ID N. 29

Università
di Genova

UNIVERSITÀ CAMPUS BIO-MEDICO DI ROMA

DEPARTMENT OF ENGINEERING

UNIVERSITÀ DEGLI STUDI DI GENOVA

DEPARTMENT OF MATHEMATICS

Italian National Ph.D. in Artificial Intelligence
Health and Life Sciences
XXXVII Cycle

Zooming out: from a voxel-wise dynamic PET modeling problem to a full FOV of implicit inverse problems

Supervisors

Prof. Federico Benvenuto
Dr. Sara Garbarino
Prof. Michele Piana

Candidate
Davide Parodi

May, 2025

To my family and friends.

Acknowledgements

Acknowledgments

I would like to thank all the professors who guided me throughout my studies and enriched my knowledge with their insights, support, and constructive advice. I am especially grateful to my supervisors at the Università Degli Studi di Genova, MIDA group: Prof. Federico Benvenuto, Dr. Sara Garbarino, Prof. Michele Piana, and Prof. Cristina Campi. I would also like to thank Prof. Dr. Gianmario Sambuceti from the Nuclear Medicine Department at Policlinico IRCCS San Martino Hospital in Genova for his invaluable interdisciplinary contribution to our work, which bridges mathematics and medicine.

I also want to thank my colleagues for the many shared moments of study, discussion, and cultural exchange. Being part of such a diverse and international environment has been a truly enriching experience.

I wish to acknowledge the informal support of friends who, directly or indirectly, contributed to making these years more manageable.

Finally, I would like to express my appreciation to my family for their continuous support during this period.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Abstract

This thesis investigates mathematical and computational methods for modeling tracer dynamics in PET imaging, with a focus on advancing quantitative analysis beyond conventional metrics. The work originates from a clinical observation in FDG-PET scans of lung cancer patients, where voxel-wise relative Patlak analysis revealed unexpected tracer release in certain tumor regions—an effect likely associated with inflammatory infiltrates. This finding highlighted the limitations of standard semi-quantitative indices like SUV and motivated a deeper exploration of kinetic modeling. We first examine compartmental models, particularly the 2-compartments and Sokoloff models, to estimate biologically meaningful kinetic parameters through systems of ODEs. These models provide a more robust framework for interpreting tracer kinetics, especially in oncology, where irreversible accumulation patterns are of interest.

Building on this foundation, we address the inverse problem of parameter estimation from noisy temporal data. We reinterpret it as an implicit inverse problem, where parameters are not directly observable but constrained by the dynamics of the system. To solve this, we propose a homotopy-based optimization strategy that gradually transitions from a highly regularized to a minimally regularized formulation. This path-following approach is coupled with gradient descent, adjoint-state gradient computation, and Newton–Raphson integration for the forward model.

The thesis concludes with a broader framing of these challenges within the emerging field of ODE learning, where data-driven techniques aim to recover unknown dynamics. By bridging clinical imaging and abstract inverse problems, this work contributes novel tools for quantitative modeling in both biomedical applications and general dynamical systems analysis.

Chapter 2

Introduction

Nuclear medicine plays a crucial role in the noninvasive investigation of physiological and pathological processes, with one of its key objectives being the reconstruction of tissue metabolism in vivo. Among the imaging modalities developed for this purpose, Positron Emission Tomography (PET) stands out for its ability to quantify the distribution of radiolabeled compounds (tracers) that participate in biological pathways. By detecting coincident γ-photon pairs produced from positron–electron annihilation events, PET enables the reconstruction of three-dimensional tracer activity within the body [127, 111, 78, 64, 79, 66, 109, 47, 134].

A widely used tracer in clinical practice is [^{18}F]fluoro-2-deoxy-D-glucose (FDG), a glucose analog that allows for indirect assessment of glucose metabolism. FDG uptake is particularly informative in oncology, where increased metabolic activity leads to hyperaccumulation of the tracer in tumors [139, 55], and in neurology, where hypometabolism often marks neurodegenerative processes [30, 103, 114]. FDG-PET's effectiveness is further supported by the favorable physical properties of the radionuclide ^{18}F, which offers a suitable compromise between image quality and practical handling. Despite its clinical value, the standard approach to PET quantification—based on the Standardized Uptake Value (SUV) - has notable limitations. SUV is a semi-quantitative metric sensitive to a range of physiological and technical factors (e.g., patient weight, blood glucose levels, acquisition timing) [61, 119]. Moreover, it provides only a snapshot of tracer accumulation, lacking the dynamic and mechanistic insight required for detailed metabolic characterization [122].

To overcome these limitations, mathematical modeling offers a more principled approach to tracer analysis. Compartmental models describe the kinetics of tracer exchange between different biological states (e.g., free, metabolized) via systems of ordinary differential equations (ODEs) [17, 41, 68, 106]. These models enable the estimation of key kinetic parameters — such as the rate of phosphorylation — through inverse problem techniques. In oncology,

for instance, the 2-compartments model and its irreversible variant (the Sokoloff model) are used to quantify FDG trapping in tumor tissues, with the phosphorylation rate serving as a proxy for glycolytic activity [33, 115]. An established method to extract information on irreversible tracer kinetics is Patlak's graphical analysis [68, 87], which under simplifying assumptions reduces the inverse problem to linear regression. While this approach facilitates interpretation, its applicability is limited to later acquisition phases and it offers only qualitative information on tracer dynamics [140, 65].

In this thesis, we revisit these modeling tools through the lens of a clinical anomaly observed in lung cancer FDG-PET scans: relative Patlak analysis at the voxel level revealed unexpected tracer release (i.e., negative slopes) in peripheral tumor regions, a phenomenon potentially linked to inflammation or necrosis. This observation motivates a more detailed investigation of metabolic heterogeneity in tumors and demonstrates the limitations of ROI-averaged or global metrics. Building on this case study, the thesis transitions from application-driven analysis to methodological innovation. Compartmental models are recast as dynamical systems governed by ODEs, and the associated inverse problems are framed in a general setting known as implicit inverse problems [77, 81, 1, 7, 76, 98], where parameters must be inferred from observations constrained by nonlinear dynamics.

We develop a novel optimization framework to solve such problems using a homotopy-based path-following approach, which incrementally relaxes the regularization imposed on the parameter estimates [84]. This strategy ensures stability and convergence in the presence of ill-posedness and noise. The optimization is performed via gradient descent, with gradients efficiently computed using the adjoint state method. The forward integration of the ODE system is handled using a Newton–Raphson-based solver to handle model nonlinearities. Finally, we contextualize this work within the broader field of ODE learning and data-driven modeling of dynamical systems. By bridging clinical imaging with modern computational techniques, this thesis contributes methodological tools applicable not only to PET but also to other domains where understanding temporal dynamics from sparse data is essential.

The thesis is structured as follows.

Chapter 3 provides a comprehensive overview of the principles underlying PET medical imaging. It covers the metabolism of glucose and FDG in the human body, the formation of PET data from photon detection, the reconstruction of images, and the estimation and clinical interpretation of the SUV.

Chapter 4 introduces compartmental analysis, presenting key models - including the Sokoloff model — as examples. Both analytical and graphical techniques for solving the associated inverse problems are discussed.

Chapter 5 focuses on a voxel-wise Patlak analysis of FDG concentration in lung cancer

patients. This study investigates localized tracer loss at tumor boundaries and was conducted in collaboration with the Department of Nuclear Medicine at IRCCS Ospedale Policlinico San Martino in Genova.

Chapter 6 explores the data-driven reconstruction of unknown dynamical systems, with a specific application to radioactive tracer kinetics in PET. It provides an overview of advanced machine learning techniques used to address such problems.

Chapter 7 presents a novel numerical and mathematical framework for solving implicit inverse problems, with a focus on stability, optimization, and integration methods tailored to nonlinear dynamical models.

Chapter 3

An overview on PET principles

This chapter explains the basic principles of performing a scan with a PET machine. PET stands for Positron Emission Tomography. PET is a functional imaging technique, which means it can image the metabolic activity of structures within the body. This type of imaging is very useful in applications of oncology or neurodegenerative diseases, where metabolic changes precede the morphological ones. The representation of functional tissue activity is accomplished through the use of a radioactive tracer, a liquid substance that is injected intravenously into the patient. The radioactive decay of the tracer emits photons that are detected by PET collimators. Then mathematical reconstruction techniques are used to obtain PET images from the measured photon counts. Finally, the PET data are normalized to provide the Standardized Uptake Value (SUV) used by nuclear medicine physicians for diagnosis or staging. Figure 3.1 show the entire PET examination process.

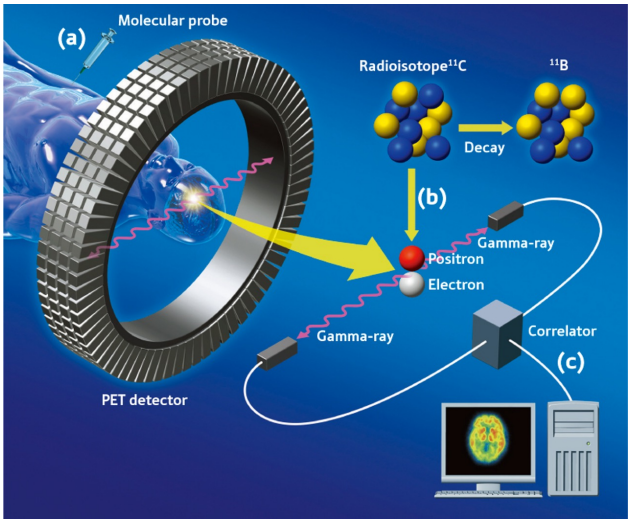

Figure 3.1: Complete process of a PET scan. (a) The radioactive tracer is injected intravenously into the patient's body and distributed throughout the body; (b) The radioactive tracer releases positrons, which, when annihilated by electrons, emit two antipodal photons that are detected by PET collimators arranged in a circular ring around the patient's body; (c) From the photon counts, the PET image is obtained through image reconstruction techniques.

In this chapter we present:

- the $[^{18}F]$fluoro-2-deoxy-D-glucose (FDG) ($C_6H_{11}FO_5$) metabolism, one of the most commonly used radioactive tracers;
- the PET data formation process;
- the principal PET image reconstruction techniques from acquired data;
- the SUV value.

3.1 Glucose and FDG metabolism

Glucose ($C_6H_{12}O_6$) is the main source of energy in our body. In oncology studies using PET scanning, the goal of the analysis is to reconstruct the pattern of tissue metabolism in the patient's body and to identify the areas of highest glucose consumption. To achieve this goal, radioactive tracers are used, i.e. radioactive substances with specific properties that are distributed in different parts of the body. These tracers, due to radioactive decay, emit photons that are detected by PET collimators, making it possible to visualize the areas of highest accumulation, providing critical information for disease diagnosis and monitoring. Tracers must therefore have properties similar to those of glucose in order to mimic its

physiological behavior as closely as possible. It is therefore crucial to know how glucose metabolism works at the cellular level.

3.1.1 Glucose metabolism

The glucose molecule enters the cell and, in order to produce energy through the process of glycolysis, must be able to cross the cell membrane. There are enzymes on the membrane, including the enzyme hexokinase (HK), which opens the membrane and allows glucose to enter the cell. Glucose undergoes a chemical reaction catalyzed by adenosine triphosphate (ATP), which adds 6 phosphate atoms (PO_4^{3-}) to glucose, forming glucose-6-phosphate ($C_6H_{11}O_6P$) and adenosine diphosphate (ADP)

$$C_6H_{12}O_6 + \text{ATP} \overset{\text{hexokinase}}{\longrightarrow} C_6H_{11}O_6P + \text{ADP}$$

Glucose-6-phosphate has several metabolic fates. It is used to produce directly usable energy through the process of glycolysis. Alternatively, it can be stored as glycogen. Other possibilities include the pentose phosphate pathway with the production of NADPH and ribose 5-phosphate, and gluconeogenesis when the cell needs to release glucose into the bloodstream (e.g., during fasting).

3.1.2 FDG metabolism

As anticipated at the beginning of this section, radioactive tracers must have glucose-like properties: in particular, the hexokinase enzyme must be able to allow the radioactive molecule to enter the cell. Early studies focused on making the constituent molecules of glucose radioactive. However, the decay characteristics of these isotopes made them either unsuitable for the timescales of PET procedures or problematic due to their potential social impact. Radioactive decay time, also known as half-life, is the time it takes for a radioactive substance to reduce its photon-emitting activity by half. For example, the half-life of radioactive oxygen (O_{15}) is extremely short, about 120 seconds, while tritium, a radioactive isotope of hydrogen, has a decay time of about 2000 years. These technical problems led to the abandonment of this first strategy.

In 1968, Dr. Josef Pacak, Zdenek Tocik and Miloslav Cerny at the Department of Organic Chemistry, Charles University (Czechoslovakia), were the first to describe the synthesis of FDG [83]. It is a radioactive molecule analogous to glucose, that is, with properties very similar to those of glucose. Later, in the 1970s, Tatsuo Ido and Al Wolf at the Brookhaven National Laboratory were the first to describe the synthesis of FDG labeled with 18F [50]. This radioactive tracer is still the most widely used PET scan in oncology applications.

Although FDG is a glucose analog, its metabolic fate is different from that of glucose. Initially, the hexokinase enzyme does not recognize the difference between FDG and glucose; therefore, FDG undergoes the same chemical reaction as glucose to form FDG-6-phosphate (FDG6P):

$$C_6H_{11}FO_5 \overset{\text{hexokinase}}{\longrightarrow} C_6H_{11}FO_{56}P + ADP$$

However, within the cytoplasm of the cell, FDG-6P does not undergo glycolysis and is not stored as an energy source as glycogen. Thus, FDG-6P is subject to two different metabolic fates:

- becomes irreversibly trapped in the cytoplasm of the cell;
- loses the 6 phosphorus atoms and exits the cell by crossing the cell membrane. It becomes available again to be metabolized by other cells.

Clinically, it is hypothesized that tumors, due to their high demand for glucose, irreversibly trap phosphorylated FDG in the cytoplasm, preventing its release [89].

3.2 Data formation process

Positron Emission Tomography [82] is the most widely used nuclear medicine imaging technique [5]. The process of generating PET data is based on the principle of radioactive decay of the tracer. This process occurs as a result of the emission of positrons from the molecule (hence the name Positron Emission Tomography). The radioactive activity of a molecule is expressed as

$$A(t) = A_0 e^{-\frac{\ln(2)t}{\frac{T_1}{2}}} \tag{3.1}$$

where $\frac{T_1}{2}$ is half-life of radioactive activity, i.e. the time required for the number of positrons emitted per unit of time to be half the initial number. Figure 3.2 shows the FDG radioactive decay curve [111].

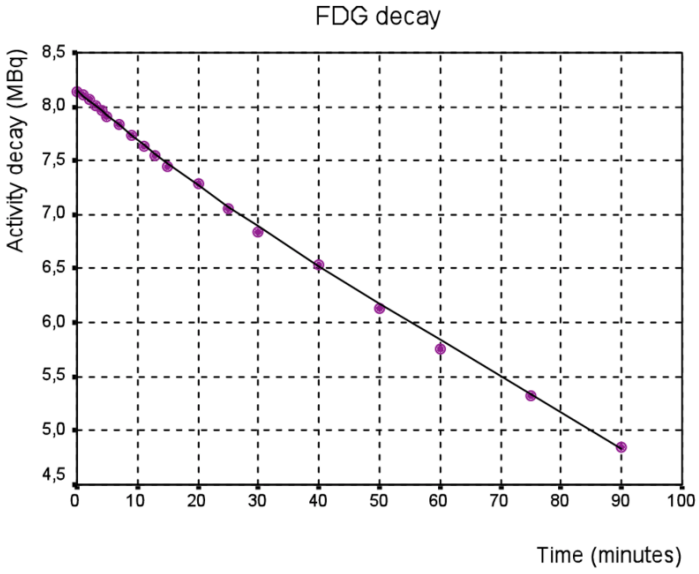

Figure 3.2: FDG radioactive decay curve

The emitted positron, being unstable, travels a short distance (on the order of about 1 mm) before interacting with an electron e^-. This interaction leads to the phenomenon of annihilation, which generates the emission of two photons γ of energy 511 KeV, emitted in opposite directions (as shown in point (b) of Figure 3.1). The line along which these two photons travel is called the LOR (Line Of Response). The photons reach the collimators of the PET machine, whose detectors are arranged in a cylindrical ring around the patient. This positioning makes it possible to detect photons emitted from different angles. When two photons belonging to the same LOR are detected simultaneously by two opposite detectors, the system records a count and stores it as raw data inside the PET machine. Data acquisition is performed in list-mode, i.e. each event is recorded with precise information about the time of detection and the detector involved. At the end of the acquisition, the sum of the counts collected by each detector represents the raw data.

From this data, by solving a inverse imaging problem, the PET image can be reconstructed, allowing analysis of the metabolism and distribution of the radioactive tracer in the patient's body.

3.2.1 Resolution degrading factors

The PET machine, like any imaging system, cannot perfectly represent reality: in fact, it acts as a signal transmission system that receives an input signal $f^0(x)$, with $x \in \mathbb{R}^n$, and transforms it into an output signal $g^0(x) \neq f^0(x)$. Several physical factors, both constructive and non-constructive, affect image quality. For example, object motion, positron range,

photon noncollinearity, and detector-related effects (including crystal width, intercrystal scattering, and intercrystal penetration).

The spatial resolution of the PET machine is the ability of the system to distinguish between two radioactive sources in close proximity. Resolution is commonly described in terms of the Point Spread Function (PSF) and its Full Width at Half Maximum (FWHM).

Assuming that the process of signal transmission and image formation can be modeled as linear processes, the output of the g^0 system can be defined by the following convolution product:

$$g^0(x) = \int_{\mathbb{R}^n} K(x, x') f^0(x') dx' \tag{3.2}$$

where $K(x, x')$ is the so called PSF. The PSF represents the response of the system to a point source. In fact, if the input signal is a point source $f^0(x) = \delta(x - x_0)$, then substituting in (3.2) we get:

$$g^0(x) = \int_{\mathbb{R}^n} K(x, x')\delta(x' - x_0)dx' = K(x, x_0) \tag{3.3}$$

The distortion effect introduced by the PSF is called *blurring*; to study it, the machine response at different sources is observed. If the distortion introduced by the PSF K is uniform at all points in the domain, then we say that the PSF is space invariant and $K(x, x_0) = \tilde{K}(x - x_0)$, otherwise we say that the PSF is space variant.

Since the PSF enlarges the point source but preserves the total radiation, the intensity at the original point is reduced in the final image, while it is increased in the neighboring areas. The quantity that characterizes how much the PSF expands the point source is the Full Width at Half Maximum (FWHM). It represents the width of the PSF at half of its maximum height. The smaller this value is, the more accurate the system is. The resolution of the PET machine is about 2 millimeters FWHM [78]. An example of the effect of PSF on a point source is shown in Figure 3.3.

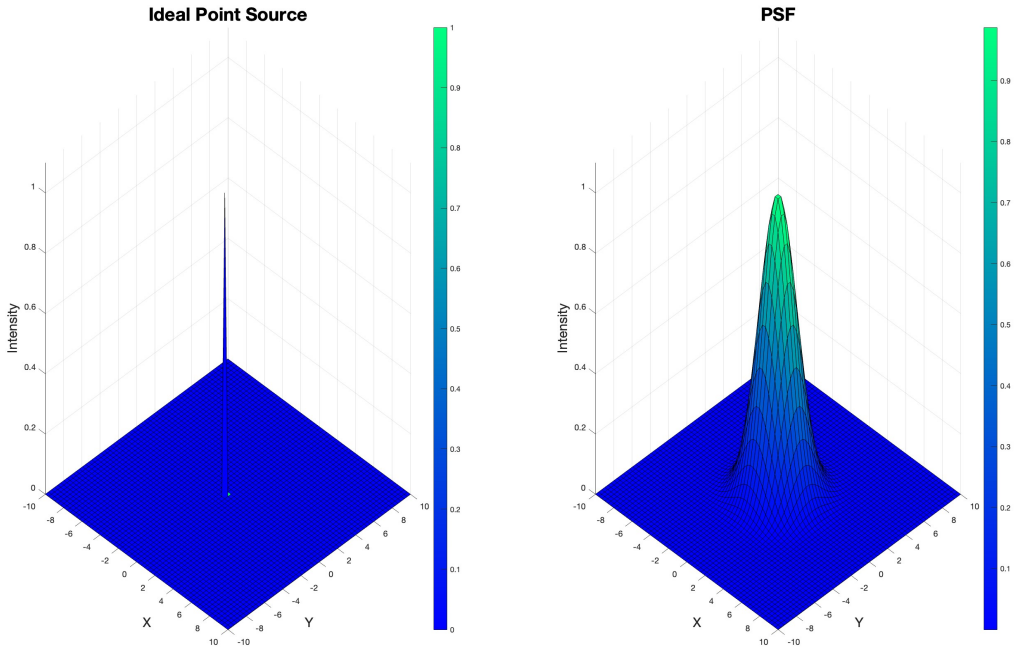

Figure 3.3: Effect of a PSF on a point source. On the left, the point source in $O = (0, 0)$ has a maximum intensity of 1. On the right, the PSF has reduced the maximum intensity value at the origin and increased the intensity value in the neighboring region.

In an image that is more complex than a simple point source, the blur phenomenon affects all points: an *aberrated* image is obtained.

Limited spatial resolution in PET leads to Partial Volume Effects (PVEs), which are generally divided into two main categories:

- spill-over effect: this phenomenon occurs when a source of high radioactive activity appears enlarged by the PET PSF. As a result, some of the radioactivity expands into adjacent regions, leading to an underestimation of the actual activity within the source itself;

- spill-in effect: this occurs when a region of low radioactive activity receives a signal from a nearby region of high intensity. This leads to an artificial overestimation of radiopharmaceutical activity in the low-intensity region.

Here are some critical issues that can degrade the data and negatively affect the quality of reconstructed PET images.

The first problem occurs at the source of the data. The positron emitted by the radioactive molecule travels a distance of about 1 mm before annihilating with an electron and

emitting two antipodal γ photons. This phenomenon introduces an uncertainty factor into the image reconstruction, since the exact point of positron emission cannot be determined.

Other issues include the phenomena of attenuation and scatter. Attenuation occurs when photons interact with electrons in body tissues via the Compton phenomenon, reducing the number of photons detected. Scattering, on the other hand, occurs when one or both photons are deflected from their original trajectory by interaction with other particles. As a result, they hit the collimators at the wrong positions, leading to the determination of an incorrect line of response. In addition, random coincidences may occur, i.e., events in which photons (scattered or unscattered) belonging to different LORs are mistakenly assigned to the same LOR.

An additional problem is the dead time of the detectors, or the blind time of the system. When a detector is hit by an excessive number of photons in a very short time interval, it is temporarily blinded for a period on the order of nanoseconds. During this time window, the detector is unable to detect new events, resulting in a loss of information and a reduction in the accuracy of the PET image.

3.3 Image reconstruction process

In this section, we address the inverse imaging problem of reconstructing the representative PET image of tissue metabolism from the photon counts measured by the PET scanner. For all the problematic aspects highlighted in subsection 3.2.1, this problem is ill-posed according to the definition given by Hadamard [43].

First, we define the Radon transform, a mathematical tool by which we can formally describe the data as a sinogram. Next, we examine the main image reconstruction methods employed in this field, starting with the analytical ones and ending with *Filtered Back Projection* algorithms (FBP) [64, 79] and those iterative ones of Bayesian nature, such as the *Maximum Likelihood Expectation Maximization* (MLEM) [66, 109] and its optimization *Ordered Subsets Expectation Maximization* (OSEM) [47].

3.3.1 Formal Description of acquired data: the sinograms

The mathematical modeling of PET data is based on the Radon transform. In order to describe it, it is necessary to provide some mathematical definitions.

- Given $\theta \in \mathcal{S}^{n-1}$ (the sphere in $\mathbb{R}^n$) and $s \in \mathbb{R}$ we define the following hyperplane:

$$H(\theta, s) = \{x \in \mathbb{R}^n, x \cdot \theta = s\} \tag{3.4}$$

- Let us define the cylinder C^n in $\mathbb{R}^n$ as:

$$C^n = \{(\theta, s), \theta \in \mathcal{S}^{n-1}, s \in \mathbb{R}\} \tag{3.5}$$

- Let us define the Schwartz space on the cylinder (3.5) as

$$\mathcal{S}(C^n) = \left\{ f \in C^\infty(C^n), \sup_{\theta \in \mathcal{S}^{n-1}, s \in \mathbb{R}} \left| s^l \frac{\partial^k \partial^j f}{\partial \theta^k \partial s^j} \right| < \infty, \forall j, l \in \mathbb{N}, k \in \mathbb{N}^n \right\} \tag{3.6}$$

Definition 3.3.1. *Given C^n (3.5) the cylinder in $\mathbb{R}^n$ and $\mathcal{S}(C^n)$ the Schwartz space on the cylinder, the Radon Transform is a bounded integral operator defined as*

$$R : \mathcal{S}(\mathbb{R}^n) \longrightarrow \mathcal{S}(C^n)$$

such that, for each $\theta \in \mathcal{S}^{n-1}$ and $s \in \mathbb{R}$

$$(Rf)(\theta, s) = \int_{H(\theta,s)} f(x)dx \tag{3.7}$$

where $H(\theta, s)$ is the hyperplane (3.4)

Writing generic vector $x \in \mathbb{R}^n$ in the coordinates identified by θ and $\theta^\perp$ the Radon Transform can be defined also as:

$$(Rf)(\theta, s) = \int_{y \in \theta^\perp} f(s\theta + y)dy \tag{3.8}$$

The Radon transform provides a formal description of the data acquired in PET imaging. In the 2-dimensional mode, the acquisition is performed by detectors arranged in a cylindrical ring around the patient. In this context, the function f in (3.7) represents the unknown metabolism to be reconstructed. Each LOR associated with photons emitted by radioactive decay corresponds to a line of integration in the (3.7) for a given value of s and a given angle θ. In this PET application we define $H(\theta, 0)$ as the origin hyperplane: $s > 0$ represents the distance from $H(\theta, 0)$ in direction θ. Consequently, the 2-dimensional recording of the PET machine can be interpreted as the set of line integrals computed as all possible values of s and θ vary. We can then describe the measured datum as the Radon transform of the unknown metabolism f that we want to reconstruct. The data formation can then be expressed by the following equation, called the imaging equation:

$$g = Rf + n \tag{3.9}$$

where:

- g is the measured data;
- R is the Radon transform as defined in (3.7);
- f is the unknown information that has to be reconstructed;
- n is the additive noise that can be deterministic or modeled by a random variable.

The goal of the inverse imaging problem is to reconstruct the unknown f function from noisy measurements of its Radon transform.

The acquired data, expressed in the variables s and θ, are stored in the so-called sinogram, as shown in the Figure 3.4.

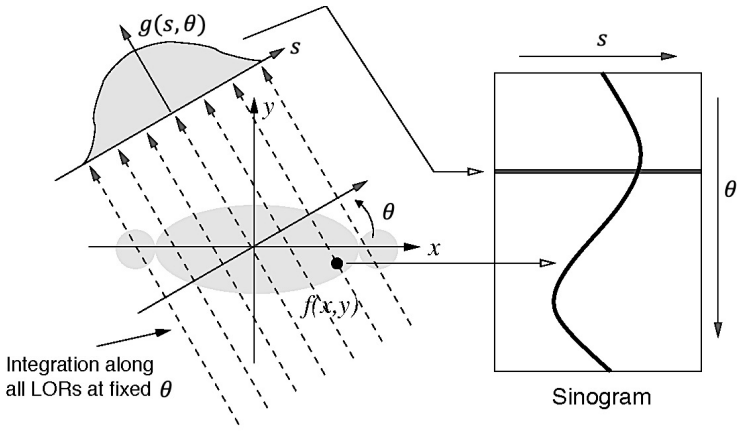

Figure 3.4: Example of a 2-dimensional PET acquisition and data expressed as a sinogram. On the left, the function $f(x, y)$ describes the quantity to be reconstructed (in this case the metabolic activity); $g(s, \theta)$ represents the measured data obtained by integration along the LORs, expressed as the Radon transform in the coordinates (s, θ). On the right is an example of a sinogram.

As shown in Figure 3.5, during the PET scan, the bed on which the patient lies moves slowly along the main axis of the body and crosses the PET detectors. The 2-dimensional acquisition is repeated for each section of the body, resulting in a series of layered scans of the patient's entire volume. For each bed scan, the same acquisition process is repeated, obtaining different sinograms, which represent the projections of the patient's metabolic activity along the different acquisition directions.

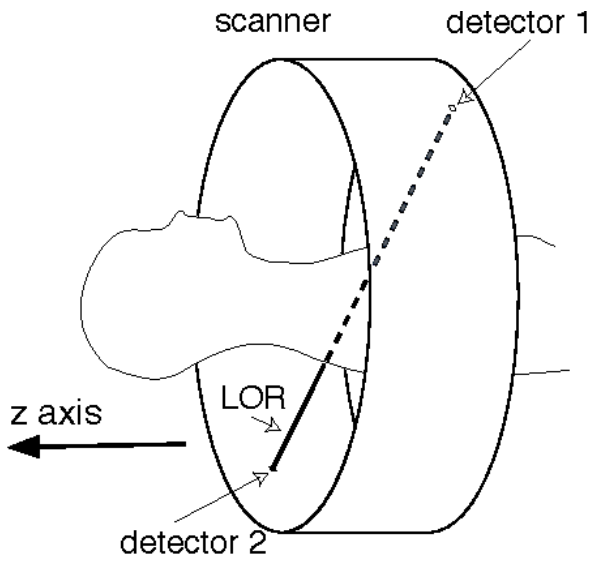

Figure 3.5: Bidimensional acquisition with PET machine. The same mode of acquisition is repeated as the bed moves along the z axis.

3.3.2 Inversion methods: analytical inverse formula and iterative techniques

Given the data formation model (3.9), the inverse imaging problem consists in reconstructing the distribution function f, defined in the Cartesian space $x \in \mathbb{R}^n$, from the data g, defined in the Radon transform space (θ, s).

Analytical inverse formula

The analytic inversion formula allows to explicitly write the unknown distribution function f in terms of the given g. To do this, it is necessary to introduce some definitions.

Definition 3.3.2. *Given $f \in \mathcal{S}(\mathbb{R}^n)$, the Fourier transform of is defined as follows:*

$$\hat{f}(\xi) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\xi} dx \tag{3.10}$$

Definition 3.3.3. *Given $f \in \mathcal{S}(\mathbb{R}^n)$ and $\alpha < n$, the Riesz's potential operator is defined through its Fourier transform as:*

$$\hat{(I^{\alpha} f)}(\xi) := |\xi|^{-\alpha} \hat{f}(\xi) \tag{3.11}$$

Definition 3.3.4. *Given $g \in \mathcal{S}(C^n)$ the backprojection operator of the Radon transform R is defined as $R^{\#} : \mathcal{S}(C^n) \to \mathcal{S}(\mathbb{R}^n)$ such that:*

$$(R^{\#}g)(x) = \int_{S^{n-1}} g(\theta, x \cdot \theta) d\theta \tag{3.12}$$

In particular, if $f \in \mathcal{S}(\mathbb{R}^n)$ holds that

$$\int_{C^n} (Rf)(\theta, s) g(\theta, s) d\theta ds = \int_{\mathbb{R}^n} f(x)(R^{\#}g)(x) dx \tag{3.13}$$

In the light of the definitions just given, it is possible to give the inversion theorem of the Radon transform.

Theorem 3.3.1. *Given $g \in \mathcal{S}(C^n)$ such that $g = Rf$ for $f \in \mathcal{S}(\mathbb{R}^n)$ and I^{α} the Riesz's potential operator with $\alpha < n$, $R^{\#}$ the backprojection operator of the Radon transform, we have that:*

$$f = \frac{1}{2}(2\pi)^{-n+1} I^{-\alpha} R^{\#} I^{\alpha-n+1} g \tag{3.14}$$

Remark 3.3.1. *We observe that, in equation* (3.14)*, parameter $\alpha < n$ is not fixed. This implies that* (3.14) *gives a family of one-parameter-dependent solutions. If a particular value of α is chosen then a solution is fixed.*

If $\alpha = 0$ is fixed then:

$$f = \begin{cases} \frac{1}{2}(2\pi)^{-n+1} R^{\#} (-1)^{\frac{n-1}{2}} g^{(n-1)} & \text{if } n \text{ is odd} \\ \frac{1}{2}(2\pi)^{-n+1} R^{\#} (-1)^{\frac{n-2}{2}} H g^{(n-1)} & \text{if } n \text{ is even} \end{cases} \tag{3.15}$$

where $H : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ is the Hilbert transform defined as

$$(Hf)(x) := \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{f(x-y)}{y} dy \tag{3.16}$$

We note that, in both cases, it is necessary to compute the numerical derivative of the datum g. This operation is always ill-posed in the sense of Hadamard [43] because of oscillations in the datum due to the presence of noise. Therefore, the inversion of the Radon transform is a numerically unstable operation. In particular, depending on whether n is even or odd, the inversion formula of the Radon transform changes both numerically and conceptually. Indeed, if n is even, the inverse of the Radon transform is unstable and non-local: the presence of the Hilbert transform operator requires the integration of $g^{(n-1)} \in \mathcal{S}(C^n)$ with respect to the second component over the whole $\mathbb{R}$. Since the inversion

operation of the Radon transform cannot conceptually depend on the dimension n, it was decided to place $\alpha = n - 3$. In this way:

$$f = -\frac{1}{2}(2\pi)^{-n+1} I^{3-n} R^{\#} g^{(2)} \tag{3.17}$$

With this choice locality is restored, but it is necessary to numerically derive g twice, which makes this inversion operation extremely ill-conditioned.

Filtered Back Projection (FBP)

For the description of this algorithm, the dimension of the space will be $n = 2$. Let us introduce the following definition and theorem.

Definition 3.3.5. *Given $\hat{f} \in \mathcal{S}(\mathbb{R}^2)$ we define the Fourier anti-transform operator applied to $\hat{f}$ as:*

$$f(x) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \hat{f}(\omega) e^{i\omega \cdot x} d\omega \tag{3.18}$$

for $x \in \mathbb{R}^2$.

Given R the Radon transform and a fixed $\theta \in \mathcal{S}^{n-1}$, we define $R_\theta(s) := R(\theta, s)$.

Theorem 3.3.2. *Given $f \in \mathcal{S}(\mathbb{R}^n)$ then*

$$\hat{(R_\theta f)}(\sigma) = (2\pi)^{\frac{n-1}{2}} \hat{f}(\sigma\theta) \tag{3.19}$$

for $\sigma \in \mathbb{R}$.

Taking advantage of the definition and theorem just given, and applying an appropriate coordinate change, it is possible to write the unknown distribution function f explicitly in terms of the Fourier anti-transform of the Fourier transform of the datum g. In particular:

$$f(x) = \frac{1}{4\pi} \int_0^{2\pi} \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} |\omega| \hat{g}(\theta, \omega) e^{i\omega x \cdot \theta} d\omega \right) d\theta \tag{3.20}$$

for $x \in \mathbb{R}^2$. By defining

$$G(\theta, s) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\omega| \hat{g}(\theta, \omega) e^{i\omega s} d\omega \tag{3.21}$$

with $\theta \in [0, 2\pi]$ and $s \in \mathbb{R}$, the filtered projections, it is possible to write (3.20) as

$$f(x) = \frac{1}{4\pi}(R^{\#} G)(x) = \frac{1}{4\pi} \int_0^{2\pi} G(\theta, x \cdot \theta) d\theta \tag{3.22}$$

From equations (3.21) and (3.22) we learn that the backprojection operator of the Radon transform $R^{\#}$ is applied to the filtered projections which are the Fourier transform of the datum g multiplied by function $|\omega|$ called *ramp filter*.

Remark 3.3.2. *Since the data is affected by noise, its Fourier transform has unwanted high frequency components: these components, multiplied by the $|\omega|$ ramp filter, are further amplified.*

The procedure used in the *Filtered Back Projection* algorithm is as follows. For each $\theta \in [0, 2\pi]$:

1. Apply the Fourier transform to the data $g(\theta, s)$ obtaining $\hat{g}(\theta, \omega)$;
2. Multiply the Fourier transform $\hat{g}(\theta, \omega)$ with $\hat{V}(\omega) = |\omega|H(\omega)$ where $H(\omega)$ is a custom frequency filter that has to attenuate high frequency components of ramp filter $|\omega|$;
3. Apply the Fourier anti-transform obtaining the filtered projections dependent on the chosen frequency filter H:

$$G_H(\theta, x \cdot \theta) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\omega| H(\omega) \hat{g}(\theta, \omega) e^{i\omega x \cdot \theta} d\omega \tag{3.23}$$

4. Apply the backprojection operator $R^{\#}$ to the filtered projections, obtaining:

$$f_{FBP}(x) = \frac{1}{4\pi} \int_{0}^{2\pi} G_H(\theta, x \cdot \theta) d\theta \tag{3.24}$$

Examples of commonly used frequency filters are reported in Figure 3.6 and are:

- Ram-Lak filter $H_{RL}(\omega) := |\omega| \chi_\Omega(\omega)$;
- Cosine filter $H_C(\omega) := |\omega| \cos\left(\pi \frac{\omega}{2}\right) \chi_\Omega(\omega)$;
- Shepp-Logan filter $H_{SL}(\omega) := |\omega| \operatorname{sinc}\left(\frac{\omega}{2}\right) \chi_\Omega(\omega)$;

where $\chi_\Omega(\cdot)$ is the characteristic function of the frequency range Ω to be maintained and sinc is the cardinal sine function such that $\operatorname{sinc}(0) = 1$ and $\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$ for $x \neq 0$

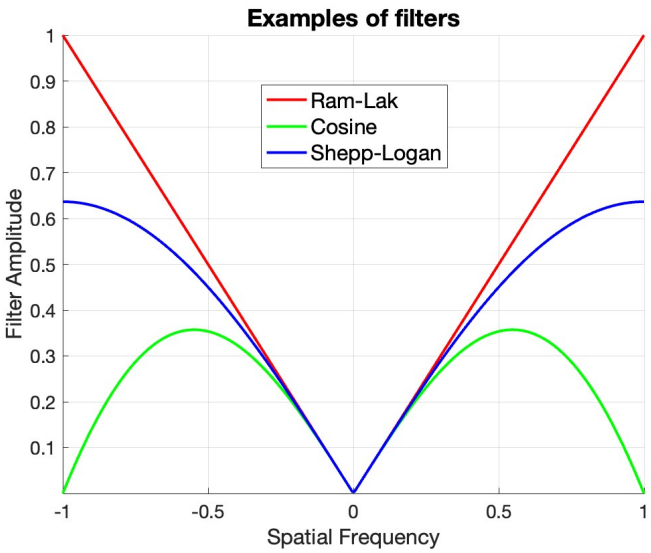

Figure 3.6: Example of the most commonly used frequency filters $H(\omega)$. x-axis reports the normalized frequency space; y-axis reports the filter amplitude.

In practical applications, only a discrete and noisy set of samples of g is available from PET scans. Consequently, the continuous formulations used in the *Filtered Back Projection* algorithm must be discretized, with integrals replaced by summations over the sampled points in Radon space.

Iterative Bayesian techniques: MLEM, OSEM

In the equation (3.9) we described the datum g as the sum of the Radon transform of the unknown distribution function f and the additive noise n. In this statistical approach based on iterative techniques [135] we assume that n is modeled by a random variable. In particular, since PET collimators count the photons that reach them, it is reasonable to assume that all quantities involved follow a Poisson-type statistic whose density function is $P_\lambda(z) = e^{-\lambda}\dfrac{\lambda^z}{z!}$ for each $z \in \mathbb{N}$.

The class of reconstruction methods we present is based on Bayes' theorem, which states that, given two events A, B, holds that

$$\pi(A|B) = \frac{\pi(B|A)\pi(A)}{\pi(B)} \tag{3.25}$$

Specifically, applying this theorem to the quantities involved in PET analysis, we obtain

$$\pi(f|g) = \pi(g|f)\pi(f) \tag{3.26}$$

where $\pi(g)$ is omitted because g is given and thus it is a deterministic quantity. The idea behind these methods is to solve a maximum likelihood problem [121]:

$$f^* := \arg\max_f \pi(g|f) \tag{3.27}$$

Since the forward model of the formation of the data (3.9) and the probabilistic Poissonian nature of the phenomenon are known, the (3.27) consists in finding the unknown distribution function f whose data generated with the forward model (3.9) is maximally compatible with g.

The form of Poisson's likelihood is

$$\pi(g|f) = \sum_{i=1}^{L} e^{-(Rf)_i} \frac{(Rf)_i^{g_i}}{g_i!} \tag{3.28}$$

where g_i is the number of counts detected by the i^{th} collimator and L is the number of collimators. Through the notation of conditional probability, we assume that the observed data g_i are realizations of random variables G_i, each with mean Rf_i, for $i = 1, \ldots, L$.

Furthermore, the solution f must be sought in a restricted subset of possible solutions, which is encoded by the prior $\pi(f)$. For example, since the function f represents the physical amount of uptake of the tracer, it can never take negative values: in probabilistic terms, this condition can be translated as $f \in S$ where $S = \{f \,|\pi(f < 0) = 0\}$.

Since the likelihood expression (3.28) and the positivity constraint on f are both convex, it is possible to transform the (3.27) into the more tractable constrained minimum problem by some algebraic operations:

$$\arg\min_{f\in S} -\ln(\pi(g|f)) = \arg\min_{f\in S} \sum_{i=1}^{L} \left\{ (Rf)_i - g_i + g_i \ln\left(\frac{g_i}{(Rf)_i}\right) + c_i \right\} \tag{3.29}$$

In this equation we observe terms:

- $(Rf)_i - g_i$ which measures the distance between the model-estimated datum and the measured datum for each collimator $i = 1, ..., L$;
- $g_i \ln\left(\frac{g_i}{(Rf)_i}\right) + c_i$ called mute-entropy term.

Equation (3.29) can also be rewritten as the following constrained minimum problem:

$$\arg\min_{f\in S} D_{KL}(g||f) \tag{3.30}$$

where D_{KL} is the Kullback-Leibler divergence, which measures the distance between two distributions.

The Karush-Kuhn-Tucker (KKT) theorem [12] allows to transform the problem (3.30) into a fixed-point problem [67]. This type of problems can be implemented by an iterative optimization algorithm known as Maximum Likelihood Expectation Maximization (MLEM) [28]. Starting from an initialization $f^{(0)}$:

$$f^{(k+1)} = \frac{f^{(k)}}{R^{\#}1} R^{\#} \frac{g}{Rf^{(k)}} \tag{3.31}$$

where 1 denotes a vector of ones having the same dimension as g.

Remark 3.3.3. *With this approach, it is possible to first reconstruct the low-frequency components of the image and then gradually enrich it with details, corresponding to the high-frequency components.*

Remark 3.3.4. *Due to noise affecting g and ill-conditioning of the backprojection operator of the Radon transform $R^{\#}$, it is necessary to find a stopping criterion to determine the optimal number of iterations.*

Remark 3.3.5. *An improvement of the MLEM method is the Ordered Subsets Expectation Maximization (OSEM) algorithm [47]. This method consists of applying the backprojection operator in equation* (3.31) *to only a subset of data, which live in the space of Radon projections: the choice of how the data are partitioned into subsets determines the result of the algorithm. In particular, from an initialization $f^{(0)}$, this algorithm is implemented as:*

$$f^{(k+1)} = \frac{f^{(k)}}{(R|_{S_p})^{\#}1} (R|_{S_p})^{\#} \frac{g}{Rf^{(k)}} \tag{3.32}$$

where S_p for $p = 1, ..., P$ are the different data subsets.

3.4 Physiological semiquantification of PET images: the SUV value

The Standardized Uptake Value (SUV) is one of the most commonly used indices to assess the aggressiveness of a tumor and is mainly used in FDG tracer analysis. This semiquantitative parameter arises from the need to introduce an objective quantity, easy to compute, that supports the visual analysis performed by nuclear medicine physicians. It expresses the amount of radio-tracer absorbed by the region or volume of interest (ROI or VOI) in the analysis time window. The SUV value is computed as:

$$SUV = \frac{Tissue\ Activity}{\frac{Corrected\ Injected\ Activity}{Body\ Mass}} \cdot \frac{1}{ROI\ Volume} \tag{3.33}$$

The unit of measure of the SUV value SUV is $[\frac{g}{ml}]$. Tissue activity and corrected injected activity are measured in $[kBq]$, body mass in $[g]$ and the volume of the ROI (region of interest) in which the SUV parameter is computed in $[ml]$. We observe that if the injected tracer were evenly distributed throughout the body, then the concentration of tracer would be equal to the ratio of administered activity to body mass. Since this is not the case, the SUV value expresses the ratio of the amount of tracer accumulated in a certain ROI to the amount of tracer that would hypothetically be present in a region of equal volume if the tracer were evenly distributed throughout the body. For example, a SUV value greater than 1 indicates preferential accumulation in a particular region.

The denominator in (3.33) is the injected activity corrected for the radioactive decay of FDG. In fact, to make the comparison between the two activities acceptable, it is necessary to estimate the injected activity at the time the tissue activity is measured. Therefore, using the formula that determines the radioactive decay of FDG:

$$CIA = IA \cdot e^{-\lambda t}$$

where $\lambda = \frac{\ln(2)}{\frac{DT}{2}}$ where DT is the FDG radioactive decay time.

Sometimes the SUV value can be presented as a pure number. In fact, if the human body is assumed to be a uniformly distributed volume, i.e. with a constant density equal to 1 $[\frac{g}{ml}]$, then:

$$SUV = \frac{Tissue\ Activity}{\frac{Corrected\ Injected\ Activity}{\frac{Body\ Mass}{Body\ Density}}} \cdot \frac{1}{ROI\ Volume}$$

A ROI-based SUV value can be obtained: the nuclear medicine physician manually draws an ROI and obtains a single SUV value for the entire selected region. On the other hand, it is also possible to obtain a voxel-based SUV value: in this case, a series of images (usually 400×400 in size) are obtained, reporting the corresponding SUV value in each voxel. In this case, the volume of the ROI corresponds to the volume of one voxel. Unexpectedly high SUV values in certain regions may suggest abnormal metabolism and should prompt further diagnostic assessment, such as biopsy.

3.4.1 Limitations of SUV in PET imaging

Evaluation by FDG-PET analysis with SUV value can have several problematic aspects due to both the limited resolution of the PET machine, the nature of the imaging itself, and the presence of biological factors that alter normal blood glucose levels.

The limited spatial resolution of FDG-PET may hinder the detection of small tumors or lesions with inherently low metabolic activity, such as neuroendocrine tumors.

Moreover, not all regions exhibiting high FDG uptake (i.e., high SUV values) can be unequivocally classified as neoplastic. FDG accumulation reflects glucose metabolism and is therefore linked to the presence of viable, metabolically active cells, regardless of whether the process is physiological or pathological. Physiological FDG uptake areas include the brain, larynx, heart, gastric fundus, intestines, kidneys, bladder, and muscles. Pathological uptake can be non-malignant (such as inflammatory lesions and benign neoplasms) or oncological. Inflammatory uptake occurs as a result of the immune system's response to a particular area. Examples include granulomatous inflammatory diseases, which can have very high SUV values. Examples of benign neoplasms include mixed salivary tumors, villous adenomas, and adrenal adenomas.

FDG-PET analysis can be significantly affected by diseases that alter the normal uptake of FDG from the blood. In diabetic conditions, analysis of FDG distribution is often problematic because the limited uptake of glucose and FDG by insulin-dependent tissues parallels the persistence of high tracer concentrations in plasma. Obviously, the ubiquitous presence of radioactive blood attenuates the difference between pathological uptake in cancer and physiological tracer retention in surrounding background tissues. Therefore, it is necessary to find methods to reduce the presence of unabsorbed FDG in the blood. For example, metformin is a widely used oral anti-diabetic drug that reduces blood glucose concentration without causing hypoglycemia [60], mainly by decreasing intestinal glucose absorption and liver glucose output [73]. It follows that PET images from metformin-treated diabetic patients are also problematic to analyze because of the metformin-induced reduction in FDG

de-phosphorylation. Therefore, techniques that can quantify this reduction may provide a significant improvement in uptake correction for metformin-treated diabetic patients [37]. Although SUV is an index that can vary dramatically depending on body metabolism and tissue-specific functions, FDG-PET imaging is becoming increasingly important as a quantitative monitor of individual response to therapy and as a tool for evaluating new drug therapies, and changes in FDG accumulation have been shown to be useful as an imaging biomarker for assessing response to therapy [134]. In fact, in the context of lymphoma, the Deauville criteria have been developed, which take advantage of the SUV value assessed on staging PET and follow-up PET [62], [51]. These criteria use a score scale from 1 to 5 to evaluate the response of tumors to treatment. These criteria are based on a comparison of the SUVmax of pathologic lesions with those found in the liver and aorta. A score of 4 or 5 indicates non-response to treatment or progression.

Chapter 4

Compartmental analysis and solution techniques

In this chapter we explain the concept of compartmental analysis. In particular, we see how it is applied to the problem of reconstructing the dynamics of a tracer with PET acquisition.

Compartmental analysis is a quantitative study of a biological phenomenon. In other words, the system of interest, in which the said biological phenomenon is to be studied at a given time interval, is divided into compartments with a homogeneous functional role with respect to it. Quantitative measurements related to the biological phenomenon are associated with each compartment as time changes. Since the compartments can communicate between each other, there are exchange coefficients, called rate constants, that rule the rate at which the functional state transition occurs from one compartment to another. The goal of compartmental analysis is to estimate the unknown rate constants between compartments from global quantitative measurements throughout the system of interest. From a mathematical point of view, quantitative measurements over time can be viewed as state variables whose dynamics are studied. Assuming that the exchange coefficients between the compartments are constant over time and space, the overall dynamics in the system of interest can be modeled as a system of differential equations expressing the equilibrium principle between the compartments.

In PET application, compartmental analysis is used to study the dynamics of radioactive tracers injected intravenously into the patient's body. In the case of FDG-PET, which is widely used in oncology, such analysis provides a more accurate estimate of cellular metabolic activity in a given region of interest (ROI) than the simple SUV value. In fact, as mentioned in the previous chapter, the SUV value only measures the uptake of the tracer, regardless of whether it is taken up by the tumor or not, without providing precise metabolic information about the neoplastic tissue. In contrast, compartmental analysis allows the study of in vivo

dynamics, i.e. the in vivo study of molecular processes. It provides absolute quantification of parameters within the ROI, i.e. quantitative estimation of physiological parameters related only to the intrinsic characteristics of the system and not to external factors such as tracer concentration or tracer mode. Indeed, in addition to FDG, compartmental analysis can be performed with other types of radioactive tracers that study other forms of metabolism or provide specific links to biomarkers to assess the expression of a particular target molecule.

In the first section we do an overview of the mathematical formulation of a compartmental analysis problem. In the second section we present some techniques (analytical and graphical) that are used to solve its inverse problem.

4.1 Compartmental Analysis: mathematical framework

In this chapter, we present the mathematical modeling of compartmental analysis with application to FDG-PET type data [41, 68]. In this case, the biological phenomenon under study is glucose metabolism. Since these analyses are used in oncology, we want to detect any changes in the physiological levels of glucose metabolism that might be caused by the presence of neoplastic tissue. By varying the type of tracer, different biological phenomena can be studied. When working with non-specific tracers (such as FDG), other forms of metabolism can be studied. When using specific tracers, the goal of the study is to evaluate the expression of a particular target molecule. The system of interest in these applications is represented by a specific organ or anatomical region. Each compartment represents a different metabolic state in which the radioactive tracer is located. The tracer concentrations in the compartments represent the state variables. The exchange coefficients between compartments encode the rate at which the tracer changes state, i.e. moves from one compartment to another: their unit of measurement is the inverse of time. The equations describing the tracer dynamics between compartments are obtained through the tracer balance equations. Since the exchange coefficients are assumed to be constant in time and space, each equation is an ordinary linear differential equation (ODE).

After injection, the tracer reaches the heart, which distributes it to all tissues and organs of the body. How the tracer is sent to all parts of the body, detected at the level of the aorta, plays a very important role in the analysis and is modeled by the so-called input function (IF). 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 [106].

The number of compartments can vary depending on the phenomenon being analyzed. In the following subsections, we show the 2-compartments model (from which the irreversible

Sokoloff model is obtained) and the generic n-compartments model.

4.1.1 Example: 2-compartments model and Sokoloff model

The simplest model, but also the most widely used in oncology applications, is the 2-compartments model. In this model, the system of interest (e.g., an organ or tissue) is divided into two compartments. The first compartment contains the FDG molecules in cells that have not yet been metabolized by the phosphorylation process - their functional state is called "free". The second compartment contains the phosphorylated FDG molecules trapped in the cell cytosol: their functional state is therefore referred to as "metabolized". Accordingly, the tissue compartments included in the model are the "Free" compartment, which we denote by "F", and the "Metabolized" compartment, which we denote by "M". A non-tissue compartment that models the FDG in blood is also included in this model: for this purpose it is denoted as "blood" or "plasma compartment" and it is denoted by "b". Associated with each of these compartments are tracer concentrations over time, $\mathcal{C}_F$, $\mathcal{C}_M$, $\mathcal{C}_b$, respectively. $\mathcal{C}_b(t)$ represents, for each time t, the so-called input function (IF): this function, obtained from the data acquired in the left ventricle at the level of the aorta, shows how FDG is distributed from the heart to each anatomical compartment. The exchange coefficients between the compartments are $k_1, k_2, k_3, k_4 \geq 0$. The coefficients k_1 and k_2 govern the exchange between the blood and the "F" compartment. k_3 and k_4 control the exchange between the "F" and "M" compartments. The graphical representation of this compartment model can be found in figures 4.1. The system of equations expressing the FDG balance between the compartments is a system of ODEs defined as:

$$\begin{cases} \dot{C}_F(t) = -(k_2 + k_3)C_F(t) + k_4 C_M(t) + k_1 C_b(t) \\ \dot{C}_M(t) = k_3 C_F(t) - k_4 C_M(t) \end{cases} \tag{4.1}$$

The initial conditions $C_F(0) = C_M(0) = 0$ complete the Cauchy problem: these conditions are really valid, since at the initial instant the tracer concentration is zero in both the "F" and "M" compartments. This system can be rewritten in a matrix form as:

$$\begin{bmatrix} \dot{C}_F \\ \dot{C}_M \end{bmatrix} = \begin{bmatrix} -(k_2 + k_3) & k_4 \\ k_3 & -k_4 \end{bmatrix} \begin{bmatrix} C_F \\ C_M \end{bmatrix} + \begin{bmatrix} k_1 C_b \\ 0 \end{bmatrix}, \quad C_F(0) = C_M(0) = 0 \tag{4.2}$$

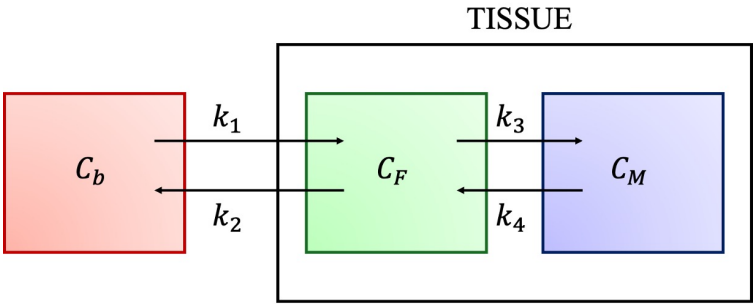

Figure 4.1: Example of a compartmental model with 2-tissue-compartments ("F" in green, "M" in blue), a blood compartment ("b" in red). Arrows indicate the direction of the four exchange coefficients k_1, k_2, k_3, k_4.

The parameter k_1 represents the influx rate of FDG from the blood into the "F" compartment. This depends on both the local blood flow and the rate of tracer extraction from the capillary to the tissue [17]. The physiological measure used is perfusion flow, which represents the volume of blood flowing in and out of a given volume (or weight) of tissue per unit time. Its unit of measurement is $\frac{mL_{blood}/min}{mL_{tissue}}$ or $\frac{mL_{blood}/min}{g_{tissue}}$. In fact, in the physiological literature, the term blood flow usually means perfusion flow. (F=blood flow or perfusion flow). The net tracer flow of a tracer is then defined as:

$$J = J_{in} - J_{out} = FC_a - FC_v \tag{4.3}$$

where the influx is the product between the blood flow F and the arterial concentration (C_a) while the outflux is the product between the blood flow and the venous concentration (C_v).

Define E the fraction of tracer extraction, i.e., the amount of tracer extracted from the blood and delivered to the tissue, as the normalized difference of the arterial-venous concentration

$$E = \frac{C_a - C_v}{C_a} \tag{4.4}$$

(4.3) can be written as

$$J = (F \cdot E)C_a = k_1 C_a \tag{4.5}$$

Equation (4.5) describes the delivery of tracer from blood to tissue. The fraction extraction E has been interpreted ([58, 95, 22]) as

$$E = 1 - e^{-\frac{PS}{F}} \tag{4.6}$$

where P is the permeability of the tracer across the capillary surface (expressed in $[\frac{cm}{min}]$)

and S is the capillary surface area per gram of tissue (expressed in $[\frac{cm^2}{g}]$).

The parameter k_3 represents the rate at which the phosphorylation process occurs, while k_4 controls the rate at which the reverse process of defosphorylation occurs. If $k_4 = 0$, you get the Sokoloff model [115]. This model is widely used in oncology [33]. In fact, it simulates the typical behavior of tumors that irreversibly accumulate FDG-6P in the cytosol of cells [89].

4.1.2 Example: n-compartments model

The n-compartments tissue model case is a direct generalization of the 2-compartments case. Figure 4.2 shows an example of such a model. For each compartment $p = 1, ..., n$ we denote the corresponding tracer concentration by C_p. The exchange coefficient k_{pq} represents the flow of tracer from compartment q to compartment p. The exchange coefficients k_{pb} and k_{bp} represent the flow from the blood to the compartment and from the compartments to the blood, respectively. In this model, we assume that each compartment can communicate with the blood. The equations governing the dynamics of the tracer are still a system of ODEs [17] and are obtained as the balance (or mass conservation) equations of the tracer. For each $p = 1, ..., n$:

$$\dot{C}_p = \sum_{q \neq p} k_{pq} C_q - \sum_{q \neq p} k_{qp} C_p + k_{pb} C_b - k_{bp} C_p \tag{4.7}$$

with initial condition $C_p(0) = 0$. Defining $k_{pp} = -(\sum_{q \neq p} k_{qp} + k_{bp})$ for each $p = 1, ..., n$, the system (4.7) can be rewritten in a matrix form as:

$$\dot{\mathbf{C}} = M\mathbf{C} + \mathbf{W}, \;\; \mathbf{C}(0) = \mathbf{0} \tag{4.8}$$

where

$$\mathbf{C} = \begin{bmatrix} C_1 \\ \cdot \\ \cdot \\ C_n \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} k_{1b} C_b \\ k_{2b} C_b \\ \cdot \\ k_{nb} C_b \end{bmatrix}, \quad M_{pq} = k_{pq}, \; p, q = 1, ..., n \tag{4.9}$$

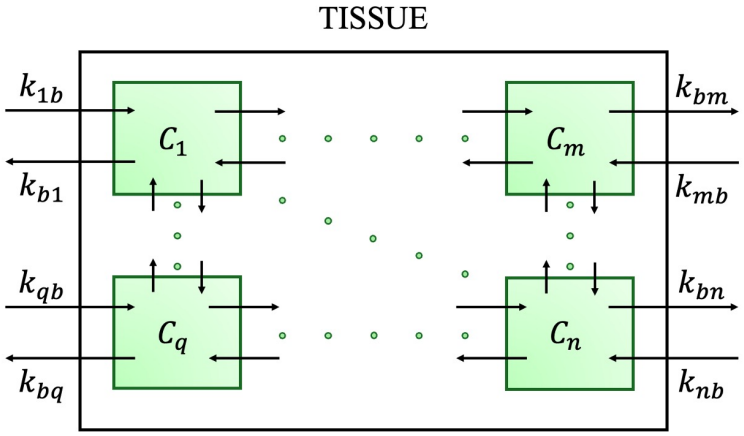

Figure 4.2: Example of a compartmental model with n-tissue compartments. Arrows indicate the direction of the exchange coefficients k_{pq}. In this model each compartment can communicate with blood.

4.2 Compartmental Analysis: the inverse problem

The goal of compartmental analysis, as explained in the introduction of this chapter, is to provide a quantitative estimate of the dynamics of a given tracer through PET acquisitions. Such dynamics, in the most general case, is determined by the system of differential equations (4.8) parameterized by the unknown k exchange coefficients. Reconstructing these coefficients implies the knowledge of the entire model and provides an estimate of tracer dynamics in all compartments. However, PET machine acquisitions cannot measure tracer concentrations in each compartment but simply provide a global measure of tracer concentration throughout the ROI during time. In particular, PET acquisitions occur in the dynamic mode. Photon counts collected during discrete time windows are used to reconstruct tracer concentration in the ROI over time, enabling a dynamic representation of tracer distribution throughout the examination. Assuming that each compartment included in the model contributes linearly and proportionally to its volume relative to the total tracer concentration measured by PET in the ROI, the data formation equation in the ROI is written as:

$$C_T(t) = \alpha \cdot \mathbf{C}(t),\ t \in \mathbb{R}_+ \tag{4.10}$$

where the vector of constants α, whose size is equal to the number of compartments included in the model, contains information about the volume of each compartment. For example, in the case of 2 compartments:

$$C_T(t) = \alpha \cdot \mathbf{C}(t) = (1 - V_b)(C_F(t) + C_M(t)) + V_b C_b(t),\ t \in \mathbb{R}_+ \tag{4.11}$$

where V_b is the volume fraction and represents the volume occupied by compartment "b". In this case, we assume that the compartments "F" and "M" have the same volume equal to $1 - V_b$. The given C_b of the input function is instead directly derived by placing a region of interest in the left ventricle at the level of the aorta.

In summary, the goal of the inverse problem of compartmental analysis is to reconstruct the exchange coefficients k between compartments from C_T, reconstructed from the time scans of PET acquisitions, and the input function C_b.

Every mathematical inverse problem is characterized by its intrinsic property of good or bad posedness in the sense of Hadamard [43]. In this problem, the most significant property is the uniqueness of the parameter set. Given a given n-compartments model and the set of measures C_T (4.10), the question is whether there exists a unique set of parameters k such that solving the forward problem (4.7) with parameters k yields the given C_T. Several results have been obtained in the general n-compartments case but the most important result for the purposes of this thesis work is the following: in the case of the 2-compartments model, the parameters k are uniquely determined by the knowledge of C_T [26].

There are two main types of techniques for reconstructing the k coefficients. One is analytical, which uses optimization methods to numerically reconstruct each k coefficient. It provides an accurate estimate of the coefficients, but is computationally expensive. The other technique is graphical and is also known as Patlak's technique [87]. This technique, which is only applicable under special conditions, is obtained by transforming the system of equations (4.7). It is much less computationally onerous than the analytical technique: however, it allows only a combination of the k parameters to be estimated, not each one. In order to be able to apply Patlak's method under more general conditions than originally envisaged, new techniques were subsequently developed [69, 48] which were still able to estimate a combination of the parameters.

While compartmental analysis provides the best possible estimate of tracer dynamics, it has a strong limitation that makes its practical application difficult: machine occupancy time. In fact, the k coefficients can only be reconstructed if the input function C_b and the tissue concentration function C_T are known at all times from time 0 of tracer injection to late times of about 50-60 minutes. This implies that the PET scanner must be occupied for about 50-60 minutes, which is an extremely long time. Techniques are currently being investigated to reduce the acquisition time while maintaining sufficient information to perform compartmental analysis.

In the following subsections, we present the analytical and graphical techniques for solving the 2-compartments problem of compartmental analysis in the case of FDG-PET analysis. We focus on a particular extension of the graphical Patlak technique, called relative Patlak

[140], which will be exploited in the work presented in Chapter 5.

4.2.1 The analytical Gauss-Newton regularized method

The analytical method consists of a Gauss-Newton optimization algorithm that aims to obtain the exchange parameters from PET scan measurements [32, 41, 54, 59, 132, 138]. In order to define the method, it is necessary to provide the data formation equation. We focus on the 2-tissue-compartments and 4 exchange coefficients k_1, k_2, k_3, k_4 case. Then the system of equations describing its dynamics is (4.1). The analytical solution of this system is:

$$\mathbf{C}(t) = k_1 \int_0^t C_b(u) e^{(t-u)M} \cdot \mathbf{e}_1 du \tag{4.12}$$

where $\mathbf{e}_1 = [1, 0]$.

The measured total concentration C_T in the region of interest is given by (4.11). Since C_b can be measured and V_b is known, we focus our interest on quantity:

$$C_{meas}(t) = \frac{C_T(t) - V_b C_b}{1 - V_b} = C_F(t) + C_M(t) \tag{4.13}$$

Given a certain set of parameters $\mathbf{k}$, for each $t \in \mathbb{R}_+$:

$$C_t(\mathbf{k}) = \alpha^T k_1 \int_0^t C_b(u) e^{(t-u)M} \cdot \mathbf{e}_1 du \tag{4.14}$$

where $\alpha = [1, 1]$, is the sum of the components of $\mathbf{C}(t)$, and it is the forward solution of the system (4.11) written as a function of the parameters $\mathbf{k}$. The inverse compartmental problem is thus reformulated as follows: given a set of $C_{measures}$ concentrations sampled over time, find the set of parameters $\mathbf{k} = \{k_1, k_2, k_3, k_4\}$ such that the corresponding solution $C_\cdot(\mathbf{k})$ (4.14) of (4.1) is compatible with the given measurements C_{meas}.

Like any mathematical inverse problem, it is necessary to study its ill-posedness in the sense of Hadamard [43]. In the literature, the uniqueness of the solution of the inverse problem [37] has been proved by exploiting some properties of the Laplace transform and of rational functions. This result is very important because the lack of uniqueness of the solution would have resulted in the problem being abandoned.

We define the following linear operator:

$$\mathcal{F}_t : \mathbb{R}_+^4 \to \mathcal{C}^{(1)}(\mathbb{R}_+, \mathbb{R})$$

$$(k_1, k_2, k_3, k_4) \to \mathcal{F}_t(\mathbf{k}) = \alpha^T k_1 \int_0^t C_b(u) e^{(t-u)M} \cdot \mathbf{e}_1 du$$

Then, by defining:

$$F_t(\mathbf{k}) = C_t(\mathbf{k}) - \mathcal{F}_t(\mathbf{k}) \tag{4.15}$$

the inverse problem becomes finding the set of parameters $\mathbf{k}$ such that $F_t \simeq 0$ in (4.15) for each $t \in \mathbb{R}_+$. The Frechét derivative in direction $\mathbf{h} \in \mathbb{R}^4$ of (4.15), for every $t \in \mathbb{R}_+$ is:

$$\frac{dF_t}{d\mathbf{k}}(\mathbf{h}) : \ \mathbb{R}^4 \to \mathcal{C}^{(1)}(\mathbb{R}_+, \mathbb{R})$$

$$\mathbf{k} \to \left(\frac{dF_t}{d\mathbf{k}}(\mathbf{h}) \right) (\mathbf{k}) = (\nabla F_t(\mathbf{k})) \cdot \mathbf{h}$$

which is the projection on vector $\mathbf{h}$ of the gradient of $F_t(\cdot)$ with respect to parameters computed in $\mathbf{k}$. Then, given an initial guess $\mathbf{k}^{(0)}$ and exploiting the development in Taylor series, we obtain:

$$F_t(\mathbf{k}^{(0)} + \mathbf{h}) \simeq F_t(\mathbf{k}^{(0)}) + \left(\frac{dF_t}{d\mathbf{k}}(\mathbf{h}) \right) (\mathbf{k}^{(0)}) = F_t(\mathbf{k}^{(0)}) + \nabla F_t(\mathbf{k}^{(0)}) \cdot \mathbf{h}$$

Leveraging a Gauss-Newton approach, since we are searching for the set of parameters that makes the functional $F_\cdot(\mathbf{k}) = 0$, we impose the value of F_t, when we move in the direction and verse of vector $\mathbf{h}$ with step size $\|\mathbf{h}\|_2^2$ from the initial guess $\mathbf{k}^{(0)}$, equal to 0: mathematically $F_t(\mathbf{k}^{(0)} + \mathbf{h}) = 0$. This implies that the vector $\mathbf{h}$ can be obtained by solving:

$$\nabla F_t(\mathbf{k}^{(0)}) \cdot \mathbf{h} = -F_t(\mathbf{k}^{(0)}) \tag{4.16}$$

for each time t. Given $\mathbf{t} = (t_1, ..., t_N)$ the sampling time such that $0 < t_1 < ... < t_N$ we define the following operator A such that

$$A : \mathbb{R}^4 \to \mathbb{R}^{N \times 4}$$

$$\bar{\mathbf{k}} \to \nabla_{\mathbf{k}} F_{\mathbf{t}}(\bar{\mathbf{k}})$$

In particular

$$A_{n,m}(\mathbf{k}) = \frac{\partial F_{t_n}}{\partial k_m}(\mathbf{k}), \quad n = 1, ..., N, \ m = 1, 2, 3, 4$$

Then we define the operator $F_{\mathbf{t}}$ such that

$$F_{\mathbf{t}} : \mathbb{R}^4 \to \mathbb{R}^N$$

$$\mathbf{k} \to F_{\mathbf{t}}(\mathbf{k})$$

Vector $\mathbf{h}$ is obtained by solving the following linear system:

$$A(\mathbf{k}^{(0)}) \cdot \mathbf{h} = -F_{\mathbf{t}}(\mathbf{k}^{(0)}) \tag{4.17}$$

$\mathbf{k}^{(0)}$ is updated as:

$$\mathbf{k}^{(1)} = \mathbf{k}^{(0)} + \mathbf{h} \tag{4.18}$$

This gives rise to an iterative algorithm described in pseudocode form in Algorithm 1.

Algorithm 1 Gauss-Newton optimization method for the 2-compartmental analysis inverse problem

1: Initial guess $\mathbf{k}^{(0)}$;
2: **for** $\ell = 0$ to N_{max} **do**
3: Find $\mathbf{h}$ by solving linear system (4.17) with parameters $\mathbf{k}^{(\ell)}$;
4: Compute $\mathbf{k}^{(\ell+1)} = \mathbf{k}^{(\ell)} + \mathbf{h}$;
5: **end for**
6: **return** $\mathbf{k}^{(N_{max})}$.

Note that the operators A and $F_{\mathbf{t}}$ depend on the measured data, which inevitably contain noise. Therefore, if the number of iterations in the Algorithm 1 is too large, numerical instability problems may occur after a large number of conditioning of the derivation operator. Then, instead of solving the linear system (4.17), the following Tikhonov regularized least squares problem is solved [39]:

$$min_{\mathbf{h}\in\mathbb{R}^4} \|A(\mathbf{k}^{(0)}) \cdot \mathbf{h} - (-F_{\mathbf{t}}(\mathbf{k}^{(0)}))\|_2^2 + \lambda\|\mathbf{h}\|_2^2 \tag{4.19}$$

This restores the numerical stability of the solution.

4.2.2 Patlak graphical method

The Patlak graphical method is an 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 or plane, allowing the estimation of a combination of the exchange coefficients [68]. This method, presented in the original paper [87] with arbitrary tracer and number n of compartments, requires that the kinetic model must satisfy several assumptions:

- the plasma compartment is the only source of tracer for the system;
- the tracer can freely flow from the plasma to a first tissue region and vice versa, and can

directly or indirectly reach any compartment of that region. This is called "exchange" or "reversible";

- the tracer can freely flow from the plasma or "reversible" tissue region to a second tissue region from which it can no longer leave. This region, grouped in a single compartment, is called the "irreversible" or "bound" region.
- tracer in the reversible region has only two paths: it either returns to the plasma or remains stored in the irreversible region.

In the case of the 2-compartments model, given the imposed assumptions, the previously presented Sokoloff model is obtained. Therefore, the system of equations describing this model is defined as follows:

$$\begin{cases} \dot{C}_F(t) = -(k_2 + k_3)C_F(t) + k_1 C_b(t) \\ \dot{C}_M(t) = k_3 C_F(t) \end{cases} \tag{4.20}$$

where the concentrations multiplied by the coefficient k_4 vanish since $k_4 = 0$. The initial conditions are $C_F(0) = C_M(0) = 0$ and it is known that $C_b(0) = 0$.

Standard Patlak plot

Starting with the system (4.20), it is required to:

- explicit C_F in the first equation;
- replace the obtained form of C_F in the second equation and compute the explicit expression of C_M by integrating from time 0 (defined as the time of tracer injection);
- reconstruct C_T as in (4.10) to the left of the equation;
- divide both sides of the equation by C_b.

Then, the following equation is obtained:

$$\frac{C_T(t)}{C_b(t)} = (1 - V_b)\frac{k_1 k_3}{k_2 + k_3}\frac{\int_0^t C_b(\tau)d\tau}{C_b(t)} + (1 - V_b)\frac{k_2}{k_2 + k_3}\frac{C_F(t)}{C_b(t)} + V_b \tag{4.21}$$

For sufficiently long times, the reversible compartment will be in complete equilibrium with the plasma compartment: the limit time beyond which this equilibrium occurs is called the

steady-state time and is denoted by t^*. For $t > t^*$ the ratio $\dfrac{C_F(t)}{C_b(t)} \simeq C^*_{Fb}$ where C^*_{Fb} is constant. Consequently, (4.21) can be rewritten as

$$\frac{C_T(t)}{C_b(t)} = K_i \frac{\int_0^t C_b(\tau)d\tau}{C_b(t)} + b, \qquad t > t^* \tag{4.22}$$

where

$$K_i = (1 - V_b)\frac{k_1 k_3}{k_2 + k_3} \tag{4.23}$$

is called the influx constant and

$$b = (1 - V_b)\frac{k_2}{k_2 + k_3} C^*_{Fb} + V_b \tag{4.24}$$

Let $\mathbf{t} = (t_0, ..., t_M)$ the sampling time be given such that $t_0 = 0$, and $t_0 < t_1 < ... < t_M$ and there exists $0 < m^* < M$ such that $t_{m^*} > t^*$. For each $m \geq m^*$ we define vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{M-m^*+1}$, as:

- $x_{m-m^*} = \dfrac{\sum_{n=1}^{m}(t_n - t_{n-1})C_b(t_n)}{C_b(t_m)} = \dfrac{\sum_{n=1}^{m}\Delta t_n C_b(t_n)}{C_b(t_m)}$;
- $y_{m-m^*} = \dfrac{C_T(t_m)}{C_b(t_m)}$;

for each $m = m^*, ..., M$. Patlak's method consists of plotting the pairs (x_m, y_m) on a Cartesian graph xy and performing a linear regression between them. Specifically

$$\hat{K}, \hat{b} = \arg\min_{K,b} \sum_{m=m^*}^{M} [y_{m-m^*} - K x_{m-m*} - b]^2 \tag{4.25}$$

where

$$\hat{K} = \frac{V^*(\mathbf{x}, \mathbf{y})}{V^*(\mathbf{x}, \mathbf{x})} \tag{4.26}$$

$$\hat{b} = \bar{\mathbf{y}}^* - \hat{K}\bar{\mathbf{x}}^* \tag{4.27}$$

where

$$\bar{\mathbf{x}}^* = \frac{1}{M - m^* + 1}\sum_{m=m^*}^{M} x_{m-m*}, \;\; \bar{\mathbf{y}}^* = \frac{1}{M - m^* + 1}\sum_{m=m^*}^{M} y_{m-m*} \tag{4.28}$$

and

$$V^*(\mathbf{x}, \mathbf{y}) = \frac{1}{M - m^* + 1}\sum_{m=m^*}^{M} x_{m-m*}y_{m-m*} - \bar{\mathbf{x}}^*\bar{\mathbf{y}}^* \tag{4.29}$$

Patlak's regression line can be drawn in the xy Cartesian as:

$$\hat{y} = \hat{K}\mathbf{x} + \hat{b} \tag{4.30}$$

where $\hat{y}$ is the prediction corresponding to $\mathbf{x}$ values.

Therefore, it is observed that the quantities $\hat{K}$ (4.26) and $\hat{b}$ (4.27) are respectively an estimate of the influx constant K_i (4.23) and the constant b (4.24) present in the equation (4.22).

Remark 4.2.1. *Although the quantities are computed for times $t_m > t^*$ knowledge of C_b from time 0 is still required to compute x_m because of the integral in* (4.22).

Remark 4.2.2. *Patlak's analysis does not provide an estimate of all coefficients k_1, k_2, k_3, k_4 of the model, but a combination of them, summarized in the influx constant K_i. What information does K_i provide and why is it important to compute its approximation $\hat{K}$? A very important parameter is the metabolic rate, defined as glucose:*

$$MRGlu = k_3 C_F - k_4 C_M = \phi k_3 C_F, \qquad \phi = 1 - \frac{k_4 C_M}{k_3 C_F} \tag{4.31}$$

This parameter represents the rate of tracer uptake in the irreversible compartment. Under equilibrium conditions $C_F = \dfrac{k_1}{k_2 + \phi k_3} C_b$ where

$$MRGlu = \frac{k_1 k_3 \phi}{k_2 + k_3 \phi} C_b$$

For $k_4 = 0$ we have that $\phi = 1$. Then:

$$MRGlu = \frac{k_1 k_3}{k_2 + k_3} C_b = \frac{K_i}{1 - V_b} C_b$$

The metabolic rate of glucose and the influx constant K_i are thus directly proportional to each other under equilibrium conditions. Calculation of K_i involves estimation of the rate of uptake into the irreversible compartment. In the application to tumors, where $k_4 = 0$, the estimation of Patlak's slope implies an estimation of the tumor voracity.

Instead, the intercept parameter is still under investigation and currently lacks a clear physical interpretation. In fact, it includes, in unknown proportions, tissue blood volume fraction, tissue tracer transport rate constants, and, for a small, usually unidentified fraction, decay-corrected IF [65].

Relative Patlak plot

Unlike the standard Patlak method, the relative Patlak method requires knowledge of all quantities involved in the relation (4.22) only from the steady state time t^* on. This is then generalizable for any time $\bar{t} > t^*$. In the paper presenting the method [140], the authors define the relation:

$$\frac{C_T(t)}{C_b(t)} = K' \frac{\int_{t^*}^{t} C_b(\tau) d\tau}{C_b(t)} + b', \qquad t > t^* \tag{4.32}$$

Note that the lower limit of the integral in the equation (4.32) is the steady state time t^* and not 0 as in (4.22). As previously done for the standard Patlak method, given $\mathbf{t}' = (t'_0, ..., t'_{M'})$ the sampling time such that $t^* < t'_0$ and $t'_0 < t'_1 < ... < t'_{M'}$, we define vectors $\mathbf{x}'$ and $\mathbf{y}'$ as:

- $x'_m = \dfrac{\sum_{n=1}^{m}(t'_n - t'_{n-1})C_b(t'_n)}{C_b(t'_m)} = \dfrac{\sum_{n=1}^{m} \Delta t'_n C_b(t'_n)}{C_b(t'_m)}$;
- $y'_m = \dfrac{C_T(t'_m)}{C_b(t'_m)}$;

for each $m = 1, ..., M$. Linear regression is performed between quantities (x'_m, y'_m) and $\hat{K}'$ and $\hat{b}'$ by solving least squares problem:

$$\hat{K}', \hat{b}' = \arg\min_{K,b} \sum_{m=1}^{M'} [y'_m - K x'_m - b]^2 \tag{4.33}$$

where

$$\hat{K}' = \frac{V(x', y')}{V(x', x')} \tag{4.34}$$

$$\hat{b}' = \bar{y} - \hat{K}' \bar{x}' \tag{4.35}$$

where $\bar{x}' = \dfrac{1}{M'} \sum_{m=1}^{M'} x'_m$ and $V(x', y') = \dfrac{1}{M'} \sum_{m=1}^{M'} x'_m y'_m - \bar{x}' \bar{y}'$. The regression line can be drawn in the xy Cartesian graph simply as:

$$\mathbf{y}' = \hat{K}' \mathbf{x}' + \hat{b}' \tag{4.36}$$

At this point, the aim is to establish a relationship between the quantities $\hat{K}'$ and $\hat{b}'$ in (4.36) and $\hat{K}$ and $\hat{b}$ in (4.26) and (4.27) respectively from standard Patlak method.

The input function C_b, after sufficient time (e.g. $t > t^*$)

$$C_b(t) = a_1 e^{-a_2 t}, \qquad t > t^* \tag{4.37}$$

with $a_1, a_2 > 0$.

Recalling (4.22) and (4.32) we observe that, for $t > t^*$ holds:

$$\frac{C_T(t)}{C_b(t)} = K\frac{\int_0^t C_b(\tau)d\tau}{C_b(t)} + b = K'\frac{\int_{t^*}^t C_b(\tau)d\tau}{C_b(t)} + b', \qquad t > t^*$$

Through analytical and numerical computations it can be shown that:

$$\frac{C_T(t)}{C_b(t)} = (1+\beta)K\frac{\int_{t^*}^t C_b(\tau)d\tau}{C_b(t)} + (K\alpha + b) = K'\frac{\int_{t^*}^t C_b(\tau)d\tau}{C_b(t)} + b', \qquad t > t^*$$

where

$$\alpha = \frac{1}{a_1}\int_0^{t^*} C_b(\tau)d\tau \tag{4.38}$$

$$\beta = \frac{1}{a_1}\int_0^{t^*} C_b(\tau)d\tau \tag{4.39}$$

Then the relations between standard Patlak slope and intercept $\hat{K}$ and $\hat{b}$ in (4.26) and relative Patlak slope and intercept $\hat{K}'$ and $\hat{b}'$ in (4.36) are

$$\hat{K}' = \hat{K}(1+\beta) \tag{4.40}$$

$$\hat{b}' = \hat{K}\alpha + b \tag{4.41}$$

Remark 4.2.3. *Since $\beta > 0$, $\hat{K}$ and $\hat{K}'$ have the same sign.*

Remark 4.2.4. *In the relative Patlak approach data are only known for times $t > t^*$. However, estimating the standard Patlak slope $\hat{K}$ from $\hat{K}'$ requires knowledge of β, which in turn depends on the input function at early time points. This means that the analysis by the relative Patlak method can only be used qualitatively, but does not allow a numerical estimation of the standard Patlak slope $\hat{K}$. However, thanks to the previous remark, it is still possible to make considerations about the physiology of the region of interest.*

4.3 ROI based analysis vs Parametric imaging

In the previous sections, we assumed that the exchange coefficients **k** between the compartments were neither time nor space dependent. The assumption that they do not depend on time is plausible: in fact, the metabolism of the patient's tissues can't change in the short time of the examination. On the other hand, the assumption that the exchange coefficients **k** do not depend on space is too strong. In fact, there is no spatial homogeneity of metabolism in different regions of a tissue.

The estimation of **k** exchange coefficients throughout the region of interest by compartmental analysis is called ROI-based analysis and provides averaged information about metabolic activity in that ROI. The main advantage of this type of analysis is that the noise in the data, normalized to the volume of the whole selected region, is attenuated. However, in many cases, such as oncology or the study of neurodegenerative diseases, where it is necessary to capture differences in metabolism within a region of interest, this type of analysis is no longer sufficient. Thus the technique of parametric imaging is used [131, 42, 34, 49]. This technique consists of generating parametric images, i.e. images of the same size as the input image, but showing for each pixel (or voxel) the value of the metabolic parameter to be analyzed. In the example of 2-compartments analysis with exchange coefficients k_1, k_2, k_3, k_4, the result obtained with parametric imaging consists of 4 static images (or series of dynamical images if a time-resolved parameters estimation is desired), one for each parameter k_i for $i = 1, ..., 4$. With parametric imaging, metabolic inhomogeneities can be captured and the tissue of interest can be analyzed more thoroughly. To perform parametric imaging, the single pixel (or voxel) becomes the region of interest and compartmental analysis is performed for each of them [85].

The main problem in using parametric imaging is the high noise that affects the data. Due to the low resolution of the PET machine and phenomena listed in 3.2.1 voxel-wise analysis is significantly more affected by noise than ROI-based analysis. For this reason, it is much more common to use simpler techniques, such as the graphical Patlak plot technique, to perform voxel-wise compartmental analysis.

Chapter 5

Localized FDG loss in lung cancer lesions

In this chapter we will present the article entitled "Localized FDG loss in lung cancer lesions" that we published on *EJNMMI Research* [85]. This article was developed in collaboration collaboration with the Department of Nuclear Medicine at the Policlinico IRCCS San Martino Hospital in Genoa, Italy.

The aim of this study is to investigate whether (and if so, where) FDG release activity from neoplastic lung tissue occurs after a certain period of time after its injection in patient's body.

Indeed, FDG shares with glucose both GLUT-facilitated transmembrane transport and hexokinase-catalyzed phosphorylation to FDG-6P that, in turn, irreversibly accumulates within the cytosol being a false substrate for the enzymes channeling glucose-6P to glycolysis or pentose phosphate pathway [89]. In agreement with this kinetic model, current guidelines strictly recommend not to shorten the time between tracer injection and scanning below 55 minutes [9, 91, 27]. By contrast, a larger variability is allowed for the maximal duration of this interval, which can be extended by a further 30 minutes beyond the optimal time of one hour [23].

This kinetic model implicitly assumes that FDG-vehiculated radioactivity cannot be lost by cancer lesions. At the time of the introduction of the FDG-PET modality, this concept fitted the notion that glucose-6P-phosphatase (G6Pase) (the enzyme hydrolyzing the sequestered glucose-6P and FDG-6P to the freely exchangeable glucose and FDG) is only expressed in liver, gut, and kidneys [36, 37]. However, more recent literature on glycogen storage diseases discovered the G6Pase isoform β (or G_6PC_3) that is ubiquitously expressed in all tissues [110] and in several tumors, including uterus, lung, breast, and colon cancer [72, 117, 107] as well as glioblastoma [80]. The activity of this enzyme and its capability to

regulate the glucose-6P/glucose ratio is extremely relevant in the maturation and proliferation of mesenchymal cells [112] and in the modulation of the host inflammatory reaction [53, 75, 113].

A large literature now documents that virtually all solid tumors are infiltrated by lymphocytes and macrophages [86]. The high G6Pase expression of these cells and their variegate distribution might thus configure tracer accumulation as a reversible process in some tissue volumes. So far, this hypothesis has never been tested since FDG accumulation rate in cancer has been most often evaluated by analyzing the time-concentration curves of relatively large volumes of interest (VOIs) whose extension did not permit the identification of any possible heterogeneity in tracer kinetics [129]. With the advancement of PET/CT technology, it is now possible to simultaneously define tracer concentrations in both the arterial blood and in the tissues with serial whole-body acquisitions. By analyzing these time activity curves (TACs), we demonstrated that, contrary to commonly accepted models, lung cancer lesions exhibit measurable and heterogeneously distributed tracer loss.

Mathematically, the study exploited the parametric imaging technique (previously explained in Section 4.3) exploiting the compartmental analysis Sokoloff model (4.20). Standard Patlak (4.22) and relative Patlak (4.32) (progressively delaying its starting time) analyses were performed: the results were compared with each other by exploiting observation on the sign of the standard and relative Patlak slope in Remark 4.2.3.

In the following sections methods and results are shown in detail.

5.1 Methods

5.1.1 Patient population and dynamic PET/CT acquisition

The study included 11 patients (9 men and 2 women, mean age 65 years, median 68 years, range 29–88 years) submitted to whole-body PET examination for staging of suspected lung cancer. Imaging was performed in the early morning, after 12 hours of fasting. After measurement of body weight and serum glucose level, an antecubital vein was cannulated, and each patient was positioned on the bed of a Siemens Flow mCT40 system (Siemens, Erlangen, Germany) to undergo the preliminary X-ray CT scanning performed according to the conventional procedure [9]. A list-mode acquisition was started soon before the bolus injection of FDG (4 $\frac{MBq}{Kg}$ body weight) with the field of view focused on the heart for 20 minutes. Immediately thereafter, acquisition mode was shifted, and eight whole body passages were performed from the skull to the mid-tights. In Supplementary Table 1 the scan time (expressed in minutes) per whole body pass is reported. A last equilibrium scan

completed the acquisition procedure to be analyzed for the clinical report.

nScan	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Patient 7	Patient 8	Patient 9	Patient 10	Patient 11
1	2	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2	2	2	2	2
3	2	2	2	2	2	2	2	2	2	2	2
4	2	2	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2	2	2	2	2
6	8	8	8	8	8	8	7	8	8	8	7
7	8	8	8	8	8	8	7	8	8	8	7
8	16	16	17	15	8	8	8	8	8	15	9

Table 5.1: Table of scan times (in minutes) for each patient and scan

Exploited reconstruction algorithm was the PSF+TOF 2i21s that combines an iterative reconstruction considering point spread function (PSF) correction and time-of-flight (TOF) information, with 2 iterations and 21 subsets; voxel size was: $(2.03642 \times 2.03642 \times 5)\ mm^3$.

5.1.2 Image analysis

The chest-centered part of the dynamic acquisition was binned according to the following frame sequence: 12×5 seconds, 12×10 seconds, 8×15 seconds, 6×30 seconds, 2×60 seconds, 5×120 seconds. By contrast, the eight subsequent whole body scans were characterized by a slightly variable time sequence due to the length of the desired field of view. The last whole body acquisition had an average duration of 14 minutes. Accordingly, the acquisition time of each slice was defined based on the DICOM metadata to perform an accurate correction of ^{18}F physical decay. An expert nuclear physician thus identified $a \geq 5\ mL$ VOI on the descending aorta to estimate the input function (IF) defined by the activity concentration in the arterial blood at all times of the dynamic chest-centered frames and in the subsequent whole body acquisitions. A further VOI was drawn to loosely surround the tumor lesion in the last scan, and a mask was created to set all the outside voxels to 0. Data were transformed into SUV images according to the conventional formulation [122] and the cancer lesion was defined as the set of all voxels with radioactivity concentration $> 40\%$ of the maximum value within the identified VOI.

5.1.3 Parametric image analysis

For each patient, two sets of parametric images were set up. For the former, a Time Activity Curve (TAC) was generated in each tumor voxel throughout the eight whole body acquisitions and six regression lines were computed considering all eight frames $(1-8)$, the last seven ones $(2-8)$, and so on, up to the last three $(6-8)$. This analysis provided a

first voxel-resolved description of FDG kinetics, made of a set of six parametric TAC images denoted, from now on, as TAC1−8, TAC2−8, ..., TAC6−8. A regression analysis performed for each parametric TAC image allowed defining the involved voxels as accumulating (with a positive slope of the regression line) and releasing (with a negative slope of the regression line). This preliminary evaluation was then compared with the conventional counterpart represented by the graphical approach to the compartmental analysis described by Patlak et al. [116], which exploits the irreversible Sokoloff model (4.20). Once the equilibrium between blood and reversible compartments is reached at time t^*, called steady state time, the irreversible accumulation within any given voxel can be described by Patlak equation (4.22). A standard Patlak analysis, performed against all the dynamical FDG-PET images, allowed the computation of a parametric series of images whose voxels contain the corresponding standard Patlak slope K_i values (4.26).

However, reproducing the descriptive evaluation reported above, this same analysis was repeated six times, once again considering tumor radioactivity concentration at all frames, in the last seven, up to the last three. For each analysis, the Input Function was estimated by considering the value of the mono-exponential (4.37) function fitting its later measured values. Accordingly, these time resolved Patlak analyses provided six parametric image series, each one reporting the K_i slope of the corresponding regression line, which was denoted as $(K_i)_{1-8}$ up to $(K_i)_{6-8}$. Once again, each voxel was identified as accumulating or releasing according to the sign of the corresponding regression slope.

5.1.4 Statistical analysis

All data are reported as mean±standard error of the mean (SEM). The statistical significance of the regression analysis in the case of the time-resolved approach was assessed by means of $R^2 > 0.3$ for both the TAC parametric images and the Patlak parametric images. In the correlation analysis $p < 0.05$ was considered statistically significant.

5.2 Results

5.2.1 Clinical data

No patient showed metastatic lesions, while homolateral lymph nodes were involved in two patients. Tumor localization and final diagnosis at the histological examination of the harvested lesion is reported in Table 5.2. The same table also reports the volumes of both surrounding VOI and included tumor lesions as well as FDG uptake data. Overall, the final

diagnosis was adenocarcinoma and squamous cell carcinoma in 7 and 4 patients, respectively. Maximal SUV was similar in the two histological types (14.5 ± 10.1 vs. 14.2 ± 8.0, respectively, $p = ns$); by contrast, tumor volume was slightly, though not significantly lower in adenocarcinoma lesions with respect to squamous ones ($20.7 \pm 17.7\ mLs$ vs. $71.3 \pm 75.9 mLs$, respectively, $p = 0.11$).

ID	Disease site	Histology	VOI (mm^3)	TV (mm^3)	TV/VOI (%)	SUV (mean±std)	SUV [min,max]	Injected dose (MBq)
1	left/upper lobe	adenocarcinoma (G2/G3)	25379.7	6676.7	26.3	3.44 ± 0.84	[2.39, 5.98]	351
2	right/upper lobe	adenocarcinoma (G2); lymph nodes involved	156756.8	21834	13.9	5.04 ± 1.13	[3.31, 8.28]	313
3	right medium lobe	squamous cell carcinoma	76636.7	28261.8	36.9	4.97 ± 1.27	[3.19, 7.97]	331
4	left/upper lobe	adenocarcinoma	109481	21108.3	19.3	9.82 ± 2.32	[6.66, 16.5]	319
5	right/upper lobe	adenocarcinoma (G2/G3)	30190.2	3068.8	10.2	8.47 ± 2.16	[5.6, 13.95]	340
6	left/bottom lobe	adenocarcinoma; lymph nodes involved	116447.9	22186.5	19.1	3.42 ± 0.85	[2.27, 5.67]	315
7	right upper lobe	adenocarcinoma	180270.4	56938.4	31.6	18.20 ± 4.55	[12.16, 30.41]	359
8	left/upper lobe	squamous cell carcinoma	526877.2	184189.3	35	12.47 ± 2.45	[7.5, 18.75]	364
9	left/upper lobe	squamous cell carcinoma	323964.1	47172.2	14.6	13.79 ± 3.13	[9.26, 23.13]	320
10	left/upper lobe	adenocarcinoma	101891.9	13000.9	12.8	3.19 ± 0.75	[2.39, 5.98]	341
11	right medium lobe	squamous cell carcinoma	119433.8	25649.2	21.5	3.78 ± 1.00	[2.82, 7.05]	297

Table 5.2: Population description: tumor localization and final diagnosis at the histological examination of the harvested lesion; average volume, average and maximal SUV values of all detected lesions.

5.2.2 Standard and time-resolved Patlak analyses confirm established results at volume level

Figure 5.1 illustrates the results of both the standard (5.1.a) and the time-resolved (5.1.b) Patlak analysis when the slope values of the regression lines are defined for the whole volume lesion. The K_i value provided by the standard procedure is nicely correlated with the average lesion SUV in all patients but one (Patient 8) (5.1.a). Similarly, the agreement between average K_i values and average SUV remained invariant regardless the starting point of the time-resolved regression analysis (5.1.b).

The result of this time-resolved analysis was further corroborated by the trend of average K_i values in the whole population (Figure 5.2.a) or in each subject (Figure 5.2.b) as a function of the starting point of regression analysis. This preliminary analysis confirmed the acknowledged model considering FDG uptake as an irreversible process when the whole lung cancer lesion is analyzed. However, this observation was not confirmed when the spatial resolution of the analysis was escalated to the voxel level (see next Subsection).

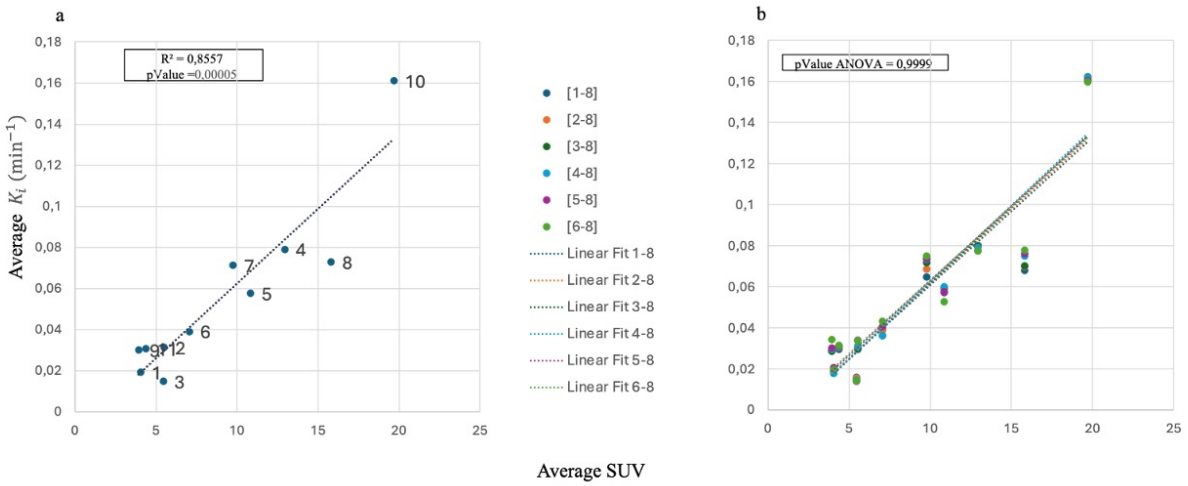

Figure 5.1: Correlation analysis between individual regional SUV and Patlak slope. (a) For each patient (labeled by their ID), the x-axis contains the SUVs averaged across the cancer volume computed at the last scan; the y-axis contains the average values of the standard Patlak slope. (b) For each patient, each colored dot represents the average value of the Patlak slope corresponding to a specific interval in the time-resolved analysis.

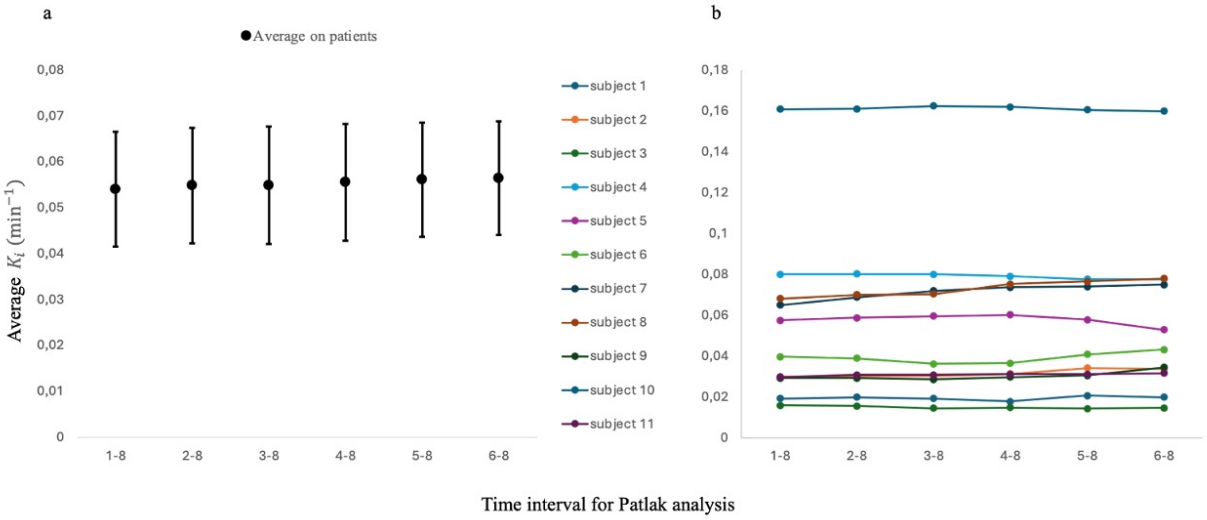

Figure 5.2: Volumetric time-resolved Patlak analysis. (a) For each time frame on the x-axis the values on the y-axis are the volumetric Patlak slopes (mean ± SEM) averaged across subjects. (b) The slopes on the y axis refer to each single subject

5.2.3 Time-resolved Patlak analysis highlights releasing voxels

The analysis of parametric maps with a voxel-wise resolution provided a significantly different picture as shown in Figure 5.3. The distribution of R^2 values (Figure 5.5) indicates that

50% of Patlak analyses achieved an $R^2 > 0.8$. By setting the R^2 threshold at 0.3, only 10% of voxel-wise Patlak analyses were excluded.

The time-resolved Patlak analysis based on progressively delaying the starting point, documented the existence of a significant amount of releasing voxels, i.e., with consistently negative K_i values. Specifically, the later the starting frame considered for the regression line computation, the higher the prevalence of cancer voxels characterized by a negative slope.

Intriguingly, "releasing" voxels were evident even in the 2–8 analysis, with the negative slope of the corresponding regression line remaining invariant when the analysis was restricted to the later images (Figure 5.4.a). In contrast, the slope for the "accumulating voxels" showed a slight but progressive increase (Figure 5.4.b).

As shown in Figure 5.6, the progressive appearance of radioactivity release was not related to an overestimation of tracer concentration in the arterial blood, since the same behavior was documented by the analysis of time-activity curves (TACs) in all voxels included in the cancer lesion. Finally, Figure 5.7 shows the correlations between the rate of releasing voxels and the cancer volume as computed using the static FDG-PET image. This finding suggests that the release of radioactivity was not significantly influenced by contamination from normal tissues, as the extent of the releasing regions was independent of tumor volume. Additionally, the location of the cancer was not displaced in the later images, as the VOI boundaries remained consistent even when the parametric image calculations were delayed, as illustrated by a representative case (Figure 5.8) and confirmed in all patients (Figure 5.9).

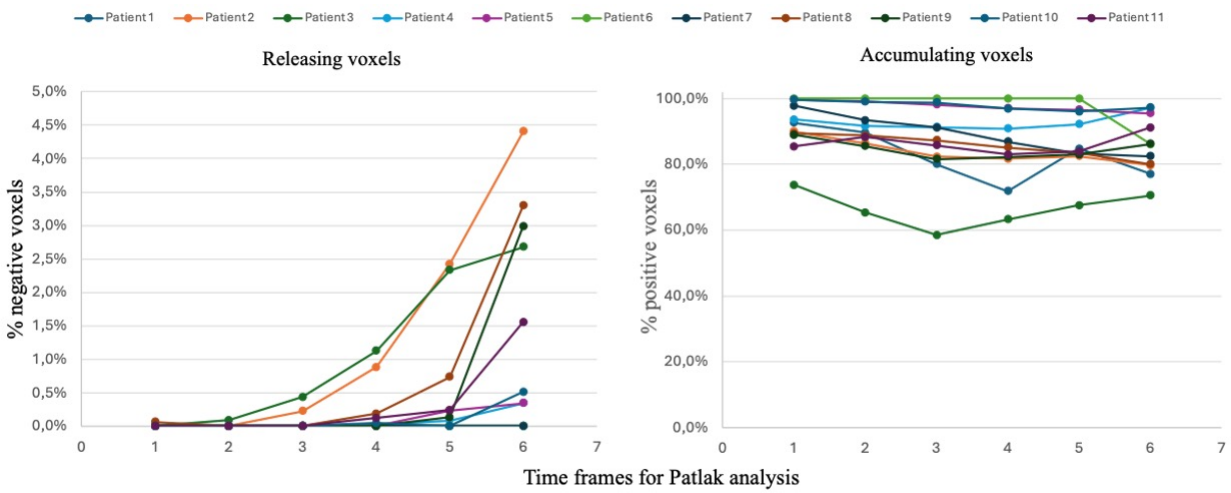

Figure 5.3: Time- and voxel-resolved Patlak analysis. The two plots show, for each patient and each time frame, the rate of voxels with negative (a) and positive (b) Patlak slope. Rates are computed with respect to the total number of voxels in the lesion

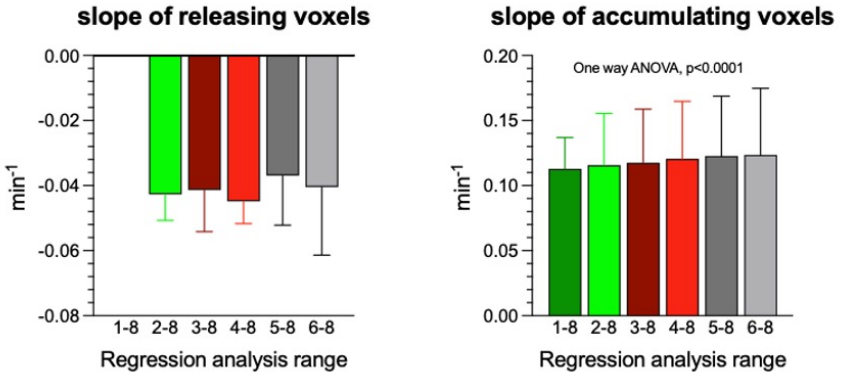

Figure 5.4: Distribution of regression line time- and voxel- resolved Patlak analysis slope in releasing (a) and accumulating voxels (b)

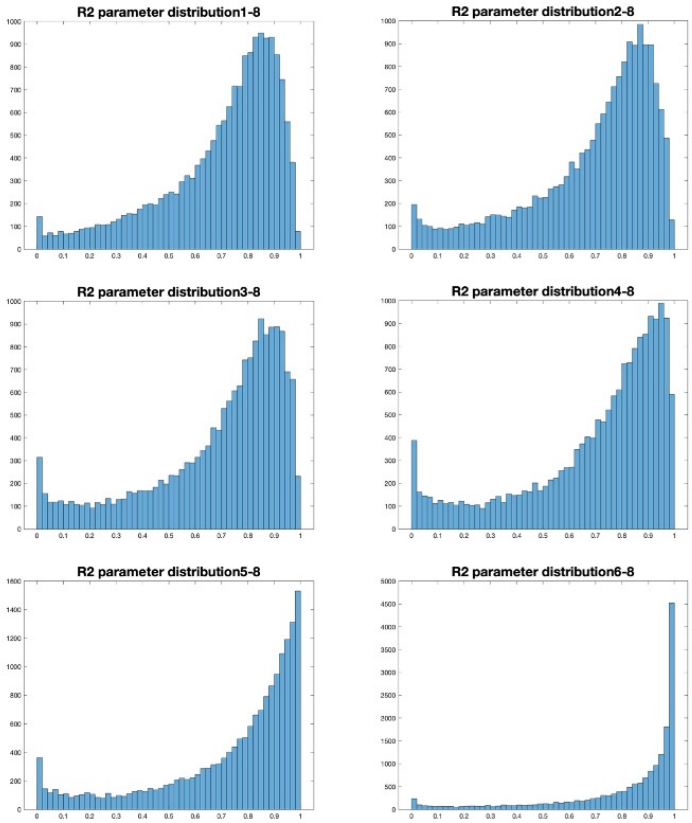

Figure 5.5: Histograms of the distribution of R^2 parameter for the different analyses.

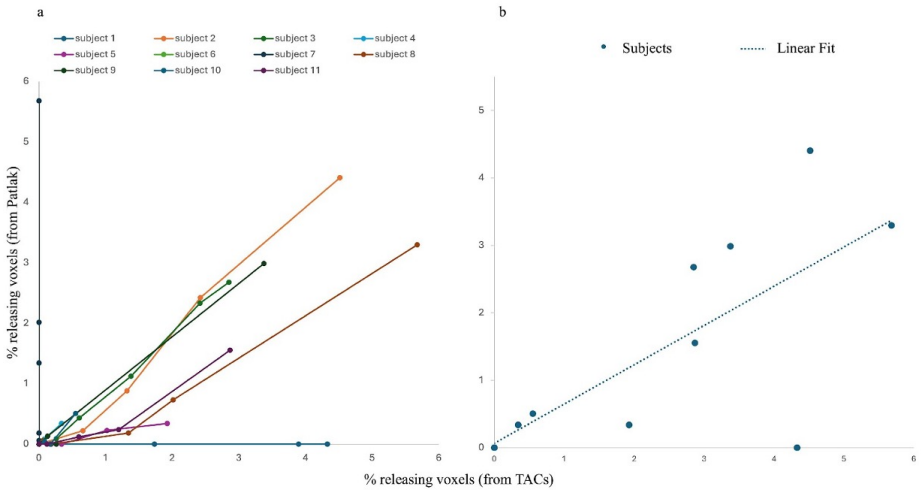

Figure 5.6: Comparison between the rate of releasing voxels computed by using the TAC slopes (x-axis) and the Patlak slopes (y-axis). (a) Each color corresponds to a subject and each point corresponds to a time frame. (b) Focus on frame 6–8, where each point corresponds to a patient

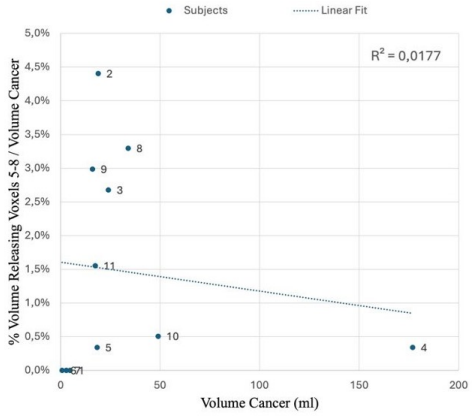

Figure 5.7: The rates of releasing voxels as computed from Patlak at interval 5–8, plotted against the cancer volume (in ml). Each patient is a dot marker. The dotted line is the linear trend on the subjects

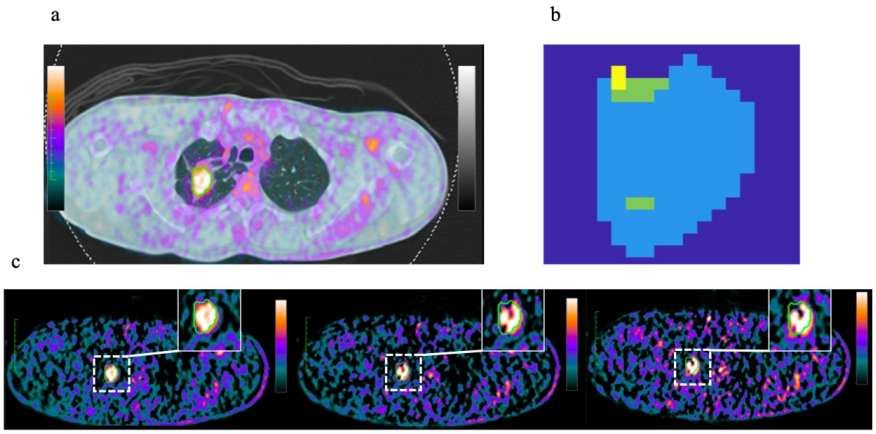

Figure 5.8: Spatial coherence of voxels. (a) WB STATIC PET tumor image coregistered with CT. (b) Parametric map showing spatial coherence of voxels that exhibited release activity in consecutive frames up to frame 6–8. The color scale indicates that release was detected in the last four (yellow) or three (green) voxels within the tumor volume (pale blue). (c) Parametric maps generated with Patlak analysis considering the interval 1–8, 3–8, and 5–8. For each image, the tumor area is identified by a square (dashed contour) and reported in its zoomed version (solid line)

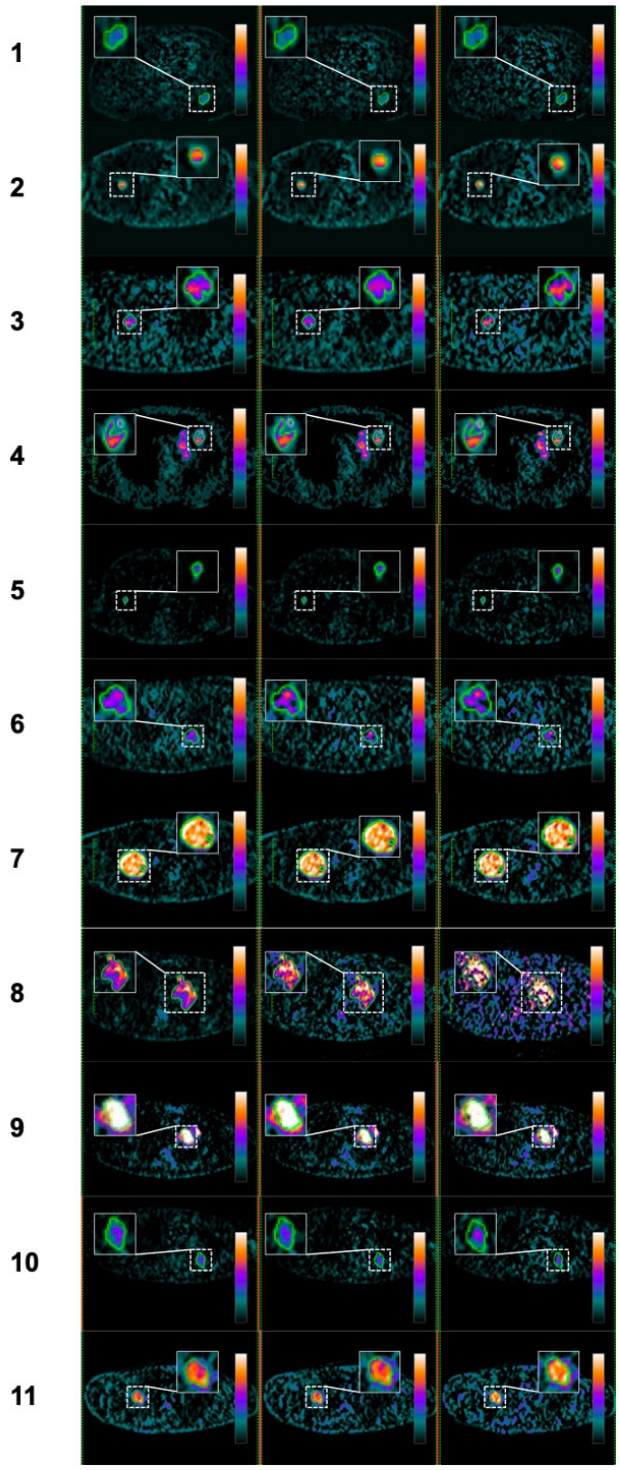

Figure 5.9: Parametric maps generated with Patlak analysis considering the frames from 1-8, from 3-8 and from 5-8, in all 11 studied patients. For each image, the tumor area is identified by a square (dashed contour) and reported in its zoomed version (solid line). The spatial coherence of the images is evident.

5.2.4 Localization of the accumulating and releasing voxels

To document whether tracer kinetics was clustered in specific cancer regions, Figure 5.10 shows the voxels representing tumor boundaries, which were identified as all voxels with faces, edges, or corners adjacent to non-cancer ones (with SUVs $< 40\%$ of maximum). By contrast, the core region was defined as the complementary tumor volume.

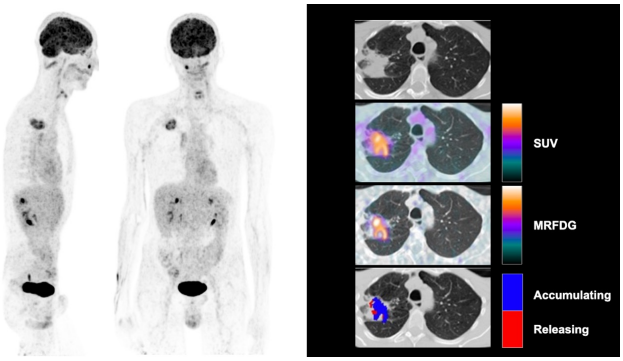

Figure 5.10: Spatial distribution of releasing and accumulating voxels coregistered with the CT scan of one of the analyzed patients. MRFDG indicates the accumulation rate of FDG K_i.

5.3 Discussion

In the present study, the conventional Patlak analysis showed heterogeneous FDG kinetics in treatment-naïve lung cancer. The tracer irreversibly accumulated in the largest part of the lesion volume; however, it was retained in a reversible pool for a measurable number of voxels. This behavior reflected a true tracer washout, as documented by the progressive decrease of the ratio between tumor and blood radioactivity concentration that ruled out significant contamination by the blood volume included in the analyzed voxel [128]. It was independent of the lesion size and relatively more represented in the border zone with respect to the lesion core. Altogether, these findings indicate that the well-documented heterogeneity of cell types populating the tumor eventually results in a heterogeneous kinetics of tracer accumulation as a possible index of a local infiltration by FDG-releasing inflammatory cells.

5.3.1 Spatial resolution and FDG kinetics

As requested by the standard procedure, the Patlak regression line was computed setting the X values according to the full-time input function, i.e., from the radio tracer injection time until the dynamic scan is recorded, and plotting the Y values after the steady state

has been reached. Due to the irreversible nature of FDG uptake, the estimated regression line in each voxel should have maintained constant values for both the slope and intercept, regardless of the starting time of the analysis (see remark 4.2.3). This tenet was not respected in those lesion voxels that showed a measurable and progressive decrease in tumor/ blood radioactivity ratio.

Due to its late onset and relatively low prevalence, the effect of tracer washout on the Patlak relationship was overshadowed by the accumulation of radioactivity in the earlier time points. Consequently, the analysis of average FDG kinetics in the entire cancer lesion showed a strong linear relationship, consistent with findings from studies using this VOI-based approach [88]. While applying the generalized Patlak model [56] could have provided a more accurate estimation of the release rate (usually defined as K_{loss}) [56], this was not feasible in our study due to the limited number of late time points (eight) available for the late phase of the tumor time-activity curve. Nevertheless, our data support the effectiveness of parametric images in highlighting the heterogeneity of cancer metabolic patterns. Furthermore, the progressive enlargement of the releasing volume gradually reduced the number of "accumulating" voxels, resulting in a small but measurable increase in their K_i.

To observe the washout phenomena, it was necessary to retain linear regressions that slightly deviated from the Patlak assumption of tissue irreversibility. The $R^2 > 0.3$ threshold ensured the validity of the analysis by excluding only 10% of the voxels in each analysis that exhibited a strongly non-linear trend (5.5). Given the small cohort and the limited number of frames available, we chose not to model these non-linear effects but to provide a phenomenological description of localized FDG-loss.

The heterogeneous tracer handling of tumor tissue documented by escalating the spatial resolution down to the voxel size closely agrees with the well documented heterogeneity of cell type populating the lesion [74, 2], and with the relatively preferential localization of inflammatory cells in proximity of tumor borders [118]. Indeed, while it is universally accepted that cancer cells display an irreversible FDG accumulation [108, 101], a similar consensus also applies to granulocytes and macrophages, whose tracer-releasing feature has been frequently proposed as a possible tool to differentiate cancer from inflammation [105].

5.3.2 Clinical implications

A small fraction of cancer volume showed a measurable FDG washout. This finding suggests that the late tracer release should scarcely affect image quality even when acquisition is postponed beyond the 50 minutes indicated by the current guidelines [9, 91, 27]. Nevertheless, this evidence might improve the capability of PET/CT in verifying the effectiveness of

treatments, such as immunotherapy, aiming to improve the host inflammatory response to cancer. Several studies report a significant increase of compartmental analysis in attributing the tracer uptake to the inflammatory infiltrates promoted by immunotherapy vs the progression of the disease [120, 130]. However, the complexity and time-consuming nature of dynamic imaging so far hampered the clinical exploitation of this approach. Nevertheless, setting up voxel-resolved parametric maps delineating the trend of tumor/blood radioactivity ratio at the late time typical of tracer washout might represent a potential window to render this analysis feasible in the routine practice. Indeed, the accuracy of this approach in providing an accurate K_i value is still matter of debate [21], its capability to identify a negative slope remaining valid when the input function is not fully determined when limited to its latest time points [140]. Obviously further studies are needed to characterize the capability of this index to identify the localization of inflammatory infiltrates.

5.3.3 Limitations

Several limitations of this study should be considered. Our study focused on lung cancer, aiming to simultaneously define the time-concentration curve in both the arterial blood and the tumor lesion. This focus inevitably affects image quality due to respiratory motion. However, the blurring caused by this motion should have uniformly affected all images after reaching a steady state, given the relatively long and constant acquisition time. Furthermore, tracer washout involved a progressively enlarging volume within the same cancer tissue location, and the automatic definition of cancer remained stable across different Patlak analyses (Figure 5.9). Therefore, chest motion should not be regarded as the sole mechanism behind the observed tracer release. On the other side, a second limitation refers to the missing analysis of cell populations and their difference between releasing and accumulating voxels. Due to the high complexity of this evaluation, this task was not attempted, although stroma elements were obviously represented in all harvested lesions. Nevertheless, this limitation only partially hampers the nature of our proof- of-concept study aimed to identify the potential capability of voxel-resolved images to recognize heterogeneous tracer kinetics within the cancer tissue.

5.4 Conclusions

In the present study, delaying the starting point of Patlak graphical analysis indicates a heterogeneous FDG kinetics within lung cancer lesions. This evidence can only be obtained by voxel-wise parametric maps limited to the analysis of the latest time points of dynamic acqui-

sition. This finding indicates that expanding the analysis of FDG uptake in the time domain might improve the informative content of PET/CT imaging adopting protocols compatible with the daily nuclear medicine practice after the introduction of novel scanners characterized by large field of view and high spatial resolution. The potential of this time-resolved approach to FDG imaging might be of relevance for the evaluation and monitoring of therapy effectiveness, mostly in patients exposed to immunotherapy in whom differentiating the inflammatory enhancement versus the progression of disease requests, so far, pronged and serial evaluations.

Chapter 6

The latent dynamics discovery inverse problem

In the previous chapters, we showed that the problem of reconstructing latent dynamics from time-series measurements is an extremely recurring problem in the world of nuclear medicine. This problem arises not only when working with known tracers, but also when testing new tracers. When the structure of the dynamic model is already known, it is sufficient to estimate the exchange coefficients between the different compartments. However, if the structure of the system is not known, the problem is greatly complicated. Moreover, the variation of the tracer concentration within each compartment could depend, even nonlinearly, on the concentrations in the other compartments. For this reason, a more general dynamics model can be expressed in the form:

$$\dot{C}_h = \sum_{q=0}^{n} \sum_{j=0}^{D-1} m_{h,q,j} \phi_j(C_q) - \sum_{q=0}^{n} \sum_{j=0}^{D-1} m_{q,h,j} \phi_j(C_h) \tag{6.1}$$

for $h = 1, ..., n$, where n denotes the maximum number of tissue compartments, and $q = 0$ refers to the blood compartment concentration C_b. $\phi_j : \mathbb{R} \to \mathbb{R}$, for $j = 0, ..., D-1$ are functions which can capture the potential nonlinear dependence of the dynamics in a given compartment on the concentrations in all the different compartments included in the model. For example some of these ϕ_j can be trigonometric or polynomial functions. $m_{h,q,j}$ and $m_{q,h,j}$, for each $h = 1, ..., n$, $q = 0, ..., n$ and $j = 0, ..., D-1$ are the exchange coefficients between compartment h and compartment q with the dependence described by function ϕ_j. In this way, for example, if $m_{\hat{h},\hat{q},j} = 0$ for every $j = 0, ..., D-1$, then $\hat{q}$ compartment does not contribute to the supply of tracer in the compartment $\hat{h}$.

More generally there are many critical data-driven problems, such as understanding cog-

nition from neural recordings, inferring climate patterns, determining stability of financial markets, predicting and suppressing the spread of disease, and controlling turbulence for greener transportation and energy [14]. Estimating their dynamics also implies the possibility of carrying out monitoring and forecasting activities.

The mathematical models that describe these type of dynamics are generally systems of ordinary differential equations (ODE). These differential equations describe the evolution over time of changes in the physical quantity of interest. Latent dynamics discovery problems consist of reconstructing systems of differential equations from temporal samples of the physical quantities involved. In this mathematical framework, they are also called ODE learning problems.

This chapter is structured as follows:

- a formal description of the mathematical framework for the latent dynamics discovery problem is given;
- an overview of the main methods used to solve ODE learning problems is provided.

6.1 A mathematical framework for ODE learning

Latent dynamics discovery problems are usually mathematically formalized in terms of ordinary differential equations (ODE) learning problems. In literature, ODE systems in this application are usually written as follows:

$$\dot{u} = \frac{du}{dt} = f(u) \tag{6.2}$$

where u is the time-depending continuous vector-valued state variable such that $u(t) \in \mathbb{R}^U$, for $t > 0$, and f represents the completely unknown time derivative function. In (6.2) the time derivative of state variable u is made explicit: nevertheless, the problem is still written in implicit form because the state variable u appears in both members of the equation.

Data are generally collected in a matrix $\mathbf{d} \in \mathbb{R}^{T \times U}$ which contains the set of measured data where $\mathbf{d} = [\mathbf{d}_0, ..., \mathbf{d}_{T-1}]$ such that $\mathbf{d}_k \in \mathbb{R}^U$ for each $k = 0, ..., T-1$. Data are often noisy. Noise is usually additive and is expressed as a random variable: each set of measured data corresponds to a different realization of noise. Given sampling time points $[t_0, ..., t_{T-1}] = \mathbf{t} \in \mathbb{R}^T$ with $t_0 < t_1 < ... < t_{T-1}$, data can be modeled as:

$$\mathbf{d}_k = \mathbf{u}_k + \boldsymbol{\eta} \tag{6.3}$$

where $\mathbf{u} \in \mathbb{R}^{T \times U}$ is such that $\mathbf{u}_{k,h} = u_h(t_k)$ for each $k = 0, ..., T-1$ and $h = 0, ..., U-1$,

and $\boldsymbol{\eta}$ is a realization of the random variable describing the noise. In general it is assumed that data $\mathbf{d}$ are affected by white Gaussian noise. Then $\boldsymbol{\eta} \sim N(\mathbf{0}, \Sigma)$ with Σ the covariance matrix.

The ODE learning problem consists of reconstructing the unknown time derivative function f from noisy measured data $\mathbf{d}$.

In the case of dynamics reconstruction problems applied to compartmental analysis on PET data (6.1), the radioactive tracer concentrations in each compartment represent a state variable of the ODE system. Again, the derivative function f is unknown but its parametric form is known since it is a linear combination of the concentrations over time in the different compartments weighted by their respective rate constants. Therefore, the problem of reconstructing the derivative function f is equivalent to estimating the exchange coefficients between compartments from the PET data. However, as noted in Chapter 4 on compartmental analysis, the model describing obtained PET data is

$$C_T(t) = (1 - V_b) \sum_{i=1}^{N} C_i(t) + V_b \sum_{i=1}^{N} C_{ie} \tag{6.4}$$

that is, it is a linear combination of the concentrations in the compartments and the input function of each compartment. This greatly complicates the dynamics estimation problem: one would have to estimate individual concentrations over time, knowing only a linear combination of them. This is a strongly ill-posed problem, since it admits infinite solutions. However, by introducing a priori knowledge of the coefficients, which can be mathematically translated into constraints on their ranges, the ill-posedness of the problem improves and it becomes possible to obtain reliable solutions.

6.2 Existing methods

This section reviews some of the best-known methods used to solve latent dynamics problems in a data-driven way. In particular the following methods will be exposed:

- Sparse Identification of Nonlinear Dynamics (SINDy) [13];
- Physics Informed Neural Networks (PINN) [20];
- Learning unknown ODE models with Gaussian processes [46].

6.2.1 Sparse Identification of Nonlinear Dynamics (SINDy)

The unknown dynamics discovery problem is approached with a sparse regression perspective ([123], [45], [52]). This method is based on the assumption that a physical phenomenon is formally simple to describe. Consequently, since the governing equations contain only a few relevant terms, they can be viewed as sparse in a high-dimensional nonlinear function space. From these observations, using (6.2) as the model describing the dynamics of the system, it follows that the derivative function f consists of only a few terms. By combining convex methods with sparsity, the resulting model is a good compromise between accuracy and complexity, avoiding its overfitting to the data.
Measured data and their measured derivatives are arranged matrices $\mathbf{d}, \dot{\mathbf{d}} \in \mathbb{R}^{T \times U}$ respectively:

$$\mathbf{d} = \begin{bmatrix} d_{1,1} & d_{1,2} & ... & d_{1,U} \\ d_{2,1} & d_{2,2} & ... & d_{2,U} \\ ... & ... & ... & ... \\ d_{T,1} & d_{T,2} & ... & d_{T,U} \end{bmatrix} \qquad \dot{\mathbf{d}} = \begin{bmatrix} \dot{d}_{1,1} & \dot{d}_{1,2} & ... & \dot{d}_{1,U} \\ \dot{d}_{2,1} & \dot{d}_{2,2} & ... & \dot{d}_{2,U} \\ ... & ... & ... & ... \\ \dot{d}_{T,1} & \dot{d}_{T,2} & ... & \dot{d}_{T,U} \end{bmatrix} \tag{6.5}$$

Then, a library $\Theta(\mathbf{d})$ of nonlinear candidate functions to describe the dynamics, evaluated in the columns of $\mathbf{d}$, is constructed. For example:

$$\Theta(\mathbf{d}) = \begin{bmatrix} | & | & | & | & | & | & | & | \\ \mathbf{1} & \mathbf{d} & \mathbf{d}^2 & \mathbf{d}^p & ... & \sin(\mathbf{d}) & \cos(\mathbf{d}) & ... \\ | & | & | & | & | & | & | & | \end{bmatrix} \tag{6.6}$$

where each column of $\Theta(\mathbf{d})$ is a $T \times U$ matrix. For example:

$$\mathbf{d}^2 = \begin{bmatrix} d_{1,1}^2 & d_{1,2}^2 & ... & d_{1,U}^2 \\ d_{2,1}^2 & d_{2,2}^2 & ... & d_{2,U}^2 \\ ... & ... & ... & ... \\ d_{T,1}^2 & d_{T,2}^2 & ... & d_{T,U}^2 \end{bmatrix}$$

Functions included in $\Theta(\mathbf{d})$ are called basis functions. In this example, polynomial and trigonometric functions are employed, but other types of functions may also be considered. The total number of basis functions in $\Theta(\mathbf{d})$ is denoted by D. Then, the dynamics can be modeled as:

$$\dot{\mathbf{d}} = \Theta(\mathbf{d})\Xi \tag{6.7}$$

where $\Xi \in \mathbb{R}^{(U \times D) \times U}$ is the matrix containing unknown coefficients $\Xi = [\xi_0, ..., \xi_{U-1}]$ where $\xi_h \in \mathbb{R}^{(U \times D) \times 1}$ is a column vector for each $h = 0, ..., U - 1$. Since it has been assumed

that the physical phenomenon under study is formally simple to describe and governing equations contain only a few relevant terms, sparsity has to be imposed on these coefficients. Therefore, a sparse regression problem is set up: the non-zero coefficients determine which basis functions remain in the model and thus describe the phenomenon. Each column of (6.7)

$$\dot{\mathbf{d}}_h = \Theta(\mathbf{d}^T)\xi_h \tag{6.8}$$

for $h = 0, ..., U-1$, is optimized independently in order to find which terms of $\Theta(\mathbf{d})$ intervene in the h^{th} component of state variable.

Since, in general only $\mathbf{d}$ is available, $\dot{\mathbf{d}}$ is numerically approximated. This is not a well-conditioned operation since data are always affected by noise. In fact, data $\mathbf{d}$ and their numerical derivatives $\dot{\mathbf{d}}$ are generally filtered: for example total variation regularization [99] is exploited to denoise the data derivative [18].

Then (6.7) can be more realistically written as

$$\dot{\mathbf{d}} = \Theta(\mathbf{d})\Xi + \mathbf{Z} \tag{6.9}$$

where $\mathbf{Z}$ is additive white noise modeled as a matrix of independent and identically distributed Gaussian entries.

The optimization algorithm exploits Least Absolute Shrinkage and Selection Operator (LASSO) [123, 45] technique which is based on ℓ_1-regularization.

A schematic representation of the SINDy algorithm applied to the Lorenz equations is shown in Figure 6.1.

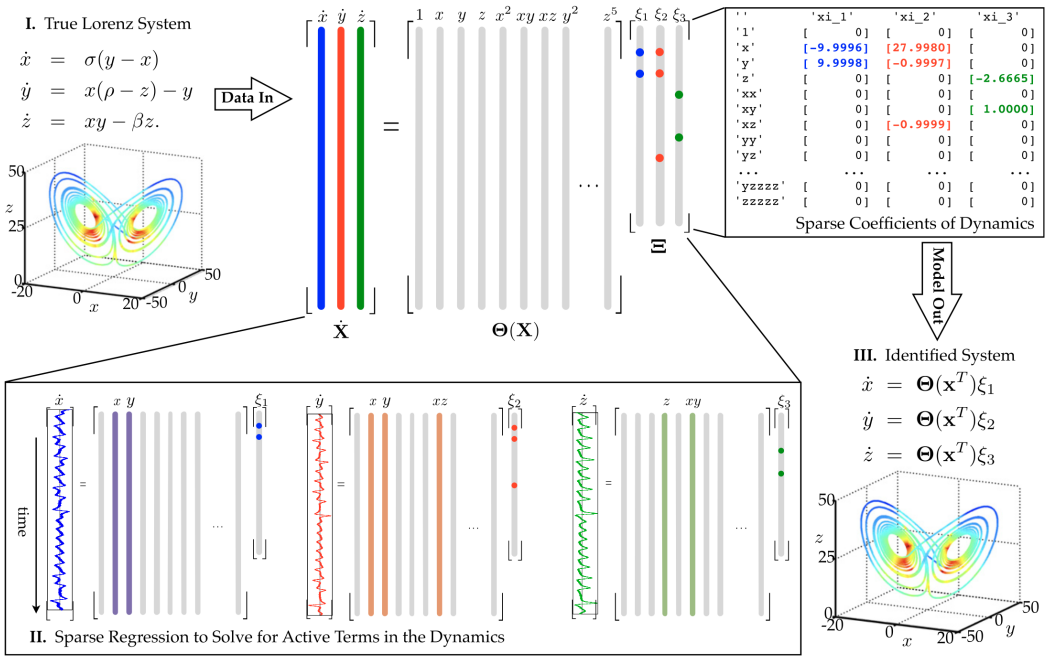

Figure 6.1: SINDy algorithm scheme applied to the Lorenz equations (image credits [13]).

An important limitation is the choice of basis functions. If the set Θ does not contain the correct functions, i.e., those that actually describe the dynamics, the algorithm will fail to identify a sparse model. As a result, the resulting model is severely over fitted, with spurious components in many entries of the Ξ matrix. In practice, this problem can be circumvented by combining the algorithm with a priori knowledge of the underlying physics of the phenomenon under study.

Another important limitation of SINDy algorithm is its strong dependence on the quality and quantity of the measured data: the numerical differentiation required to compute the derivatives to construct the governing equations is a mathematically strong ill-posed operation.

6.2.2 Physics Informed Neural Networks (PINN)

The technique described in this subsection is called PINN-SR, which stands for Physics Informed Neural Networks with Sparse Regression. It exploits neural networks combined with ℓ^0 sparse regression and automatic differentiation techniques. This is a data-driven method and allows the discovery of governing differential equations for nonlinear spatiotemporal systems that can be mathematically modeled by systems of Partial Differential Equations (PDEs), and can thus be seen as an extension of standard ODE learning techniques. Based on works inspired by SINDy, [100, 102] works based on implicit learning of weak-form PDE

models [93] and other studies [6, 10], PINN-SR algorithm aims to obtain closed-form PDEs in a constrained search space with a predefined library of PDE terms.

The governing equations can be written as:

$$\dot{u} + \mathcal{F}[u, u^2, ..., \nabla u, \nabla^2 u, ...; \boldsymbol{\lambda}] = p \tag{6.10}$$

where u, called state variable, is a U-dimensional time-depending continuous function such that $u(\mathbf{x}, t) \in \mathbb{R}^U$ for each $\mathbf{x} \in \Omega$, the space of definition, and $t \in [0, T]$. $\mathcal{F}$ is the unknown non linear time derivative function parametrized by $\boldsymbol{\lambda}$ and ∇ is the gradient operator with respect to space variable $\mathbf{x}$. The term on the right hand side of the equation p is the input source term (in common cases $p = 0$). Spatiotemporal measurements of state variable u are known. The inverse problem is to reconstruct the functional $\mathcal{F}$ in (6.10) from the measurements of the state variables u. Taking inspiration from the SINDy algorithm [13], it is reasonable to assume that a physical phenomenon is simple to describe: therefore, the equation (6.10) can be re-written in a sparse representation:

$$\dot{u} = \Phi\boldsymbol{\Lambda} \tag{6.11}$$

where $\Phi = \{1, u, u^2, ..., u_x, u_y, ..., \sin(u), \cos(u), ..\}$ is a library of symbolic functions, called basis functions, which are possible candidates to describe the unknown dynamics. The cardinality of Φ is D. $\boldsymbol{\Lambda} \in D \times U$ is the coefficient matrix. If present, the input source term p can be included in Φ. PDE initial and boundary conditions are denoted by $\mathcal{I}[\mathbf{x} \in \Omega, t = 0; u, \dot{u}] = 0$ and $\mathcal{B}[x \in \partial\Omega; u, \nabla u] = 0$ respectively. These Initial or Boundary conditions are unnecessarily either known a priori or measured since sampled data already reflects the specific I/BC.

Figure 6.2 shows the architecture of this method. The u state functions are modeled using a neural network: $\mathbf{u}^{\boldsymbol{\theta}} = \mathbf{u}(\mathbf{x}, t; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ represents the DNN trainable parameters including weights and biases. By calling $\mathcal{D}_u = \{\mathbf{u}^m | m = 1, ..., N_m\}$ the set of u spatiotemporal measurements, it is possible to define the data loss function

$$\mathcal{L}_d(\boldsymbol{\theta}; \mathcal{D}_u) = \frac{1}{N_m} \|\mathbf{u}^{\boldsymbol{\theta}} - \mathbf{u}^m\|_2^2 \tag{6.12}$$

Using the automatic differentiation technique, it is possible to compute the time derivative $\dot{\mathbf{u}}^{\boldsymbol{\theta}}$ and obtain all candidate functions contained in the set Φ, then $\boldsymbol{\Phi}^{\boldsymbol{\theta}}$. System (6.10) can be re-written in a residual form as

$$\mathcal{R}^{\boldsymbol{\theta}} := \dot{\mathbf{u}}^{\boldsymbol{\theta}} - \boldsymbol{\Phi}^{\boldsymbol{\theta}}\boldsymbol{\Lambda} \tag{6.13}$$

where $\mathcal{R}^{\boldsymbol{\theta}}$ are denoted as the PDE residuals. This residual defines how well the function $\mathbf{u}^{\boldsymbol{\theta}}$ defined by the neural network parameterized by $\boldsymbol{\theta}$ satisfies the PDE system (6.11) parameterized by $\boldsymbol{\Lambda}$.

PDE residuals (6.13) are computed in a set of random collocation points $\mathcal{D}_c = \{(\mathbf{x}_i, t_i) | \mathbf{x}_i \in \Omega, t_i \in [0, T], i = 1, ..., N_c\}$. Then, it is possible to define the physics loss function

$$\mathcal{L}_p(\boldsymbol{\theta}, \boldsymbol{\Lambda}; \mathcal{D}_c) = \frac{1}{N_c} \|\dot{\mathbf{u}}^{\boldsymbol{\theta}} - \boldsymbol{\Phi}^{\boldsymbol{\theta}} \boldsymbol{\Lambda}\|_2^2 \tag{6.14}$$

The method consists of tuning the parameters $\boldsymbol{\theta}$ of the neural network $\mathbf{u}^{\boldsymbol{\theta}}$ and $\boldsymbol{\Lambda}$ of the PDE system in such a way that $\mathbf{u}^{\boldsymbol{\theta}}$ approximates well the measured data and satisfies the PDE defined by the parameters $\boldsymbol{\Lambda}$.

The total loss function is defined as

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}; \mathcal{D}_u, \mathcal{D}_c) = \mathcal{L}_d(\boldsymbol{\theta}; \mathcal{D}_u) + \alpha \mathcal{L}_p(\boldsymbol{\theta}, \boldsymbol{\Lambda}; \mathcal{D}_c) + \beta \|\boldsymbol{\Lambda}\|_0 \tag{6.15}$$

where α weights the physics loss function and β is the regularization parameter associate to the ℓ_0 norm. The regularization term has been added in order to induce sparsity on coefficient matrix $\boldsymbol{\Lambda}$. Optimal parameters are denoted by

$$\{\boldsymbol{\theta}^\star, \boldsymbol{\Lambda}^\star\} := \arg\min_{\boldsymbol{\theta}, \boldsymbol{\Lambda}} \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\Lambda}; \mathcal{D}_u, \mathcal{D}_c) \tag{6.16}$$

Parameter optimization is implemented through the Alternating Direction Optimization (ADO) algorithm [11], which consists of alternately optimizing the parameters $\boldsymbol{\Lambda}$ of the PDE system and $\boldsymbol{\theta}$ of the neural network. In particular:

$$\hat{\boldsymbol{\Lambda}}_{k+1} := \arg\min_{\boldsymbol{\Lambda}} \mathcal{L}_p(\hat{\boldsymbol{\theta}}_k, \boldsymbol{\Lambda}; \mathcal{D}_c) + \beta \|\boldsymbol{\Lambda}\|_0 \tag{6.17}$$

$$\hat{\boldsymbol{\theta}}_{k+1} := \arg\min_{\boldsymbol{\theta}} \mathcal{L}_d(\boldsymbol{\theta}; \mathcal{D}_u) + \alpha \mathcal{L}_p(\boldsymbol{\theta}, \hat{\boldsymbol{\Lambda}}_{k+1}; \mathcal{D}_c) \tag{6.18}$$

The first optimization problem (6.17) is attached via Sequential Thresholding Ridge regression (STRidge) while the second one (6.18) can be attached through a standard DNN training algorithm.

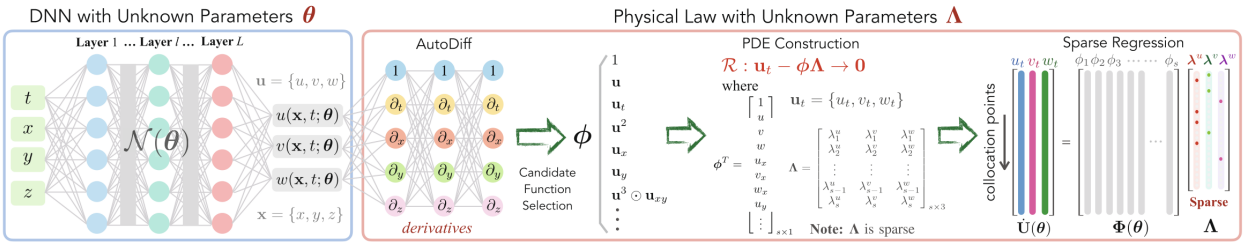

Figure 6.2: PINN algorithm scheme (image credits [19])

Remark 6.2.1. *It is possible to relax the ℓ_0 term regularization in equation* (6.15) *by ℓ_1 regularization: in this way the ill-posedness of the problem is improved. A pre-training of PINN-SR, with ℓ_1 regularization can be implemented before running the ADO algorithm, in order to obtain an initial guess. In particular a brute-force gradient-based optimization for both $\boldsymbol{\theta}$ and $\boldsymbol{\Lambda}$ can be applied.*

Remark 6.2.2. *If $r \geq 2$ independent data sets from different I/BCs are available, a "root-branch" DNN is used such that $\boldsymbol{u}_r^{\boldsymbol{\theta}} = \boldsymbol{u}(\boldsymbol{x}, t, \boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(r)})$ is the r^{th} branch of the "root-branch" neural network, where $\boldsymbol{\theta}^{(0)}$ and $\boldsymbol{\theta}^{(r)}$ are the root neural network $\mathcal{N}_0$ and r^{th} branch neural network $\mathcal{N}_r$ parameters respectively. Figure 6.3 shows the architecture of the method in this case.*

For each branch, libraries $\Phi^{(r)}$ are obtained and stacked to build a unique set $\Phi^{\boldsymbol{\theta}}$. The same stacking process is applied to $\dot{\boldsymbol{u}}_r^{\boldsymbol{\theta}}$ in order to compute the physics loss function (6.14).

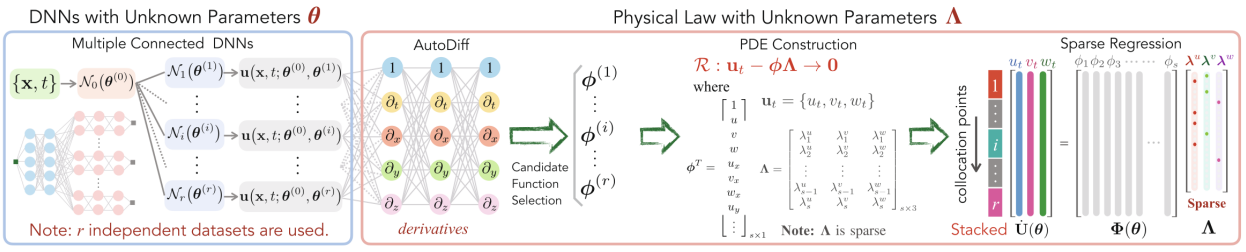

Figure 6.3: PINN algorithm scheme when $r \geq 2$ independent datasets are available (image credits [19]).

The computational cost of PINN-SR is much higher than the state of the art techniques based on the SINDy method: this is mainly due to the long training time of the DNN.

When the number r of independent data sets is quite large, the number of "root-branch" neural network parameters that need to be estimated is very large. Moreover, as the dimensionality of the spatial domain Ω increases, it is necessary to increase the number of collocation points in order to maintain sufficient accuracy: this can lead to high computational costs and problems.

Finally, the PINN-SR method is a library-based method: the library chosen must be sufficiently comprehensive (to ensure accuracy) but not unnecessarily large.

6.2.3 ODE learning with Gaussian processes

In many situations it is impossible to formulate a parametric ODE system that describes the latent dynamics of the measured data. Therefore, it is possible to introduce nonparametric modeling [46] using vector fields of Gaussian processes [40, 16, 29, 4, 71]. The dynamics equation model is

$$\dot{u} = f(u) \tag{6.19}$$

where $u(t) \in \mathbb{R}^U$ is the time-depending continuous state variable and $f : \mathbb{R}^U \to \mathbb{R}^U$ represents the completely unknown time derivative function. The forward solution of the system of ODEs is computed by an integration operation

$$u(t) = u_0 + \int_{t_0}^{t} f(u(\tau))d\tau \tag{6.20}$$

where t_0 is the initial time (usually $t_0 = 0$) and $u_0 = u(t_0)$ is the initial state. The data are time series obtained by sampling the state variable u at times $\mathbf{t} = [t_0, ..., t_{T-1}]$ and are organized in a matrix $\mathbf{d} = [\mathbf{d}_0, ..., \mathbf{d}_{T-1}] \in \mathbb{R}^{T\times U}$. Data are obtained from observation affected by Gaussian noise:

$$\mathbf{d}_k = \mathbf{u}_k + \boldsymbol{\eta}_k \tag{6.21}$$

where $\mathbf{u} \in \mathbb{R}^{T\times U}$ is such that $\mathbf{u}_{k,h} = u_h(t_k)$ for each $k = 0, ..., T-1$ and $h = 0, ..., U-1$, and $\boldsymbol{\eta}_k \sim N(\mathbf{0}, \Omega)$ with $\Omega = diag(\omega_1^2, ..., \omega_U^2)$ the covariance matrix.

By framing the problem in a Bayesian framework based on Gaussian processes, a non-parametric ODE model can be obtained. Specifically:

$$\mathbf{f} \sim \mathcal{GP}(\mathbf{0}, K) \tag{6.22}$$

that is,

- $\mathbb{E}[\mathbf{f}(\mathbf{u})] = \mathbf{0}$ for each $\mathbf{u} \in \mathbb{R}^U$;
- $\text{cov}[\mathbf{f}(\mathbf{u}, \mathbf{u}')] = K(\mathbf{u}, \mathbf{u}')$.

where K is the kernel operator and $K(\mathbf{u}, \mathbf{u}') \in \mathbb{R}^{U\times U}$.

Given a collection of state variables (in the space of state variables) $\mathbf{W} = [\mathbf{w}_1, ..., \mathbf{w}_N] \in \mathbb{R}^{N\times U}$ and their corresponding value of the vector field $\mathbf{f}(\mathbf{W}) = [\mathbf{f}(\mathbf{w}_1), ..., \mathbf{f}(\mathbf{w}_N)] \in \mathbb{R}^{N\times U}$

(in the space of derivatives) it is possible to compute the probability

$$\mathbb{P}(\mathbf{f}(\mathbf{W})) = \mathcal{N}(vec(\mathbf{f}(\mathbf{W}))|\mathbf{0}, \mathbf{K}(\mathbf{W}, \mathbf{W})) \tag{6.23}$$

where $\mathbf{K}(\mathbf{W}, \mathbf{W}) = (K(\mathbf{w}_i, \mathbf{w}_j))_{i,j=1}^{N} \in \mathbb{R}^{(N\times U)\times(N\times U)}$ is a block matrix of matrix-valued kernels $K(\mathbf{w}_i, \mathbf{w}_j)$.

Knowing exactly the vector field $\mathbf{f}(\mathbf{u})$ in all possible states $\mathbf{u}$ is impossible: however, it is necessary to know it since its integral over time domain (6.20) is needed in order to compare the estimated model with the data $\mathbf{d}$. One defines a set of inducing points $\{\mathbf{z}_i \in \mathbb{R}^U\}_{i=1}^{M}$ (or inducing locations) [92], i.e. a set of points in the space of states for which the vector field is assumed to be known, i.e. $\mathbf{z}_i, \mathbf{v}_i \in \mathbb{R}^U$ such that $\mathbf{f}(\mathbf{z}_i) = \mathbf{v}_i$ for each $i = 1, ..., M$. The vectors $\mathbf{v}_i$ are called inducing vectors. Thus, using the properties of kernel-based Gaussian processes, it is possible to estimate the vector field in all possible states $\mathbf{u}$ as

$$\mathbf{f}(\mathbf{u}) \triangleq \mathbf{K}_\theta(\mathbf{u}, \mathbf{Z})\mathbf{K}_\theta(\mathbf{Z}, \mathbf{Z})^{-1} vec(\mathbf{V}) \tag{6.24}$$

where $\mathbf{Z} = [\mathbf{z}_1, ..., \mathbf{z}_M]$ is the matrix of inducing points, $\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_M]$ is the matrix of inducing vectors, and θ are the parameters defining the kernel. Furthermore $\mathbf{K}_\theta(\mathbf{u}, \mathbf{Z}) \in \mathbb{R}^{U\times(M\times U)}$ and $\mathbf{K}_\theta(\mathbf{Z}, \mathbf{Z}) \in \mathbb{R}^{(M\times U)\times(M\times U)}$. The term $\mathbf{K}_\theta(\mathbf{u}, \mathbf{Z})$ measures the similarity between the input $\mathbf{u}$ and the inducing points $\mathbf{z}_i$ for each $i = 1, ..., M$: thus, if the input $\mathbf{u}$ is very close to an inducing point $\mathbf{z}_i$, then the value of $\mathbf{f}(\mathbf{u})$ will be similar to the corresponding inducing vector $\mathbf{v}_i$. Figure 6.4 shows an example of a vector field induced by a Gaussian process and obtained from the inducing points $\mathbf{Z}$ and the inducing vectors $\mathbf{V}$.

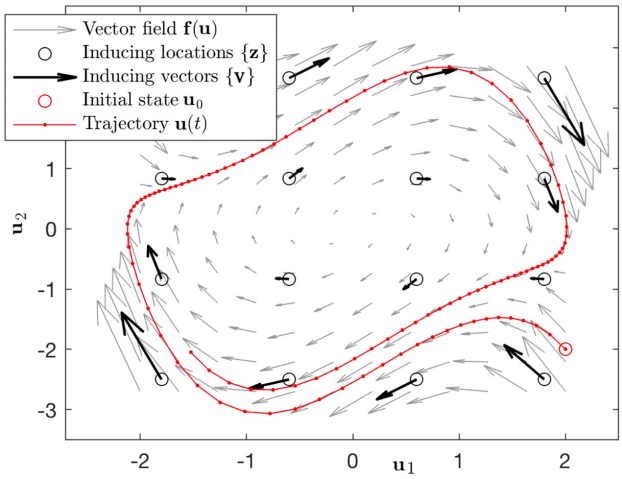

Figure 6.4: Example of interpolation using a kernel-based Gaussian process through the inducing points $\mathbf{Z}$ and the inducing vectors $\mathbf{V}$ in the case of 2-dimensional state variables $\mathbf{u}(\cdot) \in \mathbb{R}^2$ (image credits [46]).

Remark 6.2.3. *In most cases, it is assumed that f can be written in a known parametric form $\boldsymbol{f}(\boldsymbol{u};\theta,t)$, where θ are the unknown parameters to be estimated. In this way, the dynamics reconstruction problem is transformed into the problem of estimating the parameters so that the forward solution of the ODE is maximally compatible with the given data. However, the data $\boldsymbol{d}$ and the function $\boldsymbol{f}(\boldsymbol{u};\theta,\cdot)$ to be optimized belong to two conceptually different spaces: in fact, the latter lives in the space of the derivatives of the state variables. Therefore, to make the estimated model comparable with the data, the optimization process requires the computationally intensive integration operation* (6.20) *to compute the forward solution. To avoid integration computations, gradient matching techniques [125, 31, 94] compute the empirical derivatives of the data $\dot{\boldsymbol{d}}$: the latter are directly comparable to the estimated model $\boldsymbol{f}(\boldsymbol{u};\theta,\cdot)$, since they belong to the same space. However, since the data are always affected by noise, the empirical derivation is strongly ill-posed.*

In order to solve the ODE learning problem, it is necessary to reconstruct the vector field $\mathbf{f}$ by appropriately estimating the quantities on which it depends, i.e. the kernel parameters θ and the inducing vectors $\mathbf{V}$ computed at the inducing points $\mathbf{Z}$ from the set of measurements $\mathbf{d}$.

Since the problem has been framed in a Bayesian framework, it is possible to define the likelihood on the data as

$$\mathbb{P}(\mathbf{d}|\mathbf{u}_0,\mathbf{f},\Omega)=\prod_{k=0}^{T-1}\mathcal{N}(\mathbf{d}_k|\mathbf{u}_k,\Omega) \tag{6.25}$$

where $\mathbf{u}_k$, for $k=0,...,T-1$, are the forward state variables computed using integral equation (6.20) and additive noise model (6.21) with noise matrix variance Ω.

The distribution of the inducing vectors $\mathbf{V}$, defined by the inducing points $\mathbf{Z}$ and the kernel parameters θ, is given by:

$$\mathbb{P}(\mathbf{V}|\mathbf{Z},\theta)=\mathcal{N}(vec(\mathbf{V})|\mathbf{0},\mathbf{K}_\theta(\mathbf{Z},\mathbf{Z})) \tag{6.26}$$

Expressing the vector field $\mathbf{f}$ in terms of the inducing vectors $\mathbf{V}$ and the parameters θ in (6.25), it is possible to define the posterior of the dynamical model as:

$$\mathcal{P}:=\mathbb{P}(\mathbf{V},\theta,\mathbf{u}_0,\Omega|\mathbf{d})=\mathbb{P}(\mathbf{d}|\mathbf{V},\theta,\mathbf{u}_0,\Omega)\mathbb{P}(\mathbf{V}|\theta)\mathbb{P}(\theta)\mathbb{P}(\mathbf{u}_0)\mathbb{P}(\Omega) \tag{6.27}$$

where the dependence on the inducing points $\mathbf{Z}$ is omitted since these are chosen a priori and the hyperpriors $\mathbb{P}(\theta)$, $\mathbb{P}(\mathbf{u}_0)$, $\mathbb{P}(\Omega)$ are assumed to be uniform unless there is additional a priori knowledge in the respective domains. In Bayesian terms, the problem is to find the inducing vectors $\mathbf{V}$, the kernel parameters θ, the initial condition $\mathbf{u}_0$ and the matrix Ω of

variances of the noise maximizing posterior $\mathcal{P}$ (6.27); these quantities are called maximum a posteriori. By applying a logarithmic transformation and a sign change, the posterior maximization problem is transformed into the more tractable minimum problem:

$$\mathbf{V}^*, \theta^*, \mathbf{u}_0^*, \Omega^* = \arg \min_{\mathbf{V}, \theta, \mathbf{u}_0, \Omega} -\ln(\mathcal{P}) \tag{6.28}$$

This problem can be solved by the gradient descent algorithm, where the partial derivatives can be computed either analytically (as in the case of the noise Ω) or by the finite difference technique (as in the case of the parameters θ). The pseudocode 2 shows a possible algorithm to solve this problem.

Algorithm 2 Optimization algorithm for solving the ODE learning problem with non parametric model based on Gaussian processes.

1: Define inducing points $\mathbf{U}^{(0)}$ (randomly or belonging to a grid). Set initial guesses for inducing vectors $\mathbf{V}^{(0)}$, for parameters $\theta^{(0)}$, for IC $\mathbf{u}_0^{(0)}$ and noise variance matrix $\Omega^{(0)}$;
2: **for** $q = 0$ to N_{iter} **do**
3: Interpolate vector field $\mathbf{f}^{(q)}$ in the state domain with Gaussian process with (6.24) with $\mathbf{U}^{(q)}$, $\mathbf{V}^{(q)}$, $\theta^{(q)}$;
4: Compute forward solution $\mathbf{u}^{(q)}$ with (6.20) with IC $\mathbf{u}_0^{(q)}$ and vector field $\mathbf{f}^{(q)}$;
5: Gradient descent step to optimize (transformed) posterior $\mathcal{P}$ (6.27) and update $\mathbf{V}^{(q+1)}, \theta^{(q+1)}, \mathbf{u}_0^{(q+1)}, \Omega^{(q+1)}$.
6: **end for**
7: **return** $\mathbf{V}^*, \theta^*, \mathbf{u}_0^*, \Omega^*$.

Remark 6.2.4. *For example, choosing a decomposable Gaussian kernel $K_\theta(\boldsymbol{z}, \boldsymbol{z}') = k_\theta(\boldsymbol{z}, \boldsymbol{z}')I_U$ where $k_\theta(\boldsymbol{z}, \boldsymbol{z}') \in \mathbb{R}$ is the scalar kernel, then*

$$K_\theta(\boldsymbol{z}, \boldsymbol{z}') = \sigma_f^2 \, e^{-\frac{1}{2}\sum_{j=0}^{U-1} \frac{(z_j - z_j')^2}{\ell_j^2}} I_U \tag{6.29}$$

where σ_f^2 is the differential variance and $\boldsymbol{\ell} = [\ell_0, ..., \ell_{U-1}]$ are the dimension-specific length-scales. Kernel parameters are then collected as $\theta = [\sigma_f, \boldsymbol{\ell}]$.

Chapter 7

Solving Implicit Inverse Problems with Homotopy-Based Regularization Path

In the previous chapters, we have seen that the latent dynamics discovery problem is an extremely common problem in a wide variety of applications. In fact, it is possible to frame this problem within a more general category of problems: the implicit inverse problems.

Implicit inverse problems [77, 81] require the determination of a non-linear functional applied to a function representing a physical quantity, when only noisy measurements of this quantity are available. When the functional is modeled by a parametric form, the implicit inverse problem becomes the one to optimize its parameters. Due to the inherent complexity of the problem and the nonlinear nature of their constraints, implicit inverse problems are typically strongly ill-posed and, in particular, are characterized by non-uniqueness of the solution.

Typical examples of these problems are found in Ordinary or Partial Differential Equation constrained optimization, such as equation learning, shape optimization, optimal control and data assimilation problems.

The aim is to determine the parameters of an ODE or PDE system such that the corresponding solution best fits the observed data [14, 20, 46, 70, 29, 136, 96]. This involves the optimization of model parameters to match the observation and is often applied in latent dynamics discovery, where the underlying dynamics governing a system need to inferred from data. These problems are common in fields like geophysics and biological modeling, where the governing equations of a system are either unknown or poorly understood, and must be learned from data [90].

In optimal control, the goal is to determine a control function that minimizes an objective function while satisfying constraints imposed by a dynamic system of ODEs or PDEs. This type of problem occurs in a wide range of applications, from robotics (where the goal is to

optimize control inputs to achieve specific movements) to medicine (for example, optimizing drug dosages) and finance (where the objective might be to minimize financial risk) [7, 1].

Data assimilation involves integrating observational data with mathematical models to obtain the best possible representation of a physical system. This is particularly important in areas such as weather forecasting and climate modeling, where models like the Navier-Stokes equations are combined with real-world data (e.g., satellite measurements) to predict future states of the system. In methods such as variational data assimilation, the goal is to optimize the model parameters to align with the observed data, often in high-dimensional spaces [3, 57, 15].

Given the broad range of problems that can be considered as implicit inverse problems, the proposed solution methods are typically problem-specific, and the methods are often not framed within the broader context of implicit inverse problems, with few exceptions [98, 76, 7, 1]. Indeed, the strong ill-posedness of these problems often forces the development of solution methods that are restricted to the specific application.

In this chapter, we show a homotopy method [133, 3] to follow the regularization path that can potentially be applied to all specific problems that can be framed as implicit inverse problems. The method exploits a variational approach by minimizing a functional that combines a data-fitting term and a penalty term to enforce the stability of the solution. For any fixed value of the regularization parameter, we solve the problem using a gradient algorithm that leverages the adjoint state method [90] for computing large Jacobians and the Newton-Raphson technique [24] for solving nonlinear forward problems at each iteration. These consolidated numerical tools ensure the robustness of the algorithm.

Our strategy is to start with large values of the regularization parameter, ensuring that the solution exists and is unique. Then, we iteratively solve the problem with decreasing values of the regularization parameter to obtain a complete regularization path [35]. Moreover, we apply the proposed method to a latent dynamics discovery problem (ODE learning problem), where the forward model is given by a an autonomous differential equation (as equation (6.2)) characterized by few important terms that govern the dynamics. We show how this method can recover the important terms of the dynamics in the case the differential equations contains only periodic functions. As in the case of classical linear inverse problems, we show the well known semi-convergent behavior of the regularized solution family.

This work is currently submitted to arXiv [84].

The plan of this chapter is as follows. Section 7.1 provides the mathematical framework of the numerical method, which is formulated in Section 7.2. Section 7.3 describes an application related to the problem of learning ODEs. Section 7.4 concludes the chapter with a summary of our findings and conclusions.

7.1 Mathematical framework

Let $\mathcal{H}$ be a complete Hilbert space and $u \in \mathcal{H}^U$ such that

$$t \to u(t) = [u_0(t), u_1(t), ..., u_{U-1}(t)] \in \mathbb{R}^U$$

where $u_h \in \mathcal{H}$ for each $h = 0, ..., U-1$. Let

$$\mathcal{F} : \mathcal{H}^U \times \mathbb{R}^M \to \mathcal{H}^U$$

$$(u, \mathbf{m}) \to \mathcal{F}(u, \mathbf{m})$$

be an implicit parametric functional, where the parametric form is known but the values of the entries of $\mathbf{m} \in \mathbb{R}^M$ are unknown. The relationship between u and $\mathbf{m}$ can be expressed as

$$\mathcal{F}(u, \mathbf{m}) = 0. \tag{7.1}$$

We now introduce two vectors, $\mathbf{u}, \mathbf{d} \in \mathbb{R}^{T \times U}$, such that

$$u_{k,h} = u_h(t_k) \quad k = 0, ..., T-1 \;\; ; \;\; h = 0, ..., U-1, \tag{7.2}$$

and

$$d_{k,h} = u_{k,h} + \eta, \tag{7.3}$$

where η is sampled from a Gaussian distribution $\mathcal{N}(\mathbf{0}, \sigma)$. Thus, the parametric implicit inverse problem addressed in this study is to determine $\mathbf{m}$ in (7.1), given a discretized, noisy version $\mathbf{d}$ of u from (7.3) and the known parametric form of $\mathcal{F}$.

To address the intrinsic ill-posedness of the problem [43], we introduce a regularized objective function $k_{\mathbf{d},\alpha} : \mathbb{R}^{T \times U} \times \mathbb{R}^M \to \mathbb{R}$ defined as

$$(\mathbf{u}, \mathbf{m}) \to k_{\mathbf{d},\alpha}(\mathbf{u}, \mathbf{m}) = h_{\mathbf{d}}(\mathbf{u}) + \alpha g(\mathbf{m}) \;\; , \tag{7.4}$$

where the choice

$$h_{\mathbf{d}}(\mathbf{u}) = \|\mathbf{u} - \mathbf{d}\|_2^2 \tag{7.5}$$

is consistent with the Gaussian noise affecting the measurements (7.3). In (7.4), $\alpha > 0$ is the regularization parameter, and $g : \mathbb{R}^M \to \mathbb{R}$ is a regularization function [124, 99] such that

- $g(\mathbf{m}) \geq 0$ for each $\mathbf{m} \in \mathbb{R}^M$;
- $g(\mathbf{m}) = 0 \iff \mathbf{m} = \mathbf{0}$;

- $\lim_{\alpha \to 0} \mathbf{m}_\alpha = \mathbf{m}^*$.

We now define a discretized version of the implicit parametric functional in (7.1):

$$(\mathbf{u}, \mathbf{m}) \to \mathbf{F}(\mathbf{u}, \mathbf{m}) \tag{7.6}$$

so that equation (7.1) becomes

$$\mathbf{F}(\mathbf{u}, \mathbf{m}) = \mathbf{0}. \tag{7.7}$$

We denote by $\mathbf{u_m} \in \mathbb{R}^{T \times U}$ the solution to (7.7) for a given set of parameters $\mathbf{m}$.

From an operational standpoint, the solution of the discretized, regularized parametric implicit inverse problem is given by

$$\mathbf{m}_\alpha \in \arg\min_{\mathbf{m} \in \mathbb{R}^M} k_{\mathbf{d},\alpha}(\mathbf{u_m}, \mathbf{m}) \tag{7.8}$$

where $\mathbf{d}$ is the input data vector of the problem. The solution to the corresponding non-regularized discretized inverse problem is

$$\mathbf{m}^* \in \arg\min_{\mathbf{m} \in \mathbb{R}^M} h_{\mathbf{d}}(\mathbf{u_m}), \tag{7.9}$$

which is obtained by setting $\alpha = 0$ in (7.8).

7.2 Homotopy-based regularization

We propose a homotopy-based optimization scheme to regularize the implicitly constrained inverse problem defined by equation (7.9). The method is structured in two conceptual layers:

- Inner loop – adjoint-based optimization: for a fixed regularization parameter $\alpha > 0$, we solve the regularized problem via a gradient-based algorithm using the adjoint state method. This yields a sequence of approximations to the parameter vector $\mathbf{m}_\alpha$.

- Outer loop – homotopy continuation: We trace the regularization path by progressively decreasing α and using the solution obtained for α_l to initialize the optimization for α_{l+1} (a warm restart). This allows us to gradually approach the solution $\mathbf{m}^*$ of the non-regularized problem while maintaining numerical stability.

7.2.1 Inner loop: adjoint-based optimization

For a given $\alpha > 0$, we solve the regularized inverse problem by minimizing the Lagrangian:

$$\mathcal{L}_{\mathbf{d},\alpha}(\mathbf{u}, \mathbf{m}, \boldsymbol{\lambda}) = h_{\mathbf{d}}(\mathbf{u}) + \alpha g(\mathbf{m}) - \langle \boldsymbol{\lambda}, \mathbf{F}(\mathbf{u}, \mathbf{m}) \rangle_{T \times U}, \tag{7.10}$$

where $\mathbf{F}(\mathbf{u}, \mathbf{m}) = 0$ encodes the implicit constraint.

We recall the following result characterizing stationarity of the regularized problem:

Proposition 7.2.1. *Necessary condition for the solution of the discretized regularized implicit inverse problem. Given the discretized regularized parametric implicitly constrained minimum problem* (7.8) *associated to functional* $\boldsymbol{F}$*, one of its solution* $\boldsymbol{m}_\alpha$ *and the associated vector* $\boldsymbol{u}_{\boldsymbol{m}_\alpha}$*, then there exists* $\boldsymbol{\lambda}_{\boldsymbol{m}_\alpha} \in \mathbb{R}^{T \times U}$ *such that all the following conditions are fulfilled:*

1. $\boldsymbol{\lambda}_{\boldsymbol{m}_\alpha} = \left(\left[\dfrac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}} \right]^T \right)^\dagger \dfrac{\partial h_{\boldsymbol{d}}}{\partial \boldsymbol{u}} \Big|_{(\boldsymbol{u}_{\boldsymbol{m}_\alpha}, \boldsymbol{m}_\alpha)}$;
2. $\alpha \dfrac{\partial g}{\partial \boldsymbol{m}} - \langle \boldsymbol{\lambda}, \dfrac{\partial \boldsymbol{F}}{\partial \boldsymbol{m}} \rangle_{T \times U} \Big|_{(\boldsymbol{u}_{\boldsymbol{m}_\alpha}, \boldsymbol{m}_\alpha, \boldsymbol{\lambda}_{\boldsymbol{m}_\alpha})} = \boldsymbol{0}$.

Proof. Let us set $\mathbf{u}_{\mathbf{m}_\alpha} = \mathbf{u}_\alpha$ and $\boldsymbol{\lambda}_{\mathbf{m}_\alpha} = \boldsymbol{\lambda}_\alpha$. We can define the Lagrangian of the problem as

$$\mathcal{L}_{\mathbf{d},\alpha}(\mathbf{u}, \mathbf{m}, \boldsymbol{\lambda}) = h_{\mathbf{d}}(\mathbf{u}) + \alpha g(\mathbf{m}) - \langle \boldsymbol{\lambda}, \mathbf{F}(\mathbf{u}, \mathbf{m}) \rangle_{T \times U}, \tag{7.11}$$

where $\langle \boldsymbol{\lambda}, \mathbf{F}(\mathbf{u}, \mathbf{m}) \rangle_{T \times U} = \sum_{t=0}^{T-1} \sum_{h=0}^{U-1} \lambda_{t,h} \mathbf{F}(\mathbf{u}, \mathbf{m})_{t,h}$. According to Lagrange's multiplier theorem [8], if $\mathbf{m}_\alpha$ is a minimum for the discretized (7.8), then there must exist $\boldsymbol{\lambda}_\alpha \in \mathbb{R}^{T \times U}$ such that $(\mathbf{u}_\alpha, \mathbf{m}_\alpha, \boldsymbol{\lambda}_\alpha)$ is a stationary point for the Lagrangian:

$$\nabla \mathcal{L}_{\mathbf{d},\alpha}(\mathbf{u}_\alpha, \mathbf{m}_\alpha, \boldsymbol{\lambda}_\alpha) = \mathbf{0} \tag{7.12}$$

This implies that all partial derivatives must be equal to 0 at $(\mathbf{u}_\alpha, \mathbf{m}_\alpha, \boldsymbol{\lambda}_\alpha)$:

$$\frac{\partial \mathcal{L}_{\mathbf{d},\alpha}}{\partial \boldsymbol{\lambda}}(\mathbf{u}_\alpha, \mathbf{m}_\alpha, \boldsymbol{\lambda}_\alpha) = \mathbf{F}(\mathbf{u}_\alpha, \mathbf{m}_\alpha) = \mathbf{0}; \tag{7.13}$$

$$\frac{\partial \mathcal{L}_{\mathbf{d},\alpha}}{\partial \mathbf{u}}(\mathbf{u}_\alpha, \mathbf{m}_\alpha, \boldsymbol{\lambda}_\alpha) = \frac{\partial h_{\mathbf{d}}}{\partial \mathbf{u}}(\mathbf{u}_\alpha) - \langle \boldsymbol{\lambda}_\alpha, \frac{\partial \mathbf{F}}{\partial \mathbf{u}}(\mathbf{u}_\alpha, \mathbf{m}_\alpha) \rangle_{T \times U} = \mathbf{0}; \tag{7.14}$$

$$\frac{\partial \mathcal{L}_{\mathbf{d},\alpha}}{\partial \mathbf{m}}(\mathbf{u}_\alpha, \mathbf{m}_\alpha, \boldsymbol{\lambda}_\alpha) = \alpha \frac{\partial g}{\partial \mathbf{m}}(\mathbf{m}_\alpha) - \langle \boldsymbol{\lambda}_\alpha, \frac{\partial \mathbf{F}}{\partial \mathbf{m}}(\mathbf{u}_\alpha, \mathbf{m}_\alpha) \rangle_{T \times U} = \mathbf{0}. \tag{7.15}$$

Equation (7.13) is equivalent to equation (7.7) with parameters $\mathbf{m}_\alpha$. Equation (7.14) implies

that $\boldsymbol{\lambda}_\alpha$ must solve

$$\left[\frac{\partial \mathbf{F}}{\partial \mathbf{u}}(\mathbf{u}_\alpha, \mathbf{m}_\alpha)\right]^T \boldsymbol{\lambda}_\alpha = \frac{\partial h_{\mathbf{d}}}{\partial \mathbf{u}}(\mathbf{u}_\alpha), \tag{7.16}$$

which means that

$$\boldsymbol{\lambda}_\alpha = \left(\left[\frac{\partial \mathbf{F}}{\partial \mathbf{u}}\right]^T\right)^\dagger \frac{\partial h_{\mathbf{d}}}{\partial \mathbf{u}}\Big|_{(\mathbf{u}_\alpha, \mathbf{m}_\alpha)} \tag{7.17}$$

Replacing $\mathbf{u}_\alpha$ and $\boldsymbol{\lambda}_\alpha$ obtained by solving (7.13) and (7.16) respectively in (7.15), we demonstrate point (ii) of the preposition.

Remark 7.2.1. *The adjoint state method consists of computing the gradient of the Lagrangian* (7.10) *with respect to the unknown parameters* $\boldsymbol{m}$ *by computing* $\boldsymbol{\lambda}_\alpha$*, called adjoint state variables obtained by solving the linear system* (7.16)*, called adjoint state system.*

Remark 7.2.2. *This method can be applied to any* $h_{\boldsymbol{d}} : \mathbb{R}^{T\times U} \to \mathbb{R}$ *such that:*

- $h_{\boldsymbol{d}}(\boldsymbol{u}) \geq 0$*;*
- $h_{\boldsymbol{d}}(\boldsymbol{u}) = 0 \iff u_h(t_k) = d_{k,h}$ *for each* $k = 0, ..., T-1,\ h = 0, ..., U-1$.

Under these conditions $h_{\boldsymbol{d}}$ *can act as an empirical loss function [97].*

Thanks to Proposition 7.2.1, we can easily derive a gradient descent algorithm where each iteration consists of:

1. *find solution* $\boldsymbol{u}_{\boldsymbol{m}_\alpha^{(i)}}$ *of equation* (7.7)*;*
2. *find* $\boldsymbol{\lambda}_{\boldsymbol{m}_\alpha^{(i)}}$ *solution of* (7.16) *computed in* $\boldsymbol{u}_{\boldsymbol{m}_\alpha^{(i)}}, \boldsymbol{m}_\alpha^{(i)}$*;*
3. *compute* $\dfrac{\partial \mathcal{L}_{\boldsymbol{d},\alpha}}{\partial \boldsymbol{m}}(\boldsymbol{u}_{\boldsymbol{m}_\alpha^{(i)}}, \boldsymbol{m}_\alpha^{(i)}, \boldsymbol{\lambda}_{\boldsymbol{m}_\alpha^{(i)}})$ *as in* (7.15)*;*
4. *update parameters* $\boldsymbol{m}_\alpha^{(i)}$ *through gradient descent algorithm*

$$\boldsymbol{m}_\alpha^{(i+1)} = \boldsymbol{m}_\alpha^{(i)} - \tau \frac{\partial \mathcal{L}_{\boldsymbol{d},\alpha}}{\partial \boldsymbol{m}}(\boldsymbol{u}_{\boldsymbol{m}_\alpha^{(i)}}, \boldsymbol{m}_\alpha^{(i)}, \boldsymbol{\lambda}_{\boldsymbol{m}_\alpha^{(i)}}) \tag{7.18}$$

with $\tau > 0$ *the gradient step.*

Remark 7.2.3. *As* $\boldsymbol{F}$ *is non-linear, neither the forward* (7.7) *nor adjoint equations* (7.16) *are solved to full precision. Inexact solutions are used to reduce computational cost and mitigate instability.*

Remark 7.2.4. *The forward problem* (7.7) *is solved by means of the iterative Newton-Raphson (NR) method. This method is implemented as follows:*

$$\boldsymbol{u}_\alpha^{(i,k+1)} = \boldsymbol{u}_\alpha^{(i,k)} - \left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}}\right]^\dagger_{(\boldsymbol{u}_\alpha^{(i,k)}, \boldsymbol{m}_\alpha^{(i)})} \boldsymbol{F}(\boldsymbol{u}_\alpha^{(i,k)}, \boldsymbol{m}_\alpha^{(i)}), \tag{7.19}$$

where we denoted with $\boldsymbol{u}_\alpha^{(i,k)}$ *the computed solution at* i^{th} *algorithm iteration and at* k^{th} *Newton-Raphson iteration and with* $\left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}}\right]^\dagger$ *the generalized inverse of the matrix* $\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}}$*. This in turn determines the resolution of the linear system at each step:*

$$\left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}}\right]_{(\boldsymbol{u}^{(i,k)}, \boldsymbol{m}^{(i)})} \boldsymbol{v}^{(i,k)} = \boldsymbol{F}(\boldsymbol{u}^{(i,k)}, \boldsymbol{m}^{(i)}) \tag{7.20}$$

Then the solution:

$$\boldsymbol{u}_\alpha^{(i,k+1)} = \boldsymbol{u}_\alpha^{(i,k)} - \boldsymbol{v}_\alpha^{(i,k)} \tag{7.21}$$

is updated. At the end of the Newton Raphson iterative cycle, we obtain the solution $\boldsymbol{u}_\alpha^{(i)}$*.*

Remark 7.2.5. *Step (ii) of the optimization algorithm involves the resolution of the linear system* (7.16)*. For this purpose the Landweber technique ([44, 104]) is used. Recall that solving a linear system requires the inversion of the system matrix. However, if this system is ill-conditioned, its inversion can lead to computational problems. The Landweber technique aims to restore the numerical stability. It transforms the problem of inverting a matrix into an iterative algorithm: the number of iterations is the regularization parameter with respect to the inversion of the matrix under consideration. The* k^{th} *Landweber iteration is defined as follows:*

$$\boldsymbol{\lambda}_\alpha^{(i,k+1)} = \boldsymbol{\lambda}_\alpha^{(i,k)} - \tilde{w} \left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}}\right]^T \left[\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}} \boldsymbol{\lambda}_\alpha^{(i,k)} - \frac{\partial h_{\boldsymbol{d}}}{\partial \boldsymbol{u}}\right] \Big|_{(\boldsymbol{u}_\alpha^{(i)}, \boldsymbol{m}_\alpha^{(i)})} \tag{7.22}$$

where $\tilde{w}$ *is such that*

$$0 < \tilde{w} < \frac{2}{\sigma_1^2} \tag{7.23}$$

and where σ_1^2 *is the square of the largest singular value of* $\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{u}}(\boldsymbol{u}_\alpha^{(i)}, \boldsymbol{m}_\alpha^{(i)})$*.*

Remark 7.2.6. *The optimization algorithm consists of an iterative update of* $(\boldsymbol{u}_{\boldsymbol{m}_\alpha^{(i)}}, \boldsymbol{m}_\alpha^{(i)}, \boldsymbol{\lambda}_{\boldsymbol{m}_\alpha^{(i)}})$ *for* $i \geq 0$*. These quantities are never computed exactly but approximated in order to avoid numerical instability. In fact internal optimization algorithms (NR* (7.19) *and LW* (7.22)*) are never brought to convergence. At the end of each iteration of the inner loop, early stopping [137] is applied. Fixed the* i^{th} *iteration, this is done by checking the values of the loss function* $k_{\boldsymbol{d},\alpha}$ *computed in* $\boldsymbol{u}_\alpha^{(i)}$*. We define* N_{ES} *the number of iterations for early stopping*

and ϵ_{ES} *a defined threshold. Then for each* i_1, i_2 *such that* $0 \leq i - N_{ES} \leq i_1, i_2 \leq i$ *if*

$$|k_{\boldsymbol{d},\alpha}^{(i_1)} - k_{\boldsymbol{d},\alpha}^{(i_2)}| < \epsilon_{ES} \tag{7.24}$$

then the inner loop iteration is stopped.

The output of this iterative algorithm corresponding to regularization parameter α *is the set of quantities* $(\boldsymbol{u}_\alpha, \boldsymbol{m}_\alpha, \boldsymbol{\lambda}_\alpha)$.

7.2.2 Outer loop: homotopy continuation

To recover the solution of the original (non-regularized) problem, we construct a regularization path:

$$\boldsymbol{\alpha} = [\alpha_0, \alpha_1, ..., \alpha_L] \tag{7.25}$$

such that $\alpha_0 > \alpha_1 > ... > \alpha_L > 0$. For each α_l, we solve the regularized problem using the adjoint-based algorithm above, initializing from the solution at α_{l-1}:

1. Compute $\mathbf{m}_{\alpha_l}$ using the algorithm from Subsection 7.2.1;
2. Use $(\mathbf{u}_{\alpha_l}, \mathbf{m}_{\alpha_l}, \boldsymbol{\lambda}_{\alpha_l})$ to initialize the optimization for α_{l+1} [63].

Remark 7.2.7. *Since* $\alpha_0 >> 0$*, the corresponding parameters* $\boldsymbol{m}_{\alpha_0}$ *obtained at the end of the inner loop will be extremely regularized.*

Remark 7.2.8. *This strategy enables stable convergence even in ill-posed settings. It ensures that the algorithm follows a continuous path of minimizers, which typically remain in the basin of attraction of a desired local solution.*

7.2.3 Algorithmic Summary

The complete procedure is summarized in Algorithm 3. Each outer loop iteration corresponds to a step in the homotopy, while each inner loop performs adjoint-based optimization for fixed α_l.

7.3 Application to latent dynamic discovery problem

In this section, we apply the proposed algorithm to the problem of discovering latent dynamics, where the goal is to reconstruct the unknown dynamics of a set of continuous physical

Algorithm 3 Homotopy-based Adjoint Optimization

Initialize $\tau > 0$, $\mathbf{m}^{(0,0)}$
for $l = 0$ to L **do**
    Set $\alpha \leftarrow \alpha_l$
    for $i = 1$ to N_{max} **do**
        Solve $\mathbf{F}(\mathbf{u}^{(l,i)}, \mathbf{m}^{(l,i)}) = 0$ for $\mathbf{u}^{(l,i)}$ (NR (7.19) method)
        Solve adjoint equation for $\boldsymbol{\lambda}^{(l,i)}$ (LW (7.22) method)
        Compute $\nabla_{\mathbf{m}} \mathcal{L}_{\mathbf{d},\alpha}$
        Update $\mathbf{m}^{(l,i)}$ and apply early stopping (7.24)
    end for
    Set $\mathbf{m}^{(l+1,0)} \leftarrow \mathbf{m}^{(l)} \leftarrow \mathbf{m}^{(l,Niter)}$
end for
return The entire regularization path: $(\mathbf{u}^{(l)}, \mathbf{m}^{(l)}, \boldsymbol{\lambda}^{(l)})$ for $l = 0, ..., L$

quantities, denoted by u (the state variables), from their temporal samples measured at times $t_0, \ldots, t_{T-1}$ and collected in the data vector $\mathbf{d}$. We assume that the state variables are periodic with respect to the temporal samples, meaning $\mathbf{u}(t_0) = \mathbf{u}(t_{T-1})$. Inspired by the formulation in [46], the governing equation (7.1) simplifies to:

$$\mathcal{F}(u, \mathbf{m}) = \dot{u} - f(u, \mathbf{m}) = 0, \tag{7.26}$$

where the function

$$f : \mathcal{H}^U \times \mathbb{R}^M \to \mathcal{H}^U$$

$$(u, \mathbf{m}) \to f(u, \mathbf{m})$$

describes the system's dynamics. Following [14, 19, 46], we assume that the latent dynamics can be expressed as a linear combination of basis functions:

$$\Phi(u) = \{\phi_0(u), \phi_1(u), ..., \phi_{D-1}(u)\} \tag{7.27}$$

where, for each $j = 0, ..., D-1$

$$\phi_j : \mathcal{H}^U \to \mathcal{H}^U$$

$$u \to \phi_j(u) = [\phi_j(u)_0, \phi_j(u)_1, ..., \phi_j(u)_{U-1}],$$

is a vector-valued function with components $\phi_j(u)_h \in \mathcal{C}^1(\mathcal{H}^U, \mathcal{H})$ for each $h = 0, ..., U-1$. The dynamics are then modeled as:

$$f_h(u, \mathbf{m}) = \sum_{j=0}^{D-1} \sum_{h'=0}^{U-1} m_{h'+Uj,h} \phi_j(u)_{h'} \tag{7.28}$$

for each component $h = 0, ..., U-1$, where $\mathbf{m} \in \mathbb{R}^{(D\times U)\times U}$ is the unknown parameter matrix. This formulation is consistent with the working hypothesis of knowing the form of implicit parametric functional $\mathcal{F}$. This formulation is consistent with the hypothesis that the parametric form of the operator $\mathcal{F}$ is known. The corresponding dynamics can be described by the Cauchy problem:

$$\begin{cases} u(t_0) = u_0, \\ \dot{u}_h - \sum_{j=0}^{D-1} m_{jh}\phi_j(u)_h = 0 \quad \text{for each } h = 0, ..., U-1 \end{cases} \tag{7.29}$$

given an initial condition $u(t_0) = u_0 \in \mathbb{R}^U$.

By evaluating u and $\Phi(u)$ at the discrete time samples $t_0, \ldots, t_{T-1}$, we obtain the matrix $\mathbf{\Phi} \in \mathbb{R}^{T\times(D\times U)}$, with components $\phi_j(\mathbf{u})_h \in \mathbb{R}^T$. The discretized form of the functional $\mathbf{F}$ in (7.7), under the assumption of (7.26), becomes:

$$\mathbf{F}(\mathbf{u}, \mathbf{m}) = \dot{\mathbf{u}} - \langle \mathbf{\Phi}^T(\mathbf{u}), \mathbf{m} \rangle = \mathbf{0} \tag{7.30}$$

where $\dot{\mathbf{u}} \in \mathbb{R}^{T\times U}$ contains the discrete time derivatives of the state variables.

To promote sparsity in the learned dynamics and reduce model complexity, we apply L_1-regularization to the coefficient matrix $\mathbf{m}$. In order to enable the use of the adjoint method, we use a smooth approximation:

$$g(\mathbf{m}) = \sum_{p=1}^{M} \sqrt{m_p + \epsilon^2}, \tag{7.31}$$

with $\epsilon > 0$ small, as suggested in [25, 126]. While this yields nearly sparse solutions, true sparsity is enforced by a hard thresholding step [38] applied after each parameter update:

$$HT(m_p^{(l,i)}, \alpha_l) = \begin{cases} 0, & \text{if } |m_p^{(l,i)}| < \dfrac{\alpha_l}{2} \\ m_p^{(l,i)} & \text{otherwise} \end{cases} \tag{7.32}$$

where α_l is the regularization parameter. This step is implemented at line 7 of Algorithm 3.

7.3.1 Results

We evaluated the proposed method in a synthetic setting for the one-dimensional case ($U = 1$). Experiments were carried out by varying the set of basis functions, the ground truth parameters vector and the percentage of noise on synthetic data. To obtain the data in simulation, given an initial condition u_0, the analytical solution is obtained by solving the

Cauchy forward problem (7.29) using the Python ODE solver *solveivp* in *scipy.integrate* package. Then, noisy data are computed by adding Gaussian noise with varying standard deviations as in (7.3).

First, we present simulation results for the base 1 and base 5 (see Table (7.1)) through error analysis and graphical visualization of the obtained solutions. Next, we propose a theoretical discussion: the developed method allows us to observe semi-convergence behavior, typical of linear inverse problems, as well as the emergence of bifurcations along the regularization path. These phenomena highlight the ill-posedness of the problem, particularly the non-uniqueness of the solution issue.

Remark 7.3.1. *The proposed method works in a non-inverse crime framework. In fact, the technique used to solve the forward problem during the data generation phase differs from the one* (7.19) *used to solve the same problem in the inverse problem.*

Table 7.1 shows the set of basis functions $\mathbf{\Phi}$, which includes constant, polynomial and trigonometric terms that can be used to approximate any periodic function.

Table 7.1: List of basis function sets used in experiments.

	basis functions $\Phi(u)$
basis 1	$\cos(u)$,$\cos(2u)$, $\cos(3u)$, $\cos(4u)$, $\cos(5u)$, $\cos(6u)$
basis 2	1, $\cos(u)$, $\cos(2u)$, $\cos(3u)$, $\cos(4u)$, $\cos(5u)$, $\cos(6u)$
basis 3	$\cos(u)$, $\sin(u)$, $\cos(2u)$, $\sin(2u)$,.., $\cos(6u)$, $\sin(6u)$
basis 4	1, $\cos(u)$, $\sin(u)$, $\cos(2u)$, $\sin(2u)$,..., $\cos(6u)$, $\sin(6u)$
basis 5	u, $\cos(u)$, $\cos(2u)$, $\cos(3u)$, $\cos(4u)$, $\cos(5u)$, $\cos(6u)$
basis 6	u, $\cos(u)$, $\sin(u)$, $\cos(2u)$, $\sin(2u)$,..., $\cos(6u)$, $\sin(6u)$
basis 7	1, u, $\cos(u)$, $\sin(u)$, $\cos(2u)$, $\sin(2u)$,..., $\cos(6u)$, $\sin(6u)$

Results were obtained with:

- $u(t_0) = 0.2$;
- $\mathbf{t}_k = \frac{2\pi k}{T-1}$ for $k = 0, ..., T-1$ and $T = 100$ the time samples;
- $\eta \sim N(0, \sigma)$ the additive white Gaussian noise, where we considered three values for σ, i.e., $\sigma = 0.01, 0.1, 0.2$.
- $\boldsymbol{\alpha}$ a vector of $L = 100$ logarithmically equi-spaced points in the interval $[10^0, 10^{-6}]$;
- $\tau = 10^{-3}$ gradient step;
- $N_{max} = 1000$ maximum number of Lagrangian optimization algorithm with the adjoint state method;

- $R_{max} = 50$ maximum number of NR iterations;
- $L_{max} = 100$ maximum number of LW iterations;
- $N_{ES} = 5$ number of checked loss function values for early stopping technique.

Given a regularization parameter α_l with $l = 0, ..., 99$, we define the relative error between the associated parameters $\mathbf{m}^{(l)}$ and the ground truth $\mathbf{m}$ as:

$$RE_{\mathbf{m}^{(l)},\mathbf{m}} = \frac{\|\mathbf{m}^{(l)} - \mathbf{m}\|_2}{\|\mathbf{m}\|_2} \tag{7.33}$$

Similarly, the relative error between the solution $\mathbf{u}^{(l)}$ of the Cauchy problem (7.29) associated with $\mathbf{m}^{(l)}$ and the solution $\mathbf{u}$ associated with the ground truth $\mathbf{m}$ is given by:

$$RE_{\mathbf{u}^{(l)},\mathbf{u}} = \frac{\|\mathbf{u}^{(l)} - \mathbf{u}\|_2}{\|\mathbf{u}\|_2} \tag{7.34}$$

For each ground truth and noise level, we generated $n = 20$ realizations of the noise and applied the algorithm. For each trial $q = 1, ..., n$, we define

$$\alpha_q^* = \arg \min_{\alpha^{(l)}, l=0,...,99} RE_{\mathbf{m}_q^{(l)},\mathbf{m}} \tag{7.35}$$

To keep the notation simple, we will omit the noise level dependency. This will be clear and explicit in every situation.

We define $\mathbf{m}_q^*$ the set of parameters associated with the best regularization parameter α_q^* (7.35). $\mathbf{u}_q^*$ is the solution curve of the Cauchy problem (7.29) associated with $\mathbf{m}_q^*$. Then the relative mean square errors and standard deviations are computed as:

$$M_{\mathbf{m}} = \frac{1}{n}\sum_{q=1}^{n} RE_{\mathbf{m}_q^*,\mathbf{m}}, \quad \sigma_{\mathbf{m}} = \sqrt{\frac{1}{n}\sum_{q=1}^{n}\left(RE_{\mathbf{m}_q^*,\mathbf{m}} - M_{\mathbf{m}}\right)^2} \tag{7.36}$$

and

$$M_{\mathbf{u}} = \frac{1}{n}\sum_{q=1}^{n} RE_{\mathbf{u}_q^*,\mathbf{u}}, \quad \sigma_{\mathbf{u}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(RE_{\mathbf{u}_q^*,\mathbf{u}} - M_{\mathbf{u}}\right)^2} \tag{7.37}$$

7.3.2 Simulation Results

Base 1

We show results for base 1 in Table 7.1. Ground truth coefficient vectors used are:

$$\mathbf{m}_1 = (1, -1, 0, 0, 0, 0) \tag{7.38}$$

$$\mathbf{m}_2 = (-1.5, 1.5, -1.5, 1, -1, 0) \tag{7.39}$$

Figure 7.1 shows the generated synthetic data for $\mathbf{m}_1$ and $\mathbf{m}_2$ under the three noise levels $\boldsymbol{\sigma} = [0.01, 0.1, 0.2]$.

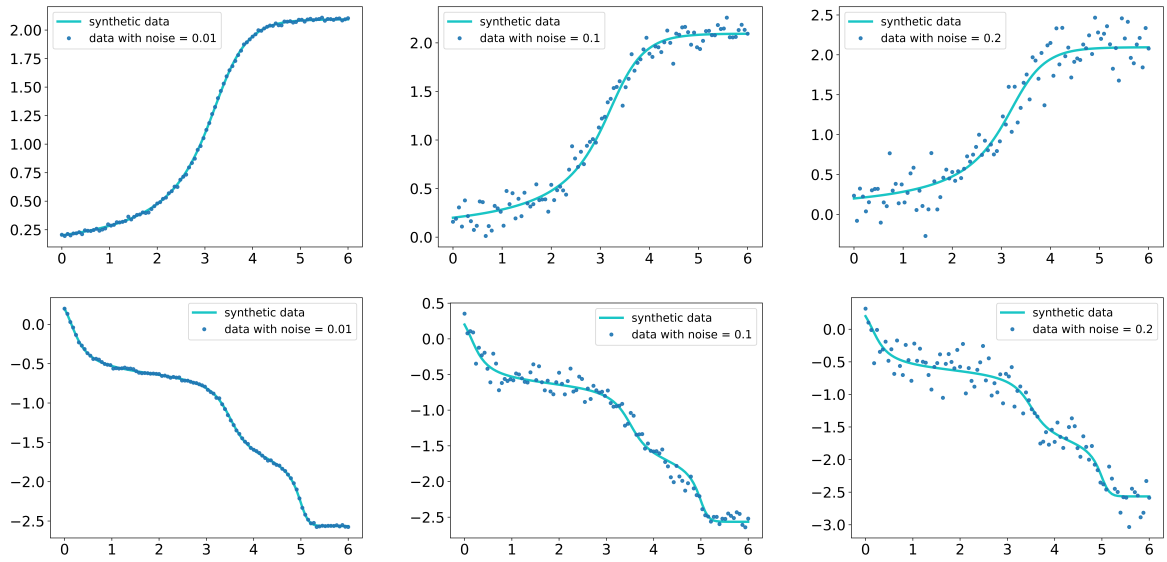

Figure 7.1: Base 1, $\mathbf{m}_1$ (first row), $\mathbf{m}_2$ (second row) synthetic data (light blue line) and noisy discrete synthetic data (blue dots) with different amount of noise: from left to right $\boldsymbol{\sigma} = [0.01, 0.1, 0.2]$. x-axis shows time. y-axis shows the values of the state variable.

Results for this base are summarized in Table 7.2 and illustrated via violin plots and regularization paths in Figures 7.2 and 7.3 .

Table 7.2: Base 1. Parameters and solutions mean relative error table. Each entry reports the mean relative error ± standard deviation (7.36) (7.37). Each row is a different ground truth, each column is a different level of noise $\boldsymbol{\sigma} = [0.01, 0.1, 0.2]$

	Noise levels		
M	Low	Medium	High
$\mathbf{m}_1$	0.01 ± 0.00	0.02 ± 0.02	0.02 ± 0.01
$\mathbf{m}_2$	0.11 ± 0.00	0.12 ± 0.02	0.11 ± 0.05
$\mathbf{u}_1$	0.01 ± 0.00	0.05 ± 0.02	0.04 ± 0.09
$\mathbf{u}_2$	0.03 ± 0.00	0.05 ± 0.01	0.04 ± 0.16

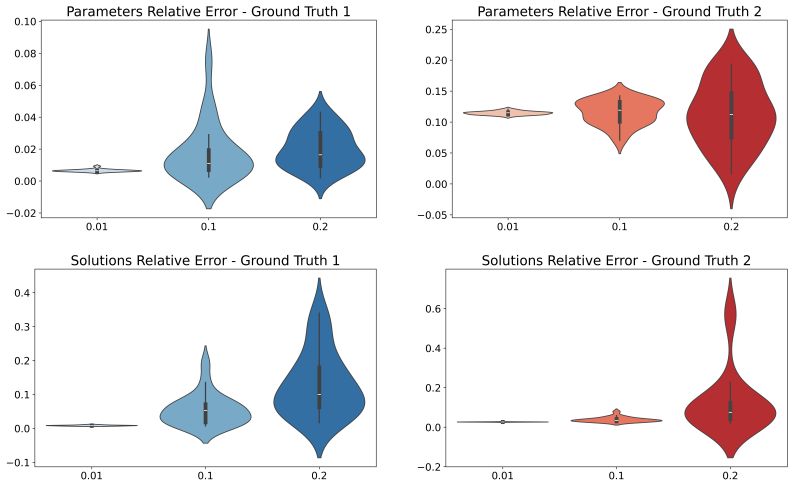

Figure 7.2: Base 1, violin plots of relative errors (7.33), (7.34) computed with best regularization parameter α_q^* for each trial $q = 1, ..., n$. First row shows results about parameters $\mathbf{m}$. Second row shows results about solutions $\mathbf{u}$. x-axis reports different levels of noise $\boldsymbol{\sigma} = [0.01, 0.1, 0.2]$. y-axis reports the relative error.

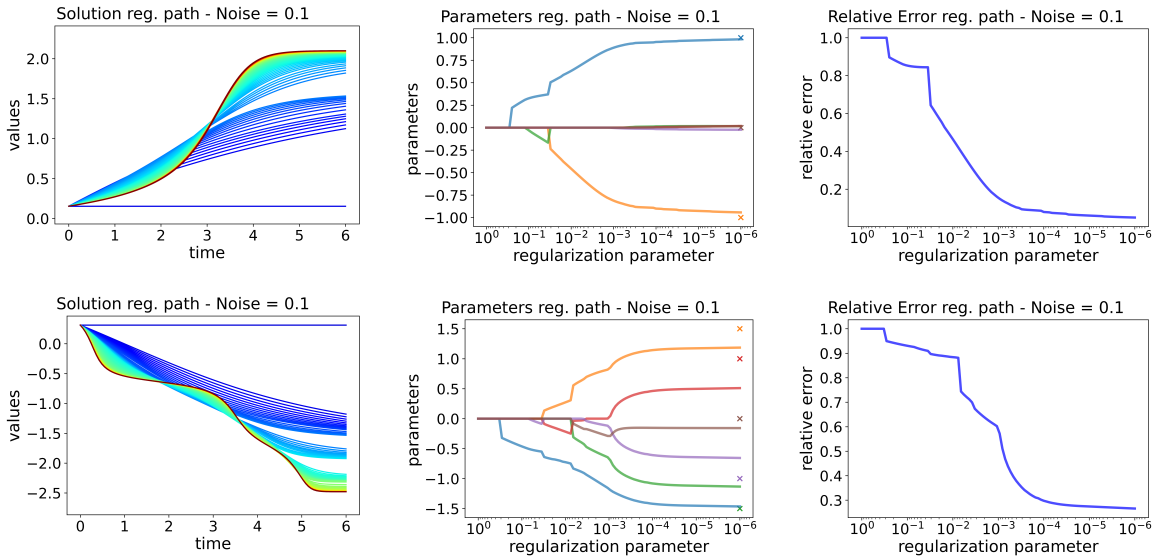

Figure 7.3: Base 1 regularization paths with level of noise $\sigma = 0.1$. First row shows results for $\mathbf{m}_1$ (7.38), second row shows results for $\mathbf{m}_2$ (7.39). The first column reports the regularization paths of the solutions. The colors of the curves transition from blue ($l = 0$) to red ($l = 99$). The x-axis reports the time, the y-axis the value of the state variable. The second column reports the regularization paths of the parameters. x-axis shows the regularization parameter, y-axis shows the parameter values. The ground truth values are symbolized by a cross at the last regularization parameter; the curves and crosses of the same color correspond to the same parameter. The third column shows the regularization path of the relative error (7.33). x-axis shows the regularization parameter, y-axis shows the relative error.

From Table 7.2 and the violin plots shown in Figure 7.2, it can be observed that parameter reconstruction is more accurate in the $\mathbf{m}_1$ case, that is, when the ground truth exhibits a sparser distribution. In particular, the relative error associated with $\mathbf{m}_1$ is approximately one order of magnitude lower than that observed in the $\mathbf{m}_2$ case. Another noteworthy aspect is the robustness of the proposed method to noise: the average relative error does not vary significantly as the noise level in the data increases, indicating a degree of stability in the reconstruction process. Despite the differing performance in estimating the two parameter sets, the relative errors on the recovered solutions remain low in both cases. Specifically, by analyzing the first column of Figure 7.3, it is evident that the regularization paths of the solutions converge to synthetic data presented in Figure 7.1. The second and third columns of the same figure further highlight the smoothness of the regularization path, which is characterized by a consistent and stable evolution without pronounced oscillations.

Overall, the conducted experiments provide valid proof of the self-consistency of the proposed method. The procedure is designed to promote sparsity in the solution of the inverse problem's parameter set by virtue of the regularization function's (7.31) specific

form. This feature is evident in the results: the best performance is observed when the ground truth has a sparse configuration.

Base 5

We show results for base 5 in Table 7.1. Ground truth coefficient vectors used are:

$$\mathbf{m}_1 = (0.5, 1, 0, 0, -0.5, 0, 0) \tag{7.40}$$

$$\mathbf{m}_2 = (0.6, -0.8, -0.2, 0.6, 0.3, 0, -0.2) \tag{7.41}$$

As in the previous case, we present the results using the same pattern and consider two sets of ground truth parameters. The first set is characterized by a sparse configuration, and the second set is characterized by a less sparse distribution.

Figure 7.4 shows the generated synthetic data for $\mathbf{m}_1$ (7.40) and $\mathbf{m}_2$ (7.41) under the three noise levels $\boldsymbol{\sigma} = [0.01, 0.1, 0.2]$.

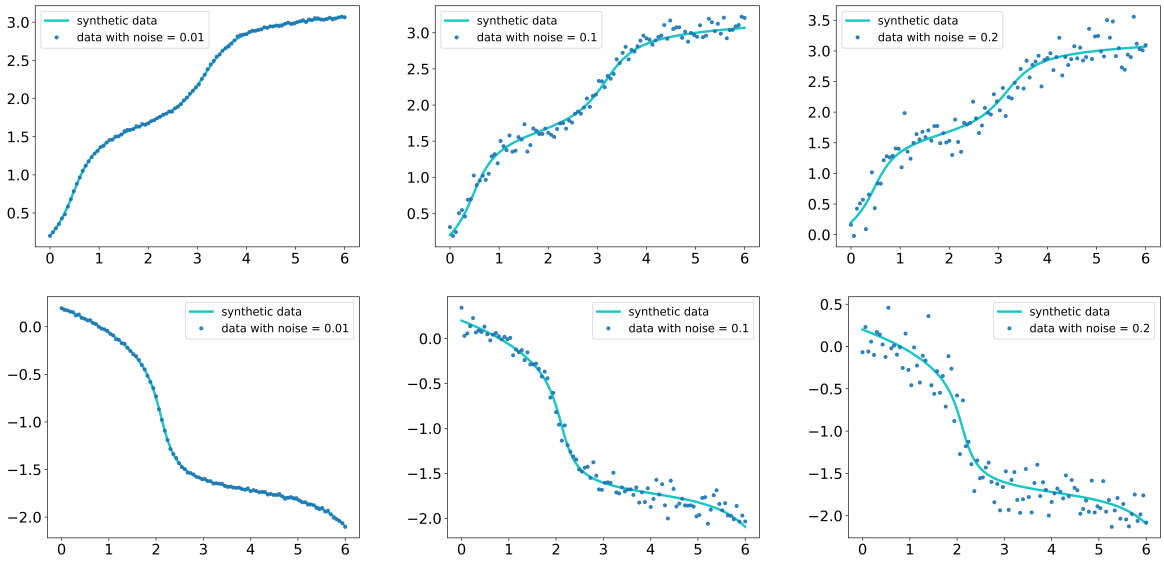

Figure 7.4: Base 5, $\mathbf{m}_1$ (first row), $\mathbf{m}_2$ (second row) synthetic data (light blue line) and noisy discrete synthetic data (blue dots) with different amount of noise: from left to right $\boldsymbol{\sigma} = [0.01, 0.1, 0.2]$. x-axis shows time. y-axis shows the values of the state variable.

Results for this base are summarized in Table 7.3 and illustrated via violin plots and regularization paths in Figures 7.5 and 7.7.

Table 7.3: Base 5. Parameters and solutions mean relative error table. Each entry reports the mean relative error ± standard deviation (7.36) (7.37). Each row is a different ground truth, each column is a different level of noise $\boldsymbol{\sigma} = [0.01, 0.1, 0.2]$

	Noise levels		
M	Low	Medium	High
$\mathbf{m}_1$	0.00 ± 0.00	0.03 ± 0.02	0.08 ± 0.06
$\mathbf{m}_2$	0.13 ± 0.00	0.13 ± 0.03	0.15 ± 0.06
$\mathbf{u}_1$	0.01 ± 0.00	0.16 ± 0.15	0.63 ± 0.94
$\mathbf{u}_1$	0.03 ± 0.00	0.09 ± 0.05	0.31 ± 0.22

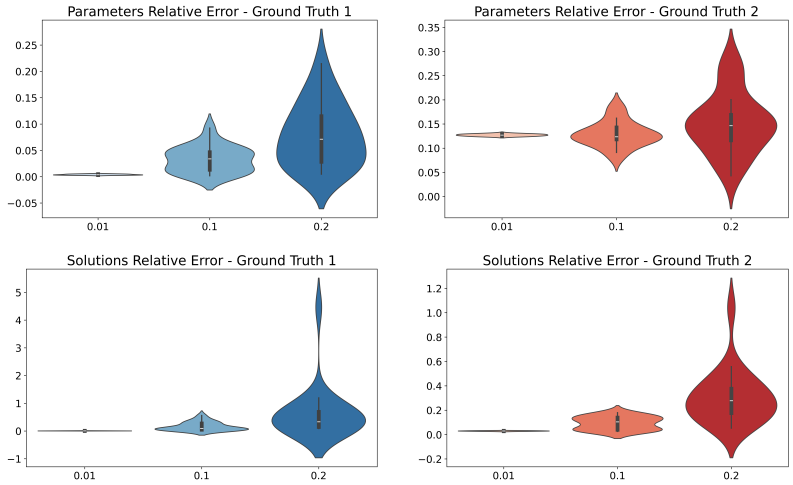

Figure 7.5: Base 5, violin plots of relative errors (7.33), (7.34) computed with best regularization parameter α_q^* for each trial $q = 1, ..., n$. First row shows results about parameters $\mathbf{m}$. Second row shows results about solutions $\mathbf{u}$. x-axis reports different levels of noise $\boldsymbol{\sigma} = [0.01, 0.1, 0.2]$. y-axis reports the relative error.

From Table 7.3, it can be observed that the parameters were successfully reconstructed. Once again, the sparser parameter set $\mathbf{m}_1$ (7.40) was recovered more accurately than the less sparse one, $\mathbf{m}_2$ (7.41). Moreover, the robustness of the reconstruction with respect to noise is preserved. The reconstruction errors of the forward solutions of the Cauchy problem (7.29) remain low for both the low ($\sigma = 0.01$) and the medium ($\sigma = 0.1$) noise levels. However, for higher noise levels ($\sigma = 0.2$), the solutions are no longer reconstructed accurately for either ground truth. As shown in Figure 7.6, this discrepancy arises from the fact that, although the data dynamics are well reconstructed, the incorrect initial condition causes the solution to diverge onto a different trajectory when solving the forward Cauchy problem (7.29). This phenomenon, caused by noise in the observed data, can be mitigated by applying smoothing techniques during the pre-processing phase.

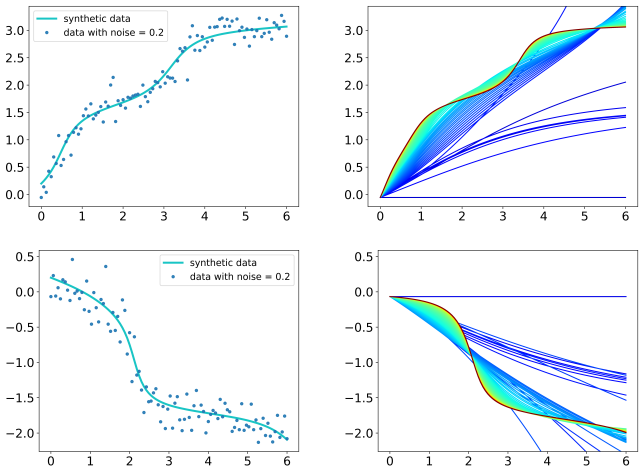

Figure 7.6: Base 5, $\mathbf{m}_1$ (7.40) (first row), $\mathbf{m}_2$ (7.41) (second row). First column shows synthetic data (light blue line) and noisy discrete synthetic data (blue dots) with noise level $\sigma = 0.1$.

Figure 7.7 displays the regularization paths of the solutions, parameters, and relative errors for both ground truths, $\mathbf{m}_1$ (7.40) and $\mathbf{m}_2$ (7.41). The regularization paths are generally smooth, except for an initial settling phase observed in $\mathbf{m}_1$. This behavior highlights the stability of the obtained solutions. Furthermore, while in the case of $\mathbf{m}_1$ the parameters converge accurately to their true values, in the case of the less sparse $\mathbf{m}_2$, the parameters converge stably but only approximately to the expected values.

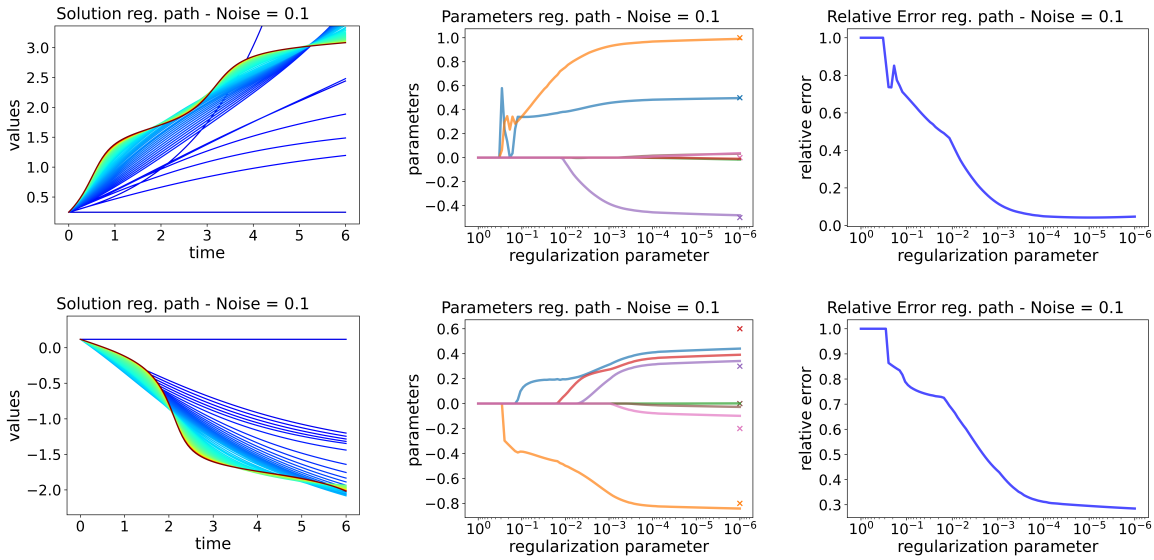

Figure 7.7: Base 5 regularization paths with level of noise $\sigma = 0.1$. First row shows results for $\mathbf{m}_1$ (7.40), second row shows results for $\mathbf{m}_2$ (7.41). The first column reports the regularization paths of the solutions. The colors of the curves transition from blue ($l = 0$) to red ($l = 99$). The x-axis reports the time, the y-axis the value of the state variable. The second column reports the regularization paths of the parameters. x-axis shows the regularization parameter, y-axis shows the parameter values. The ground truth values are symbolized by a cross at the last regularization parameter; the curves and crosses of the same color correspond to the same parameter. The third column shows the regularization path of the relative error (7.33). x-axis shows the regularization parameter, y-axis shows the relative error.

7.3.3 Theoretical discussion

Semi-convergent behavior

In this subsection, we demonstrate that the proposed method, when applied to this nonlinear inverse problem, exhibits a semi-convergent behavior — a phenomenon typically associated with linear inverse problems. To this end, we consider the results obtained by fixing basis 1 from Table 7.1, the sparse ground truth $\mathbf{m}_1$ defined in (7.38), and a noise level of $\sigma = 0.2$. With this noise level fixed, we generated $n = 20$ independent realizations, each corresponding to a distinct perturbation, thereby producing a diverse set of observed data. The choice of a relatively high noise level is driven by the intention to work in a framework where regularization becomes essential. Indeed, in the presence of significant noise, non-regularized inversion may capture and amplify spurious noise components, leading to numerically unstable solutions. Furthermore, the selected value of $\sigma = 0.2$ is sufficiently large to induce meaningful variability across realizations, while still ensuring that the perturbed data remain within a reasonable neighborhood of the noise-free reference. This setup enables a consistent and

informative evaluation of the algorithm's behavior under a fixed noise condition.

Figures 7.8 and 7.9 present two panels of size 5×4 that collect the results of the 20 trials performed. Specifically, Figure 7.8 shows the regularization paths of the relative error on the parameters, computed as in (7.33), while Figure 7.9 shows the evolution of the parameters themselves along the regularization path. Each panel box corresponds to a different realization of the noise. Images in the same position in both panels refer to the same realization of the noise.

Analysis of the results shows that, in some cases, the method behaves in a convergent manner. In other cases, semi-convergence is observed, which is a phenomenon typical of ill-posed linear inverse problems. When convergence occurs, the relative error decreases monotonically, even for very small values of the regularization parameter. This suggests that, although the problem is affected by noise, it is not critically unstable. This implies that parameter reconstruction remains stable even without strong regularization. In contrast, in cases of semi-convergence, the relative error initially decreases as the regularization parameter decreases. It then reaches a minimum and increases again for too small values of the regularization parameter. This behavior reflects the numerical instability of the problem. For low levels of regularization, the inversion algorithm starts reconstructing the spurious components introduced by the noise. Consequently, the estimated parameters deviate from the optimal ground truth configuration, which worsens the reconstruction. This behavior is also clearly seen in the regularization parameter paths shown in Figure 7.9. In cases of semi-convergence, the optimal parameter configuration, which is sparser and more consistent with the non-zero components of the ground truth, does not correspond to the minimum values of the regularization parameter. Rather, it lies in an intermediate zone of the regularization path.

In spite of this, a relevant point to note is that the forward solution of the Cauchy problem (7.29), obtained with the parameters along the regularization path, continues to approximate the data better and better effectively even when the semi-convergent phenomenon occurs. This behavior underscores the ill-posed nature of the problem and highlights the importance of regularization in balancing fidelity to the data and numerical stability in parameter estimation.

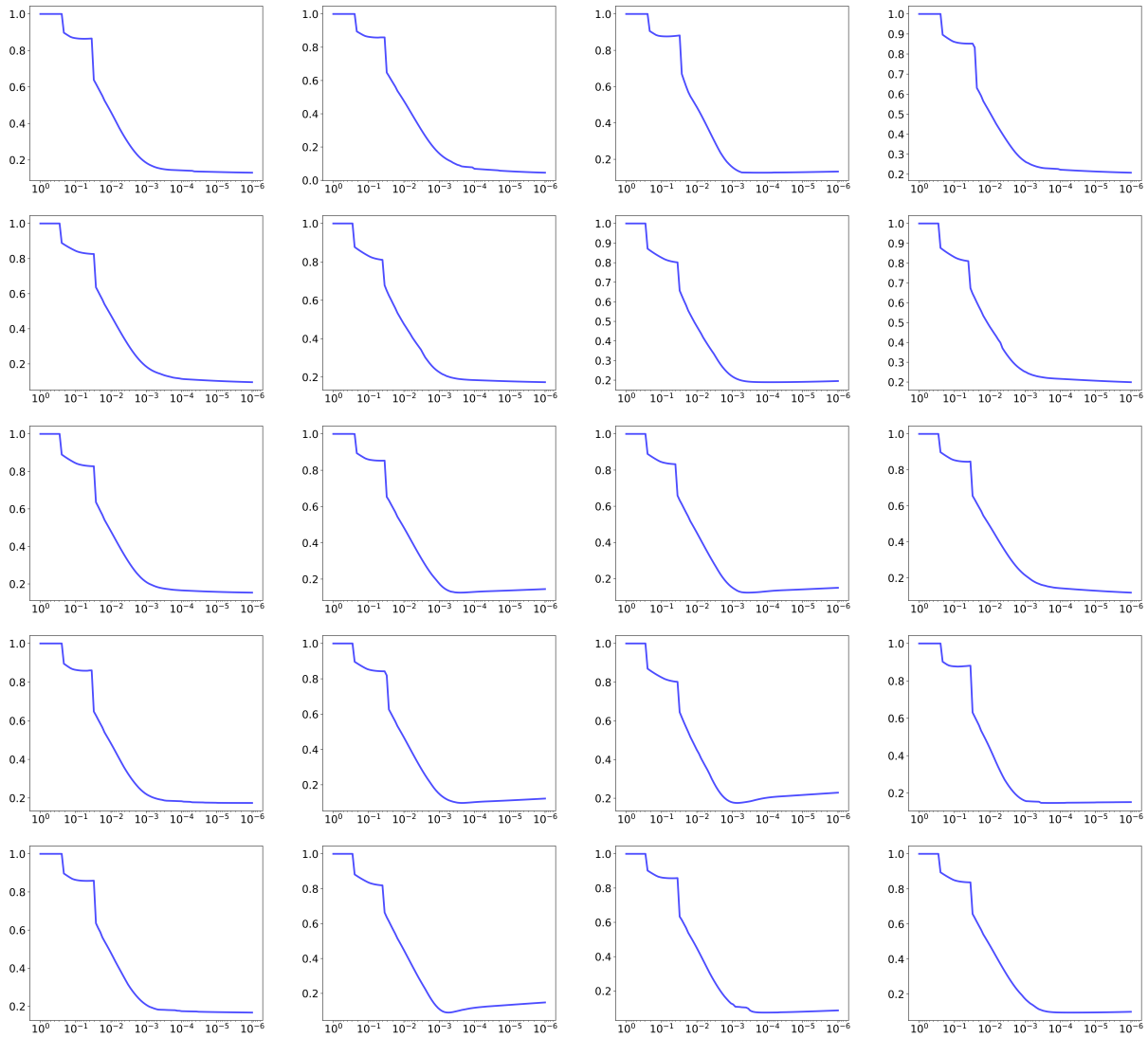

Figure 7.8: Base 1, regularization path of the relative error (7.33) with noise level $\sigma = 0.2$. x-axis shows the regularization parameter, y-axis shows the relative error.

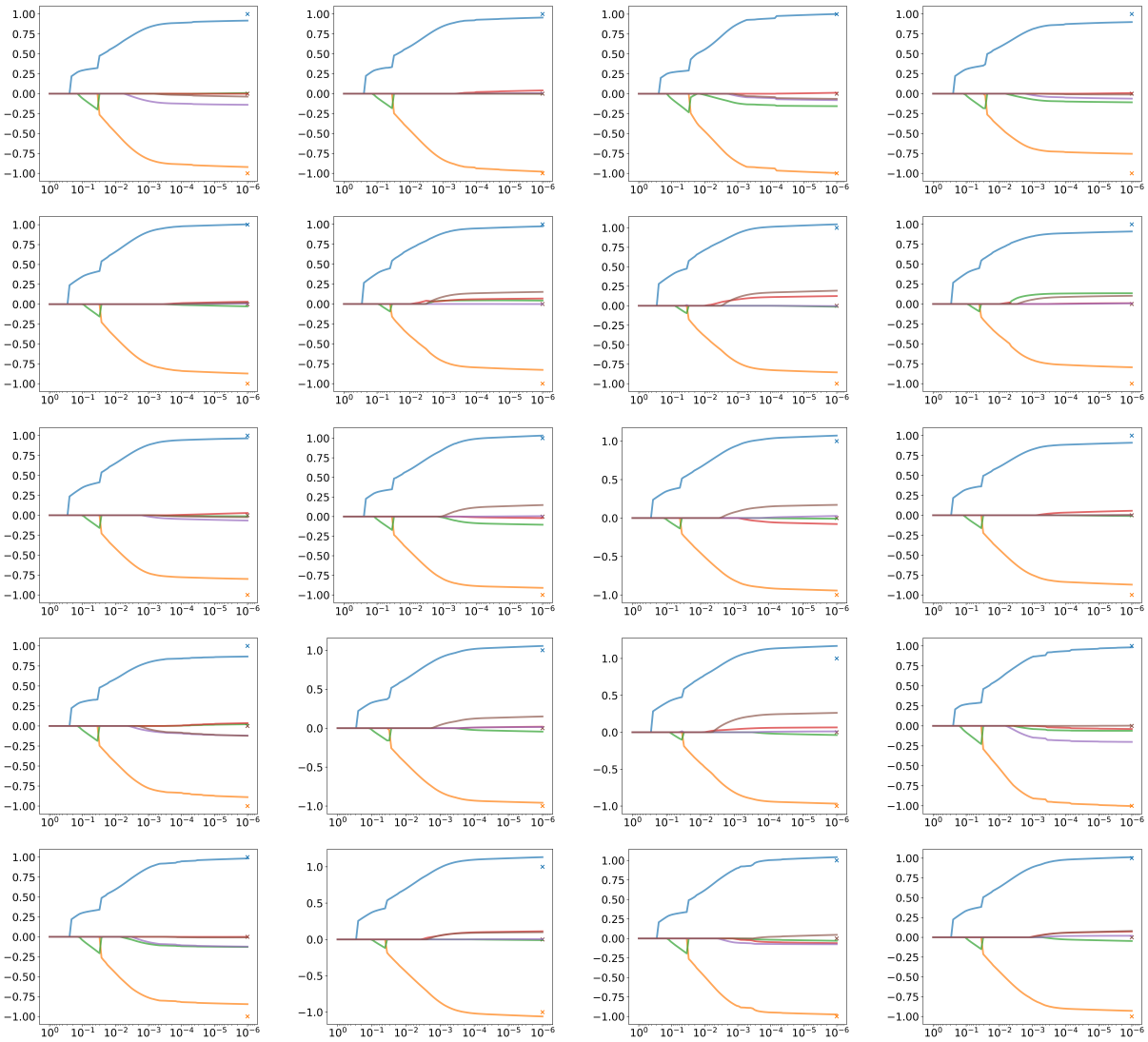

Figure 7.9: Base 1, regularization paths of the parameters. x-axis shows the regularization parameter, y-axis shows the parameter values. The ground truth values are symbolized by a cross at the last regularization parameter; the curves and crosses of the same color correspond to the same parameter.

Bifurcations of the regularization path

In this section, we investigate the influence of algorithmic hyperparameters on the regularization path, highlighting their impact on the solution process. It is important to emphasize that hyperparameters themselves must be carefully tuned or adapted depending on the specific case. In this analysis, we focus on the gradient step size τ as the primary hyperparameter of interest. We conducted a series of experiments varying the value of τ: specifically, $\tau = 10^{-2}$ and $\tau = 10^{-3}$.

We used base 2 from Table (7.1), and the following ground truth parameter vector:

$$\mathbf{m} = (0, 1, -1, 0, 0, 0, 0) \tag{7.42}$$

Figure 7.10 shows synthetic and noisy data.

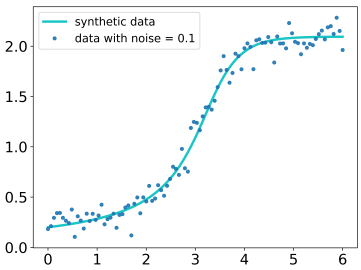

Figure 7.10: Base 2, $\mathbf{m}$ (7.42) synthetic data (light blue line) and noisy discrete synthetic data (blue dots) with level of noise $\sigma = 0.1$. x-axis shows time. y-axis shows the values of the state variable.

Figure 7.11 displays the resulting regularization paths corresponding to these settings. From this figure, it is clear that the gradient step size τ significantly affects the behavior of the regularization paths. In the case of $\tau = 10^{-2}$, the paths are notably unstable: the larger step size causes more drastic parameter updates at each iteration. As a result, the parameters exhibit sharp fluctuations. However, starting from a regularization parameter $\alpha \sim 10^{-3}$, the magnitude of these oscillations begins to decrease, and the parameters gradually stabilize near the ground truth values. The observed zigzag behavior reflects the fact that the parameters oscillate around the minimum without reaching it, due to the excessive step size. The corresponding solutions to the associated Cauchy problem (7.29) also display significant oscillations, although they eventually settle into a good approximation of the observed data.

Conversely, when using a smaller step size $\tau = 10^{-3}$, the regularization paths become much smoother and more gradual. However, because the gradient updates are (in this specific case) too small, the algorithm can't reconstruct the ground truth parameters. Instead, the optimization converges to an alternative set of parameters. Despite this, the resulting solutions to the forward problem exhibit a smooth trajectory and are in strong agreement with the observed data.

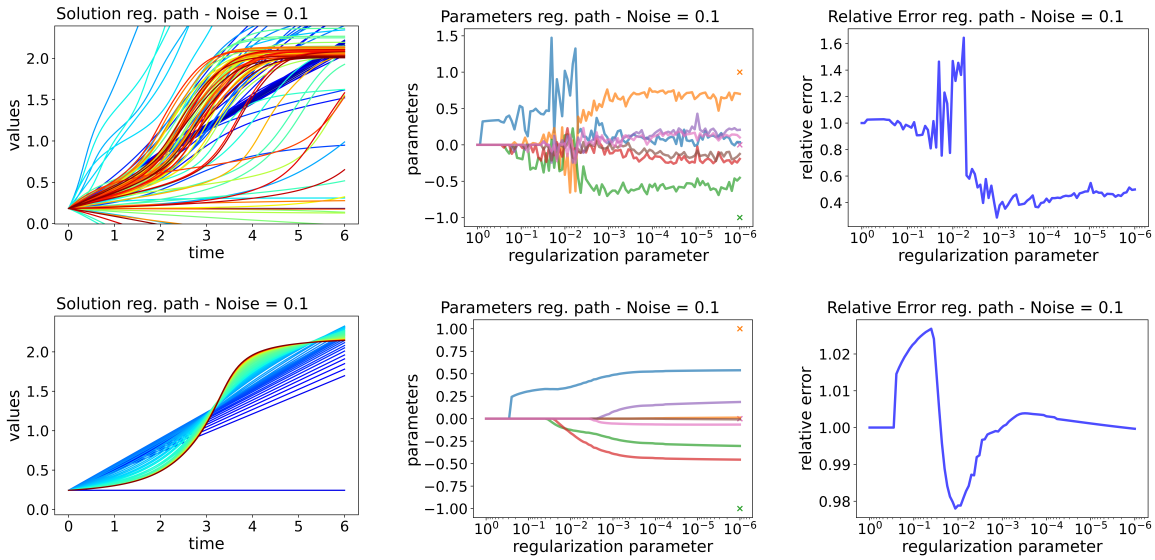

Figure 7.11: Base 2 regularization paths with ground truth parameter $\mathbf{m}$ (7.42), level of noise $\sigma = 0.1$. First row shows results gradient step $\tau = 10^{-2}$, second row shows results for $\tau = 10^{-3}$. The first column reports the regularization paths of the solutions. The colors of the curves transition from blue ($l = 0$) to red ($l = 99$). The x-axis reports the time, the y-axis the value of the state variable. The second column reports the regularization paths of the parameters. x-axis shows the regularization parameter, y-axis shows the parameter values. The ground truth values are symbolized by a cross at the last regularization parameter; the curves and crosses of the same color correspond to the same parameter. The third column shows the regularization path of the relative error (7.33). x-axis shows the regularization parameter, y-axis shows the relative error.

These results illustrate a trade-off: with $\tau = 10^{-2}$, the optimization progresses more quickly and drives the parameters closer to the minimum, but the updates are too aggressive, leading to persistent oscillations around the solution (a phenomenon often referred to as "zigzagging"). On the other hand, $\tau = 10^{-3}$ yields a much more stable regularization path, but converges to a different set of parameters whose associated solutions to the Cauchy problem (7.29) still fit the data well; this highlights the ill-posedness and non-uniqueness inherent in the inverse problem.

7.4 Conclusions

In this Chapter, we proposed a homotopy-based regularization path method for solving implicit inverse problems. In particular we wanted determine a non-linear functional applied to physical quantities directly from their noisy measurements. Assuming that this functional had a parametric form (7.1), the implicit inverse problem became an optimization

parameters problem. By introducing a regularized objective function (7.4), the problem has been reframed as a discretized, regularized, constrained minimization problem (7.8). Fixed the regularization parameter, a variational approach was used and a gradient descent-based algorithm (7.18) was obtained. To compute large Jacobians (7.15) efficiently, the adjoint state method was employed (7.16). To obtain a solution to the non-regularized constrained problem (7.9), we applied a continuation strategy based on the regularization path and warm-start initialization (7.25). This process culminated in the proposed algorithm summarized in pseudocode (3).

The method was then specialized to a latent dynamics discovery problem, where the unknown relationships were modeled as systems of ordinary differential equations (ODEs) and the data were given as time series. The non-linear unknown functional was expressed as a linear combination of non-linear basis functions (7.27), as in (7.28), and the forward problem was formulated as the Cauchy problem (7.29). Numerical experiments in a synthetic, one-dimensional setting validated the effectiveness of the proposed method.

A key limitation of the approach lies in the design of the candidate basis function library. The basis must be expressive enough to capture the dynamics but not so large as to introduce unnecessary complexity. Additionally, the approach currently requires manual computation of the basis function derivatives, which can be burdensome as the basis grows.

Future work will focus on enhancing algorithmic efficiency, particularly when extending the basis function set. This includes exploring automatic differentiation to ease the derivative computation process and methods to scale the basis function library systematically. Furthermore, we aim to extend the methodology to more complex dynamic systems, such as those involving multiple physical quantities, spatiotemporal phenomena described by partial differential equations (PDEs), and systems with inequality constraints.

Chapter 8

Thesis Conclusions

This thesis began with a focused application into an unexpected phenomenon observed in FDG-PET imaging of lung cancer patients and gradually expanded to explore general mathematical and computational frameworks for inverse problems. Each chapter represented a step along this path - from clinical motivation and physiological modeling to formulation of dynamical systems, optimization and regularization theory.

In Chapter 3, we introduced the fundamental principles of PET. We described the biological role of glucose and its analog FDG, the physical mechanisms underlying PET signal generation, and the challenges associated with image reconstruction and quantification. Both analytical and iterative methods for tomographic reconstruction were reviewed, and the clinical use of SUV was discussed critically in light of its interpretative limitations.

Chapter 4 focused on compartmental modeling as a method for quantitatively interpreting PET tracer kinetics. We presented classical models such as the Sokoloff model and framed the associated inverse problems of parameter estimation. Different solution strategies were explored and we also compared parametric imaging with traditional ROI-based approaches, emphasizing their respective advantages and trade-offs.

In Chapter 5, we applied Patlak graphical techniques in a voxel-wise fashion to dynamic FDG-PET data from lung cancer patients, obtained thanks to a collaboration with the Department of Nuclear Medicine at IRCCS Ospedale Policlinico San Martino in Genova. This study revealed localized washout phenomena, manifesting as negative slopes in Patlak plots, predominantly located in tumor peripheries. These findings highlighted underlying metabolic heterogeneity and suggested a spatial association with inflammatory infiltration.

Chapter 6 introduced a more general mathematical setting by interpreting compartmental models as latent dynamical systems governed by ODEs. We presented the ODE learning paradigm, reviewed common computational approaches (such as embedding neural networks, sparse identification techniques or Bayesian frameworks with ODEs), and analyzed the chal-

lenges posed by applying these tools to compartmental inverse problems.

Chapter 7 constituted the most abstract and general contribution of the thesis. We reformulated the learning of tracer dynamics as an implicit inverse problem, where the unknown is defined via a nonlinear functional with a known parametric form. We proposed a novel two-level solution strategy: an inner optimization loop solves a constrained problem using the Lagrangian formalism, Newton–Raphson method, and adjoint-state sensitivities; an outer homotopy loop traces a solution path from strong regularization (well-posed) to minimal regularization (ill-posed). This homotopy-based approach was tested on one-dimensional problems with noisy and sparse data, demonstrating its ability to recover functional parameters using various basis expansions. We also analyzed critical phenomena such as semi-convergence and bifurcations in the regularization path, which serve as indicators of ill-posedness and non-uniqueness of the inverse problem.

The theoretical and numerical results presented in this thesis open several directions for future work. A natural extension involves applying the implicit regularization method to spatiotemporal problems in higher dimensions, incorporating more realistic physiological constraints. Another promising direction is the integration of machine learning tools—such as neural architectures, automatic differentiation, and symbolic regression—to automate the selection of functional structures and model parametrizations. This would allow for the simultaneous identification of both the form and the parameters of the governing dynamics directly from data.

The research developed in this thesis has led to the following publications:

- Davide Parodi, Edoardo Dighero, Giorgia Biddau, Francesca D'Amico, Matteo Bauckneht, Cecilia Marini, Sara Garbarino, Cristina Campi, Michele Piana, and Gianmario Sambuceti. *Localized FDG loss in lung cancer lesions.* EJNMMI Research, 14(1):102, 2024.

- Davide Parodi. *Solving Implicit Inverse Problems with Homotopy-Based Regularization Path.* arXiv preprint doi:2505.19608, 2025.

Bibliography

[1] Alen Alexanderian. Optimal experimental design for infinite-dimensional bayesian inverse problems governed by pdes: A review. *Inverse Problems*, 37(4):043001, 2021.

[2] Nasser K Altorki, Geoffrey J Markowitz, Dingcheng Gao, Jeffrey L Port, Ashish Saxena, Brendon Stiles, Timothy McGraw, and Vivek Mittal. The lung microenvironment: an important regulator of tumour growth and metastasis. *Nature Reviews Cancer*, 19(1):9–31, 2019.

[3] Jorn Baayen, Bernhard Becker, Klaas-Jan van Heeringen, Ivo Miltenburg, Teresa Piovesan, Julia Rauw, Matthijs den Toom, and Jesse VanderWees. An overview of continuation methods for non-linear model predictive control of water systems. *IFAC-PapersOnLine*, 52(23):73–80, 2019.

[4] David Barber and Yali Wang. Gaussian processes for bayesian estimation in ordinary differential equations. In *International conference on machine learning*, pages 1485–1493. PMLR, 2014.

[5] Frank M Bengel, Peter Ueberfuhr, Nina Schiepel, Stephan G Nekolla, Bruno Reichart, and Markus Schwaiger. Effect of sympathetic reinnervation on cardiac performance after heart transplantation. *New England Journal of Medicine*, 345(10):731–738, 2001.

[6] Jens Berg and Kaj Nyström. Data-driven discovery of pdes in complex datasets. *Journal of Computational Physics*, 384:239–252, 2019.

[7] Maïtine Bergounioux, Élie Bretin, and Yannick Privat. How to position sensors in thermo-acoustic tomography. *Inverse Problems*, 35(7):074003, 2019.

[8] Dimitri P Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[9] Ronald Boellaard, Roberto Delgado-Bolton, Wim JG Oyen, Francesco Giammarile, Klaus Tatsch, Wolfgang Eschner, Fred J Verzijlbergen, Sally F Barrington, Lucy C

Pike, Wolfgang A Weber, et al. Fdg pet/ct: Eanm procedure guidelines for tumour imaging: version 2.0. *European journal of nuclear medicine and molecular imaging*, 42:328–354, 2015.

[10] Gert-Jan Both, Subham Choudhury, Pierre Sens, and Remy Kusters. Deepmod: Deep learning for model discovery in noisy data. *Journal of Computational Physics*, 428:109985, 2021.

[11] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.

[12] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[13] Steven Brunton, Joshua Proctor, and Nathan Kutz. Sparse identification of nonlinear dynamics (sindy). In *APS Division of Fluid Dynamics Meeting Abstracts*, pages L8–005, 2016.

[14] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, 113(15):3932–3937, 2016.

[15] Martin Burger, Jan-Frederik Pietschmann, and Marie-Therese Wolfram. Data assimilation in price formation. *Inverse Problems*, 36(6):064003, 2020.

[16] Ben Calderhead, Mark Girolami, and Neil Lawrence. Accelerating bayesian inference over nonlinear differential equations with gaussian processes. *Advances in neural information processing systems*, 21, 2008.

[17] Richard E Carson. Tracer kinetic modeling in pet. In *Positron emission tomography: Basic sciences*, pages 127–159. Springer, 2005.

[18] Rick Chartrand. Numerical differentiation of noisy, nonsmooth data. *International Scholarly Research Notices*, 2011(1):164564, 2011.

[19] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.

[20] Zhao Chen, Yang Liu, and Hao Sun. Physics-informed learning of governing equations from scarce data. *Nature communications*, 12(1):6136, 2021.

[21] Zixiang Chen, Zhaoping Cheng, Yanhua Duan, Qiyang Zhang, Na Zhang, Fengyun Gu, Ying Wang, Yun Zhou, Haining Wang, Dong Liang, et al. Accurate total-body ki parametric imaging with shortened dynamic 18f-fdg pet scan durations via effective data processing. *Medical Physics*, 50(4):2121–2134, 2023.

[22] Christian Crone. The permeability of capillaries in various organs as determined by use of the 'indicator diffusion'method. *Acta physiologica scandinavica*, 58(4):292–305, 1963.

[23] Eleonore H de Groot, Nieky Post, Ronald Boellaard, Nils RL Wagenaar, Antoon TM Willemsen, and Jorn A van Dalen. Optimized dose regimen for whole-body fdg-pet imaging. *EJNMMI research*, 3:1–11, 2013.

[24] Jean-Pierre Dedieu. *Newton-Raphson Method*, pages 1023–1028. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015.

[25] Michel Defrise, Christian Vanhove, and Xuan Liu. An algorithm for total variation regularization in high-dimensional linear problems. *Inverse Problems*, 27(6):065002, 2011.

[26] Fabrice Delbary, Sara Garbarino, and Valentina Vivaldi. Compartmental analysis of dynamic nuclear medicine data: models and identifiability. *Inverse Problems*, 32(12):125010, 2016.

[27] Dominique Delbeke, R Edward Coleman, Milton J Guiberteau, Manuel L Brown, Henry D Royal, Barry A Siegel, David W Townsend, Lincoln L Berland, J Anthony Parker, Karl Hubner, et al. Procedure guideline for tumor imaging with 18f-fdg pet/ct 1.0. *Journal of nuclear Medicine*, 47(5):885–895, 2006.

[28] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22, 1977.

[29] Frank Dondelinger, Dirk Husmeier, Simon Rogers, and Maurizio Filippone. Ode parameter inference using adaptive gradient matching with gaussian processes. In *Artificial intelligence and statistics*, pages 216–228. PMLR, 2013.

[30] Juergen Dukart, Karsten Mueller, Annette Horstmann, Barbara Vogt, Stefan Frisch, Henryk Barthel, Georg Becker, Harald E Möller, Arno Villringer, Osama Sabri, et al. Differential effects of global and cerebellar normalization on detection and differentiation of dementia in fdg-pet studies. *Neuroimage*, 49(2):1490–1495, 2010.

[31] Stephen P Ellner, Yodit Seifu, and Robert H Smith. Fitting population dynamic models to time-series data by gradient matching. *Ecology*, 83(8):2256–2270, 2002.

[32] Neil D Evans, Rachel J Errington, Michael J Chapman, Paul J Smith, Michael J Chappell, and Keith R Godfrey. Compartmental modelling of the uptake kinetics of the anti-cancer agent topotecan in human breast cancer cells. *International Journal of Adaptive Control and Signal Processing*, 19(5):395–417, 2005.

[33] Alan J Fischman and Nathaniel M Alpert. Fdg-pet in oncology: there's more to it than looking at pictures, 1993.

[34] Kirk A Frey, Vjera A Holthoff, Robert A Koeppe, Douglas M Jewett, Michael R Kilbourn, and David E Kuhl. Parametric in vivo imaging of benzodiazepine receptor distribution in human brain. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 30(5):663–672, 1991.

[35] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1, 2010.

[36] 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. *Physics in Medicine & Biology*, 59(10):2469, 2014.

[37] Sara Garbarino, Valentina Vivaldi, Fabrice Delbary, Giacomo Caviglia, Michele Piana, Cecilia Marini, Selene Capitanio, Iolanda Calamia, Ambra Buschiazzo, and Gianmario Sambuceti. A new compartmental method for the analysis of liver fdg kinetics in small animal models. *EJNMMI research*, 5:1–9, 2015.

[38] Zoltan German-Sallo. Nonlinear wavelet denoising of data signals. *UbiCC J*, 6:895–900, 2011.

[39] Gene H Golub, Per Christian Hansen, and Dianne P O'Leary. Tikhonov regularization and total least squares. *SIAM journal on matrix analysis and applications*, 21(1):185–194, 1999.

[40] Thore Graepel. Solving noisy linear operator equations by gaussian processes: Application to ordinary and partial differential equations. In *ICML*, volume 3, pages 234–241, 2003.

[41] Roger N Gunn, Steve R Gunn, and Vincent J Cunningham. Positron emission tomography compartmental models. *Journal of Cerebral Blood Flow & Metabolism*, 21(6):635–652, 2001.

[42] Roger N Gunn, Adriaan A Lammertsma, Susan P Hume, and Vincent J Cunningham. Parametric imaging of ligand-receptor binding in pet using a simplified reference region model. *Neuroimage*, 6(4):279–287, 1997.

[43] Jacques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, pages 49–52, 1902.

[44] Martin Hanke, Andreas Neubauer, and Otmar Scherzer. A convergence analysis of the landweber iteration for nonlinear ill-posed problems. *Numerische Mathematik*, 72(1):21–37, 1995.

[45] Trevor Hastie. The elements of statistical learning: data mining, inference, and prediction, 2009.

[46] Markus Heinonen, Cagatay Yildiz, Henrik Mannerström, Jukka Intosalmi, and Harri Lähdesmäki. Learning unknown ode models with gaussian processes. In *International conference on machine learning*, pages 1959–1968. PMLR, 2018.

[47] H Malcolm Hudson and Richard S Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE transactions on medical imaging*, 13(4):601–609, 1994.

[48] Masanori Ichise and James R Ballinger. From graphical analysis to multilinear regression analysis of reversible radioligand binding. *Journal of Cerebral Blood Flow & Metabolism*, 16(4):750–751, 1996.

[49] Masanori Ichise, Jeih-San Liow, Jian-Qiang Lu, Akihiro Takano, Kendra Model, Hiroshi Toyama, Tetsuya Suhara, Kazutoshi Suzuki, Robert B Innis, and Richard E Carson. Linearized reference tissue parametric imaging methods: application to [11c] dasb positron emission tomography studies of the serotonin transporter in human brain. *Journal of Cerebral Blood Flow & Metabolism*, 23(9):1096–1112, 2003.

[50] T Ido, C-N Wan, Va Casella, JS Fowler, AP Wolf, M Reivich, and DE Kuhl. Labeled 2-deoxy-d-glucose analogs. 18f-labeled 2-deoxy-2-fluoro-d-glucose, 2-deoxy-2-fluoro-d-mannose and 14c-2-deoxy-2-fluoro-d-glucose. *Journal of Labelled Compounds and Radiopharmaceuticals*, 14(2):175–183, 1978.

[51] Emmanuel Itti, Michel Meignan, Alina Berriolo-Riedinger, Alberto Biggi, Amanda F Cashen, Pierre Véra, Hervé Tilly, Barry A Siegel, Andrea Gallamini, René-Olivier Casasnovas, et al. An international confirmatory study of the prognostic value of early pet/ct in diffuse large b-cell lymphoma: comparison between deauville criteria and δsuvmax. *European journal of nuclear medicine and molecular imaging*, 40:1312–1320, 2013.

[52] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, et al. *An introduction to statistical learning*, volume 112. Springer, 2013.

[53] Hyun Sik Jun, Yuk Yin Cheung, Young Mok Lee, Brian C Mansfield, and Janice Y Chou. Glucose-6-phosphatase-β, implicated in a congenital neutropenia syndrome, is essential for macrophage energy homeostasis and functionality. *Blood, The Journal of the American Society of Hematology*, 119(17):4047–4055, 2012.

[54] Mustafa E Kamasak, Charles A Bouman, Evan D Morris, and Ken Sauer. Direct reconstruction of kinetic parameter images from dynamic pet data. *IEEE transactions on medical imaging*, 24(5):636–650, 2005.

[55] Sanghee Kang, Eui Hyun Kim, Jun-Eul Hwang, Ji-Hyun Shin, Yun Seong Jeong, Sun Young Yim, Eun Wook Joo, Young Gyu Eun, Dong Jin Lee, Bo Hwa Sohn, et al. Prognostic significance of high metabolic activity in breast cancer: Pet signature in breast cancer. *Biochemical and biophysical research communications*, 511(1):185–191, 2019.

[56] Nicolas A Karakatsanis, Yun Zhou, Martin A Lodge, Michael E Casey, Richard L Wahl, Habib Zaidi, and Arman Rahmim. Generalized whole-body patlak parametric imaging for enhanced quantification in clinical pet. *Physics in Medicine & Biology*, 60(22):8643, 2015.

[57] Jan D Keller and Roland Potthast. Ai-based data assimilation: Learning the functional of analysis estimation. *arXiv preprint arXiv:2406.00390*, 2024.

[58] Seymour S Kety. The theory and applications of the exchange of inert gas at the lungs and tissues. *Pharmacological Reviews*, 3:1–41, 1951.

[59] Yuichi Kimura, Mika Naganawa, Miho Shidahara, Yoko Ikoma, and Hiroshi Watabe. Pet kinetic analysis—pitfalls and a solution for the logan plot. *Annals of nuclear medicine*, 21:1–8, 2007.

[60] Teresa B Klepser and Michael W Kelly. Metformin hydrochloride: an antihyperglycemic agent. *American journal of health-system pharmacy*, 54(8):893–903, 1997.

[61] Annemieke C Kole, Omgo E Nieweg, Jan Pruim, Anne MJ Paans, John Th M Plukker, Harald J Hoekstra, Heimen Schraffordt Koops, and Willem Vaalburg. Standardized uptake value and quantification of metabolism for breast cancer imaging with fdg and l-[1-11c] tyrosine pet. *Journal of Nuclear Medicine*, 38(5):692–696, 1997.

[62] Lale Kostakoglu, Heiko Schöder, Jeffrey L Johnson, Nathan C Hall, Lawrence H Schwartz, David J Straus, Ann S LaCasce, Sin-Ho Jung, Nancy L Bartlett, George P Canellos, et al. 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? *Leukemia & lymphoma*, 53(11):2143, 2012.

[63] Jan Kukačka, Vladimir Golkov, and Daniel Cremers. Regularization for deep learning: A taxonomy. *arXiv preprint arXiv:1710.10686*, 2017.

[64] Leonid A Kunyansky. A new spect reconstruction algorithm based on thenovikov explicit inversion formula. *Inverse problems*, 17(2):293, 2001.

[65] Eric Laffon and Roger Marthan. Is patlak y-intercept a relevant metrics?, 2021.

[66] Kenneth Lange and Jeffrey A Fessler. Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Transactions on Image Processing*, 4(10):1430–1438, 1995.

[67] Henri Lanteri, Muriel Roche, and Claude Aime. Penalized maximum likelihood image restoration with positivity constraints: multiplicative algorithms. *Inverse problems*, 18(5):1397, 2002.

[68] Jean Logan. A review of graphical methods for tracer studies and strategies to reduce bias. *Nuclear medicine and biology*, 30(8):833–844, 2003.

[69] Jean Logan, Joanna S Fowler, Nora D Volkow, Alfred P Wolf, Stephen L Dewey, David J Schlyer, Robert R MacGregor, Robert Hitzemann, Bernard Bendriem, S John Gatley, et al. Graphical analysis of reversible radioligand binding from time—activity measurements applied to [n-11c-methyl]-(-)-cocaine pet studies in human subjects. *Journal of Cerebral Blood Flow & Metabolism*, 10(5):740–747, 1990.

[70] Marco Lorenzi and Maurizio Filippone. Constraining the dynamics of deep probabilistic models. In *International Conference on Machine Learning*, pages 3227–3236. PMLR, 2018.

[71] Benn Macdonald. *Statistical inference for ordinary differential equations using gradient matching*. PhD thesis, University of Glasgow, 2017.

[72] Cecilia Marini, Silvia Ravera, Ambra Buschiazzo, Giovanna Bianchi, Anna Maria Orengo, Silvia Bruno, Gianluca Bottoni, Laura Emionite, Fabio Pastorino, Elena Monteverde, et al. Discovery of a novel glucose metabolism in cancer: The role of endoplasmic reticulum beyond glycolysis and pentose phosphate shunt. *Scientific reports*, 6(1):25092, 2016.

[73] Michela Massollo, Cecilia Marini, Massimo Brignone, Laura Emionite, Barbara Salani, Mattia Riondato, Selene Capitanio, Francesco Fiz, Alessia Democrito, Adriana Amaro, et al. Metformin temporal and localized effects on gut glucose metabolism assessed using 18f-fdg pet in mice. *Journal of Nuclear Medicine*, 54(2):259–266, 2013.

[74] Corbin E Meacham and Sean J Morrison. Tumour heterogeneity and cancer cell plasticity. *Nature*, 501(7467):328–337, 2013.

[75] Daniela Melis, Fortunata Carbone, Giorgia Minopoli, Claudia La Rocca, Francesco Perna, Veronica De Rosa, Mario Galgani, Generoso Andria, Giancarlo Parenti, and Giuseppe Matarese. Cutting edge: increased autoimmunity risk in glycogen storage disease type 1b is associated with a reduced engagement of glycolysis in t cells and an impaired regulatory t cell function. *The Journal of Immunology*, 198(10):3803–3808, 2017.

[76] Stanisław Migórski, Akhtar A Khan, and Shengda Zeng. Inverse problems for nonlinear quasi-variational inequalities with an application to implicit obstacle problems of p-laplacian type. *Inverse Problems*, 35(3):035004, 2019.

[77] Klaus Mosegaard and Camilla Rygaard-Hjalsted. Probabilistic analysis of implicit inverse problems. *Inverse problems*, 15(2):573, 1999.

[78] William W Moses. Fundamental limits of spatial resolution in pet. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 648:S236–S240, 2011.

[79] Frank Natterer. Inversion ofthe attenuated radon transform. *Inverse problems*, 17(1):113, 2001.

[80] Carol A Nelson, Jennifer Q Wang, Irwin Leav, and Paul D Crane. The interaction among glucose transport, hexokinase, and glucose-6-phosphatase with respect to

3h-2-deoxyglucose retention in murine tumor models. *Nuclear medicine and biology*, 23(4):533–541, 1996.

[81] Elias D Nino-Ruiz, Carlos Ardila, and Rafael Capacho. Local search methods for the solution of implicit inverse problems. *Soft Computing*, 22:4819–4832, 2018.

[82] John M Ollinger and Jeffrey A Fessler. Positron-emission tomography. *Ieee signal processing magazine*, 14(1):43–55, 1997.

[83] Josef Pacák, Zdeněk Točík, and Miloslav Černỳ. Synthesis of 2-deoxy-2-fluoro-d-glucose. *Journal of the Chemical Society D: Chemical Communications*, 2:77–77, 1969.

[84] Davide Parodi, Federico Benvenuto, Sara Garbarino, and Michele Piana. Solving implicit inverse problems with homotopy-based regularization path, 2025.

[85] Davide Parodi, Edoardo Dighero, Giorgia Biddau, Francesca D'Amico, Matteo Bauckneht, Cecilia Marini, Sara Garbarino, Cristina Campi, Michele Piana, and Gianmario Sambuceti. Localized fdg loss in lung cancer lesions. *EJNMMI research*, 14(1):102, 2024.

[86] Paulina Pathria, Tiani L Louis, and Judith A Varner. Targeting tumor-associated macrophages in cancer. *Trends in immunology*, 40(4):310–327, 2019.

[87] Clifford S Patlak and Ronald G Blasberg. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. generalizations. *Journal of Cerebral Blood Flow & Metabolism*, 5(4):584–590, 1985.

[88] Silvia Peppicelli, Elena Andreucci, Jessica Ruzzolini, Francesca Bianchini, and Lido Calorini. Fdg uptake in cancer: a continuing debate. *Theranostics*, 10(7):2944, 2020.

[89] Michael E Phelps. Positron computed tomography studies of cerebral glucose metabolism in man: theory and application in nuclear medicine. In *Seminars in nuclear medicine*, volume 11, pages 32–49. Elsevier, 1981.

[90] R-E Plessix. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503, 2006.

[91] Quantitative FDG-PET Technical Committee Quantitative Imaging Biomarkers Alliance. UPICT Oncology FDG-PET/CT Protocol, Online. Accessed: March 2024.

[92] Joaquin Quinonero-Candela and Carl Edward Rasmussen. A unifying view of sparse approximate gaussian process regression. *Journal of machine learning research*, 6(Dec):1939–1959, 2005.

[93] Maziar Raissi. Deep hidden physics models: Deep learning of nonlinear partial differential equations. *Journal of Machine Learning Research*, 19(25):1–24, 2018.

[94] Jim O Ramsay, Giles Hooker, David Campbell, and Jiguo Cao. Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69(5):741–796, 2007.

[95] Eugene M Renkin. Transport of potassium-42 from blood to tissue in isolated mammalian skeletal muscles. *American Journal of Physiology-Legacy Content*, 197(6):1205–1210, 1959.

[96] Victor Rielly, Kamel Lahouel, Ethan Lew, Nicholas Fisher, Vicky Haney, Michael Wells, and Bruno Jedynak. Mock: an algorithm for learning nonparametric differential equations via multivariate occupation kernel functions. *stat*, 1050:8, in press.

[97] Lorenzo Rosasco, Ernesto De Vito, Andrea Caponnetto, Michele Piana, and Alessandro Verri. Are loss functions all the same? *Neural computation*, 16(5):1063–1076, 2004.

[98] Dimitri Rothermel and Thomas Schuster. Solving an inverse heat convection problem with an implicit forward operator by using a projected quasi-newton method. *Inverse Problems*, 37(4):045014, 2021.

[99] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992.

[100] Samuel H Rudy, Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Data-driven discovery of partial differential equations. *Science advances*, 3(4):e1602614, 2017.

[101] Gianmario Sambuceti, Vanessa Cossu, Matteo Bauckneht, Silvia Morbelli, AnnaMaria Orengo, Sonia Carta, Silvia Ravera, Silvia Bruno, and Cecilia Marini. 18 f-fluoro-2-deoxy-d-glucose (fdg) uptake. what are we looking at?, 2021.

[102] Hayden Schaeffer. Learning partial differential equations via data discovery and sparse optimization. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2197):20160446, 2017.

[103] Philip Scheltens, Bart De Strooper, Miia Kivipelto, Henne Holstege, Gael Chételat, Charlotte E Teunissen, Jeffrey Cummings, and Wiesje M van der Flier. Alzheimer's disease. *The Lancet*, 397(10284):1577–1590, 2021.

[104] Otmar Scherzer. Convergence criteria of iterative methods based on landweber iteration for solving nonlinear problems. *Journal of Mathematical Analysis and Applications*, 194(3):911–933, 1995.

[105] Orazio Schillaci. Use of dual-point fluorodeoxyglucose imaging to enhance sensitivity and specificity. In *Seminars in nuclear medicine*, volume 42, pages 267–280. Elsevier, 2012.

[106] KC Schmidt and FE Turkheimer. Kinetic modeling in positron emission tomography. *The Quarterly Journal Of Nuclear Medicine And Molecular Imaging*, 46(1):70, 2002.

[107] Mara Scussolini, Matteo Bauckneht, Vanessa Cossu, Silvia Bruno, Anna Maria Orengo, Patrizia Piccioli, Selene Capitanio, Nikola Yosifov, Silvia Ravera, Silvia Morbelli, et al. G6pase location in the endoplasmic reticulum: Implications on compartmental analysis of fdg uptake in cancer cells. *Scientific reports*, 9(1):2794, 2019.

[108] Lalitha K Shankar, John M Hoffman, Steve Bacharach, Michael M Graham, Joel Karp, Adriaan A Lammertsma, Steven Larson, David A Mankoff, Barry A Siegel, Annick Van den Abbeele, et al. Consensus recommendations for the use of 18f-fdg pet as an indicator of therapeutic response in patients in national cancer institute trials. *Journal of Nuclear Medicine*, 47(6):1059–1066, 2006.

[109] Lawrence A Shepp and Yehuda Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE transactions on medical imaging*, 1(2):113–122, 2007.

[110] Jeng-Jer Shieh, Chi-Jiunn Pan, Brian C Mansfield, and Janice Yang Chou. A glucose-6-phosphate hydrolase, widely expressed outside the liver, can explain age-dependent resolution of hypoglycemia in glycogen storage disease type ia. *Journal of Biological Chemistry*, 278(47):47098–47103, 2003.

[111] Susana Evaristo de Oliveira Branco Silva. *Small animal PET imaging using GATE Monte Carlo simulations: implementation of physiological and metabolic information.* PhD thesis, Universidade de Lisboa (Portugal), 2010.

[112] Sang Wan Sim, Tae Sub Park, Sung-Jo Kim, Byung-Chul Park, David A Weinstein, Young Mok Lee, and Hyun Sik Jun. Aberrant proliferation and differentiation of

glycogen storage disease type ib mesenchymal stem cells. *FEBS letters*, 592(2):162–171, 2018.

[113] Sang Wan Sim, David A Weinstein, Young Mok Lee, and Hyun Sik Jun. Glycogen storage disease type ib: role of glucose-6-phosphate transporter in cell metabolism and function. *FEBS letters*, 594(1):3–18, 2020.

[114] Nadja Smailagic, Marco Vacante, Chris Hyde, Steven Martin, Obioha Ukoumunne, and Christos Sachpekidis. 18f-fdg pet for the early diagnosis of alzheimer's disease dementia and other dementias in people with mild cognitive impairment (mci). *The Cochrane database of systematic reviews*, 2015(1):CD010632, 2015.

[115] Louis Sokoloff. Modeling metabolic processes in the brain in vivo. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 15(S1):1–11, 1984.

[116] Sara Sommariva, Giacomo Caviglia, Gianmario Sambuceti, and Michele Piana. Mathematical models for fdg kinetics in cancer: a review. *Metabolites*, 11(8):519, 2021.

[117] Sara Sommariva, Mara Scussolini, Vanessa Cossu, Cecilia Marini, Gianmario Sambuceti, Giacomo Caviglia, and Michele Piana. The role of endoplasmic reticulum in in vivo cancer fdg kinetics. *PLoS One*, 16(6):e0252422, 2021.

[118] Agnes J Staal-van den Brekel, EFM Wouters, FBJM Thunnissen, and WA Buurman. Expression of e-selectin, intercellular adhesion molecule (icam)-1 and vascular cell adhesion molecule (vcam)-1 in non-small-cell lung carcinoma. *Virchows Archiv*, 428:21–27, 1996.

[119] Yoshifumi Sugawara, Kenneth R Zasadny, Alex W Neuhoff, and Richard L Wahl. Reevaluation of the standardized uptake value for fdg: variations with body weight and methods for correction. *Radiology*, 213(2):521–525, 1999.

[120] Yishan Sun, Liming Xiao, Yanmei Wang, Changping Liu, Li Cao, Wei Zhai, Bo Wang, Shupeng Yu, and Jun Xin. Diagnostic value of dynamic 18f-fdg pet/ct imaging in non-small cell lung cancer and fdg hypermetabolic lymph nodes. *Quantitative Imaging in Medicine and Surgery*, 13(4):2556, 2023.

[121] Albert Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, 2005.

[122] Joseph A Thie. Understanding the standardized uptake value, its methods, and implications for usage. *Journal of Nuclear Medicine*, 45(9):1431–1434, 2004.

[123] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288, 1996.

[124] Andreĭ Nikolaevich Tikhonov and VIAK Arsenin. Solutions of ill-posed problems. *(No Title)*, 1977.

[125] James M Varah. A spline least squares method for numerical parameter estimation in differential equations. *SIAM Journal on Scientific and Statistical Computing*, 3(1):28–46, 1982.

[126] Curtis R Vogel. *Computational methods for inverse problems*. SIAM, 2002.

[127] Duccio Volterrani, P Erba, Giuliano Mariani, et al. *Fondamenti di medicina nucleare Tecniche e applicazioni*. Springer-Verlag Italia, 2010.

[128] Dennis Vriens, Jonathan A Disselhorst, Wim JG Oyen, Lioe-Fee de Geus-Oei, and Eric P Visser. Quantitative assessment of heterogeneity in tumor metabolism using fdg-pet. *International Journal of Radiation Oncology* Biology* Physics*, 82(5):e725–e731, 2012.

[129] Dennis Vriens, Eric P Visser, Lioe-Fee de Geus-Oei, and Wim JG Oyen. Methodological considerations in quantification of oncological fdg pet studies. *European journal of nuclear medicine and molecular imaging*, 37:1408–1425, 2010.

[130] DaQuan Wang, Bo Qiu, QianWen Liu, LiangPing Xia, SongRan Liu, ChaoJie Zheng, Hui Liu, YiWen Mo, Xu Zhang, YingYing Hu, et al. Patlak-ki derived from ultra-high sensitivity dynamic total body [18f] fdg pet/ct correlates with the response to induction immuno-chemotherapy in locally advanced non-small cell lung cancer patients. *European Journal of Nuclear Medicine and Molecular Imaging*, 50(11):3400–3413, 2023.

[131] Guobao Wang, Arman Rahmim, and Roger N Gunn. Pet parametric imaging: past, present, and future. *IEEE transactions on radiation and plasma medical sciences*, 4(6):663–675, 2020.

[132] Wenli Wang, Jens-Christoph Georgi, Sadek A Nehmeh, Manoj Narayanan, Timo Paulus, Matthieu Bal, Joseph O'Donoghue, Pat B Zanzonico, C Ross Schmidtlein,

Nancy Y Lee, et al. Evaluation of a compartmental model for estimating tumor hypoxia via fmiso dynamic pet imaging. *Physics in Medicine & Biology*, 54(10):3083, 2009.

[133] Layne T Watson and Raphael T Haftka. Modern homotopy methods in optimization. *Computer Methods in Applied Mechanics and Engineering*, 74(3):289–305, 1989.

[134] Wolfgang A Weber. Assessing tumor response to therapy. *Journal of nuclear medicine*, 50(Suppl 1):1S–10S, 2009.

[135] Miles N Wernick and John N Aarsvold. *Emission tomography: the fundamentals of PET and SPECT*. Elsevier, 2004.

[136] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

[137] Yuan Yao, Lorenzo Rosasco, and Andrea Caponnetto. On early stopping in gradient descent learning. *Constructive Approximation*, 26(2):289–315, 2007.

[138] James WT Yates. Structural identifiability of physiologically based pharmacokinetic models. *Journal of pharmacokinetics and pharmacodynamics*, 33:421–439, 2006.

[139] A Zhu, D Lee, and H Shim. Metabolic pet imaging in cancer detection and therapy reponse. *NIH*, 38(1):1–23, 2011.

[140] Yang Zuo, Jinyi Qi, and Guobao Wang. Relative patlak plot for dynamic pet parametric imaging without the need for early-time input function. *Physics in Medicine & Biology*, 63(16):165004, 2018.