

UNIVERSITÀ DEGLI STUDI DI GENOVA - ITALY



---

Ph.D. IN MATHEMATICS AND APPLICATIONS

SARA GARBARINO

**Compartmental analysis  
in nuclear medicine:  
an inverse problem approach**

Advisor

**Prof. Michele Piana**

---

22 April 2015



*Keep calm and add a penalty term*



# Contents

<b>Introduction</b>	<b>i</b>
<b>1 Glucose metabolism</b>	<b>1</b>
1.1 Clinical aspects . . . . .	1
1.1.1 Structure and physiology . . . . .	1
1.1.2 Pathologies . . . . .	3
1.1.3 (impractical) glucose imaging . . . . .	4
1.2 FDG physiology . . . . .	4
1.2.1 Similarities and differences . . . . .	5
1.2.2 Problematic aspects . . . . .	6
<b>Part I: from photons to images</b>	<b>9</b>
<b>2 PET data formation</b>	<b>11</b>
2.1 Physical principles . . . . .	11
2.1.1 Resolution . . . . .	13
2.1.2 Attenuation . . . . .	14
2.1.3 Scattered Events . . . . .	15
2.1.4 Accidental Coincidences . . . . .	15
2.1.5 Detector Deadtime . . . . .	16
2.2 Data formation . . . . .	16
2.3 microPET scanner . . . . .	17
2.3.1 Albira system . . . . .	17
<b>3 PET imaging inverse problem</b>	<b>21</b>
3.1 PET imaging . . . . .	22
3.1.1 Two-dimensional imaging . . . . .	22

3.1.2	Fully three-dimensional imaging . . . . .	23
3.1.3	Imaging equation . . . . .	24
3.2	Inversion methods . . . . .	25
3.2.1	Analytical inverse formula . . . . .	25
3.2.2	Iterative techniques . . . . .	31
<b>Part II: from images to physiology</b>		<b>35</b>
<b>4</b>	<b>Compartmental models in a formal perspective</b>	<b>37</b>
4.1	Generalities on compartmental analysis . . . . .	38
4.1.1	Analytical method . . . . .	39
4.1.2	Graphical methods . . . . .	40
4.2	The $n$ -compartment system . . . . .	41
4.2.1	Properties of $n$ -compartment systems . . . . .	43
4.2.2	The $n$ -compartment catenary case . . . . .	45
<b>5</b>	<b>Compartmental inverse problem</b>	<b>47</b>
5.1	Identifiability . . . . .	47
5.1.1	Definitions . . . . .	47
5.1.2	Results on a $n$ -compartment system . . . . .	48
5.1.3	Study of a 2-compartment catenary system . . . . .	51
5.1.4	Study of a 3-compartment catenary system . . . . .	54
5.2	The Gauss-Newton Regularized method . . . . .	57
<b>Part III: from micro- to macro- physiology (and vice versa)</b>		<b>65</b>
<b>6</b>	<b>Application to real data: the liver problem</b>	<b>67</b>
6.1	Liver physiology . . . . .	67
6.1.1	Dual Input . . . . .	67
6.1.2	Pharmacological stimuli . . . . .	69
6.2	The liver compartmental problem . . . . .	69
6.2.1	Model equation . . . . .	71
6.2.2	Regularized Gauss Newton method on data . . . . .	72
6.2.3	Results . . . . .	74
<b>7</b>	<b>Application to real data: the kidneys problem</b>	<b>83</b>
7.1	Kidneys physiology . . . . .	84
7.1.1	Diabetes . . . . .	84

7.1.2	Reabsorption and excretion . . . . .	84
7.2	The urinary compartmental problem . . . . .	85
7.2.1	Model equation . . . . .	86
7.2.2	A Maximum Likelihood - Steepest Descent method . . . . .	90
7.2.3	Results . . . . .	93
<b>Bibliography</b>		<b>103</b>
<b>A Inverse problem theory</b>		<b>115</b>
A.1	The generalized inverse . . . . .	117
A.2	Compact linear operator . . . . .	119
<b>B Regularization methods</b>		<b>123</b>
<b>C Proof of Theorem 5.2</b>		<b>133</b>





# Introduction

One of the major challenges in the application of nuclear medicine to oncology is the quantification of glucose metabolism in cancerous tissues: its been known since the first decades of the XX century that tumor cells use a lot more glucose than normal cells and that cancer thrives on glucose. Thus methods that could quantify this abnormal uptake of glucose could be in principle extremely useful to make diagnosis, prognosis and clinical treatment of tumors.

Standardized Uptake Value (SUV), a relative measure of glucose-like-tracer uptake, is currently the most popular index used in clinical nuclear medicine studies to distinguish between cancerous and healthy tissues and to assess the efficacy of therapies. However, despite its popularity, the reliability of this simple index is somewhat controversial, because of its variability with respect to a large amount of different physiological and imaging factors. Furthermore, this index does not allow for an accurate description of the glucose metabolism in cells and therefore its interpretation is hampered when the patient, affected by cancer, is also affected by other disorders that alter glucose metabolism at the cellular level, making hard to accurately quantify the sugar uptake by cancer cells. The most popular pathology of this kind is *diabetes mellitus*, that is a group of metabolic diseases characterized by high blood glucose levels over a prolonged period. This condition modifies the diagnostic interpretation of SUV in diabetic patients with cancer because the ubiquitous presence of glucose in blood smoothens the difference between the pathologic uptake in cancer and the physiological tracer retention in the background surrounding tissue. For these reasons, methods that could locally quantify the flow and absorption of glucose in organs like kidneys or liver, thus potentially expressing a pathological behavior with respect to glucose metabolism, could be extremely useful for cancer diagnosis and treatment.

But what do we mean when we talk about *nuclear medicine data*? Essentially we talk about dynamic data of a specific metabolic activity (i.e., the metabolism of a specific molecule). These data are acquired via a functional nuclear medicine imaging technique, often coupled with a morphological imaging modality, in order

to dynamically obtain functional and anatomical images of the metabolism under study. Such images can be acquired, for example, by imaging the radioactive decay of isotopes bound to known molecules, with suitable metabolic properties. These imaging techniques are known as nuclear medicine techniques, and among them, Positron Emission Tomography (PET) is the most sensitive and specific one. PET is able to detect the time dependent spatial distribution of pico-molar quantities of a labeled tracer, which is diffused into a living body or one of its parts. In general, this modality aids diagnosis as well as treatment planning and monitoring of different diseases by producing images of functional processes inside the body. For these reasons, PET, in principle, could image glucose metabolism by means of the following procedure:

1. glucose is tagged by means of a positron emitter and the tagged glucose tracer is then injected into the blood;
2. blood delivers the glucose tracer into tissues and organs;
3. each emitted positron annihilates with a tissue electron, emitting two photons, in opposite directions, along a perfect straight line;
4. the distribution of radiation detected by collimators is a signature of glucose distribution in time and space (see Figure 2.1 for a pictorial representation).

In chemical terms, step 1. could be implemented, in principle, in three ways, according to the three possible modalities glucose could be tagged by a radioactive isotopes. However, none of them is practical, essentially owing to the fact that in all three cases the half lives of the glucose isotopes are not appropriate for dynamic imaging or they imply the necessity of an on-site cyclotron, which is rather expensive and needs for a nontrivial maintenance.

So *how can we image glucose metabolism?* The idea is to find a different radionuclide, analog to glucose that could mimic its behavior at cellular level. The most popular glucose analog is [ $^{18}\text{F}$ ]fluoro-2-deoxy-D-glucose (FDG); indeed a radioactivity elimination half-life of 110 minutes makes PET with FDG widely exploitable from a clinical viewpoint. In this sense, the uptake of FDG by tissue is a marker for the tissue uptake of glucose, which in turn is closely correlated with certain types of tissue metabolism: after the injection, FDG is taken up by high-glucose-using cells such as brain, kidneys, liver and cancer cells, and PET scanner can form 3-D images of the distribution of FDG within the body.

From a physiological viewpoint, although FDG is a glucose analog, there are several differences between the uptake and metabolism of FDG compared with normal glucose, that must be taken into account when we image glucose metabolism

by using FDG. From a computational viewpoint, the analysis of PET experiments involve two kinds of problems. *First*, signal processing techniques must be applied in order to reconstruct the time dependence and location of tracer concentration from the measured radioactivity. Of course, measured PET data are corrected to take into account scattering issues, physical decay of radioactive tracer, dead time of detector and attenuation. *Second*, these dynamic PET data can be processed in order to estimate physiological parameters that describe the functional behavior of the inspected tissues and the distribution of tracer between their different constituents.

When we deal with a PET experiment, we essentially deal with two different inverse problems: the first one is the reconstruction of the concentration of tracer, that leads to a quite classical imaging inverse problem, connected to the inversion of the Radon transform. We have to pass from the measured radioactivity, stored in sinograms, to dynamic images of tracer concentration. This is done by means of classical inversion methods, that rely on the analytical *filtered back projection* formula or *expectation-maximization*-based techniques. On the other hand, the second inverse problem is the quantification of tracer physiology and kinetics. We have to pass from images of tracer concentration to local metabolism markers and descriptors of the flow of tracer: following a model originated in the framework of pharmacokinetics, it is assumed that there are functionally separated pools of tracer (*compartments*), exchanging tracer between each other. Therefore the distribution and flow of tracer is described by a *compartmental system*.

The framework of compartmental analysis identifies different compartments in the physiological system of interest, each one characterized by a specific and homogeneous functional role. With the help of a global observation of the organ along time in PET-scan images, we aim at getting information on the radioactive tracer exchange rates between compartments. From a mathematical viewpoint, the time dependent concentrations of tracer in each compartment constitutes the state variables that can be determined from PET data. The time evolution of the state variables is modeled by a system of differential equations for the concentrations, enforcing tracer balance during exchange processes. As a first approximation, or when it is known that the exchange rates depend neither on time nor on the concentrations, the evolution of the system is governed by a linear system of ordinary differential equations with constant coefficients. The constant coefficients describe the input/output rate of tracer for each compartment and represent the physiological parameters assessing the systems metabolism; hence they are the unknown to be estimated. In this sense we have to solve an inverse problem.

The mathematical reduction of compartmental systems has classically relied on graphical methods, which consist in quite heuristic manipulation of the system of ODEs in order to find an easier equation that, by means of linear regression tech-

niques, allows to estimate combinations of the exchange coefficients. These methods (*Patlak* or *Logan methods*) have an obvious drawback: if there are several exchange coefficients to be estimated in the compartmental model, i.e. the organ or tissue we are modeling has a complex physiological behavior, or if more than one organ is involved in the compartmental system, we are able to estimate just few combinations of rate coefficients. Therefore, in some sense, we lose the capability of quantifying the markers of the cellular-level metabolism, which is however the intrinsic power of compartmental models.

In order to overcome these limitations, in this work we introduce a very general approach to compartmental analysis, based on the theory of the inverse problems. Actually, from a modeling viewpoint we introduce a systematic way to describe the systems and to write the model equations in a general form; from a methodological viewpoint we provide, for the first time, general results on the uniqueness of the solution of compartmental inverse problems, and accurate optimization techniques for their reduction; from the point of view of applications we are able to estimate, with a good degree of accuracy, many exchange coefficients involved in the FDG metabolism, by taking also into account the reversibility of FDG chemical process in cells, and provide satisfactory results on real data of murine models.

The thesis is structured as follows.

Chapter 1 describes the glucose metabolism in a body, focusing attention on problems arising from pathologies, as cancer and diabetes, in the first paragraph. The second paragraph provides an overview on FDG metabolism, its problematic aspects and similarities and differences with the glucose metabolism.

Chapters 2 and 3 are devoted to the *first* PET problem: specifically, Chapter 2 describes the PET imaging forward problem, providing an overview on the physical principles beyond a PET experiment and focusing on the data formation problem. Chapter 3 is devoted to the PET imaging inverse problem: the first paragraph introduces and provides details on 2D and 3D PET imaging modalities; the second paragraph describes both the analytical and the iterative inversion approaches for the reduction of the PET imaging inverse problem.

Chapters 4 and 5 concern the *second* problem: in two paragraphs, Chapter 4 describes generalities on compartmental models, classical results reported in literature and a formal description of general  $n$ -compartment systems. Chapter 5 is devoted to description of results on the uniqueness of the solution of the compartmental inverse problem. The first paragraph provides our results on the study of the inverse problem, in a general  $n$ -dimensional case and with a focus on the most popular 2- and 3-dimensional cases. The second paragraph describes a method for the reduction of the inverse problem: a multidimensional Gauss-Newton with a penalty term, accounting for the ill-posedness of the problem is described and details on the

implementation are provided.

Chapter 6 and 7 are devoted to applications of compartmental models to small-animal real data. Chapter 6 concerns the hepatic physiology, provides details on the specific glucose metabolism and thus presents the liver compartmental problem. The model equations for the forward and the inverse problems are described, and results on glucose metabolism in the hepatic system are provided. Chapter 7 describes the renal physiology, and again details on the the importance of kidneys behavior in the understanding of glucose metabolism are described. Thus the urinary compartmental problem is discussed and solved, and results on both synthetic and real data are shown.

A few, rather technical contents are discussed in the appendices. Appendix A describes the inverse problem framework, and standard results on the generalized inverse and compact linear operator are shown; Appendix B is dedicated to classical regularization techniques; Appendix C provides the formal proof of a uniqueness theorem, presented in Chapter 5.



# Chapter 1

## Glucose metabolism

Energy is required for the normal functioning of the organs in the body. Many tissues can use fat or protein as an energy source, but *all* of them use primarily glucose. Some of them also, such as brain and red blood cells, can use *only* glucose. Glucose is the cornerstone of life, and in this sense, imaging glucose metabolism would help understanding specific organ metabolism and its eventual alteration in order to make diagnosis, prognosis and clinical treatment of different possible disorders or diseases.

### 1.1 Clinical aspects

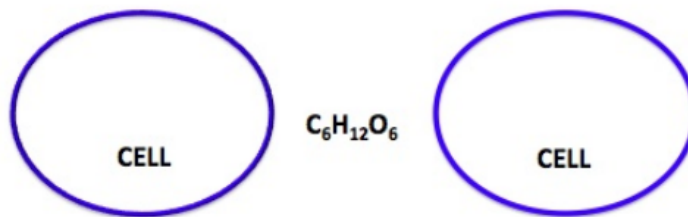
#### 1.1.1 Structure and physiology

Glucose is a simple aldonic monosaccharide that is absorbed directly into the bloodstream during digestion: it has the molecular formula  $C_6H_{12}O_6$  and it is a ubiquitous supply of power in life. It is used as an energy source in most organisms, from bacteria to humans: its use, indeed, may be by either aerobic respiration, anaerobic respiration, or fermentation. In human body, glucose is the key source of energy, through aerobic respiration, providing about 16 kilojoules of food energy per gram.

The metabolism of a glucose molecule is extremely complicated and involve several different pathways. However, using the very simplified scheme in Figure 1.1 one can describe the way glucose is metabolized by cells according to a two-way destiny: 1) it can remain in the intracellular space, essentially the space between cells, at hypothetical disposal of all the cells about, or 2) it can be phosphorylated by hexokinase (becoming glucose-6-phosphate, G6P) and entering a specific cell membrane. Phosphorylation is the addition of a phosphate group to an organic molecule, and

hexokinase is an enzyme that phosphorylates sugars, forming phosphate. In living body glucose is the most important substrate of hexokinases, and G6P the most important product of such a reaction.

first path



second path

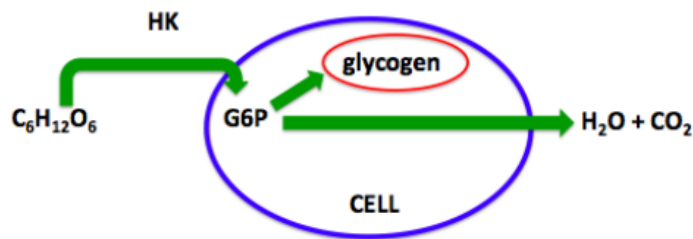


Figure 1.1: A two-way destiny.

Once glucose, in his G6P form, has entered the cell membrane it can be both stored as glycogen or - if it is needed immediately upon entering the cell to supply energy - it begins the metabolic process of glycolysis, being “burnt” from the cell in order to produce energy (becaming  $H_2O + CO_2$ ). If blood glucose levels are high, the body needs a way to store the excess glucose, and so, as previously stated, after being converted to G6P, the molecule can be turned into glycogen. When the body needs glucose for energy, glycogen phosphorylase can cleave away a molecule from the glycogen chain. The cleaved molecule can be converted into G6P by phosphoglucomutase. Next, the phosphoryl group on G6P can be cleaved by glucose6phosphatase (G6Pase) so that a free glucose can be formed. This free glucose can pass through membranes and can enter the bloodstream to travel to other places in the body.



### 1.1.2 Pathologies

There exist two important pathologies that can affect glucose metabolism in human body. The first one is *diabetes mellitus*, that is a group of metabolic disease in which there are high blood sugar levels over a prolonged period, and, if left untreated, can cause many complications and even lead to death. The second one is *cancer*, which, in few words, is a group of diseases involving abnormal cell growth with the potential to invade or spread to other parts of the body, and in some sense, it is strongly connected to glucose metabolism.

#### Diabetes<sup>1</sup>

Blood sugar concentrations are controlled by different hormones: insulin is the one related to diabetes. If the concentration of glucose in the blood is too high, insulin is secreted by the pancreas and stimulates the transfer of glucose into the cells, especially in the liver and muscles. Diabetes is due to either the pancreas not producing enough insulin or the cells of the body not responding properly to the insulin produced: as a result, glucose concentration in blood is altered.

#### Cancer<sup>2</sup>

One of the most logical and way to treat cancer is obviously to cut it off the supply of food, starving tumors and cancer cells with a lack of glucose. Consequently, a strategy for the selective starvation of tumors has been one of the major goals in oncology over the past decades. For this purpose, information that could improve our knowledge of glucose metabolism of cancer cells, could be extremely useful in cancer defeating. It is been known from almost a hundred of years that tumor cells use a lot more glucose than normal cells and that cancer cells thrive on glucose. Indeed, in his work of 1924 [110], Warburg stated: “the prime cause of cancer is the replacement of the respiration of oxygen in normal body cells by fermentation, and fermentation requires sugar”. Thus cancer requires sugar, and since its metabolism is approximately eight times greater than the metabolism of normal cells, cancer needs *a lot* of glucose; and while normal cells need oxygen, cancer cells despise it. Warburg’s hypothesis was in fact that cancer growth was caused when cancer

---

<sup>1</sup>As of 2014, an estimated 387 million people have diabetes worldwide: this is equal to 8.3% of the adult population, with equal rates in both women and men. In the years 2012 to 2014, diabetes is estimated to have resulted in 1.5 to 4.9 million deaths per year. Furthermore, the global economic cost of diabetes in 2014 was estimated to be \$ 612 billion USD.

<sup>2</sup>Glucose cell metabolism: it’s all about cancer [111]!

cells converted glucose into energy without using oxygen: essentially that cancer cells increase glucose uptake in aerobic conditions. In addition, many cancer cells hyperactivate insulin receptors and thus evidences exists that chronically elevated blood glucose and insulin levels facilitate tumor genesis and worsen the outcome in cancer patients. Furthermore, talking about connections between diabetes and cancer, Braunstein observed in 1921 that in the diabetic patients that developed cancer, glucose secretion in the urine disappeared, adding another link between the glucose metabolism and the cancer one.

Therefore, *imaging glucose metabolism would help diagnosis, prognosis and therapy of cancer.*

### 1.1.3 (impractical) glucose imaging

Positron Emission Tomography (PET) is a functional medical imaging modality<sup>3</sup> that, in principle, could image glucose metabolism by means of a specific procedure. First, glucose ( $C_6H_{12}O_6$ ) is tagged by means of a positron emitter and the tagged glucose tracer is then injected into the blood. Each emitted positron annihilates with a tissue electron, emitting light along a perfect straight line and collimators collect the emitted light, whose distribution is a signature of glucose distribution in time and space.

This procedure could be implemented, in principle, in three ways, i.e. the three possible ways to tag glucose:  $O_{15}$ , which is a radioactive isotope of oxygen, *Tritium*, radioactive isotope of hydrogen, and  $C_{11}$ , of carbonium. Unfortunately, Tritium is an almost stable isotope, since decays in 2000 years, rendering it totally impractical; furthermore the half-lives of  $O_{15}$  and  $C_{11}$  (120 seconds and 20 minutes respectively) imply the the necessity of an on-site cyclotron, which is rather expensive and needs for a nontrivial maintenance.

So *how can we image glucose metabolism?* The idea is to find a different radionuclide, analog to glucose, that could mimic, in the best way, its behavior at cellular level. For this purpose,  $[^{18}F]$ fluoro-2-deoxy-D-glucose has been synthesized, in 1968, by Dr. Josef Pacak in Charles University, Czechoslovakia.

## 1.2 FDG physiology

$[^{18}F]$ fluoro-2-deoxy-D-glucose (FDG) is a radio-pharmaceutical used in PET. FDG has a radioactivity elimination half-life of 110 minutes, rendering PET with FDG

---

<sup>3</sup>See Chapter 2 for details.

widely exploitable from a clinical viewpoint. Chemically FDG is a glucose analog, with the positron-emitting radioactive isotope fluorine-18 substituted for the hydroxyl group at the 2' position in the glucose molecule. In this sense, the uptake of FDG by tissue is a marker for the tissue uptake of glucose.

### 1.2.1 Similarities and differences

FDG is one of the most utilized tracers for PET applications to oncology. FDG-PET [32, 46, 92] is a PET modality in which FDG is used as a tracer to identify several kinds of tumors, since malignant cells typically exhibit an increased FDG uptake in PET scans. FDG-PET has represented a breakthrough in the estimate of cancer metabolic activity and it is successfully used not only for the characterization of cancer infiltration but also for the monitoring of the tissue response to chemio- or radio-therapy. In fact, glucose consumption by malignant cells tends to increase [65], rendering FDG uptake a valuable index of cancer aggressiveness and patient outcome [1]. In this sense, the uptake of FDG by tissue is a marker for the tissue uptake of glucose, which in turn is closely correlated with certain types of tissue metabolism: after FDG is injected into a patient, PET scanner can form 3-D images of the distribution of FDG within the body on a discrete set of time values. After the injection, FDG is taken up by high-glucose-using cells such as brain, kidneys, liver and cancer cells, where phosphorylation prevents the glucose from being released again from the cells, once it is absorbed. Indeed, such caging mechanism acts also for FDG, forming [ $^{18}\text{F}$ ]-FDG-6-phosphate, (FDG-6P), which is trapped inside cells. Again, also for FDG a dephosphorylation mechanism of FDG-6P exists, making FDG again available and freely exchangeable between bloodstream and cells.

Although FDG is a glucose analog, there are several differences between the uptake and metabolism of FDG compared with normal glucose. Both molecules, for example, are transported into cells by the same proteins and both are phosphorylated. However, FDG is trapped within the cell as a consequence of phosphorylation and therefore it cannot be further metabolized (see Figure 1.2); unlabeled glucose, on the contrary, is either stored as glycogen or rapidly metabolized (see a pictorial representation in Figure 1.1). So if we want to be able to image glucose metabolism by using FDG, we have to take into account this fundamental difference; but with FDG we are now able to follow glucose path at least up to a certain point.

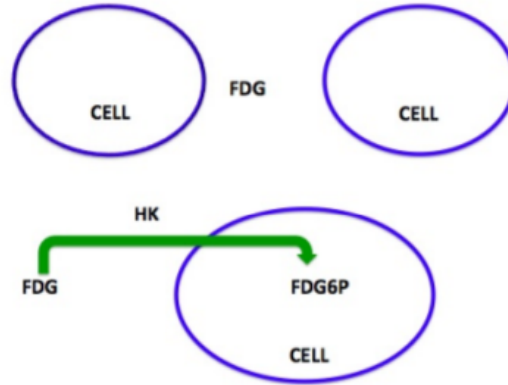


Figure 1.2: A one-way destiny.

### 1.2.2 Problematic aspects

In addition to cancer detection and staging, FDG-PET imaging is becoming more important as a quantitative monitor of individual response to therapy and an evaluation tool for new drug therapies and changes in FDG accumulation have been shown to be useful as an imaging biomarker for assessing response to therapy [112]. However, FDG uptake in tumors is related in a quite complex manner to the proliferative activity of malignant tissue and to the number of viable tumor cells [34] and besides the previously described differences between glucose and FDG, there are potential additional difficulties in FDG imaging. Indeed FDG uptake in tissues is an index that can vary dramatically depending on the body metabolism and specific tissue functionalities.

Under diabetic condition, for example, analysis of FDG distribution is often problematic, as the limited uptake of glucose and FDG by insulin dependent tissues is paralleled by the persistence of high tracer concentrations in plasma. Obviously, the ubiquitous presence of radioactive blood smoothens the difference between the pathologic uptake in cancer and the physiological tracer retention in background surrounding tissues. Accordingly, interventions able to increase urinary FDG loss might improve image quality in the large number of diabetic patients submitted to PET/CT for cancer studies. However, testing this hypothesis asks to first verify whether and to what extent a fraction of urinary FDG is indeed reabsorbed by renal tubule, implying the availability of methods able to describe the different mechanisms tuning FDG kinetics within the urinary apparatus.

On the other hand, Metformin is a very used oral antidiabetic drug that reduces blood glucose concentration without causing hypoglycemia [52], mostly by decreasing intestinal glucose absorption and glucose delivery by the liver [71]. It follows that PET images of diabetic patient, treated with metformin, are also problematic to analyze, because of the reduction of the FDG dephosphorylation induced by metformin. It follows also that methods able to quantify this reduction can provide notable improvement in the correction of the uptakes for metformin-treated diabetic patients [28].

So, even if FDG uptake is a valuable index of cancer aggressiveness, clinical activity asks to represent this variable in a standard and normalized way, as Standard Uptake Value (SUV), an index that should be strictly correlated with the glucose metabolism; unfortunately, it has been observed that differences between FDG and glucose metabolism can reflect to an erroneous SUV interpretation and then a tumor misplacing or under-estimation. On the other hand, the estimate of the SUV itself is hampered by the interferences of several variables modulating tracer availability independently of tumor metabolism [19].

The SUV is commonly defined as the ratio between the tissue radioactivity concentration [kBq/g] and the injected activity [kBq] divided by the body weight [g], at a time point  $t$ :

$$SUV(t) = \frac{C(t)}{\frac{C_{injected}(t)}{W}}. \quad (1.1)$$

The primary problem with the SUV is that it is subject to too many sources of variability which are hard to control or even to take into account, as body composition and habits, length of the uptake period, kidney functionality and hepatocytes functions in dephosphorylation process [45].

### Renal excretion

Kidney function plays a relevant role in the quantification of SUV and in the correlation between glucose and FDG, as, differently from the tracked glucose, FDG is scarcely reabsorbed in renal tubule and is largely excreted in the urines [100], which implies that there is radioactivity in the bladder. We point out that this issue has a significant consequence as far as the effectiveness of the imaging procedure is concerned: in fact, the more radioactivity is in the bladder the less amount of tracer is available for tumor identification. Urinary loss of glucose and of FDG might modify the diagnostic interpretation of PET/CT studies in patients with cancer, especially if the patient suffers of diabetes [79].

Methods able to efficiently describe the renal physiology and the different mechanism of excretion-reabsorption in kidneys can be useful to normalize SUV rendering this index more reliable from a prognostic viewpoint [26, 27].

### **Hepatocyte behavior**

The liver plays a crucial role in the maintenance of glycemic levels despite a large variability in glucose consumption in the different body tissues. To this purpose, liver metabolism is regulated as to counterbalance the hyperglycemic response to food intake and to adapt serum glucose concentration to whole body needs under fasting periods. This buffer function is of pivotal importance since - once entered in the cells of almost all mammalian - glucose is sequestered by its phosphorylation to G6P. The irreversible nature of this reaction also applies to FDG whose accumulation as FDG-6P accounts for the potential of this tracer in the study of brain and tumor metabolism [103]. Coherently with its role in maintaining metabolic homeostasis, this simplified model does not apply to hepatocytes whose G6P content -either newly synthesized or coming from glycogen breakdown - can be dephosphorilated by the glucose6phosphatase (G6Pase). This reaction replenishes cytosol content of glucose to be released into the bloodstream [76, 78].

Furthermore, the Deauville criterion [53, 21], which is a visual interpretation of the PET scans - worldwide used to diagnose the presence and the aggressiveness of lymphoma - is a criterion to classify the FDG uptake, expressed in SUV, of the potentially-lymphoma-tissue. The threshold for the diagnosis of a lymphoma is whether such a SUV is higher or lower then the liver SUV.

For these reasons, methods able to describe efficiently liver physiology, to take into account the hepatocyte capability to dephosphorilate FDG and to correct liver FDG uptake, and SUV, accordingly, can be useful to render SUV more reliable from a diagnostic viewpoint [28].

# Part I: from photons to images

The next two chapters describe the *first* inverse problem found when dealing with PET and glucose imaging: how to pass from the measured PET data to a comprehensible image of the body under scans. PET detects photons, emitted by the annihilation of a tracer-decayed-positron with a tissue-electron: photons are collected by specific collimators and mathematically encode position and time of positron decay, that is what we want to reconstruct. The inverse problem of determining dynamic images of tracer distribution in the body is ill-posed and hence not easy to handle, from a mathematical viewpoint.

Chapter 2 deals with the PET data formation forward problem, describing *how* decay of a specific radionuclide (essentially FDG, but it applies in general) can lead to data collected into sinograms, and *where* information on position and time of decays are encoded.

Chapter 3 describes the PET imaging inverse problem of retrieving a dynamic image of radionuclide distribution from collections of emitted photons. The equation that describes the PET imaging problem is related to the well known *Radon transform*, so that the inverse problem of recovering the PET image is an *inversion of Radon transform* problem.





# Chapter 2

## PET data formation

Medical imaging is commonly thought as a method to look at anatomical details and structures in human body: in this sense, x-ray Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) - the most known morphological medical imaging techniques - yield very detailed images of such structures. However, it could be extremely useful to acquire also images of physiologic functions in human body, possibly combined with anatomy. This kind of images can be obtained, for example, by imaging the radioactive decay of isotopes constrained to known molecules, with suitable properties related to metabolism. The imaging techniques of this class are known as *nuclear medicine techniques*, and among them, Positron Emission Tomography is the most sensitive and specific. [7, 20, 36, 102, 66, 62, 63, 64, 72, 77, 94, 119].

This Chapter is devoted to the description of the physical principles beyond PET and to the derivation of the equation for data formation: essentially, the PET *forward* problem. The two inverse problems we have to deal with, in order to be able to describe the glucose metabolism, are handled in Chapter 3 and Chapter 4. A mathematical description of the inverse problem world and the related reduction is found in Appendix A and Appendix B.

### 2.1 Physical principles

Positron Emission Tomography (PET) [86] is an imaging technique capable of detecting the time dependent spatial distribution of pico-molar quantities of a labeled tracer which is diffused into a living body or one of its parts, such as a specific organ or tissue. In general, this modality aids diagnosis as well as treatment planning and

monitoring of different diseases by producing images of functional processes in the body. A PET study begins with the injection of a radiopharmaceutical. The scan takes place after a delay ranging from seconds to minutes, to allow for transport and uptake by the organ of interest. When the radio-isotopes decays, it emits a positron ( $\beta^+$ ), which travel a short distance (typically 1mm) before annihilating with an electron ( $\beta^-$ ). The annihilation produces two high-energy (511 keV) photons propagating in opposite directions, along a line called Line Of Response (LOR). The

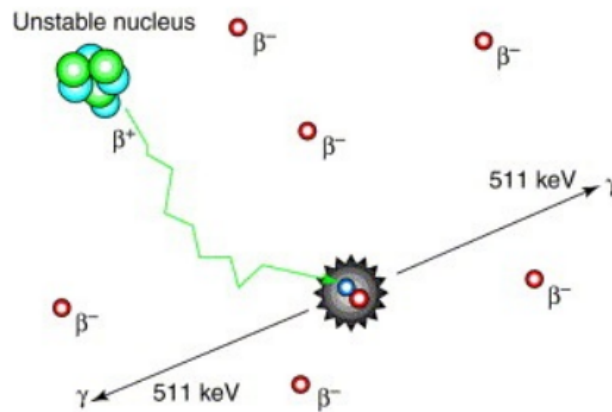


Figure 2.1: Pictorial representation of the decay of the radio-isotope and the annihilation with an electron, producing two 511 keV  $\gamma$  rays.

gamma rays are detected by photomultiplier tubes when they reach a scintillator in the scanning device, suitably positioned. The scintillator crystal - photomultiplier devices are called *detectors*, and typically they are combined to form a cylindrical ring, that surround the body or the organ under PET study.

If two photons are received within a short timing window, and are not scattered, an event is recorded along the line connecting the two detectors that have detected the gamma rays. Summing many such events results in quantities that approximate line integrals through the radio-isotope distribution, which is the quantity that we want to reconstruct, in order to provide images of the radio-pharmaceutical activity in the body. The number and the quality of the counts detected can heavily affect the validity of this approximation: the more reliable counts one is able to collect, the more accurate approximation of the line integrals is obtained, which leads to a better reconstruction of the radio-isotope distribution. For this reason, in addition to a timing window, an energy window is implemented, in order to collect photons

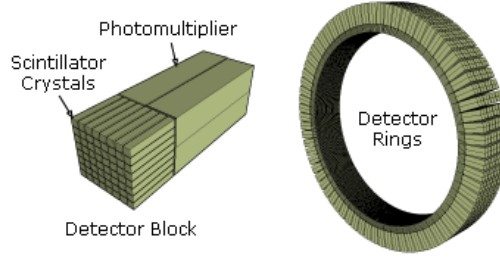


Figure 2.2: Example of a PET detector ring, divided into blocks, made of scintillator crystals.

not only at exactly 511 keV, but in a small energy interval around. Furthermore, corrections for scatter phenomena and accidental coincidence events can be added, to collect more photons and improve the number of counts.

Next sections are devoted to a brief overview on the physical principles beyond the PET image formation: for a complete discussion on these phenomena see [80, 86, 50, 16].

### 2.1.1 Resolution

The main physical characteristics that control the spatial resolution in PET scanners are essentially the width of the crystal in a detector block, the positron range, the penetration of gamma rays and the sampling error. The dominant factor is the first one, since the position of photon interaction within the crystal is not determined. The rectangles in Figure 2.3 represent two opposite set of crystals (three crystals placed on opposite sides of PET detector ring). When a pair of photons reach the two crystals, the coincidence rate between the detector pair in question (the central part of rectangles) describes a triangle. Thus the Response Function for this LOR is a triangle whose full width at half maximum (fwhm) is  $d/2$ , where  $d$  is the width of the crystal. Furthermore, there is a resolution problem connected to positron range. The positron, once it is ejected from the nucleus of the radio-isotope, travels some millimeters before it annihilates with an electron: thus the position where the photons are created differs from the position of the nucleus, and there is necessarily some blurring effect, whose magnitude is bound to the radio-isotope species. Another effect that can degrades spatial resolution is the penetration of gamma rays into the detector ring. The 511 keV photons often penetrate some distance before they interact and are detected: in this sense they

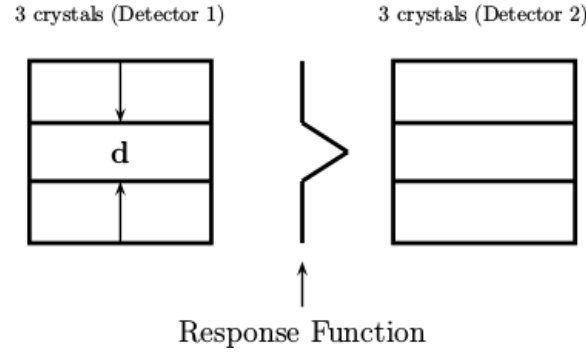


Figure 2.3: Coincidence Response Function between two detector element (the central part of rectangles).  $y$ -axis of Response Function is position and  $x$ -axis is the coincidence rate.

may interact in a crystal different from the one that they impinge upon, and so get assigned to the wrong crystal. The consequent blurring effect is in the radial direction and depends on the detector material. The final effect that we have to take into account to talk about spatial resolution in PET is the sampling error. To collect the total number in coincident events recorded by a LOR, it is necessary to sample the Field Of View (FOV) of the scanner and count the total coincidences for all the pixels in the FOV. Figure 2.4 represents the collection of all the LORs acquired during a PET experiment (cylindric scanner), and it is clear how this sampling error effect is especially pronounced near the center of the scanner. All these effects can be collected together in the so called *intrinsic resolution*. In clinical practice one has to add statistical noise that can be furthermore amplified by the image-reconstruction algorithm and degrades the resolution of images. So the *reconstructed spatial resolution* depends also on the reconstruction algorithm used, and occurs in modifying the equation for the intrinsic resolution by other factors.

### 2.1.2 Attenuation

At 511 keV there are just two possible interactions: photoelectric absorption and Compton scatter, but the incidence of the first phenomenon can be considered as negligible for photons in body tissue [17]. In a Compton interaction a photon interacts with an electron that deflects its path, causing a loss of energy. In PET, most scattered photons are deflected out of the FOV and are lost: the effect of these interactions is termed *attenuation*. The probability of a photon not to interact with

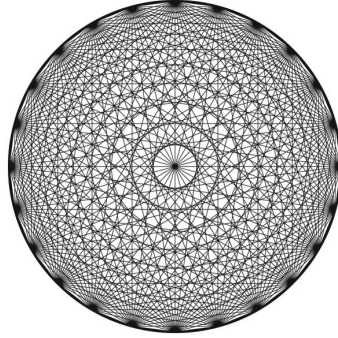


Figure 2.4: The lines of response in a cylindric PET scanner. The sampling depends on the position in the field of view: at the center the pixels are very well sampled, while nearby pixels are not.

an electron as it propagates along its LOR is called *narrow-beam attenuation* and it is independent of the position of the annihilation along the LOR, rendering possible a simple precorrection of the data [50].

### 2.1.3 Scattered Events

Those photon-electron annihilations for which one or both photons are scattered, but both are still detected, are the *scattered* events, as shown in Figure 2.5 (left). These events are evidently misplaced because of the non collinearity of the photons' path, but, since photons lose a fraction of their energy when they are scattered, they can be discriminated from unscattered photons by measuring the energy field in the crystal. Although this measure is just an approximation, it can be used as a simple threshold to reject a significant fraction of the scattered detectors, in order to improve the effectiveness of energy discrimination and the accuracy of scatter correction [31].

### 2.1.4 Accidental Coincidences

Given the large number of scattered photons and the relatively small solid angle subtended by the detector ring, it is quite evident that for many annihilations only one of the photons will be detected. These events are termed *singles*. If two singles arising from separate annihilations are detected within the same coincidence tim-

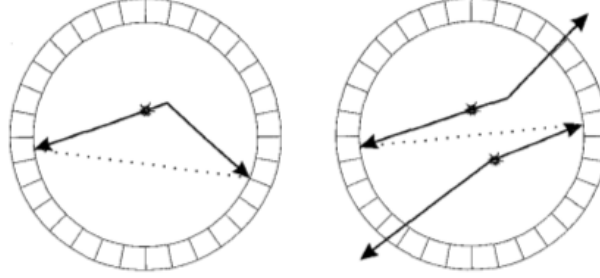


Figure 2.5: Diagram of a scattered event (left) and an accidental coincidence (right).

ing window, they will be recorded as shown in Figure 2.5 (right) and these events are termed *accidental coincidences* or randoms. The incident singles rate are proportional to the amount of injected isotope thus determining - together with the detector dead-time - the upper limit on the injected dose in the study.

### 2.1.5 Detector Deadtime

The counting rate of a PET scanner is bounded by the time required to process an event: each event processing begins with detection of a photon impulse by the detector involved, thus the impulse is integrated for some time interval and the scanner performs position calculations and energy discrimination. The involved detector is *dead* to new events during this time, and it is not able to detect and measure any of the photons that collide its crystal.

## 2.2 Data formation

If statistical effects are ignored, all the previous physical factors can be incorporated into a model for the total number of recorded events, leading to

$$Y = \gamma_d (\eta^t \eta^a M + \eta^r r + \eta^f f), \quad (2.1)$$

where  $Y$  is the total number of recorded events,  $M$  is the number of annihilations with photons emitted along the LOR  $L(s, \phi)$ ,  $\eta^t$  is the probability of detection for true events,  $\eta^a$  is the narrow-beam attenuation,  $\eta^r$  is the probability of detection for accidental coincidences,  $r$  is the number of accidental coincidences,  $\eta^f$  is the

probability of detection for scattered events,  $f$  is the number of scattered events,  $\gamma_d$  is the probability of an event not being lost due to deadtime and dependence on  $(s, \phi)$  is omitted but implied. The effects included in Eq. (2.1) have been characterized in the last years and details about the algorithms and techniques that have been implemented can be found in [107, 108, 87, 73, 116, 11, 24, 15].

## 2.3 microPET scanner

The bridge between discoveries at molecular level and implementation of relevant clinically diagnostics is represented by small-animal models, murine models in particular, and a large number of small-animal positron emission tomography (microPET) [56, 51, 91] systems have been developed and have become commercially available during the past years with this purpose. Although in mice less attenuation and scatter than in humans can be expected, these phenomena are still relevant [39, 38], in addition, when quantitative information on small target size is needed, microPET images can suffer of loss of resolution. It is thus necessary to increase the sensibility of scanners by developing more powerful detectors and new materials, in order to ensure enough spatial resolution for microPET imaging: in the past twenty years progresses have been great: nowadays, spatial resolution of microPET scanner is around 2-3 mm.

### 2.3.1 Albira system

The Albira scanner, illustrated in Figure 2.6, is an imaging system used at IRCCS San Martino-IST for medical research on rodents [13] and it is the scanner we have used in the past years for all the experiments performed on mice, to study glucose metabolism. With this scanner, data can be acquired in PET, SPECT or CT mode, and these three acquisition systems can also be used together in order to obtain better functional, structural and combined images.

The *Albira* microPET scanner at Dipartimento di Medicina Nucleare, IRCCS San Martino of Genova, from Bruker technology, is made up of two different detectors rings, to acquire oblique planes in a fully 3-D mode<sup>1</sup>. Each ring is made of 8 identical detectors, each including a scintillation crystal and a photo-detection system, forming an octagon. The crystals are trapezoidal in shape,  $40 \times 40 \text{ mm}^2$  in the front and  $50 \times 50 \text{ mm}^2$  in the back, as in Figure 2.7. This system has an axial FOV of 40 mm and a transaxial FOV of 80 mm in diameter. The crystals top

<sup>1</sup>Please refer to 3.1.2 for details about 3D acquisition.



Figure 2.6: Albira microPET system (Bruker Corporation).

faces are discretized into  $300 \times 300$  pixels and there is a precise order for labeling the 16 detectors, which can be seen in Figure 2.8. When a positron annihilation

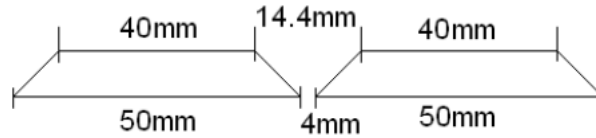


Figure 2.7: Pictorial representation of Albira crystals, with detectors' number ranging from 0 to 15.

occurs, the two scattered photons are collected by a couple of detectors, and the  $(x, y)$  coordinates of the two impact points are stored. Not all pairs of detectors are allowed for a coincidence event. Referring to Figure 2.8, the detector 0 can only be coupled with the three detectors belonging to the same ring and located on the opposite side (detectors 3, 4 and 5), and with their three counterparts belonging to the other ring (detectors 11, 12 and 13): by repeating this coupling for all the detectors, we have 48 possible pairs of detector, recording coincidences in the FOV. For full details on Albira system refer to [96]. A typical image acquired with Albira PET system is shown in Figure 2.9.



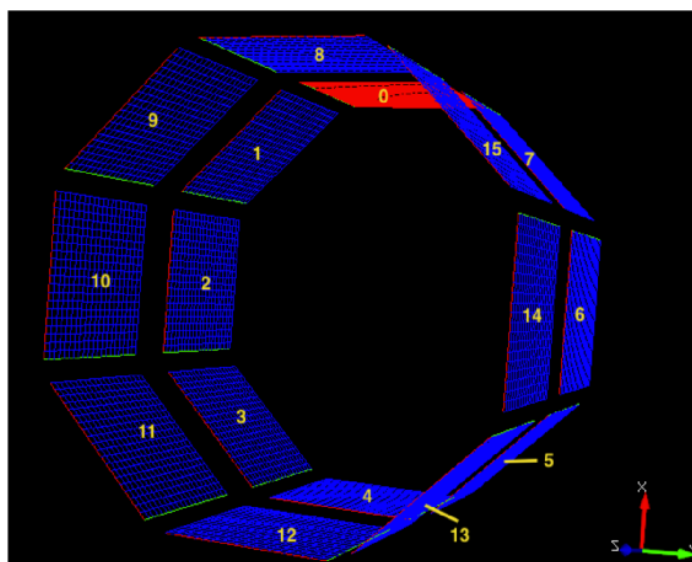


Figure 2.8: Detectors and their orientation in Albira.

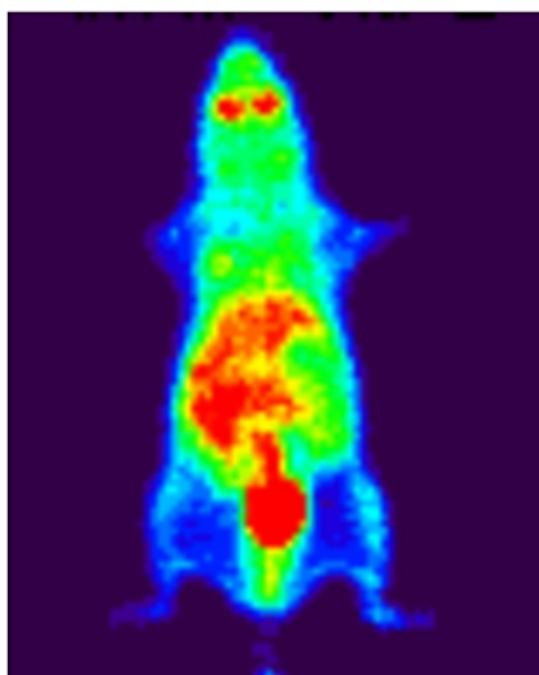


Figure 2.9: Mice acquired with Albira, using FDG.

# Chapter 3

## PET imaging inverse problem

The goal of PET image reconstruction is to recover the time dependent radio-tracer concentration from the PET measurement, collected into the so called *sinograms*. Such a problem is a well known *inverse problem*<sup>1</sup>, the first one we have to deal with in order to be able to describe the glucose metabolism inside the body, which is the aim of this chapter. The description of the methods for the reduction of the PET imaging inverse problem is the aim of the present chapter. For a description of the second inverse problem (how to pass from PET images to a quantitative description of the glucose metabolism) please refer to Chapter 4.

For Positron Emission Tomography there exist essentially two different approaches to image reconstruction: one is based on the analytical inverse formula, and the other one is based on iterative techniques. With regard to the first approach, *filtered back projection* (FBP) was until few years ago the most used method of image reconstruction in clinical practice [55, 84], thanks to its speed that allows for almost on-line reconstruction. The approaches based on iterative techniques are essentially statistical method [58, 113], require specific assumption on data noise (Gaussian or Poisson), and differ on the balancing between accuracy of reconstruction and speed of reconstruction. Among the iterative techniques we mention the Maximum Likelihood Expectation Maximization (MLEM) technique [60, 99], while in clinical practice the most commonly used iterative technique [37] is the Ordered Subsets EM (OSEM) [35], which is a simple but powerful modification of MLEM, capable of maintaining almost all the accuracy, while notably increasing the speed of reconstruction.

---

<sup>1</sup>For theory and methods for the reduction of Inverse Problems, please refer to Appendix A and Appendix B.

This chapter is devoted to the PET image reconstruction problem, with a brief review of different acquisition methods and a more detailed description of the two main inversion methods: FBP and ML-based techniques.

## 3.1 PET imaging

### 3.1.1 Two-dimensional imaging

LORs lying within a specified imaging plane - called *direct* planes - are the only LORs considered by two-dimensional PET imaging; the acquired data are collocated along LORs through the object, called  $f(x, y)$ , as shown in Figure 3.1. The LORs are organized into sets of projections, i.e. line integrals for all  $s$  for a fixed direction  $\phi$ . The collection of all projections  $p(s, \phi)$  forms the two-dimensional sinogram of the object  $f(x, y)$ , because in the case of a point source,  $s$  varies sinusoidally with  $\phi$  when it is stored in an array with columns indexed by  $s$  and the rows by  $\phi$ . A sinogram for a general object will be the superposition of all sinusoids corresponding to each point of activity in the object (see Figure 3.2).

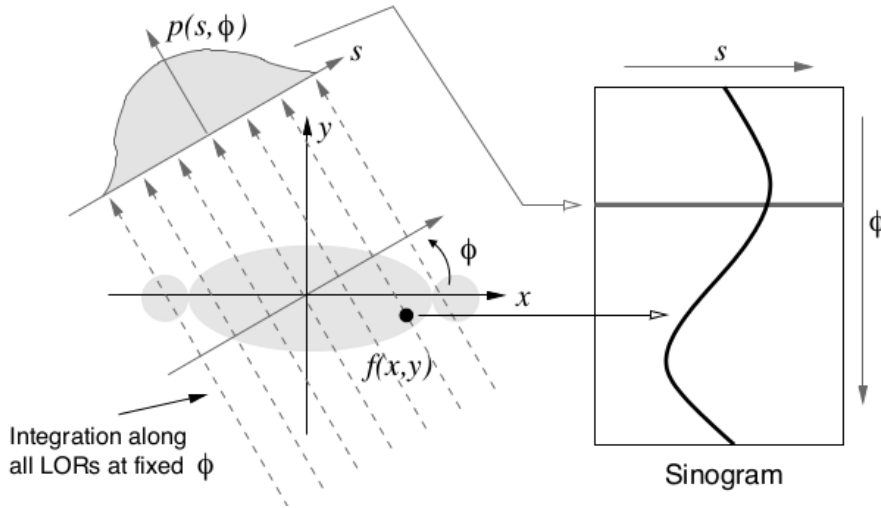


Figure 3.1: A projection,  $p(s, \phi)$  formed from integration along all parallel LORs at an angle  $\phi$ . The projections are organized into a sinogram.



Figure 3.2: During data acquisition, the object of emission activity on the left forms the sinogram on the right.

In order to image a volumetric object, it is necessary to repeat the 2D acquisition for multiple axial ( $z$ ) slices: thus, when the sinogram for each value of  $z$  is reconstructed, it is possible to stack the image planes together to form a three-dimensional image of  $f(x, y, z)$ . Although this can be considered a form of three-dimensional imaging, it is different from the *fully* three-dimensional acquisition model described in the next section.

### 3.1.2 Fully three-dimensional imaging

In fully three-dimensional PET imaging, it is possible to acquire both the direct planes as well as the line-integral data lying on different and oblique imaging planes: it is clear that fully 3D PET scanners increase sensitivity, and thus lower the photon counting noise, improving the quality of the reconstructed image. On the other hand, fully 3D measurements require more storage with data set sizes approximately  $10^3$  times larger than 2D measurements and thus more computational burden; moreover, fully 3D measurements contain significantly more scattered events. Advances in hardware storage, software implementation and parallel computation, coupled with progress in scatter correction techniques, have been of help in alleviating these problems [85, 115].

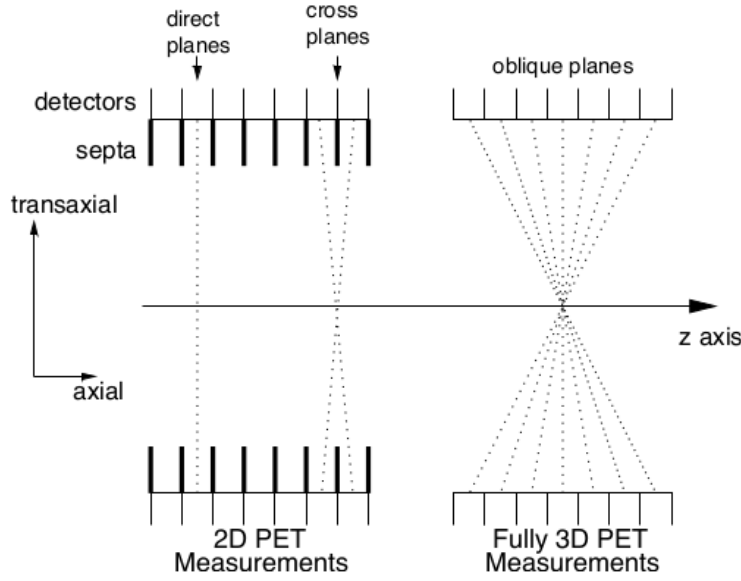


Figure 3.3: Comparison of fully 3D and 2D PET measurements. In 2D mode, scanner only collects direct and cross planes. In fully 3D mode, scanner collects also oblique planes.

### 3.1.3 Imaging equation

A compact representation of the PET imaging system is given by the following linear relationship, that accounts also for the noise:

$$\mathbf{p} = \mathbf{R}\mathbf{f} + \mathbf{n}. \quad (3.1)$$

Here  $\mathbf{p}$  is the set of observation (the sinogram),  $\mathbf{R}$  is the known system model (accounting for all the effects described in Section 2.1),  $\mathbf{f}$  is the unknown image and  $\mathbf{n}$  is the noise. The goal of reconstruction is to use the projections  $\mathbf{p}$  through the unknown object to find the image  $\mathbf{f}$ : the data values can be modeled either as deterministic or stochastic (random) variables.

If we assume that the datum is deterministic, and is not affected by statistical noise, then  $\mathbf{n}$  is a deterministic number. If  $\mathbf{n}$  is known, we can find the exact solution for the image (a deterministic image from a deterministic set of data). Methods involving the inverse of the discrete Radon transform are used to address this problem,

providing an analytical mathematical solution for the image  $\mathbf{f}$  from known projections,  $\mathbf{p}$ . In this sense, the deterministic assumption allows for a fast, predictable and analytical solution; obviously, the noise structures in the observations are disregarded by these methods. Consequently, this condition can lead to images with low resolution and poor noise properties (often in the form of streaking artifacts).

Actually, the data values are intrinsically stochastic, due to several factors: the positron decay process, the effects of attenuation, the addition of scattered and random events, and the photon detection process. As a consequence,  $\mathbf{n}$  can be described by a random process (random noise) making it impossible to find the exact solution for the problem. Therefore, *estimation* techniques are needed, which in the field of tomographic reconstruction are solved iteratively, leading to good approximate solution only through some form of regularization.

## 3.2 Inversion methods

### 3.2.1 Analytical inverse formula

In this subsection we describe the analytic 2D image reconstruction problem in PET<sup>2</sup>. To this purpose we introduce some notations and definitions which are used to find  $\mathbf{f}$  such that  $\mathbf{p} = \mathbf{R}\mathbf{f}$  (for the moment we forget about the additive deterministic noise).

**Definition 3.1.** Let  $f \in L^1(\mathbb{R}^n)$ ; the Fourier transform of  $f$  is defined as

$$\hat{f}(\xi) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-ix\xi} dx.$$

**Definition 3.2.** A Schwartz space  $\mathcal{S}(\mathbb{R}^n)$  is defined as

$$\mathcal{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n) : \sup_{x \in \mathbb{R}^n} |x^k D^l f| < +\infty, \forall k \in \mathbb{Z}_+^n, \forall l \in \mathbb{Z}_+^n\},$$

where  $k$  and  $l$  are multindex in  $\mathbb{Z}_+^n$ .

---

<sup>2</sup>This problem has been widely discussed and described over the last decades by the mathematical community, so we are not going into details. For a complete discussion please refer to [84].

**Remark 3.1.** If  $f \in \mathcal{S}(\mathbb{R}^n)$  it may be shown that also  $\hat{f} \in \mathcal{S}$ . Furthermore, defining  $\tilde{f}(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{ix\xi} d\xi$ , we have  $\hat{\tilde{f}} = \hat{\hat{f}} = f$  (i.e. the Fourier transform is bijective).

**Definition 3.3.** Let  $C^n := \{(s, \phi) : s \in \mathbb{R}, \phi \in \mathbf{S}^{n-1}\}$  be the cylinder in  $\mathbb{R}^n$ , where  $\mathbf{S}^{n-1}$  is the sphere in  $\mathbb{R}^{n-1}$ ; define the Schwartz space on the cylinder as

$$\mathcal{S}(C^n) := \{f \in \mathcal{C}^\infty(C^n) : \sup_{\phi \in \mathbf{S}^{n-1}, s \in \mathbb{R}} |(1+|s|^k) \frac{\partial^l}{\partial s^l} (Df)(s, \phi)| < +\infty, \forall k \in \mathbb{Z}_+ \forall l \in \mathbb{Z}_+\}$$

### The Radon Transform

**Definition 3.4.** The Radon transform is a bounded integral operator defined as

$$\begin{aligned} \mathcal{R} : \mathcal{S}(\mathbb{R}^n) &\rightarrow \mathcal{S}(C^n) \\ f &\mapsto \int_{x \cdot \phi = s} f(x) dx = \int_{\phi^\perp} f(s\phi + y) dy. \end{aligned} \quad (3.2)$$

and it is the integral of  $f \in \mathcal{S}(\mathbb{R}^n)$  over the hyperplane perpendicular to  $\phi$  with (signed) distance  $s$  from the origin.

**Definition 3.5.** A projection of  $f$  at  $\bar{\phi} \in [0, \pi)$  is  $(\mathcal{R}_{\bar{\phi}} f)(s) = \mathcal{R}f(s, \bar{\phi})$ , where  $\mathcal{R}_\phi : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R})$ .

Our PET problem, in this framework, can be seen as the problem of finding  $f \in \mathcal{S}(\mathbb{R}^2)$  such that

$$p = \mathcal{R}f \quad (3.3)$$

where  $p \in \mathcal{S}(C^2)$ , and  $p$  it is known, since it is the set of the line integrals of  $f$ , i.e. the projection measured by PET. Our image reconstruction problem is thus to invert the Radon transform. Unfortunately, in this framework, the Radon transform defines a *compact* operator between (weighted)  $L^2$ -spaces, therefore<sup>3</sup> the inverse Radon transform is not bounded and naive solution procedures imply numerical instability. At some stage a regularization step is needed.

There exists a fundamental property that links the Fourier transform with the Radon transform of suitable functions:

<sup>3</sup>For a description of the inverse problem theory please refer to Appendix A



**Theorem 3.2** (Fourier Slice Theorem). *If  $f \in \mathcal{S}(\mathbb{R}^n)$ , then  $\mathcal{R}_\phi f(v) = (2\pi)^{\frac{n-1}{2}} \hat{f}(v\phi)$ .*

This theorem is essential in order to prove that there exists the inverse of the Radon transform and that it is given by the equation described by Theorems 3.3 and 3.5 (we do not show the proofs here but they can be found in [9]).

### Backprojection and Filtered Backprojection

An essential step in image reconstruction is backprojection, which is the adjoint to forward projection process that forms the projections of the object.

**Definition 3.6.** *The backprojection of  $\mathcal{R}$  is*

$$\begin{aligned} \mathcal{R}^\# : \mathcal{S}(\mathbb{C}^n) &\rightarrow \mathcal{S}(\mathbb{R}^n) \\ p &\mapsto \mathcal{R}^\# p, \end{aligned}$$

$$\text{where } (\mathcal{R}^\# p)(x) = \int_{\mathbf{S}^{n-1}} p(x \cdot \phi, \phi) d\phi.$$

**Theorem 3.3** (Inversion of the Radon transform). *If  $f \in \mathcal{S}(\mathbb{R}^n)$  and  $p = \mathcal{R}f$ , then*

$$f(x) = \frac{1}{2} (2\pi)^{1-n} (\mathcal{I}^\alpha \mathcal{R}^\# \mathcal{I}^{\alpha-n+1} p)(x), \quad (3.4)$$

where  $\mathcal{I}$  is the linear Riesz potential, defined as  $(\mathcal{I}^\alpha f)^\wedge(\xi) = |\xi|^{-\alpha} \hat{f}(\xi)$  for  $f \in \mathcal{S}(\mathbb{R}^n)$  and  $\alpha < n$ .

**Remark 3.4.** *We want to make some remarks.*

1. Since  $\alpha < n$ , then  $(\mathcal{I}^\alpha f)^\wedge \in L^1(\mathbb{R}^n)$ ; therefore we can perform the inverse Fourier transform and define  $\mathcal{I}^{-\alpha}$  such that  $\mathcal{I}^{-\alpha} \mathcal{I}^\alpha f = f$ ;
2. the formula for the inversion of the Radon transform is a one-parameter family of formulas, varying for  $\alpha < n$ ;
3. if  $\alpha = 0$ , we have  $f(x) = \frac{1}{2} (2\pi)^{1-n} \mathcal{R}^\# \mathcal{I}^{1-n} p(x)$  where  $\mathcal{I}^{1-n}$  acts on a function of  $\mathbb{R}^1$ . For  $g \in \mathcal{S}(\mathbb{R})$  it is useful to define the Hilbert transform as

$$(\mathcal{H}g)(x) := \frac{1}{\pi} \int_{\mathbb{R}} \frac{g(y)}{x - y} dy \quad (3.5)$$

and its Fourier transform as

$$(\mathcal{H}g)^\wedge(\xi) = -i\operatorname{sgn}(\xi)\hat{g}(\xi). \quad (3.6)$$

Thus we have

$$(\mathcal{I}^{1-n}g)^\wedge(\xi) = |\xi|^{n-1}\hat{g}(\xi) = (\operatorname{sgn}(\xi))^{n-1}\xi^{n-1}\hat{g}(\xi) \quad (3.7)$$

therefore

$$\mathcal{I}^{1-n}g = \mathcal{H}^{n-1}g^{(n-1)} \quad (3.8)$$

and hence

$$f(x) = \frac{1}{2}(2\pi)^{1-n}\mathcal{R}^\#\mathcal{H}^{n-1}p^{(n-1)} \quad (3.9)$$

where the  $(n-1)$ th derivative is taken with respect to the first argument. Using

$$\mathcal{H}^{n-1} = \begin{cases} (-1)^{(n-2)/2}\mathcal{H}, & n \text{ even} \\ (-1)^{(n-1)/2}, & n \text{ odd} \end{cases} \quad (3.10)$$

equation (3.9) can be rewritten as

$$f(x) = \frac{1}{2}(2\pi)^{1-n} \begin{cases} (-1)^{(n-2)/2} \int_{\mathbf{S}^{n-1}} \mathcal{H}p^{(n-1)}(x \cdot \phi, \phi) d\phi, & n \text{ even} \\ (-1)^{(n-1)/2} \int_{\mathbf{S}^{n-1}} p^{(n-1)}(x \cdot \phi, \phi) d\phi, & n \text{ odd.} \end{cases} \quad (3.11)$$

4. The fact that  $\mathcal{H}$  appears only in the even cases has an important consequence: since  $p(x \cdot \phi, \phi)$  is the integral of  $f$  over the hyperplane perpendicular to  $\phi$  that contains  $x$ , for  $n$  odd (3.11) can be evaluated if  $p(y \cdot \phi, \phi)$  is known for  $\phi \in \mathbf{S}^{n-1}$  and  $y \in I_x$ , i.e. if the integrals of  $f$  along all hyperplanes meeting a neighborhood of  $x$  are known. Thus the problem of reconstructing  $f$ , in the odd cases, is local, in the sense that the function is determined at some point by the integrals along the hyperplanes through a neighborhood of that point. That is not true if  $n$  is even, since  $\mathcal{H}$  is not local: thus the problem of reconstructing  $f$ , in the even cases, is not local, in the sense that computing the function at some point requires the integrals along all hyperplanes meeting the support of the function.

5. As previously stated, in the framework of interest (PET reconstruction problem) the Radon transform is a compact operator, i.e. the inverse is not bounded and therefore the formula (3.11) is not exploitable as it is; we need a regularization step.

**Theorem 3.5** (Filtered backprojection). *Let  $f \in \mathcal{S}(\mathbb{R}^n)$  and  $p = \mathcal{R}f$ ; then*

$$f(x) = \frac{1}{2}(2\pi)^{1-n}(\mathcal{I}^\alpha \mathcal{R}^\# \mathcal{I}^{\alpha-n+1}(k_c * p))(x).$$

where  $*$  denotes the convolution and  $k_c$  is a ramp filter that filters the projections before applying the backprojection operator.

This theorem guarantees that it is possible to filter the projections before the backprojection step and obtain again the objective function  $f$ , yielding a exploitable inversion formula for the Radon transform, that could be used in practice. Regularization is obtained by means of an optimal choice of the cut-off frequency  $c$  [9].

**Remark 3.6.** *Let us focus on the 2D case: in this setting the body to be imaged lies at the center of a line (the Line Of Response) connecting two detectors, and we consider just the direct planes (see Figure 3.3 left). We introduce a natural system of coordinates  $(x_1, x_2)$  with origin in the positron annihilation point; therefore a variable coordinate system  $\boldsymbol{\theta}, \boldsymbol{\theta}^\perp$  is given by the vector  $\boldsymbol{\theta} = (\cos(\phi), \sin(\phi))$  indicating the direction of the LOR and by  $\boldsymbol{\theta}^\perp = (-\sin(\phi), \cos(\phi))$ , orthogonal to  $\boldsymbol{\theta}$ . Given  $f$  defined on  $\mathbb{R}^2$ , let us consider the inverse Fourier Transform*

$$f(\mathbf{x}) = (2\pi)^{-1/2} \int \hat{f}(\boldsymbol{\xi}) e^{i\boldsymbol{\xi} \cdot \mathbf{x}} d\boldsymbol{\xi} \quad (3.12)$$

Since  $\boldsymbol{\xi} = \xi \boldsymbol{\theta}$ , we have that

$$f(\mathbf{x}) = (2\pi)^{-1/2} \int_0^{2\pi} d\phi \int_0^{+\infty} d\xi \hat{f}(\xi \boldsymbol{\theta}) e^{i\xi \boldsymbol{\theta} \cdot \mathbf{x}}. \quad (3.13)$$

This integral can be split into

$$f(\mathbf{x}) = (2\pi)^{-1/2} \left( \int_0^\pi \int_0^{+\infty} \hat{f}(\xi, \phi) e^{i\xi \boldsymbol{\theta} \cdot \mathbf{x}} \xi d\xi d\phi + \int_0^\pi \int_0^{+\infty} \hat{f}(\xi, \phi + \pi) e^{i\xi \boldsymbol{\theta} \cdot \mathbf{x}} \xi d\xi d\phi \right), \quad (3.14)$$

from  $\hat{f}(\xi, \phi + \pi) = \hat{f}(\xi, \phi)$  and from the Fourier Slice Theorem, it follows that

$$f(\mathbf{x}) = (2\pi)^{-1/2} \int_0^\pi \int_{-\infty}^{+\infty} (\mathcal{R}_\theta f)(\xi) |\xi| e^{i\xi\theta \cdot \mathbf{x}} d\xi d\phi. \quad (3.15)$$

Equation (3.15) is the filtered back projection formula for the 2 dimensional case. It states that if the Radon Transform of a function is available, the original function can be exactly restored by filtering the Radon Transform by means of a filter, whose frequency response in  $|\xi|$  and then backprojecting it by means of the backprojection operator (that in 2D is  $(\mathcal{R}^\# p_\theta)(\mathbf{x}) = \int_0^\pi p_\theta(\mathbf{x} \cdot \boldsymbol{\theta}) d\phi$ ).

By summarizing, let us consider the following algorithm for the inversion of the Radon transform by means of the filtered back projection in 2D:

- given the measured projections  $p_\theta(s)$ , compute  $\hat{p}_\theta(\xi)$ ;
- multiply times the ramp filter  $|\xi|$ ;
- compute the inverse Fourier Transform;
- apply  $\mathcal{R}^\#$ .

We observe that the algorithm is very fast, since it only requires the computation of two Fast Fourier Transform (FFT) plus the application of  $\mathcal{R}^\#$ .

### Three-Dimensional Reprojection Algorithm (3DRP)

All the previous statements are valid if we have a two-dimensional imaging scanner; in fully three-dimensional imaging there are essentially two differences with the two-dimensional case: spatial dependence of the scanner and redundancy of the data. In this modality, the activity in the center of the axial FOV is more visible to the scanner than to the activity at the edge, causing spatial variance, which notably complicates the use of analytical techniques. Luckily, fully 3D data contain redundancies, since, from an analytical viewpoint, just one single slice of data is needed to reconstruct an image, and this property can be exploited to make FBP also available in fully three-dimensional acquisition. For details please refer to [49].

### 3.2.2 Iterative techniques

Improvements over the analytical approach can be achieved by iterative statistical techniques, that can account for the noise structure in the data and thus are able to describe the problem's model with more reliability. On the other hand, these improvements come at the cost of an increased complexity in mathematical formulation and a loss of speed in reconstruction algorithm. With respect to the exact inversion formulas that are easily obtained in the analytical case, if a statistical technique is applied to the PET reconstruction problem, we have to give up the idea of “solution” and must settle for “estimation”.

#### Maximum Likelihood Expectation Maximization

Maximum Likelihood (ML) is a standard statistical approach for estimating the object  $\mathbf{f}$  from data  $\mathbf{p}$  (refer to equation (3.1)), which is based on the assumption that given data  $\mathbf{p}$  are observed values of the random process  $\mathbf{P}$  with mean  $\mathbf{Rf}$ <sup>4</sup>. Maximum Likelihood has the advantage that ML estimates can be effectively constrained to be non negative, which are exactly the constraints we have in a image-reconstruction problem. In the case of Poisson noise, essentially the case in which data are corrupted by “counting” noise, as in the PET case, in which data are collected by measuring photons in finite time interval (as “counting” them) the constrained ML approach leads to the well known the Expectation Maximization (EM) algorithm [18]. Recall that we need to find  $\mathbf{f}$ , from the knowledge of  $\mathbf{p}$ , corrupted by Poisson noise, such that  $\mathbf{p} = \mathbf{Rf}$ , and subject to  $\mathbf{f} \geq 0$ : by applying the MLEM algorithm to this specific problem, a simple iterative equation is obtained:

**Definition 3.7** (MLEM iterative algorithm).

$$\mathbf{f}^{(n+1)} = \frac{\mathbf{f}^{(n)}}{\mathbf{R}^T \mathbf{1}} \mathbf{R}^T \frac{\mathbf{p}}{\mathbf{Rf}^{(n)}} \quad (3.16)$$

where  $n$  represents the iteration step in the estimation procedure, and  $\mathbf{f}^{(n)}$  is the estimate of  $\mathbf{f}$  at step  $n$ .

Let us consider therefore the following algorithm:

- start with an initial guess  $\mathbf{f}^{(0)}$ ;
- forwardly project  $\mathbf{f}^{(0)}$  into the projection domain;

---

<sup>4</sup>For detail on this statistical technique see Appendix B.

- compare these projections with  $\mathbf{p}$ ;
- thanks to that comparison, form a multiplicative correction factor for each projection;
- backproject this factor into the image domain;
- get a correction factor for the initial guess estimate;
- this correction factor is thus multiplied times the image estimate, divided by a weighting term
- a new image estimate is then reentered in the algorithm as the next image

The algorithm goes on, approaching the maximum likelihood solution. At each guess step, the MLEM algorithm allows the low frequency components of the image to appear first, and, as the ML estimate is approached, higher frequency definition is resolved in the image. The convergence rate of MLEM is image dependent and the algorithm usually requires 20-50 iterations to reach an acceptable solution. Obviously, since MLEM requires one forward projection and one backprojection at each step, the overall processing time is greater than the corresponding FBP time, but the image reconstruction is potentially more accurate.

### Ordered Subset Expectation Maximization

With the purpose of reducing the reconstruction time of MLEM, Ordered Subsets Expectation Maximization (OSEM) was introduced in 1994 [35]. OSEM uses subsets of the entire data set for each image, updated accordingly to the following definition:

**Definition 3.8** (OSEM iterative algorithm).

$$\mathbf{f}^{(n+1)} = \frac{\mathbf{f}^{(n)}}{(\mathbf{R}_{|S_a})^T \mathbf{1}} (\mathbf{R}_{|S_a})^T \frac{\mathbf{p}}{\mathbf{R} \mathbf{f}^{(n)}} \quad (3.17)$$

The backprojection step sum only over a subset  $S_a$  of a total of  $A$  projection subset, thus reducing notably the reconstruction time that is required by MLEM. When there is only one subset, OSEM is obviously the same of MLEM. There are many techniques to decide how to divide the projection space into subsets; many of them use non-overlapping subsets, by dividing the projection into sets with different view.

**Remark 3.7.** *OSEM is a computationally convenient method that resembles MLEM but, even if so, there is no guarantee of convergence to the ML solution. In practice, however, its convergence is similar to MLEM and is  $A$  times faster. This speed increase comes at the expense of slightly more image variance at the same bias level when compared with MLEM [57].*

### Three-Dimensional Iterative Reconstruction

In principle, it is possible to extend iterative methods to fully three-dimensional PET acquisition and reconstruction: in this case the data model  $\mathbf{p}$  based on 3D measurements, the image model  $\mathbf{f}$  is a three-dimensional volume and the system model  $\mathbf{R}$  relates voxel elements to fully three-dimensional projections. The statistical relationship apply as well to 3D PET measurements, leading to the same optimization algorithm. In clinical practice, the challenge for fully 3D iterative reconstruction is the increased computational request. Fully 3D reconstructions increase the image sizes from  $10^4$  pixels to  $10^5$  voxels, while the data set sizes pass from  $10^4$  to as much as  $10^7$  entries. As a consequence, the system model is computed for  $10^{12}$  combinations, with respect to  $10^8$  combinations with 2D PET.





## Part II: from images to physiology

Now that we are able to reconstruct dynamic PET images from sinograms, the following chapters describe the *second* inverse problem we deal with: how to pass from dynamic PET images to specific descriptors of tracer (and glucose) physiology? Dynamic PET data can be processed in order to estimate physiological parameters that describe the functional behavior of the inspected tissues and the flow of tracer between their different constituents: following a model originated in the framework of pharmacokinetics, it is assumed that there are functionally separated pools of tracer, exchanging tracer between each other, referred to as *compartments*. Therefore the flow of tracer is described by a compartmental system.

Chapter 4 focuses on the evaluation of the tracer density of each compartment and the exchange rates between compartments.

Chapter 5 discusses the so called *compartmental inverse problem*, namely, that of determining the unknown rate coefficients from PET images.



# Chapter 4

## Compartmental models in a formal perspective

This chapter is devoted to the description of compartmental models, to be regarded as tools which provide a quantitative description of a family of biological phenomena, occurring within a given organ, e.g., tracer (glucose) flows. Indeed, the framework of compartmental analysis identifies different compartments in the physiological system of interest, each one characterized by a specific and homogeneous functional role. With the help of a global observation of the organ along time, PET-scan images, one aims at getting information on the radioactive tracer exchange rates between compartments. From a mathematical viewpoint, the time dependent concentrations of tracer in each compartment constitute the state variables that can be determined from PET data. The time evolution of the state variables is modeled by a system of differential equations for the concentrations, expressing the principle of tracer balance during exchange processes. As a first approximation or when it is known that the exchange rates do not depend on time or the concentrations, the concentration evolution in the system is governed by a linear system of ordinary differential equations with constant coefficients. The constant coefficients describe the input/output rate of tracer for each compartment and represent the physiological parameters assessing the system's metabolism; hence they are the unknown to be estimated as solutions to an inverse problem.

In the first section we give a very introductory and pictorial idea of *what we talk about when we talk about* [14] pharmacokinetics and compartmental models, reviewing classical compartmental systems, in the mathematical framework. It has not been an easy work, because from a biological and physiological viewpoint, these

methods are widely used and “roughly” understood, but a formal mathematical characterization, involving also study of well-posedness, was almost faraway. The second section is devoted to a more formal description of compartmental models.

## 4.1 Generalities on compartmental analysis

Pharmacokinetics is the study of absorption, distribution and elimination of drugs, and the study of pharmacokinetics has essentially three objectives: prediction, description and prescription. In all of its application, pharmacokinetics utilizes the concept of *compartment*: thus the analysis of tracer kinetics in tissues whose activities have been measured with FDG-PET is typically based on compartmental models [30, 67]. This conceptual framework identifies different compartments in the physiological system of interest, each one characterized by a specific (and homogeneous) functional role (therefore a compartment is not necessarily represented by a specific organ or an anatomical district).

As any other tracer, FDG is injected into the system with a concentration mathematically modeled by the so-called Input Function (IF), which is the FDG concentration in the bloodstream. After injection, FDG, carried by the blood, perfuses the organs and tissues of the body. Every tissue or organ can be resolved into different compartments, as in Figure 4.1, correlated to a tracer-specific-functional behavior (for example cells containing phosphorylated FDG and interstitial fluid containing free FDG). The time dependent concentrations of FDG in each compartment constitute the state variables that can be determined from PET data. The time evolution of the state variables (the kinetics of the system) is modeled by a linear system of ordinary differential equations (ODEs) for the concentrations, expressing the conservation of tracer during flow between compartments. The (constant) micro-parameters describing the input/output rate of tracer for each compartment are called exchange coefficients or rate constants. They represent the physiological parameters describing the system’s metabolism and are the unknowns to be estimated. In this settings it is assumed that no diffusive effects can take place during the FDG perfusion, and tracer is just carried by blood. Furthermore FDG is assumed to be uniformly distributed in each compartment at each instant [97].

For the mathematical reduction of such a compartmental system, there exist essentially two approaches, in literature. An analytical approach according to which a formal analytic solution of the system of ODEs describing the model is evaluated, followed by application of an optimization scheme in order to find the exchange parameters by comparison with image PET data [109, 23, 117, 30, 46, 48]. A graphical

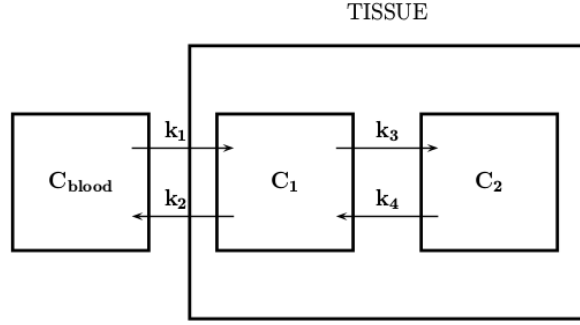


Figure 4.1: Example of 2-compartment compartmental system with 1 blood Input Function ( $C_{blood}$ ) and 4 exchange coefficients, typical for FDG metabolism of phosphorylation-dephosphorylation. From a physiological viewpoint,  $C_1$  represent the FDG that is freely in the interstitial space,  $C_2$  the FDG that is trapped inside cells;  $k_1$  and  $k_2$  the exchanges between blood and the compartmental system;  $k_3$  represent the phosphorylation process, and  $k_4$  the inverse de-phosphorylation process.

approach consisting in the manipulation of the system of ODEs in order to find an easier equation that in some sense can be fitted with a line on plane, allowing the estimation of a combination of the exchange coefficients [67, 89].

#### 4.1.1 Analytical method

The state variable of the 2-compartmental model described in Figure 4.1 are the tracer concentration in  $C_1$  and in  $C_2$ ; the kinetic process in the system is initialized by the IF  $C_{blood}$ , representing the tracer concentration in blood. For sake of simplicity in the following we will denote  $C_{blood}$  with the most common  $C_b$ . The four constant coefficients between compartments in contact are denoted as  $k_1, k_2, k_3, k_4$  and we assume that  $k_i \geq 0$  for all  $i = 1, \dots, 4$ . Conservation of tracer exchanged between compartments leads to the following system of linear ordinary differential equations with constant coefficients:

$$\begin{cases} \dot{C}_1(t) = -(k_2 + k_3)C_1(t) + k_4C_2(t) + k_1C_b(t) \\ \dot{C}_2(t) = k_3C_1(t) - k_4C_2(t) \end{cases} \quad (4.1)$$

with initial condition  $C_1(0) = C_2(0) = 0$ . Giving such equations, compartmental analysis requires the determination of the tracer coefficients  $k_1, \dots, k_4$  by utilizing

measurements of the tracer concentration provided by nuclear imaging and applying an optimization scheme for the solution of the inverse problem. In nuclear imaging experiments, the reconstructed images can provide information on the tracer concentration in the tissue under examination (typically an organ of interest) as well as in the input arterial blood (that provides the Input Function  $C_b$ ) as measured in the left ventricle. Specifically, an acquisition sequence is set up providing count data sets collected at subsequent time intervals. For each data set, an image reconstruction algorithm is applied, Regions of Interest (ROIs) are drawn within the left ventricle and the organ of interest, and the corresponding tracer concentrations are computed. Obviously, the tracer concentration in such an organ is an estimate of  $C_1 + C_2$  plus the tracer carried by the blood contained in the organ's vascular system, leading to an estimation of a kind of weighted sum of concentrations:

$$\tilde{C}_{exp} = (1 - V)(C_1 + C_2) + VC_b, \quad (4.2)$$

where  $V$  is such a weight, representing the blood fraction with which the organ's vascular system is supplied. The weight can either be estimated by means of the optimization procedure (in addition to the tracer coefficients), or taken as known from physiological consideration on the organ under examination. The optimization scheme that is typically implemented [2], minimizes the functional

$$\mathcal{C} = \|(1 - V)(C_1 + C_2) + VC_b - \tilde{C}_{exp}\|, \quad (4.3)$$

where  $C_1$  and  $C_2$  here are the analytical solutions of the direct problem (4.1), computed at each time point, and  $\|\cdot\|$  indicates an appropriate norm (typically the euclidean distance or a weighted  $L^2$ -norm).

### 4.1.2 Graphical methods

Graphical techniques provide simple methods for the analysis of data from tracer studies. They ensure considerable ease of computation, if compared to the optimization of individual model parameters occurring in the solution of the differential equations generally used to describe the binding of tracers. The theoretical work of Patlak [89] which was applied to irreversible tracers (i.e.  $k_4 = 0$ ) formed the basis for extensions of graphical techniques to reversibly ( $k_4 \neq 0$ ) binding tracers.

Graphical analysis converts the model equations into one equation evaluated at the time points corresponding to the scanning times and provides fewer parameters, namely a *slope* and *intercept*: the slope can be interpreted in terms of a combination of model parameters for some model structure. Graphical methods also require

an Input Function. The method for irreversible tracers was developed first - the theoretical foundation was provided by Patlak [89], the extension to reversible systems developed by Logan [68] was based on the original work of Patlak. Further refinements have been made by Ichise [40]. The main problems with the use of graphical methods are the bias in the estimated parameters due to noise, and the fact that these methods cannot describe accurately the metabolism of organs with complex physiology, such as kidneys or liver, because of the few parameters that can be estimated.

Mathematically, it is possible to rearrange Equation (4.1) in order to write another (easier) equation, depending just on two parameters: if the tracer injected, and the system is not reversible we have

$$\frac{\tilde{C}_{exp}}{C_b} = K_p \frac{\int_0^t C_b}{C_b} + V \quad (4.4)$$

while if the system is reversible, we have:

$$\frac{\int_0^t \tilde{C}_{exp}}{C_b} = K_l \frac{\int_0^t C_b(\tau) d\tau}{\tilde{C}_{exp}} + V \quad (4.5)$$

where  $K_p = \frac{k_1}{k_2}$  and  $K_l = \frac{k_1 k_3}{k_2 + k_3}$  and, coupled with  $V$ , are the only parameters that graphical methods can estimate [89, 68, 40].

## 4.2 The $n$ -compartment system

We want to consider now a general  $n$ -dimensional system, composed of  $n$  compartments as in Figure 4.2. Remark that although the generic term *compartment* is used, it does not necessary mean that each compartment is contained in a physical compartment distinguishable from the others. In fact, this even constitutes one of the main issues in the inverse problem of getting information on the system, since each compartment can not be individually observed. A radioactive tracer is injected to a patient and for a compartment  $p \in [1, n]$ ,  $C_p$  denotes the non-negative concentration function of the tracer in the compartment. The compartment  $p$  receives the radioactive tracer from the outside world at a constant non-negative rate  $k_{pe}$  and a non-negative concentration function  $C_{pe}$  and it excretes the tracer at a constant non-negative rate  $k_{ep}$  in the outside world. The constant non-negative rate at which the compartment  $p$  receives the tracer from a compartment  $q \neq p$  is denoted  $k_{pq}$ . The concentration functions  $(C_{pe})_{p \in [1, n]}$  are supposed to be continuous. The

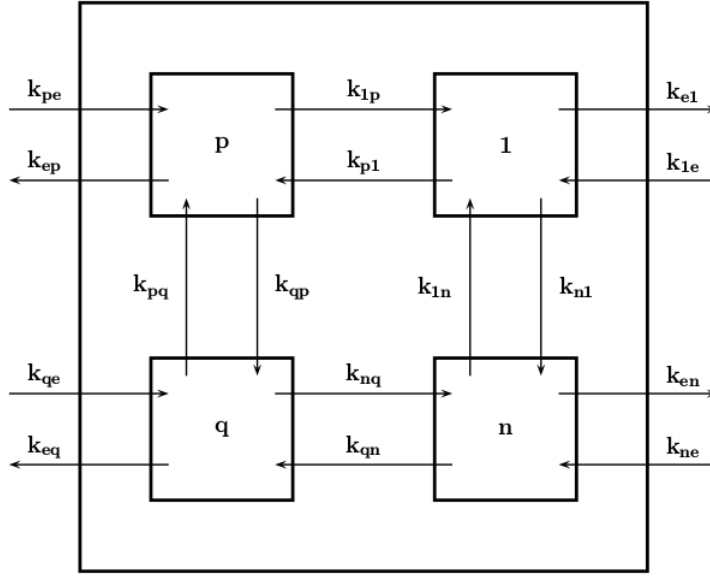


Figure 4.2:  $n$ -compartment system, just some of the allowed flows are drawn, for sake of legibility.

evolution of the tracer concentrations in each compartment is then governed by the following linear system of ordinary differential equations with constant coefficients

$$\dot{C}_p = \sum_{q=1}^n k_{pq} C_q + k_{pe} C_{pe}, \quad p \in [1, n], \quad (4.6)$$

with the initial conditions

$$C_p(0) = 0, \quad p \in [1, n], \quad (4.7)$$

where, for  $p \in [1, n]$ ,  $k_{pp} = -\left(\sum_{q \neq p} k_{qp} + k_{ep}\right)$  and  $k_{ep}$  are termed excretion rates. Synthetically, we have

$$\dot{C} = MC + W, \quad C(0) = 0, \quad (4.8)$$

where

$$C = \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix}, \quad W = \begin{pmatrix} k_{1e} C_{1e} \\ \vdots \\ k_{ne} C_{ne} \end{pmatrix}, \quad (4.9)$$

and the matrix  $M$  is given by

$$M_{pq} = k_{pq}, \quad p, q \in [1, n]. \quad (4.10)$$



PET-scan images provide the total amount of radioactive tracer in the organ. Each compartment contributes to the intensity of the image linearly with respect to the amount of tracer in the compartment, that is its volume times the tracer concentration. Hence, in practice, PET-scan data gives access to

$$\tilde{C}(t) = \alpha \cdot C(t), \quad t \in \mathbb{R}_+, \quad (4.11)$$

where  $\alpha \in \mathbb{R}_+^{*n}$  is a known constant vector. Thus, the inverse problem we consider is to recover the exchange rates  $\underline{k} \in \mathbb{R}^{n^2+n}$ , where

$$\underline{k} = \begin{cases} k_{ep}, & p \in [1, n^2], p \equiv 1 \pmod{n+1}, \\ k_{p-n \lfloor \frac{p-1}{n} \rfloor, 1 + \lfloor \frac{p-1}{n} \rfloor}, & p \in [1, n^2], p \not\equiv 1 \pmod{n+1}, \\ k_{(p-n^2)e}, & p \in [n^2+1, n^2+n], \end{cases} \quad (4.12)$$

with measures of  $\tilde{C}$ . Note however that some exchange rates may be *a priori* known. In the following of the document, for a positive integer  $n$  and  $k \in \mathbb{R}^{n^2+n}$ , we denote by  $\hat{k} \in \mathbb{R}^{n^2}$  the first  $n^2$  components of  $k$  and  $\check{k} \in \mathbb{R}^n$  the last  $n$  components of  $k$ . For a positive integer  $n$ , we denote by  $\mathfrak{M}$  the following linear operator

$$\begin{aligned} \mathfrak{M} : \mathbb{R}^{n^2} &\rightarrow M_n(\mathbb{R}) \\ H &\mapsto \mathfrak{M}(H), \end{aligned} \quad (4.13)$$

where for all  $H \in \mathbb{R}^{n^2}$

$$\mathfrak{M}(H)_{pq} = \begin{cases} -H_{1+(n+1)(p-1)} - \sum_{\substack{p'=1 \\ p' \neq p}}^n H_{p+n(p'-1)}, & p, q \in [1, n], p = q, \\ H_{p+n(q-1)}, & p, q \in [1, n], p \neq q, \end{cases} \quad (4.14)$$

so that for all  $H \in \mathbb{R}_+^{n^2}$ ,  $\mathfrak{M}(H)$  is the matrix defined in (4.10) for the parameters  $H \in \mathbb{R}^{n^2}$ .

### 4.2.1 Properties of $n$ -compartment systems

In this subsection, we recall some properties of  $n$ -compartment systems which will be useful for our study of two particular cases. First of all, for general  $n$ -compartment systems, we have the following theorem [33]

**Theorem 4.1.** *Consider  $H \in \mathbb{R}_+^{n^2}$ . Then the eigenvalues of the matrix  $M = \mathfrak{M}(H)$  as defined in (4.13) have a non-positive real part and if an eigenvalue has a zero*

real part, then the eigenvalue is 0, moreover,  $\dim(\ker(M)) = m_0$  where  $m_0$  is the multiplicity of 0. In addition, the solution  $C$  to

$$\dot{C} = MC + W, \quad C(0) = C_0, \quad (4.15)$$

where  $C_0 \in \mathbb{R}_+^n$  and  $W : \mathbb{R}_+ \rightarrow \mathbb{R}_+^n$ , verifies  $C_p(t) \geq 0$  for all  $p \in [1, n]$  and  $t \in \mathbb{R}_+$ .

**Remark 4.2.** As remarked in [33], physically, it is conservation of mass applied to the system (4.15) where  $W = 0$ , which insures that the eigenvalues of  $M$  have a non-positive real part and that the multiplicity of the eigenvalue 0 is the dimension of the nullspace of  $M$  (the solutions are bounded) and that the only possible eigenvalue with a zero real part is 0 (if any oscillations occur, they are damped). The positiveness of  $C$  is simply the fact that the concentrations are positive quantities.

Some additional properties on the system matrix  $M$  lead to more restrictions on its eigenvalues. In particular, we have [33]

**Theorem 4.3.** Consider  $H \in \mathbb{R}_+^{n^2}$  and denote by  $M = \mathfrak{M}(H)$ . If  $M$  is irreducible, then 0 is an eigenvalue of  $M$  if and only if  $k_{ep} = 0$  for  $p \in [1, n]$ , where for  $p \in [1, n]$ ,  $k_{ep} = H_{1+(n+1)(p-1)}$  (no excretion). In other terms, 0 is an eigenvalue of  $M$  if and only if the  $n$ -compartment system of exchange rates  $H \in \mathbb{R}_+^{n^2}$  and without input is closed. Moreover, in that case, 0 is a simple eigenvalue.

**Theorem 4.4.** Consider  $H \in \mathbb{R}_+^{n^2}$ , denote by  $M = \mathfrak{M}(H)$  and suppose that there exist  $(a_p)_{p \in [1, n]}$  in  $\mathbb{R}_+^{*n}$  such that for all  $p, q \in [1, n], p \neq q$

$$k_{pq}a_q = k_{qp}a_p, \quad (4.16)$$

where for all  $p, q \in [1, n], p \neq q$ ,  $k_{pq} = H_{p+n(q-1)}$ , then  $M$  is diagonalizable and its eigenvalues are real and non-positive. Note that the condition (4.16) does not depend on the excretion rates  $k_{ep} = H_{1+(n+1)(p-1)}$  for  $p \in [1, n]$ .

**Remark 4.5.** Again, as remarked in [33], the equation (4.16) is the principle of de-tailed balance for the  $n$ -compartment closed system of exchange rates  $(k_{pq})_{p, q \in [1, n], p \neq q}$ .

Theorem 4.4 applies in particular to catenary system we will use in the following. We end this subsection recalling a result on the multiplicity of the eigenvalues of the matrix of a  $n$ -compartment system [88]

**Theorem 4.6.** *Consider a sign-symmetric cycle-free connected  $n$ -compartment system. Then an eigenvalue  $\lambda$  of the system matrix  $M$  is multiple if and only if there exists a compartment  $p$ , directly connected to at least three other compartments  $p_1, p_2, p_3$ , such that  $\lambda$  is an eigenvalue of the matrices  $N_1, N_2, N_3$  of the respective connected subsystems containing  $p_1, p_2, p_3$  and not  $p$ .*

### 4.2.2 The $n$ -compartment catenary case

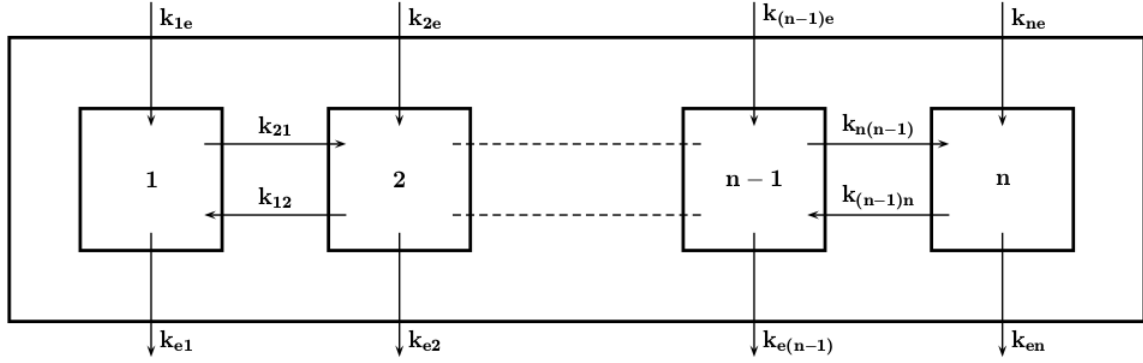


Figure 4.3:  $n$ -compartment catenary system

A  $n$ -compartment catenary system is composed of  $n$  compartments such that

$$\begin{aligned} k_{ep} &\geq 0, & p &\in [1, n], \\ k_{pq} &> 0, & p, q &\in [1, n], |p - q| = 1, \\ k_{pq} &= 0, & p, q &\in [1, n], |p - q| > 1, \\ k_{pe} &\geq 0, & p &\in [1, n]. \end{aligned}$$

According to the theorems 4.1, 4.3, 4.4, 4.6 of the previous subsection 4.2.1, we have the following theorem

**Theorem 4.7.** *The matrix  $M$  of a  $n$ -compartment catenary system is diagonalizable and its eigenvalues are real, non-positive and simple. Moreover, 0 is an eigenvalue of  $M$  if and only if the system with no input is closed, that is  $k_{ep} = 0$  for all  $p \in [1, n]$ .*



# Chapter 5

## Compartmental inverse problem

In this chapter we give a formal description of the compartmental inverse problem. We regarded the constant coefficients which describe the input/output rate of tracer for each compartment and form the entries of the system matrix as the unknown to be estimated. A linear combination of the concentrations, i.e. of the solutions of the linear system, is regarded as given by PET data. Hence we have to solve an inverse problem. One of the most important questions in an inverse problem is: how much information on the system can one get with such a global observation? In this chapter, after recalling some definitions and results, we partially answer the previous question in a general and brief introductory study and then apply the results in two particular cases of common models: catenary system with 2 and 3 compartments.

### 5.1 Identifiability

We are now concerned with the problem of whether the rate coefficients can be determined uniquely from a given set of data.

#### 5.1.1 Definitions

We denote by  $\mathfrak{W}$  the following linear operator

$$\begin{aligned} \mathfrak{W} : C^0(\mathbb{R}_+, \mathbb{R})^n &\rightarrow L(\mathbb{R}^n, C^0(\mathbb{R}_+, \mathbb{R})^n) \\ \check{C} &\mapsto \mathfrak{W}(\check{C}), \end{aligned} \tag{5.1}$$

where for all  $\check{C} \in C^0(\mathbb{R}_+, \mathbb{R})^n$ ,  $H \in \mathbb{R}^n$  and  $p \in [1, n]$ ,  $[\mathfrak{W}(\check{C})(H)]_p = H_p \check{C}_p$ , so that for all vector of input concentrations functions  $\check{C} = (C_{pe})_{p \in [1, n]} \in C^0(\mathbb{R}_+, \mathbb{R}_+)^n$

and  $H \in \mathbb{R}_+^n$ ,  $\mathfrak{W}(\check{C})(H)$  is the vector defined in (4.9) for the input concentrations functions  $\check{C}$  and the parameters  $H$ . We denote by  $\mathfrak{C}$  the following function

$$\begin{aligned} \mathfrak{C} : C^0(\mathbb{R}_+, \mathbb{R})^n &\rightarrow L(\mathbb{R}^{n^2+n}, C^1(\mathbb{R}_+, \mathbb{R})^n) \\ \check{C} &\mapsto \mathfrak{C}(\check{C}), \end{aligned} \quad (5.2)$$

where for all  $\check{C} \in C^0(\mathbb{R}_+, \mathbb{R})^n$  and  $k \in \mathbb{R}^{n^2+n}$ ,  $C = \mathfrak{C}(\check{C})(k) \in C^1(\mathbb{R}_+, \mathbb{R})^n$  is the unique solution to

$$\dot{C} = MC + W, \quad C(0) = 0, \quad (5.3)$$

where  $M = \mathfrak{M}(\hat{k})$  and  $W = \mathfrak{W}(\check{C})(\hat{k})$ . For  $\alpha \in \mathbb{R}_+^{*n}$ ,  $\mathfrak{C}^\alpha$  is the function defined by

$$\begin{aligned} \mathfrak{C}^\alpha : C^0(\mathbb{R}_+, \mathbb{R})^n &\rightarrow L(\mathbb{R}^{n^2+n}, C^1(\mathbb{R}_+, \mathbb{R})) \\ \check{C} &\mapsto [k \mapsto \alpha^T(\mathfrak{C}(\check{C})(k))]. \end{aligned} \quad (5.4)$$

We now recall the definitions of identifiability [23, 75]. Consider a positive integer  $n$ , input concentration functions  $\check{C} \in C^0(\mathbb{R}_+, \mathbb{R}_+)^n$ ,  $\alpha \in \mathbb{R}_+^{*n}$  and a subset  $\Omega$  of  $\mathbb{R}_+^{n^2+n}$  of admissible exchange rates  $k$ . Denote by  $\tilde{C} = \mathfrak{C}^\alpha(\check{C})$  and by  $\tilde{C}_\Omega$  the restriction of  $\tilde{C}$  to  $\Omega$ .

**Definition 5.1.**

1. The model of equations (4.8)-(4.11) is said *globally identifiable* at  $k \in \Omega$ , if  $\tilde{C}_\Omega^{-1}(\{\tilde{C}(k)\}) = \{k\}$ .
2. The model of equation (4.8)-(4.11) is said *locally identifiable* at  $k \in \Omega$  if there exists  $\epsilon > 0$  such that  $\tilde{C}_{\Omega, k, \epsilon}^{-1}(\{\tilde{C}(k)\}) = \{k\}$  where  $\tilde{C}_{\Omega, k, \epsilon}$  denotes the restriction of  $\tilde{C}_\Omega$  to the open ball  $B_{k, \epsilon}$  of  $\Omega$ , with center  $k$  and radius  $\epsilon$ .
3. The model of equation (4.8)-(4.11) is said *structurally globally identifiable* if it is globally identifiable at all  $k \in \Omega$ .
4. The model of equation (4.8)-(4.11) is said *structurally locally identifiable* if it is locally identifiable at all  $k \in \Omega$ .

### 5.1.2 Results on a $n$ -compartment system

It is hard to find a characterization for identifiability of a general  $n$ -compartmental system, but we can still find some weak results in particular cases, leading to more

precise identifiability properties.

Consider  $n$ -compartment systems where for  $p \in [1, n]$ , the input concentration functions  $C_{pe}$  from  $\mathbb{R}_+$  to  $\mathbb{R}_+$  are Laplace-transformable. For  $p \in [1, n]$ ,  $r_{pe} = \inf \left\{ r \in \mathbb{R} : \int_0^{+\infty} e^{-rt} C_{pe}(t) dt < +\infty \right\}$  will denote the abscissa of convergence of  $C_{pe}$ . Note that in practical applications, the input concentration functions are bounded so that they are Laplace-transformable and for  $p \in [1, n]$  we have  $r_{pe} \leq 0$ . The Laplace transform will be denoted by  $\mathcal{L}$  and for  $r \in \mathbb{R}$ , we denote by  $\mathbb{C}_r = \{z \in \mathbb{C} : \Re(z) > r\}$ .

**Lemma 5.1.** *Consider a positive integer  $n$ , Laplace-transformable input concentration functions  $\check{C} = (C_{pe})_{p \in [1, n]} \in C^0(\mathbb{R}_+, \mathbb{R}_+)^n$ ,  $\alpha \in \mathbb{R}_+^{*n}$  and exchange rates  $k \in \mathbb{R}_+^{n^2+n}$ . Then, for  $\alpha \in \mathbb{R}_+^{*n}$ ,  $\tilde{C} = \mathfrak{C}^\alpha(\check{C})(k)$  is Laplace-transformable and its abscissa of convergence  $r$  verifies*

$$r \leq r_m = \max(\{r_m^e\} \cup \text{sp}(M)), \quad (5.5)$$

where

$$r_m^e = \max(\{r_{pe} : p \in [1, n]\}), \quad (5.6)$$

and  $\text{sp}(M)$  denotes the spectrum of the matrix  $M = \mathfrak{M}(\hat{k})$ . Moreover for all  $z \in \mathbb{C}_{r_m}$ , we have

$$\mathcal{L}\tilde{C}(z) = \sum_{p=1}^n \alpha_p k_{pe} \frac{Q_p(z)}{P(z)} \mathcal{L}C_{pe}(z), \quad (5.7)$$

where the unitary polynomial  $P$  of degree  $n$  is the characteristic polynomial of  $M$  and for  $p \in [1, n]$ ,  $Q_p$  is a unitary polynomial of degree  $n-1$  given by

$$Q_p(z) = \frac{\alpha^T \text{adj}(zI_n - M) e_p}{\alpha_p}, \quad (5.8)$$

where  $\text{adj}(zI_n - M)$  denotes the adjugate matrix of  $zI_n - M$ ,  $I_n$  is the  $n$ -dimensional identity matrix and  $e_p$  is the  $p$ -th versor in  $\mathbb{R}^n$ .

*Proof.* The first statement of the lemma is obvious. In addition, since  $C = \mathfrak{C}(\check{C})(k)$  verifies (4.8), for all  $z \in \mathbb{C}_{r_m}$ , we have

$$\mathcal{L}\tilde{C}(z) = \sum_{p=1}^n k_{pe} \alpha^T (zI_n - M)^{-1} e_p \mathcal{L}C_{pe}(z). \quad (5.9)$$

Since

$$(zI_n - M)^{-1} = \frac{\text{adj}(zI_n - M)}{\det(zI_n - M)} = \frac{\text{adj}(zI_n - M)}{P(z)}, \quad (5.10)$$

where  $P$  is the characteristic polynomial of  $M$ , we have

$$\mathcal{L}\tilde{C}(z) = \sum_{p=1}^n k_{pe} \frac{\alpha^T \text{adj}(zI_n - M)e_p}{P(z)} \mathcal{L}C_{pe}(z). \quad (5.11)$$

Moreover, it can be easily remarked that the only cofactors of  $zI_n - M$  of degree  $n - 1$  in  $z$  are the diagonal cofactors and that the coefficient of the monomial  $z^{n-1}$  in these cofactors is 1, so that for all  $p \in [1, n]$ , the polynomial  $\alpha^T \text{adj}(zI_n - M)e_p$  in  $z$  is of degree  $n - 1$  and its leading coefficient is  $\alpha_p$ .  $\square$

For a positive integer  $n$ , we denote by  $\mathfrak{P}$  the function

$$\begin{aligned} \mathfrak{P} : \mathbb{R}^{n^2} &\rightarrow \mathbb{R}[X] \\ H &\mapsto \det(XI_n - \mathfrak{M}(H)), \end{aligned} \quad (5.12)$$

and for  $\alpha \in \mathbb{R}_+^{*n}$ , we denote by  $\mathfrak{Q}^\alpha$  the function

$$\begin{aligned} \mathfrak{Q}^\alpha : \mathbb{R}^{n^2} &\rightarrow \mathbb{R}[X]^n \\ H &\mapsto \left( \frac{\alpha^T \text{adj}(XI_n - \mathfrak{M}(H))e_p}{\alpha_p} \right)_{p \in [1, n]}. \end{aligned} \quad (5.13)$$

For  $r \in \mathbb{R}$ , we denote by  $\mathcal{M}(\mathbb{C}_r)$  the field of meromorphic functions on  $\mathbb{C}_r$  and by  $\mathbb{R}(X)$  the field of rational fractions in the indeterminate  $X$  with coefficients in  $\mathbb{R}$ . In the following, for  $r \in \mathbb{R}$ ,  $\mathcal{M}(\mathbb{C}_r)$  will be regarded as a vector space over  $\mathbb{R}(X)$ . As an immediate consequence of lemma 5.1, we have

**Corollary 5.1.** *Consider a positive integer  $n$  and Laplace-transformable input concentration functions  $\check{C} = (C_{pe})_{p \in [1, n]} \in C^0(\mathbb{R}_+, \mathbb{R}_+)^n$  such that  $C_{pe} = 0$  for  $p \in [1, n] \setminus \{p_l : l \in [1, k]\}$ , where  $k \geq 1$  and  $(p_l)_{l \in [1, k]}$  are  $k$  distinct elements of  $[1, n]$ . We recall that for exchange rates  $k \in \mathbb{R}_+^{n^2+n}$ ,  $\check{k}$  denotes the last  $n$  components of  $k$ , that is, for  $p \in [1, n]$ ,  $\check{k}_p = k_{pe}$ . Consider  $\alpha \in \mathbb{R}_+^{*n}$  and denote by  $\check{C} = \mathfrak{C}^\alpha(\check{C})$ . If the functions  $(\mathcal{L}C_{pe})_{l \in [1, k]}$  of  $\mathcal{M}(\mathbb{C}_{r_m^e})$  are linearly independent over  $\mathbb{R}(X)$ , where  $r_m^e$  is defined by (5.6), then for  $k \in \mathbb{R}_+^{n^2+n}$ , the exchange rates  $(\check{k}_{pe})_{l \in [1, k]}$  are uniquely determined by the function  $\check{C}(k)$ , that is, for all  $k' \in \mathbb{R}_+^{n^2+n}$ , we have  $\check{C}(k') = \check{C}(k)$*



only if  $\check{k}_{p_{le}} = \check{k}'_{p_{le}}$  for all  $l \in [1, k]$ . More precisely, we have  $\tilde{C}(k') = \tilde{C}(k)$  if and only if the previous condition holds and for all  $l \in [1, k]$

$$\frac{\mathfrak{Q}^\alpha(k')_{p_l}}{\mathfrak{P}(k')} = \frac{\mathfrak{Q}^\alpha(k)_{p_l}}{\mathfrak{P}(k)}. \quad (5.14)$$

Otherwise, in a more general case, we have

**Corollary 5.2.** *Consider a positive integer  $n$  and Laplace-transformable input concentration functions  $\check{C} = (C_{pe})_{p \in [1, n]} \in C^0(\mathbb{R}_+, \mathbb{R}_+)^n$ , then, there exist  $k$  distinct elements  $(p_l)_{l \in [1, k]}$  of  $[1, n]$ , where  $k \in [0, n]$  ( $k = 0$  if all input concentration functions are identically zero) such that the functions  $(\mathcal{L}C_{p_{le}})_{l \in [1, k]}$  of  $\mathcal{M}(\mathbb{C}_{r_m^e})$  are linearly independent over  $\mathbb{R}(X)$ , where  $r_m^e$  is defined by (5.6), and such that for all  $p \in [1, n] \setminus \{p_l : l \in [1, k]\}$ ,  $\mathcal{L}C_{pe} \in \text{span}(\{\mathcal{L}C_{p_{le}} : l \in [1, k]\})$ . Consider  $\alpha \in \mathbb{R}_+^{*n}$  and denote by  $\tilde{C} = \mathfrak{C}^\alpha(\check{C})$ . For all  $k \in \mathbb{R}^{n^2+n}$  and  $z \in \mathbb{C}_{r_m}$ , where  $r_m$  is defined by (5.5), we then have*

$$\mathcal{L}\tilde{C}(z) = \sum_{l=1}^k F_l^{\check{C}, k, \alpha}(z) \mathcal{L}C_{p_{le}}(z), \quad (5.15)$$

where for all  $l \in [1, k]$ ,  $F_l^{\check{C}, k, \alpha} \in \mathbb{R}(X)$ . Moreover, for  $k, k' \in \mathbb{R}^{n^2+n}$ , we have  $\tilde{C}(k) = \tilde{C}(k')$  if and only if for all  $l \in [1, k]$

$$F_l^{\check{C}, k', \alpha} = F_l^{\check{C}, k, \alpha}. \quad (5.16)$$

### 5.1.3 Study of a 2-compartment catenary system

We introduce now the case of a 2-compartment catenary system. The low degree of complexity of this 2-compartmental model, allows nuclear medical doctors to use it in post-processing the images, without high computational request. For this reason, this model is very popular in literature, handling with FDG-PET scans in particular, in organs in which the FDG metabolism can be simplified in just two chemical phases: a free intracellular phase and a phosphorylated one. We recall that the evolution equations in a 2-compartment catenary system as in Figure 5.1 are given by

$$\begin{aligned} \dot{C}_1 &= -(k_{e1} + k_{21})C_1 + k_{12}C_2, \\ \dot{C}_2 &= k_{21}C_1 - k_{12}C_2, \end{aligned}$$

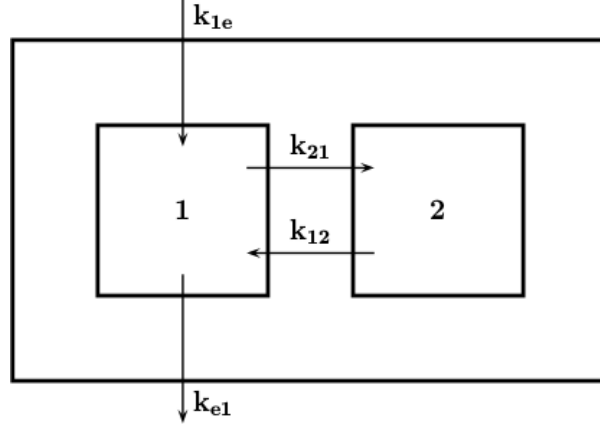


Figure 5.1: 2-compartment catenary model

with initial conditions

$$C_1(0) = 0, \quad C_2(0) = 0.$$

That is

$$\dot{C} = MC + W, \quad C(0) = 0, \quad (5.17)$$

where  $M \in M_2(\mathbb{R})$ ,  $C \in C^1(\mathbb{R}_+, \mathbb{R})^2$ ,  $W \in C^0(\mathbb{R}_+, \mathbb{R})^2$  are given by

$$M = \begin{pmatrix} -(k_{e1} + k_{21}) & k_{12} \\ k_{21} & -k_{12} \end{pmatrix}, \quad C = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}, \quad W = k_{1e} \begin{pmatrix} C_{1e} \\ 0 \end{pmatrix} = k_{1e} C_{1e} e_1.$$

PET-scan images allow to access  $C^* = (1 - V)(C_1 + C_2) + VC_{1e}$  where  $V$  is a known fraction in  $(0, 1)$  and represent the blood fraction in the tissue under examination. Since  $C_{1e}$  is the concentration of tracer in blood, it is directly measurable by means of ROIs drawn over the left ventricle. So it is possible to rewrite previous equation as  $\tilde{C} = (1 - V)(C_1 + C_2)$ , where  $\tilde{C} = C^* - VC_{1e}$  which is accessible from PET-scans. The inverse problem consists in recovering  $k_{1e}$ ,  $k_{e1}$ ,  $k_{21}$ ,  $k_{12}$  from the knowledge of  $V$ ,  $C_{1e}$  and

$$\tilde{C} = \alpha \cdot C, \quad \text{where } \alpha = \begin{pmatrix} 1 - V \\ 1 - V \end{pmatrix}. \quad (5.18)$$

More generally, we will study the identifiability of this 2-compartment catenary system on  $\Omega = \mathbb{R}_+^4$ . For sake of simplicity, in the following of the section, the exchange rates of the system are denoted by  $a, b, c, k$  where

$$a = k_{e1}, \quad b = k_{21}, \quad c = k_{12}, \quad k = k_{1e}.$$

$M$  and  $W$  are then rewritten

$$M = \begin{pmatrix} -(a+b) & c \\ b & -c \end{pmatrix}, \quad W = kC_{1e}e_1.$$

We recall that according to Theorem 4.7,  $M$  is diagonalizable and its eigenvalues are simple and negative. We suppose that  $C_{1e}$  is bounded and not identically zero. Hence, according to Corollary 5.1, we first have the following lemma

**Lemma 5.2.**  *$k$  is uniquely determined by the knowledge of  $\tilde{C}$ .*

Now, in order to know the solutions of the inverse problem, we have to search for the matrices  $M_x \in M_2(\mathbb{R})$

$$M_x = \begin{pmatrix} -(x_a + x_b) & x_c \\ x_b & -x_c \end{pmatrix},$$

where  $x_a, x_b, x_c \in \mathbb{R}_+^*$ , such that  $F_x = F$ , with  $F = Q/P$  and  $F_x = Q_x/P_x$ , where  $P$  is the characteristic polynomial of  $M$ ,  $P_x$  is the characteristic polynomial of  $M_x$  and  $Q, Q_x$  are defined by

$$Q(X) = \alpha^T \text{adj}(X - M)e_1, \quad Q_x(X) = \alpha^T \text{adj}(X - M_x)e_1.$$

That is

$$\begin{aligned} P(X) &= X^2 + (a + b + c)X + ac, \\ Q(X) &= (1 - V)(X + b + c), \\ P_x(X) &= X^2 + (x_a + x_b + x_c)X + x_ax_c, \\ Q_x(X) &= (1 - V)(X + x_b + x_c). \end{aligned}$$

First of all, write the rational fraction by splitting the polynomial  $P$

$$F(X) = \frac{(1 - V)(X + b + c)}{(X - \lambda_1)(X - \lambda_2)},$$

where  $\lambda_1 = -\frac{a+b+c+\sqrt{(a+b+c)^2-4ac}}{2}$  and  $\lambda_2 = -\frac{a+b+c-\sqrt{(a+b+c)^2-4ac}}{2}$ . Remark that in this simple example, we can easily verify the statement Theorem 4.7. Indeed, we have  $0 < (a + b - c)^2 + 4bc = (a + b + c)^2 - 4ac < a + b + c$  so that  $\lambda_1$  and  $\lambda_2$  are real, negative and distinct. Moreover, we have  $\lambda_1, \lambda_2 \neq -(b + c)$ , hence, the rational fraction  $F$  is in its irreducible form. Since  $\deg P_x = \deg P = 3$  and  $\deg Q_x = \deg Q = 2$ ,  $F_x = F$  only if the rational fraction  $F_x$  is irreducible too (but

as we have just seen, since  $x_a, x_b, x_c > 0$ , it is always irreducible). In addition, since the leading coefficients of  $P_x$  and  $P$  are identical, as well as those of  $Q_x$  and  $Q$ ,  $F_x = F$  if and only if  $P_x = P$  and  $Q_x = Q$ , that is

$$\begin{aligned}x_a + x_b + x_c &= a + b + c, \\x_a x_c &= c, \\x_b + x_c &= b + c.\end{aligned}$$

Hence  $F_x = F$  if and only if  $x_a = a$ ,  $x_b = b$ ,  $x_c = c$ , that is  $M_x = M$ . Consequently, we have the following theorem

**Theorem 5.1.** *The model of equations (5.17)-(5.18) is structurally globally identifiable.*

#### 5.1.4 Study of a 3-compartment catenary system

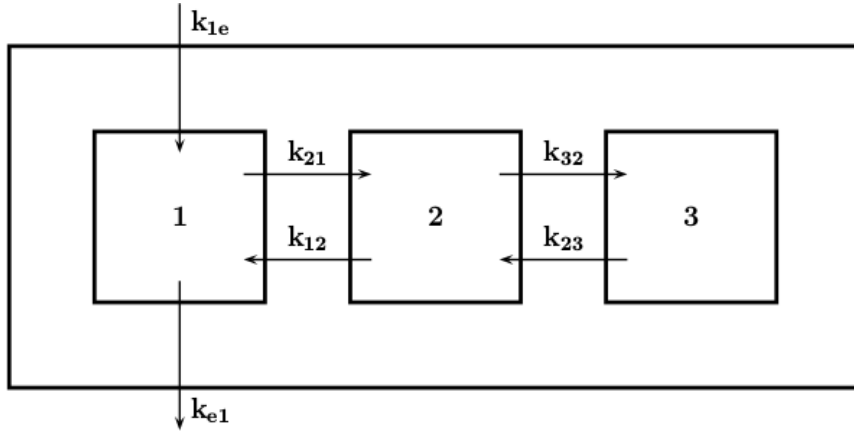


Figure 5.2: 3-compartment catenary model

We now introduce a 3-compartment catenary system, which is the classical evolution of the previous 2-compartmental model. In this model, the third compartment usually accounts for the presence of a blood stream that carries tracer inside the organ under examination. In 2-compartment models, this is approximated by using directly the tracer concentration in input  $C_{1e}$ . We recall that the evolution equations

in a 3-compartment catenary system as in Figure 5.2 are given by

$$\begin{aligned}\dot{C}_1 &= -(k_{e1} + k_{21}) C_1 + k_{12} C_2 + k_{1e} C_{1e}, \\ \dot{C}_2 &= k_{21} C_1 - (k_{12} + k_{32}) C_2 + k_{23} C_3, \\ \dot{C}_3 &= k_{32} C_2 - k_{23} C_3,\end{aligned}$$

with initial conditions

$$C_1(0) = 0, \quad C_2(0) = 0, \quad C_3(0) = 0.$$

That is

$$\dot{C} = MC + W, \quad C(0) = 0, \quad (5.19)$$

where  $M \in M_3(\mathbb{R})$ ,  $C \in C^1(\mathbb{R}_+, \mathbb{R})^3$ ,  $W \in C^0(\mathbb{R}_+, \mathbb{R})^3$  are given by

$$\begin{aligned}M &= \begin{pmatrix} -(k_{e1} + k_{21}) & k_{12} & 0 \\ k_{21} & -(k_{12} + k_{32}) & k_{23} \\ 0 & k_{32} & -k_{23} \end{pmatrix}, \\ C &= \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix}, \quad W = k_{1e} \begin{pmatrix} C_{1e} \\ 0 \\ 0 \end{pmatrix} = k_{1e} C_{1e} e_1.\end{aligned}$$

PET-scan images allow to access  $\tilde{C} = VC_1 + (1 - V)(C_2 + C_3)$  where  $V$  is a known fraction in  $(0, 1)$ , representing the blood fraction in the tissue under examination. The inverse problem consists in recovering  $k_{1e}$ ,  $k_{e1}$ ,  $k_{21}$ ,  $k_{12}$ ,  $k_{32}$ ,  $k_{23}$  from the knowledge of  $V$ ,  $C_{1e}$  and

$$\tilde{C} = \alpha \cdot C, \quad \text{where } \alpha = \begin{pmatrix} V \\ 1 - V \\ 1 - V \end{pmatrix}. \quad (5.20)$$

Since it is common to use a third compartment in order to account for the presence of a blood stream in the organ, it makes sense to generalize  $\alpha$  of the 2-compartmental model in this way, in which the third compartment is effectively weighted by this blood fraction  $V$ . More generally, we will study the identifiability of this 3-compartment catenary system on  $\Omega = \mathbb{R}_+^*$ <sup>6</sup>. For sake of simplicity, in the following of the section, the exchange rates of the system are denoted by  $a, b, c, d, e, k$  where

$$a = k_{e1}, \quad b = k_{21}, \quad c = k_{12}, \quad d = k_{32}, \quad e = k_{23}, \quad k = k_{1e}.$$

$M$  and  $W$  are then rewritten

$$M = \begin{pmatrix} -(a+b) & c & 0 \\ b & -(c+d) & e \\ 0 & d & -e \end{pmatrix}, \quad W = kC_{1e}e_1.$$

We recall that according to Theorem 4.7,  $M$  is diagonalizable and its eigenvalues are simple and negative. We suppose that  $C_{1e}$  is bounded and not identically zero. Hence, according to Corollary 5.1, we first have the following lemma

**Lemma 5.3.**  *$k$  is uniquely determined by the knowledge of  $\tilde{C}$ .*

Using the properties of the  $n$ -compartment catenary system given in Theorem 4.7 and the Corollary 5.1 of the preliminary results, in the special case of one input concentration function, we can write the following theorem, whose proof (quite long) is found in Appendix C.

**Theorem 5.2.** *If  $V \geq 1/2$ , the model of equations (5.19)-(5.20) is structurally globally identifiable. Otherwise, the model is neither structurally globally nor structurally locally identifiable. However, the model is locally identifiable at points  $(a, b, c, d, e, k) \in \mathbb{R}_+^{*6}$  such that*

$$a \neq \frac{1-2V}{2V(1-V)} \left( (c+d+e)V + (1-V)b \pm \sqrt{\Delta_Q} \right).$$

*More precisely, in that case, the model is globally identifiable at points  $(a, b, c, d, e, k) \in \mathbb{R}_+^{*6}$  such that one of the following exclusive conditions of Lemma C.4 holds: (1), (2.1), (2.2.1), (2.2.2.1), (3.1.1), (3.1.2.1.1), (3.1.2.2), 3.2.1, (3.2.2.1), (3.2.2.2.1), (3.3.1.1), (3.3.2.1). Otherwise, the inverse problem has two distinct solutions. In the case where*

$$a = \frac{1-2V}{2V(1-V)} \left( (c+d+e)V + (1-V)b \pm \sqrt{\Delta_Q} \right),$$

*the model is neither globally nor locally identifiable. More precisely, the set of solutions of the inverse problem is a curve of  $\mathbb{R}^6$ .*

In next Section we describe the method for the solution of the compartmental inverse problem of recover (4.12) from PET measures of (4.11).

## 5.2 The Gauss-Newton Regularized method

We introduce here the more general case of a  $n$ -compartment system to describe the Newton algorithm that we have implemented to actually and computationally find the solution (or solutions, if uniqueness do not holds) of the compartmental inverse problem. It is a classical Gauss-Newton method to search roots of a function, in the multi-dimensional case, but with a regularization term that accounts for the not well-posedness of the problem. We notice that, either when it is known that no uniqueness holds or when no information on uniqueness is known, the robustness and accuracy of the algorithm might greatly rely on the choice of a good initial guess.

Recall the equations governing the  $n$ -compartment model:

$$\dot{C} = MC + W, \quad C(0) = 0,$$

where

$$C = \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix}, \quad W = \begin{pmatrix} k_{1e}C_{1e} \\ \vdots \\ k_{ne}C_{ne} \end{pmatrix},$$

and  $M$  is

$$M_{pq} = k_{pq}, \quad p, q \in [1, n].$$

PET-scan data gives access to

$$\tilde{C}(t) = \alpha \cdot C(t), \quad t \in \mathbb{R}_+,$$

so the inverse problem is to recover the exchange rates  $\underline{k} \in \mathbb{R}^{n^2+n}$ , where

$$\underline{k} = \begin{cases} k_{ep}, & p \in [1, n^2], p \equiv 1 \pmod{n+1}, \\ k_{p-n\lfloor \frac{p-1}{n} \rfloor, 1+\lfloor \frac{p-1}{n} \rfloor}, & p \in [1, n^2], p \not\equiv 1 \pmod{n+1}, \\ k_{(p-n^2)e}, & p \in [n^2+1, n^2+n], \end{cases} \quad (5.21)$$

with measures of  $\tilde{C}$ . Denote by  $\hat{k} \in \mathbb{R}^{n^2}$  the first  $n^2$  components of  $k$  and  $\check{k} \in \mathbb{R}^n$  the last  $n$  components of  $k$ . For a positive integer  $n$ , recall the definition of  $\mathfrak{M}$ :

$$\begin{aligned} \mathfrak{M} : \mathbb{R}^{n^2} &\rightarrow M_n(\mathbb{R}) \\ H &\mapsto \mathfrak{M}(H), \end{aligned}$$

where for all  $H \in \mathbb{R}^{n^2}$

$$\mathfrak{M}(H)_{pq} = \begin{cases} -H_{1+(n+1)(p-1)} - \sum_{\substack{p'=1 \\ p' \neq p}}^n H_{p+n(p'-1)}, & p, q \in [1, n], p = q, \\ H_{p+n(q-1)}, & p, q \in [1, n], p \neq q, \end{cases}$$

Consider now the linear operator  $\text{Vec} : M_n(\mathbb{R}) \rightarrow \mathbb{R}^{n^2}$  stacking the columns of a matrix  $A$  into a column vector  $\text{Vec}(A)$ , that is

$$\text{Vec}(A)_p = A_{p-n\lfloor \frac{p-1}{n} \rfloor, 1+\lfloor \frac{p-1}{n} \rfloor}, \quad p \in [1, n^2]. \quad (5.22)$$

The operator  $\text{Vec} \circ \mathfrak{M}$  is linear from  $\mathbb{R}^{n^2}$  into itself, we denote by  $S \in M_{n^2}(\mathbb{R})$  its matrix. It is a  $n^2 \times n^2$  sparse matrix with  $2n^2 - n$  non-zero elements which are equal to  $\pm 1$ . More precisely,  $n^2 - n$  entries of  $S$  are equal to 1 and  $n^2$  entries of  $S$  are equal to  $-1$ .

$$S_{pq} = \begin{cases} 1, & p, q \in [1, n^2], p = q, p \not\equiv 1 \pmod{n+1}, \\ -1, & p, q \in [1, n^2], p \equiv 1 \pmod{n+1}, p \equiv q \pmod{n}, \\ 0, & \text{otherwise.} \end{cases} \quad (5.23)$$

For instance, in the case where  $n = 2$ , we have

$$S = \begin{pmatrix} -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix},$$

and in the case where  $n = 3$ , we have

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

We denote by  $\mathfrak{W}$  the following linear operator

$$\begin{aligned} \mathfrak{W} : C^0(\mathbb{R}_+, \mathbb{R})^n &\rightarrow L(\mathbb{R}^n, C^0(\mathbb{R}_+, \mathbb{R})^n) \\ \check{C} &\mapsto \mathfrak{W}(\check{C}), \end{aligned}$$

where for all  $\check{C} \in C^0(\mathbb{R}_+, \mathbb{R})^n$ ,  $H \in \mathbb{R}^n$  and  $p \in [1, n]$ ,  $[\mathfrak{W}(\check{C})(H)]_p = H_p \check{C}_p$ , so that for all vector of input concentrations functions  $\check{C} = (C_{pe})_{p \in [1, n]} \in C^0(\mathbb{R}_+, \mathbb{R}_+)^n$



and  $H \in \mathbb{R}_+^n$ ,  $\mathfrak{W}(\check{C})(H)$  is the vector defined in (4.9) for the input concentrations functions  $\check{C}$  and the parameters  $H$ . We denote by  $\mathfrak{C}$  the following function

$$\begin{aligned} \mathfrak{C} : C^0(\mathbb{R}_+, \mathbb{R})^n &\rightarrow L(\mathbb{R}^{n^2+n}, C^1(\mathbb{R}_+, \mathbb{R})^n) \\ \check{C} &\mapsto \mathfrak{C}(\check{C}), \end{aligned}$$

where for all  $\check{C} \in C^0(\mathbb{R}_+, \mathbb{R})^n$  and  $k \in \mathbb{R}^{n^2+n}$ ,  $C = \mathfrak{C}(\check{C})(k) \in C^1(\mathbb{R}_+, \mathbb{R})^n$  is the unique solution to

$$\dot{C} = MC + W, \quad C(0) = 0,$$

where  $M = \mathfrak{M}(\hat{k})$  and  $W = \mathfrak{W}(\check{C})(\check{k})$ . For  $\alpha \in \mathbb{R}_+^{*n}$ ,  $\mathfrak{C}^\alpha$  is the function defined by

$$\begin{aligned} \mathfrak{C}^\alpha : C^0(\mathbb{R}_+, \mathbb{R})^n &\rightarrow L(\mathbb{R}^{n^2+n}, C^1(\mathbb{R}_+, \mathbb{R})^n) \\ \check{C} &\mapsto [k \mapsto \alpha^T(\mathfrak{C}(\check{C})(k))]. \end{aligned}$$

### Algorithm description

We consider a  $n$ -compartment system as in Figure 4.2 with known input concentration functions  $\check{C} = (C_{pe})_{p \in [1, n]} \in C^0(\mathbb{R}_+, \mathbb{R}_+)^n$  and  $\alpha \in \mathbb{R}_+^{*n}$ . We recall that the problem we are interested in consists in recovering the exchange rates  $k \in \Omega \subset \mathbb{R}^{n^2+n}$ , where  $k$  is given by (5.21), from the knowledge of  $\tilde{C}(k)$ , where  $\tilde{C} = \mathfrak{C}^\alpha(\check{C})$ . We recall that

$$\tilde{C}(k)(t) = \alpha \cdot C(k)(t), \quad t \in \mathbb{R}_+,$$

where  $C = \mathfrak{C}(\check{C})$ . The concentrations vector  $C(k)$  is the solution of the ordinary differential equations

$$\dot{C}(k) = \mathfrak{M}(\hat{k})C(k) + W(\check{k}), \quad C(k)(0) = 0,$$

where  $W = \mathfrak{W}(\check{C})$ . The solution  $C(k)$  to the system of ordinary differential equations is given by

$$C(k)(t) = \int_0^t \exp((t - \tau)\mathfrak{M}(\hat{k}))W(\check{k})(\tau) d\tau, \quad t \in \mathbb{R}_+,$$

hence

$$\tilde{C}(k)(t) = \alpha^T \int_0^t \exp((t - \tau)M(\hat{k}))W(\check{k})(\tau) d\tau, \quad t \in \mathbb{R}_+.$$

$\tilde{C} : \mathbb{R}^{n^2+n} \rightarrow C^1(\mathbb{R}_+, \mathbb{R})^n$  can be easily seen to be differentiable and even analytic. In order to use a Newton algorithm, we need to compute its Fréchet derivative. More precisely, considering  $t \in \mathbb{R}_+$ , we will compute the gradient of  $\tilde{C}_t$  for all  $t \in \mathbb{R}_+$ , with respect to  $k$ , where for all  $k \in \mathbb{R}^{n^2+n}$ ,  $\tilde{C}_t(k) = \tilde{C}(k)(t)$ . For all  $k \in \mathbb{R}^{n^2+n}$ , the

Fréchet derivative  $\frac{d\tilde{C}}{dk}(k)$  of  $\tilde{C}$  at  $k$ , bounded operator from  $\mathbb{R}^{n^2+n}$  to  $C^1(\mathbb{R}_+, \mathbb{R})$  is then given by

$$\begin{aligned} \frac{d\tilde{C}}{dk}(k) : \mathbb{R}^{n^2+n} &\rightarrow C^1(\mathbb{R}_+, \mathbb{R}), \\ H &\mapsto \left[ t \mapsto \nabla \tilde{C}_t(k) \cdot H \right]. \end{aligned}$$

For all  $t \in \mathbb{R}_+$ , the gradient of  $\tilde{C}_t$  is given by

$$\nabla \tilde{C}_t = \begin{pmatrix} \nabla_{\hat{k}} \tilde{C}_t \\ \nabla_{\check{k}} \tilde{C}_t \end{pmatrix}, \quad (5.24)$$

where  $\nabla_{\hat{k}}$  denotes the gradient with respect to  $\hat{k}$  and  $\nabla_{\check{k}}$  denotes the gradient with respect to  $\check{k}$ . Since  $\tilde{C}_t$  is linear with respect to  $\check{k}$ , we simply have for all  $k \in \mathbb{R}^{n^2+n}$

$$\nabla_{\check{k}} \tilde{C}_t(k) = \left( \int_0^t \check{C}(\tau) \odot \exp((t-\tau)\mathfrak{M}(\hat{k})^T) d\tau \right) \alpha, \quad (5.25)$$

where  $\odot$  denotes the Khatri-Rao product. Compute now the gradient of  $\tilde{C}_t$  with respect to  $\hat{k}$ . Writing  $\tilde{C}_t$  as

$$\tilde{C}_t(k) = \alpha^T \mathcal{F}_t(\check{k}, \mathfrak{M}(\hat{k})), \quad k \in \mathbb{R}^{n^2+n},$$

where for all  $k \in \mathbb{R}^{n^2+n}$  and  $N \in M_n(\mathbb{R})$

$$\mathcal{F}_t(\check{k}, N) = \int_0^t \exp((t-s)N) W(\check{k})(s) ds,$$

we have for all  $k \in \Omega$  and  $N, H \in \mathbb{R}^{n^2}$

$$\nabla_{\hat{k}} \tilde{C}_t(k) \cdot H = \alpha^T \frac{\partial \mathcal{F}_t}{\partial N} \left( \left( \check{k}, \mathfrak{M}(\hat{k}) \right); \frac{d\mathfrak{M}}{d\hat{k}}(\hat{k}; H) \right),$$

and since  $\mathfrak{M} : \mathbb{R}^{n^2} \rightarrow M_n(\mathbb{R})$  is linear

$$\nabla_{\hat{k}} \tilde{C}_t(k) \cdot H = \alpha^T \frac{\partial \mathcal{F}_t}{\partial N} \left( \left( \check{k}, \mathfrak{M}(\hat{k}) \right); \mathfrak{M}(H) \right). \quad (5.26)$$

Hence, to compute the gradient of  $\tilde{C}_t$  with respect to  $\hat{k}$ , we need to compute the Fréchet derivative of  $\mathcal{F}_t$  with respect to the second variable  $N \in M_n(\mathbb{R})$ . The Fréchet derivative of the exponential function can be written in this way

$$\frac{d \exp}{dN}(N; H) = \int_0^1 \exp(\tau N) H \exp((1-\tau)N) d\tau, \quad N, H \in M_n(\mathbb{R}).$$

Hence, for all  $k \in \Omega$  and  $N, H \in M_n(\mathbb{R})$

$$\frac{\partial \mathcal{F}_t}{\partial N}((\check{k}, N); H) = \int_0^t (t-s) \left( \int_0^1 \exp(\tau(t-s)N) H \exp((1-\tau)(t-s)N) d\tau \right) W(\check{k})(s) ds,$$

that is, with the change of variables  $\tau(t-s) \rightarrow \tau$

$$\frac{\partial \mathcal{F}_t}{\partial N}((\check{k}, N); H) = \int_0^t \left( \int_0^{t-s} \exp(\tau N) H \exp((t-s-\tau)N) d\tau \right) W(\check{k})(s) ds.$$

Hence

$$\frac{\partial \mathcal{F}_t}{\partial N}((\check{k}, N); H) = \int_0^t \left( \int_0^{t-\tau} \exp(\tau N) H \exp((t-s-\tau)N) W(\check{k})(s) ds \right) d\tau,$$

so that making the change of variables  $(t-\tau) \rightarrow \tau$ , we get

$$\frac{\partial \mathcal{F}_t}{\partial N}((\check{k}, N); H) = \int_0^t \exp((t-\tau)N) H \left( \int_0^\tau \exp((\tau-s)N) W(\check{k})(s) ds \right) d\tau.$$

In other words

$$\frac{\partial \mathcal{F}_t}{\partial N}((\check{k}, N); H) = \int_0^t \exp((t-\tau)N) H \mathcal{F}_\tau(N) d\tau.$$

Writing the previous formula in a more convenient way for computations, we have

$$\frac{\partial \mathcal{F}_t}{\partial N}((\check{k}, N); H) = \left( \int_0^t \mathcal{F}_\tau(N)^T \otimes \exp((t-\tau)N) d\tau \right) \text{Vec}(H),$$

where  $\otimes$  denotes the Kronecker product and the linear operator  $\text{Vec} : M_n(\mathbb{R}) \rightarrow \mathbb{R}^{n^2}$ , stacking the column of a matrix into a column vector, is defined in (5.22). Consequently, from (5.26), we have for all  $k \in \mathbb{R}^{n^2+n}$  and  $H \in \mathbb{R}^{n^2}$

$$\nabla_{\check{k}} \tilde{C}_t(k) \cdot H = \alpha^T \left( \int_0^t C(k)(\tau)^T \otimes \exp((t-\tau)\mathfrak{M}(\hat{k})) d\tau \right) \text{Vec}(\mathfrak{M}(H)).$$

Recalling that  $S$ , defined in (5.23), denotes the matrix of the linear operator  $\text{Vec} \circ \mathfrak{M} : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$ , we then have

$$\nabla_{\check{k}} \tilde{C}_t(k) \cdot H = \alpha^T \left( \int_0^t C(k)(\tau)^T \otimes \exp((t-\tau)\mathfrak{M}(\hat{k})) d\tau \right) SH,$$

so that

$$\nabla_{\hat{k}} \tilde{C}_t(k) = S^T \left( \int_0^t C(k)(\tau) \otimes \exp((t - \tau)\mathfrak{M}(\hat{k})^T) d\tau \right) \alpha. \quad (5.27)$$

Hence, using (5.24), (5.25), (5.27)

$$\nabla \tilde{C}_t(k) = \int_0^t \begin{pmatrix} S^T \left( C(k)(\tau) \otimes \exp((t - \tau)\mathfrak{M}(\hat{k})^T) \right) \alpha \\ \check{C}(\tau) \odot \exp((t - \tau)\mathfrak{M}(\hat{k})^T) \alpha \end{pmatrix} d\tau.$$

In other words, for all  $k, H \in \mathbb{R}^{n^2+n}$  and  $t \in \mathbb{R}_+$ , we have

$$\left[ \frac{d\tilde{C}}{dk}(k; H) \right] (t) = \left( \int_0^t \begin{pmatrix} S^T \left( C(k)(\tau) \otimes \exp((t - \tau)\mathfrak{M}(\hat{k})^T) \right) \alpha \\ \check{C}(\tau) \odot \exp((t - \tau)\mathfrak{M}(\hat{k})^T) \alpha \end{pmatrix} d\tau \right) H.$$

Consider a known function  $\tilde{C}_{exp} \in C^1(\mathbb{R}_+, \mathbb{R})$ , the measures of  $\tilde{C}(k^{\text{exact}})$  where  $k^{\text{exact}}$  are the real unknown exchange rates of the compartmental system. Let  $k^0 \in \Omega$  be an initial guess, then the Newton algorithm consists in solving the linear equation with unknown  $H^0 \in \Omega$

$$\left[ \frac{d\tilde{C}}{dk}(k^0; H^0) \right] (t) = \tilde{C}_{exp}(t) - \tilde{C}(k^0)(t), \quad \text{for all } t \in \mathbb{R}_+. \quad (5.28)$$

Then, increment the value  $k^0$ , giving  $k^1 = k^0 + H^0$  and iterate the process.

### Implementation details

The equation (5.28) may have no solution, moreover, in real applications, one has only the measured data  $\tilde{C}_{exp}$  for a finite number of sampling time points  $t_1, \dots, t_m \in \mathbb{R}_+$  and the data may be noisy. The discretized Newton algorithm consists in solving the linear system (5.28) by Tikhonov regularization. The non-regularized discretized system is given by

$$\begin{pmatrix} [\nabla \tilde{C}_{t_1}(k^0)]^T \\ \vdots \\ [\nabla \tilde{C}_{t_m}(k^0)]^T \end{pmatrix} H^0 = \begin{pmatrix} \tilde{C}_{exp}(t_1) - \tilde{C}_{t_1}(k^0) \\ \vdots \\ \tilde{C}_{exp}(t_m) - \tilde{C}_{t_m}(k^0) \end{pmatrix},$$

that is, denoting by  $A_0$  the matrix

$$A_0 = \begin{pmatrix} [\nabla \tilde{C}_{t_1}(k^0)]^T \\ \vdots \\ [\nabla \tilde{C}_{t_m}(k^0)]^T \end{pmatrix},$$

and by  $Y^0$  the vector

$$Y^0 = \begin{pmatrix} \tilde{C}_{exp}(t_1) - \tilde{C}_{t_1}(k^0) \\ \vdots \\ \tilde{C}_{exp}(t_m) - \tilde{C}_{t_m}(k^0) \end{pmatrix},$$

the non-regularized discretized system can be written as

$$A_0 H^0 = Y^0. \quad (5.29)$$

Solving the system (5.29) by Tikhonov regularization consists in finding the solution  $H^0$  to

$$(rI + A_0^T A_0) H^0 = A_0^T Y^0,$$

where  $r$  is a regularization parameter. As previously described, the value  $k^0$  is incremented, giving  $k^1 = k^0 + H^0$  and the process is iterated. Note that the regularization parameter can be different at each iteration step.



## Part III: from micro- to macro-physiology (and viceversa)

The next two chapters describe two important macro-compartmental models: Chapter 6 is concerned with flow of tracer (glucose) in the hepatic system, Chapter 7 with tracer circulation in the renal system. In both cases a somewhat global flow of tracer needs to be considered, in order to obtain realistic and useful results.

Specifically, we have to take into account the fact that the liver is perfused by blood (and tracer) coming from two independent sources: the hepatic artery and the portal vein. As usual, arterial concentration of tracer is obtained from measurements over the left ventricle, but determination of tracer concentration in the portal vein is a very difficult task. We take the view that it is reconstructed through measurements ultimately involving gut. As a consequence, the compartmental analysis cannot be restricted to the discussion of what happens inside some small subregion of liver. Indeed, as already observed, ROIs are taken over the entire liver and, in addition, over gut.

As to the kidneys, globality comes, e.g., from the need of considering the bladder as a significant component of the renal system, since it is deeply involved in the flow of tracer. In addition, we have to consider also volume changes of bladder (and the related effects on tracer concentration) resulting from continuous input of urine.

Under the aforementioned realistic conditions the basic compartmental scheme is supplemented by suitable modeling assumptions so as to obtain inverse problems that, on the one hand, can be effectively solved for the unknown exchange coefficients (microparameters), while on the other provide results that may be regarded as satisfactory from the physiological viewpoint. In this regard the analysis of microcirculation of tracer in the hepatic and renal systems is shown to provide consistent and important results for understanding the global behavior of the liver and the kidneys under normal and pathological conditions, and to predict drug-induced behavior in glucose metabolism in both liver and kidneys.

Uniqueness results can be obtained quite straightforwardly for the hepatic system, while it is much more difficult to provide analogous results for the renal system. The complex analysis of the hepatic flow, which in our framework involves 4 tissue compartments and 2 input functions, is divided into two different 2-compartment catenary systems (guts and liver to be described in succession), and therefore uniqueness issues can be analyzed separately. For each 2-compartment system (despite the presence, in one of them, of 2 different IF), the statement of Theorem 5.1 holds, thus providing the uniqueness of the solution of the associated inverse problem.

On the other hand, the urinary system, as we modeled it, can not be reduced to a standard compartmental model, for which uniqueness issues have been provided (see Chapter 5). The anomaly of the renal system is due to the presence of the volume-varying-bladder-compartment and of a kidney-embedded anatomical compartment that accounts for the water re-absorption in renal tubules and that gives rise to nonstandard variations in tracer concentration. Because of these unusual conditions the formulation of the inverse problem for the renal system does not comply with the acquirements of Theorem 5.1. Nevertheless, thanks to the suitable modeling assumptions and a powerful numerical technique, we are able to solve the renal inverse problem and estimate all the unknown exchange coefficients.



# Chapter 6

## Application to real data: the liver problem

Hepatic glucose metabolism is of extreme relevance in a number of clinical syndromes and, in particular, in diabetes [6, 5, 44, 11, 15]. Several authors evaluated the potential of PET with FDG in this setting, both in humans and animals [5, 42]. However, a standard method has not been developed so far because of the peculiarities of liver function and anatomy. From a physiological viewpoint, hepatocytes can dephosphorylate glucose-6-phosphatase (G6P) to deliver glucose back to the bloodstream and this feature also applies to FDG-6-phosphatase (FDG6P) preventing the accumulation kinetics typical of most tissues and nicely suited to compartmental analysis. Furthermore, liver is supplied of blood in two different ways: through the Hepatic Artery and the Portal Vein, leading to an unusual 2-Input Functions compartmental model.

### 6.1 Liver physiology

#### 6.1.1 Dual Input

From the anatomical viewpoint, the dual blood supply from Hepatic Artery (HA) and Portal Vein (PV) further hampers the definition of tracer delivery to hepatic tissue [82]. Actually, limiting the analysis to the only arterial Input Function (HA IF) inevitably causes systematic underestimation of tracer distribution [4, 47, 71]. Accordingly, several attempts have been made to estimate the dual-input IFs from

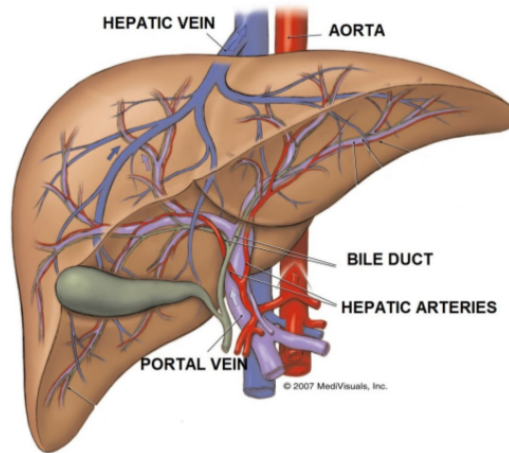


Figure 6.1: Pictorial representation of a human liver. It is possible to notice the Dual Input peculiarity of liver: blood is transported into the liver by both the Hepatic Artery and the Portal Vein.

dynamic PET data [47, 76, 78]. However, no method successfully solved this problem so far, since the portal vein, that arrives at liver through the gut, is hard to recognize in PET images and suffers from a relevant partial-volume effect related to vessel size and respiratory movements. According to these limitations, several authors tried to indirectly predict PV tracer concentration by the arterial IF. More precisely, one approach applied ad hoc algorithms to compute the PV IF as heuristic modification of the arterial one [43, 69]. A second approach directly and invasively sampled PV IF [43, 42, 44] and then determined the dual IF as a weighted superposition of the two activities [54, 82, 83, 81, 93, 114]. However, both methods are difficult to apply in a non invasive setting, while their potential under non-physiological conditions has been scarcely explored.

We have studied [28] a computation-based, model-dependent method for the use of FDG to study glucose kinetics in liver, in order to be able to determine the dual input typical of liver physiology just by means of computation, thus avoiding: 1) the use of invasive catheter-based measurement procedures and, 2) the major drawback of methods relying on the sole arterial IF, which are not able to account for the intrinsic variability of perfusion and glucose consumption in splanchnic organs. Other attempts to compute the PV IF have been performed, but those cases rely on a heuristic modification of the arterial IF which is obtained once for all and applied

to all animal models. On the contrary, our method is intrinsically model-dependent, since it utilizes *compartmental analysis of the gut* to directly estimate the PV IF, which, in its turn, is combined with the arterial IF to perform the compartmental analysis of the liver.

### 6.1.2 Pharmacological stimuli

As we previously stated (Section 1.2.2), the liver plays a crucial role in the maintenance of glycemic levels despite a large variability in glucose consumption in the different body tissues. To this purpose, liver metabolism is regulated as to counterbalance the hyperglycemic response to food intake and to adapt serum glucose concentration to whole body needs under fasting periods. This buffer function is of pivotal importance since once entered in the cytosol of almost all mammalian cells glucose is sequestered by its phosphorylation to G6P. The irreversible nature of this reaction also applies to FDG whose accumulation as FDG-6P accounts for the potential of this tracer in the study of brain and tumor metabolism [103].

We have studied [28] the changes of the hepatic FDG metabolism in mice treated with interventions able to blunt or enhance liver G6P de-phosphorylation and glucose delivery, in order to validate our algorithm by testing its performances in such a real experiment. In order to blunt liver G6P de-phosphorylation we treated mice with metformin (see 1.2.2), while to enhance such a de-phosphorylation we essentially “starved” mice for 48 hours preceding the microPET acquisition.

## 6.2 The liver compartmental problem

To overcome the Dual Input limitations and to be able to quantify the blunting/enhancing effects on hepatic FDG metabolism, we coupled the capability of Albira microPET to provide FDG Time Activity Curves (TACs) in virtually all organs with the formulation of a compartmental model for the hepatic system using the three descriptors of tracer kinetics in the liver:

1. time concentration curve in the arterial system (HA IF);
2. time concentration curve in the whole liver;
3. time concentration curve in the whole gut as a source of PV blood.

We used this knowledge as input datum of a regularized Gauss-Newton algorithm (see Section 5.2) that reduces a 2-compartment compartmental model in which gut tracer concentration is used to compute PV IF, while liver activity is used to describe tracer kinetics in the hepatic system.

We used a 2-compartment compartmental model to describe gut kinetics and another 2-compartment compartmental model for the liver kinetics (see Section 5.1.3): in both cases, thanks to the results on identifiability previously described, we are insured on the uniqueness of the solution of the inverse problems associated to both the compartmental systems. Figure 6.2 shows a representation of the compartmental model used for our problem.

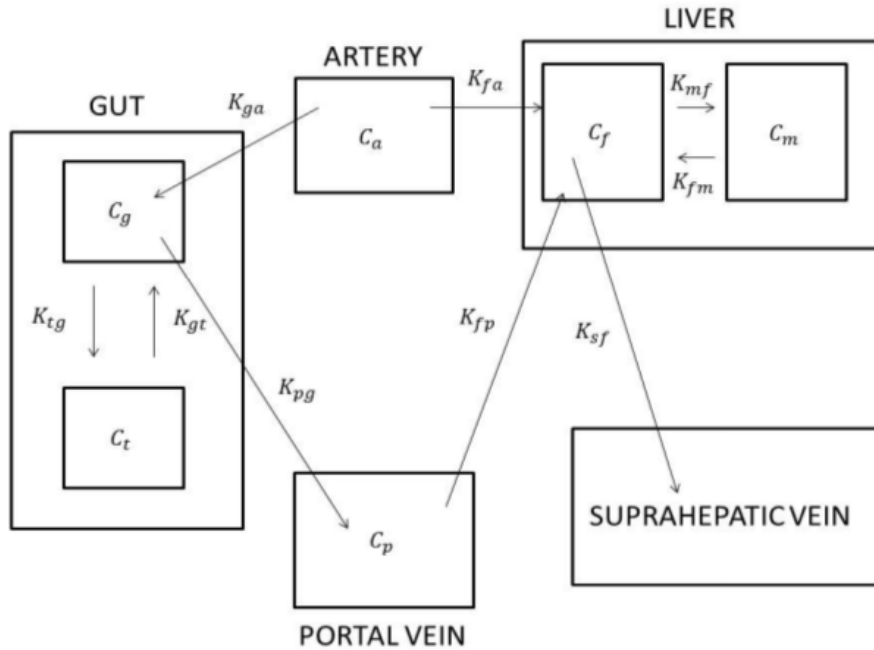


Figure 6.2: Compartmental model for gut and liver.

### 6.2.1 Model equation

#### Gut-PV subsystem

This subsystem was considered to have one arterial IF, one output function to PV and two compartments: the free tracer in gut (denoted with  $g$ ) and the trapped FDG-6P (denoted with  $t$ ). The mathematical model at the basis of our compartmental analysis relied on the balance of the tracer activities (i.e., concentrations per unit volume) between the different compartments. This balance leads to the classical system of ordinary differential equations. In the specific case of the gut tracer kinetics, the compartmental model is made of the following differential equations with vanishing initial data:

$$\dot{\mathbf{C}}_G = \mathbf{A}_G \mathbf{C}_G + \mathbf{e}_1 k_{ga} C_a \quad (6.1)$$

with

$$\mathbf{C}_G(0) = 0 \quad (6.2)$$

where

$$\mathbf{A}_G = \begin{pmatrix} -(k_{tg} + k_{pg}) & k_{gt} \\ k_{tg} & -k_{gt} \end{pmatrix} \quad (6.3)$$

and

$$\mathbf{C}_G = \begin{pmatrix} C_g \\ C_t \end{pmatrix}. \quad (6.4)$$

In these equations,  $C_g$ ,  $C_t$  and  $C_a$  represent the tracer concentrations in the free compartment, in the phosphorylated compartment and in arterial blood, respectively;  $k_{ij}$  denote the rate coefficients to target compartment  $i$  from source compartment  $j$ .

For the portal vein PV we have another compartment and, assuming that in such a compartment the blood flow is constant, the activity conservation leads to a third differential equation for  $C_p$ :

$$\dot{C}_p = k_{pg} C_g - k_{fp} C_p. \quad (6.5)$$

or, equivalently (since conservation of carrier fluid holds we have  $k_{fp} = k_{pg}$ )

$$\dot{C}_p = k_{pg} C_g - k_{pg} C_p. \quad (6.6)$$

#### Liver subsystem

Also liver can be described as a two-compartment model, one consisting of non-metabolized, free, tracer (compartment  $C_f$ ), and one consisting of phosphorylated,

metabolized tracer (compartment  $C_m$ ), where, as in the case of the gut, de-phosphorylation is explicitly allowed. Tracer inputs through HA and PV (compartments  $a$  and  $p$ , respectively) are modeled in an independent manner, i.e., there is no mixing of blood from the two vessels before entrance into liver. On the contrary, possible heterogeneities in perfusion source may occur as to consider the most general input model. Accordingly, exchange coefficients for HA blood ( $k_{fa}$ ) and PV blood ( $k_{fp}$ ) are considered as independent. On the other hand, FDG delivery to PV blood is considered to occur only from the gut, tracer exchange throughout the vein is negligible and thus the whole tracer amount entered into PV is available for liver uptake. We now denote with  $C_f$  and  $C_m$  the concentration of the free and metabolized FDG pools, respectively, with  $k_{sf}$  the rate coefficient from the free compartment to the venous efflux to the supra-hepatic vein  $s$ ,  $k_{mf}$  the exchange coefficient from the FDG to the FDG-6P pool, and  $k_{fm}$  the exchange coefficient for the inverse process. Then the usual assumption on the conservation of activities provides

$$\dot{\mathbf{C}}_L = \mathbf{A}_L \mathbf{C}_L + \mathbf{e}_1(k_{fa}C_a + k_{pg}C_p) \quad (6.7)$$

with

$$\mathbf{C}_L(0) = 0 \quad (6.8)$$

where

$$\mathbf{A}_L = \begin{pmatrix} -(k_{mf} + k_{sf}) & k_{fm} \\ k_{mf} & -k_{fm} \end{pmatrix} \quad (6.9)$$

and

$$\mathbf{C}_G = \begin{pmatrix} C_f \\ C_m \end{pmatrix}. \quad (6.10)$$

### 6.2.2 Regularized Gauss Newton method on data

MicroPET data provide information on the overall concentrations in ROIs drawn on the gut and liver throughout the whole acquisition. Therefore, denoting with  $\tilde{C}_G$  and  $\tilde{C}_L$  such experimental concentrations, we can write the following two equations for the microPET data:

$$\alpha \cdot \mathbf{C}_G = \tilde{C}_G \quad (6.11)$$

and

$$\alpha \cdot \mathbf{C}_L = \frac{\tilde{C}_L - V(\frac{1}{4}C_a + \frac{3}{4}C_p)}{1 - V} \quad (6.12)$$

where

$$\alpha = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad (6.13)$$

the numerical values  $\frac{1}{4}$  and  $\frac{3}{4}$  indicate the rate of arterial and venous contributions to the hepatic blood content  $V$  per unit volume [93, 101]. Further we can fix  $V$  and assume for it the physiologically sound value of 0.3 [47].

In both problems we have four unknowns to be estimated (coefficients  $k_{ij}$  of the compartmental system) and the same data equation. Our problems can be synthesized in a general form as

**Problem 6.1.** *We need to recover the vector  $\mathbf{k} := (k_1, k_2, k_3, k_4) \in \mathbb{R}_+^4$  from the knowledge of  $\alpha \cdot \mathbf{C}(\mathbf{k}) \in C^1([0, T], \mathbb{R})$  where  $\mathbf{C}(\mathbf{k})$  is the solution to*

$$\dot{\mathbf{C}}(\mathbf{k}) = \mathbf{M}(\tilde{\mathbf{k}})\mathbf{C}(\mathbf{k}) + k_1\mathbf{v} + \mathbf{u} \quad (6.14)$$

where  $T$  is the total acquisition time, with  $\mathbf{C}(\mathbf{k})(0) = 0$ , where  $\mathbf{v}$  and  $\mathbf{u}$  are known continuous functions and where for  $\mathbf{k} \in \mathbb{R}_+^4$ ,  $\tilde{\mathbf{k}} \in \mathbb{R}_+^3$  denotes the vector of the last three components of  $\mathbf{k}$ , that is, for all  $\mathbf{k} \in \mathbb{R}_+^4$

$$\mathbf{M}(\tilde{\mathbf{k}}) = \begin{pmatrix} -(k_2 + k_3) & k_4 \\ k_2 & -k_4 \end{pmatrix} \quad (6.15)$$

Specifically:

- (i) Gut:  $k_1 = k_{ga}$ ,  $k_2 = k_{pg}$ ,  $k_3 = k_{tg}$ ,  $k_4 = k_{gt}$ ,  $\mathbf{u} = 0$  and  $\mathbf{v} = C_a$
- (ii) Liver:  $k_1 = k_{fa}$ ,  $k_2 = k_{sf}$ ,  $k_3 = k_{mf}$ ,  $k_4 = k_{fm}$ ,  $\mathbf{u} = k_{pg}C_p$  and  $\mathbf{v} = C_a$

Since we have

$$\mathbf{C}(\mathbf{k})(t) = \int_0^t e^{((t-\tau)\mathbf{M}(\tilde{\mathbf{k}}))} (k_1\mathbf{v}(\tau) + \mathbf{u}(\tau)) d\tau, \quad (6.16)$$

the problems for gut and liver become:

**Problem 6.2.** *We need to recover  $\mathbf{k}$  from the knowledge of*

$$\begin{aligned} \alpha \cdot \mathbf{C}(\mathbf{k}) : [0, T] &\rightarrow \mathbb{R}, \\ t &\mapsto \alpha \cdot \int_0^t e^{((t-\tau)\mathbf{M}(\tilde{\mathbf{k}}))} (k_1\mathbf{v}(\tau) + \mathbf{u}(\tau)) d\tau. \end{aligned} \quad (6.17)$$

Thus we can apply the Gauss-Newton method as discussed in Section 5.2. We can summarize it in an algorithmic way as follow:

1. measured data:  $D \in C^1([0, T], \mathbb{R}_+)$
2. initial guess:  $\mathbf{k}^0$
3. Newton step: solving in  $\mathbf{h}^0 \in \mathbb{R}^4$

$$\left[ \frac{d\mathcal{F}}{d\mathbf{k}}(\mathbf{k}^0; \mathbf{h}^0) \right] (t) = D(t) - \alpha \cdot \mathbf{C}(\mathbf{k}^0)(t), \text{ for all } t \in [0, T]. \quad (6.18)$$

where  $\mathcal{F} : \mathbb{R}_+^4 \rightarrow C^1([0, T], \mathbb{R}_+)$  such that  $[\mathcal{F}(\mathbf{k})](t) = \alpha \cdot \mathbf{C}(\mathbf{k})(t)$

4. increment the value  $\mathbf{k}^0$ , giving  $\mathbf{k}^1 = \mathbf{k}^0 + \mathbf{h}^0$  and iterate the process.

Be recalling that equation (6.18) may have no solution, we need to regularize the Gauss Newton method (as discussed in Section 5.2) in the following way (also discretization):

1. we have measured data  $D$  for sampling time points  $t_1, \dots, t_N$
2. equation (6.18) becomes  $\mathbf{F}\mathbf{h} = D$
3. solving the linear system (6.18) by Tikhonov regularization, i.e.

$$(r\mathbf{I} + \mathbf{F}^T\mathbf{F})\mathbf{h} = \mathbf{F}^T D \quad (6.19)$$

4. increment the value  $\mathbf{k}^0$ , giving  $\mathbf{k}^1 = \mathbf{k}^0 + \mathbf{h}^0$  and iterate the process.

### 6.2.3 Results

#### Numerical validation

An Albira microPET system is currently operational at the IRCCS San Martino IST, Genova. Experiments with mice are currently performed at this site, by using different tracers, mainly for applications to oncology and to FDG physiology in the renal and liver systems [70, 26, 27, 28]. In this section we describe the performance of our approach to compartmental analysis in the case of synthetic data simulated by mimicking Albira acquisitions for FDG-PET experiments.

In order to produce the synthetic data we initially choose meaningful values for tracer coefficients for gut model, i.e.  $k_{tg}$ ,  $k_{pg}$ ,  $k_{gt}$  and  $k_{ga}$ . With these selected values we can solve (6.1) in term of  $\mathbf{C}_G$ : such a solution of the Cauchy problem is then sampled in time, on a time interval  $[t_1, t_N]$  that correspond to the total acquisition time with Albira and in correspondence of time points typical of experiments with Albira in



this application context. The IF  $C_b$  has been obtained by fitting with a gamma variate function a set of real measurements acquired from a healthy mouse in a very controlled experiment. Then we have simulated a physiological reliable value for the volume of a mouse gut ( $V_G = 0.4 \text{ cm}^3$ ) and computed the activity of radioactive tracer in gut as  $\mathbf{C}_G V_G$  in order to affect such data by Poisson noise; we have then computed back the *noisy gut concentration* and applied our algorithm in order to reconstruct the exchange coefficients. Once the simulation is been performed, we have used the reconstructed value for  $k_{pg}$  to solve equation (6.6) and find the portal vein concentration  $C_p$ : this concentration is then used, with a set of synthetic values for  $k_{mf}$ ,  $k_{sf}$ ,  $k_{fm}$  and  $k_{fa}$ , to solve (6.7) in term of  $\mathbf{C}_L$ . Again the concentration is sampled in time and then a reliable value for the liver of mouse is simulated ( $V_L = 0.1 \text{ cm}^3$ ) to produce values for the activity of radio-tracer in liver (useful to affect it by Poisson noise), produce the *noisy liver concentration* and then apply our algorithm. We have performed different simulations, varying both the initial guess (starting point of the Gauss-Newton Regularized method), the ground truth set of tracer coefficients, the noise on the data, the regularization parameter: in order to perform this huge set of simulation, we have built up a Matlab Graphic User Interface (GUI) that, slightly modified, has been used also in real cases, for the automatic analysis of compartmental model.

We present here two different simulation experiments, with two different but reliable sets of values for the tracer kinetic parameters: the two sets for the tracer kinetic parameters are represented in the ground truth (g.t.) row in Table 6.1 and in Table 6.2. Comparison with the g.t. values for this parameters provides limits about the reliability of the model and of the inversion procedure: the standard deviations are systematically small, thus showing the notable numerical stability of the iterative reconstruction scheme with respect to noise (mean and standard deviations are computed over 50 runs of the same problem with different - random - initial guess). For both the experiments we also applied to the data a standard multi-linear fitting approach [3, 10, 41] that utilizes the Levenberg-Marquardt technique for least-squares minimization. Tables 6.1 and 6.2 show that our approach produces significantly smaller uncertainties (with the same computational burden).

**Remark 6.1.** *We are aware that the “trial and error” approach to find the regularization parameter is not the best choice to find an optimal value for this parameter. Different techniques can be implemented, as, for example, using the SVD of the operator describing the inverse problem, and choose the regularization parameter according to singular values of the operator. We just started with the study of a similar approach, that could make our GUI really an automatic interface.*

	$k_{ga}$	$k_{tg}$	$k_{pg}$	$k_{gt}$	$k_{fa}$	$k_{mf}$	$k_{sf}$	$k_{fm}$
g. t.	1.00	0.10	1.30	0.20	1.50	0.01	1.80	2.00
R-GN	$1.00 \pm 0.06$	$0.11 \pm 0.02$	$1.31 \pm 0.02$	$0.20 \pm 0.01$	$1.50 \pm 0.06$	$0.02 \pm 0.01$	$1.82 \pm 0.03$	$2.07 \pm 0.08$
LS	$1.00 \pm 0.07$	$0.08 \pm 0.04$	$1.32 \pm 0.02$	$0.19 \pm 0.03$	$1.53 \pm 0.08$	$0.03 \pm 0.03$	$1.77 \pm 0.10$	$1.89 \pm 0.12$

Table 6.1: Ground truth (g.t.) and reconstructed values for tracer kinetic parameters and for the first simulated experiment, with both our method (Regularized Gauss Newton) and a standard least-squares approach (LS). We also used  $10^4$  as a regularization parameter for both gut and liver (trial and error approach), and to stop the iterative algorithm we checked the difference in the approximation of the reconstructed concentration with the original one, using a threshold as a stopping criterion. Here such value is  $10^{-4}$  for both gut and liver. Means and standard deviations are computed over 50 different runs of the algorithm, with different (random) initial guess.

	$k_{ga}$	$k_{tg}$	$k_{pg}$	$k_{gt}$	$k_{fa}$	$k_{mf}$	$k_{sf}$	$k_{fm}$
g. t.	1.50	0.01	1.00	0.02	1.00	0.10	0.60	0.02
R-GN	$1.48 \pm 0.04$	$0.01 \pm 0.01$	$1.02 \pm 0.03$	$0.02 \pm 0.01$	$1.03 \pm 0.05$	$0.09 \pm 0.02$	$0.62 \pm 0.03$	$0.02 \pm 0.01$
LS	$1.48 \pm 0.05$	$0.01 \pm 0.01$	$0.95 \pm 0.07$	$0.02 \pm 0.02$	$1.03 \pm 0.07$	$0.11 \pm 0.03$	$0.67 \pm 0.10$	$0.02 \pm 0.02$

Table 6.2: Ground truth (g.t.) and reconstructed values for tracer kinetic parameters and for the first simulated experiment, with both our method (Regularized Gauss Newton) and a standard least-squares approach (LS). We also used  $10^4$  as a regularization parameter for both gut and liver (trial and error approach), and to stop the iterative algorithm we checked the difference in the approximation of the reconstructed concentration with the original one, using a threshold as a stopping criterion. Here such value is  $10^{-4}$  for both gut and liver. Means and standard deviations are computed over 50 different runs of the algorithm, with different (random) initial guess.

### Real data

21 mice were subdivided into three groups according to pre-imaging protocol: controls ( $n=7$ ) received no treatment, STS ( $n=7$ ) were starved with free access to water for 48 hours and MTF ( $n=7$ ) were treated with metformin (750 mg/Kg per day) for one month. All mice underwent a dynamic microPET study for 50 min after FDG injection through a tail vein. In vivo imaging was performed according to a protocol validated in our lab [71]. To ensure a steady state of substrate and hormones governing glucose metabolism, all animals were studied after six hours fasting. Mice were weighted and anesthesia was induced by intra-peritoneal administration of ketamine/xylazine (100 and 10 mg/kg, respectively). Serum glucose level was tested and animals were positioned on the bed of Albira, whose two-ring configuration permits to cover the whole animal body in a single bed position. A dose of 3-4 MBq of FDG was then injected through a tail vein, soon after the start of a list mode acquisition lasting 50 minutes. Acquisition was reconstructed using the following framing rate:  $10 \times 15$  sec,  $5 \times 30$  sec,  $2 \times 150$  sec,  $6 \times 300$  sec,  $1 \times 600$  sec. PET data were reconstructed using a Maximum Likelihood Expectation Maximization method (MLEM). Thereafter, each image dataset was reviewed by an experienced observer who recognized three regions of interest (ROIs) encompassing left ventricle, gut and liver, respectively (see Figure 6.3).

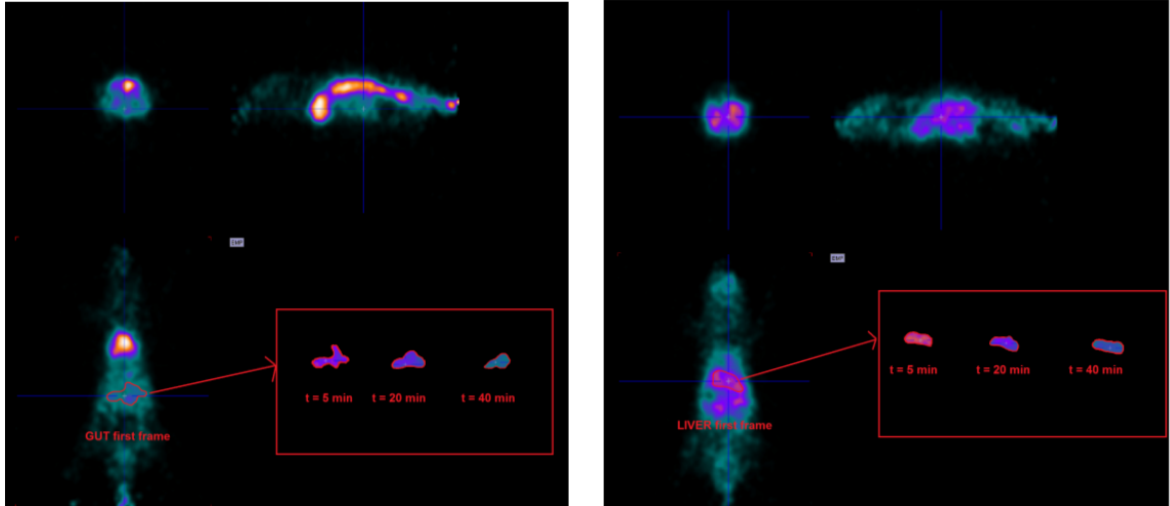


Figure 6.3: Examples of ROIs on gut and liver at different time points. Left panel ROIs on gut. Right panel ROIs on liver.

	Weight [g]	Glycaemia [mg/dL]	SUV gut [1]	SUV liver [1]
control	$24.4 \pm 1.8$	$131.1 \pm 12.1$	$1.07 \pm 0.36$	$1.07 \pm 0.33$
MTF	$23.9 \pm 1.3$	$89.1 \pm 18.6$	$2.49 \pm 0.97$	$1.03 \pm 0.31$
STS	$23.7 \pm 1.6$	$54.0 \pm 24.3$	$1.51 \pm 0.39$	$1.17 \pm 0.27$

Table 6.3: Means and standard deviations for weight, glycaemia and SUV over the three sets of mice (control, metformin, starvation)

	$k_{fa}$	$k_{mf}$	$k_{sf}$	$k_{fm}$	$k_{fp}(= k_{pg})$
control	$2.180 \pm 0.568$	$0.003 \pm 0.003$	$2.510 \pm 0.510$	$0.154 \pm 0.059$	$2.291 \pm 0.530$
MTF	$2.112 \pm 0.420$	$0.004 \pm 0.003$	$2.478 \pm 0.389$	$0.041 \pm 0.026$	$2.068 \pm 0.757$
STS	$2.384 \pm 0.689$	$0.095 \pm 0.051$	$3.801 \pm 1.650$	$0.563 \pm 0.221$	$2.135 \pm 0.646$

Table 6.4: Means and standard deviations for liver kinetics parameters over the three sets of mice (control, metformin, starvation)

The experiments were completed in all animals and no side effects occurred at the drug dosage used. Body weight was stable and glycaemia was reduced only in STS mice (Table 6.3). In agreement with our previous experience [69], response of whole body glucose metabolism was characterized by a relative reduction in FDG clearance and whole body glucose consumption in both MTF and STS animals. This systemic effect was paralleled by a peculiar response of tracer distribution in the different investigated organs. In fact, MTF treatment was followed by an increase in gut tracer retention, and thus in higher gut average SUV with respect to control ( $1.07 \pm 0.36$  VS  $1.51 \pm 0.39$ ,  $p < 0.01$  VS control mice). On the contrary, FDG liver accumulation was eventually similar, independently of experimental conditions (Table 6.3).

In each animal the arterial TAC was obtained from a ROI drawn within the left ventricle and the PV TAC was provided as output of the compartmental analysis concerned with the gut subsystem, as shown, for an example, in Figure 6.4: TAC in PV well fitted with expectations based on previous literature [54, 82] indicating a good precision of our estimation of the dual input for the liver system.

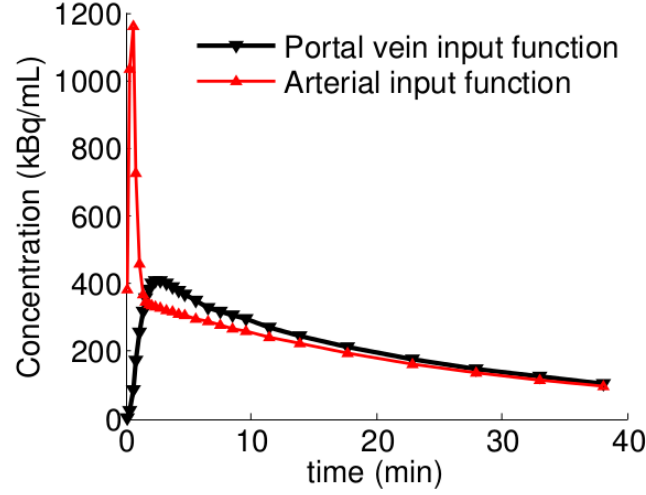


Figure 6.4: Comparison between HA and PV in one of the mice.

In our compartmental analysis, rate constants of FDG handling in liver were indicated as  $k_{mf}$  and  $k_{fm}$  that represent, respectively, the exchange coefficients from the free compartment to the phosphorylated one and its reciprocal. According to the values of  $k_{fm}$  given in Table 6.4 and Figure 6.5 (top), glucose accumulation in liver was left virtually unaltered by MTF ( $0.004 \pm 0.003$ ,  $p = ns$  vs control mice), while it was markedly increased by STS ( $0.095 \pm 0.051$ ,  $p < 0.05$  vs control mice and MTF group), as expected as a consequence of fasting condition. The rate coefficient  $k_{mf}$  describes de-phosphorylation of FDG-6P inside the liver. According to Table 6.4 and Figure 6.5 (bottom), this process was markedly reduced by MTF ( $0.041 \pm 0.026$  vs  $0.154 \pm 0.059$ ,  $p < 0.01$  vs control mice). On the contrary,  $k_{fm}$  was increased by STS to  $0.563 \pm 0.221$  ( $p < 0.01$  vs control mice). This last result gives evidence to the occurrence of an increase of free FDG accumulation as a natural consequence of a significant caloric restriction. Table 6.4 also reports a marked increase of  $k_{sf}$  and a slight increase of  $k_{fa}$  in STS models, which corresponds to a significant release of FDG from mice liver to blood stream and a slightly less significant input increase from blood to liver, induced by starvation. The effect of MTF treatment on these coefficients was negligible. Finally, no relevant change in the values of  $k_{fp}$  for the three groups was observed. This study introduces a computation-based, model-dependent method for the use of FDG to study glucose kinetics in liver, whose main

advantages are that it is able to determine the dual input typical of liver physiology just by means of computation, thus avoiding both the use of invasive catheter-based measurement procedures and the major drawback of methods relying on the sole arterial IF, which are not able to account for the intrinsic variability of perfusion and glucose consumption in splanchnic organs.

From the physiological viewpoint, the description of the hepatic FDG handling provided by our compartmental analysis is pictured in Figure 6.6 where the black arrows indicate all the tracer kinetic parameters computed by the method and changes in the thickness of such arrows indicate variations of the coefficient values.

From a numerical viewpoint, this task was accomplished applying the regularized Newton algorithm to the inverse problems associated to the multi-dimensional Cauchy problems describing the two-compartment gut and liver systems, respectively. In this setting, the computation of this algorithm is rather effective, since the matrix differentiation step required at some stage can be realized analytically, thus avoiding time consuming numerical differentiation.

From a mathematical viewpoint, we are ensured about the uniqueness of the solution of the liver inverse problem, thanks to results presented in Chapter 5, for the 2-compartment catenary systems. The complex liver physiological system has been studied by means of two different 2-compartment catenary system, describing gut and liver, respectively, and a third compartment accounting for the presence of the portal vein, which has a central role in liver physiology. This modeling allowed for a quite straightforward application of the uniqueness results that we have proved and that are presented in the previous chapter.

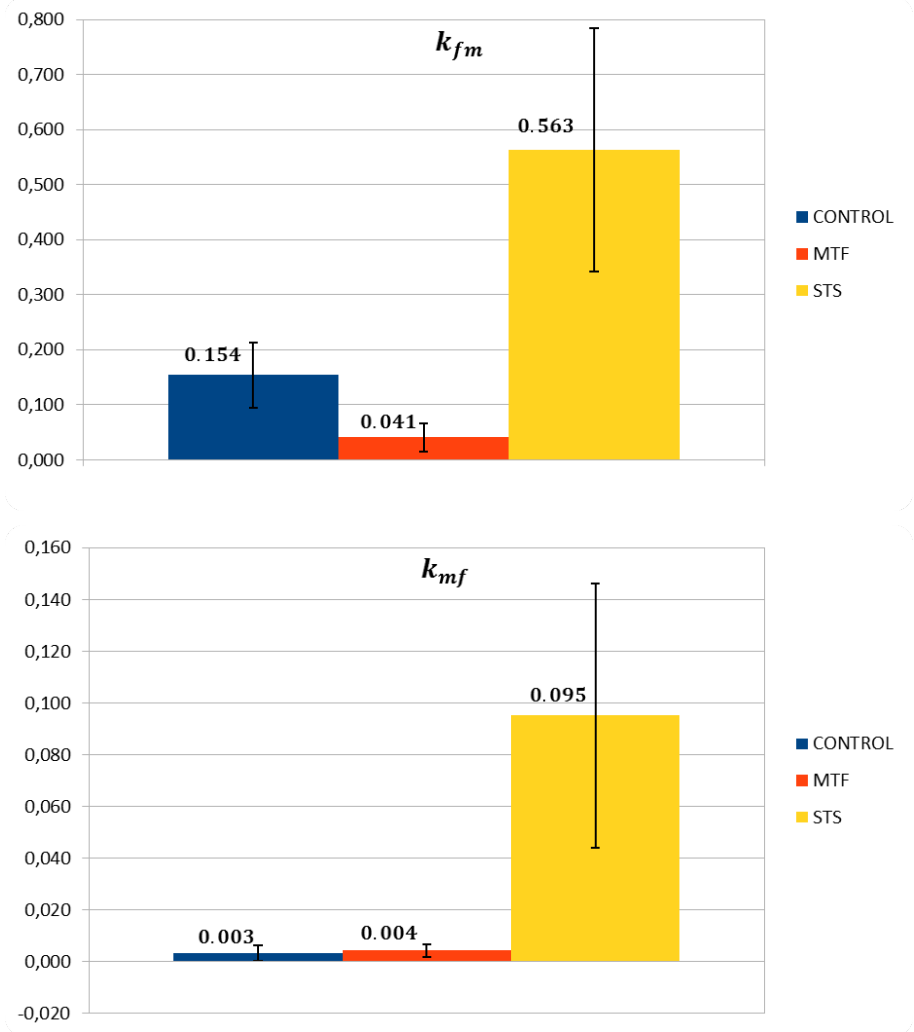


Figure 6.5: Comparison between the values of  $k_{fm}$  (top) and  $k_{mf}$  (bottom) in each set of mice. Each bar represents the mean over a specific set of mice, while the black line represents standard deviation.

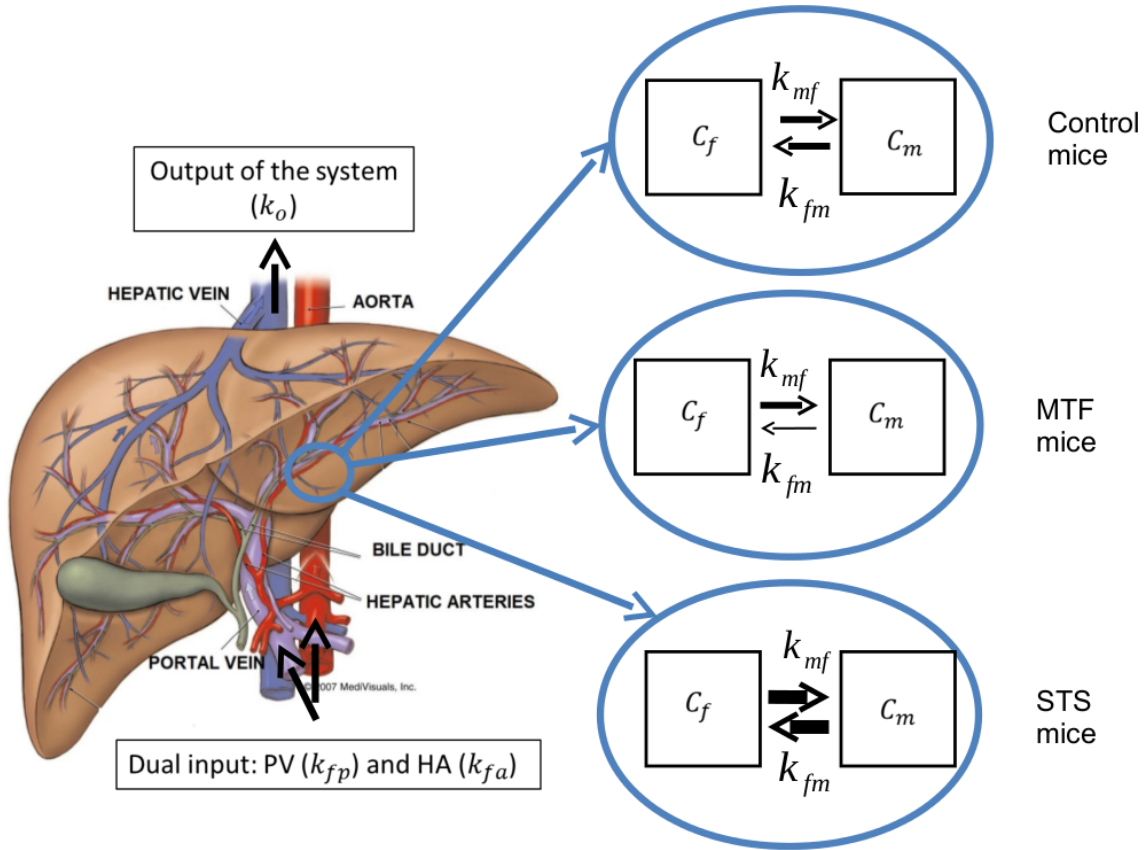


Figure 6.6: Picture of the FDG metabolism as designed by the novel compartmental model. The black arrows correspond to the computed tracer kinetic parameters. The thickness variation in the arrows corresponds to the variation of the coefficient values as reconstructed by the Gauss Newton approach.



## Chapter 7

# Application to real data: the kidneys problem

Clinical activity asks to represent glucose consumption by malignant cells as Standardized Uptake Value (SUV) (see subsection 1.2.2) whose estimation is hampered by the interference of several variables modulating tracer availability independently of tumor metabolism [19]. Among these factors, kidney function plays a relevant role as, differently from the tracked glucose, FDG is scarcely reabsorbed in renal tubule and is largely excreted in the urines [100]. Furthermore, variations in FDG concentration due to water re-absorption in renal tubules and increase of bladders volume during the FDG excretion process have to be taken in account, in order to describe the renal FDG metabolism.

In this chapter we introduce a novel compartmental model describing the excretion of FDG in the renal system, with 3 renal compartment and 1 output compartment (the bladder), to take into account the loss of FDG in urines. The peculiarity of this model allowed us to solve numerically the inverse problem without use of the Gauss-Newton regularized method. Actually, the reconstruction of the tracer kinetic parameters is obtained by solving the Maximum Likelihood problem iteratively, using a non-stationary Steepest Descent approach that explicitly accounts for the Poisson nature of nuclear medicine data.

## 7.1 Kidneys physiology

### 7.1.1 Diabetes

Urinary loss of glucose (typical of diabetes) and of FDG (that is not reabsorbed by renal tubule) might modify the diagnostic interpretation of PET/CT studies in diabetics patients with cancer. Under this condition, analysis of FDG distribution is often problematic, as the limited uptake of glucose and FDG by insulin dependent tissues is paralleled by the persistence of high tracer concentrations in plasma. Obviously, the ubiquitous presence of radioactive blood smoothens the difference between the pathologic uptake in cancer and the physiological tracer retention in background surrounding tissues. Accordingly, interventions able to increase urinary FDG loss might improve image quality in the large number of diabetic patients submitted to PET/CT for cancer studies. However, testing this hypothesis asks to first verify whether and to what extent a fraction of urinary FDG is indeed reabsorbed by tubule, thus in turn requires the availability of methods able to describe the different mechanisms tuning FDG kinetics within the renal compartments. Thus we have to determine the efficiency of urinary excretion.

### 7.1.2 Reabsorption and excretion

The standard way to quantitatively determine the efficiency of urinary excretion in nuclear medicine is based on either *measurements of the mean renal clearance* or the application of *compartmental analysis techniques*. Mean renal clearance is defined as the ratio between the asymptotic activity in the bladder and the integral over the whole time interval of the plasma input function [95]. On the other hand, FDG compartmental analysis typically involves two compartments: one where the tracer is free and one where the tracer is trapped inside the cells due to 6-phosphorylation. However, the potential of these conventional approaches is notably hindered by several factors: 1) it is difficult to estimate urinary tracer loss in renal pelvis, due to the fact that the urinary flow is extremely high and can be variable during the experiment; 2) FDG concentration is influenced by fluid and nutrients re-absorption anatomically occurring in the tubule; 3) it is not possible to quantitatively estimate the renal excretion without involving a pool for the tracer anatomically represented by the bladder.

As stated for the study of the hepatic compartmental problem, we again studied the changes of renal FDG metabolism in mice treated with interventions able to blunt kidneys G6P de-phosphorylation and glucose delivery, in order to validate our algorithm by testing its performances in real experiments.

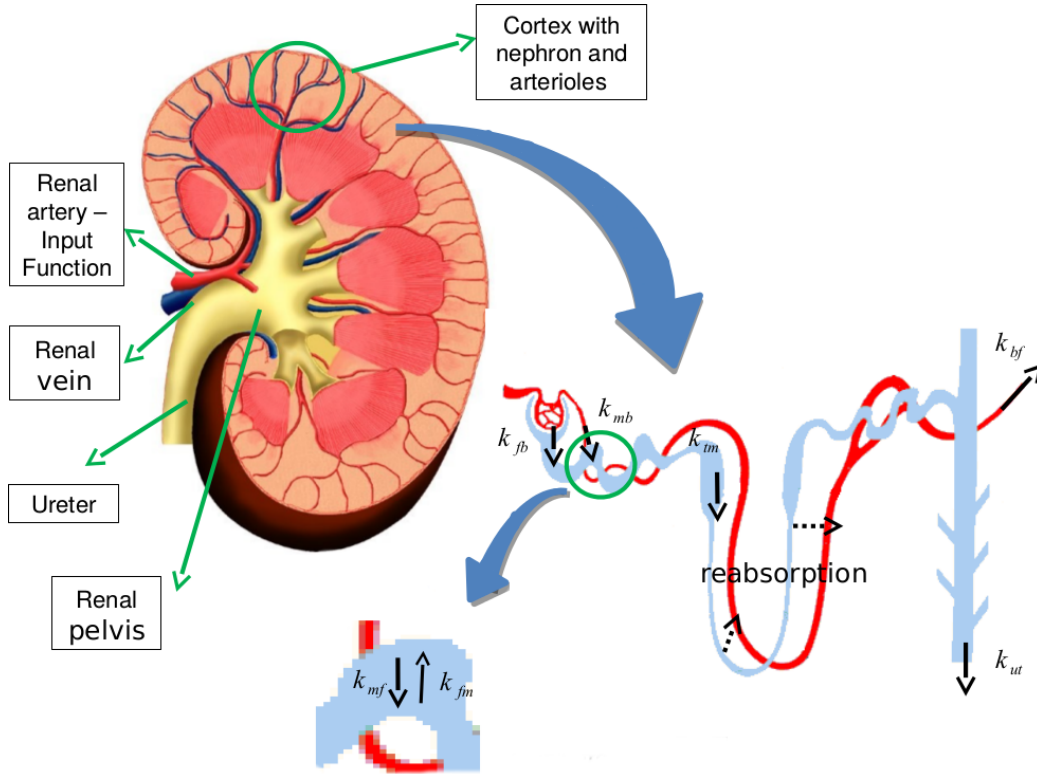


Figure 7.1: Pictorial representation of a human kidney. It is possible to notice the very complex renal FDG metabolism coupled with the very complex renal physiology itself.

## 7.2 The urinary compartmental problem

We now describe the FDG excretion process by taking advantage of microPET studies in mice, using a system that simultaneously covers the whole animal body, allowing the measurement of the tracer amount sequestered over time in the bladder [26]. Further, we introduce a novel compartmental model (see Figure 7.2) involving:

1. an extravascular compartment accounting for free tracer, whose exchange with blood (compartment  $C_b$ ) is allowed ( $C_f$ );
2. a compartment involving the metabolized, 6-phosphorylated FDG ( $C_m$ );
3. the tubular compartment, in which tracer concentration can vary because of water reabsorption ( $C_t$ );

4. the urinary pool representing the ultimate tracer excretion that is measured in the bladder ( $C_u$ ).

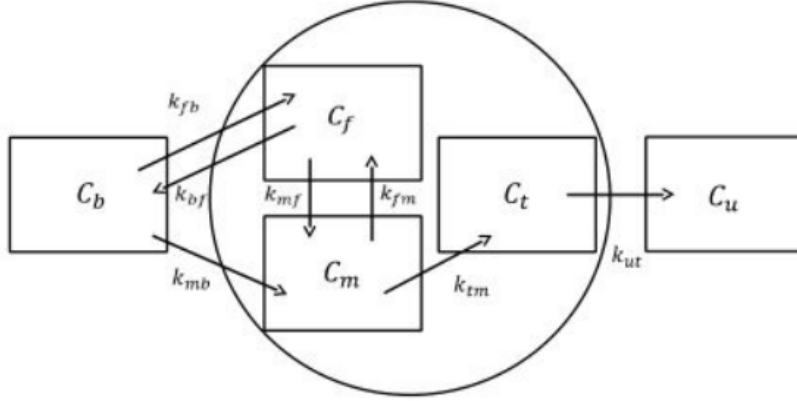


Figure 7.2: Compartmental model for the urinary system: in the circle the 3-compartment compartmental system for kidneys, and as output the bladder compartment  $C_u$ .

### 7.2.1 Model equation

#### The forward problem

Standard compartmental analysis [30, 46, 92, 97, 98, 26] models the renal tracer kinetics by means of two functionally (but not spatially) distinct compartments: the tissue in which the tracer is free,  $C_f$ , and the metabolized component  $C_m$ . Since we want to quantitatively assess the role of mechanisms transporting FDG back to metabolism, here we also include the re-absorption compartment  $C_t$ , anatomically identified with the tubular system; finally, the urine  $C_u$  represents the pool where the tracer is accumulated and is anatomically identified with the bladder. The tracer concentrations  $C_f$ ,  $C_m$ ,  $C_t$  and  $C_u$  associated to these four compartments are well-defined state variables measured in Bq ml<sup>-1</sup>. These state variables satisfy the following Cauchy problem, expressing the balance of tracer between compartments:

$$\dot{\mathbf{C}} = \mathbf{MC} + \mathbf{KC}_b \quad (7.1)$$

$$\frac{d}{dt}(V_u C_u) = F_{ut} C_t \quad (7.2)$$

with initial conditions  $\mathbf{C}(0) = C_u(0) = 0$  and where

$$\mathbf{C} = \begin{pmatrix} C_f \\ C_m \\ C_t \end{pmatrix}, \quad (7.3)$$

$$\mathbf{M} = \begin{pmatrix} -(k_{bf} + k_{mf}) & k_{fm} & 0 \\ k_{mf} & -(k_{fm} + k_{tm}) & 0 \\ 0 & k_{tm} & -k_{ut} \end{pmatrix}, \quad (7.4)$$

$$\mathbf{K} = \begin{pmatrix} k_{fb} \\ k_{mb} \\ 0 \end{pmatrix}. \quad (7.5)$$

The constants  $k_{ij}$  ( $\text{min}^{-1}$ ) represent the usual exchange coefficients (rate constants or tracer kinetic parameters) between compartments in contact and are positive and the plus, minus signs characterize incoming and outgoing fluxes, respectively. The peculiarity of equation (7.2) with respect to equation (7.1) is in the fact that the bladder  $C_u$  is the only compartment whose volume changes in time. In particular, this equation states that the urine accumulates in  $C_u$  at a rate which is constant (because of the resting state of mice), with a bulk flow  $F_{ut}$  ( $\text{ml min}^{-1}$ ) entering  $C_u$  from  $C_t$ . Finally,  $C_b$ , representing the tracer concentration in the blood, is the canonical Input Function (IF) and plays the role of trigger for the process of tracer kinetics in the system. This corresponds to considering a Single Input Model, which is more feasible than the Dual Input Model [105, 106] for application of microPET renal analysis.

### Global balance equations

By combining equation (7.1) and equation (7.2) it is rather straightforward to obtain

$$\begin{aligned} \dot{C}_f + \dot{C}_m = & -k_{bf}(C_f + C_m) + \left(\frac{k_{bf}}{k_{tm}} - 1\right) \left[ \frac{d^2}{dt^2}(V_u C_u) + \right. \\ & \left. + \frac{k_{ut}}{F_{ut}} \frac{d}{dt}(V_u C_u) \right] + (k_{fb} + k_{mb})C_b \end{aligned} \quad (7.6)$$

and

$$\begin{aligned} \frac{d^2}{dt^2}(V_u C_u) = & \frac{F_{ut} k_{tm} D}{k_{ut}(k_{bf} - k_{tm})} (C_f + C_m) + \\ & + \frac{F_{ut} k_{tm} (k_{fm} + k_{tm} + k_{mf})}{k_{ut}(k_{bf} - k_{tm})} (\dot{C}_f + \dot{C}_m) + \\ & - \frac{1}{k_{ut}} \frac{d^3}{dt^3} (V_u C_u) - \frac{F_{ut} \xi}{(k_{bf} - k_{tm}) k_{ut}} \dot{C}_b, \end{aligned} \quad (7.7)$$

where

$$D := k_{mf} + k_{tm} + k_{bf} k_{fm} + k_{bf} k_{tm} \quad (7.8)$$

and

$$\xi := (k_{fm} + k_{tm} + k_{mf}) k_{fb} + (k_{mf} + k_{bf} + k_{mf}) k_{mb}. \quad (7.9)$$

We now observe that the total radioactivity concentration in the kidneys,  $C_K$ , can be written (in a standard way, as seen in Chapter 4) as

$$C_K - \eta_b C_b = \alpha \cdot \mathbf{C} \quad (7.10)$$

where

$$\alpha = \begin{pmatrix} 1 - \eta_b - \eta_t \\ 1 - \eta_b - \eta_t \\ \eta_t \end{pmatrix} \quad (7.11)$$

and where  $\eta_t$  and  $\eta_b$  are the fractions of kidney volumes  $V_K$  occupied, respectively, by the tubular compartment and the blood. Unlike from the approach that we have applied in the analysis of the liver system, here we want to obtain some *global balance equation* for the measured data (i.e. the radio-isotope activity), using the peculiarity of this specific compartmental model (that essentially allows a direct connection between two different organs in just one compartmental system <sup>1</sup>) in order to exploit the characteristics of noise on data.

---

<sup>1</sup>For the liver case we also had a direct connection between two different organs, gut and liver, but we distinguished the compartmental models, using a standard 2-compartment system for gut and a (nonstandard - Dual Input) 2-compartment system for liver; then we reduced them separately and in cascade through the Gauss Newton Regularized method. Here we consider two different connected organs (kidneys and bladder) as belonging to the same compartmental model. Accordingly we have to consider two different sets of measurements: concentration (or activity) on kidneys and concentration (or activity) on bladder: this peculiarity allows us to build up a different reduction algorithm which is very specific for this compartmental system.

Integrating once equation (7.6) and twice equation (7.7) in time and using equation (7.10) leads to

$$\begin{aligned} A_K = & V_K \eta_b C_b - V_K k_{bf} \int_0^t C_K(\tau) d\tau + \\ & + V_K [\eta_b k_{bf} + (1 - \eta_b - \eta_t)(k_{fb} + k_{mb})] \int_0^t C_b(\tau) d\tau + \\ & + V_K \left( k_{bf} \frac{\eta_t}{F_{ut}} + \beta k_{ut} \right) V_u C_u + V_K \left( \frac{\eta_t}{F_{ut}} + \beta \right) \frac{d}{dt} (V_u C_u) \end{aligned} \quad (7.12)$$

and

$$\begin{aligned} A_u = & \frac{1}{\gamma} \left[ \frac{F_{ut} k_{tm} k_{mf}}{1 - \eta_b - \eta_t} \int_0^t d\tau \int_0^\tau d\tau' C_K(\tau') + \right. \\ & + (F_{ut} k_{tm} k_{mb} - \frac{F_{ut} k_{tm} k_{mf}}{1 - \eta_b - \eta_t} \eta_b) \int_0^t d\tau \int_0^\tau d\tau' C_b(\tau') + \\ & \left. - \frac{d}{dt} (V_u C_u) - \left( \frac{k_{tm} k_{mf} \eta_t}{1 - \eta_b - \eta_t} + (k_{fm} + k_{tm} + k_{mf}) k_{ut} \right) \int_0^t (V_u C_u)(\tau) d\tau \right] \end{aligned} \quad (7.13)$$

where

$$A_K = V_K C_K, \quad (7.14)$$

$$\beta := \frac{1 - \eta_b - \eta_t}{F_{ut}} \left( \frac{k_{bf}}{k_{tm}} - 1 \right), \quad (7.15)$$

$$A_u = V_u C_u. \quad (7.16)$$

and

$$\gamma := k_{ut} + k_{fm} + k_{tm} + k_{mf}. \quad (7.17)$$

$A_K$  and  $A_u$  denote the total radioactivity in kidneys and bladder, respectively. Both equation (7.12) and equation (7.13) can be simplified by observing that the terms containing  $d/dt(C_u V_u)$  are much smaller than the terms in the sums at the right hand sides. Therefore the (approximated) *global balance equations* for the tracer in this compartmental model are

$$A_K = x_1 V_K C_b + x_2 \int_0^t A_K + x_3 V_K \int_0^t C_b + x_4 V_K V_u C_u \quad (7.18)$$

and

$$A_u = z_1 \int_0^t \int_0^\tau C_K + z_2 \int_0^t \int_0^\tau C_b + z_3 \int_0^t V_u C_u, \quad (7.19)$$

where the column vectors of macroparameters,  $x$  and  $z$ , are defined by the equations

$$(x_1, x_2, x_3, x_4)^T := \left( \eta_b, \quad -k_{bf}, \quad \eta_b k_{bf} + (1 - \eta_b - \eta_t)(k_{fb} + k_{mb}), \quad k_{bf} \frac{\eta_t}{F_{ut}} + \beta k_{ut} \right)^T \quad (7.20)$$

and

$$(z_1, z_2, z_3)^T := \frac{1}{\gamma} \left( \frac{F_{ut} k_{tm} k_{mf}}{1 - \eta_b - \eta_t}, F_{ut} k_{tm} k_{mb} - \frac{F_{ut} k_{tm} k_{mf}}{1 - \eta_b - \eta_t}, \right. \\ \left. - \frac{k_{tm} k_{mf} \eta_t}{1 - \eta_b - \eta_t} - (k_{fm} + k_{tm} + k_{mf}) k_{ut} \right)^T \quad (7.21)$$

with  $T$  denoting the matrix transpose.

In equations (7.18) and (7.19) the activities  $A_K$  and  $A_u$  represent the input experimental data. They are obtained by drawing Region of Interests (ROIs) on PET images at different time points, in correspondence with kidneys and bladder, respectively. The Time Activity Curve (TAC)  $C_b$  is obtained in an analogous way, using ROIs on the left ventricle. On the other hand, the equation solutions  $x$  and  $z$  contain the physiological unknowns of the process, i.e. the tracer coefficients  $k_{bf}, k_{mf}, k_{fm}, k_{fb}, k_{tm}, k_{mb}, k_{ut}$ , together with the volume fractions  $\eta_b$  and  $\eta_t$  and the flux  $F_{ut}$ . However, before considering the numerical solution of the two equations, we introduce two physiological constraints that reduce the number of these unknowns. First, we observe that according to the assumption of stationarity the constant flux rate of the fluid into bladder can be estimated from the measured bladder volumes  $V_u(t_f)$ , at the final time  $t_f$ , and  $V_u(\bar{t})$ , at an intermediate time  $\bar{t}$ . It is found that

$$F_{ut} = \frac{V_u(t_f) - V_u(\bar{t})}{t_f - \bar{t}}. \quad (7.22)$$

It follow that  $F_{ut}$  is given by the experimental knowledge of  $V_u$  at different time points. Second, we have to account for the outgoing flux rate of free fluid from the compartment  $C_t$ . Table 1 in [74] shows that the tracer flow toward the bladder is around two orders of magnitude smaller than the re-absorbed flow. Since the volume of the re-absorption compartment is a measurable anatomical volume, the tracer balance equation in tubule implies  $k_{tm} = 10^2 k_{ut}$ . Moreover it follows from the definitions that  $k_u t \eta_t V_K = F_{ut}$ .

### 7.2.2 A Maximum Likelihood - Steepest Descent method

PET activities are acquired at different time points  $t_1, \dots, t_N$  according to pre-defined acquisition paradigms. This implies the following discretization of equations



(7.18) and (7.19):

$$y_i = A_R(t_i), \quad w_i = A_u(t_i), \quad i = 1, 2, \dots, N; \quad (7.23)$$

where  $y$  and  $w$  are  $N$ -dimensional column vectors;  $\mathbf{A}$  is the  $N \times 4$  matrix with rows

$$V_K C_b(t_i), \quad \int_0^{t_i} A_K d\tau, \quad V_K \int_0^{t_i} C_b d\tau, \quad V_R V_u C_u(t_i) \quad (7.24)$$

$i = 1, 2, \dots, N$ : the  $N \times 3$  matrix  $\mathbf{B}$  with rows given by

$$\int_0^{t_i} \int_0^\tau C_K d\tau d\tau', \quad \int_0^{t_i} \int_0^\tau C_b d\tau d\tau', \quad \int_0^{t_i} V_u C_u d\tau, \quad (7.25)$$

$i = 1, 2, \dots, N$  (according to the definitions 7.24 and 7.25 the elements of  $\mathbf{A}$  and  $\mathbf{B}$  are positive). Therefore the discretized version of equations (7.18) and (7.19) for  $x$  takes the form

$$y = \mathbf{A}x \quad (7.26)$$

while eq 7.19 for  $z$  yields

$$w = \mathbf{B}z. \quad (7.27)$$

The optimization problem defined by equations (7.26) and (7.27) is characterized by two peculiarities that make it significantly different with respect to the more traditional frameworks. First, most compartmental models of the renal system include just two functional compartments, both anatomically embedded (but not anatomically distinguishable) in the kidneys. On the contrary here we have a third (functional and anatomical) compartment represented by the tubular structures and describing the re-absorbing process of the tracer inside the kidneys; and a fourth anatomical compartment, the bladder, which simply accumulates the excreted tracer during the output process. This implies that the number of unknowns in the present case is much *higher* and that traditional optimization schemes may not work properly for this model. The second peculiarity of equations (7.26) and (7.27) is that in both cases the data are activities characterized by *Poisson statistics*. This has an important technical implication since this piece of a priori information can be exploited in the parameter identification process by applying a *maximum likelihood* (see Appendix B) approach in which the likelihood is provided by the Kullback Leibler divergence.

By repeating standard computations typical of ML-problems (see Appendix B), the problem of finding  $x$  and  $z$  can be rewritten in the form (we define this approach

to optimization in the case of equation (7.26) and the extension to equation (7.27) is straightforward):

$$\tilde{x} = \arg \min_{x \in \mathbb{R}^4} D_{KL}(y, x) \quad (7.28)$$

where  $D_{KL}(y, x) = \frac{2}{N} \sum_{i=1}^N y_i \log \frac{y_i}{(\mathbf{A}x)_i} + (\mathbf{A}x)_i - y_i + y_i \log(y_i)$  is the Kullback Leibler divergence. In a large part of cases, physical motivations imply that the components of parameters  $x$  has to be non-negative; this is not one of these cases, so we can't use the well-known method of *constrained* ML (EM, see Appendix B), but we are still interested in building up an iterative algorithm that could reaches the minimum of  $D_{KL}$  in a finite number of iterations.

If  $\nabla D_{KL}$  is the gradient of the Kullback Leibler divergence, we can write the steepest descent method associated to our problem as

$$x^{(k+1)} = x^{(k)} - \tau^{(k)} \nabla D_{KL}(y, x), \quad (7.29)$$

where  $\tau^{(k)}$  is the step size and it is allowed to change every iteration. Thus we have the following Theorem [27]

**Theorem 7.1.** *Let  $U(x^{(k)}) := \frac{y}{\mathbf{A}x^{(k)}}$  and  $V(x^{(k)}) := \mathbf{A}^T \mathbf{1}$ ,  $\forall k$ . Then if*

$$0 < \tau < \min_i \frac{-\mathbf{A}x^{(k)}}{\mathbf{A}(U(x^{(k)}) - V(x^{(k)}))} \quad (7.30)$$

*the algorithm 7.29 can reach the minimum of  $D_{KL}$  in a finite number of iteration.*

The proof of the theorem follows immediately by exploiting the constraints we have on  $\tau$ , essentially:

- (i)  $\tau^{(k)} > 0$  for all  $k$ ;
- (ii)  $(\mathbf{A}x^{(k+1)})_i > 0$  for every  $k$  and for all  $i = 1, \dots, N$ ;
- (iii)  $D_{KL}(y, x^{(k+1)}) \leq D_{KL}(y, x^{(k)})$  for every  $k$ .

**Remark 7.2.** *The algorithm we implemented for the solution of equation (7.26) reads as follows: given  $x^{(k)}$ ,*

1. choose  $\tau^{(k)}$  such that (7.1) holds;

2. compute  $x^{(k+1)}$  in (7.29);
3. if  $D_{KL}(y, x^{(k+1)}) \leq D_{KL}(y, x^{(k)})$  go on with the iterations. Otherwise
4. decrease  $\tau^{(k)}$  until  $D_{KL}(y, x^{(k+1)}) \leq D_{KL}(y, x^{(k)})$  and when this happens, go on with the iterations.

The scheme is stopped when  $|D_{KL}(x^{(k+1)}) - D_{KL}(x^{(k)})| \cdot |x^{(k+1)} - x^{(k)}| < \epsilon$ , with  $\epsilon$  an appropriate threshold.

### 7.2.3 Results

#### Numerical validation

An Albira microPET system produced by Carestream Health is currently operational at the IRCCS San Martino IST, Genova. Experiments with mice are currently performed at this site, by using different tracers, mainly for applications to oncology and to FDG physiology in the renal and liver systems [70, 26, 27, 28]. In this section we describe the performance of our approach to compartmental analysis in the case of synthetic data simulated by mimicking Albira acquisitions for FDG-PET experiments.

In order to produce the synthetic data we initially choose meaningful values for  $V_K$ ,  $V_u(t_0)$  and  $F_{ut}$ ; then we started from six g.t. values for the tracer kinetic parameters  $k_{bf}$ ,  $k_{mf}$ ,  $k_{fm}$ ,  $k_{fb}$ ,  $k_{mb}$  and  $k_{ut}$  and for  $\eta_b$ . We also exploited the two relations  $k_{tm} = 10^2 k_{ut}$  and (7.22). With these selected values we can solve equation (7.1) in terms of  $C_m$  and  $C_f$  and  $C_t$ ; then we solve (7.2) and find  $V_u C_u$ . These solutions of the Cauchy problem are sampled in time, on a time interval  $[t_1, t_N]$  that correspond to the total acquisition time with Albira and in correspondence of time points typical of experiments with Albira in this application context. The IF  $C_b$  has been obtained by fitting with a gamma variate function a set of real measurements acquired from a healthy mouse in a very controlled experiment. Then, using vectors corresponding to the discretization of  $C_m$ ,  $C_f$ ,  $C_t$  and  $C_u$ , it is possible to compute the matrices **A** and **B**, and the data vectors  $y$  and  $w$  in the same equations (7.26) and (7.27). Finally we affected the data by Poisson noise and applied our algorithm in order to reconstruct the exchange coefficients and  $\eta_b$ . We point out that this kind of test is specific for the validation of the physiological model, since the synthetic data have been computed by integrating equations different from those of the model (i.e., no inverse crime has been performed).

	$k_{bf}$	$k_{ut}$	$k_{mf}$	$k_{fb}$	$k_{fm}$	$k_{mb}$	$\eta_b$
g. t.	0.60	1.30	0.60	1.00	0.20	0.30	0.20
ML-SD	$0.63 \pm 0.04$	$1.31 \pm 0.05$	$0.60 \pm 0.01$	$1.01 \pm 0.03$	$0.21 \pm 0.02$	$0.32 \pm 0.03$	$0.22 \pm 0.04$
LS	$0.67 \pm 0.16$	$1.34 \pm 0.31$	$0.66 \pm 0.12$	$1.12 \pm 0.19$	$0.26 \pm 0.10$	$0.32 \pm 0.11$	$0.19 \pm 0.09$

Table 7.1: Ground truth (g.t.) and reconstructed values for tracer kinetic parameters and  $\eta_b$  for the first simulated experiment, with both our method (Maximum Likelihood-Steepest Descent) and a standard least-squares approach (LS). Here  $V_K = 0.15$  ml,  $F_{ut} = 0.04$  ml min<sup>-1</sup>,  $V_u(t_1) = 0.16$  ml. The signal-to-noise ratios for the activities  $A_K$  and  $A_u$  are 12.4 dB and 26.3 dB, respectively.

We realized two different simulation experiments, with two different but reliable sets of values for the tracer kinetic parameters. For both the simulation experiment we set  $V_K = 0.15$  ml,  $F_{ut} = 0.004$  ml/min and  $V_u(t_1) = 0.16$  ml: the two sets for the tracer kinetic parameters are represented in the ground truth (g.t.) row in table 7.1 and in table 7.1. The signal-to-noise ratios adopted for these tests are 12.4 dB for (7.26) and 26.3 dB for (7.27) in the case of the first experiment and 13.6 dB and 27.7 dB for the same equations, in the second experiment. Comparison with the g.t. values for this parameters provides limits about the reliability of the model and of the inversion procedure. Our approach reconstructs the input tracer kinetic parameters with relative errors never larger than 10%. Further, the standard deviations are systematically small, thus showing the notable numerical stability of the iterative reconstruction scheme with respect to noise (mean and standard deviations are computed over 50 runs of the same problem with different (random) initialization vectors  $x^{(0)}$  and  $z^{(0)}$ ). For both the experiments we also applied to the data a standard multi-linear fitting approach [3, 10, 41] that assumes Gaussian noise on data and utilizes the LevenbergMarquardt technique for least-squares minimization. Tables 7.1 and 7.2 show that our approach produces significantly smaller uncertainties (with the same computational burden).

### Real data of metformin treated mice

We apply now this compartmental approach to the investigation of the effects that metformin produces on the glucose excretion process. For our experiment we consider a control group (n=12) and a group (of the same size) in which FDG injection was performed after one month of high dose metformin treatment (750 mg/Kg body weight daily). In all animals 3700 KBq of FDG were injected into a tail vein after a

	$k_{bf}$	$k_{ut}$	$k_{mf}$	$k_{fb}$	$k_{fm}$	$k_{mb}$	$\eta_b$
g. t.	0.40	1.00	0.70	0.90	0.02	0.10	0.15
ML-SD	$0.41 \pm 0.03$	$1.02 \pm 0.04$	$0.73 \pm 0.03$	$0.92 \pm 0.03$	$0.02 \pm 0.01$	$0.09 \pm 0.01$	$0.16 \pm 0.02$
LS	$0.46 \pm 0.19$	$1.09 \pm 0.20$	$0.72 \pm 0.12$	$0.89 \pm 0.11$	$0.02 \pm 0.02$	$0.11 \pm 0.06$	$0.17 \pm 0.06$

Table 7.2: Ground truth (g.t.) and reconstructed values for tracer kinetic parameters and  $\eta_b$  for the second simulated experiment, with both our method and a standard least-squares approach (LS). Here  $V_K = 0.15$  ml,  $F_{ut} = 0.04$  ml min<sup>-1</sup>,  $V_u(t_1) = 0.16$  ml. The signal-to-noise ratios for the activities  $A_K$  and  $A_u$  are 13.6 dB and 27.7 dB, respectively.

fasting period of six hours during a dynamic list mode acquisition ( $10 \times 15s + 1 \times 22s + 4 \times 30s + 5 \times 60s + 2 \times 150s + 5 \times 300s$ ). The images have been reconstructed by applying an Ordered Subset Expectation Maximization (OSEM) iterative algorithm [35] while ROIs have been drawn over the left ventricle, kidneys and bladder in order to measure the activities  $A_K$  and  $A_u$  (see Figure 7.3). ROIs have been also drawn over the left ventricle in order to compute the IF.

Table 7.3 and Table 7.4 contain the reconstructed values of the tracer coefficients for the 12 untreated mice and the 12 metformin-treated mice. Means and standard deviations are computed by using 50 runs of the code for 50 different initializations  $x^{(0)}$  and  $z^{(0)}$  (the initialization is performed component-wise for these two vectors, by picking up numbers in the interval  $[0, 1]$  with uniform distribution). Figure 7.5 contains the average and the standard deviations of the values of the tracer coefficients computed over the treated and untreated mice. This plot shows that  $k_{fp}$  decreases of around one order of magnitude when the mice are treated with metformin while the variations of the other tracer coefficients are statistically negligible. Figure 7.4 studies the correlation between the urinary clearance (ml mm<sup>-1</sup>) and one of the tracer coefficients,  $k_{mb}$ , which indicates the efficiency with which the tracer from blood reaches the metabolized compartment. This correlation coefficient is very high in all cases, i.e. by considering separately non-treated and treated mice and by looking at all animals as a unique model set.

From the physiological viewpoint, the description of the renal FDG handling provided by our compartmental analysis is pictured in Figure 7.6 where the black arrows indicate all the tracer kinetic parameters computed by the method and changes in the thickness of such arrows indicate variations of the coefficient values. Figure 7.4 shows that there exists a strong correlation between  $k_{mb}$  and the mean urinary clear-

ance, and so that glomerular filtration (expressed in our compartmental framework just by  $k_{mb}$ ) is the major determinant of possible tracer elimination.

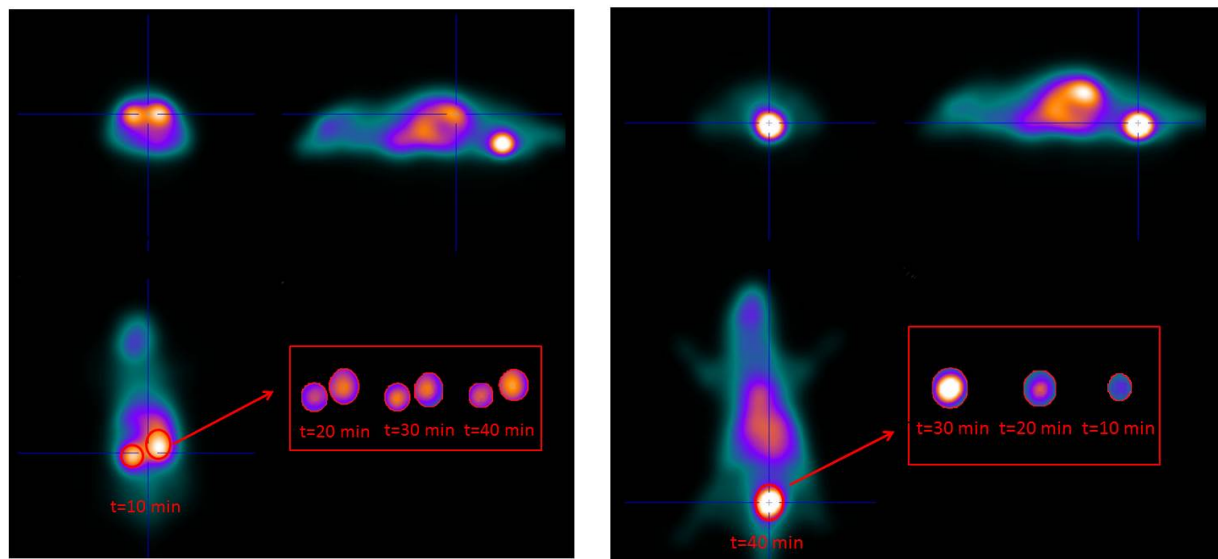


Figure 7.3: Examples of ROIs on kidneys and bladder at different time points. Left panel ROIs on kidneys. Right panel ROIs on bladder.

	$k_{bf}$	$k_{ut}$	$k_{mf}$	$k_{fb}$	$k_{fm}$	$k_{mb}$	$\eta_b$	$\eta_t$
m1	$0.32 \pm 0.06$	$0.30 \pm 0.04$	$0.20 \pm 0.05$	$(6.21 \pm 0.23)10^{-4}$	$(1.61 \pm 0.41)10^{-3}$	$0.08 \pm 0.01$	$0.19 \pm 0.02$	$0.03 \pm 0.01$
m2	$1.29 \pm 0.12$	$1.09 \pm 0.08$	$0.36 \pm 0.06$	$(9.96 \pm 0.34)10^{-4}$	$0.52 \pm 0.04$	$1.32 \pm 0.07$	$0.24 \pm 0.02$	$0.03 \pm 0.01$
m3	$0.21 \pm 0.06$	$0.22 \pm 0.04$	$0.16 \pm 0.03$	$(6.91 \pm 0.23)10^{-4}$	$(2.59 \pm 0.13)10^{-4}$	$0.46 \pm 0.09$	$0.19 \pm 0.02$	$0.09 \pm 0.02$
m4	$0.23 \pm 0.07$	$0.26 \pm 0.09$	$0.02 \pm 0.01$	$(5.01 \pm 3.14)10^{-4}$	$0.05 \pm 0.01$	$1.32 \pm 0.09$	$0.16 \pm 0.04$	$0.04 \pm 0.01$
m5	$0.11 \pm 0.02$	$0.10 \pm 0.02$	$0.10 \pm 0.03$	$(7.71 \pm 0.35)10^{-3}$	$(9.11 \pm 4.32)10^{-4}$	$0.37 \pm 0.06$	$0.14 \pm 0.03$	$0.02 \pm 0.01$
m6	$0.58 \pm 0.07$	$0.56 \pm 0.05$	$0.02 \pm 0.01$	$0.03 \pm 0.01$	$0.02 \pm 0.01$	$0.80 \pm 0.09$	$0.28 \pm 0.08$	$0.03 \pm 0.01$
m7	$0.40 \pm 0.06$	$0.47 \pm 0.11$	$0.13 \pm 0.02$	$(9.60 \pm 3.02)10^{-3}$	$0.23 \pm 0.07$	$1.00 \pm 0.14$	$0.24 \pm 0.06$	$0.05 \pm 0.02$
m8	$0.43 \pm 0.05$	$0.35 \pm 0.05$	$0.20 \pm 0.03$	$(6.44 \pm 2.01)10^{-4}$	$0.10 \pm 0.01$	$1.88 \pm 0.10$	$0.21 \pm 0.03$	$0.03 \pm 0.01$
m9	$0.50 \pm 0.06$	$0.61 \pm 0.07$	$0.11 \pm 0.02$	$(4.39 \pm 2.72)10^{-4}$	$0.09 \pm 0.02$	$0.81 \pm 0.06$	$0.26 \pm 0.03$	$0.04 \pm 0.01$
m10	$0.31 \pm 0.05$	$0.41 \pm 0.04$	$0.18 \pm 0.03$	$(7.01 \pm 3.32)10^{-4}$	$0.06 \pm 0.01$	$0.08 \pm 0.01$	$0.31 \pm 0.03$	$0.05 \pm 0.02$
m11	$0.28 \pm 0.02$	$0.29 \pm 0.03$	$0.21 \pm 0.02$	$(3.87 \pm 1.57)10^{-4}$	$0.04 \pm 0.01$	$0.43 \pm 0.03$	$0.22 \pm 0.01$	$0.03 \pm 0.01$
m12	$0.38 \pm 0.06$	$0.30 \pm 0.04$	$0.12 \pm 0.01$	$(5.04 \pm 2.09)10^{-4}$	$0.24 \pm 0.03$	$0.66 \pm 0.05$	$0.27 \pm 0.02$	$0.04 \pm 0.01$

Table 7.3: Tracer coefficients and volume rates for 12 healthy untreated mice representing the control group. The uncertainties are computed as standard deviations over 50 runs of the algorithm with 50 different random initialization values.

	$k_{bf}$	$k_{ut}$	$k_{mf}$	$k_{fv}$	$k_{fm}$	$k_{mb}$	$\eta_b$	$\eta_t$
m13	$0.62 \pm 0.08$	$0.34 \pm 0.05$	$0.04 \pm 0.01$	$0.04 \pm 0.02$	$(8.59 \pm 1.56)10^{-4}$	$0.41 \pm 0.04$	$0.23 \pm 0.02$	$0.02 \pm 0.01$
m14	$0.12 \pm 0.02$	$0.14 \pm 0.01$	$0.02 \pm 0.01$	$0.12 \pm 0.02$	$(3.65 \pm 0.23)10^{-4}$	$0.65 \pm 0.10$	$0.17 \pm 0.01$	$0.01 \pm 0.01$
m15	$1.90 \pm 0.12$	$0.24 \pm 0.01$	$0.08 \pm 0.02$	$0.42 \pm 0.09$	$0.03 \pm 0.01$	$1.84 \pm 0.16$	$0.30 \pm 0.05$	$0.01 \pm 0.01$
m16	$0.51 \pm 0.09$	$0.52 \pm 0.11$	$0.15 \pm 0.04$	$0.20 \pm 0.02$	$0.01 \pm 0.01$	$0.90 \pm 0.10$	$0.27 \pm 0.07$	$0.06 \pm 0.02$
m17	$0.20 \pm 0.05$	$0.11 \pm 0.02$	$0.01 \pm 0.01$	$0.11 \pm 0.02$	$(1.57 \pm 0.24)10^{-5}$	$0.42 \pm 0.07$	$0.31 \pm 0.04$	$0.02 \pm 0.01$
m18	$0.11 \pm 0.02$	$0.17 \pm 0.01$	$0.04 \pm 0.01$	$0.15 \pm 0.04$	$0.01 \pm 0.01$	$0.70 \pm 0.13$	$0.25 \pm 0.09$	$0.07 \pm 0.02$
m19	$0.73 \pm 0.13$	$0.72 \pm 0.09$	$0.10 \pm 0.02$	$0.22 \pm 0.05$	$0.02 \pm 0.01$	$0.06 \pm 0.01$	$0.21 \pm 0.07$	$0.04 \pm 0.01$
m20	$0.38 \pm 0.08$	$0.36 \pm 0.04$	$0.06 \pm 0.01$	$0.01 \pm 0.01$	$(1.64 \pm 0.48)10^{-4}$	$0.91 \pm 0.17$	$0.31 \pm 0.07$	$0.08 \pm 0.02$
m21	$0.49 \pm 0.12$	$0.39 \pm 0.05$	$0.13 \pm 0.02$	$0.01 \pm 0.01$	$(1.71 \pm 0.09)10^{-3}$	$1.23 \pm 0.20$	$0.22 \pm 0.07$	$0.09 \pm 0.03$
m22	$0.36 \pm 0.08$	$0.40 \pm 0.10$	$0.05 \pm 0.01$	$0.02 \pm 0.01$	$(3.76 \pm 0.21)10^{-4}$	$0.42 \pm 0.12$	$0.19 \pm 0.02$	$0.08 \pm 0.03$
m23	$0.39 \pm 0.10$	$0.44 \pm 0.11$	$0.21 \pm 0.05$	$0.11 \pm 0.02$	$(3.97 \pm 0.21)10^{-4}$	$1.85 \pm 0.20$	$0.30 \pm 0.03$	$0.11 \pm 0.01$
m24	$0.13 \pm 0.03$	$0.12 \pm 0.01$	$0.11 \pm 0.02$	$0.01 \pm 0.01$	$(5.62 \pm 0.31)10^{-4}$	$0.06 \pm 0.01$	$0.23 \pm 0.03$	$0.06 \pm 0.01$

Table 7.4: Tracer coefficients and volume rates for 12 mice treated with metformin. The uncertainties are computed as standard deviations over 50 runs of the algorithm with 50 different random initialization values.



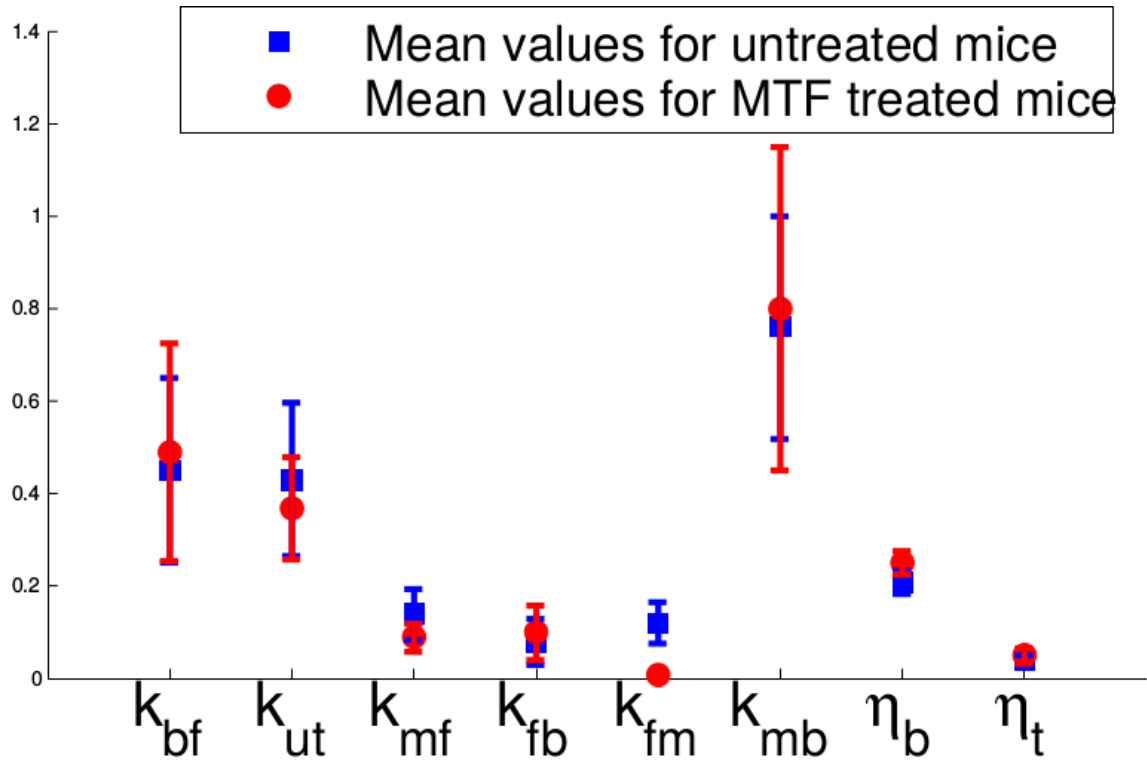


Figure 7.4: Means and standard deviations for the coefficients reconstructed via the maximum likelihood approach ( $k_{bf}, k_{ut}, k_{mf}, k_{fb}, k_{fm}, k_{mb}, \eta_b$  and  $\eta_t$ ) over the 2 sets of mice. In blue the control group of the healthy untreated mice, in red the group of metformin treated mice. It is possible to notice that  $k_{fm}$  is significantly reduced in metformin treated mice.

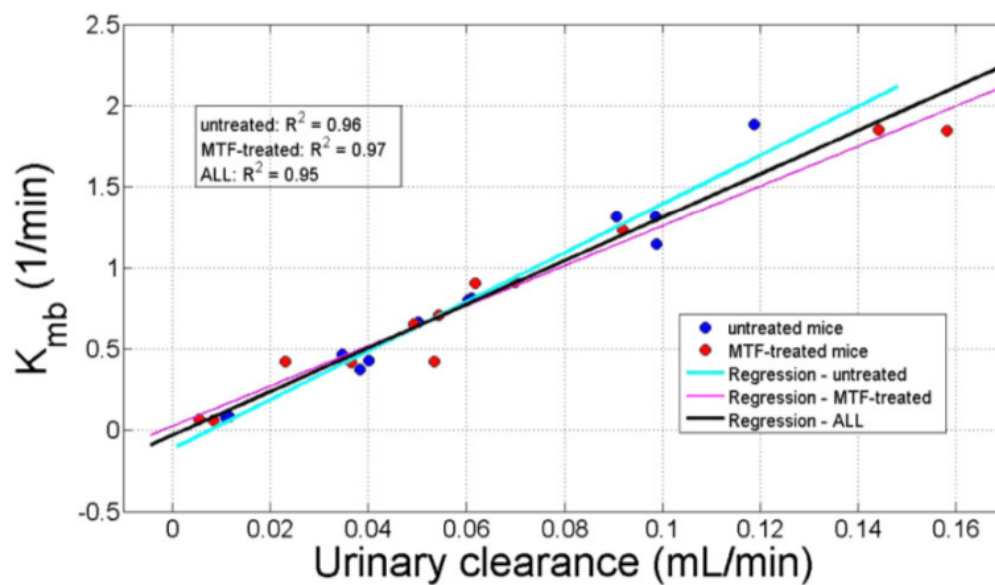


Figure 7.5: Correlation between the average clearance and the rate coefficients over the control group of 12 healthy untreated mice (light blue line); over the set of the 12 metformin-treated mice (magenta line); over the whole set of mice (black line).

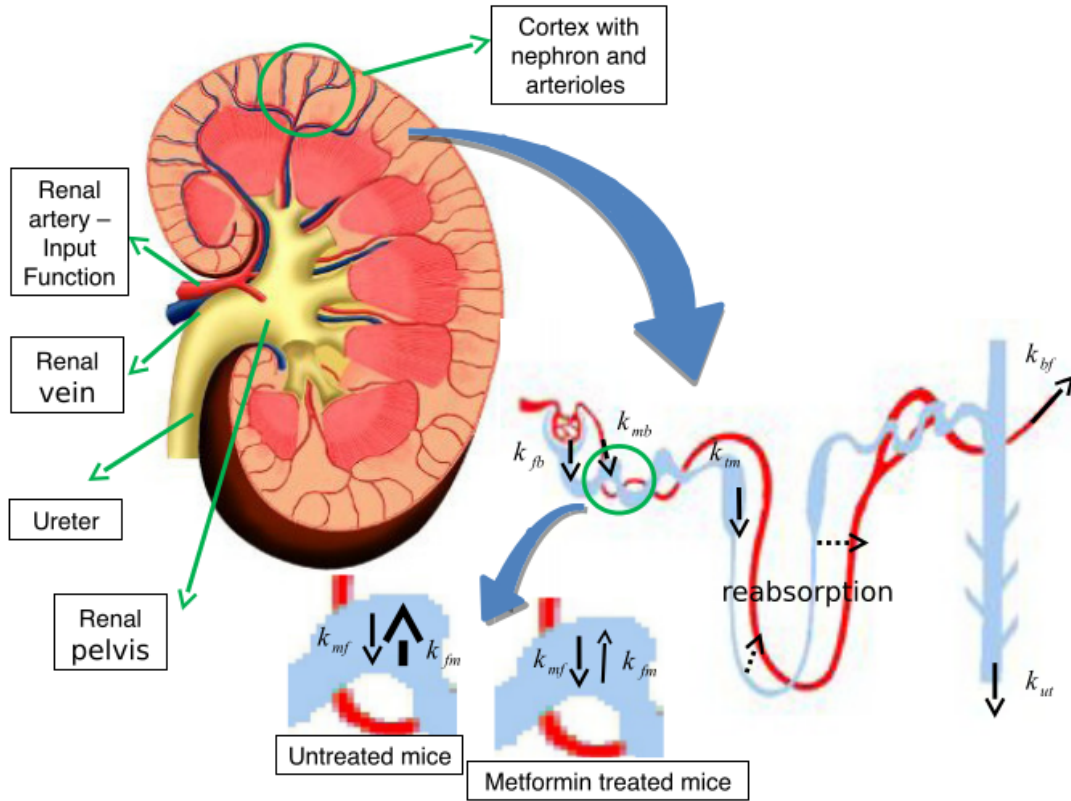


Figure 7.6: Picture of the FDG excretion process as designed by the novel compartmental model. Arterioles are represented in red and tubules in light blue. The black arrows correspond to the computed tracer kinetic parameters. The thickness variation in the arrows corresponds to the variation of the coefficient values as reconstructed by the maximum likelihood approach.



# Bibliography

- [1] S. Adams, R. Baum, T. Stuckensen, T. Bitter, and G. Hoer. Prospective comparison of fdg-pet with conventional imaging modalities (ct, mri, us) in lymph node staging of head and neck cancer. *Eur. J. Nucl. Med.*, 25:255–260, 1998.
- [2] T. Ahearn, R. Staff, T. Redpath, and S. Semple. The use of levenberg marquardt curve fitting algorithm in pharmacokinetic modelling of dce mri data. *Phys. Med. Biol.*, 50:72–85, 2005.
- [3] N. Alpert, R. Badgaiyan, E. Livni, and A. Fischman. A novel method for noninvasive detection of neuromodulatory changes in specific neurotransmitter system. *Neuroimage*, 19:1049–1060, 2003.
- [4] M. Jalving and J. Gietema, J. Lefrandt, S. Jong, A. Reyners, R. Gans, and E. Vries. Metformin: taking away the candy for cancer? *Eur. J. Cancer*, 46:2369–2380, 2010.
- [5] A. Basu, R. Basu, P. Shah, A. Vella, C. Johnson, K. Nair, M. Jensen, W. Schwenk, and R. Rizza. Evaluation of the effect of glucose ingestion and kinetic model configurations of fdg in the normal liver. *J. Nucl. Med.*, 35:818–823, 1994.
- [6] A. Basu, R. Basu, P. Shah, A. Vella, C. Johnson, K. Nair, M. Jensen, W. Schwenk, and R. Rizza. Effects of type 2 diabetes on the ability of insulin and glucose to regulate splanchnic and muscle glucose metabolism: evidence for a defect in hepatic glucokinase activity. *Diabetes*, 49:272–283, 2000.
- [7] F. Bengel, P. Ueberfuhr, N. Schiepel, S. Nekolla, B. Reichart, and M. Schwaiger. Effect of sympathetic reinnervation on cardiac performance after heart transplantation. *New Eng. J. Med.*, 345(10):731–738, 2001.

- [8] F. Benvenuto and M. Piana. Regularization of multiplicative iterative algorithms with nonnegative constraint. *Inverse Problems*, 30(3):035012, 2014.
- [9] M. Bertero and P. Boccacci. *Introduction to Inverse Problems in Imaging*. Institute of Physics Publishing, Bristol,.
- [10] G. Blomqvist. On the construction of functional maps in positron emission tomography. *J. Cereb. Blood Flow Metab.*, 4:629–632, 1984.
- [11] C. Bouman and K. Sauer. A unified approach to statistical tomography using coordinate descent optimization. *IEEE Trans. Imag. Proc.*, 5(3):480–492, 1996.
- [12] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge: University Press.
- [13] Bruker. *Albira Imaging, Bruker Albira Imaging System User Manual*. 2012.
- [14] R. Carver. *What do we talk about when we talk about love*. Knopf, 1981.
- [15] M. Casey and E. Hoffman. Quantification in positron emission tomography: a technique to reduce noise in accidental coincidence measurements and coincidence efficiency calibration. *J. Comput. Asst. Tomogr.*, 7(5):845–850, 1986.
- [16] M. Daube-Witherspoon and R. Carson. Unified deadtime correction model for pet. *IEEE Trans. Med. Imag.*, 10(3):267–275, 1991.
- [17] C. Davisson and R. Evans. Gamma-ray absorption coefficients. *Rev. Mod. Phys.*, 24(2):79, 1952.
- [18] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the em algorithm. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, pages 1–38, 1977.
- [19] C. Diederichs, L. Staib, G. Glatting, H. Beger, and S. Reske. Fdg pet: elevated plasma glucose reduces both uptake and detection rate of pancreatic malignancies. *J. Nucl. Med.*, 39:1030–1033, 1998.
- [20] M. Doubrovin, V. Ponomarev and T. Beresten, J. Balatoni, W. Bornmann, R. Finn, J. Humm, S. Larson, M. Sadelain, and R. Blasberg. Imaging transcriptional regulation of p53-dependent genes with positron emission tomography in vivo. *Proceedings of the National Academy of Sciences*, 98(16):9300–9305, 2001.

- [21] Itti E, M. Meignan, A. Berriolo-Riedinger, A. Biggi, A. Cashen, P. Véra, H. Tilly, B. Siegel, A. Gallamini, and R. Casasnovas. An international confirmatory study of the prognostic value of early pet/ct in diffuse large b-cell lymphoma: comparison between deauville criteria and  $\delta$ suvmax. *Eur. J. Nucl. Med. Mol. Imag.*, 40(9):1312–1320, 2013.
- [22] H. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, 1996.
- [23] N. Evans, R. Errington, M. Chapman, P. Smith, M. Chappell, and K. Godfrey. Compartmental modelling of the uptake kinetics of the anti-cancer agent topotecan in human breast cancer cells. *Internat. J. Adapt. Control Signal Process.*, 19(5):395–417, 2005.
- [24] J. Fessler. Hybrid poisson/polynomial objective functions for tomographic image reconstruction from transmission scans. *IEEE Trans. Imag. Proc.*, 4(10):1439–1450, 1995.
- [25] V. Fridman. Methods of successive approximations for fredholm integral equation of the first kind. *Uspekhi Mat. Nauk.*, 11(1):233–234, 1956.
- [26] S. Garbarino, G. Caviglia, M. Brignone, M. Massollo, G. Sambuceti, and M. Piana. Estimate of fdg excretion by means of compartmental analysis and ant colony optimization of nuclear medicine data. *Comput. Math. Methods. Med.*, 2013:793142, 2013.
- [27] S. Garbarino, G. Caviglia, G. Sambuceti, F. Benvenuto, and M. Piana. A novel description of fdg excretion in the renal system: application to metformin treated models. *Phys. Med. Biol.*, 59:2469–2484, 2014.
- [28] S. Garbarino, V. Vivaldi, F. delbary, G. Caviglia, M. Piana, C. MArini, S. Capitanio, I. Calamia, A. Buschiazzo, and G. Sambuceti. A new compartmental method for the analysis of liver fdg metabolism. *Eur. J. Nucl. Med. Mol. Imag.*, *submitted*.
- [29] C. Groetsch. *Generalized Inverses of Linear Operators: Representation and Approximation*. Dekker, New York, 1997.
- [30] R. Gunn, S. Gunn, and V. Cunningham. Positron emission tomography compartmental models. *J. Cereb. Blood Flow Metab.*, 21:635–652, 2001.

- [31] R. Harrison, D. Haynor, and T. Lewellen. Dual energy window scatter corrections for positron emission tomography. In *Nuclear Science Symposium and Medical Imaging Conference, 1991., Conference Record of the 1991 IEEE*, pages 1700–1704. IEEE, 1991.
- [32] T. Hays and G. Segall. A mathematical model for the distribution of fluorodeoxyglucose in humans. *J. Nucl. Med.*, 40:1358–1366, 1999.
- [33] J. Hearon. Theorems on linear systems. *Ann. New York Acad. Sci.*, 108:36–68, 1963.
- [34] K. Higashi, A. Clavo, and R. Wahl. Does fdg uptake measure proliferative activity of human cancer cells? in vitro comparison with dna flow cytometry and tritiated thymidine uptake. *J. Nucl. Med.*, 34:414419, 1993.
- [35] H. Hudson and R. Larkin. Accelerated image reconstruction using ordered subsets of projection. *IEEE Trans. Med. Imaging*, 13:601–609, 1994.
- [36] B. Hutton. Cardiac single-photon emission tomography: is attenuation correction enough? *Eur. J. Nucl. Med. Mol. Imaging*, 24(7):713–715, 1997.
- [37] B. Hutton, H. Hudson, and F. Beekman. A clinical perspective of accelerated statistical reconstruction. *Eur. J. Nucl. Med.*, 24(7):797–808, 1997.
- [38] A. Hwang, B. Franc, G. Gullberg, and B. Hasegawa. Assessment of the sources of error affecting the quantitative accuracy of spect imaging in small animals. *Phys. Med. Biol.*, 53:2233–2252, 2008.
- [39] A. Hwang and B. Hasegawa. Attenuation correction for small animal spect imaging using x-ray ct data. *Med. Phys.*, 32:2799–2804, 2005.
- [40] M. Ichise and JR Ballinger. From graphical analysis to multilinear regression analysis of reversible radioligand binding. *J. Cereb. Blood Flow Metab.*, 16:750–752, 1996.
- [41] M. Ichise, J. Liow, J. Lu, A. Takano, K. Model, H. Toyama, T. Suhara, K. Suzuki, R. Innis, and R. Carson. Linearized reference tissue parametric imaging methods: application to [<sup>11</sup>C]dasm positron emission tomography studies of the serotonin transporter in human brain. *J. Cereb. Blood Flow Metab.*, 23:1096–1112, 2003.



- [42] P. Iozzo, F. Geisler, V. Oikonen, M. Maki, T. Takala, O. Solin, E. Ferrannini, J. Knuuti, and P. Nuutila. Insulin stimulates liver glucose uptake in humans: an 18f-fdg pet study. *J. Nucl. Med.*, 44:682–689, 2003.
- [43] P. Iozzo, K. Hallsten, V. Oikonen, K. Virtanen, J. Kemppainen, O. Solin, E. Ferrannini, J. Knuuti, and P. Nuutila. Insulin-mediated hepatic glucose uptake is impaired in type 2 diabetes: evidence for a relationship with glycemic control. *J. Clin. Endocrinol. Metab.*, 88:2055–2060, 2003.
- [44] P. Iozzo, K. Hallsten, V. Oikonen, K. Virtanen, R. Parkkola, J. Kemppainen, O. Solin, F. Lonnqvist, E. Ferrannini, J. Knuuti, and P. Nuutila. Effects of metformin and rosiglitazone monotherapy on insulin-mediated hepatic glucose uptake and their relation to visceral fat in type 2 diabetes. *Diabetes Care*, 26:2069–2074, 2003.
- [45] J. Keyes Jr. Suv: Standard uptake or silly useless value? *J. Nucl. Med.*, 36:1836–1839, 1995.
- [46] M. Kamasak, C. Bouman, E. Morris, and K. Sauer. Direct reconstruction of kinetic parameter images from dynamic pet data. *IEEE Trans. Med. Imaging*, 24:636–650, 2005.
- [47] S. Keiding. Bringing physiology into pet of the liver. *J. Nucl. Med.*, 53:425–433, 2012.
- [48] Y. Kimura, M. Naganawa, M. Shidara, Y. Ikoma, and H. Watabe. Pet kinetic analysis pitfalls and a solution for the logan plot. *Ann. Nucl. Med.*, 21:1–8, 2007.
- [49] P. Kinahan and J. Rogers. Analytic 3d image reconstruction using all detected events. *IEEE Trans. Nuclear Science*, 36(1):964–968, 1989.
- [50] P. Kinahan, D. Townsend, T. Beyer, and D. Sashin. Attenuation correction for a combined 3d pet/ct scanner. *Med. Phys.*, 25(10):2046–2053, 1998.
- [51] S. Kis, I. Lajtos, M. Emri, L. Tron, G. Opposits, T. Bkki, Gy. Hegyesi, J. Imrek, I. Valastyn, J. Molnr, D. Novk, , and L. Balkay. Performance test of the minipet-ii small animal scanner according to the nema nu-4 standard. *IEEE Nucl. Sci. Symp. Conf. Record*, page 31853189, 2009.
- [52] T. Klepser and M. Kelly. Metformin hydrochloride: an antihyperglycemic agent. *Am. J. Health Syst. Pharm.*, 54:839–903, 1997.

- [53] L. Kostakoglu, H. Schoder, and J. Johnson. Interim fdg pet imaging in stage i/ii non bulky hodgkin lymphoma: Would using combined pet and ct criteria better predict response than each test alone. *Leuk Lymphoma*, 53:2143–2150, 2012.
- [54] N. Kudomi, M. Jarvisalo, J. Kiss, R. Borra, A. Viljanen, T. Viljanen, T. Sauvunen, J. Knuuti, H. Iida, P. Nuutila, and P. Iozzo. Non-invasive estimation of hepatic glucose uptake from [18f]fdg pet images using tissue-derived input functions. *Eur. J. Nucl. Med. Mol. Imaging*, 36:2014–2026, 2009.
- [55] L. Kunyansky. A new spect reconstruction algorithm based on the novikov explicit inversion formula. *Inverse problems*, 17(2):293, 2001.
- [56] E. Lage, J. Vaquero, A. Sisniega, S. Espana, G. Tapias, M. Abella, A. Rodriguez-Ruano, J. Ortuno, A. Udas, and M. Desco. Design and performance evaluation of a coplanar multimodality scanner for rodent imaging. *Phys. Med. Biol.*, 54:75427–5441, 2009.
- [57] D. Lalush and B. Tsui. Performance of ordered-subset reconstruction algorithms under conditions of extreme attenuation and truncation in myocardial spect. *J. Nucl. Med.*, 41:737–744, 2000.
- [58] D. Lalush and M. Wernick. Iterative image reconstruction. *Emission tomography: the fundamentals of PET and SPECT*, pages 443–472, 2004.
- [59] L. Landweber. An iteration formula for fredholm integral equations of the first kind. *Amer. J. Math*, 73:615–624, 1951.
- [60] K. Lange and J. Fessler. Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Trans. Image Process.*, 4(10):1430–1438, 1995.
- [61] H. Lanteri, M. Roche, and Claude C. Aime. Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms. *Inverse problems*, 18(5):1397–1419, 2002.
- [62] D. Lardinois, W. Weder, T. Hany, E. Kamel, S. Korom, B. Seifert, G. von Schulthess, and H. Steinert. Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. *New Eng. J. Med.*, 348(25):2500–2507, 2003.

- [63] M. Lauritzen and J. Olesen. Regional cerebral blood flow during migraine attacks by xenon-133 inhalation and emission tomography. *Brain*, 107(2):447–461, 1984.
- [64] B. Lee, H. Wellman, A. Siddiqui, H. Park, B. Mock, R. Worth, M. Edwards, and J. Krepshaw. Hipdm-spect in patients with medically intractable complex partial seizures: ictal study. *Arch. Neurol.*, 45(4):397–402, 1988.
- [65] P. Lindholm, H. Minn, S. Leskinen-Kallis, J. Bergman, U. Ruotsalainen, and H. Joensuu. Influence of the blood glucose concentration on fdg uptake in cancer pet study. *J. Nucl. Med.*, 34:1–6, 1993.
- [66] L. Junck, J. Olson, B. Ciliax, R. Koeppe, G. Watkins, D. Jewett, P. McKeever, D. Wieland, M. Kilbourn, Simon S. Starosta-Rubinstein, et al. Pet imaging of human gliomas with ligands for the peripheral benzodiazepine binding site. *Annals of neurology*, 26(6):752–758, 1989.
- [67] J. Logan. A review of graphical methods for tracer studies and strategies to reduce bias. *Nucl. Med. Biol.*, 30(8):833–844, 2003.
- [68] J. Logan, J. Fowler, N. Volkow, A. Wolf, S. Dewey, and D. Schlyer. Graphical analysis of reversible radioligand binding from time-activity measurements applied to n11c methylcocaine pet studies in human subjects. *J. Cereb. Blood Flow Metab.*, 10(5):740–747, 1990.
- [69] C. Marini, B. Salani, M. Massollo, A. Amaro, A. Esposito, A. Orengo, S. Capitanio, L. Ermionite, M. Riondato, G. Bottoni, C. Massara, S. Boccardo, M. Fabbi, C. Campi, S. Ravera, G. Angelini, S. Morbelli, M. Cilli, R. Cordera, M. Truini, D. Maggi U. Pfeffer, and G. Sambuceti. Direct inhibition of hexokinase activity by metformin at least partially impairs glucose metabolism and tumor growth in experimental breast cancer. *Cell Cycle*, 12:3490–3499, 2013.
- [70] M. Massollo, C. Marini, M. Brignone, L. Ermionite, B. Salani, M. Riondato, S. Capitanio, F. Fiz, A. Democrito, A. AMaro, S. Morbelli, M. Piana, D. Maggi, M. Cilli, U. Pfeffer, and G. Sambuceti. Metformin temporal and localized effects on gut glucose metabolism assessed using 18f-fdg pet in mice. *J. Nucl. Med.*, 54:259–266, 2013.
- [71] M. Massollo, C. MArini, M. Brignone, L. Ermionite, B. Salani, M. Riondato, S. Capitanio, F. Fiz, A. Democrito, S. Morbelli, M. Piana, D. Maggi, M. Cilli, U. Pfeffer, and G. Sambuceti. Metformin temporal and localized effects on

- gut glucose metabolism assessed using fdg in mice. *J. Nucl. Med.*, 54:259–266, 2013.
- [72] J. Mazziotta, M. Phelps, J. Pahl, S. Huang, L. Baxter, W. Riege, J. Hoffman, D. Kuhl, A. Lanto, and J. Wapenski. Reduced cerebral glucose metabolism in asymptomatic subjects at risk for huntington’s disease. *New Eng. J. Med.*, 316(7):357–362, 1987.
  - [73] S. Meikle, M. Dahlborm, and S. Cherry. Attenuation correction using count-limited transmission data in positron emission tomography. *J. Nucl. Med.*, 34(1):143–144, 1993.
  - [74] P. Meneton, I. Ichikawa, T. Inagami, and J. Schnermann. Renal physiology of the mouse. *Am. J. Physiol. Renal Physiol.*, 278:F339–F351, 2000.
  - [75] H. Miao, X. Xia, A. Perelson, and H. Wu. On identifiability of nonlinear ODE models and applications in viral dynamics. *SIAM Rev.*, 53(1):3–39, 2011.
  - [76] C. Minassian, N. Daniele, J. Bordet, C. Zitoun, and G. Mithieux. Liver glucose-6 phosphatase activity is inhibited by refeeding in rats. *J. Nutr.*, 125:2727–2732, 1995.
  - [77] S. Minoshima, K. Frey, R. Koeppe, N. Foster, and D. Kuhl. A diagnostic approach in alzheimer’s disease using three-dimensional stereotactic surface projections of fluorine-18-fdg pet. *J. Nucl. Med.*, 36(7):1238–1248, 1995.
  - [78] G. Mithieux, H. Vidal, C. Zitoun, N. Bruni, N. Daniele, and C. Minassian. Liver glucose-6 phosphatase activity is inhibited by refeeding in rats; glucose-6-phosphatase mrna and activity are increased to the same extent in kidney and liver of diabetic rats. *Diabetes*, 45:891–896, 1996.
  - [79] J. Moran, H. Lee, and M. Blauford. Optimization of urinary fdg excretion during pet imaging. *J. Nucl. Med.*, 40:1352–1357, 1999.
  - [80] W. Moses. Fundamental limits of spatial resolution in pet. *Nucl. Instrum. Methods Phys. Res. A*, 648:S236–S240, 2011.
  - [81] O. Munk, L. Bass, H. Feng, and S. Keiding. Determination of regional flow by use of intravascular pet tracers: microvascular theory and experimental validation for pig livers. *J. Nucl. Med.*, 44:1862–1870, 2003.

- [82] O. Munk, L. Bass, K. Roelsgaard, D. Bender, S. Hansen, and S. Keiding. Liver kinetics of glucose analogs measured in pigs by pet: importance of dual-input blood sampling. *J. Nucl. Med.*, 42:795–801, 2001.
- [83] O. Munk, S. Keiding, and L. Bass. Impulse-response function of splanchnic circulation with model-independent constraints: theory and experimental validation. *Am. J. Physiol. Gastrointest. Liver Physiol.*, 285:G671–G680, 2003.
- [84] F. Natterer. Inversion of the attenuated radon transform. *Inverse problems*, 17(1):113, 2001.
- [85] J. Ollinger. Model-based scatter correction for fully 3d pet. *Phys. Med. Biol.*, 41(1):153–176, 1996.
- [86] J. Ollinger and J. Fessler. Positron-emission tomography. *IEEE Signal Process. Mag*, 14:1053–5888, 1997.
- [87] M. Palmer, J. Rogers, M. Bergstrom, M. Beddoes, and B. Pade. Transmission profile filtering for positron emission tomography. *IEEE Trans. Nucl. Sci.*, 33(1):478–481, 1986.
- [88] S. Parter. On the eigenvalues and eigenvectors of a class of matrices. *J. Soc. Indust. Appl. Math.*, 8:376–388, 1960.
- [89] C. Patlak and Ronald G R. Blasberg. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. generalizations. *J. Cereb. Blood Flow Metab.*, 5(4):584–590, 1985.
- [90] W. Patterson. *Iterative Methods for the Solution of a Linear Operator Equation in Hilbert Space - A Survey*. Springer, Berlin, Heidelberg, New York, third edition.
- [91] M. Chen D. B. Stout Q. Bao, D. Newport and A. F. Chatziioannou. Performance evaluation of the inveon dedicated pet preclinical tomograph based on the nema nu-4 standards. *J. Nucl. Med.*, 50:401–408, 2009.
- [92] H. Qiao, J. Bai, Y. Chen, and J. Tian. Kidney modelling for fdg excretion with pet. *Int. J. Biomed. Imaging*, 2007:63234, 2007.
- [93] S. Rani, S. Nemanich, N. Fettig, and K. Shoghi. Kinetic analysis of fdg in rat liver: effect of dietary intervention on arterial and portal vein input. *Nucl. Med. Biol.*, 40:537–546, 2013.

- [94] E. Reiman, M. Fusselman, P. Fox, and M. Raichle. Neuroanatomical correlates of anticipatory anxiety. *Science*, 243(4894):1071–1074, 1989.
- [95] A. Rescigno. Clearance, turnover time, and volume of distribution. *Pharm. Res.*, 35:189–193, 1997.
- [96] F. Sanchez, L. Moliner, C Correcher, A González, A Orero, M Carles, A Soriano, MJ Rodriguez-Alvarez, LA Medina, and F Mora. Small animal pet scanner based on monolithic lyso crystals: performance evaluation. *Med. Phy.*, 39(2):643–653, 2012.
- [97] K. Schmidt and F. Turkheimer. Kinetic modeling in positron emission tomography. *Q. J. Nucl. Med. Mol. Imaging*, 46:70–85, 2002.
- [98] Y. Shen, H. Liu, and P. Shi. Limited view pet reconstruction of tissue radioactivity maps. *Comput. Med. Imaging Graph.*, 34:142–148, 2010.
- [99] L. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imag.*, 1(2):113–122, 1982.
- [100] P. Shreve, Y. Anzai, and R. Wahl. Pitfalls in oncologic diagnosis with fdg pet imaging: physiological and benign variants. *Radiographics*, 19:61–77, 1999.
- [101] L. Slimani, N. Kudomi, V. Oikonen, M. Jarvisalo, J. Kiss, A. Naum, R. Borra, A. Viljanen, H. Sipila, E. Ferrannini, T. Savunen, P. Nuutila, and P. Iozzo. Quantification of liver perfusion with [(15)o]h(2)o-pet and its relationship with glucose metabolism and substrate levels. *J. Hepatol.*, 48:974–982, 2008.
- [102] E. Smith, J. Jonides, and R. Koepp. Dissociating verbal and spatial working memory using pet. *Cerebral Cortex*, 6(1):11–20, 1996.
- [103] L. Sokoloff, M. Reivich, C. Kennedy, M. Des Rosiers, C. Patlak, K. Pettigrew, O. Sakurada, and M. Shinohara. The [14c]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure and normal values in the conscious and anesthetized albino rat. *J. Neurochem.*, 28:897–916, 1977.
- [104] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia: Society for Industrial and Applied Mathematics.
- [105] K. Tichuaner, K. Samkoe, K. Klubben, T. Hasan, and B. Pogue. Advantages of a dual-tracer model over reference tissue models for binding potential measurement in tumors. *Phys. Med. Biol.*, 57:6647–6659, 2012.

- [106] G. Tomasi, S. Kimberley, L. Rosso, E. Aboagye, and F. Turkheimer. Double-input compartmental modelling and spectral analysis for the quantification of positron emission tomography data in oncology. *Phys. Med. Biol.*, 57:1899–1906, 2012.
- [107] D. Townsend, P. Valk, and M. Maisey. *Positron emission tomography*. Springer-Verlag London Limited, 2005.
- [108] Y. Vardi, L. Shepp, and L. Kaufman. A statistical model for positron emission tomography. *J. Amer. Statist. Assoc.*, 80(389):8–20, 1985.
- [109] W. Wang, J-C. Georgi, and J. Humm. Evaluation of a compartmental model for estimating tumor hypoxia via fmiso dynamic pet imaging. *Phys. Med. Biol.*, 10:3083–3099, 2009.
- [110] O. Warburg. Über den stoffwechsel der carcinomzelle. *Naturwissenschaften*, 12(50):1131–1137, 1924.
- [111] P. Ward and C. Thompson. Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. *Cancer cell*, 21(3):297–308, 2012.
- [112] W. Weber. Assessing tumor response to therapy. *J. Nucl. Med.*, 50(Suppl 1):1S–10S, 2009.
- [113] M. Wernick and J. Aarsvold. *Emission tomography: the fundamentals of PET and SPECT*. Academic Press, 2004.
- [114] M. Winterdahl, O. Munk, M. Sorensen, F. Mortensen, and S. Keiding. Hepatic blood perfusion measured by 3-minute dynamic 18f-fdg pet in pigs. *J. Nucl. Med.*, 52:1119–1124, 2011.
- [115] S. Wollenweber. Parameterization of a model-based 3-d pet scatter correction. *IEEE Trans. Nuclear Science*, 49(1):722–727, 2002.
- [116] M. Xu, W. Luk, P. Cutler, and W. Digby. Local threshold for segmented attenuation correction of pet imaging in the thorax. *IEEE Trans. Nucl. Sci.*, 41:1532–1537, 1993.
- [117] J. Yates. Structural identifiability of physiologically based pharmacokinetic models. *J. Pharmacokinet. Pharmacodyn.*, 33:421–439, 2006.
- [118] D. Youla. *Mathematical theory of image restoration by the method of convex projection*. Image recovery: Theory and Applications.

- [119] Y. Yu, A. Annala, J. Barrio and T. Toyokuni, N. Satyamurthy, M. Namavari, S. Cherry, M. Phelps, H. Herschman, and S. Gambhir. Quantification of target gene expression by imaging reporter gene expression in living animals. *Nat. Med.*, 6(8):933–937, 2000.



# Appendix A

## Inverse problem theory

Inverse problems are quite a recent argument for the mathematical community. The first example of images obtained by solving a mathematical inverse problem was the CT (computed tomography) computer-generated images, that was provided by the first X-ray computed tomography machine, installed in 1971 in Great Britain. From that point important advances were made both the theory and the practice of inverse problems, and they go hand in hand with the technological development of the last decades, in particular in the medical diagnostic field. But what is an inverse problem? From the mathematical viewpoint, this concept has a certain degree of ambiguity: first of all the existence of an *inverse* problem, necessarily implies the existence of *direct* problem (which is commonly referred to as forward problem), which has to be strictly correlate with the inverse one. Thus two problems are inverse of one another when the formulation of each involves the solution of the other; in this sense, they can be considered as duals, which means that it is possible to exchange the role of the data with the one of the unknowns in order to pass from one problem to the other one. How is it thus possible to decide what is the direct problem and what is the inverse one? Mathematically, we define forward the problem of computing the effects of given causes, and inverse the problem related to the opposite process, i.e. to find the unknown causes from known consequences. In this sense, it is possible to say, in a naive way, that a forward problem is a problem oriented along a cause-effect sequence, while the associated inverse one reverses this cause-effect sequence. This classification seems to naturally introduce an important difference between the two problems: the direct problem usually enjoys good property, that corresponds to the definition of *well-posed* problem, while a typical property of inverse problem is the *ill-posedness*.

The basic concept of a well-posed problem was introduced by Hadamard in a

paper published in 1902, and in 1923 he gave the following definition of a well-posed problem:

1. **existence**: the solution exists for each given data;
2. **uniqueness**: the solution is unique for each given data;
3. **stability**: the solution depends continuously on the data.

Vice versa, a problem is ill-posed if it does not verify one of the three conditions.

From the mathematical viewpoint, this is meaningful only if it is enclosed in a suitable framework: thus let us define a normed space  $X$ , in which live the causes, a normed space  $Y$ , in which live the effects and an operator, say  $A$ , which transforms any elements of  $X$  into an element of  $Y$ . The solution of the direct problem typically causes a loss of information: as a consequence of this loss of information we find the ill-posedness of the inverse problems, i.e. the problem of determining the  $x \in X$  corresponding to a given  $y \in Y$ .

The definition of well-posedness became:

**Definition A.1** (Well-Posedness). *Let  $X$  and  $Y$  normed space,  $A$  an operator from  $X$  to  $Y$  and  $f \in X$  and  $g \in Y$ . The equation*

$$Af = g \tag{A.1}$$

*is well-posed if it has the following properties*

1. *existence*:  $\forall g \in Y$  exist at least an  $f \in X$  such that  $Af = g$ ;
2. *uniqueness*:  $\forall g \in Y$  exists only one  $f \in X$  such that  $Af = g$ ;
3. *stability*: for every succession  $f_n \in X$  such that  $\lim_{n \rightarrow +\infty} Af_n = Af$ , we have  $\lim_{n \rightarrow +\infty} f_n = f$ .

*The equations that does not satisfy one of these properties are called ill-posed.*

Such properties can be rewritten in a compact form, using an operational notation, as

1.  $R(A) = Y$ ;
2.  $N(A) = 0$ ;

3.  $A^{-1} : Y \rightarrow X$  is continuous

where  $R(A)$  and  $N(A)$  are, respectively, the range and the kernel of  $A$ . From now on we will consider the inverse problem with  $A \in L(X, Y)$ , the space of bounded linear operator between Hilbert spaces  $X$  and  $Y$ .

## A.1 The generalized inverse

The basic theory of best-approximate solutions of linear operator equations and of the generalized inverse of linear operator (between Hilbert spaces) is briefly provided here, but for a complete treatment, please refer to [29].

**Definition A.2.** *let  $A : X \rightarrow Y$  be a bounded linear operator,*

*(i)  $f \in X$  is called least-squares solution of (A.1) if*

$$\|Af - g\| = \inf\{\|Ax - g\| : x \in X\} \quad (\text{A.2})$$

*(ii)  $f \in X$  is called generalized solution of (A.1) if  $f$  is a least squares solution of (A.1) and*

$$\|f\| = \inf\{\|x\| : x \text{ is a solution of } Af = g\} \quad (\text{A.3})$$

*and we call it  $f^\dagger$ .*

The generalized solution is defined as the least-squares solution of minimal norm and this notion is related to the generalized inverse of  $A$  (the Moore-Penrose inverse):

**Definition A.3.** *The Moore-Penrose (generalized) inverse  $A^\dagger$  of  $A \in L(X, Y)$  is defined as the unique linear extension of  $\tilde{A}^{-1}$  to*

$$D(A^\dagger) := R(A) + R(A)^\perp \quad (\text{A.4})$$

*with*

$$N(A^\dagger) := R(A)^\perp \quad (\text{A.5})$$

*where*

$$\tilde{A} := A|_{N(A)^\perp} : N(A)^\perp \rightarrow R(A) \quad (\text{A.6})$$

$A^\dagger$  is well-defined: since  $N(\tilde{A}) = 0$  and  $R(\tilde{A}) = R(A)$ ,  $A^\dagger$  exists. Due to (A.5) and the requirement that  $A^\dagger$  is linear, for any  $g \in D(A^\dagger)$  with the unique representation  $g = g_1 + g_2$ ,  $g_1 \in R(A)$ ,  $g_2 \in R(A)^\perp$ ,  $A^\dagger g$  has to be  $\tilde{A}^{-1}g_1$ . The connection between least-squares solutions and the Moore-Penrose inverse is provided by this theorem (whose proof can be found in [29]):

**Theorem A.1.** *Let  $g \in D(A^\dagger)$ . Then,  $Af = g$  has a unique generalized solution, given by*

$$f^\dagger := A^\dagger g \quad (\text{A.7})$$

and  $f^\dagger \in N(A)^\perp$ .

**Proposition A.1.** *Let  $S$  and  $Q$  be the orthogonal projectors onto  $N(A)$  and  $\overline{R(A)}$ , respectively. Then  $R(A^\dagger) = N(A)^\perp$ , and the “four Moore-Penrose equations” hold:*

$$AA^\dagger A = A; \quad (\text{A.8})$$

$$A^\dagger AA^\dagger = A^\dagger; \quad (\text{A.9})$$

$$A^\dagger A = I - S; \quad (\text{A.10})$$

$$AA^\dagger = Q|_{D(A^\dagger)}. \quad (\text{A.11})$$

**Theorem A.2.** *Let  $R(A)$  be closed; then, there exists an  $m > 0$  such that  $\|Af\|_Y \geq m\|f\|_X$ ,  $\forall f \in N(A)^\perp$ .*

The following proposition binds the concept of well-posedness of the generalized inverse with the closeness of the range of the operator.

**Proposition A.2.**  *$A^\dagger$  is bounded if and only if  $R(A)$  is closed.*

Therefore, if  $R(A)$  is closed, by using  $A^\dagger$  it is possible to transform the original problem in a well-posed one, in which the solution exists, it is unique (for theorem A.1) and continuously depends from the data (for theorem A.2); on the other hand if  $R(A)$  is not closed the theorem A.2 does not hold and so the continuous dependence is lost. By defining  $A^*$  the adjoint operator of  $A$ , the generalized solution can be characterized by the normal equation:

**Theorem A.3.** *Let  $g \in D(A^\dagger)$ ; then  $f \in X$  is the least-squares solution of  $Af = g$  if and only if the normal equation*

$$A^*Af = A^*g \quad (\text{A.12})$$

*holds.*

It follows that  $A^\dagger g$  is the solution of  $A^*Af = A^*g$  of minimal norm, i.e.

$$A^\dagger = (A^*A)^\dagger A^*$$

It is possible to observe that, if  $g \notin D(A^\dagger)$ , no least-squares solution of  $Af = g$  exists: the introduction of the concept of generalized inverse solution does not always leads to solvable problems. In this setting, operator equations with a *compact* linear operator will play a crucial role.

## A.2 Compact linear operator

**Definition A.4.** *Let  $X, Y$  Hilbert spaces.  $A : X \rightarrow Y$  linear operator is compact if  $\forall \{f_n\}_{n \in \mathbb{N}} \in X$  bounded  $\{Af_n\}_{n \in \mathbb{N}}$  has a converging sub-succession in  $Y$ .*

**Theorem A.4.** *If  $A$  is a compact operator, then  $A$  is limited.*

If  $A : X \rightarrow X$ , two important theorems hold:

**Theorem A.5.** *Let  $A : X \rightarrow X$  be a compact operator where  $X$  is a Hilbert space of infinite dimension. Defining the spectrum of  $A$  as*

$$\sigma(A) = \{\lambda \in \mathbb{C} : \lambda I - A \text{ not invertible}\}$$

*then*

- (i)  $0 \in \sigma(A)$ ;
- (ii) each element of  $\sigma(A) \setminus \{0\}$  is eigenvalue of  $A$ ;
- (iii) if  $\sigma(A)$  is not finished then  $\sigma(A)$  is numerable and it has zero as the only accumulation point.

**Theorem A.6.** Let  $A : X \rightarrow X$  be a compact and self-adjoint operator. Let  $\forall \{\lambda_n\}_{n=1}^N$ , where  $N = \mathbb{N} \cup \{+\infty\}$ , be the succession of the eigenvalues of  $A$  different from 0, each one taken a number of times equal to multiplicity of the eigenvalues. Then there is an orthogonal system  $\forall \{u_n\}_{n=1}^N$  of eigenvectors such that

$$Af = \sum_{n=1}^N \lambda_n (f, u_n)_X u_n$$

which is named spectral representation of the operator.

This theorem is essentially the generalization of the singular value decomposition to infinite dimension. For a generic compact operator  $A : X \rightarrow Y$  the following singular values decomposition holds:

**Theorem A.7.** Let  $A : X \rightarrow Y$  be a compact operator. Then there are two orthogonal systems  $\{u_n\}_{n=1}^N$  and  $\{v_n\}_{n=1}^N$  and a succession  $\{\sigma_n\}_{n=1}^N$  where  $\{\sigma_n, u_n, v_n\}$  is the singular system of  $A$ , such that

$$(i) \quad \sigma > 0 \forall n \quad \sigma_n \rightarrow 0;$$

$$(ii) \quad Af = \sum_{n=1}^N \sigma_n (f, u_n)_X v_n, \quad f \in X;$$

$$(iii) \quad A^*g = \sum_{n=1}^N \sigma_n (g, v_n)_Y u_n, \quad g \in Y.$$

**Proposition A.3.** Let  $A$  be a compact operator and  $\{\sigma_n, u_n, v_n\}$  its singular system, then

$$f^\dagger = \sum_{n=1}^N \frac{(g, v_n)_Y}{\sigma_n} u_n \tag{A.13}$$

**Theorem A.8** (Picard Criterion).  $Af = g$  with  $A$  compact operator has a solution if

$$1. \quad g \in \overline{R(A)} = N(A^*)^\perp;$$

$$2. \quad \sum_{n=1}^N \frac{|(g, v_n)|^2}{\sigma_n^2} < +\infty$$

The condition 2. is the criterion for the existence of the generalized solution. It says that the best-approximate solution of  $Af = g$  exists only if the coefficients  $(g, v_n)$  decays fast enough relative to the singular values  $\sigma_n$ . The following theorem gives a rule to determine if a problem with a compact operator is well-posed

**Theorem A.9.** *A is a compact operator. Then  $R(A)$  is closed if and only if  $R(A)$  has finite dimension.*

**Remark A.10.** *Putting together theorem A.2 and theorem A.9 it is straightforward to obtain*

*A compact operator is invertible  $\Leftrightarrow R(A)$  has finite dimension.*

*So a problem  $Af = g$  with a compact operator  $A$  is well-posed if and only if  $R(A)$  has a finite dimension.*

So, basically, the idea that must be kept in mind if one wants to solve an inverse problem is to renounce to the exact solution and to resort to methods that look for an approximation of the solution by solving a modified version of the original problem that allows us to transform it in a well-posed problem. These family of methods are known as *regularization methods*.





# Appendix B

## Regularization methods

Regularization is the approximation of a ill-posed problem by a family of neighbouring well-posed problems, approximating the general solution  $f^\dagger = A^\dagger g$  of  $Af = g$ , for a specific  $g$  in a situation that *exact data* are not known precisely, but that only an approximation  $g_\delta$  - the noisy data - with  $\|g - g_\delta\|$  is available. In the ill-posed case,  $A^\dagger g_\delta$  is certainly not a good approximation of  $A^\dagger g$  due to the unboundeness of  $A^\dagger$  (even if it exists). The approximation needed is a  $f_\delta^\alpha$ , which depends continuously on the noisy data, in order to be computed in a stable way, and has also the property that as the noise level  $\delta$  decreases to zero and the regularization parameter  $\alpha$  is chosen appropriately, then  $f_\delta^\alpha$  tends to  $f^\dagger$ .

In order to talk about regularization, it makes more sense to look not only at the equation  $Af = g$  for a specific  $g$ , but to consider  $Af = g$  as a collection of equations for every  $g \in R(A)$  or for every  $g \in D(A^\dagger)$ . In this situation, it is possible to talk about *regularizing the solution operator*  $A^\dagger$ . Intuitively, a regularization of  $A^\dagger$  should be the replacement of the unbounded operator  $A^\dagger$  by a parameter-dependent family  $\{R^\alpha\}$  of continuous operators. As approximation of  $f^\dagger$ , we then take  $f_\delta^\alpha := R^\alpha g_\delta$ , which can then be computed in a stable way (since  $R^\alpha$  is assumed to be continuous). The parameter  $\alpha$  should be choose following some parameter choice rules that could be linked with either  $g_\delta$  or some a-priori information about the exact data  $g$ . Therefore, let define regularization operators, as in [22] for the whole collection of equations

$$Af = g, \quad g \in D(A^\dagger) \tag{B.1}$$

but parameter choice rules for a specific equation out of this collection. These consideration lead to the following definition

**Definition B.1.** Let  $A : X \rightarrow Y$  be a bounded linear operator between Hilbert spaces. For every  $\alpha > 0$  let  $R^\alpha : Y \rightarrow X$  be a continuous (not necessary linear) operator. The one-parameter family  $\{R^\alpha\}$  is called a regularization operator if, named  $P$  a projection of  $Y$  onto  $\overline{R(A)}$ , then

$$\lim_{\alpha \rightarrow 0} \|R^\alpha - f^\dagger\|_X = 0$$

for each  $g$  such that  $Pg \in R(A)$ . Furthermore a pair  $(R^\alpha, \alpha)$  is called a (convergent) regularization method.

The following questions now arise:

- (i) How to build up regularization operators?
- (ii) How to choose the parameter in order to have a convergent method?
- (iii) How to do that in an *optimal* way?

The previous questions can be partially answered, for linear operators:

**Proposition B.1.** Let  $R^\alpha$  be a continuous operator  $\forall \alpha > 0$ . Then the family  $\{R^\alpha\}$  is a regularization for  $A^\dagger$  if

$$R^\alpha \rightarrow A^\dagger \text{ pointwise on } D(A^\dagger) \text{ as } \alpha \rightarrow 0. \quad (\text{B.2})$$

In this case, there exists, for every  $g \in D(A^\dagger)$ , a parameter choice rule such that  $(R^\alpha, \alpha)$  is a convergent regularization method for solving  $Af = g$ .

**Proposition B.2.** Let  $\{R^\alpha\}$  be a linear regularization,  $f^\alpha$  defined as  $f^\alpha := R^\alpha g$ ,  $\forall g \in Y$ . Then

$$\{f^\alpha\} \text{ converges to } A^\dagger g \text{ as } \alpha \rightarrow 0 \text{ for } g \in D(A^\dagger)$$

and if

$$\sup\{\|AR^\alpha\| : \alpha > 0\} < +\infty$$

then

$$\|x^\alpha\| \rightarrow +\infty \text{ as } \alpha \rightarrow 0 \text{ for } g \notin D(A^\dagger)$$

Proposition B.1 shows that, if equation (B.2) holds, there exists a choice rule for the convergence of the regularization method, that can be characterized as follows:

**Proposition B.3.** *Let  $\{R^\alpha\}$  be a linear regularization; for every  $g \in D(A^\dagger)$ , let the choice rule  $\alpha$  be a function of the noise level  $\delta$ ,  $\alpha : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ . Then  $(R^\alpha, \alpha)$  is a convergent regularization method if and only if*

$$\lim_{\delta \rightarrow 0} \alpha(\delta) = 0 \quad (\text{B.3})$$

and

$$\lim_{\delta \rightarrow 0} \delta \|R^{\alpha(\delta)}\| = 0 \quad (\text{B.4})$$

hold.

These propositions suggest that for each noise level  $\delta$  an optimal  $\alpha_{opt}$  exists and this value should minimize the following quantity:

$$\|R^\alpha g_\delta - f^\dagger\|_X \leq \|R^\alpha A f - f^\dagger\|_X + \|R^\alpha \delta\|_X.$$

The first term of the right-hand side represents the approximation error due to the use of  $R^\alpha$  instead of  $A^\dagger$ . By definition of regularization method, this term goes to zero as  $\alpha$  goes to zero. The second term is the *noise propagation error on the data* and it can be considered as a measure of the error on the regularized solution  $f^\alpha$  due to the noise on the data: fixed the noise level this term goes to infinity as  $\alpha$  goes to zero. These antithetic behaviors of the two errors suggest the crucial importance of the right choice of  $\alpha$  and introduce the typical trend of the error in the regularization method known as semi-convergence.

By summarizing, the regularization algorithms generate a family of approximate solutions which converge to the true object in the case of noise-free data and has the semi-convergence property in the case of noisy data.

There are many regularization methods, one of the most known is certainly the *Tikhonov* method. The main idea of this method is to modify the functional to be minimized: from  $\|A f - g\|_Y$  to

$$\|A f - g\|_Y^2 + \alpha \|f\|_X^2 \quad (\text{B.5})$$

with  $\alpha$  the appropriate regularization parameter. It is easy to show that the problem of minimizing equation (B.5) is equivalent to solve

$$(A^* A f + \alpha I) f^\alpha = A^* g. \quad (\text{B.6})$$

So the Tikhonov method is a regularization method with

$$R^\alpha = (A^* A f + \alpha I) A^*. \quad (\text{B.7})$$

The following theorem allows to define a regularization solution as a filtered version of the generalized solution by using the singular values decomposition  $(\sigma_n, u_n, v_n)$ :

**Theorem B.1.**

$$R^\alpha g = \sum_{n=1}^N \frac{w_n^\alpha}{\sigma_n} (g, v_n) u_n \quad (\text{B.8})$$

where  $\{w_n^\alpha\}_{n=1}^N$  are called spectral windows and they obey to the following properties:

1.  $\forall \alpha > 0, 0 \leq w_n^\alpha \leq 1$ ;
2.  $\forall n \in \mathbb{N}, \lim_{\alpha \rightarrow 0} w_n^\alpha = 1$ ;
3.  $\forall \alpha > 0, \exists c(\alpha)$  such that  $\frac{w_n^\alpha}{\sigma_n} \leq c(\alpha), \forall n \in \mathbb{N}$ .

In the case of the Tikhonov method the regularized solution can be written using the singular values decomposition as

$$R^\alpha g = \sum_{n=1}^N \frac{\sigma_n}{\sigma_n^2 + \alpha} (g, v_n) u_n, \quad (\text{B.9})$$

so the filters are

$$w_n^\alpha = \frac{\sigma_n^2}{\sigma_n^2 + \alpha}. \quad (\text{B.10})$$

Here we briefly summarize the most used regularized iterative methods. For an exhaustive treatment, see [22].

## Landweber

The Landweber method is an iterative regularization method for the approximation of  $f^\dagger$ , based on a transformation of the normal equation

$$A^* A f = A^* g, \quad (\text{B.11})$$

whose unique solution is  $f^\dagger$  (according to theorem B.1), into an equivalent fixed point equations like

$$f = f + A^*(g - A f). \quad (\text{B.12})$$

The Landweber iteration can be supplied with an initial guess  $f^{(0)}$ , and the iteration computes further approximation  $\{f^{(n)}\}$  recursively

$$f^{(n+1)} = f^{(n)} + \tau A^*(g - Af^{(n)}) \quad (\text{B.13})$$

where  $\tau$  is the *relaxation parameter*. By introducing the nonlinear operator

$$\begin{aligned} T : X &\rightarrow Y \\ f &\mapsto f + \tau A^*(g - Af) \end{aligned} \quad (\text{B.14})$$

it is clear that the set of solution of the least-square problem  $\|Af - g\| = \text{minimum}$  (that are solutions of eq. (B.11) if  $g \in R(A) \oplus R(A)^\perp$ , thanks to theorem A.1) coincides with the set of the fixed points of the operator  $T$ . The method of Landweber is thus the method of successive approximations for determining the fixed point of the operator  $T$ , that is  $T(f^{(n)}) = f^{(n+1)}$ . There exists different way to find an optimal choice for the regularization parameter  $\tau$  and different results on the convergence of the sequence  $\{f^{(n)}\}$  to a fixed point of  $T$ , that we do not present here, but an exhaustive description can be found here [59, 25, 118]. Furthermore, it is possible to show that the Landweber method is effectively a regularization method, in the case in which  $A$  is an operator that admits singular value decomposition:

$$f^{(n)} = \sum_{k=1}^N w_n^k \frac{(g, u_k)}{\sigma_k} v_k + \sum_{k=1}^N (1 - \tau \sigma_k^2)^n (f^{(0)}, v_k) + P_N f^{(0)} \quad (\text{B.15})$$

where  $w_n^k = 1 - (1 - \tau \sigma_k^2)^n$  and  $P_N$  is the projection onto  $N(A)$ . So if  $f^{(0)} = 0$ ,  $f^{(n)}$  is a filtered version of  $f^\dagger$  with filters  $w_n^k$ , respecting conditions of theorem B.1 if  $\tau \in \left[0, \frac{2}{\sigma_1}\right]$ . Therefore Landweber method is a regularization method and the number of iteration  $n$  plays the role of the regularization parameter.

Actually, in the case of noisy data one has to solve

$$Af + \delta = g \quad (\text{B.16})$$

In that situation the method, in general, does not converge. In such a case, however, the method has a peculiar property of the regularization methods, the *semiconvergence*. Let  $\|f^{(n)} - f^\dagger\|_X$  be the restoration error; it is the addition of two terms, the approximation error and the noise propagation error, that behave oppositely with respect to the increasing of the regularization parameter  $n$ . Thus the restoration error is the addition of a decreasing function with an increasing function, so, as  $n$  increases, the restoration error is decreasing first and increasing afterwards and

hence it has a minimum for a particular value of  $n$ ; this value represents the optimal number of iterations. At such a value corresponds the regularized solution, that is the best approximation of  $f^\dagger$ . In the case of noisy real data it is possible to calculate such an optimal  $n$  by using methods known as *stopping rules*, the most important one is the *discrepancy principle*. In this method, if a precise estimate  $\delta$  of the energy of the noise is known, the optimal estimate for  $n$  is the value of  $n = n(\delta)$  such that the discrepancy of the corresponding regularized solution is just equal to  $\epsilon^2$ , i.e.

$$\|Af^{(n)} - g\|_Y = \epsilon. \quad (\text{B.17})$$

The simplest stopping rule can thus be obtained by an extension of the discrepancy principle: let  $b \geq 1$  given, then if there exists a value  $\tilde{n}$  such that

$$\|Af^{(\tilde{n})} - g\|_Y > b\epsilon \quad (\text{B.18})$$

and

$$\|Af^{(\tilde{n}+1)} - g\|_Y \leq b\epsilon, \quad (\text{B.19})$$

then  $\tilde{n}$  can be taken as an estimate of the optimal choice for  $n$  and  $f^{(\tilde{n})}$  can be taken as an estimate of the best approximation of  $f^\dagger$  provided by the Landweber method. It is possible to prove that if  $b > 1$  and if the noise on the data tends to zero, so that  $\epsilon \rightarrow 0$  and  $\|f^{(\tilde{n})} - f^\dagger\| \rightarrow 0$ .

Another interesting feature of Landweber method is that it can be used, after some appropriate modifications, for solving least-squares problems including *a priori* information on the behavior of the solution. The most convenient way for expressing *a priori* information is to state that the object must belong to some given subset  $C$  of the space of all possible objects and a common feature of the above such a set is to impose it to be closed and convex. Then, if it is known that the solution of the restoration problem must belong to some closed and convex  $C$ , it is quite natural to look, among all the possible objects satisfying this condition, for that or those which minimize the discrepancy, i.e. the solutions of the following constrained problem

$$\|Af - g\|_Y = \text{minimum}, \quad f \in C. \quad (\text{B.20})$$

The following theorem holds

**Theorem B.2.** *Named  $J = \overline{A(C)}$ , if  $P_J g \in A(C)$ , then  $\{f^{(n)}\}_{n=1}^\infty$  weakly converges to the solution of (B.20)*

Therefore, the modified Landweber method, also called the projected Landweber method is the following

$$f_{|C}^{(n+1)} = P_C \left[ f_{|C}^{(n)} + \tau(A^*g - A^*A f_{|C}^{(n)}) \right] \quad (\text{B.21})$$

with  $\tau \in \left[0, \frac{1}{\sigma_1}\right]$ . At each step the method consists of a Landweber iteration followed by a projection onto the convex set  $C$ .

**Remark B.3.** *One of the most used projection is onto the set of non-negative function, which is always true if the object  $f$  we want to reconstruct is an image, coming from some medical imaging techniques.*

Note that if  $g$  is noisy, no result of semiconvergence has been proved.

### Conjugate Gradient Descent

Although it is easy to apply, the Landweber method often displays a slow rate of convergence, as all the steepest descent methods. The conjugate gradient (CG) method was developed in order to obtain better rates of convergence; such a method is a modification of the steepest descent method and it is known as one of the most powerful algorithms for the solution of the selfadjoint, positive (semi)definite well-posed linear equations. Given the normal equation

$$A^*A f = A^*g \quad (\text{B.22})$$

of an ill-posed problem, the main idea of this method is to minimize the discrepancy functional

$$\epsilon^2(f; g) = \|A f - g\|_Y^2 \quad (\text{B.23})$$

which is equivalent to the minimization of the functional

$$\eta(f; g) = \frac{1}{2}(A^*A f, f) - (A^*A g, g). \quad (\text{B.24})$$

Indeed, by using the properties of the scalar product, the following equation holds:

$$\epsilon^2(f; g) = 2\eta(f; g) + \|g\|^2$$

and it is easy to show [9] that

$$\nabla_f \eta(f; g) = \frac{1}{2} \nabla_f \epsilon^2(f; g) = A^*A f - A^*g, \quad (\text{B.25})$$

so that equivalent formulation coincides with the negative residual associated with the least-squares equation  $r = A * g - A * Af$ . Thus, minimizing the discrepancy functional is equivalent to solve the normal equation (B.22). Given an approximation  $f^{(n)}$  for the object, in a neighborhood of this point the functional  $\eta(f; g)$  decreases most rapidly in the direction of the negative gradient, i.e. the direction of the negative residual  $r$ . If  $r^{(n)} = A * g - A * Af^{(n)}$ , thus the Landweber method can be written as follows

$$f^{(n+1)} = f^{(n)} + \tau r^{(n)}. \quad (\text{B.26})$$

and so at each iteration the solution is modified in the negative gradient direction. The main difference between the CG and gradient methods like Landweber is the choice of the descent direction in order to minimize  $\eta$ : in CG the *conjugate gradient direction* is chosen, instead of the gradient direction. For a detailed description of the method see [22, 90].

## Maximum Likelihood

One of the classical approaches to parameter estimation in statistics, that provides also estimates of the solution for statistical inverse problems, is Maximum likelihood (ML). Once the statistical distribution of data is known, its application is straightforward and is widely used in many different fields [104]. Furthermore, ML estimates can be effectively constrained to be non negative. In the case of Poisson noise, the constrained ML approach leads to the Expectation Maximization (EM) algorithm [18], while in the case of white Gaussian noise, the constrained ML approach leads to the Iterative Space Reconstruction Algorithm (ISRA) [16]. Given a forward problem in the form  $g = Af$ , a ML approach for estimating  $f$  given  $g$  is based on the assumption that the data  $g$  is an observed value of random vector  $G$  with mean  $Af$ . The density function of  $G$  is given by the joint probability density function  $p_\eta(g, Af)$  where  $\eta$  indicates a generic random process. When this density is thought of as a function of  $f$  given  $g$  we call it the likelihood and we write

$$\mathcal{L}_g(f) = p_\eta(g, Af). \quad (\text{B.27})$$

Once the matrix  $A$  and the data  $g$  are given, ML obtains the solution of  $f$  when the likelihood reaches its maximum value. However, in the majority of cases physical motivations regarding the nature of the problem imply that the components of parameter  $x$  has to be non-negative. Consequently, the ML estimator is constrained as

$$\hat{f} = \arg \max_{x \in C} \mathcal{L}_g(f) \quad (\text{B.28})$$



where  $C$  is the non-negative subset of possible solution. Usually, it is more convenient to minimize the negative logarithm of the likelihood instead of maximizing it. The constrained ML problem is therefore equivalent to

$$\hat{f} = \arg \min_{x \in C} L_g(f) \quad (\text{B.29})$$

where  $L_g(f) = -\log(\mathcal{L}_g(f))$ . Indeed, when both the negative logarithm of the likelihood function and the constraint are convex, the necessary and sufficient conditions for  $f$  to be the constrained ML estimator are the Karush-Kuhn-Tucker (KKT) conditions [12], which in this particular case take the form

$$f \nabla L_g(f) = 0, \quad x \geq 0 \quad (\text{B.30})$$

Splitting the gradient into the positive part  $V(x)$  and the negative part  $U(x)$  transforms (B.30) into a fixed point equation [61]; then applying the successive approximation method leads to the multiplicative iterative algorithm

$$f^{(n+1)} = f^{(n)} \frac{U(f^{(n)})}{V(f^{(n)})}. \quad (\text{B.31})$$

When  $\eta$  is a vector of independent and identically distributed Gaussian variables (we are thinking at a discrete case now) with mean given by  $Af$  and variance equal to  $\sigma^2$ , the constrained ML problem (B.28) is equivalent to the minimization of the least-squares function

$$\|Af - g\|^2, \quad (\text{B.32})$$

under the non-negativity constraint. Consequently, KKT conditions (B.30) lead to the algorithm

$$f^{(n+1)} = f^{(n)} \frac{A^T g}{A^T A f^{(n)}} \quad (\text{B.33})$$

which is known as Iterative Space Reconstruction Algorithm (ISRA) and it is convergent to the constrained minimum of (B.32).

Analogously, when  $\eta$  is a vector of independent and identically distributed Poisson variables with parameter given by  $Af$ , the negative logarithm of the likelihood is equivalent to the Kullback Leibler divergence

$$g \log \frac{g}{Af} + Af - g. \quad (\text{B.34})$$

Hence, the constrained ML problem is equivalent to the minimization of such a divergence under the non-negativity constraint. In the Poisson case the KKT conditions lead to the following iterative algorithm

$$f^{(n+1)} = \frac{f^{(n)}}{A^T \mathbf{1}} A^T \frac{g}{A f^{(n)}}, \quad (\text{B.35})$$

which is known as Expectation Maximization.

The noise corrupting the data  $g$  makes the constrained ML solution not physically acceptable. In this case one can get an estimation of the unknown signal exploiting *regularization*, but the classical definition of regularization does not explicitly use statistical concepts, because takes place in the framework of functional analysis. In order to provide a statistical definition of regularization and stopping rule we suggest the paper [8].

# Appendix C

## Proof of Theorem 5.2

In order to prove Theorem 5.2 we first need to search for the matrices  $M_x \in M_3(\mathbb{R})$

$$M_x = \begin{pmatrix} -(x_a + x_b) & x_c & 0 \\ x_b & -(x_c + x_d) & x_e \\ 0 & x_d & -x_e \end{pmatrix},$$

where  $x_a, x_b, x_c, x_d, x_e \in \mathbb{R}_+^*$ , such that  $F_x = F$ , with  $F = Q/P$  and  $F_x = Q_x/P_x$ , where  $P$  is the characteristic polynomial of  $M$ ,  $P_x$  is the characteristic polynomial of  $M_x$  and  $Q, Q_x$  are defined by

$$Q(X) = \alpha^T \text{adj}(X - M)e_1, \quad Q_x(X) = \alpha^T \text{adj}(X - M_x)e_1.$$

That is

$$P(X) = X^3 + (a + b + c + d + e)X^2 + (ac + ad + ae + bd + be + ce)X + ace$$

$$Q(X) = VX^2 + (V(c + d + e) + (1 - V)b)X + (Vce + (1 - V)(d + e)b)$$

$$P_x(X) = X^3 + (x_a + x_b + x_c + x_d + x_e)X^2 +$$

$$+ (x_ax_c + x_ax_d + x_ax_e + x_bx_d + x_bx_e + x_cx_e)X + x_ax_cx_e$$

$$Q_x(X) = VX^2 + (V(x_c + x_d + x_e) + (1 - V)x_b)X + (Vx_cx_e + (1 - V)(x_d + x_e)x_b)$$

We first begin with some preliminary results about the polynomials  $P$  and  $Q$  and the rational fraction  $F$ . The discriminant of  $Q$  is given by

$$\begin{aligned} \Delta_Q &= (V(c + d + e) + (1 - V)b)^2 - 4V(Vce + (1 - V)(d + e)b) = \\ &= ((-c + d + e)V - (1 - V)b)^2 + 4cdV^2, \end{aligned}$$

thus  $0 < \Delta_Q < V(c+d+e) + (1-V)b$ . Hence  $Q$  has two distinct real negative roots

$$\mu_q = -\frac{(c+d+e)V + (1-V)b + (-1)^q \sqrt{\Delta_Q}}{2V}, \quad q \in \{1, 2\},$$

Hence, the rational fraction  $F$  can be written as

$$F(X) = \frac{V(X - \mu_1)(X - \mu_2)}{(X - \lambda_1)(X - \lambda_2)(X - \lambda_3)},$$

where  $\lambda_1 < \lambda_2 < \lambda_3 < 0$  are the eigenvalues of  $M$ . In the following, we will need to know the irreducible form of the rational fraction  $F$ , that is, the number of roots common to  $P$  and  $Q$ . For the moment, we have three possibilities.

1.  $P$  and  $Q$  have no common root, *i.e.* they are coprime, so that  $F$  is irreducible,
2.  $P$  and  $Q$  have one common root  $\lambda_p = \mu_q, p \in \{1, 2, 3\}, q \in \{1, 2\}$ , hence

$$F = \frac{V(X - \mu_{q'})}{(X - \lambda_{p'})(X - \lambda_{p''})}, \quad \text{where } p', p'' \in \{1, 2, 3\} \setminus \{p\}, q' \in \{1, 2\} \setminus \{q\},$$

3.  $P$  and  $Q$  have two common roots, *i.e.*  $Q|P, \mu_1 = \lambda_p, \mu_2 = \lambda_{p'}, p, p' \in \{1, 2, 3\}, p \neq p'$ , hence

$$F = \frac{V}{X - \lambda_{p''}}, \quad \text{where } p'' \in \{1, 2, 3\} \setminus \{p, p'\}.$$

$P$  and  $Q$  are not coprime if and only if their resultant  $\text{res}(P, Q)$  is 0. The resultant  $\text{res}(P, Q)$  of  $P$  and  $Q$  is given by

$$\begin{aligned} \text{res}(P, Q) = & -b^2cd \left( V(1-V)^2a^2 - (1-V)(1-2V)(V(c+d+e) + (1-V)b)a + \right. \\ & \left. + (1-2V)^2(Vce + (1-V)(d+e)b) \right). \end{aligned}$$

In particular, we can see that if  $V = 1/2$ ,  $\text{res}(P, Q) = -a^2b^2cdV(1-V)^2 \neq 0$ , otherwise, it can be easily remarked that

$$\text{res}(P, Q) = -b^2cd(1-2V)^2Q \left( -\frac{1-V}{1-2V}a \right).$$

Hence, for  $V \neq 1/2$ ,  $\text{res}(P, Q) = 0$  if and only if  $-\frac{1-V}{1-2V}a$  is a root of  $Q$ . In particular, if  $V > 1/2$ , then  $-\frac{1-V}{1-2V}a > 0$ . However, the roots of  $Q$  are negative, thus  $\text{res}(P, Q) \neq 0$ . We recall that the same result holds for  $V = 1/2$ , thus we have the following lemma

**Lemma C.1.** *If  $V \geq 1/2$ , then  $P$  and  $Q$  are coprime so that  $F$  is irreducible.*

If now  $V < 1/2$  and  $-\frac{1-V}{1-2V}a$  is a root of  $Q$ , that is

$$a = \frac{1-2V}{2V(1-V)} \left( (c+d+e)V + (1-V)b + (-1)^q \sqrt{\Delta_Q} \right), \quad q \in \{1, 2\},$$

then  $P(X) = (X - \mu_q)\check{P}(X)$ , where

$$\check{P}(X) = X^2 + \frac{(2-3V)(c+d+e) + (1-V)b - (-1)^q \sqrt{\Delta_Q}}{2(1-V)} X + \frac{1-2V}{1-V} ce,$$

that is

$$\check{P}(X) = X^2 + \left( -\frac{V}{1-2V}a + b - \frac{1-2V}{1-V}(c+d+e) \right) X + \frac{1-2V}{1-V} ce.$$

Since  $Q(X) = (X - \mu_q)\check{Q}(X)$ , with

$$\check{Q}(X) = V \left( X + \frac{(c+d+e)V + (1-V)b - (-1)^q \sqrt{\Delta_Q}}{2V} \right),$$

that is

$$\check{Q}(X) = V \left( X - \frac{1-V}{1-2V}a + \frac{1-V}{V}b + c + d + e \right),$$

we have

$$F(X) = \frac{V \left( X - \frac{1-V}{1-2V}a + \frac{1-V}{V}b + c + d + e \right)}{X^2 + \left( -\frac{V}{1-2V}a + b - \frac{1-2V}{1-V}(c+d+e) \right) X + \frac{1-2V}{1-V} ce}.$$

$P$  and  $Q$  have two common roots if and only if  $\check{P}(\mu_{q'}) = 0$  where  $q' \in \{1, 2\} \setminus \{q\}$ . However

$$\begin{aligned} \check{P}(\mu_{q'}) &= \frac{1-2V}{V^2} \left( bV + (1-V)d - \frac{V(1-V)}{1-2V}e \right) = \\ &= -\frac{1-2V}{2V^2} ((-c+d+e)V - (1-V)b + (-1)^q \sqrt{\Delta_Q}). \end{aligned}$$

Since  $\Delta_Q = ((-c+d+e)V - (1-V)b)^2 + 4cdV^2$ , we have  $\sqrt{\Delta_Q} > |(-c+d+e)V - (1-V)b|$ , hence  $\check{P}(\mu_{q'}) \neq 0$ . We thus have the following lemma

**Lemma C.2.** *If  $V < 1/2$ ,  $P$  and  $Q$  have at most one common root. This happens if and only if  $-\frac{1-V}{1-2V}a$  is a root of  $Q$ , that is*

$$a = \frac{1-2V}{2V(1-V)} \left( (c+d+e)V + (1-V)b + (-1)^q \sqrt{\Delta_Q} \right), \quad q \in \{1, 2\},$$

where

$$\begin{aligned} \Delta_Q &= (V(c+d+e) + (1-V)b)^2 - 4V(Vbc + (1-V)(d+e)b), \\ &= ((-c+d+e)V - (1-V)b)^2 + 4cdV^2. \end{aligned}$$

In this case, we have  $F = \check{Q}/\check{P}$ , where  $\check{P}$  and  $\check{Q}$  are coprime and

$$\begin{aligned} \check{P}(X) &= X^2 + \left( -\frac{V}{1-2V}a + b - \frac{1-2V}{1-V}(c+d+e) \right) X + \frac{1-2V}{1-V}ce, \\ \check{Q}(X) &= V \left( X - \frac{1-V}{1-2V}a + \frac{1-V}{V}b + c + d + e \right). \end{aligned}$$

We are now ready to seek the solutions of the inverse problem and study the identifiability of the system. We first treat the case where  $P$  and  $Q$  are coprime and finish the study with the case where they are not.

### Coprimality of $P$ and $Q$

We suppose here that  $P$  and  $Q$  are coprime. Since  $F$  is irreducible and  $\deg P_x = \deg P = 3$  and  $\deg Q_x = \deg Q = 2$ ,  $F_x = F$  only if the rational fraction  $F_x$  is irreducible too. In addition, since the leading coefficients of  $P_x$  and  $P$  are identical, as well as those of  $Q_x$  and  $Q$ ,  $F_x = F$  if and only if  $P_x = P$  and  $Q_x = Q$ , that is

$$\begin{aligned} V(x_c + x_d + x_e) + (1-V)x_b &= V(c+d+e) + (1-V)b, \\ Vx_cx_e + (1-V)(x_d + x_e)x_b &= Vce + (1-V)(d+e)b, \\ x_a + x_b + x_c + x_d + x_e &= a + b + c + d + e, \\ (x_c + x_d + x_e)x_a + (x_d + x_e)x_b + x_cx_e &= (c+d+e)a + (d+e)b + ce, \\ x_ax_cx_e &= ace. \end{aligned}$$

Suppose first that  $V = 1/2$ , then we have

$$x_b + x_c + x_d + x_e = b + c + d + e, \quad (\text{C.1a})$$

$$(x_d + x_e)x_b + x_c x_e = (d + e)b + ce, \quad (\text{C.1b})$$

$$x_a + x_b + x_c + x_d + x_e = a + b + c + d + e, \quad (\text{C.1c})$$

$$(x_c + x_d + x_e)x_a + (x_d + x_e)x_b + x_c x_e = (c + d + e)a + (d + e)b + ce, \quad (\text{C.1d})$$

$$x_a x_c x_e = ace. \quad (\text{C.1e})$$

Subtracting (C.1a) and (C.1c), we get  $x_a = a$ , so that the previous system (C.1) is equivalent to

$$x_a = a, \quad (\text{C.2a})$$

$$(x_d + x_e)x_b = (d + e)b, \quad (\text{C.2b})$$

$$x_b + x_c + x_d + x_e = b + c + d + e, \quad (\text{C.2c})$$

$$(x_c + x_d + x_e)x_a + (x_d + x_e)x_b = (c + d + e)a + (d + e)b, \quad (\text{C.2d})$$

$$x_c x_e = ce. \quad (\text{C.2e})$$

Doing (C.2d)–(C.2b) to replace (C.2d), we get

$$x_a = a, \quad (\text{C.3a})$$

$$(x_d + x_e)x_b = (d + e)b, \quad (\text{C.3b})$$

$$x_b + x_c + x_d + x_e = b + c + d + e, \quad (\text{C.3c})$$

$$x_c + x_d + x_e = c + d + e, \quad (\text{C.3d})$$

$$x_c x_e = ce. \quad (\text{C.3e})$$

Thus, subtracting (C.3c) and (C.3d), we get  $x_b = b$  so that the previous system (C.3) is equivalent to

$$x_a = a, \quad (\text{C.4a})$$

$$x_b = b, \quad (\text{C.4b})$$

$$x_d + x_e = d + e, \quad (\text{C.4c})$$

$$x_c + x_d + x_e = c + d + e, \quad (\text{C.4d})$$

$$x_c x_e = ce. \quad (\text{C.4e})$$

Then, subtracting (C.4d) and (C.4c), we get  $x_c = c$ , the equation (C.4e) then gives  $x_e = e$  and finally (C.4c) gives  $x_d = d$ , so that  $M_x = M$ . Suppose now that  $V \neq 1/2$ .

We have

$$V(x_c + x_d + x_e) + (1 - V)x_b = V(c + d + e) + (1 - V)b, \quad (\text{C.5a})$$

$$Vx_c x_e + (1 - V)(x_d + x_e)x_b = Vce + (1 - V)(d + e)b, \quad (\text{C.5b})$$

$$x_a + x_b + x_c + x_d + x_e = a + b + c + d + e, \quad (\text{C.5c})$$

$$(x_c + x_d + x_e)x_a + (x_d + x_e)x_b + x_c x_e = (c + d + e)a + (d + e)b + ce, \quad (\text{C.5d})$$

$$x_a x_c x_e = ace. \quad (\text{C.5e})$$

Doing (C.5a)–(C.5c) to replace (C.5a), followed by (1 – 2V)(C.5c)–(C.5a) to replace (C.5c), then by (1 – V)(C.5d)–(C.5b) to replace (C.5d), the system (C.5) is equivalent to

$$-Vx_a + (1 - 2V)x_b = -Va + (1 - 2V)b, \quad (\text{C.6a})$$

$$Vx_c x_e + (1 - V)(x_d + x_e)x_b = Vce + (1 - V)(d + e)b, \quad (\text{C.6b})$$

$$(x_c + x_d + x_e) + \frac{1 - V}{1 - 2V}x_a = (c + d + e) + \frac{1 - V}{1 - 2V}a, \quad (\text{C.6c})$$

$$(x_c + x_d + x_e)\frac{1 - V}{1 - 2V}x_a + x_c x_e = (c + d + e)\frac{1 - V}{1 - 2V}a + ce, \quad (\text{C.6d})$$

$$x_c x_e \frac{1 - V}{1 - 2V}x_a = ce \frac{1 - V}{1 - 2V}a. \quad (\text{C.6e})$$

The equations (C.6c), (C.6d), (C.6e) are verified if and only if  $\frac{x_c + x_d + x_e \pm \sqrt{(x_c + x_d + x_e)^2 - 4x_c x_e}}{2}$  and  $\frac{1 - V}{1 - 2V}x_a$  are the roots of the polynomial  $R$  of degree 3

$$R(X) = \left( X - \frac{1 - V}{1 - 2V}a \right) \check{R}(X), \quad \text{where } \check{R}(X) = X^2 - (c + d + e)X + ce.$$

Hence, the system (C.6) is equivalent to the one obtained by including the equation  $R\left(\frac{1 - V}{1 - 2V}x_a\right) = 0$ , although the new system is redundant considering equations (C.6c), (C.6d), (C.6e) and the newly included one. The system (C.6) is then equivalent to

$$-Vx_a + (1 - 2V)x_b = -Va + (1 - 2V)b, \quad (\text{C.7a})$$

$$Vx_c x_e + (1 - V)(x_d + x_e)x_b = Vce + (1 - V)(d + e)b, \quad (\text{C.7b})$$

$$(x_c + x_d + x_e) + \frac{1 - V}{1 - 2V}x_a = (c + d + e) + \frac{1 - V}{1 - 2V}a, \quad (\text{C.7c})$$

$$(x_c + x_d + x_e)\frac{1 - V}{1 - 2V}x_a + x_c x_e = (c + d + e)\frac{1 - V}{1 - 2V}a + ce, \quad (\text{C.7d})$$

$$x_c x_e \frac{1 - V}{1 - 2V}x_a = ce \frac{1 - V}{1 - 2V}a, \quad (\text{C.7e})$$

$$R\left(\frac{1 - V}{1 - 2V}x_a\right) = 0. \quad (\text{C.7f})$$



The solution  $x_a = a$  to the equation (C.7f) leads to  $x_b = b, x_c = c, x_d = d, x_e = e$ , that is  $M_x = M$ . In addition, since  $0 < (-c + d + e)^2 + 4cd = (c + d + e)^2 - 4ce < (c + d + e)^2$ , the roots of  $\check{R}$  are distinct, real and positive. However, for  $V > 1/2$  and  $x_a > 0$ , we have  $\frac{1-V}{1-2V}x_a < 0$ , then  $\check{R}\left(\frac{1-V}{1-2V}x_a\right) \neq 0$  and the unique solution of (C.7f) is  $x_a = a$ , so that the unique solution of the inverse problem is  $M_x = M$ . We recall that we had the same result for  $V = 1/2$ . In addition, we also recall that according to lemma C.1,  $P$  and  $Q$  are always coprime for  $V \geq 1/2$ , thus, we have the following lemma

**Lemma C.3.** *If  $V \geq 1/2$ , the inverse problem has a unique solution.*

Suppose now that  $V < 1/2$ , then (C.7f) has at least two distinct solutions since the roots of  $\check{R}$  are distinct and at most three distinct solutions. Consider for the moment the solutions of (C.7) in  $\mathbb{R}^5$  regardless of their sign and denote by  $x_a^\pm = \frac{1-2V}{2(1-V)} \left( c + d + e \pm \sqrt{(c + d + e)^2 - 4ce} \right)$  the solutions of  $\check{R}\left(\frac{1-V}{1-2V}x_a\right) = 0$ . These solutions lead to two solutions  $(x_a^\pm, x_b^\pm, x_c^\pm, x_d^\pm, x_e^\pm)$  to (C.7), provided  $b + \frac{V}{1-2V}(x_a^\pm - a) \neq 0$  and  $(1 - 2V)bc - (1 - V)(x_a^\pm - a) \left( b - \frac{V}{1-2V}a \right) \neq 0$ , they are given by

$$x_a^\pm = \frac{1 - 2V}{2(1 - V)} \left( c + d + e \pm \sqrt{(c + d + e)^2 - 4ce} \right), \quad (\text{C.8a})$$

$$x_b^\pm = b + \frac{V}{1 - 2V}(x_a^\pm - a), \quad (\text{C.8b})$$

$$x_c^\pm = \frac{(1 - 2V)bc - (1 - V)(x_a^\pm - a) \left( b - \frac{V}{1-2V}a \right)}{(1 - 2V)x_b^\pm}, \quad (\text{C.8c})$$

$$x_d^\pm = \frac{bQ \left( -\frac{1-V}{1-2V}a \right) \left( \frac{1-2V}{1-V}c - x_a^\pm \right)}{(1 - 2V)x_b^{\pm 2}x_c^\pm}, \quad (\text{C.8d})$$

$$x_e^\pm = a \left( \frac{1 - V}{1 - 2V} \right)^2 \frac{x_a^\mp}{x_c^\pm}. \quad (\text{C.8e})$$

Remark that if both solutions  $(x_a^+, x_b^+, x_c^+, x_d^+, x_e^+)$  and  $(x_a^-, x_b^-, x_c^-, x_d^-, x_e^-)$  exist, we have

$$x_c^+ x_d^+ x_c^- x_d^- = -16cde^2(1 - V)^2 \frac{\Xi^2}{\Upsilon^2},$$

where

$$\begin{aligned} \Xi = & a(c + d + e)(1 - V)(1 - 2V)^2 - b(d + e)(1 - V)(1 - 2V)^2 + ab(1 - 2V)(1 - V)^2 + \\ & - ceV(1 - 2V)^2 - a^2V(1 - V)^2, \end{aligned}$$

and

$$\Upsilon = ((c + d + e)V(1 - 2V) + 2b(1 - V)(1 - 2V) - 2aV(1 - V))^2 - V^2(1 - 2V)^2 ((c + d + e)^2 - 4ce)^2.$$

Hence  $x_c^+ x_d^+ x_c^- x_d^- \leq 0$  and we have  $x_c^+ x_d^+ x_c^- x_d^- = 0$  if and only if  $x_d^+ = 0$  or  $x_d^- = 0$ , otherwise  $x_c^+ x_d^+ x_c^- x_d^- < 0$ . Moreover, since  $x_a^\pm > 0$ , according to equation (C.7e),  $x_c^+$  and  $x_e^+$  have the same sign as well as  $x_c^-$  and  $x_e^-$ . In addition,  $x_c^+$ ,  $x_d^+$ ,  $x_e^+$  are not all negative and  $x_c^-$ ,  $x_d^-$ ,  $x_e^-$  neither, since  $\frac{x_c^\pm + x_d^\pm + x_e^\pm \pm \sqrt{(x_c^\pm + x_d^\pm + x_e^\pm)^2 - 4x_c^\pm x_e^\pm}}{2}$  are roots of the polynomial  $R$ , thus positive. Hence, if  $x_d^\pm \neq 0$ , one and only one of  $x_c^+$ ,  $x_d^+$ ,  $x_c^-$ ,  $x_d^-$  is negative so that we have one of the two exclusive cases:  $x_c^+ > 0$  and  $x_d^+ > 0$  or  $x_c^- > 0$  and  $x_d^- > 0$ . Consequently, the inverse problem has at most two solutions. More precisely, considering (C.8) and the previous remark, we have the following lemma

**Lemma C.4.** *If  $V < 1/2$  and the polynomials  $P$  and  $Q$  are coprime, that is*

$$a \neq \frac{1 - 2V}{2V(1 - V)} \left( (c + d + e)V + (1 - V)b \pm \sqrt{\Delta_Q} \right),$$

*then, we have one of the exclusive cases:*

- 1)  $\frac{V}{1-2V}x_a^+ \leq \frac{V}{1-2V}a - b$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .
- 2)  $\frac{V}{1-2V}x_a^- \leq \frac{V}{1-2V}a - b < \frac{V}{1-2V}x_a^+$  and
  - 2.1)  $(1 - 2V)bc - (1 - V)(x_a^+ - a)(b - \frac{V}{1-2V}a) \leq 0$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .
  - 2.2)  $0 < (1 - 2V)bc - (1 - V)(x_a^+ - a)(b - \frac{V}{1-2V}a)$  and
    - 2.2.1)  $((-c + d + e)V - (1 - V)b)^2 + 4cdV^2 < \left( (c + d + e)V + (1 - V)b - \frac{2V(1-V)}{1-2V}a \right)^2$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .
    - 2.2.2)  $\left( (c + d + e)V + (1 - V)b - \frac{2V(1-V)}{1-2V}a \right)^2 < ((-c + d + e)V - (1 - V)b)^2 + 4cdV^2$  and
      - 2.2.2.1)  $a = x_a^+$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .
      - 2.2.2.2)  $a \neq x_a^+$ . Then the inverse problem has two solutions:  $(a, b, c, d, e)$  and  $(x_a^+, x_b^+, x_c^+, x_d^+, x_e^+)$ .
- 3)  $\frac{V}{1-2V}a - b < \frac{V}{1-2V}x_a^-$  and
  - 3.1)  $b = \frac{aV(1-V)(x_a^+ - a)}{(1-2V)((1-V)(x_a^+ - a) - (1-2V)c)}$  and

3.1.1)  $b - \frac{V}{1-2V}a \leq 0$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .

3.1.2)  $0 < b - \frac{V}{1-2V}a$  and

3.1.2.1)  $((-c + d + e)V - (1 - V)b)^2 + 4cdV^2 < \left( (c + d + e)V + (1 - V)b - \frac{2V(1-V)}{1-2V}a \right)^2$   
and

3.1.2.1.1)  $a = x_a^-$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .

3.1.2.1.2)  $a \neq x_a^-$ . Then the inverse problem has two solutions:  $(a, b, c, d, e)$  and  $(x_a^-, x_b^-, x_c^-, x_d^-, x_e^-)$ .

3.1.2.2)  $\left( (c + d + e)V + (1 - V)b - \frac{2V(1-V)}{1-2V}a \right)^2 < ((-c + d + e)V - (1 - V)b)^2 + 4cdV^2$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .

3.2)  $b = \frac{aV(1-V)(x_a^- - a)}{(1-2V)((1-V)(x_a^- - a) - (1-2V)c)}$  and

3.2.1)  $0 \leq b - \frac{V}{1-2V}a$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .

3.2.2)  $b - \frac{V}{1-2V}a < 0$  and

3.2.2.1)  $((-c + d + e)V - (1 - V)b)^2 + 4cdV^2 < \left( (c + d + e)V + (1 - V)b - \frac{2V(1-V)}{1-2V}a \right)^2$ .

Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .

3.2.2.2)  $\left( (c + d + e)V + (1 - V)b - \frac{2V(1-V)}{1-2V}a \right)^2 < ((-c + d + e)V - (1 - V)b)^2 + 4cdV^2$   
and

3.2.2.2.1)  $a = x_a^+$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .

3.2.2.2.2)  $a \neq x_a^+$ . Then the inverse problem has two solutions:  $(a, b, c, d, e)$  and  $(x_a^+, x_b^+, x_c^+, x_d^+, x_e^+)$ .

3.3)  $b \neq \frac{aV(1-V)(x_a^\pm - a)}{(1-2V)((1-V)(x_a^\pm - a) - (1-2V)c)}$  and

3.3.1)  $\left( (c + d + e)V + (1 - V)b - \frac{2V(1-V)}{1-2V}a \right)^2 < ((-c + d + e)V - (1 - V)b)^2 + 4cdV^2$   
and  $0 < (1 - 2V)bd - (1 - V)(x_a^+ - a) \left( b - \frac{V}{1-2V}a \right)$  and

3.3.1.1)  $a = x_a^+$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .

3.3.1.2)  $a \neq x_a^+$ . Then the inverse problem has two solutions:  $(a, b, c, d, e)$  and  $(x_a^+, x_b^+, x_c^+, x_d^+, x_e^+)$ .

3.3.2)  $((-c + d + e)V - (1 - V)b)^2 + 4cdV^2 < \left( (c + d + e)V + (1 - V)b - \frac{2V(1-V)}{1-2V}a \right)^2$   
and  $0 < (1 - 2V)bd - (1 - V)(x_a^- - a) \left( b - \frac{V}{1-2V}a \right)$  and

3.3.2.1)  $a = x_a^-$ . Then the inverse problem has a unique solution:  $(a, b, c, d, e)$ .

3.3.2.2)  $a \neq x_a^-$ . Then the inverse problem has two solutions:  $(a, b, c, d, e)$  and

$$(x_a^-, x_b^-, x_c^-, x_d^-, x_e^-).$$

### Non-coprimality of $P$ and $Q$

We suppose here that  $P$  and  $Q$  are not coprime. Then, according to lemma C.1,  $V < 1/2$ , moreover, according to lemma C.2,  $-\frac{1-V}{1-2V}a$  is a root of  $Q$ , that is

$$a = \frac{1-2V}{2V(1-V)} \left( (c+d+e)V + (1-V)b + (-1)^q \sqrt{\Delta_Q} \right), \quad q \in \{1, 2\},$$

where

$$\begin{aligned} \Delta_Q &= (V(c+d+e) + (1-V)b)^2 - 4V((1-V)(d+e)b + Vce) = \\ &= ((-c+d+e)V - (1-V)b)^2 + 4cdV^2. \end{aligned}$$

Moreover,  $F = \check{Q}/\check{P}$  where  $\check{P}$  and  $\check{Q}$  are coprime and

$$\begin{aligned} \check{P}(X) &= X^2 + \left( -\frac{1-2V}{1-V}(c+d+e) + b - \frac{V}{1-2V}a \right) X + \frac{1-2V}{1-V}ce, \\ \check{Q}(X) &= V \left( X + c + d + e + \frac{1-V}{V}b - \frac{1-V}{1-2V}a \right). \end{aligned}$$

$F_x = F$  only if  $P_x$  and  $Q_x$  are not coprime, that is  $-\frac{1-V}{1-2V}x_a$  is a root of  $Q_x$

$$x_a = \frac{1-V}{2V(1-2V)} \left( (x_c+x_d+x_e)V + (1-V)x_b + (-1)^{q_x} \sqrt{\Delta_{Q_x}} \right), \quad q_x \in \{1, 2\},$$

where

$$\begin{aligned} \Delta_{Q_x} &= (V(x_c+x_d+x_e) + (1-V)x_b)^2 - 4V((1-V)(x_d+x_e)x_b + Vx_cx_e) = \\ &= ((-x_c+x_d+x_e)V - (1-V)x_b)^2 + 4x_cx_dV^2. \end{aligned}$$

So that,  $F_x = \check{Q}_x/\check{P}_x$  where  $\check{P}_x$  and  $\check{Q}_x$  are coprime and

$$\begin{aligned} \check{P}_x(X) &= X^2 + \left( -\frac{1-2V}{1-V}(x_c+x_d+x_e) + x_b - \frac{V}{1-2V}x_a \right) X + \frac{1-2V}{1-V}x_cx_e, \\ \check{Q}_x(X) &= V \left( X + x_c + x_d + x_e + \frac{1-V}{V}x_b - \frac{1-V}{1-2V}x_a \right). \end{aligned}$$

Since  $\check{Q}_x/\check{P}_x$  and  $\check{Q}/\check{P}$  are irreducible and the leading coefficients of  $\check{P}_x$  and  $\check{P}$  are identical, as well as those of  $\check{Q}_x$  and  $\check{Q}$ ,  $F_x = F$  only if  $\check{P}_x = \check{P}$  and  $\check{Q}_x = \check{Q}$ . Hence,  $F_x = F$  if and only if

$$V \left( \frac{1-V}{1-2V} x_a \right)^2 + ((1-V)(x_d + x_e)x_b + Vx_c x_e) + \\ -(V(x_c + x_d + x_e) + (1-V)x_b) \frac{1-V}{1-2V} x_a = 0, \quad (\text{C.9a})$$

$$x_c x_e = ce, \quad (\text{C.9b})$$

$$-\frac{1-2V}{1-V}(x_c + x_d + x_e) + x_b + \\ -\frac{V}{1-2V} x_a = x_b - \frac{V}{1-2V} a \\ -\frac{1-2V}{1-V}(c + d + e), \quad (\text{C.9c})$$

$$V(x_c + x_d + x_e) + (1-V)x_b - \frac{V(1-V)}{1-2V} x_a = V(c + d + e) + \\ + (1-V)b - \frac{V(1-V)}{1-2V} a. \quad (\text{C.9d})$$

Hence, doing (C.9d)−(1−V)(C.9c) to replace (C.9c), followed by (1−V)(C.9d)−V(C.9c) to replace (C.9d), we get

$$V \left( \frac{1-V}{1-2V} x_a \right)^2 + ((1-V)(x_c + x_e)x_b + Vx_d x_e) + \\ -(V(x_c + x_d + x_e) + (1-V)x_b) \frac{1-V}{1-2V} x_a = 0, \quad (\text{C.10a})$$

$$x_c x_e = ce, \quad (\text{C.10b})$$

$$x_c + x_d + x_e = c + d + e, \quad (\text{C.10c})$$

$$(1-2V)x_b - Vx_a = (1-2V)b - Va. \quad (\text{C.10d})$$

Replacing  $x_a$  in (C.10a) by  $a + \frac{1-2V}{V}(x_b - b)$  given by (C.10d), we get

$$(x_c + h)x_b = \frac{V}{1-V}(h^2 + (x_c + x_d + x_e)h + x_c x_e), \quad (\text{C.11a})$$

$$x_c x_e = ce, \quad (\text{C.11b})$$

$$x_c + x_d + x_e = c + d + e, \quad (\text{C.11c})$$

$$(1-2V)x_b - Vx_a = (1-2V)b - Va, \quad (\text{C.11d})$$

where  $h = \frac{1-V}{V}b - \frac{1-V}{1-2V}a$ . Then, replacing  $x_c + x_d + x_e$  by  $c + d + e$  and  $x_c x_e$  by  $ce$  in (C.11a), we get

$$\left(x_c + \frac{1-V}{V}b - \frac{1-V}{1-2V}a\right)x_b = \left(c + \frac{1-V}{V}b - \frac{1-V}{1-2V}a\right)b, \quad (\text{C.12a})$$

$$x_c x_e = ce, \quad (\text{C.12b})$$

$$x_c + x_d + x_e = c + d + e, \quad (\text{C.12c})$$

$$(1-2V)x_b - Vx_a = (1-2V)b - Va. \quad (\text{C.12d})$$

The set of solutions  $(x_a, x_b, x_c, x_d, x_e)$  to (C.12) is a curve given by the intersection of two cylindrical hypersurfaces over hyperbolas ((C.12a), (C.12b)) and two hyperplanes ((C.12c), (C.12d)). If  $c + \frac{1-V}{V}b - \frac{1-V}{1-2V}a = 0$ , the cylindrical hypersurface given by (C.12a) degenerates into two hyperplanes whose respective equations are  $x_b = 0$  and  $x_c = c$  so that the set of solutions of (C.12) is composed of the line of equations  $(1-2V)x_b - Vx_a = (1-2V)b - Va, x_c = c, x_d = d, x_e = e$  and the hyperbola of equations  $x_a = 0, x_b = 0, x_c + x_d + x_e = c + d + e, x_c x_e = ce$ . Considering the inverse problem, we thus have the following lemma

**Lemma C.5.** *If  $V < 1/2$  and the polynomials  $P$  and  $Q$  are not coprime, then, if  $c + \frac{1-V}{V}b - \frac{1-V}{1-2V}a = 0$ , the set of solutions  $(x_a, x_b, x_c, x_d, x_e)$  to the inverse problem is given by the open half-line of  $\mathbb{R}^5$  of parametric equation*

$$t \mapsto (a + (1-2V)t, b + Vt, c, d, e), \quad t \in \left(-\frac{b}{V}, +\infty\right).$$

*Otherwise, the set of solutions  $(x_a, x_b, x_c, x_d, x_e)$  to the inverse problem is given by the curve of  $\mathbb{R}^5$  of parametric equation*

$$t \mapsto \left(a + \frac{(1-2V)(c-t)b}{Vt + (1-V)b - \frac{V(1-V)}{1-2V}a}, \frac{c + \frac{1-V}{V}b - \frac{1-V}{1-2V}a}{t + \frac{1-V}{V}b - \frac{1-V}{1-2V}a}b, t, c + d + e - t - \frac{ce}{t}, \frac{ce}{t}\right),$$

where

$$t \in \begin{cases} \left(\beta^-, \min\left(\beta^+, \frac{1-V}{1-2V}a + \frac{(1-2V)bc}{(1-2V)b - Va}\right)\right), & \text{if } Va - (1-2V)b < 0, \\ \left(\max\left(\frac{1-V}{1-2V}a - \frac{1-V}{V}b, \beta^-\right), \beta^+\right), & \text{if } 0 \leq Va - (1-2V)b < (1-2V)Vc, \\ \left(\beta^-, \min\left(\beta^+, \frac{1-V}{1-2V}a - \frac{1-V}{V}b\right)\right), & \text{if } (1-2V)Vc < Va - (1-2V)b, \end{cases}$$

where

$$\beta^\pm = \frac{c + d + e \pm \sqrt{(c + d + e)^2 - 4ce}}{2}.$$