \begin{pmatrix} 1.98 & 11.334 & 7.5 & 3.186 & 1 & 1 & 1 & 0 \end{pmatrix}$,
$v_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$.

All solutions are parameterized by positive parameters α_1, α_2. In particular, α_2 is related to the fact that we can not know the capacitance of the unmeasured node. Moreover, D depends only on α_1, so that the P component of the solution T is unique apart from an unknown scaling factor.

At this point, we compute the rotation component Q of parameterization (2.6). Let $P = \sqrt{M}, G$ be solutions of (2.7) and let T be defined as in (2.6). Then, substituting T in the second and third of (2.4), we obtain the following conditions

$$\begin{aligned} &\hat{C}PQ\sqrt{G} = C \\ &\sqrt{G}Q^T P^{-1}\hat{B} = G^{-1}B. \end{aligned} \tag{2.15}$$

These conditions can be rewritten as $QZ = W$, where

$$W = \begin{pmatrix} P\hat{C}^T & P^{-1}\hat{B} \end{pmatrix}, Z = \begin{pmatrix} \sqrt{G}^{-1}C^T & \sqrt{G}^{-1}B \end{pmatrix}. \tag{2.16}$$

Note that, by the second, the third and the fouth of (2.7), $Z^TZ = W^TW$, so that Z and W have the same rank.

In the following computations, it is convenient to assume that Z and W are full column-rank since, in this case, their left inverses Z^+ and W^+ are well-defined. If Z and W are not full column-rank, it is possible to reduce them to full column-rank matrices by right multiplying them by a suitable matrix L, as a consequence of the following simple algebraic property.

Proposition 5. *Let n,m,r be positive natural numbers, $Z,W \in \mathbb{R}^{n\times m}$ with $Z^TZ = W^TW$, $Q \in \mathbb{R}^{n\times n}$ and let $L \in \mathbb{R}^{m\times r}$ be such that ZL is full column-rank and Im $ZL =$ Im Z, then the following statements are equivalent:*

i) $QZ = W$

ii) $QZL = WL$.

Proof. i) $\Rightarrow$ ii). Obvious.

ii) $\Rightarrow$ i). By Proposition 8, being $Z^TZ = W^TW$, there exists $\hat{Q} \in O(n)$ such that $\hat{Q}Z = W$. Since Im $ZL =$ Im Z, there exists $M \in \mathbb{R}^{r\times m}$ such that $Z = ZLM$. Then, $WLM = \hat{Q}ZLM = \hat{Q}Z = W$. Then i) is obtained by right-multiplying ii) by M. □

In the following, we will assume that Z is full column-rank. In fact, if this is not the case, it is sufficient to pick L such that ZL is full column-rank and to redefine $Z = ZL, W = WL$.

The following proposition shows that, if (2.7) holds, there always exists an orthonormal matrix Q that satisfies $QZ = W$. We can distinguish two cases. First, in the trivial case in which rank $Z = n$ (that is, Z is full row-rank), the solution is unique, as shown in the following Proposition.

Proposition 6. *Let P,G be a solution of* (2.7)*, let W,Z defined as in* (2.16) *be such that rank $W = n$. Set $T = PQ\sqrt{G}$. Then, T is a solution of Problem 1 if and only if*

$$Q = WZ^{-1}. \tag{2.17}$$

Proof. ($\Rightarrow$) By assumption $QZ = W$ has a solution, then, since Z is full rank, it is invertible and $Q = WZ^{-1}$. ($\Leftarrow$) Assume that Q is given by (2.17) and set $T = PQ\sqrt{G}$, where P and G are a solution of (2.7). Note that (2.7) implies that $\hat{A}P^2 = P^2\hat{A}^T$, that is $P^{-1}\hat{A}P = P\hat{A}^TP^{-1}$, so that $P^{-1}\hat{A}P$ is symmetric. Then, also matrix

$$GT^{-1}\hat{A}T = G\sqrt{G^{-1}}Q^TP^{-1}\hat{A}PQ\sqrt{G} = \sqrt{G}Q^TP^{-1}\hat{A}PQ\sqrt{G}$$

is symmetric, proving the first of (2.4). Moreover, $QZ = WZ^{-1}Z = W$, which implies conditions (2.15). □

If rank $Z < n$, the solution Q of $QZ = W$ is not unique, since, if Q satisfies $QZ = W$ and $\hat{Q}$ is any orthonormal matrix such that $\hat{Q}Z = Z$, then also $Q\hat{Q}Z = W$. In fact, the following proposition shows that the set of possible solutions Q is parameterized by $O(n - \text{rank } Z)$.

Proposition 7. *Let P, G be a solution of* (2.7)*, let W, Z be defined as in* (2.16) *with rank $Z < n$ and let $\bar{W}$, $\bar{Z}$ be orthonormal matrices such that Im $\bar{W} = (Im\ W)^{\perp}$, Im $\bar{Z} = (Im\ Z)^{\perp}$. Set $T = PQ\sqrt{G}$. Then, T is a solution of Problem 1 if and only if*

$$Q = WZ^{+} + \bar{W}\bar{U}\bar{Z}^T, \tag{2.18}$$

where $\bar{U} \in O(n - rank\ Z)$.

Proof. ($\Rightarrow$) Let P, G be a solution of (2.7) and set $T = PQ\sqrt{G}$. Then, Q satisfies (2.15) or, equivalently, $QZ = W$. Then, the thesis follows from Proposition 9.

($\Leftarrow$) It is the same as the proof of the necessity of Proposition (6), with the difference that, in this case, $QZ = WZ^{+}Z + \bar{W}\bar{Q}\bar{Z}^TZ = W$, which implies conditions (2.15). □

Example 1 (continued)**.** *We consider again Example 1, and we choose a particular solution for P and G by setting $\alpha_1 = \alpha_2 = 1$ in* (2.14)*. We make this choice in order to have $G = I$. We apply Proposition* (7) *to find the rotation component Q. In this case, B*

is not present and $\dim(Im\, W)^{\perp} = 1$. *For this reason* $\bar{U} \in O(1)$, *since* $O(1) = \{-1, 1\}$, *there are two possible solutions for Q, given by*

$$Q_1 = WZ^{+} + \bar{W}\bar{Z}^{T}, Q_2 = WZ^{+} - \bar{W}\bar{Z}^{T}.$$

The corresponding transformation matrices T_1, T_2 *are obtained from* (2.6) *and the symmetric part S of* (2.4) *by relation* $S_i = GT_i^{-1}\hat{A}T_i$, $i = 1, 2$, *that is*

$$S_1 = \begin{pmatrix} -4 & 1 & 1 & 2 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -2 & 1 \\ 2 & 0 & 1 & -3 \end{pmatrix}, S_2 = \begin{pmatrix} -4 & 1 & 1 & -2 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -2 & -1 \\ -2 & 0 & -1 & -3 \end{pmatrix}.$$

These are all the solutions of Problem 1 that correspond to the chosen values for P and G. Note that only S_1 *is Metzler, so that it is the only solution of Problem 2, moreover* $S_1 = S$, *where S is in* (2.13). *Figures 2.4 and 2.5 represent the graphs associated with matrices* S_1 *and* S_2. *In these and in next graph figures, red nodes denote unmeasured outputs.*

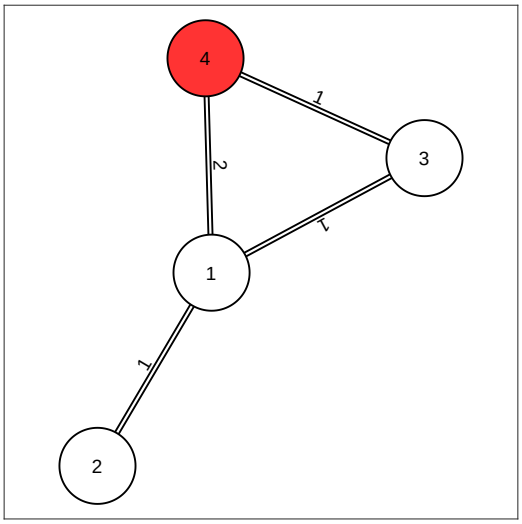


Figure 2.4: Graph representation of S_1

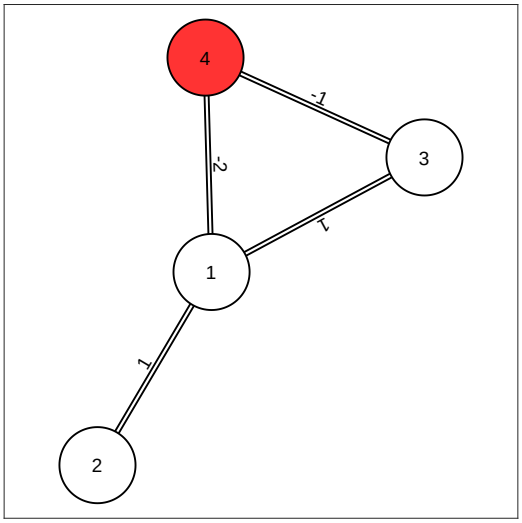


Figure 2.5: Graph representation of S_2

2.3 Discussion of Problems 2 and 3

Proposition 7 shows that, in the general case, Problem 1 has multiple solutions. We introduced Problems 2 and 3 in order to find specific solutions that satisfy additional properties. Consider a solution of form (2.6) and assume that P and G are fully known. Then, Problem 2 consists in finding an orthonormal matrix Q such that $T^{-1}\hat{A}T$ is Metzler, or equivalently, finding an orthonormal matrix $\bar{U}$ that satisfies the following equation

$$\begin{aligned} &\left(\sqrt{G}^{-1}Q^T P^{-1}\hat{A}PQ\sqrt{G}\right)_{i,j} \geq 0, i \neq j \\ &Q = WZ^+ + \bar{W}\bar{U}\bar{Z}^T \end{aligned} . \tag{2.19}$$

Note that Problem (2.19) is non-convex due to the orthonormality constraint on $\bar{U}$ (i.e., $\bar{U}^T\bar{U} = I$). Anyway, because of Proposition 7, the dimension of $\bar{U}$ may be small, so that, in some cases, solving (2.19) can still be a simple task.

Problem 3 adds the requirement of minimizing $\|S\|_0$, or, equivalently, minimizing

$\|T^{-1}\hat{A}T\|_0$. Since the minimization of the zero-norm is a difficult task, as commonly done (see, for instance, [32]), one can use the 1-norm as a sparsity-promoting objective function, obtaining the following problem:

$$\min_{\bar{U}} \left\| \sqrt{G}^{-1} Q^T P^{-1} \hat{A} P Q \sqrt{G} \right\|_1 \text{ such that (2.19) holds.} \tag{2.20}$$

We can rewrite this problem more explicitly as

$$\begin{aligned} &\min_{\bar{U}} \left\| \sqrt{G}^{-1} Q^T P^{-1} \hat{A} P Q \sqrt{G} \right\|_1 \\ &\text{such that} \\ &\left(\sqrt{G}^{-1} Q^T P^{-1} \hat{A} P Q \sqrt{G} \right)_{i,j} \geq 0, i \neq j \\ &Q = W Z^+ + \bar{W} \bar{U} \bar{Z}^T \\ &\bar{U} \bar{U}^T = I. \end{aligned} \tag{2.21}$$

2.3.1 Overall algorithm for Problem 3

Leveraging Proposition 4, we can formulate the following algorithm for solving Problem 3. Here the problem data are the identified model $\hat{A}$, $\hat{B}$, $\hat{C}$ and the required input and output matrices B, C. The final output is given by matrices P, G and $\bar{U}$, that give transformation T by (2.6) and (2.18).

- Solve Problem (2.8) in order to find a solution P, G. Note, that, by Remark 4, in many cases, this solution is unique up to a scaling factor. As said, if the second or thid condition in Problem 1 is not present, Problem (2.8) is convex and can be solved with standard solvers for this class of problems, such as GUROBI or CPLEX, as we did in our tests. Otherwise, a nonlinear solver such as SNOPT or IPOPT can be used.

- Solve Problem (2.21) using a nonlinear local search algorithm (we used IPOPT in our simulations). In our tests we used different randomly generated initial conditions $\bar{U}_0$ for $\bar{U}$ and selected the best solution.

Some remarks are in order on the choice of the initial conditions $\bar{U}_0$ for $\bar{U}$. Note that $O(n)$ has two connected components given by $\{e^S A, S \in \mathbb{R}^{n\times n} : S = -S^T, A \in$

$\{I, M\}\}$, where M is the diagonal matrix with all ones on the diagonal apart from a term -1 on the first element and S is a skew-symmetric matrix. This comes from the facts the set of skew-symmetric matrices is the Lie algebra of $O(n)$ and that I and M belong to each of the two separate connected components of $O(n)$. Hence, we can generate a random initial guess $\bar{U}_0$ for $\bar{U}$ by setting $\bar{U} = e^S A$, where S is a random skew-symmetric matrix and A is randomly chosen between I and M.

2.3.2 Case of data affected by noise

Real input and output data are affected by noise. This causes an error in the identification of A, B, C matrices. Because of this, Problem 1 and equations (2.8) may not have an exact solution. To take this fact into account, Problem (2.8) may be substituted with the following relaxed one:

$$\begin{cases} \min \|\hat{C}VDV^T\hat{C}^T - CG^{-1}C^T\|^2 + \|\hat{B}^T W^T D^{-1} W\hat{B} - B^T G^{-1} B\|^2 + \|\hat{C}\hat{B} - CG^{-1}B\|^2 \\ \text{subject to} \\ D > 0 \\ \Lambda D = D\Lambda \\ G > 0 \end{cases} . \tag{2.22}$$

Note that if the second or third conditions (the requirement on B or C) are not present in (2.4), Problem (2.22) becomes a convex one. In fact if, for instance, the third conditions is missing, by setting $H = G^{-1}$ the objective function reduces to the convex one $\|\hat{C}VDV^T\hat{C}^T - CHC^T\|^2$.

2.4 Examples

In this section, we consider some examples of larger dimension. We considered instances of Problem 1, obtained by randomly generating some autonomous systems in form (2.1), we call these systems the "true" systems. Since the systems are autonomous, we did not consider condition $T^{-1}\hat{B} = G^{-1}B$ in Problem 1. This simplifies

the computation of matrices P and G of form (2.6), since we can obtain them by solving convex Problem (2.9). To generate the problem instances, we used the following procedure. Given a number of states n, we set $G = I$ and $S = -\mathscr{I} K \mathscr{I}^T$, where $\mathscr{I}$ is the incidence matrix of a randomly generated connected graph of n vertices and K is a diagonal matrix of randomly generated conductances (with integer values). Then, we computed a random transformation matrix $\tilde{T}$ and set $\hat{A} = \tilde{T}^{-1} A \tilde{T}$, $\hat{C} = C\tilde{T}$. Given a number of outputs m, we considered as output matrix $C \in \mathbb{R}^{m \times n}$ the projection matrix on the first m state components. In all examples, $\hat{A}$ is diagonalizable, with distinct eigenvalues, and condition (2.12) is satisfied. Hence, for the generic case, P and the first m component along the diagonal of G have only one solution, up to a positive scaling factor. The remaining elements of diagonal matrix G are undetermined, since they do not appear in Problem (2.10). For simplicity, we chose the scaling factor such that the reconstructed G is the identity. We solved Problem 3 with the algorithm presented in Section 2.3.1 and computed the corresponding transformation matrix T and the reconstructed matrix S as $\hat{S} = GT^{-1}\hat{A}T$. Then, we compared matrix S of the true system with the reconstructed one $\hat{S}$. Note that, for any solution G and P, by Proposition (7), there are multiple choices of Q that solve Problem 1. Hence, in general, S is going to be different from $\hat{S}$. We considered the following two cases.

2.4.1 Case 1: $n = 10$, $m = 8$

In this example, we do not measure the potential of the last 2 nodes, that is, matrix C in (2.4) is the projection on the first 8 nodes. Generically (see Remark 4), P and G are unique (up to a positive scaling), apart from the last two components of the diagonal of G, that are undetermined. By Proposition (7), since $\dim \ker C = 2$, the component Q of (2.6) has multiple solutions, parameterized by $\bar{U} \in O(2)$. The algorithm in Section 2.3.1 allows finding one among such solutions. Fig. 2.6 is the graph associated with S while Fig. 2.7 is the one associated with the reconstructed $\hat{S}$. Note that the two graphs are similar but not identical. That is, at the end of our procedure, we found a reconstructed system of form (2.1) which solves Problem 2 (and, approximately, Problem 3), but is different from the true system. This is unavoidable since, by Proposition (7), there are multiple systems that solve Problem 1 and, in general,

there may be multiple solutions also of Problems 2 and 3.

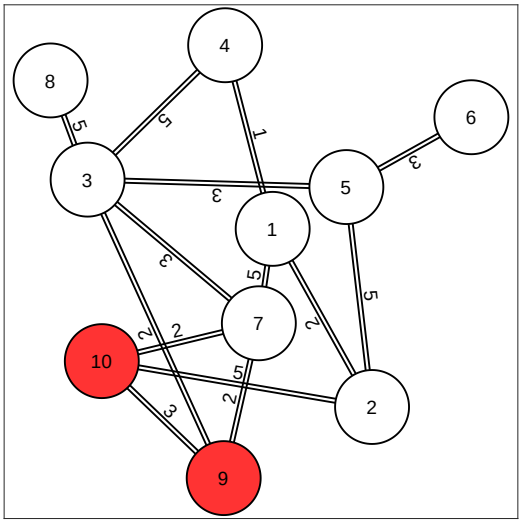


Figure 2.6: Case 1: true connections

2.4.2 Case 2: $n = 12$, $m = 6$

In this example, we measure the potential of 6 nodes out of 12. Again, Fig. 2.8 refers to the true matrix S while Fig. 2.9 refers to the reconstructed $\hat{S}$. In this case, Q has multiple solutions parameterized by $\bar{U} \in O(6)$. Again, the reconstructed matrix is different from the true one, namely, at the end of our procedure, we found one of the multiple solutions that solve Problem 2.

2.5 Conclusion

Resistive-capacitive (RC) networks are used to model various systems in engineering, physics or biology. In this chapter, we considered in detail the structured identification Problem 1 and found its complete solution set. A key element of the proposed approach is the use of the polar decomposition (2.6). In particular, we noticed

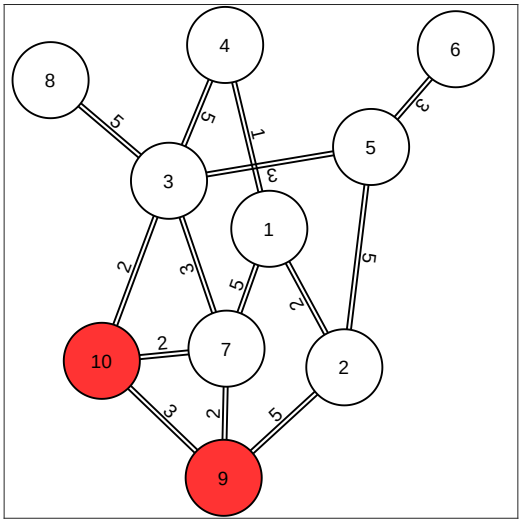


Figure 2.7: Case 1: reconstructed connections

that matrices P and Q of this form can be found separately and that, in some cases (see Remark 3), P corresponds to the solutions of a convex problem. Moreover, (see Proposition (7)) all solutions for matrix Q can be parameterized by the orthonormal matrix group. From these results, we derived an algorithm (see Section 2.3.1) for reconstructing the network connections with partial information.

Appendix

The following is a well-known property of Gram matrices (see for instance Theorem 3.1 of [33]) that we will use in the proof of Proposition 2.

Proposition 8. *Let $A, B \in \mathbb{R}^{n\times m}$ be such that $A^TA = B^TB$, then there exists $Q \in O(n)$ such that $A = QB$.*

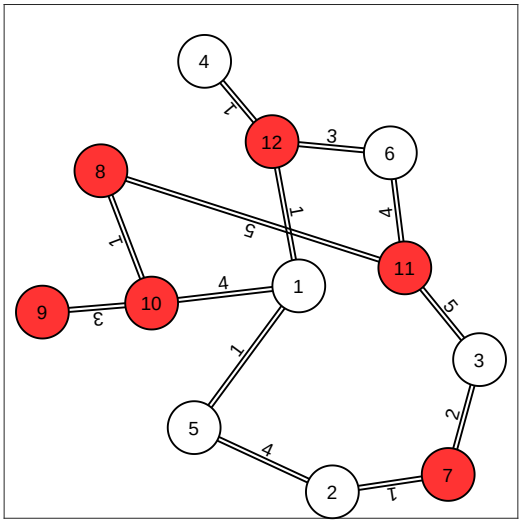


Figure 2.8: Case 2: true connections

Proof of Proposition 2

($\Rightarrow$) Assume that (2.4) has a solution T, S, G, then $GT^{-1}\hat{A}T$ is symmetric, which implies that $GT^{-1}\hat{A}T = T^T\hat{A}^T T^{-T}G$, which, setting $M = TG^{-1}T^T$, implies the first of (2.7). The second of (2.4) implies $T^T\hat{C}^T = C^T$, thus $\hat{C}TG^{-1}T^T\hat{C}^T = CG^{-1}C^T$, that is the second of (2.7). Similarly, the third of (2.4) implies $\hat{B}^T T^{-T}GT^{-1}\hat{B} = B^T G^{-1}B$, that is the third of (2.7). Finally, by the second and third of (2.4) $CG^{-1}B = \hat{C}TT^{-1}\hat{B} = \hat{C}\hat{B}$.

($\Leftarrow$) Assume that (2.7) has a solution M, G. Let $UU^T = M$ be the Cholesky decomposition of M. The second, third and fourth conditions of (2.7) imply that

$$\begin{pmatrix} \hat{C}U, \\ \hat{B}^T U^{-T} \end{pmatrix} \begin{pmatrix} U^T\hat{C}^T, & U^{-1}\hat{B} \end{pmatrix}$$
$$= \begin{pmatrix} CG^{-1/2} \\ B^T G^{-1/2} \end{pmatrix} \begin{pmatrix} G^{-1/2}C^T, & G^{-1/2}B \end{pmatrix}.$$

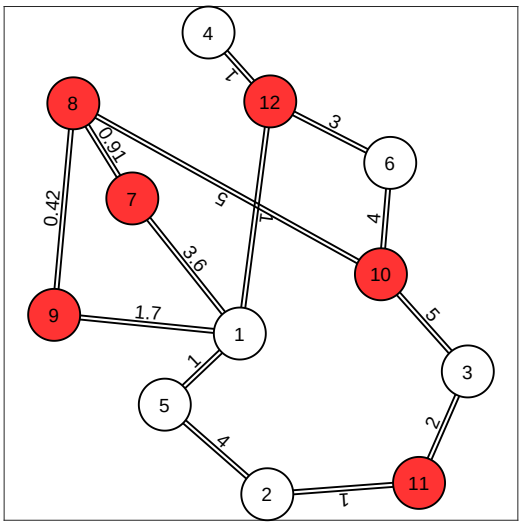


Figure 2.9: Case 2: reconstructed connections

Then, by Proposition 8, there exists an orthonormal matrix Q such that $\begin{pmatrix} \hat{C}U \\ \hat{B}^T U^{-T} \end{pmatrix} Q = \begin{pmatrix} CG^{-1/2} \\ B^T G^{-1/2} \end{pmatrix}$, that is

$$\hat{C}UQ = CG^{-1/2}$$
$$\hat{B}^T U^{-T} Q = B^T G^{-1/2}$$

and, setting $T = UQG^{1/2}$, it follows that $\hat{C}T = C$ and $T^{-1}\hat{B} = B$. Finally, $\hat{A}M = M\hat{A}^T$ implies that $\hat{A} = M\hat{A}^T M^{-1} = UU^T\hat{A}^T U^{-T}U^{-1}$ and $GT^{-1}\hat{A}T = GG^{-1/2}Q^T U^{-1}\hat{A}UQG^{1/2} = G^{1/2}Q^T U^{-1}UU^T\hat{A}^T U^{-T}U^{-1}UQG^{1/2} = G^{1/2}Q^T U^T\hat{A}^T U^{-T}QG^{-1/2}G = T^T\hat{A}^T T^{-T}G$. Hence $GT^{-1}\hat{A}T$ is symmetric, which proves the first of (2.4).

The following proposition is a property of orthonormal transformations.

Proposition 9. *Let $A, B \in \mathbb{R}^{n\times m}$ with $A^TA = B^TB$ and rank $A = m < n$, let $\bar{A}, \bar{B} \in \mathbb{R}^{n\times(n-m)}$ be such that $\bar{A}^T\bar{A} = I$, $\bar{B}^T\bar{B} = I$, $\bar{B}^TB = 0$, $\bar{A}^TA = 0$, and let $Q \in \mathbb{R}^{n\times n}$. Let $\mathscr{R} = \{Q \in O(n) : QA = B\}$ and $\mathscr{S} = \{BA^+ + \bar{B}U\bar{A}^T, U \in O(n-m)\}$, then $\mathscr{R} = \mathscr{S}$.*

Proof. (Proof that $\mathscr{R} \subset \mathscr{S}$.)

Let $Q \in \mathscr{R}$, then $QA = B$ and $B^T Q\bar{A} = (QA)^T Q\bar{A} = A^T\bar{A} = 0$. Hence, $Q\bar{A}$ is orthogonal to B and the image of $Q\bar{A}$ belongs to the image of $\bar{B}$. This implies that there exists a matrix U such that $Q\bar{A} = \bar{B}U$. Moreover, $\bar{A}^T Q^T Q\bar{A} = I = U^T\bar{B}^T\bar{B}U$ so that $\bar{B}U \in O(n-m)$. Then, $Q\begin{pmatrix} A & \bar{A} \end{pmatrix} = \begin{pmatrix} B & \bar{B}U \end{pmatrix}$. Note that $\begin{pmatrix} A^+ \\ \bar{A}^T \end{pmatrix}\begin{pmatrix} A & \bar{A} \end{pmatrix} = I$, so that $\begin{pmatrix} A^+ \\ \bar{A}^T \end{pmatrix} = \begin{pmatrix} A & \bar{A} \end{pmatrix}^{-1}$ and $Q = \begin{pmatrix} B & \bar{B}U \end{pmatrix}\begin{pmatrix} A^+ \\ \bar{A}^T \end{pmatrix} = BA^+ + \bar{B}U\bar{A}^T$.

(Proof that $\mathscr{S} \subset \mathscr{R}$.)

Let $U \in O(n-m)$, note that

$$\begin{pmatrix} B^T \\ U^T\bar{B}^T \end{pmatrix}\begin{pmatrix} B & \bar{B}U \end{pmatrix} = \begin{pmatrix} B^TB & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} A^TA & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} A^T \\ \bar{A}^T \end{pmatrix}\begin{pmatrix} A & \bar{A} \end{pmatrix},$$

then, by Proposition 8, there exists $Q \in O(n)$ such that $\begin{pmatrix} B & \bar{B}U \end{pmatrix} = Q\begin{pmatrix} A & \bar{A} \end{pmatrix}$, so that $QA = B$. Finally, since $\begin{pmatrix} A^+ \\ \bar{A}^T \end{pmatrix} = \begin{pmatrix} A & \bar{A} \end{pmatrix}^{-1}$, it follows that

$$Q = \begin{pmatrix} B & \bar{B}U \end{pmatrix}\begin{pmatrix} A^+ \\ \bar{A}^T \end{pmatrix} = BA^+ + \bar{B}U\bar{A}^T. \qquad \square$$

Chapter 3

Time-aware Data Clustering Approach

In industrial pharmaceutical plants, the data collected by sensors used to monitor the manufacturing processes are typically recorded for years [34]. In particular, in sterile plants, Cleaning In Place (CIP) and Leak Test (LT) processes are almost always fully automated. More generally, the history of the process data of an industrial plant can be used to monitor the behaviour of the plant components [35, 36, 37]. The most commonly used strategies aim at training anomaly classifiers and building predictive maintenance algorithms based on those data [38, 39]. The efficiency of the developed models is strongly related to the collected data and the type of components to be monitored [40].

Predictive maintenance aims at predicting the degradation of a machine, typically associated with anomalies of its components. However, it is very difficult to predict imminent failures with high accuracy: in-depth knowledge of the specific system and its characteristics is often necessary in order to build an accurate prediction model [41]. However, in many real cases the process data lack fundamental information to accurately capture the deterioration level of a specific component. In particular, the collected data typically provide information only on a large set of components. Thorough knowledge of system physics may not be sufficient to understand the cause of a

change in the observed sensor signals [42]. In fact, it may be easy to identify a change in the state of the overall machine, but it is often difficult to attribute the cause of this change to a particular component of the machine.

In this chapter, we develop a semi-automatic approach, which allows to evaluate a Health Indicator (HI) for multiple machine components of an industrial freeze-dryer. We focus our analysis on two different freeze-drying processes, namely, the CIP and the LT. We show that the DBSCAN clustering method is useful to pre-process the collected data in order to identify anomalies, thus leading to a robust classification system. The used dataset consist, respectively, of (i) the water flow rate signal of a spray used in the CIP and (ii) the pressure signal recorded during the LT. Both signals are extracted from the historical data of the same machine.

This chapter is structured as follows. In Section 3.1, we present the two freeze-drying analysed processes. In Section 3.2, we develop a semi-automatic method to analyze the water flow and the pressure signal in order to build a HI. In Section 3.3 and in Section 3.4 we present the obtained results for the CIP and the LT, respectively. In Section 3.5, we draw our conclusions.

3.1 System Background

Freeze-drying or lyophilization, is a process largely used in the pharmaceutical field, since its operational conditions guarantee that the final product, despite the shape transformation, keeps all its initial qualities and preserves them over time. Industrial freeze-drying are designed to reach and maintain specific temperature and pressure conditions needed for the process to be successful. In addition to the lyophilization cycle, other automated processes are run in the freeze-dryer with the aim of cleaning, sterilizing or testing its integrity. Among these additional processes, CIP and LT are of interest in this work.

Cleaning In Place

During the CIP process the freeze-dryer is cleaned with purified water: chamber walls, shelves, condenser walls, and condenser plates are sprayed by nozzles. The nozzle that sprays hot water in the condenser is subject to strong mechanical and thermal stresses—in fact, the temperature difference between the water introduced into the condenser and the steel of the condenser can reach 150°K. This stress may cause leaks in the spray tube, which would lead to incorrect machine washing: this is a critical issue for a sterile machine. It is thus necessary to constantly monitor the status of the spray to keep the machine operating correctly. The components of the studied system are shown in Fig. 3.1. During the CIP process, valve 1 remains open while valve 2 opens and closes three times in a fixed time interval (equal to 50 sec). During this period, the water pushed by the pump flows to the nozzles and enters into the machine. The Water Flow Rate Sensor (WFRS), used to monitor the process, measures the water flow rate (dimension: $\left[\mathrm{m}^3/\mathrm{h}\right]$): the WFRS signal is used to calculate the total amount of water that has been sprayed in the freeze-dryer. The sampling rate of the WFRS is 1 sample/sec.

The only process signal that can be used to monitor the status of the sprays is the WFRS signal associated with the water flowing into the machine. During the CIP, different sprays are activated in disjoint time intervals, so that the water flow signal is representative of the single spray that pours water into the freeze dryer at the corresponding time instant.

The water flow rate poured from a spray depends on two factors:

- the structural conditions of the spray, which depend on the deterioration of its steel components;
- the performance of the pump that pushes water into the nozzles, which can vary its thrust force depending on the state of health of the pump.

Even though the variations in the WFRS signal are due to both factors mentioned above, the component most prone to anomalies is the condenser spray, because of thermal stress.

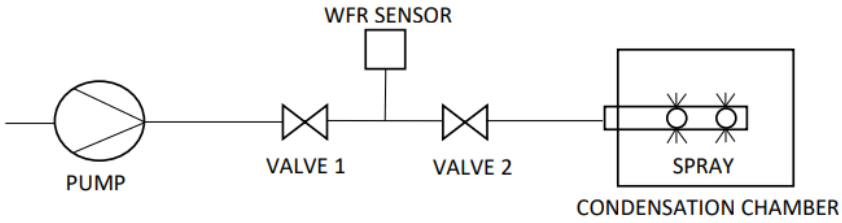


Figure 3.1: P&ID of the water supply for the spray.

This study is based on the historical data of the CIP process carried out from January 2015 to January 2020 through an industrial freeze-dryer of the production plant of GSK in San Polo di Torrile (Parma, Italy).

Leak Test

The LT is necessary to measure the sealing of the freeze-dryer. In fact, because of strong thermal variations, microscopical cracks can appear, especially in tubes and support structures. These cracks may cause a leak, i.e., an influx of gas into the drying chamber.

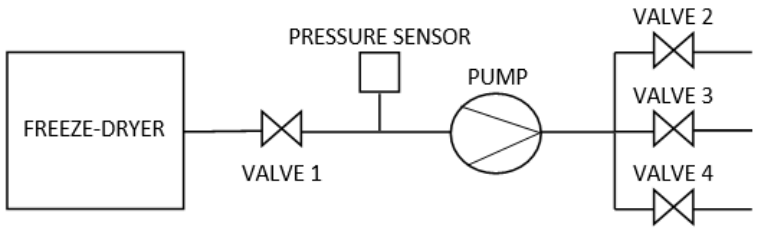


Figure 3.2: P&ID of LT main components.

The process signals that can be used to monitor the status of the freeze-dryer

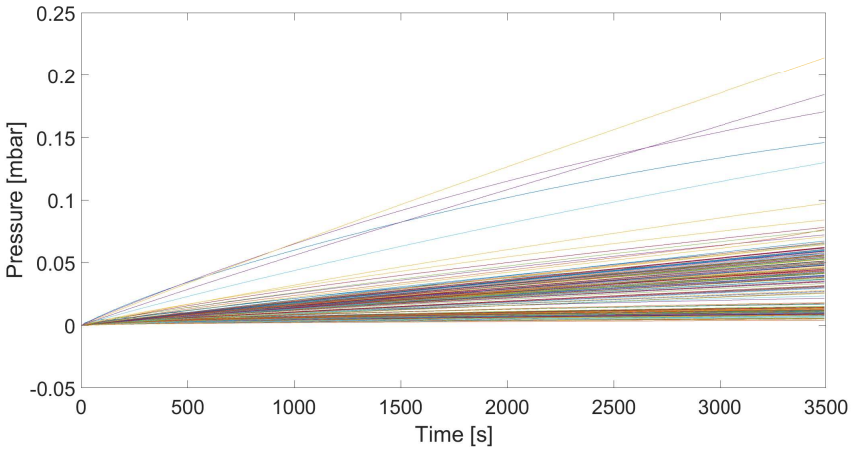


Figure 3.3: LT pressure signals.

sealing are the pressure signals. In the machine there are many pressure sensors, those that serve to control the leak test process are inside the main chamber of the freeze dryer, this study focuses instead on the pressure sensor that is inside the tube that connects the freeze dryer with the vacuum pump. The signals related to this sensor are represented in Fig. 3.3. Data are extracted by the pressure sensor shown in Fig. 3.2 together with the considered section of the plant, which includes a vacuum pump and four vacuum valves. We underline that, in this case, the pressure increase measured during the LT cannot be associated with a single component, but it depends on the health of multiple components which all contribute to lyophilizer sealing.

The analysed historical data refer to the LT carried out from January 2016 to January 2020 on the freeze-dryer mentioned for the CIP at the end of Subsection 3.1.

3.2 The Proposed Data Analysis Approach

The objective of this chapter is to describe a general strategy, which allows to obtain a time-series data segmentation expedient to classify the machine component's operational conditions, so that one can predict the evolution of the health status of the

analysed component over time. In particular, we will show that by applying clustering methods it is possible to identify cycles with anomalous behavior, which may be due to a process performed under non-standard conditions or a machine failure. In order to reach this goal, the proposed strategy uses a standard feature extraction method combined with the clustering of system states over time: in other words, we consider time-aware state clustering. As will be shown in Section 3.3 and in Section 3.4, DBSCAN will turn out to be the most effective time series clustering method, using the process cycle number as a fundamental feature. In the remainder of this section we sketch the main “ingredients" of our approach.

3.2.1 Feature Extraction and Monotonicity

The HI is a one-dimensional indicator that quantifies the “healthiness” of the system under consideration. Standard strategies for the implementation of this indicator are known in the literature [43]. However, its applicability and accuracy depend heavily on the features computed on the signals obtained from the sensors monitoring the system. There is no a priori rule for feature extraction, as this depends on the specific situation of interest and requires a thorough knowledge of the underlying process in order to focus on the most important aspects of available sensor signals.

Smoothing

Since sensor signals are often noisy, “smoothing” them is expedient to better highlight the underlying trend. In order to do this, a causal moving median filter with a window of 6 elements is applied to the extracted features, obtaining the following smoothed signal (associated with the most recent time epoch of the window):

$$f_{smooth}(i) = \text{median}[f(i), f(i-1), ..., f(i-5)] \tag{3.1}$$

where $f(i)$ is the value of feature f in the i-th cycle, $i = 6, ..., N$, where N is the number of all available cycles. Note that, for $i < 6$, $f_{smooth}(i) = \text{median}[f(i), ..., f(1)]$.

Once the features are extracted and smoothed, their "potential" to predict the degradation of one system component must be evaluated. This "potential" is quantified in terms of monotonicity, which is given by the following formula:

$$\text{monotonicity}(f,N) = \left| \sum_{i=1}^{N-1} \frac{\text{sgn}\left[f(i+1) - f(i)\right]}{N-1} \right| \tag{3.2}$$

where $f(i)$ is the value of feature f in the i-th cycle [44]. By definition, the monotonicity is between 0 and 1 and provides a measure of how well a feature describes the system evolution.

3.2.2 DBSCAN

As anticipated, our strategy to obtain a time-series data segmentation for the classification of the machine operational conditions revolve around the use of the DBSCAN. DBSCAN is a density-based clustering method, which requires two parameters: (i) the minimum amount of elements per cluster, denoted as minPts, and (ii) the distance ε that specifies the radius of a neighborhood of a given point in the cluster [44, 45].

The main advantages of this method are the following:

1. it is not necessary to know the number of clusters in advance;

2. it is possible to identify the outliers.

3.3 CIP

With the purpose of computing the HI and highlighting the condenser spray anomalies, we choose to extract the following three intuitive temporal features from the WFRS signal. An illustrative representation of such features is shown in Fig. 3.4.

- Feature 1 is the time interval between the instant at which the water flow rate overcomes the threshold of $0.3\,\text{m}^3/\text{h}$ and the time instant corresponding to its return below this threshold.

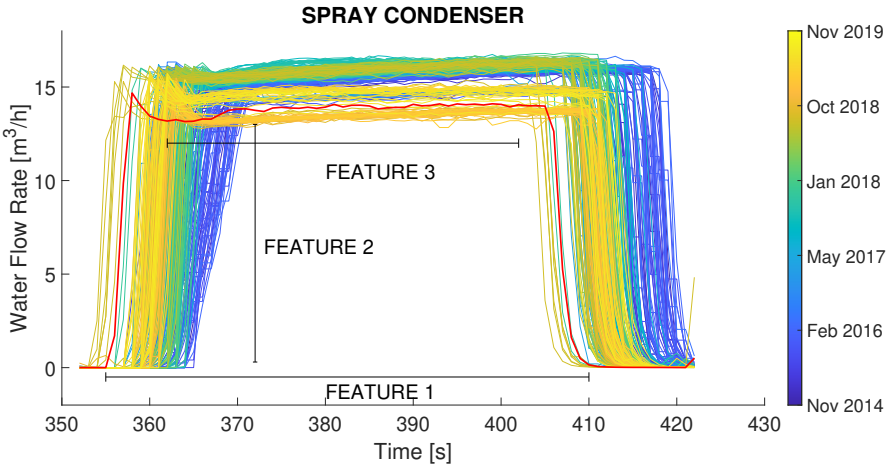


Figure 3.4: WFRS Signal and time intervals associated with possible features.

- Feature 2 measures the time taken by the system to reach the maximum capacity: it corresponds to the time interval between the instant at which the water flow rate overcomes $0.3\,\mathrm{m}^3/\mathrm{h}$ and the instant at which it reaches the value $13\,\mathrm{m}^3/\mathrm{h}$.

- Feature 3 corresponds to the average value of the WFRS signal during the "steady-state," i.e., the time interval during which the water flow rate is maximum. In order to extract this feature, we identify the MidPoint (MP) between the instant at which the water flow rate exceeds $12\,\mathrm{m}^3/\mathrm{h}$ and the instant at which the water flow rate returns below this threshold. Then, we calculate the average value of the WFRS signal in an interval equal to 40 sec centered at the MP.

The monotonicity values of the three temporal features are shown in Table 3.1. It can be noted that the only sufficiently monotonous feature is Feature 3. For this reason, we decide to extract other (common) statistical features, listed in Table 3.2, from the WFRS signal. In this table, $\{v(t)\}_{t=1}^{n}$ corresponds to the WFRS signal in the (n-sample) interval of Feature 3.

Table 3.1: Three Features Monotoniicty

	Feature 1	**Feature 2**	**Feature 3**
Monotonicity	0.01	0.01	0.56

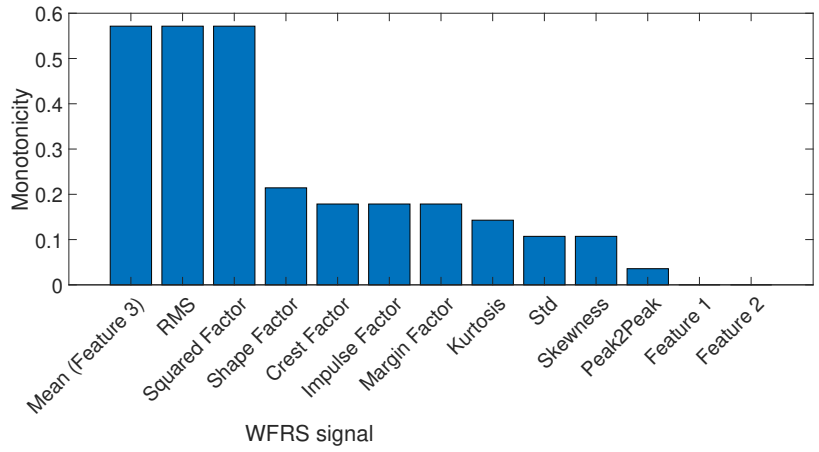


Figure 3.5: Monotonicity of all the features.

We can calculate the values of the statistical features outlined above for each CIP cycle, thus obtaining these features as functions of time (an epoch corresponds to a CIP cycle). In this case as well, we apply smoothing according to Equation (3.1).

In Fig. 3.5, the monotonicity of the extracted statistical features (all statistical features of WFRS, Feature 1 and Feature 2) is shown. It can be observed that, while almost all values are below 0.2, the only three features that have good monotonicity are “Mean,” “RMS,” and “Square Factor.” Since these three features have a correlation of 99%, one of them is representative of the other two. We will thus use Feature 3 (the “Mean”) as the only feature of the WFRS signal relevant for our analysis and we will refer to it as “FlowMean” feature. It can be expected that FlowMean, properly

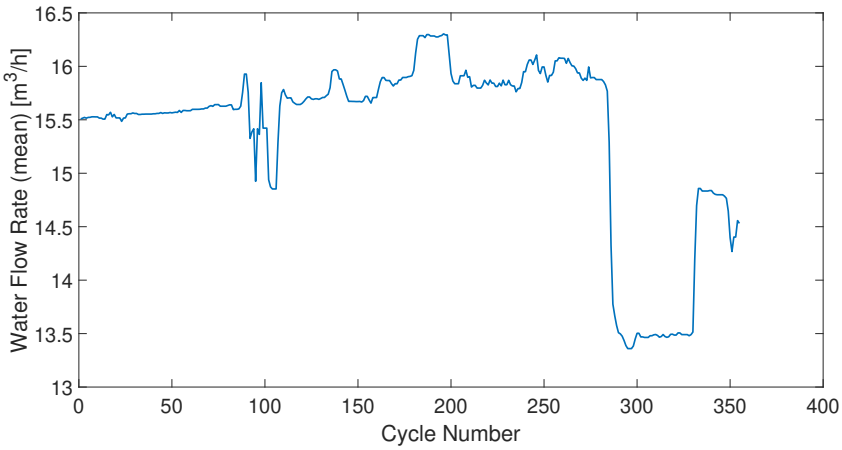


Figure 3.6: FlowMean, as a function of the cycle numbers smoothed with a causal moving median with a six steps window.

smoothed with the moving median filter in Equation (3.1), will be monotonically increasing. This is confirmed by the results in Fig. 3.6, where the value of FlowMean is shown as a function of the cycle number. Our goal is to determine a HI which is representative of the "healthiness" of the spray watering system. Therefore, we would like to obtain a HI monotonically related to the health of the freeze-dryer. In particular, a sudden variation of the feature corresponds to a system modification carried out by the maintenance technicians. Unfortunately, the HI shown in Fig. 3.6 does not have the desired behaviour, as its values vary significantly. In this case, it is not possible to identify a clear trend representing the slow degradation of the system. In order to overcome this limitation, we introduce a time-aware clustering approach, based on the use of DBSCAN, to properly pre-process FlowMean.

3.3.1 DBSCAN-based Data Clustering

DBSCAN is applied to two features:

- the average value of the WFRS signal (FlowMean);

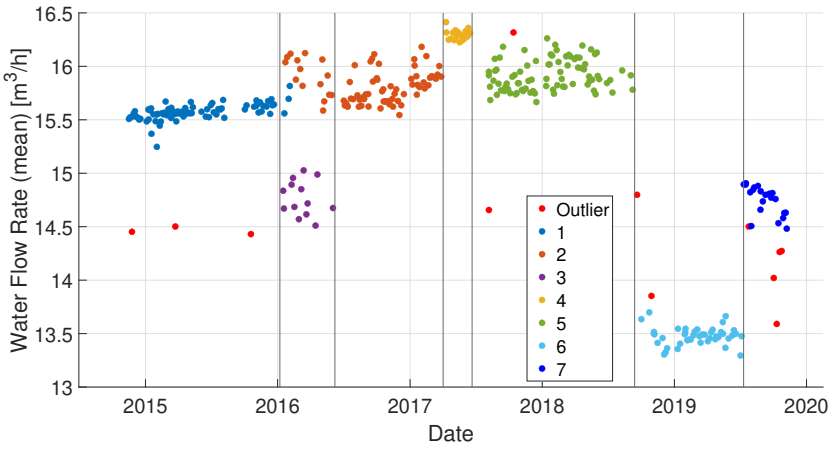


Figure 3.7: Clusters found by the DBSCAN (MinPts $= 5$, $\varepsilon = 17$). CIP process.

- the cycle number of the CIP process (denoted as CycleNumber).

In order to make these two features comparable to each other, we normalize them. In this specific case, we first normalize FlowMean and, then, multiply it by a factor of 100. This "heuristic" normalization makes FlowMean have a size comparable to that of Cycle Number. The chosen (MinPts, ε) configuration for the DBSCAN is MinPts $= 5$ and $\varepsilon = 17$. In Fig. 3.7, the outcome of DBSCAN based clustering of the historical data is shown. It is important to note that the borders between adjacent clusters correspond to modifications carried out in the system (e.g., maintenance acts). Nevertheless, a "strange" behaviour is shown by cluster 3, as it overlaps temporally with cluster 2, but the two clusters have clear distinct FlowMean values. The physical cause of this anomaly is not clear, but it is remarkable that this anomalous behaviour is detected by an automatic method. It can be noted that the DBSCAN also manages to identify, in addition to the anomalous cluster, all the isolated points, classifying them as outliers.

Cleaning data from outliers and detecting anomalies are fundamental for the construction of a robust HI. In Fig. 3.8, the HI of the condenser spray, obtained after discarding outliers and anomalous clusters through the DBSCAN algorithm, is shown.

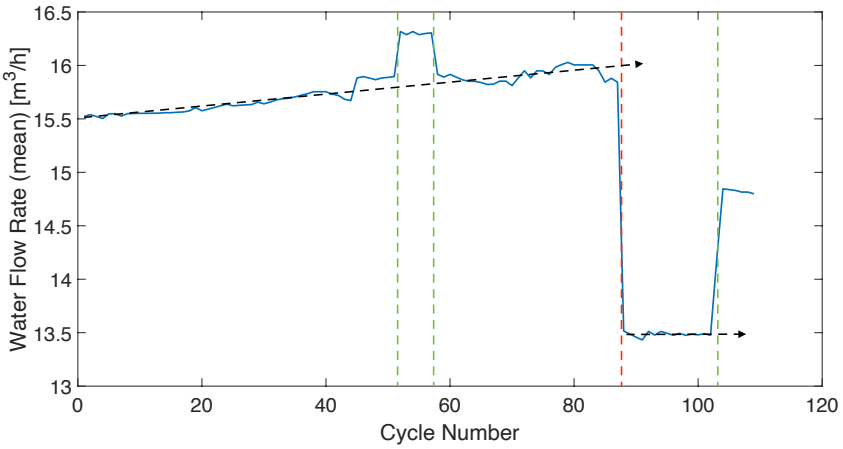


Figure 3.8: Health Indicator of the spray

The vertical green lines are associated with maintenance acts carried out on the pump. In particular, the first two maintenances are the events delimiting cluster 4 in Fig. 3.7. The state of the system is altered with the first maintenance and, then, returns to the starting state with the second. These two maintenance acts are very close in time and change the status of the pump, not of the spray. Therefore, in order to describe the spray deterioration one could eliminate cluster 4 to build a more robust HI. The red vertical line of Fig. 3.8 represents the only maintenance carried out on the spray—in particular, a leak was identified and welded. It can be seen that the indicator grows linearly from the initial data instant to the welding date (red vertical line at cycle 88): after this event, the slope of the straight line that interpolates the values of the HI is approximately equal to 0, which indicates that the leak has been eliminated (i.e., the machine state remains approximately the same). The green line at cycle 104 refers to a process modification intervention on the water distribution system, which

intentionally leads to an increase in water flow rate from that cycle onward.

3.3.2 DBSCAN versus K-Means

In order to justify the use of the DBSCAN we compare its performance with that of another relevant clustering algorithm, namely K-Means. K-Means partions a set of n observations into k clusters, so that the intercluster similarity is minimized and the intracluster similarity is maximized. Similarity is expressed as the mean value of the observations in a cluster [46]. Unlike DBSCAN, the number of clusters k must be set a priori. This is critical for the application at hand, especially when the amount of data to be clustered keeps on increasing with real-time data acquisition and the number of clusters is expected to increase over time.

In Fig. 3.9, the clusters identified by K-Means are shown. It can be observed that for $k = 7$, i.e., for a value of k equal to the number of clusters found by DBSCAN in Fig. 3.7, the detected clusters differ from the ones predicted by DBSCAN. In particular, it can be noticed that cluster 4 in Fig. 3.7 is not correctly detected by K-Means; rather, it is included in cluster 7 of Fig. 3.7 with many other cycles belonging to the two adjacent states. This highlights a major problem of k-means: if the clusters representing the machine state have very different sizes (in terms of number of cycles), k-means is not able to separate the data correctly.

Moreover, unlike DBSCAN K-Means cannot automatically identify the outliers. However, cluster-based or distance-based methods, which allow to remove the outliers and can be used together with K-Means, have been proposed [47]. Nevertheless, using these methods requires to set additional parameters, such as the cardinality of the k-nearest neighbors set N_P. Therefore, it can be concluded that, for our problem, the most efficient clustering approach is DBSCAN, since it can simultaneously identify the right machine states and the outliers, without any further action.

3.4 Leak Test

The same semi-automatic HI computation strategy applied to the CIP process and described in Section 3.2 can be applied also to the analysis of LT. In this case as

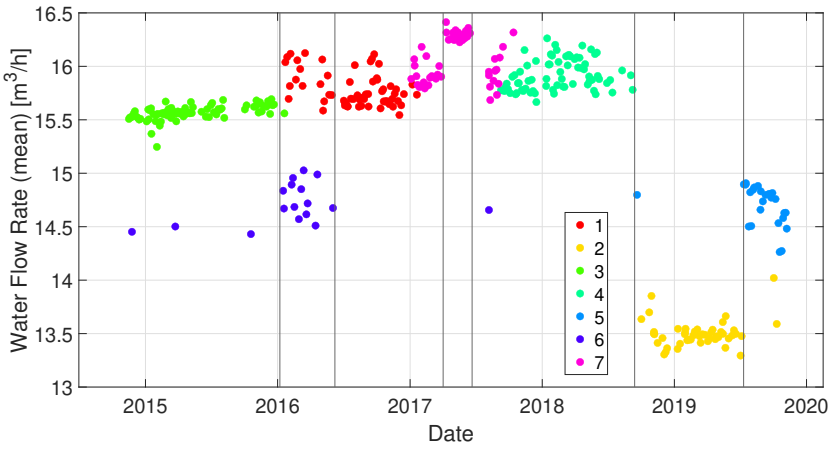


Figure 3.9: Clusters found by the K-Means (k=7). CIP process.

well, the only feature taken into account is the mean value of the pressure signal. In fact, in Section 3.3 it has been observed that the mean value (of the WFRS signal) was a feature sufficient to describe the whole process. We will refer to this feature as “PressureMean." After smoothing by means of the moving median filter modelled in Equation (3.1), the monotonicity can be computed and turns out to be 0.5 (the results are not shown here for the sake of conciseness). This monotonicity value justifies the use of the PressureMean feature for anomaly detection.

As happened for the FlowMean feature in Fig. 3.6, in this case as well the properly smoothed PressureMean is not an effective HI. In fact, as shown in Fig. 3.10, there are significant oscillations that prevent a clear identification of the system behaviour.

At this point, DBSCAN can be applied to the two following relevant features:

- the average pressure (PressureMean);
- the cycle number of the Leak Test (CycleNumber).

In order to make these two features comparable to each other, the same ”heuristic” normalization method previously adopted for FlowMean is applied to PressureMean.

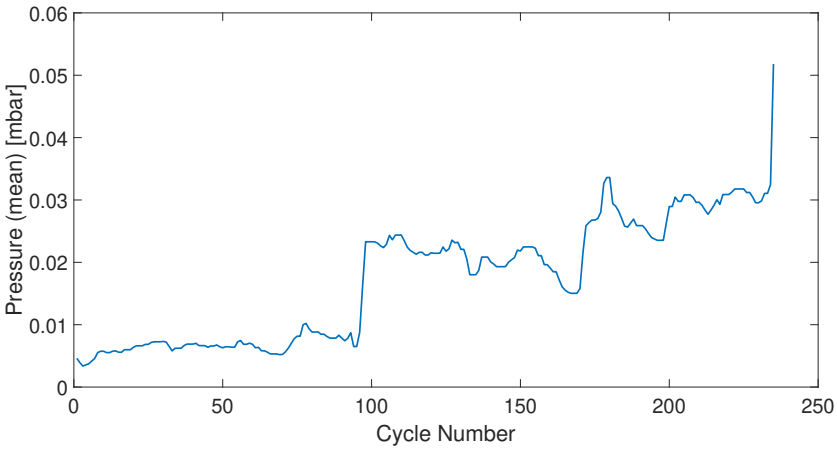


Figure 3.10: PressureMean, as a function of the cycle numbers smoothed with a causal moving median with a six steps window.

In the LT case, a good machine state identification is obtained setting MinPts $= 5$ and $\varepsilon = 17$: the corresponding cluster data are shown in Fig. 3.11. The chosen (MinPts, ε) configuration is the same which allows to obtain, for the CIP, the data clusters shown in Fig. 3.7. It can be noticed that the clusters' separations correspond to repair activities carried out on the analysed freeze-dryer and, consequently, to changes in the machine operational conditions. The outliers can be detected and removed in order to compute a robust HI, which is shown in Fig. 3.12. It can be observed that three time intervals, associated with the evolution of the state of the machine, can be clearly identified: during each of these interval the HI remains relatively stable. In correspondence to the separation instants between adjacent intervals, changes were made to the machine was subject to changes which led to a degradation (higher HI).

It is remarkable that the same DBSCAN-based semiautomatic method can be used to describe the evolution of the state of two different components, starting from two signals of different nature, namely, water flow rate and pressure. But for the different natures of the used sensors, the used metholodogy (including their parametric values) was the same.

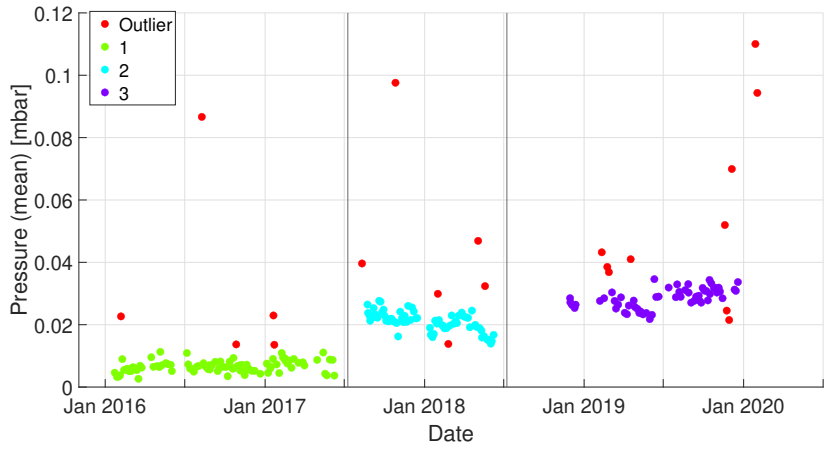


Figure 3.11: Clusters found by the DBSCAN (MinPts $= 5, \varepsilon = 17$). Leak Test process.

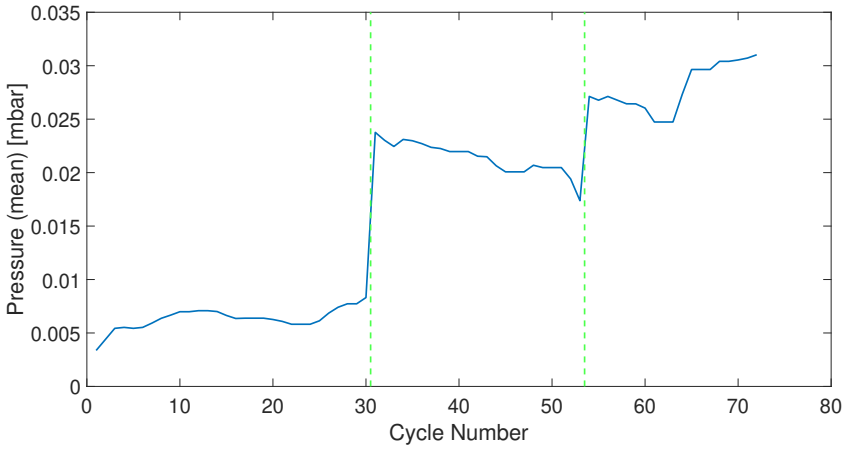


Figure 3.12: Health Indicator which describes the deterioration of the multiple components contribute to the sealing of the machine.

3.5 Conclusion

In this chapter, we have proposed a time-aware clustering approach to the derivation of an HI of an industrial pharmaceutical machine (namely, a freeze-dryer), with refer-

ence to two different signals acquired during two different automated processes. Our results show that, on the basis of a time-aware DBSCAN-based clustering, we can obtain a temporal data “segmentation" into clusters which characterize different operational conditions of the analyzed machine. A comparison with another clustering method has been carried out, highlighting the validity of our approach. By combining the proposed clustering approach with knowledge of the maintenance acts, it is possible to identify outliers and distinguish anomalous clusters from those useful to describe the degradation of the monitored components.

Table 3.2: Time Domain Statistical Features of a time-discrete signal $\{v(t)\}_{t=1}^{n}$. The mean of the WFRS signal is Feature 3.

Feature	Math expression
Mean	$\mu = \frac{1}{n}\sum_{t=1}^{n} v(t)$
Standard Deviation (Std)	$\sigma = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(v(t)-\mu)^2}$
Skewness	$\frac{\frac{1}{n}\sum_{t=1}^{n}(v(t)-\mu)^3}{\sigma^3}$
Kurtosis	$\frac{\frac{1}{n}\sum_{t=1}^{n}(v(t)-\mu)^4}{\sigma^4}$
Peak2Peak	$\max\{v(t), t=1...n\} - \min\{v(t), t=1...n\}$
Root Mean Square (RMS)	$\sqrt{\frac{1}{n}\sum_{t=1}^{n} v^2(t)}$
Crest Factor	$\frac{\max\limits_{t=1...n} v(t)}{RMS}$
Shape Factor	$\frac{RMS}{\frac{1}{n}\sum_{t=1}^{n}\lvert v(t)\rvert}$
Impulse Factor	$\frac{\max\limits_{t=1...n} v(t)}{\frac{1}{n}\sum_{t=1}^{n}\lvert v(t)\rvert}$
Margin Factor	$\frac{\max\limits_{t=1...n} v(t)}{\left(\frac{1}{n}\sum_{t=1}^{n}\lvert v(t)\rvert\right)^2}$
Squared Factor	$\sum_{t=1}^{n} v^2(t)$

Chapter 4

Leak Detection and Diagnosis

Leaks in lyophilizers can be classified in two groups: external and internal. *External leaks* are due to openings of the lyophilization chamber to the external environment. They can be due to the deterioration of the chamber isolation valves or to cracks in the tubes that carry the fluid used for thermal regulation. *Internal leaks* are due to cracks in hollow components located inside the chamber, such as support structures. There is a fundamental difference between the two types of leaks. External leaks contaminate the chamber since they are associated with the non-sterile external environment. Instead, internal leaks are not contaminating, since they are associated with gas flows coming from the interior of the chamber sterile environment.

Because of this difference, it is important to separate external and internal leak contributions. This is often difficult, in fact, citing [48], "in most cases both phenomena will occur simultaneously so that separating the two causes is often difficult if not impossible." As will be shown by the mathematical model presented in Section 4.1, flows associated with internal and external leaks vary differently in time. In fact, external leaks are constant, while internal leaks decrease exponentially. However, their correct detection and separation is difficult for the following reasons.

- Internal leaks originating from very small cracks are almost constant in time and, hence, indistinguishable from external ones.
- Leaks are estimated from special low pressure sensors. The measured signals

are often irregular and affected by noise and artifacts.

In this chapter, we address the problem of identifying and separating internal and external leaks, using the lyophilization pressure signals acquired in multiple leak detection tests, spanning various months of operation. A key element of the proposed method is the use of multiple leak detection tests. This allows to identify more precisely the model parameters and to estimate the time evolution of internal and external leaks.

Chapter outline. In Section 4.2, we propose an optimization-based method for the identification of the model parameters from a sequence of leak detection cycles. In Section 4.3, we present a method for automatically identifying new leaks. In Section 4.4, we present experimental results obtained from a real data set provided by GlaxoSmithKline's production plant in San Polo di Torrile (Parma, Italy). Finally, Seciton 4.5 concluded the chapter.

Statement of contribution. The problem of leak identification in lyophilizers has not received much attention in the literature. Indeed, to the best of our knowledge, this is the first work that addresses this problem by using data acquired from multiple leak detection cycles. The main novelties of this chapter can be summarized as follows:

- The development of the simplified model (4.5), that describes internal and external leaks.
- The combined use of multiple leak tests, obtained in several months of operation. This allows to identify model parameters more accurately. In fact, pressure signals used for leak detection are noisy and subject to artifacts, and the data obtained from a single test are not sufficient for a precise model identification. Moreover, the use of multiple cycles allows finding trends in model parameters and detecting new leaks.

Note that we do not claim that the used regularization techniques (which will be described by equation (4.10)) is new, since similar methods are used in Magnetic Resonance Imaging (MRI). In fact, MRI problems are described by a mathematical model which is equivalent to the one used to describe leaks (which will be modeled with equation (4.5)).

Notation. Given a matrix $A = (a_{ij}) \in \mathbb{R}^{n\times m}$, we denote by $\|A\|_{1,1} = \sum_{i=1}^{n}\sum_{j=1}^{m}|a_{ij}|$ its entrywise 1-norm. Given matrices $A = (a_{ij}) \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{p\times q}$, their Kronecker product is $A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{bmatrix} \in \mathbb{R}^{(np)\times(mq)}$.

4.1 Leak Modeling

4.1.1 A Simplified Leak Model

In this subsection, we present a simplified model for the lyophilization chamber pressure p during a leak test. In the machine there are many pressure sensors, those that serve to control the leak test process are inside the main chamber of the freeze dryer, this study focuses instead on the pressure sensor that is inside the freeze-dryer main chamber. The signals related to this sensor are represented in Fig. 4.1. As depicted in

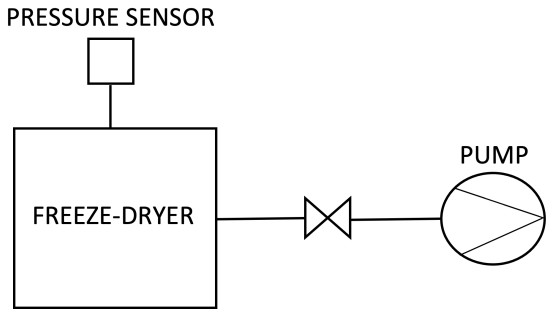


Figure 4.1: Simplified P&ID of LT components involved.

Fig. 4.2, we assume that the lyophilizer is modelled as a closed chamber containing an ideal gas at low pressure. During evacuation and leak test preparation, a vacuum pump extracts gas from the chamber. The pump is stopped during the leak test phase and, in the case of perfect sealing, the chamber pressure remains constant. However, the internal pressure may increase over time due to leaks.

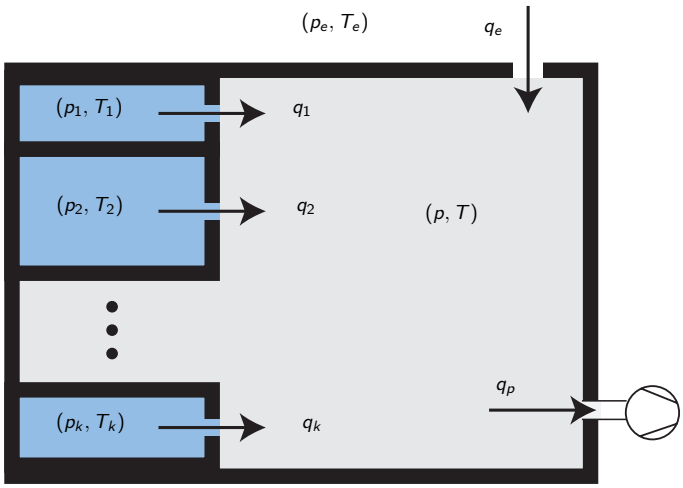


Figure 4.2: Lyophilization chamber model, including k containers that originate the internal leaks.

In our model, we separate internal and external leaks. External leaks are modelled as an aggregate inflow q_e from the external environment. Internal leaks are modelled as k separate inflows $q_1, \ldots, q_k$ from k internal containers, embedded in the lyophilization chamber and connected to this by small orifices, corresponding to cracks. We consider only one aggregate external leak, since, as we will see below, external leaks are constant over time, so that their contributions can be represented by a single term. On the other hand, internal leaks decrease, over time, with different rates (depending on the embedded containers' sizes and on the cracks areas) and must be considered separately.

The chamber gas consists of a mixture of nitrogen (N_2), that fills the chamber before the beginning of the leak test, and unwanted atmospheric air, due to external leaks. From the ideal gas equation, the chamber pressure $p(t)$ satisfies:

$$p(t)V = [n_N(t) + n_A(t)]RT \tag{4.1}$$

where: V is the chamber volume (dimension: $[l]$), R (dimension: $[J/mol \cdot K]$)is the ideal gas constant; T is the chamber temperature (dimension: $[K]$), assumed to be

constant during leak detection (this is justified by the fact that the system is kept at thermal equilibrium). Finally, $n_N(t)$ and $n_A(t)$ are the numbers of nitrogen and air moles in the chamber at time t. We denote by $q_i(t)$ the inflow corresponding to internal leak i at time t. Hence, $n_N(t), n_A(t)$ satisfy the following mass balance equations:

$$\begin{aligned} M_A \dot{n}_A(t) &= q_e(t) \\ M_N \dot{n}_N(t) &= \sum_{i=1}^{n} q_i(t) \end{aligned} \tag{4.2}$$

where M_N, M_A are nitrogen and air molar masses (dimension: $[Kg/mol]$), and $q_e(t)$ is the external inflow, at time t.

Since external air pressure is much higher than chamber pressure, we can assume that the external gas flow is choked, so that it is computed by (3.20) of [49] as

$$q_e = \frac{A p_E}{\sqrt{RT_e}} F(\gamma)$$

where: $F(\gamma) = \sqrt{\gamma}\left(1 + \frac{\gamma-1}{2}\right)^{-\frac{\gamma+1}{2(\gamma-1)}}$; p_E is the external air pressure (dimension: $[mbar]$); T_e is the external air temperature (dimension: $[K]$), γ is the gas heat capacity ratio; and A is the fissure area (dimension: $[m^2]$). We assume that external air pressure and temperature are constant (air temperature is usually regulated), so that external air flow can be assumed to be proportional to the fissure area A, which represents the sum of all areas of external leakage. We assume that A is constant during each leak detection cycle, but may vary among different cycles. We also assume that the k flows corresponding to the internal leaks are choked. The internal flow originated from a container is choked if its pressure is at least 1.9 times the chamber pressure (see for instance Table 3.2 of [49]). Since the chamber is kept at a very low pressure during leak test preparation and Leak Test (LT) phases, and the volume of the chamber is much higher than the volumes of the embedded containers, a non-choked flow from a container to the chamber would be negligible. Therefore, we only consider choked flows. Note that this assumption greatly simplifies the model, since choked flows depend only on upstream pressure (in our case, the pressure of the k containers). Furthermore, we assume that the temperature of each container is constant: this

is justified by the fact that, during LT preparation and LT phases, gas flows are very small, so that the system can be considered at thermal equilibrium at all times.

Because of the above assumptions, we model the outflows due to internal leaks as

$$q_i(t) = \frac{A_i p_i(t)}{\sqrt{RT_i}} F(\gamma) \quad i = 1, ..., k$$

where: p_i is the pressure in the i-th container (dimension: $[mbar]$); A_i is the corresponding orifice area (dimension: $[m^2]$) i.e., the size of the associated crack; and T_i is the air temperature inside the container (dimension: $[K]$). According to the ideal gas law, $p_i(t)V_i = n_i(t)RT_i$, where n_i is the number of nitrogen moles inside the i-th container and V_i is its volume. Then,

$$\dot{n}_i(t) = -\frac{q_i(t)}{M_N},$$

and

$$\dot{p}_i(t) = -\frac{q_i(t)RT_i}{V_i M_N} = -p_i(t)\frac{A_i\sqrt{RT_i}}{V_i M_N} F(\gamma)$$

so that, setting

$$\tau_i^{-1} = \frac{A_i\sqrt{RT_i}}{V_i M_N} F(\gamma),$$

we obtain

$$q_i(t) = -c_i e^{-\frac{t}{\tau_i}},$$

where $c_i \in \mathbb{R}$ is a positive constant, that represents the flow from the i-th internal container at $t = 0$.

Hence, according to (4.1) and (4.2), when the void pump is not activated, the chamber pressure obeys the following equation:

$$\dot{p}(t) = \frac{q_e RT}{M_A V} + \sum_{i=1}^{k} \frac{c_i RT}{M_N V} e^{-\frac{t}{\tau_i}} \tag{4.3}$$

which can be rewritten, more concisely, as

$$\dot{p}(t) = l + \sum_{i=1}^{k} \frac{1}{\tau_i} a_i e^{-\frac{t}{\tau_i}} \tag{4.4}$$

with proper definitions of l and $\{a_i\}_{i=1}^{k}$. The factor $\{\tau_i\}_{i=1}^{k}$ guarantees the normalization condition

$$\int_0^{+\infty} \frac{1}{\tau_i} e^{-\frac{t}{\tau_i}} dt = 1, \quad \forall \tau_i.$$

In this way, in equation (4.4), a_i represents the overall pressure increase due to the i-th internal leak, over an infinite time horizon. In Fig. 4.3, the increments of the chamber pressure, due to external leaks and to one internal leak, are directly compared. Since

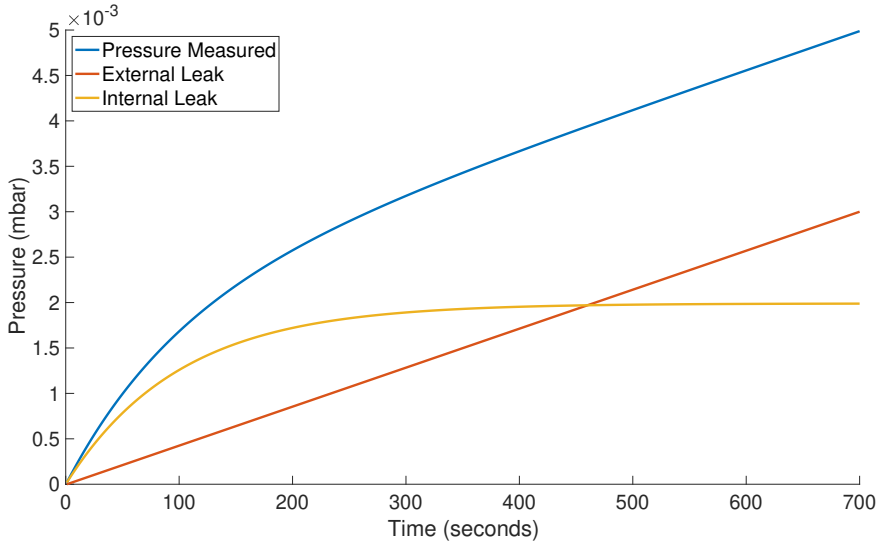


Figure 4.3: Comparison of the pressure increment in the chamber caused by external leaks and by an internal leak.

internal leaks decrease exponentially, they are more relevant in the LT preparation phase than in the LT phase.

Note that fitting model (4.4) to a measured pressure signal p corresponds to decomposing p in exponential components: this is a recurring problem in the literature (see, for instance, [50]).

Measuring chamber pressure p is problematic for a number of reasons. First, since the chamber pressure is very low, a precise measurement is difficult to achieve. The pressure signal is affected by large noise and exhibits a significant offset that

drifts with time. For this reason, the employed low pressure sensors need frequent re-calibrations. Moreover, the pressure signal often presents artifacts, such as the spikes reported in Fig. 4.4, of unclear origin, that need be removed from the data used for identification.

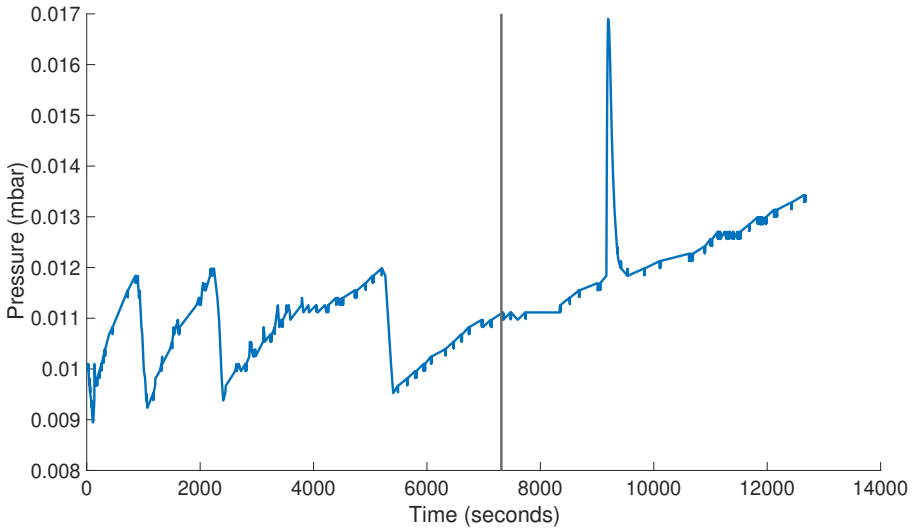


Figure 4.4: A typical raw chamber pressure signal. The black vertical line separates the LT preparation phase from the LT phase. The signal is noisy and a spike is present during the LT phase.

Roughly speaking, the main problem that we consider is the following one. the problem that we consider is the following one. *Given the chamber pressure signals obtained from various successive leak tests, identify the evolution over time of the parameters of model* (4.4)*, predicting possible future failures before the plant needs be stopped for repairs.*

4.1.2 Spectral Decomposition of Leaks

The identification of parameters l and $\{a_i\}_{i=1}^{k}$ in (4.4) presents two main challenges: (i) we do not know the number N of internal leaks; (ii) the model depends non-

linearly on parameters $\{\tau_i\}_{i=1}^{k}$. To overcome those challenges, we consider the following alternative model:

$$\dot{p}(t) = l + \sum_{i=1}^{n} \frac{1}{\tau_i} a_i e^{-\frac{t}{\tau_i}}, \tag{4.5}$$

where: n is a *preassigned* number of components; and $\tau_1, \ldots, \tau_n$ are *preassigned* time constants (dimension: [s]. We choose a sufficiently large value of n and we set the values of $\{\tau_i\}_{i=1}^{n}$ so that they encompass all time-constants that are relevant to our identification problem (typically between 10^2 s and 10^5 s). The choice of n is taken on the basis of the application context: if n is too small, there is not a sufficient granularity to distinguish the different leaks; on the contrary, if n is very large, the computational time increases considerably due to the many variables introduced into the problem.

Note that model (4.5) represents a spectral decomposition of $\dot{p}$, since it describes $\dot{p}$ as the sum of a constant term l and various exponential curves, with different time constants. Mathematically, the problem of finding l and $\{a_i\}$ from (4.5) can also be considered a discrete inverse Laplace transform problem (see [51], [52] for a more detailed discussion).

4.2 Parameter Identification

Let $\{t_1, \ldots, t_r\}$ be the times corresponding to the r chamber pressure samples. We denote by $\dot{p}(t_i)$ a numerical approximation of the pressure derivative at time t_i. We simply use Euler's forward approximation $\dot{p}(t_i) = p(t_{i+1}) - p(t_i)/t_{i+1} - t_i$, also other finite difference approximations could be used. On a single LT, we can identify the parameters in model (4.5) by solving the following optimization problem:

$$\begin{array}{ll} \min\limits_{l,a_i} & \sum_{j=1}^{r} \left[\dot{p}(t_j) - l - \sum_{i=1}^{n} a_i f_i(t_j)\right]^2 \\ \text{subject to} & a_i \geq 0 \;\; i = 1, \ldots, n \\ & l \geq 0 \end{array} \tag{4.6}$$

with $f_i(t) = e^{-\frac{t}{\tau_i}}/\tau_i \;\; (t \geq 0)$. As mentioned above, the main advantage of model (4.5), with respect to (4.4), is that the parameters $\{\tau_i\}$ are preassigned and so the model is

linear with respect to the remaining parameters $l, a_1, \ldots, a_n$. As a consequence, Problem (4.6) is a linear programming one and can be solved at low computational cost. However, due to its large number of variables, Problem (4.6) can be ill-conditioned. Later, we will return to this point.

We do not focus on a single LT, but on a sequence of them, in order to be able to capture trends in the evolution of internal and external leaks and to detect possible future failures. We denote by l_q the external leak of cycle q and by $a_{i,q}$ the value of a_i for cycle q. In this way, we can reformulate Problem (4.6) as follows:

$$\begin{array}{ll} \min\limits_{l_q, a_{i,q}} & \sum_{q=1}^{m} \sum_{j=1}^{r} \left| \dot{p}_q(t_j) - l_q - \sum_{i=1}^{n} a_{i,q} f_i(t_j) \right| \\ \text{subject to} & a_{i,q} \geq 0 \;\; i = 1, \ldots, n; \;\; q = 1, \ldots, m \\ & l_q \geq 0 \;\; q = 1, \ldots, m. \end{array} \tag{4.7}$$

where m is the total number of processes. In the objective function, we have replaced the squared 2-norm by the 1-norm, that is more robust with respect to outliers. Since Problem (4.7) is still a linear programming one, with a larger set of variables with respect to (4.6), it is convenient to rewrite it in the following matrix form:

$$\begin{array}{ll} \min\limits_{A,L} & \|\dot{P} - \mathbb{1} \otimes L - FA\|_{1,1} \\ \text{subject to} & A \geq 0, L \geq 0 \end{array} \tag{4.8}$$

where: $\dot{P} \in \mathbb{R}^{r \times m}$; $(\dot{P}_{j,q}) = \dot{p}_q(t_j)$; $\mathbb{1} = [1, \ldots, 1]^T \in \mathbb{R}^r$; $L = [l_1, \ldots, l_m] \in \mathbb{R}^{1 \times m}$; $F \in \mathbb{R}^{r \times n}$; $(F_{ji}) = f_i(t_j)$; $A \in \mathbb{R}^{n \times m}$; $(A_{i,q}) = a_{i,q}$; and $\|\cdot\|_{1,1}$ is the elementwise matrix 1-norm. The inequalities in (4.8) are intended component-wise.

Remark 5. *A problem mathematically similar to* (4.7) *is addressed in the literature in the context of inversion of Nuclear Magnetic Resonance (NMR) data (see for instance [53] or [54]).*

In our identification procedure, we consider only sampling times t_i in which the void pump is switched off. Admissible sampling times belong to the complete LT phase, but also to those intervals of LT preparation phase in which the void pump is not active. For instance, in Fig. 4.5: the blue line represents the chamber pressure signal in LT preparation and LT phases; orange lines correspond to the samples used

for identification. As anticipated above, we exclude the samples in which the pump is active and those corresponding to pressure spikes (a phenomenon of unclear origin that is not described by model (4.4)). The LT preparation phase is very important for a correct identification of internal leaks, since the flows due to internal leakages are stronger (and more informative) than in the later LT phase, because of the exponential decay. A method that can be used to find samples with positive derivative is described, for example, in [55].

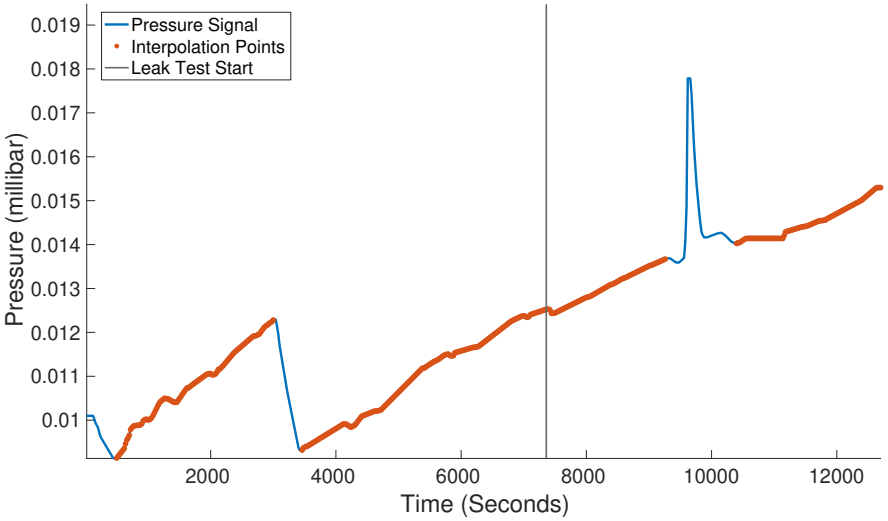


Figure 4.5: A raw chamber pressure signal. The orange part of the curve corresponds to the samples used for identification.

4.2.1 Regularization Techniques

In general, Problem (4.7) is ill-conditioned, due to the large number of variables (namely, $m(n+1)$) and the fact that functions $f_i(t)$, $f_j(t)$ are similar if τ_i and τ_j are close. This problem is well-known in the literature on NMR inversion (a problem that, as said, is mathematically analogous to (4.7)). For a more detailed discussion, see for instance [51] or [56]. As we will see later, this problem can be solved by regu-

larization techniques. Furthermore, we would like to find a solution of (4.7) in which A is a sparse matrix. In fact, it is reasonable to assume that only a few closed containers contribute to internal leaks. A possible solution would be to limit the number of non-zero elements of each column of A: this corresponds to a bound on the so-called zero norm of the columns of A. However, it is well-known that this is computationally not convenient, since the resulting optimization problem would be an NP-hard one. As typically done, we add to the cost function in (4.7) a sparsity-promoting penalty term proportional to the 1-norm of A, obtaining the following problem:

$$\begin{array}{ll} \min\limits_{A,L} & \|\dot{P} - \mathbb{1} \otimes L - FA\|_{1,1} + \lambda_1 \|A\|_{1,1} \\ \text{subject to} & A \geq 0, L \geq 0 \end{array} \tag{4.9}$$

where $\lambda_1 \geq 0$ is a regularization real parameter.

In Fig. 4.6, we show the solution of Problem (4.9) corresponding to real LT data provided by GlaxoSmithKline. Namely, in Fig. 4.6 we show the entries of matrix A and vector L in the solution of (4.9): the x-axis represents the different leak detection cycles and the y-axis the different time constants. The rectangle at position (q,i) (where $q = 1,...,m$ and $i = 1,...,n$) has a color that depends on the value of $a_{q,i}$ in the solution of Problem (4.9), according to the color scale on the right-side of the chart. The last row of the chart represents the external leaks vector L and the fuchsia vertical line corresponds to a maintenance. As can be observed in Fig. 4.6, internal leaks vary significantly from one cycle to the next one: this is not realistic. In fact, formations of new internal leaks are rare events and existing leaks increase slowly. To reduce the coefficients variations between consecutive cycles, we add two regularization terms to Problem (4.9), obtaining

$$\begin{array}{ll} \min\limits_{A,L} & f(A,L) = \|\dot{P} - \mathbb{1} \otimes L - FA\|_{1,1} \\ & +\lambda_1 \|A\|_{1,1} + \lambda_2 \|AM\|_{1,1} + \lambda_3 \|LM\|_{1,1} \\ \text{subject to} & A \geq 0, L \geq 0 \end{array} \tag{4.10}$$

where $M = \begin{bmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{bmatrix} \in \mathbb{R}^{m\times(m-1)}$ is a regularization matrix. Note that $\|AM\|_{1,1} = \sum_{i=1}^{n}\sum_{q=1}^{m}\left|a_{i,q+1} - a_{i,q}\right|$, $\|LM\|_{1,1} = \sum_{q=1}^{m}\left|l_{q+1} - l_q\right|$. The addition of these two terms to the objective function of (4.10) penalizes variations in coefficients $\{a_{i,q}\}$ and $\{l_q\}$ from one cycle to the next one. We can assign different weights to terms λ_2 and λ_3 since we expect larger variations in external leaks $\{l_q\}$ than in internal leaks coefficients $\{a_{i,q}\}$. In fact, external leaks can be triggered by many causes, such as the deterioration of a bolt or a sealing valve, and new external leaks can emerge rather frequently. Instead, internal leaks are mainly related to the development of micro-cracks in the chamber internal support structures and are rare, at least according to the maintenance data at our disposal.

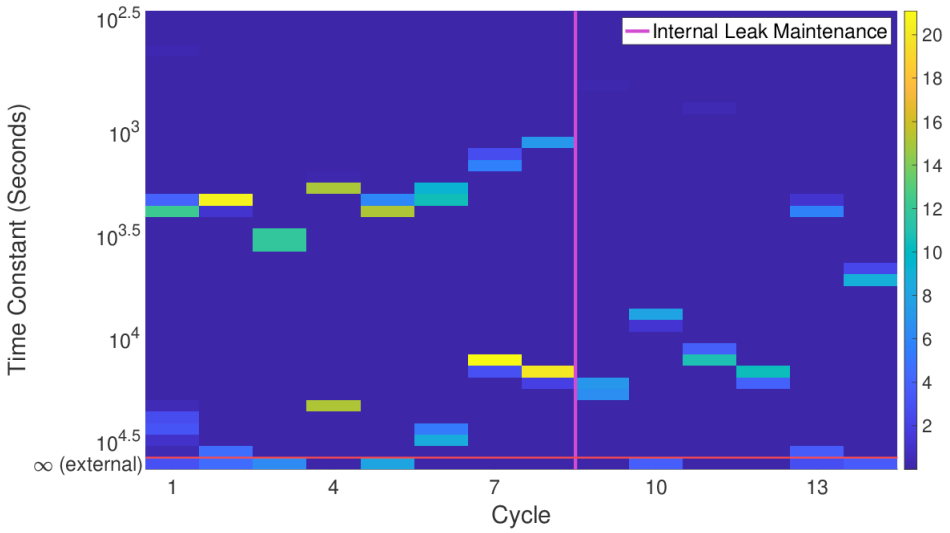


Figure 4.6: Representation of the solution of Problem (4.9).

Fig. 4.7 shows the solution of Problem (4.10) on the same dataset used for Fig. 4.6.

Note that the leak coefficients $\{a_{i,q}\}$ remain almost constant over time up to the maintenance, that has fixed two internal leaks (two horizontal lines are eliminated).

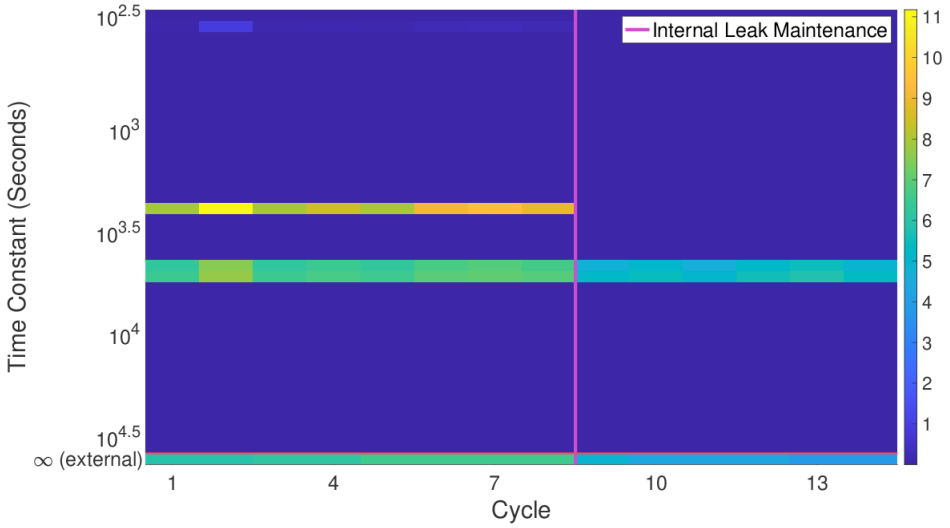


Figure 4.7: Representation of the solution of the regularized Problem (4.10).

Remark 6. *Similar regularization strategies are adopted in NMR inversion (see, for instance, [57]). In particular, various authors use penalty terms to obtain similar values on neighboring pixels on the x and y axis of the reconstructed image. The regularization methods presented in this chapter are indeed inspired by the literature on NMR inversion.*

4.2.2 Choice of Regularization Parameters

The choice of the regularization parameters $\lambda_1, \lambda_2, \lambda_3$ greatly influences the solution of Problem (4.10). Let $g^*(\lambda_1, \lambda_2, \lambda_3) = \|\dot{P} - \mathbb{1} \otimes L - FA\|_{1,1}$ be the chamber pressure estimation error, where A and L are the solution of problem (4.10), for the given values of $\lambda_1, \lambda_2, \lambda_3$.

If $\lambda_1, \lambda_2, \lambda_3$ are too large, the obtained optimal solution has a large error term $g^*(\lambda_1, \lambda_2, \lambda_3)$. On the other hand, if $\lambda_1, \lambda_2, \lambda_3$ are too small, they do not sufficiently

regularize the solution. In the literature, a typical strategy for choosing regularization coefficients consists in selecting them so that

$$g^*(\lambda_1,\lambda_2,\lambda_3) = (1+\alpha)g(0,0,0) \tag{4.11}$$

where α is a positive constant. Note that $g^*(\lambda_1,\lambda_2,\lambda_3)/g^*(0,0,0)$ represents the relative increment of the model error norm due to regularization terms. By choosing α sufficiently small, one ensures that regularization terms do not increase significantly the error norm with respect to the non-regularized solution. In our numerical experiments, we chose $\alpha = 0.2$. In order to determine the relative contributions of λ_1 and λ_2, one can set the ratios $k_2 = \lambda_2/\lambda_1$ and $k_3 = \lambda_3/\lambda_1$. Then, equation (4.11) can be solved by bisection with respect to λ_1.

4.3 Changepoints Detection

Faults or maintenance works generally cause large variations in the parameters of model (4.4). Instead, we expect that the parameters of model (4.4) do not vary significantly between one cycle and the next one if no faults or maintenances have occurred between them. In this section, we present a different approach for the identification of the parameters of model (4.4) based on this observation. Namely, we assume that the parameter values remain the same from one cycle to the next one, unless a cycle is classified as a "changepoint." We want to detect such changepoints, that we are expected to correspond to faults or maintenances.

Reference [58] presents a survey of algorithms for offline detection of multiple changepoints in multivariate time series. In the following, we apply the optimal partitioning method presented in [59].

Given m leak detection cycles, set $C \subset \{2,\ldots,m\}$ is a new optimization variable that denotes the set of changepoints. In Problem (4.10), we require that the parameters are the same between cycles $j-1$ and j, provided that j is not a changepoint:

$$\begin{aligned} &a_{i,j-1} = a_{i,j}\, i = 1,\ldots,n, j \notin C \\ &l_{j-1} = l_j, j \notin C\,. \end{aligned} \tag{4.12}$$

Furthermore, we change the objective function in Problem (4.10) by adding the following penalty term:

$$\bar{f}(A,L) = f(A,L) + |C|\beta \tag{4.13}$$

where $|C|$ is the cardinality of C and β is a positive constant. The added term $|C|\beta$ penalizes solutions with a larger number of changepoints. In general, increasing parameter β reduces the number of changepoints. Later, we will discuss the choice of β in more detail.

The optimal partitioning algorithm presented in [59] allows to found the optimal set C of changepoints. The complexity of this algorithm is quadratic with respect to the total number of cycles m. This is due to the fact that the algorithm needs to solve a number of problems of form (4.10), which is quadratic with respect to m.

4.4 Experiments on Real Data

We ran tests on real leak detection tests, carried out on an industrial freeze-dryer of GlaxoSmithKline's production plant in San Polo di Torrile (Parma). The freeze-dryer internal volume V is equal $16000l$. Each leak test cycle is associated with a chamber pressure signal, recorded with a sampling time of 20 s. The evacuation phase has a variable duration, due to various factors, such as vacuum pumps efficiency. On average, it lasts about 60 min. The second phase (pressure stabilization) always lasts 120 min, while the third phase (leak test) always lasts 90 min. As previously anticipated, in Problem (4.10) we consider only sampling times in which the void pumps are switched off, i.e., those intervals of the leak test preparation phase in which the pumps are switched off and the entire LT phase.

We consider $m = 34$ consecutive leak detection cycles, carried out over a period of approximately 8 months. This particular data set represents an interesting case study, since, in this time period, the lyophilizer had one maintenance for an internal leak and two maintenances for external leaks. In model (4.5), we consider $n = 39$ different time constants, logarithmically spaced between $10^{2.5}$ s and $10^{4.5}$ s. We set $\alpha = 0.2$ in (4.11) and $k_2 = 20$, $k_3 = 10$.

Fig. 4.8 shows the model fit for a single leak detection cycle. The x-axis represents the time elapsed from the beginning of the leak test preparation phase and the y-axis represents pressure differences from the beginning of the leak test preparation phase. The blue curve is the measured chamber pressure, while the red line corresponds to the samples times used in model (4.10). Note that these samples cover the complete LT phase and the portions on the LT preparation phase in which the void pump is off. The green curve represents the identified chamber pressure curve $p_q(t)$, where q is the cycle index, namely

$p_q(t) = \int_0^t (l_q + \sum_{i=1}^n a_{i,q} f_i(\tau))\, d\tau$. Function p_q would be an estimate of the chamber pressure during leak test and leak test preparation if the *void pump were not working*. Note also that the effect of internal leaks is much more evident in LT preparation phase, while the pressure growth during the LT phase is mainly due to external leaks.

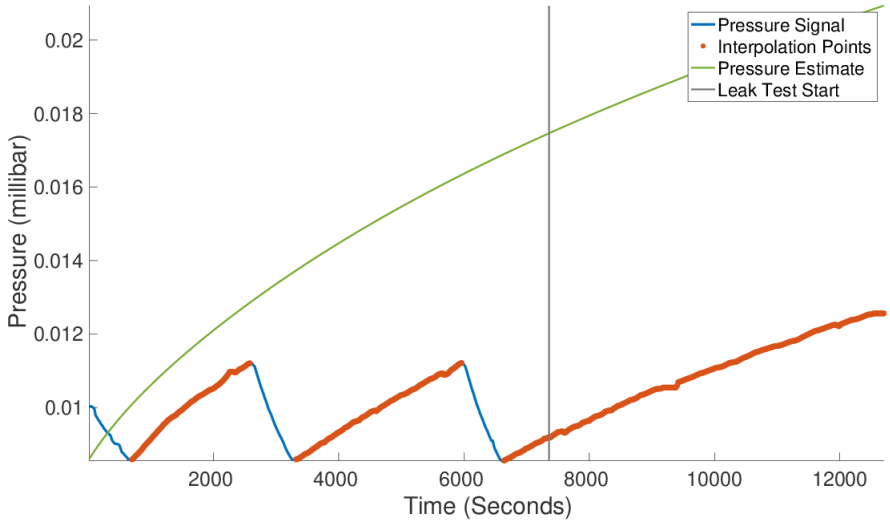


Figure 4.8: Estimated pressure curve (in green) obtained from the growth points (in orange) of the pressure signal

The histogram in Fig. 4.9 represents the obtained values of $\{a_{i,q}\}$ and $\{l_q\}$, corresponding to the decomposition of the chamber pressure derivative according to

model (4.5). The x-axis represents the time constants $\{\tau_i\}$, in logarithmic scale and decreasing order. The heights of the bars correspond to weigths $\{a_{i,q}\}$ in the fitted model, with the exception of the first bar, which corresponds to the external leak l_q and is denoted on the x axis by ∞, since a constant term is associated with an infinite time constant. Note that most of the variables are set to 0. This fact is due to the sparsity regularization term $\lambda_1\|A\|_{1,1}$ in Problem (4.10). The histogram shows 5 leaks, which is still an excessively large number. In practice, we do not expect that the chamber has more than 2 or 3 leaks. We are currently working on alternative methods to reduce the number of non-zero elements in matrix A without increasing prohibitively the complexity of the resulting optimization problem.

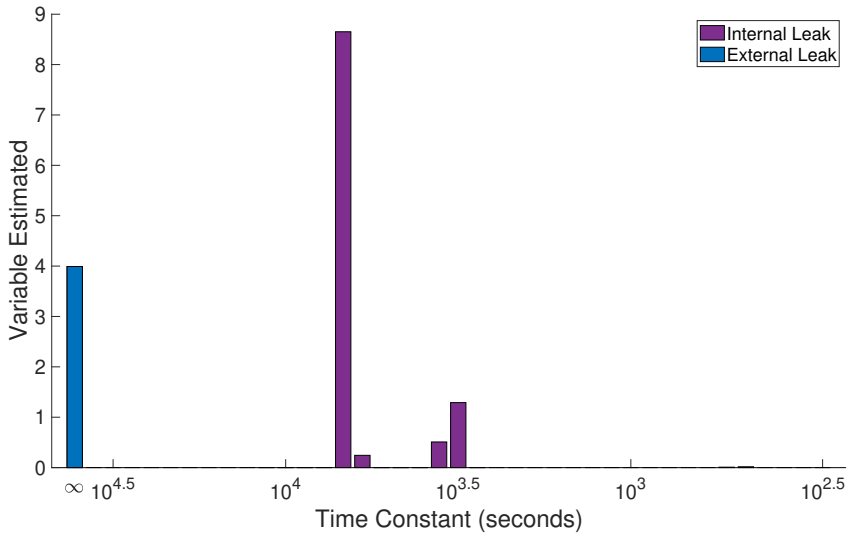


Figure 4.9: Estimation of parameters l_q and $a_{q,i}$ from a single cycle.

Fig. 4.10 shows the overall solution of Problem (4.10), including all leak detection cycles. In particular, the x-axis indicates the cycle number and the y-axis the time constants. The rectangle at position (q,i) has a color that depends on the value of $a_{q,i}$ in the solution of Problem (4.10), according to the color scale represented on the right-side of the chart. The last row of the chart represents matrix L, associated with

external leaks. The three vertical lines represent maintenances carried out on the machine. In particular, the fuchsia line corresponds to a repair on an internal leak, while the two green lines are related to repairs on external leaks. Note that, after the maintenance for the internal leak, various rectangles in the center of the figure decrease in color intensity. This indicates that this maintenance decreased internal leaks. In the same way, the intensity of the external leaks (associated with the color of the lowest row of the plot) decreases after the two repairs, marked by the green lines.

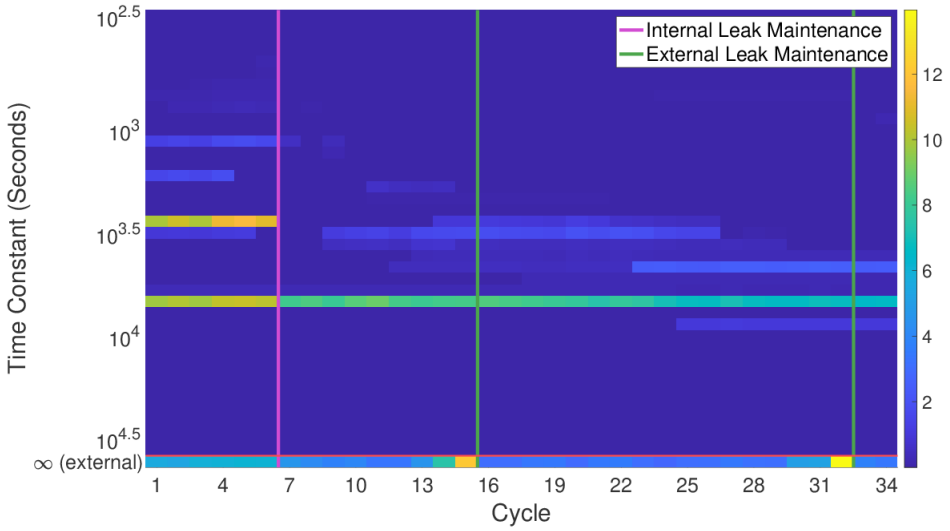


Figure 4.10: Representation of the solution of Problem (4.10).

Fig. 4.11 shows two graphs which are useful for monitoring the status of internal and external leaks and predicting imminent sealing problems. In both graphs, the x-axis represents the cycle number. The first chart represents the estimated external leak parameters $\{l_q\}$, for $q = 1,\dots,m$. The second one reports the sum $\sum_{i=1}^{n} a_{i,q}$ of all the values of the internal leaks parameters estimated for each cycle q. This representation allows to separately view the trend of internal and external leaks. It is important to remark that it is possible to notice growth trends in the external leak values a few cycles before they are sufficiently large to constitute a serious contamination and

lead to a machine halt. This allows to conduct a preventive maintenance that avoids stopping production. Moreover, the aggregate sum of internal leaks is very high in all the cycles that precede the maintenance associated with the fuchsia line. In fact, internal leaks are hardly detected by the standard method of calculating the leak rate (described in Section 1.4) since, during the LT phase, the overall gas outflow due to internal leaks is very small. Inspecting previous leak detection cycles, we noticed that the machine had had undetected internal leaks for more than a year before the maintenance.

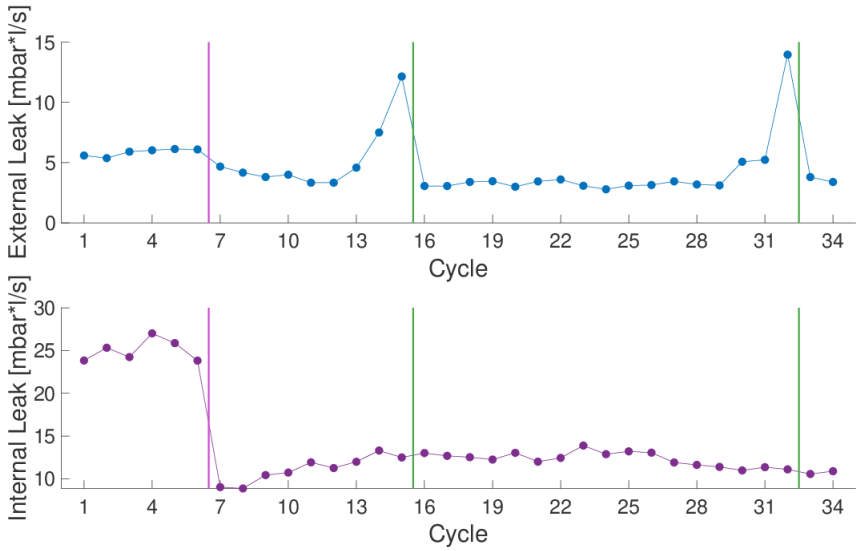


Figure 4.11: Evolution of internal and external leaks in multiple cycles.

Finally, Fig. 4.12 represents the leak rate computed according to the standard method described in Section 1.4, using the identified model. In other words, the pressure difference in (1.1) is computed by integrating the identified model in the LT phase, that is $\Delta p = \int_{t_i}^{t_f} \left(l_q + \sum_{i=1}^{n} a_{i,q} f_i(\tau)\right) d\tau$, where $[t_i, t_f]$ is the time interval corresponding to the leak test phase and terms l_q, $a_{i,q}$ correspond to the solution of (4.10). We represent separately the contribution to this integral of the external leaks $\{l_q\}$ and the internal ones $\sum_{i=1}^{n} a_{i,q} e^{-\frac{\tau}{\tau_i}}$. In Fig. 4.12 the x-axis represents the cycle num-

ber, while the y-axis represents the leak rate. The color bars indicate the two leak contributions. Blue bars represent external leaks and purple bars the internal ones. The horizontal red line is used to indicate the critical threshold, after which the leak test is considered unsuccessful and the machine is stopped for repairs. As before, the three vertical lines correspond to maintenances.

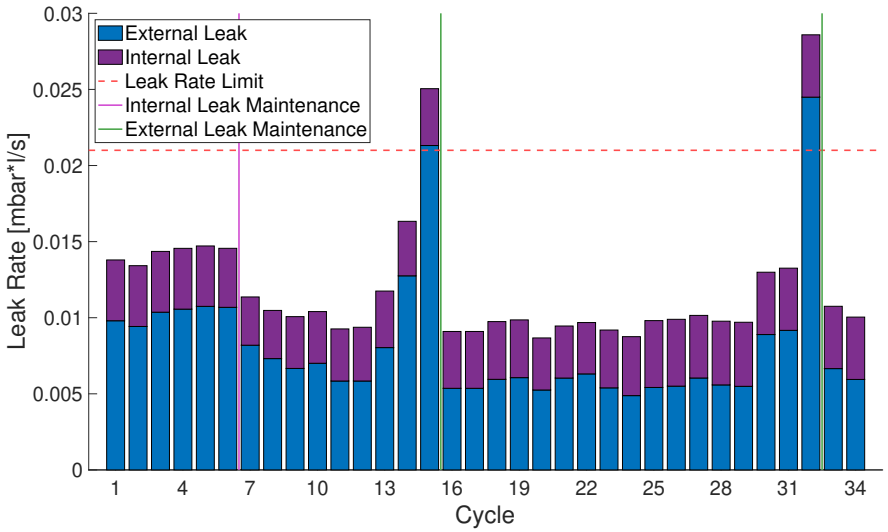


Figure 4.12: Estimated leak rates for various cycles.

We applied the changepoints detection method, presented in Section 4.3, to the same dataset. To choose the penalty term β in (4.13), we followed the method in [60]. Namely, with the optimal partitioning algorithm we solved various instances of Problem (4.10), with added constraint (4.12) and modified objective function (4.13). In Fig. 4.13, the graph on the left shows the value of the objective function (without the penalizing term $|C|\beta$) as a function of $|C|$ (the number of changepoints). The method presented in [60] consists in choosing for β the value corresponding to the most significant increase in the number of changepoints. This corresponds to the "elbow" on the left graph of Fig. 4.13, represented by a blue dot. Following this method, we set the penalty term to $\beta^* = 9800$. The graph on the right side of Fig. 4.13 repre-

sents the penalty constant β on the y-axis, while the x-axis shows the corresponding number of changepoints.

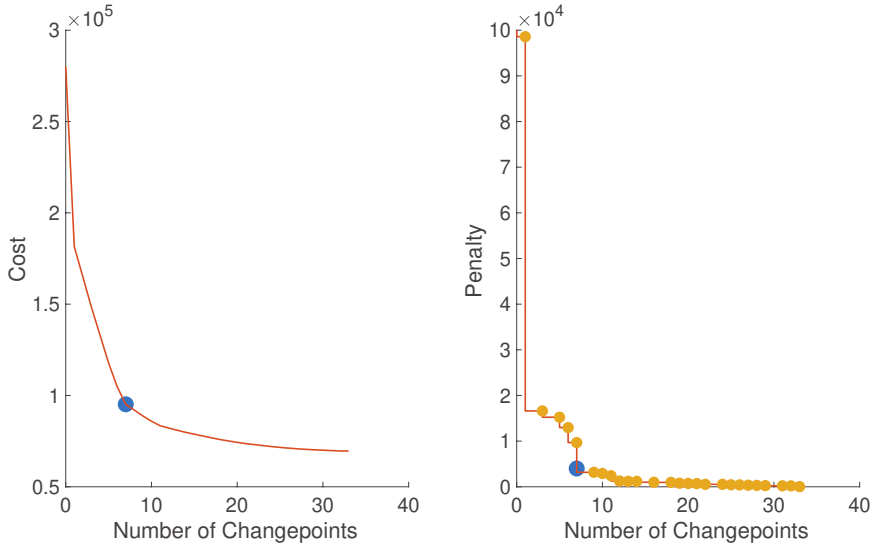
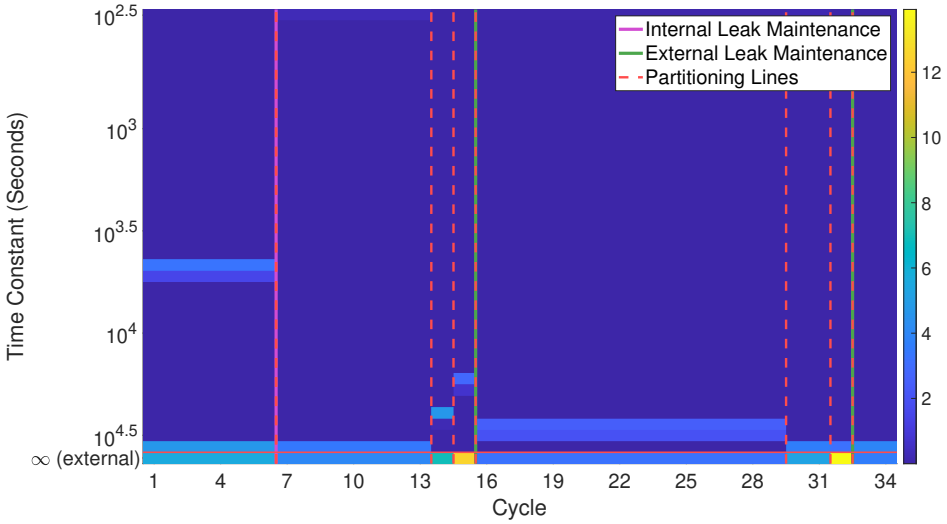
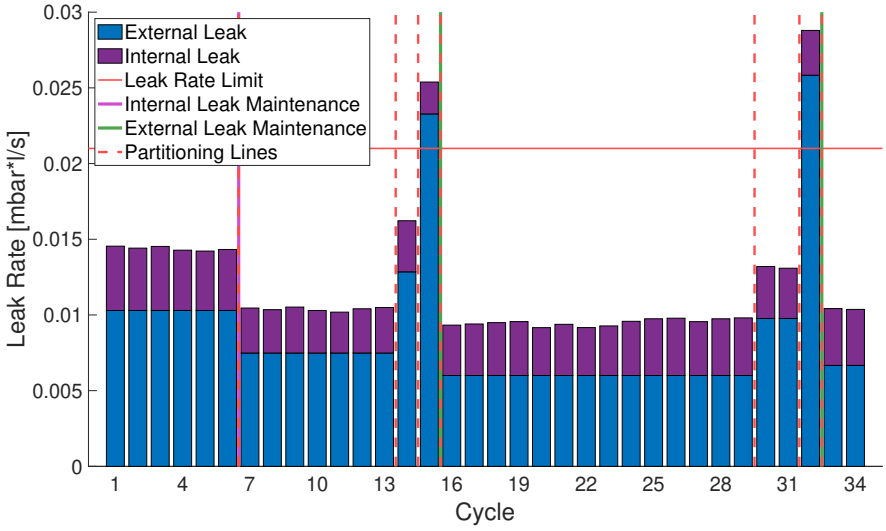


Figure 4.13: Elbow of Cost vs Number of Changepoints.

Figures 4.14 (a) and (b) have the same meaning of Figures 4.10 and 4.12, with the difference that the results have been obtained by applying the optimal partitioning algorithm to Problem (4.10), with added constraint (4.12) and modified objective function (4.13), with $\beta^* = 9800$. The red dashed vertical lines indicate changepoints. Note that the green vertical lines, representing maintenances, are preceded by changepoints. This shows that changepoints detection may allow revealing initial faults of the freeze-dryer components, before the machine state becomes critical. In particular, in case a changepoint is detected on the last cycle, it is advisable to activate an alarm to alert maintainers of possible faults.



(a) Representation of the solution of Problem (4.13).



(b) Estimated leak rates for various cycles in Problem (4.13).

Figure 4.14: Solutions of Problem (4.13) applied to LT.

4.5 Conclusion

We presented a simple mathematical model for internal and external lyophilizer leaks (in equation (4.5)). We proposed a leak identification method from multiple tests, based on the solution of a regularized optimization problem (see (4.10)). Experimental results shows that the proposed method allows identifying internal and external leaks and estimating their evolution in time. Future research will involve searching for alternative methods to enforce the sparsity of matrix A in (4.10) and to regularize the problem solution.

Conclusions

In this thesis we present the pharmaceutical freeze-dryer machine and explained some of the automated processes that take place in the machine. Then, we considered the problem of recovering the network connection structure from the measured input-output data of a resistive-capacitive (RC) circuit. We have shown a method to solve the network reconstruction problem. The model considered is general enough to describe various phenomena occurring within traditional industrial machines, such as thermal systems and physical network systems. Then, we focused our study on freeze-drying machines and used historical real data from their sensor signals to create diagnostic models. We presented a time-aware clustering-based approach for analyzing sensor data with the aim of predicting the temporal evolution of the health status of a machine component in a pharmaceutical plant. Finally, we described a mathematical model for freeze-dryer leaks and addressed the problem of identifying and separating internal and external leaks. We proposed a leak identification method based on the use of multiple leak detection tests. The developed algorithms have been implemented as monitoring systems used by GSK.

List of Publications

International Journal

- G. Calzavara, L. Consolini, J. Kavaja, Structured identification for network reconstruction of RC-models. Systems & Control Letters, 2021, Vol. 147, 104849.
- G. Calzavara, M. Iori, M. Locatelli, M. C. Moreira, T. Silveira, Mathematical models and heuristic algorithms for pallet building problems with practical constraints. Annals of Operations Research, 2021, pp. 1-32.

International Conference

- G. Calzavara, E. Oliosi, G. Ferrari, A Time-aware Data Clustering Approach to Predictive Maintenance of a Pharmaceutical Industrial Plant. 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 454-458. IEEE.
- G. Calzavara, L. Consolini, G. Ferrari, Leak Detection and Classification in Pharmaceutical Freeze-Dryers: 60th IEEE Conference on Decision and Control, IEEE.

Bibliography

[1] T. Zonta, C. A. da Costa, R. da Rosa Righi, M. J. de Lima, E. S. da Trindade, G. P. Li, Predictive maintenance in the industry 4.0: A systematic literature review, Computers & Industrial Engineering (2020) 106889.

[2] S. Sajid, A. Haleem, S. Bahl, M. Javaid, T. Goyal, M. Mittal, Data science applications for predictive maintenance and materials science in context to industry 4.0, Materials Today: Proceedings (2021).

[3] B. Ding, Pharma industry 4.0: Literature review and research opportunities in sustainable pharmaceutical supply chains, Process Safety and Environmental Protection 119 (2018) 115–130.

[4] I. C. Reinhardt, J. C. Oliveira, D. T. Ring, Current perspectives on the development of industry 4.0 in the pharmaceutical sector, Journal of Industrial Information Integration 18 (2020) 100131.

[5] J. Moyne, Y. Qamsane, E. C. Balta, I. Kovalenko, J. Faris, K. Barton, D. M. Tilbury, A requirements driven digital twin framework: Specification and opportunities, IEEE Access 8 (2020) 107781–107801.

[6] N. Mehta, A. Pandit, Concurrence of big data analytics and healthcare: A systematic review, International journal of medical informatics 114 (2018) 57–65.

[7] N. S. Arden, A. C. Fisher, K. Tyner, X. Y. Lawrence, S. L. Lee, M. Kopcha, Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future, International Journal of Pharmaceutics 602 (2021) 120554.

[8] S. Vaidya, P. Ambad, S. Bhosle, Industry 4.0–a glimpse, Procedia manufacturing 20 (2018) 233–238.

[9] A. Pesqueira, M. J. Sousa, Á. Rocha, Big data skills sustainable development in healthcare and pharmaceuticals, Journal of Medical Systems 44 (11) (2020) 1–15.

[10] M. A. Meyer, Healthcare data scientist qualifications, skills, and job focus: a content analysis of job postings, Journal of the American Medical Informatics Association 26 (5) (2019) 383–391.

[11] J. D. Mellor, et al., Fundamentals of freeze-drying., Academic Press Inc.(London) Ltd., 1978.

[12] G.-W. Oetjen, P. Haseley, Freeze-drying, John Wiley & Sons, 2004.

[13] H. Rottländer, W. Umrath, G. Voss, Fundamentals of leak detection, Leybold GMBH (ed) Cat 199 (2016) 37.

[14] L. M. Hardwick, S. L. Nail, J. Jarman, K. Hasler, T. Hense, A proposed rationale and test methodology for establishment of acceptance criteria for vacuum integrity testing of pharmaceutical freeze dryers, European Journal of Pharmaceutics and Biopharmaceutics 85 (2) (2013) 236–239.

[15] R. P. Silveira, J. O. Trierweiler, M. Farenzena, H. C. Teixeira, Systematic approaches for pi system™ data compression tuning, IFAC Proceedings Volumes 45 (15) (2012) 309–313, 8th IFAC Symposium on Advanced Control of Chemical Processes. doi:https://doi.org/10.3182/20120710-4-SG-2026.00137.
URL https://www.sciencedirect.com/science/article/pii/S147466701630461X

[16] A. van der Schaft, Modeling of physical network systems, Systems & Control Letters 101 (2017) 21 – 27, jan C. Willems Memorial Issue, Volume 2. doi:https://doi.org/10.1016/j.sysconle.2015.08.013.

URL http://www.sciencedirect.com/science/article/pii/S0167691115001814

[17] G. B. Ermentrout, D. H. Terman, Mathematical foundations of neuroscience, Vol. 35, Springer Science & Business Media, 2010.

[18] W. M. Haddad, V. Chellaboina, Q. Hui, Nonnegative and compartmental dynamical systems, Princeton University Press, 2010.

[19] J. Z. Hearon, A monotonicity theorem for compartmental systems, Mathematical Biosciences 46 (3) (1979) 293 – 300. doi:https://doi.org/10.1016/0025-5564(79)90074-9.
URL http://www.sciencedirect.com/science/article/pii/0025556479900749

[20] L.-L. Xie, L. Ljung, Estimate physical parameters by black box modeling, in: Proceedings of the 21st Chinese Control Conference, 2002, pp. 673–677.

[21] P. A. Parrilo, L. Ljung, Initialization of physical parameter estimates, IFAC Proceedings Volumes 36 (16) (2003) 1483–1488.

[22] G. Mercère, O. Prot, J. A. Ramos, Identification of parameterized gray-box state-space systems: From a black-box linear time-invariant representation to a structured one, IEEE Transactions on Automatic Control 59 (11) (2014) 2873–2885. doi:10.1109/TAC.2014.2351853.

[23] C. Yu, L. Ljung, M. Verhaegen, Identification of structured state-space models, Automatica 90 (2018) 54 – 61. doi:https://doi.org/10.1016/j.automatica.2017.12.023.
URL http://www.sciencedirect.com/science/article/pii/S0005109817306106

[24] O. Prot, G. Mercère, Combining linear algebra and numerical optimization for gray-box affine state-space model identification, IEEE Transactions on Automatic Control (2019). doi:10.1109/TAC.2019.2942567.

[25] C. Yu, L. Ljung, A. Wills, M. Verhaegen, Constrained subspace method for the identification of structured state-space models, IEEE Transactions on Automatic Control (2019). doi:10.1109/TAC.2019.2957703.

[26] J. Gonçalves, S. Warnick, Necessary and sufficient conditions for dynamical structure reconstruction of lti networks, IEEE Transactions on Automatic Control 53 (7) (2008) 1670–1674.

[27] P. E. Paré, V. Chetty, S. Warnick, On the necessity of full-state measurement for state-space network reconstruction, in: 2013 IEEE Global Conference on Signal and Information Processing, IEEE, 2013, pp. 803–806.

[28] H. J. van Waarde, P. Tesi, M. K. Camlibel, Identifiability of undirected dynamical networks: a graph-theoretic approach, IEEE control systems letters 2 (4) (2018) 683–688.

[29] H. J. van Waarde, P. Tesi, M. K. Camlibel, Topology reconstruction of dynamical networks via constrained lyapunov equations, IEEE Transactions on Automatic Control 64 (10) (2019) 4300–4306.

[30] M. Fazlyab, V. M. Preciado, Robust topology identification and control of lti networks, in: 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2014, pp. 918–922.

[31] M. Nabi-Abdolyousefi, M. Mesbahi, Network identification via node knockout, IEEE Transactions on Automatic Control 57 (12) (2012) 3214–3219.

[32] E. J. Candes, M. B. Wakin, S. P. Boyd, Enhancing sparsity by reweighted l1 minimization, Journal of Fourier analysis and applications 14 (5-6) (2008) 877–905.

[33] R. A. Horn, I. Olkin, When does a* a= b* b and why does one want to know?, The American mathematical monthly 103 (6) (1996) 470–482.

[34] J. S. Barrett, S. P. Koprowski, The epiphany of data warehousing technologies in the pharmaceutical industry., International Journal of Clinical Pharmacology and Therapeutics 40 (3) (2002) 3–13, pmid: 11911607.

[35] X. Wang, Data mining and knowledge discovery for process monitoring and control, Springer Science & Business Media, 2012.

[36] J. Chen, K.-C. Liu, On-line batch process monitoring using dynamic pca and dynamic pls models, Chemical Engineering Science 57 (1) (2002) 63–75, doi: 10.1016/S0009-2509(01)00366-9.

[37] J. MacGregor, A. Cinar, Monitoring, fault diagnosis, fault-tolerant control and optimization: Data driven methods, Computers & Chemical Engineering 47 (2012) 111–120, doi: 10.1016/j.compchemeng.2012.06.017.

[38] Beebe, S. Raymond, Predictive maintenance of pumps using condition monitoring, Elsevier, 2004.

[39] R. A. A. Habeeb, F. Nasaruddin, A. Gani, I. A. T. Hashem, E. Ahmed, M. Imran, Real-time big data processing for anomaly detection: A survey, International Journal of Information Management 45 (2019) 289–307, doi: 10.1016/j.ijinfomgt.2018.08.006.

[40] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. d. P. Francisco, J. P. Basto, S. G. Alcalá, A systematic literature review of machine learning methods applied to predictive maintenance, Computers & Industrial Engineering 137 (106024), doi: 10.1016/j.cie.2019.106024 (November 2019).

[41] H. Skima, K. Medjaher, N. Zerhouni, Accelerated life tests for prognostic and health management of mems devices., in: Second European Conference of the Prognostics and Health Management (PHM) Society., Vol. 2, 2014, pp. 1–7.

[42] T. Wang, Trajectory similarity based prediction for remaining useful life estimation, Ph.D. thesis, University of Cincinnati (August 2010).

[43] K. Javed, R. Gouriveau, N. Zerhouni, P. Nectoux, Enabling health monitoring approach based on vibration data for accurate prognostics, IEEE Transactions on Industrial Electronics 62 (1) (2015) 647–656, doi: 10.1109/TIE.2014.2327917.

[44] J. Coble, J. W. Hines, Identifying optimal prognostic parameters from data: A genetic algorithms approach, in: Annual Conference of the Prognostics and Health Management (PHM) Society, Vol. 1, 2009.

[45] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al., A density-based algorithm for discovering clusters in large spatial databases with noise, in: Knowledge Discovery and Data Mining, Second International Conference on (KDD-96), Vol. 96, 1996, pp. 226–231, url: https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf.

[46] J. Yadav, M. Sharma, A review of k-mean algorithm, in: International Journal of Engineering Trends and Technology (IJETT), Vol. 4, 2013, pp. 2972–2976, url: ijettjournal.org/volume-4/issue-7/IJETT-V4I7P139.pdf.

[47] R. Pamula, J. K. Deka, S. Nandi, An outlier detection method based on clustering, in: Emerging Applications of Information Technology, International Conference on, Vol. 0, 2011, pp. 253–256, doi: 10.1109/EAIT.2011.25.

[48] A. Roth, Vacuum technology, Elsevier, 2012.

[49] J. John, T. Keith, Gas Dynamics, Pearson Prentice Hall, 2006.

[50] A. A. Istratov, O. F. Vyvenko, Exponential analysis in physical phenomena, Review of Scientific Instruments 70 (2) (1999) 1233–1257.

[51] P. Berman, O. Levi, Y. Parmet, M. Saunders, Z. Wiesman, Laplace inversion of low-resolution nmr relaxometry data using sparse representation methods, Concepts in Magnetic Resonance Part A 42 (3) (2013) 72–88.

[52] I. Craig, A. Thompson, W. J. Thompson, Practical numerical algorithms why laplace transforms are difficult to invert numerically, Computers in Physics 8 (6) (1994) 648–653.

[53] I. J. Day, On the inversion of diffusion nmr data: Tikhonov regularization and optimal choice of the regularization parameter, Journal of Magnetic Resonance 211 (2) (2011) 178–185. doi:https://doi.org/10.1016/j.jmr.2011.05.014.

[54] D. Kumar, T. D. Nguyen, S. A. Gauthier, A. Raj, Bayesian algorithm using spatial priors for multiexponential t2 relaxometry from multiecho spin echo mri, Magnetic resonance in medicine 68 (5) (2012) 1536–1543.

[55] A. Levant, Robust exact differentiation via sliding mode technique, automatica 34 (3) (1998) 379–384.

[56] L. Ying, D. Xu, Z.-P. Liang, On tikhonov regularization for image reconstruction in parallel mri, in: The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 1, IEEE, 2004, pp. 1056–1059.

[57] M. Zimmermann, A.-M. Oros-Peusquens, E. Iordanishvili, S. Shin, S. D. Yun, Z. Abbas, N. J. Shah, Multi-exponential relaxometry using l_1-regularized iterative nnls (merlin) with application to myelin water fraction imaging, IEEE transactions on medical imaging 38 (11) (2019) 2676–2686.

[58] C. Truong, L. Oudre, N. Vayatis, Selective review of offline change point detection methods, Signal Processing 167 (2020) 107299.

[59] B. Jackson, J. Scargle, D. Barnes, S. Arabhi, A. Alt, P. Gioumousis, E. Gwin, P. Sangtrakulcharoen, L. Tan, T. T. Tsai, An algorithm for optimal partitioning of data on an interval, IEEE Signal Processing Letters 12 (2) (2005) 105–108. doi:10.1109/LSP.2001.838216.

[60] M. Lavielle, Using penalized contrasts for the change-point problem, Signal processing 85 (8) (2005) 1501–1510.

Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisors Prof. Luca Consolini and Prof. Gianluigi Ferrari who have supported and guided me over the last few years, with competence and patience. Their guidance helped and motivated me in all the time of my research. Thanks also to Gaia and the entire IoTLab, the pandemic has kept us away but certainly, without you, this PhD would not have been as exciting and fun. In addition, without their precious support, it would not have been possible to complete my research. Last but not least, I would like to thank my family, always ready to support and encourage me at any time in my life, and for always believing in me.