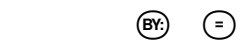
Estimation of Kinetic Parameters From List-Mode Data Using an Indirect Approach

ESTIMATION OF KINETIC PARAMETERS FROM LIST-MODE DATA USING AN INDIRECT APPROACH

by

Joseph Christian Ortiz



A Dissertation Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA
2016

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Joseph Christian Ortiz, entitled Estimation of Kinetic Parameters from List-Mode Data Using an Indirect Approach, and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: 08/24/2016

Eric Clarkson

Date: 08/24/2016

Matthew Kupinski

Date: 08/24/2016

Lars Furenlid

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.
I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: 08/24/2016

Dissertation Director: Eric Clarkson

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

SIGNED: Joseph Christian Ortiz

ACKNOWLEDGEMENTS

I would like to thank God for giving me wisdom and strength throughout my schooling.

I would like to thank my advisor, Dr. Eric Clarkson for his patience, support, and guidance throughout my years at CGRI, and never giving up on me. I would also like to thank Dr. Matthew Kupinski for sharing his simulations expertise in the clutch. Finally I would like to thank Dr. Lars Furenlid, an individual who was always looking out for me, and for always being approachable and willing to shoot the breeze about just about anything.

I would like to thank Dr. Art Gmitro for providing my funding in my time of need, as well as Dr. Harrison Barrett for the beginning of my grad student career.

Thank you to Dr. Luca Caucci for the frequent discussions and assistance. I would especially like to thank Dr. Helen Fan, Dr. Vaibhav Bora, and Dr. Cecile Chaix for being wonderful sources of information, and for providing support with various aspects of my project, and for proofreading my dissertation extensively.

A special thank you to Chris Summitt, who was been there since day one of this remarkable journey, he is truly my brother from another mother, and none of this would have been possible without him. Also, a big thank you to Soha Namnabat for his patience in explaining EM problems and always having words of encouragement

Thank you to Merry Warner, Liz Hague, Christy Barber and Lori Zimmerman for keeping me honest, and never sugar coating things, and thank you to my lab mates: Ling, Xin, Neil, Joy, Stephen, Nasrin, Tomoe, and big thank you to all those previous students who worked tirelessly on FastSPECT II.

Thank you to my parents, Jose and Genevieve, and my siblings Cyprian and Crystal for believing in me first. Finally, and most importantly, thank you to my wife Marissa and my children Ezekiel, Ixchella and Mr. Quetzal for all the love, support, and encouragement. You gave me the motivation I needed to finish.

DEDICATION

To my lovely wife Marissa and my children Ixchella, Ezekiel, and Mr. Quetzal.

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

TABLE OF CONTENTS – *Continued*

LIST OF FIGURES

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

ABSTRACT

This dissertation explores the possibility of using an imaging approach to model classical pharmacokinetic (PK) problems. The kinetic parameters which describe the uptake rates of a drug within a biological system, are parameters of interest. Knowledge of the drug uptake in a system is useful in expediting the drug development process, as well as providing a dosage regimen for patients. Traditionally, the uptake rate of a drug in a system is obtained via sampling the concentration of the drug in a central compartment, usually the blood, and fitting the data to a curve. In a system consisting of multiple compartments, the number of kinetic parameters is proportional to the number of compartments, and in classical PK experiments, the number of identifiable parameters is less than the total number of parameters. Using an imaging approach to model classical PK problems, the support region of each compartment within the system will be exactly known, and all the kinetic parameters are uniquely identifiable. To solve for the kinetic parameters, an indirect approach, which is a two part process, was used. First the compartmental activity was obtained from data, and next the kinetic parameters were estimated. The novel aspect of the research is using listmode data to obtain the activity curves from a system as opposed to a traditional binned approach. Using techniques from information theoretic learning, particularly kernel density estimation, a non-parametric probability density function for the voltage outputs on each photo-multiplier tube, for each event, was generated on the fly, which was used in a least squares optimization routine to estimate the compartmental activity. The estimability of the activity curves for varying noise levels as well as time sample densities were explored. Once

an estimate for the activity was obtained, the kinetic parameters were obtained using multiple cost functions, and the compared to each other using the mean squared error as the figure of merit.

CHAPTER 1

Basic Pharmacology

1.1 Overview

The field of pharmacology is the study of the interaction of a drug with a system, and is typically divided into two sections: pharmacodynamics (PD) and pharmacokinetics (PK). A capable pharmacologist has an understanding of the drugs, and the precise amount needed in order for the drug to beneficial. In the pharmaceutical industry, knowledge of the PD/PK processes is imperative to speeding up the drug discovery and development phase. Proper modeling allows for testing the therapeutic ranges of a particular drug, and the drugs with the highest probability of success are where resources are focused. It is a cost effective method used by large pharmaceutical companies to maximize returns on investments, given that the cost and time necessary to bring new drugs to market has increased [101], [69] [94]. Furthermore, integration of data from PD/PK models allows for drug development strategy to be continuously updated, increasing the efficiency at the pre-clinical development stages.

In this chapter we will discuss some basic pharmacodynamics, with the bulk of the chapter devoted to pharmacokinetics. We will introduce concepts and math using a single-compartment model, and generalize to a multi-compartment model. Furthermore, we will investigate the use of imaging techniques and the potential applications to classical pharmacokinetic modeling. Much of the information from this chapter was derived from various online references as well as short courses and

literature [77] [13] [3] [82].

1.2 Pharmacodynamics

Pharmacodynamics is primarily concerned with the effect that a drug has on a system, with the effects directly related to the concentration of the drug at a specific site. A drug administered to a system will cause a response, be it a therapeutic or an unwanted one. The specific effect a drug has on a particular organ is unknown, however knowledge of the location of the administered drug and the effect on the surrounding organs is sufficient to model the response of the system to a drug. Meticulous monitoring of the dose-response curve is necessary to determine when the maximal efficacy will be achieved and what is the corresponding dose. System variables such as age, disease, and gender for example, will complicate the effects a drug has, causing observable shifts of the dose-response curves.

1.2.1 Drug/Receptor Binding

A drug interacts with a system by targeting and attaching to cell receptors, which are proteins located on the surface of the cell that trigger specific responses within the cell. To maximize the efficiency of the drug, ideally the density of receptors at the site of interaction is large. The drug-receptor binding process is fully described by the law of mass action given by the following expression, where k_1 and k_2 represent the rate at which the binding occurs and dissociates, D represents the drug or ligand, and R represents the receptor.

$$D + R \quad {}_{k_1} \leftrightarrow^{k_2} \quad DR \tag{1.1}$$

For the drug-receptor interaction, at equilibrium the rate at which the drug binds

with the receptors and the rate at which it dissociates is equal. Hence, $k_2[DR] = k_1[D+R]$, and the equilibrium dissociation and equilibrium association constants are $k_d = k_2/k_1 = [D+R]/[DR]$ and $k_a = 1/k_d$. If the total concentration of receptors is given by $R_{total} = R + DR$ (bound and unbound), the percentage of receptors that bind with the drug, can be represented via a scaled form of the Michaelis-Menten shown in equation 1.2 equation [97] [24].

$$f_{DR} = \frac{DR}{R + DR} = \frac{D}{k_d + D} \tag{1.2}$$

Plotting scaled Michaelis-Menten Eq. 1.2 shows that as the concentration of the drug increases, there becomes a point at which the effectiveness of the drug saturates.

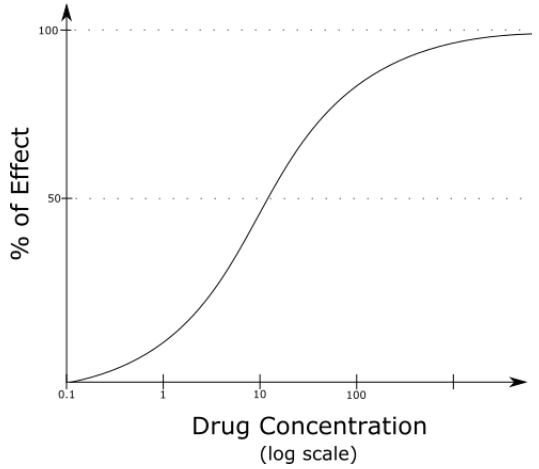


Figure 1.1: Michaelis-Menten plot showing the effectiveness of a drug (binding to receptors) as the concentration is increased.

There is a small region called the therapeutic region in which the drug is actually useful. below the threshold the drug has no effect and above the threshold the drug becomes toxic [95] [30] [33].

1.2.2 Agonists and Antagonists

A high concentration of receptors is not enough to ensure an administered drug will produce a desired response within a system. The drug itself is either categorized as being an agonist or an antagonist. An agonist drug is designed to bind with specific receptors in a cell and produce a positive response, while an antagonist drug binds to specific receptors without producing any effect. It's main purpose is to block targeted receptors.

When dealing with agonists, there are two main characteristics, the affinity and the efficacy, which deal with the probability the drug binds to the receptor, and the ability for a drug to cause a response, respectively. Furthermore, when an agonist bonds with receptors, if it produces a maximum response it is characterized as a full agonist, as opposed to a partial agonist, which even when all receptors are occupied, does not produce a full response. Most drugs have a high affinity, and typically as seen in Fig. 1.1 for a specific drug, as the concentration is increased, so will the efficacy. The rate of efficacy between different agonists will occur at different concentrations due to the affinity of each of the agonists. A drug that reaches max efficacy at a lower drug concentration is considered to have a higher affinity, and is thus more potent.

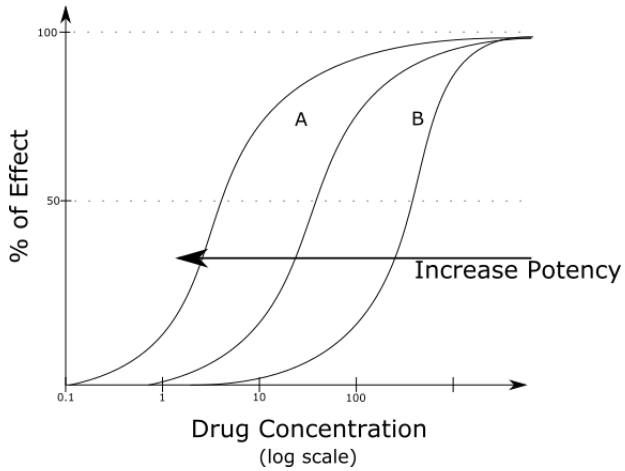


Figure 1.2: Diagram illustrating the drug dose and response. the two curves denoted by (A) and (B), are two drugs administered to a system. (A) and (B) are two drugs with different doses, with the same efficacy. Drug (A) a has a higher affinity, hence it is the more potent of the two.

Antagonists also have a high affinity and tend to readily bond to receptors. However, they do not produce an effect while bonded to the receptor. They do however compete with the agonist for the same receptor site, known as a competitive antagonist, and will block the agonist from fulfilling the therapeutic requirements. A competitive antagonist is a reversible process, and increasing the concentration of agonists will cause the antagonist-receptor bonds to be broken. A non-competitive antagonist, on the other hand, is an irreversible process. Non-competitive antagonists occupy the receptor sites, reducing the density of free receptors in a system, reducing the efficacy of the system.

1.3 Pharmacokinetics

Pharmacokinetics is primarily concerned with drug uptake in a system, which is a collective term that includes the individual rates of **A**bsorption, **D**istribution, **M**etabolization, and **E**xcretion (**ADME**). What is observed in a PK experiment is the concentration of a particular drug in a system as a function of time. By combining the observed concentrations with a mathematical model, the ADME rates can obtained, and more importantly, quantified. The most common type of model is the compartmental model, which partitions the system into several sections. Furthermore, compartmental models are generally considered to be deterministic in nature due to the fact the models are chosen such that they have the best agreement with the observed data, and the value of the model is determined by how well the uptake parameters are able to be extracted.

1.3.1 Compartmental Modeling in Pharmacokinetics

In order to gain some insight regarding the basic function of the the compartmental model, a single compartmental model will first be analyzed. Consider a drug being administered to a human subject. For simplicity, the organs, bones, vascular system, blood, etc can be all grouped into one single compartment. The method in which the drug is administered is not important for the moment, but will become significant as the complexity of the model increases. However, initially the assumption that will be made is that immediately after a drug is administered the system, the drug is instantaneously distributed to all the individual parts of the body, and an equilibrium state is reached. An example of a measured activity as a function of time can be seen in Fig. 1.3.

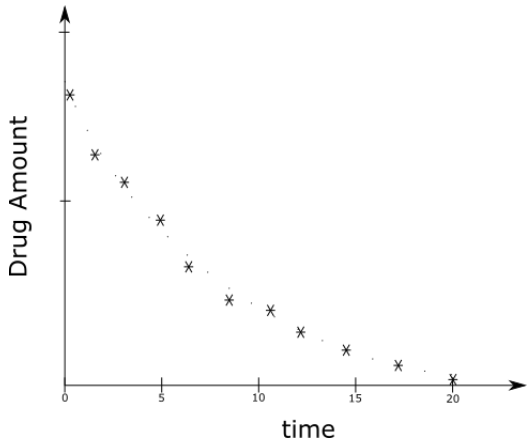


Figure 1.3: Diagram illustrating an example of the experimental data recorded for the activity of a drug in time within a system.

Typically in PK experiments, the data is recorded first, then a model is selected, and finally using a mathematical toolbox, the model is fitted to the data.

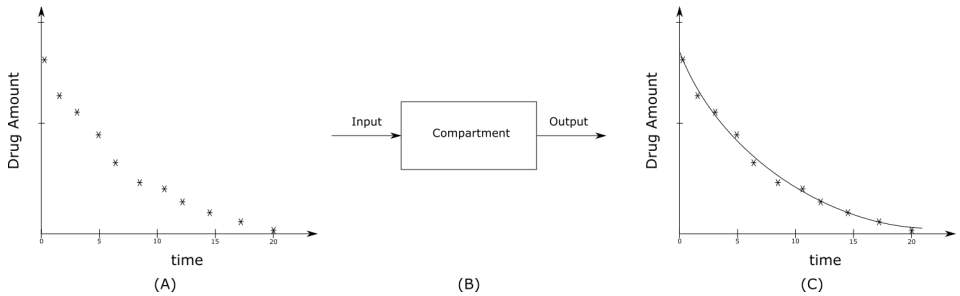


Figure 1.4: Diagram illustrating the steps involved in fitting the measured activity data from a pharmacokinetic experiment [1].

For the activity curve in Fig. 1.3, the best choice of a model would be one that has some sort of exponential decay.

$$\frac{da(t)}{dt} = -ka(t) \tag{1.3}$$

The expression from Eq. 1.3 describes amount of a drug in the compartment, via a linear, first-order elimination process, which is commonly used in PK experiments. The variable $a(t)$ represents to activity or concentration of the drug in the system, and k is the first order elimination constant. As time increases, the drug will eventually be eliminated from the system, at a rate k, that is proportional to the amount that remains in the system. A useful metric in PK experiments, that is also used in pharmacodynamics, is the concentration of the drug within a system, which given the volume of distribution and the amount of the drug administered as the input, can be expressed as

$$C = \frac{\text{amount of drug in system}}{\text{volume of system}} = \frac{a}{V} \tag{1.4}$$

and often, the concentration and activity will be interchanged without a loss of generality since they only differ by a scaling factor. In an ideal situation, the drug concentration within the compartment will be sampled at a constant rate. However in practice, the sampling is very coarse, and does not occur in regular time intervals. Often in clinical applications a sparse number of measurements after the drug has been administered to a patient will be taken. Hence the concentration of the plasma must be predicted, which is why a correct model that properly describes the system dynamics is important. From the dynamic model given by Eq. 1.3, an expression for the time course of the concentration can easily be solved for,

$$C(t) = C_0 \exp(-kt) \tag{1.5}$$

and by taking the natural log of the expression from Eq. 1.5, a linear relationship

can be obtained which allows for the concentration in time to be easily predicted.

$$\ln C(t) = C_0 + -kt \tag{1.6}$$

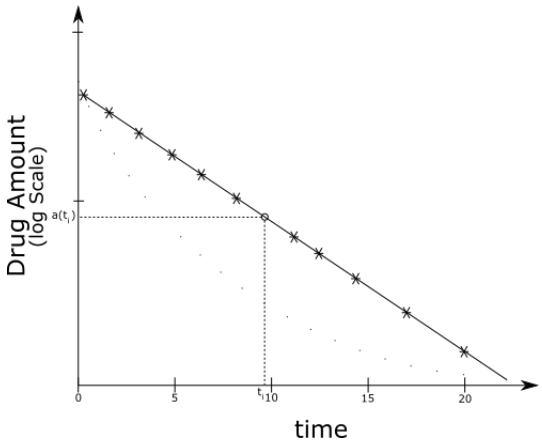


Figure 1.5: Diagram illustrating the logarithm of the model predicted by the data from Fig. 1.3.

1.4 Imaging Approach to PK Modeling

Using an imaging system such as FastSPECT II, the spatio-temporal activity distribution of a radiotracer in an object is measured. The activity distribution $f(\mathbf{r}, t) = \sum_{l=1}^{L} a_l(\mathbf{r}, t)$, is the sum of the individual compartments, or organs, that make up the object. Knowledge of the support region, S_l, of each individual compartment can be obtained from a CT scan a priori. This knowledge is used to integrate over the spatial component of each compartmental activity as shown in Eq. 1.7,

$$a_l(t) = \int_{S_l} a_l(\mathbf{r}, t) d\mathbf{r} \tag{1.7}$$

which happens to be similar in form to the temporal drug concentration observed in classical pharmacokinetics. In an imaging system, the measured $f(\mathbf{r}, t)$ is proportional to the radioactivity of the object. The number of photons collected from the object is proportional to the amount or concentration of the tracer within the system. If we assume there exists a relationship between the radioactivity of the object measured using imaging, and the concentration of the drug by sampling fluids, an imaging approach can be used in modeling classical PK experiments. An imaging approach is desirable because it allows access to previously inaccessible compartments within the system, which allows for a more accurate representation of the tracer kinetics.

1.4.1 Radiotracer

The tracer used in FastSPECT II for calibration as well as the one commonly used in imaging experiments is a technetium (^{99m}Tc) labeled radio-pharmaceutical. It is used because of the relatively long half life (6 hours), and the relatively low patient radiation exposure risk. Furthermore it is cheap, and easily attaches to ligands.

1.4.2 MOBY Object Phantom

The 4-D MOBY mouse model, originally developed by Paul Segars, is a program used to generate whole body digital mouse phantoms for use in small animal imaging applications. MOBY utilizes non-uniform rational b-splines or NURBS for short, to design and model the complex surfaces to create realistic phantoms, orders of magnitude better than voxelized, or pure mathematical phantoms [73]. A unique feature of the MOBY phantom is that the cardiac and respiratory system can also be programed in the phantom causing motion as well as anatomical variations resembling

an actual animal. Furthermore, it is a valuable tool which allows for the evaluation of the capabilities of an imaging system.

Using the MOBY phantom allows for the support region of each organ within the object to be known exactly. The object can be represented by a linear combination of the basis elements, and the voxels are partitioned into compartments. In the context of this research, the assumption made is that the activity is constant throughout the compartment, and the specific amount of activity contained in each compartment is proportional to the concentration of the tracer, $C_l = a_l/S_l$. If the l^{th} compartment within the object contains N_l voxels, the activity in each voxel, $\alpha_{n\in S_l}$, within compartment l is also be constant. The object $\mathbf{f}$ is a vector that contains N_{basis} elements, where $f_n = \alpha_n$. Hence, the activity in each compartment is represented as

$$a_l(t) = \sum_{n\in S_1} f_n(t) \tag{1.8}$$

where the activity of each individual voxel of the object is given by $f_n = a_l/N_l$ for $n \in S_l$. To simplify the notation, an $L \times N$ matrix $\mathbf{D}$ is created, which is a binary matrix that specifies the voxels within each individual compartment. Another assumption made is that the object being imaged is stationary, and the compartment boundaries do not change in time. The only parameter that changes is the activity distribution in the object. The sole purpose of the matrix $\mathbf{D}$ is to sum up the voxels within the individual compartments of the object, which yields the activity in each of the compartments. This allows Eq. 1.8 to be written in a more compact form.

$$\mathbf{a}(t) = \mathbf{D}\mathbf{f}(t) \tag{1.9}$$

Given the activity is uniformly distributed in each compartment, the object can also

be written in terms of the activity

$$\mathbf{f} = \tilde{\mathbf{D}}\mathbf{a} \tag{1.10}$$

with

$$\tilde{\mathbf{D}} = \mathbf{D}^T \mathrm{diag}(1/N_1, 1/N_2, \cdots, 1/N_L) \tag{1.11}$$

This relationship will be used in a later section.

1.4.3 Multi-Compartmental Pharmacokinetics

In Sec. 1.3.1, the individual compartments of the system were all grouped into a single compartment for simplicity of modeling. In a multi-compartment analysis, elements of the system that share similar characteristics are also grouped into compartments, which allows for a reduction in the number of parameters that need to be estimated, while still being able to obtain information regarding the uptake of the tracer. For example, the highly perfused organs of the body such as the heart, liver, kidney and spleen have very similar uptake characteristics, and are typically grouped together into a single compartment referred to as the central compartment. On the other hand, the fat tissues, as well as the muscle tissues, and tumors, are grouped into another compartment referred to as the peripheral compartment. A drug will enter the central compartment, be circulated through the central and peripheral compartments, and leave the central compartment at which it will be permanently removed from the system.

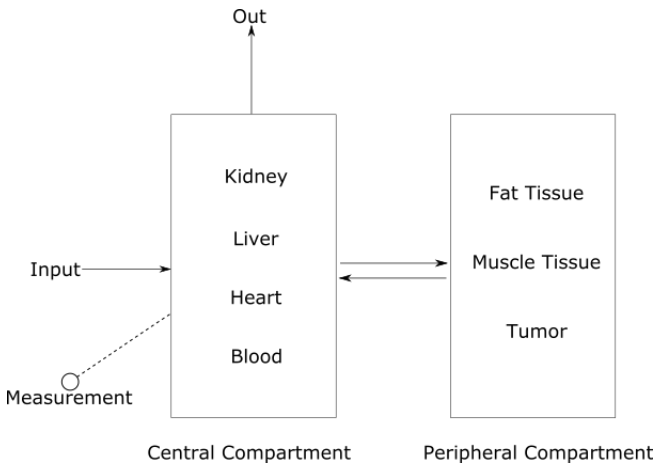


Figure 1.6: Illustration showing typical organs that compose the central and peripheral compartments. This would be an example of a system modeled as a two-compartment model.

Within a compartment, the assumption is that when the drug enters the compartment is kinetically homogeneous, and the rate of flow in and out of the compartment is the same for all of the elements within the group. In PK experiments, typically access to only one compartment is possible, such as the blood, to which the concentration of the drug is measured. The advantages of using an imaging approach to solve classical PK problems is that specific compartments within the system that normally would be inaccessible can be studied, and the kinetics of the tracer can be obtained. The basic pharmacokinetics equations introduced in Sec. 1.3.1 are still applicable, and can be generalized to multiple compartments easily. Instead of a single drug activity denoted by $a(t)$, what will be used is a spatio-temporal vector of activities $\mathbf{a}(\mathbf{r}, t)$, where each of the elements of the vector indicate the activity

within each individual compartment.

$$\mathbf{a}(\mathbf{r},t) = \begin{pmatrix} a_1(\mathbf{r},t) \\ a_2(\mathbf{r},t) \\ \vdots \\ a_L(\mathbf{r},t) \end{pmatrix} \tag{1.12}$$

With the addition of more compartments, the spatial component in the activity is also included since the compartment occupies a three-dimensional space, within a system. However, since the support region S_l, of each of the individual compartments is known, the compartments are non-overlapping, and the activity remains constant throughout the compartment, the spatial component of the activity can be integrated out, leaving the vector of compartmental activities as a function of time only.

$$a_l(t) = \int_{S_l} a_l(\mathbf{r},t)\, d\mathbf{r} \tag{1.13}$$

To model the temporal tracer evolution in a system consisting of multiple compartments, a vectorized form of the linear kinetic differential equation from Eq. 1.3 is used.

$$\frac{d\mathbf{a}(t)}{dt} = \mathbf{K}\mathbf{a}(t) + \mathbf{I}(t) \tag{1.14}$$

The key components of Eq. 1.14 are $\mathbf{K}$, and $\mathbf{I}(t)$, the kinetic matrix and the delivery method of the drug into the system. Typically $\mathbf{I}(t)$ is a constant, bolus input, that is administered to a single compartment.

Kinetic Matrix

The $L \times L$ kinetic matrix, $\mathbf{K} = \{k_{i,j} : i, j = 1, \cdots L\}$, governs the uptake of the drug within the system, and each of the individual elements $k_{i,j}$ represent the rate of partial flow of the drug between, or out of each compartment. Depending on the chosen system model, there will be some elements of $\mathbf{K}$ that will be equal to zero. The importance of the model will be discussed in a later section, but in order for the model to be physically realizable, there are some standard assumptions that $\mathbf{K}$ must adhere to.

1. $k_{i,j} \geq 0 \quad \text{for} \quad i \neq j$: The off diagonal elements represent the rate of partial flow of the drug between compartments, and are non-negative.

2. $k_{i,i} \leq 0$: The diagonal elements represent the rate of partial flow out of the individual compartments, and are non-positive.

3. $|k_{ii,}| \geq \sum_{i \neq j} k_{i,j}$: In order to ensure that activity within the system is conserved, the sum of the elements of the column not including the diagonal element, must be at most equal to the magnitude of the diagonal element.

4. $\det(\mathbf{K}) \neq 0$: The kinetic matrix must be invertible.

The condition that $\mathbf{K}$ is invertible ensures the drug entering the system does not accumulate, and eventually leaves the after being administered, and on a microscopic level, no compartments within the system act as traps. Furthermore, $\mathbf{K}$ is diagonalizable which allows for it to be represented in terms of it's eigenvalues λ_l and eigenvectors $\mathbf{v}_l$. The familiar eigenvalue equation for $\mathbf{K}$ is written as follows, where

the assumption shall be made the eigenvalues of $\mathbf{K}$ are real and positive.

$$\mathbf{K}\mathbf{v}_l = -\lambda_l \mathbf{v}_l \tag{1.15}$$

Solving the kinetic equation from Eq. 1.14, yields the compartmental activity as a function of time.

$$\mathbf{a}(t) = \int_0^t e^{\mathbf{K}(t-s)}\mathbf{I}(s)ds. \tag{1.16}$$

Modeling the activity within the system requires an integration of an exponentiated matrix, which can be approximated using either a Taylor series expansion, or Laplace methods. The significance of rewriting the kinetic matrix in the eigenvalue basis allows for the simplification of Eq. 1.16. The kinetic matrix written in terms of the eigenvalues and eigenvectors is

$$\mathbf{K} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}. \tag{1.17}$$

The input function in terms of the eigenvector basis is written as $\mathbf{I}(t) = \mathbf{V}\boldsymbol{\alpha}(t)$, and Eq. 1.16 becomes

$$\mathbf{a}(t) = \mathbf{V}\int_0^t e^{-\mathbf{\Lambda}(t-s)}\boldsymbol{\alpha}(s)ds = \sum_l \mathbf{v}_l \int_0^t e^{-\lambda_l(t-s)}\alpha_l(s)ds \tag{1.18}$$

for a system consisting of L compartments with a constant bolus input, the activity in each compartment is a linear combination of L decaying exponential functions, after the bolus input.

CHAPTER 2

Gamma Ray Physics and Detectors

Before elaborating on the how imaging can be used to analyze classical PK problems, the imaging systems and theory must be introduced. Chapter 2 is concerned with the physics behind a gamma ray as well as the various detectors used in gamma ray imaging applications. Chapter 3 is concerned with the theory behind the imaging process, and Chapter 4 introduces the actual imaging system that uses concepts discussed in Chapter 2 and Chapter 3.

2.1 Overview

Gamma-ray detectors have a variety of applications in areas ranging from clinical and high-energy physics to homeland security. Gamma-ray detectors can be broadly classified two main categories – spectrometers and imaging systems. At the Center for Gamma Ray Imaging (CGRI), gamma ray detector are primarily used in imaging systems. In a gamma-ray imaging system, the signal from a gamma-ray interaction is converted to an electrical signal that can be recorded and further processed using a variety of techniques to estimate properties the object being imaged.

The outline of this chapter is as follows, in order to gain some insight in how a gamma ray imaging system works, the physics behind gamma rays and their interaction with matter is discussed in Sec. 2.2. In Sec. 2.3, the different types of gamma-ray detector are discussed, with an emphasis on the scintillator detector.

Much of the mathematical derivations in this chapter have been derived and discussed in greater detail in [58], [6], [5], [61], [19], [21], [10], and [76].

2.2 Interaction of Gamma Rays with Matter

Gamma rays are high-energy photons which are emitted due to nuclear processes, with energy typically on the order of 100keV or more, and have wavelengths smaller than 10^{-12} meters. Due to the high-energy associated with gamma rays, they are classified, as ionizing radiation. A gamma-ray interaction in a material triggers a complicated cascade of events. The gamma-ray interaction can result in either the photon being absorbed completely (photoelectric interaction), or partially (Compton scattering), or for gamma-ray energies greater than 1022 keV the creation of a positron-electron pair. In the next few sections these topics will be touched on briefly.

2.2.1 Photoelectric Interaction

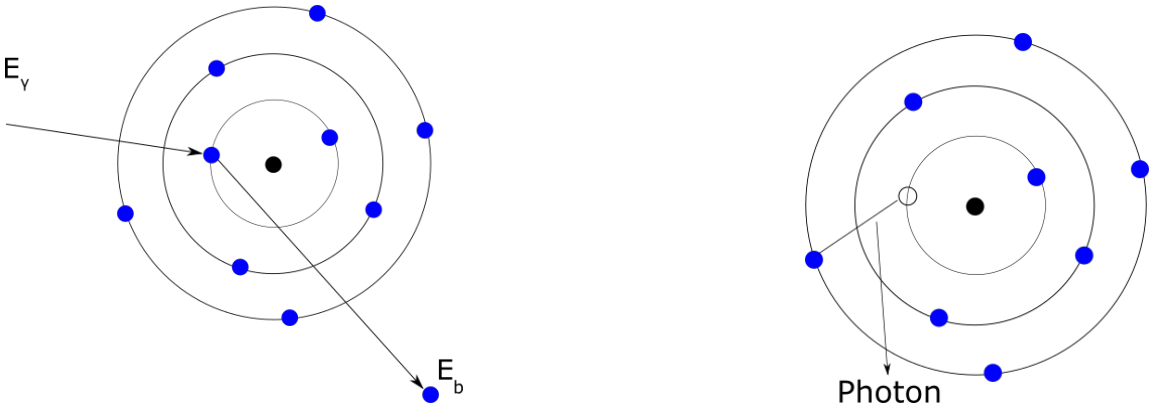
During a photoelectric interaction, all the energy of an incident gamma ray is transferred to an electron of an atom. Each atom has of a number of energy-levels or shells, that might be occupied by electrons. Electrons on different shells are bound to the nucleus with different binding energies (E_b). The gamma-rays interact more strongly with the inner shell electrons which have higher binding energies..

A gamma-ray photon with energy E_γ, cannot interact with electrons whose binding is more than E_γ. In a photoelectric gamma-ray interaction, the interacting electron is ejected from the atom with energy E_e, leaving a vacancy in it's place. The kinetic energy of the ejected electron, is equal to the difference between the energy

of the incident gamma ray photon and the binding energy of the electron.

$$E_e = E_\gamma - E_b \tag{2.1}$$

The vacancy caused by the ejected electron is filled by an electron from an outer shell, and in the process the binding energy is released in the form of X-rays (see Fig.2.1b or Auger electrons (see Fig.2.2). The Auger electron is ejected from the atom when an outer shell electron fills the inner shell vacancy, and energy of the inner shell vacancy is transferred to the Auger electron. The Auger electron with its energy higher than the ionization potential further ionizes atoms in the material. The emitted X-ray can either escape the material, or be reabsorbed to further ionize the material. The photoelectric effect is the most probable gamma-ray interaction at low gamma-ray energies up to a few hundred keVs.



(a) Collision with incident gamma ray.

(b) Generation of photon.

Figure 2.1: Example of a gamma ray colliding and transferring it's energy to an electron. The result is the electron is free from it's shell, leaving in place a vacancy equal to the original binding energy of the freed electron.

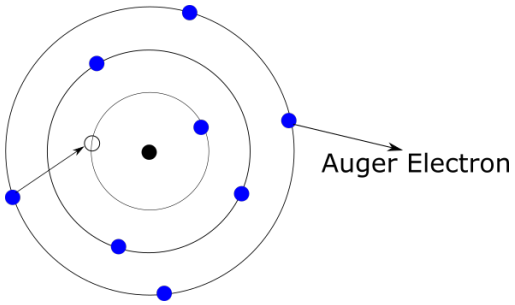


Figure 2.2: Energy released from an electron in an outer shell filling an inner shell ionizes the atom further, creating an Auger electron.

2.2.2 Compton Scattering

In a Compton scattering event, the gamma-ray photon partially transferring energy to an electron, and scatters. The relationship between the energy of the scattered gamma ray and the scattering angle is given by the equation below.

$$E_{\gamma}^{sc} = \frac{E_{\gamma}^{in}}{(1 + \frac{E_{\gamma}^{in}}{m_e c^2}(1 - cos(\theta_{sc})))} \tag{2.2}$$

The ejected electron has energy equal to the difference between the initial gamma-ray energy and the energy of the scattered gamma-ray. The differential cross section from which a photon is scattered by an electron can be described using the Klein-Nishina formula. The Compton effect dominates gamma-ray interactions in the mid-hundred keVs.

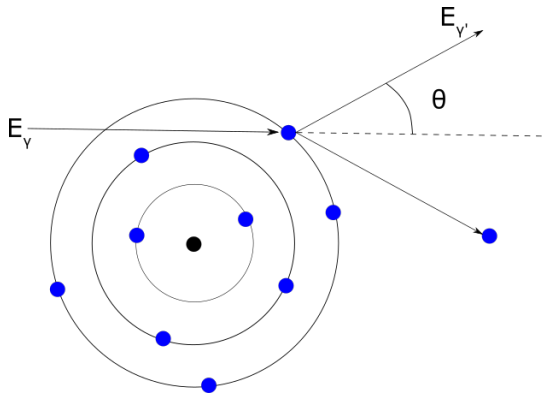


Figure 2.3: Compton scattering, incident partially transfers energy to an electron. The photon is scattered, with an energy less th/¿an the incident energy.

2.2.3 Positron-Electron Pair Production

When the incident gamma-ray energy is $\geq$ 1.02MeV, in the presence of a nucleaus, the gamma-ray energy can be converted to an electron-positron pair. This type of interaction is called pair production. The minimum energy required for pair production is 1.022MeV, which corresponds to twice the rest energy of an electron or positron (.511MeV).

$$E_{thresh} = 2m_0c^2 + KE(e^-) + KE(e^+) \tag{2.3}$$

When the gamma ray is exactly equivalent to twice the rest-mass energy of the electron, a positron-electron pair is created with zero kinetic energy. Any additional energy in the gamma-ray photon is transferred to the electron-positron pair's kinetic energy.

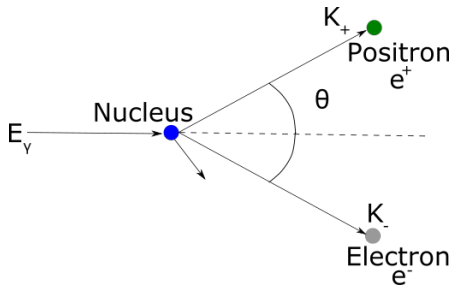


Figure 2.4: Positron and electron pair creation given the incident gamma ray energy is greater than 1.022MeV. Part of the kinetic energy will be transferred to the nucleus.

After either a Compton scattering event, or a photoelectric interaction, the resulting high energy electron propagates in the material producing more hole-electron pairs, and lattice vibrations which produce optical photons (light). Finally, after the energy of the photons has reduced dramatically, the electron hole pairs recombine via a non-radiative process and produce photons or heat.

2.2.4 Total Attenuation

When a wave enters a material, the attenuation of the wave as a function of the penetration inside the material is given by the Beer-Lambert law.

$$I(x) = I_0 \exp(-\mu x) \tag{2.4}$$

In the case of a homogeneous and isotropic material, the linear attenuation coefficient equals $\mu = \sigma n$, where σ is the cross-sectional area of scattering events of the incident photon, and n is the electron density in the material. For gamma-rays, the scattering coefficient from Eq. 2.4, is the sum of the photoelectric, Compton and the pair production cross-sections.

2.3 Gamma Ray Detectors

In a gamma-ray detector, gamma rays are not directly recorded, rather, the gamma-ray interaction produces a cascade of electrons and ionizes the material. This charge produced by a gamma-ray interaction is either measured directly in semiconductor or gas detectors, or measured indirectly via scintillation photons in scintillation detectors.

2.3.1 Gas Detector

In a gas detector, a thin wire that is positively biased, is positioned in the center of a gas filled chamber as shown in Fig. 2.5. The gas chamber walls acts as the cathode. When a photon is incident on the gas chamber, it ionizes the gas atoms, producing free electrons that move towards the center anode, and positively charged ions which move towards the cathode. This charge migration results in a current flow through the circuit, which is recorded.

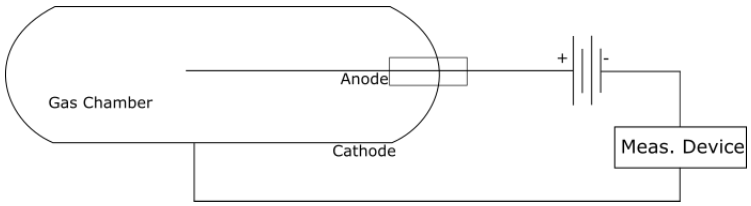


Figure 2.5: Gas detector. Gamma rays incident in the gas chamber cause ionization, and the electrons are collected by the anode located at the center.

Depending on the applied voltage, the detector exhibits different properties as shown in Fig. 2.6. At lower voltages, the detector becomes an ionization chamber – the ionized electrons are collected, and the output signal is proportional to the incident energy of the gamma ray. At higher voltages, the free electrons ionize other atoms

as they are accelerated towards the anode. Thus, the system has a gain, but the output current is still proportional to the energy of the incident radiation. Finally, at very high voltages, the ionized electrons are accelerated towards the anode at such a high velocity, that the secondary electrons produced further ionize other electrons. The point at which the ionization of atoms occurring within the chamber saturates, the output is independent of the initial photon energy. This is the basis behind the Geiger-Mueller Counter, which counts the number of particles entering the tube per second.

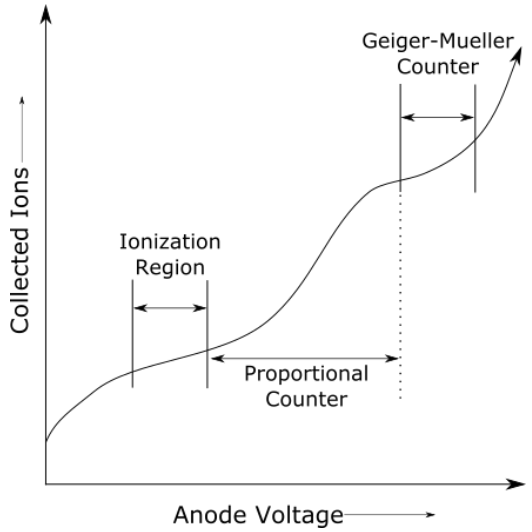


Figure 2.6: Trends in the output of the gas detector as the anode voltage is increased.

2.3.2 Semiconductor Detector

Semiconductor detectors have a PN diode structure which is operated in reverse bias. When a gamma-ray interaction occurs in depleted region of the PN junction, electron-hole pairs are produced. With the reverse bias applied to the PN junction, the electrons are attracted towards the N region, while the holes are attracted towards

the P region. This results in a current flow. This current flow is proportional to the energy of the incident gamma-ray energy, and is read out by electronics for further processing. A desirable characteristic of the semiconductor detector compared to a gas detector, is a higher energy resolution.

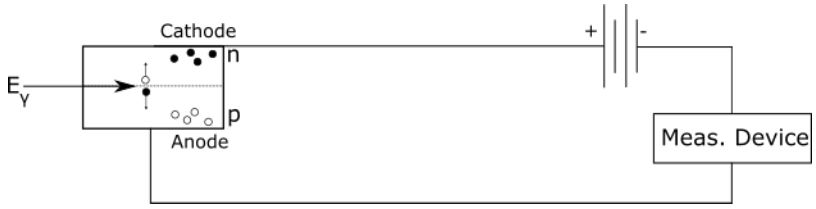


Figure 2.7: Schematic of semiconductor detector operating in reverse bias. An incident gamma ray produces a current proportional to the incident gamma ray energy.

2.3.3 Scintillation Detector

A scintillator is a material that emits visible photons when it interacts with gamma-ray photons. A scintillation gamma-ray detector consists of a scintillator, which emits visible photons, and a low-light photon detector like Photo Multiplier Tube (PMT), or photodiodes. Gamma-ray detectors for imaging application also have a light pipe between the scintillator and the photodetector to spread the scintillation photons to multiple photodetectors to improve the spatial resolution [39], [38].

The gamma ray scintillation detectors used extensively at CGRI are used as photon counting devices and can be modeled using Poisson statistics.

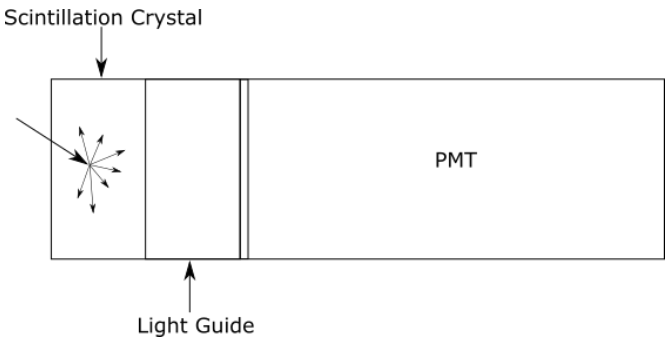


Figure 2.8: Scintillation spectrometer. A single PMT is attached to the scintillator crystal. The incident gamma ray releases a spray of visible photons.

Scintillator Material

A gamma-ray interaction in a material produces a large number of electron-hole pairs. These electron-hole pairs de-exite non-radiatively or radiatively. In most materials, the radiative decay process has lower probability, however, in scintillators a large fraction of the electron-hole pairs de-excite radiatively.

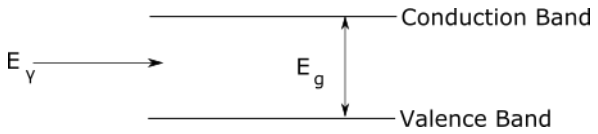


Figure 2.9: Scintillation detector. The photons released as the electrons decay to the ground state are small in number and lie outside the visible range.

A doped scintillator material such as sodium iodide, has a small amount of dopant such as thallium. These dopants, also called activators, add energy levels associated with the activators within the forbidden region of the base crystal. At the activator levels, there is a high probability of the exciton recombining radiatively. Furthermore, the activators are chosen such that the emitted photon lies within the band gap of the crystal. However, the scintillation process is an inefficient one, and only about 10%

of the energy deposited from the gamma-ray photon produces optical scintillation photons, and further only a fraction of the emitted scintillation photons are detected by the photodetector.

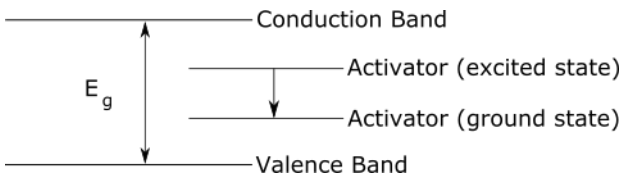


Figure 2.10: Scintillation detector with impurities added. The photons released in the decay from the activator excited state to the activator ground state are lower in energy than in the case of the pure crystal. More important, the resulting photons are within the visible region.

Photo-multiplier Tubes

Since the number of optical photons produced is small, the purpose of the PMT is two-fold: to efficiently detect incoming photons, and to provide very high gain to be able to use electronics to read the output. As shown in Fig. 2.11, the PMT consists of a photocathode, focusing electrodes, a series of dynodes, followed by an anode, all packed into a vacuum tube maintained externally at a high voltage. An optical photon from the scintillation material striking the face of the PMT will eject a photoelectron from the photocathode. The photoelectron is then focused onto a series of plates called dynodes, which are maintained at a positive voltage relative to the photocathode. Furthermore, each of the dynodes in the chain increases in voltage as the chain progresses, which aids in the production of secondary electrons. Eventually the photoelectrons will reach the anode at the end of the PMT, where they are converted to a voltage to be processed.

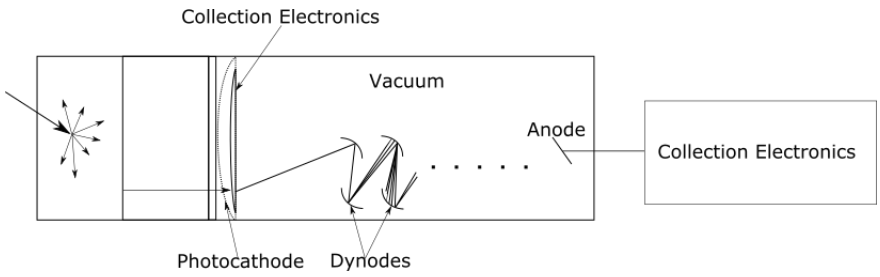


Figure 2.11: Photo-multiplier tube. The point of the PMT is two-fold: 1. amplify the incident photon, 2. convert the signal to a digital output.

Multi-PMT Detector

In order to obtain spatial information regarding the location at which the gamma-ray took place, an array of PMT's must be utilized. A widely used system that utilizes a multi-PMT configuration is the Anger camera, which consists of a scintillation crystal and a hexgonal-array of PMT's. The position of a scintillation event can be estimated using Anger logic. As shown in Fig. 2.12, the basic premise behind Anger logic uses the centroid position of the distribution of signal amplitudes from all the PMT's that are above threshold value to estimate the interaction position. The expression to find the interaction position $\hat{x}$, along an axis is given in Eq. 2.5.

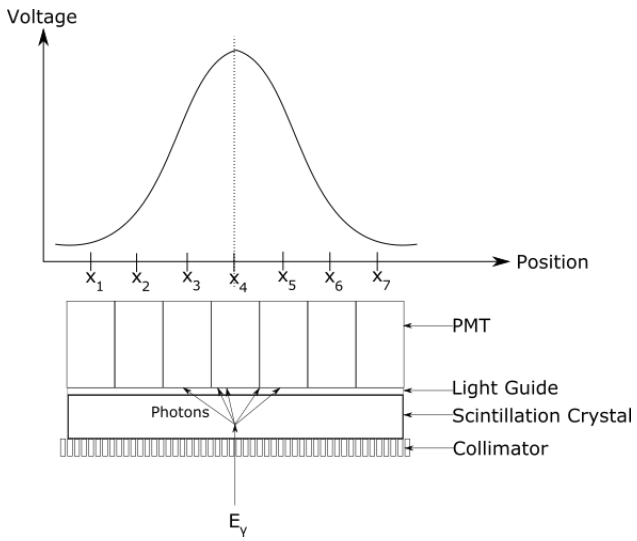


Figure 2.12: Basic theory of how an anger camera operates. An incident photon is converted to multiple optical photons via the scintillation crystal. The optical photons are then detected using the PMT's and a weighted distribution of the PMT outputs is obtained. The centroid of the weighted distribution is the position of interaction of the original gamma photon.

$$\hat{x} = \frac{\sum_i x_i w(x_i)}{\sum_i w(x_i)} \tag{2.5}$$

The cameras used at CGRI consist of a 3×3 array of PMT's as shown in Fig. 2.13, and a maximum likelihood (ML) method is used in order to estimate the interaction position and the energy of the incident gamma-ray. Details of the maximum-likelihood approach to position estimation will be discussed in Sec. 2.6.1.

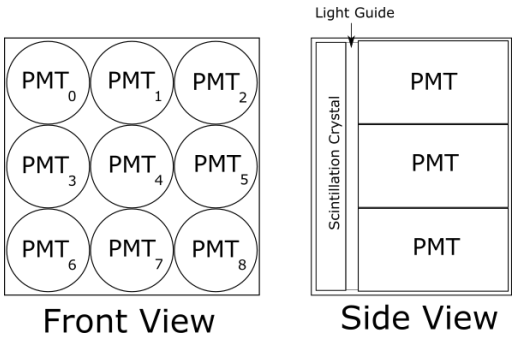


Figure 2.13: Detector used at CGRI consisting of a 3×3 array of PMT's.

2.4 Gamma-ray Imaging Systems

The gamma-ray detectors described in the previous section can be used to estimate the position of interaction and the energy of the gamma-ray. However, the information that we seek from a gamma-ray imaging system is the position from which the gamma-ray was emitted. To gather information about the origin of a gamma-ray, we must add constraints in our imaging system. Two common ways to constraint the gamma-rays that are impinging on a detector are using a parallel-hole collimator, and using a pinhole.

2.4.1 Parallel-Hole Collimator

The parallel-hole collimator is used in systems such as the Anger camera. The collimator only allows gamma-ray photons with a small range of angles to interact with the detector. The point of interaction on the gamma-ray detector and the angle information from the collimator give us enough information to estimate the line along which the gamma-ray was emitted. A parallel-hole collimator consisting of an array of extremely long tubes and small diameters would provide the best spatial

resolution, but the trade off in such a configuration would be a severe reduction in the sensitivity, which would ultimately limit the performance of the system.

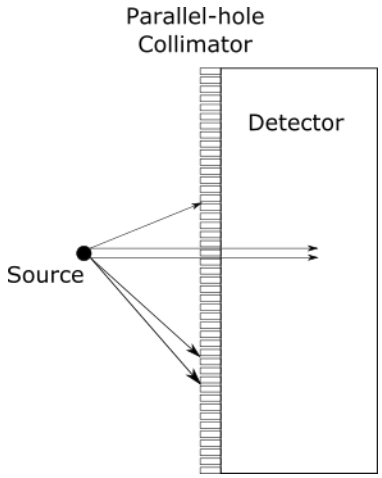


Figure 2.14: Schematic showing the effect of adding a parallel-hole collimator in front of the detector. The parallel-hole collimator limits the angles of the source that the detector collects. One advantage of using a parallel-hole collimator is the image on the detector is upright.

2.4.2 Pinholes

An alternative to the parallel-hole collimator is the pinhole, which projects an inverted version of the object onto the detector as shown in Fig. 2.15. The point of interaction on the gamma-ray detector and the pinhole position give us enough information estimate the line along which the gamma-ray was emitted. The spatial resolution is typically better with a pinhole because the object of interest can be magnified, as long as the distance from the object to pinhole is less than the pinhole to detector distance. Hence the pinholes are desirable in imaging contexts, because the only parameters are the diameter of the pinhole D_{ph}, the object to pinhole distance (z_1), and pinhole to detector distance (z_2). As the size of the pinhole decreases,

the spatial resolution increases at the expense of the sensitivity.

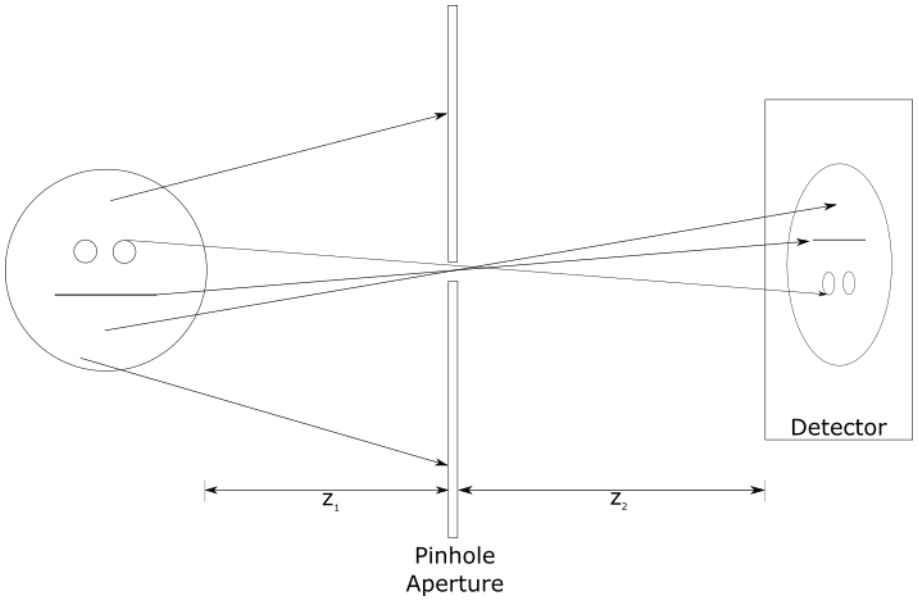


Figure 2.15: Basic imaging process of the pinhole aperture, which limits the amount of photons allowed to pass through to the detector. The variable z_1 and z_2 indicate the object to pinhole and pinhole to detector distances.

2.4.3 Multi-pinhole Systems

A system with multiple pinholes allows for an in increase in sensitivity while keeping the pinhole the same size. However with more pinholes also comes the problem of multiplexing, which is an overlap of the images of the object on the detector. The problem that arises due to multiplexing is there is an uncertainty in the pinhole that the photon passed through, hence a loss of object information. However if the system can be constructed such that the images do not overlap, the sensitivity will increase linearly with the number of pinholes.

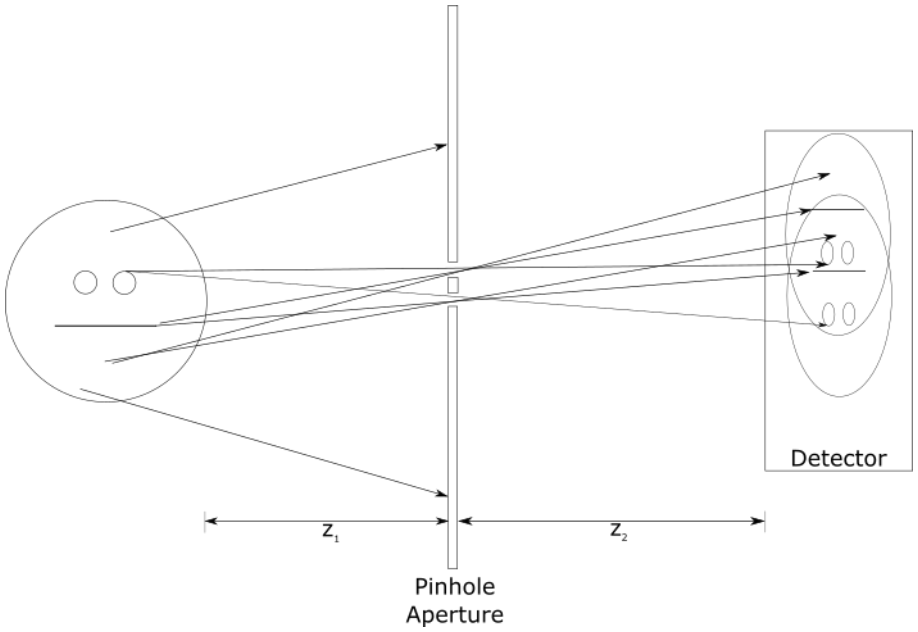


Figure 2.16: Multiplexing effect resulting in the use of multiple pinholes for imaging.

The intrinsic spatial resolution, which is the contribution to the spatial resolution from the detectors and electronics, is the ability of the detector to localize a scintillation event on the surface of the detector. While the imaging optics in front of the detector have a heavy bearing on the spatial resolution of the system, other factors such as the intensity of the gamma ray, or the thickness of the scintillation crystal are also important. Weak gamma rays with low energy cause statistical fluctuations in the output of the system due to a low number of optical photons being produced. Another key aspect that decreases the intrinsic resolution is the thickness of the scintillation crystal. With a thicker crystal, the spread of light produced by a scintillation event is larger, which results in a larger fluctuations in estimating the point of interaction of the gamma ray photon.

2.5 Statistics in the Scintillation Gamma-Ray Detector

In the previous sections, the physics of the scintillation gamma-ray detector, and gamma-ray imaging systems were introduced. In this section, the mathematics and statistics used to model gamma-ray imaging systems are discussed in some detail.

2.5.1 Emission of Optical Scintillation Photons

When a gamma-ray interacts in a scintillation material, it triggers a complicated cascade process which produces optical photons. We assume that all the optical photons are emitted isotropically from the point of gamma-ray interaction ($\mathbf{r} = (x, y, z)$). For a given gamma-ray interaction, if we assume that the optical photons from the scintillation process are emitted independently, then we can model the total number optical photons as a Poisson random variable N_{opt}.

$$P(N_{opt}|\mathcal{E}) = \frac{(\bar{N}_{opt}(\mathcal{E}))^{N_{opt}}}{N_{opt}!} \exp(-\bar{N}_{opt}(\mathcal{E})). \tag{2.6}$$

Here, $\mathcal{E}$ is the energy deposited in the gamma-ray interaction (in keV). The average number of optical photons produced increase with the energy deposited in the gamma-ray interaction. For small changes in the deposited gamma-ray energy, most material can be assumed to have a linear relationship between the number of optical photons produced and the deposited gamma-ray energy.

$$\bar{N}_{opt} = \eta_{sc}\mathcal{E} \tag{2.7}$$

Here, η_{sc} is a material dependent constant, with units of number of photons / keV.

2.5.2 Conversion of Scintillation Optical Photons to Photoelectrons in PMTs

The optical photons from the scintillation event are detected using PMTs. In general, a fraction of the emitted optical photons, $\alpha_j(\boldsymbol{r})$ will be incident on the photo-sensitive part of the j^{th} PMT. If we ignore all scattering and reflections, then $\alpha_j(\boldsymbol{r})$ is given by the solid angle subtended at the point of interaction $\mathbf{r} = (x, y, z)$ in the scintillator by the j^{th} PMT.

Additionally, in a PMT, a fraction η_j of the incident optical photons are converted into photoelectrons. Here, η_j is the quantum efficiency of the j^{th} PMT. A high value of η_j is very desirable in order to maximize the signal to noise ratio. One way to maximize η_j is to match the spectral response of the photocathode with the emission spectra of the scintillation material.

The probability of a photon incident on the j^{th} PMT, and the probability of a photoelectron being produced given an optical photon reaches the j^{th} PMT, are both binomial selection processes. According to the binomial selection theorem – the binomial selection of a Poisson distribution yields a Poisson distribution [6]. Therefore, the photoelectrons distribution on the j^{th} detector for a gamma-ray interaction at $\mathbf{r}$ with energy $\mathcal{E}$, will also have a Poisson distribution given by

$$p(N_j|\mathbf{r}, \mathcal{E}) = \frac{(\bar{N}_j(\mathbf{r}, \mathcal{E}))^{N_j}}{N_j!} \exp(-\bar{N}_j(\mathbf{r}, \mathcal{E})), \tag{2.8}$$

where the mean $\bar{N}_j(\mathbf{r}, \mathcal{E})$ is given by

$$\bar{N}_j(\mathbf{r}, \mathcal{E}) = \eta_j \ \alpha(\boldsymbol{r})_j \bar{N}_{opt}(\mathcal{E}). \tag{2.9}$$

Furthermore, since the number of photoelectrons in a PMT is Poisson, and the number of photoelectrons on each PMT are independent, the multivariate representation

for measuring all photoelectron for all the J PMTs is given by the multinomial law.

$$p(\{N_j\}|\mathbf{r},\mathcal{E}) = \prod_{j=1}^{J} p(N_j(\mathbf{r},\mathcal{E})|\mathbf{r},\mathcal{E}). \tag{2.10}$$

2.5.3 PMT Gain and Gain Variance

The photoelectrons produced in the photocathode of a PMT represent a very small signal, therefore, they are amplified by 4-6 orders of magnitude to produced a voltage that can be processed with electrons. The PMT's do not amplify all the individual photoelectrons by the same gain, but each photoelectron is amplified by a random gain. The PMT output is futher amplified using an electronic amplifier. If we neglect electronic noise, and denote the PMT's gain as G^{PMT} and the electronics gains as G^{elec}, then we can write the voltage signal as

$$V_j = \sum_{j=1}^{N_j} G_j^{elec} G_j^{PMT} N_j. \tag{2.11}$$

The total signal V_j is the sum of the statistically independent signals from many photoelectrons [5]. By using the central limit theorem, the resulting PDF for the voltage output $\mathbf{V}$ is approximately normal, and specified by the mean and variance. If we neglect the gain variance of the PMT (noise), and combine the two gains, $G_j = G_j^{elec} G_j^{PMT}$, we can rewrite Eq. 2.11 as

$$V_j = N_j G_j. \tag{2.12}$$

Thus, the voltage readout of the j^{th} PMT is proportional to the number of photoelectrons produced at the photocathode, where the constant of proportionality is

the inverse of the gain of the j^{th} PMT. NINT is an operator which returns an integer closest to the nearest integer resulting from the division, we rewrite Eq.2.12 as

$$g_j = \text{NINT}\{\frac{V_j}{G_j}\}. \tag{2.13}$$

This result will be utilized in later chapters, when estimating parameters from list-mode data.

With the output from the PMT's given by Equation 2.13, an equivalent scaled Poisson model of Equation 2.10 can be written as the following

$$p(\{\mathbf{g}\}|\mathbf{r}, \mathcal{E}) = \prod_{j=1}^{J} \frac{(\bar{N}_j)^{g_j}}{g_j!} \exp(-\bar{N}_j), \tag{2.14}$$

which gives the probability of measuring the data $\mathbf{g}$, given a gamma ray with incident energy $\mathcal{E}$ interacted at position $\mathbf{r}$ on the detector face.

2.6 Maximum Likelihood Estimation

A probability density function of the form $p(\mathbf{x}|\boldsymbol{\theta})$ specifies the probability that the data $\mathbf{x}$ is observed given the set of parameters $\boldsymbol{\theta}$. In estimation tasks, since the data is known, the inverse problem arises – finding the values of $\boldsymbol{\theta}$ that produced the data $\mathbf{x}$. The likelihood function represents this problem, and it can be rewritten in similar notation as Lehovich [61],

$$\mathcal{L}(\boldsymbol{\theta}|\mathbf{x}) = p(\mathbf{x}|\boldsymbol{\theta}) \tag{2.15}$$

The maximum likelihood estimation (MLE) maximizes the likelihood function to estimate the most likely values of the parameter $\boldsymbol{\theta}$ that resulted in the observed data.

Mathematically, it can be written in terms of the underlying parameters [6].

$$\hat{\boldsymbol{\theta}}_{MLE} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}(p(\mathbf{x}|\boldsymbol{\theta})) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}(\mathcal{L}(\boldsymbol{\theta}|\mathbf{x})) \tag{2.16}$$

The log function is monotonic, therefore, maximizing the log-likelihood function is equivalent to maximizing the likelihood. It can also be more efficient, for example, if our data set consists of a set of independent and identically distributed random variables, each with an associated PDF, taking the log of the likelihood product of the individual PDF's has the effect of turning the embedded products to sums. The log function also simplifies exponents which are common in a number of PDFs.

$$\hat{\boldsymbol{\theta}}_{MLE} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}(ln(p(\mathbf{x}|\boldsymbol{\theta}))) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}(ln(\boldsymbol{l}(\boldsymbol{\theta}|\mathbf{x}))). \tag{2.17}$$

The score is defined as the gradient of the log-likelihood, with respect to the parameters,

$$\mathbf{s}(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}} ln(\boldsymbol{l}(\boldsymbol{\theta}|\mathbf{x})) \tag{2.18}$$

The Fisher information matrix is the covariance matrix of the score. The F_{ij} element of the Fisher information matrix is given by the expression.

$$F_{ij} = \langle s_i s_j \rangle_{\mathbf{x}|\boldsymbol{\theta}} \tag{2.19}$$

where the angle brackets indicate an ensemble average. The diagonal elements of the inverse of the Fisher information matrix are the Cramèr-Rao bound. The Cramèr-Rao bound is the lower bound on the variance of an unbiased estimator.

$$Var\{\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}\} \geq \mathbf{F}^{-1}(\boldsymbol{\theta})_{ii}, \tag{2.20}$$

If an estimator achieves the Cramér-Rao bound, then the estimator is said to be efficient.

2.6.1 Position Estimation via ML Estimation

Consider a gamma ray which deposits energy $\mathcal{E}$ in a scintillation crystal at position $\mathbf{r}$ generating detector outputs $\mathbf{V}$. A multivariate normally distributed likelihood model with mean $\bar{\mathbf{V}}$ and covariance matrix $\mathbf{K}_V$ can be used to model it.

$$p(\mathbf{V}|\mathbf{r},\mathcal{E}) \approx \mathcal{N}(\bar{\mathbf{V}}(\mathbf{r},\mathcal{E}),\mathbf{K}_V) \tag{2.21}$$

For a given position of interaction $\mathbf{r}$ and energy deposited $\mathcal{E}$, the voltage outputs from the PMT's can be assumed to be statistically independent. However, in the typical gamma-ray detectors, the energy and depth z are tightly coupled, and due to noise, it is hard to accurately estimate both of them. To solve this problem, during the calibration process, a collimated source with gamma rays of a known energy $\mathcal{E}$ excite the detector along a two-dimensional grid, and the mean and covariance of the signal are obtained for each position. The key concept is that the statistical quantities are obtained for a two-dimensional grid, averaging over the z coordinate, which is the depth of interaction. Furthermore, by using the normalized PMT output from Eq. 2.13, the scaled Poisson model from Equation 2.14 can be used, and the only parameters that need to be estimated are the two-dimensional point of interaction on the detector face, $\mathbf{r}_d$. The corresponding simplified log-likelihood version of Equation 2.21 can be written in the following form, which is often utilized at CGRI.

$$\hat{\mathbf{r}}_{MLE} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\{\ln \prod_{j=1}^{J}(\frac{(\bar{N}_j)^{g_j}}{g_j!}\exp(-\bar{N}_j))\}$$

$$\hat{\mathbf{r}}_{MLE} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\{\sum_{j=1}^{J}(g_j \ln \bar{N}_j(\mathbf{r}_d) - \bar{N}_j(\mathbf{r}_d)) - \sum_{j=1}^{J}(g_j!)\}$$

$$\hat{\mathbf{r}}_{MLE} \approx \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\{\sum_{j=1}^{J}(g_j \ln \bar{N}_j(\mathbf{r}_d) - \bar{N}_j(\mathbf{r}_d))\} \tag{2.22}$$

CHAPTER 3

Imaging Theory

3.1 Imaging Theory

The imaging process as described by Barrett and Myers [6], is a linear mapping between Hilbert spaces. In our discussion, we will assumes the class of discrete to discrete mappings only, which for modeling a system on the computer is typical, and the relationship between an object and image known via the system matrix.

3.1.1 H Matrix

The system matrix is a linear, shift-invariant operator, which is a representation of the imaging system that incorporates the imperfections from both the system due to misalignments as well as detector non-linearities [58], and is essential for the forward and inverse problems. The H matrix is a property of the system itself, and can be obtained analytically [88] [8]‘[11], by Monte Carlo methods [12] [79], or directly measured [18] [39] [78], which is the method of choice at CGRI. The advantage of directly measuring the system matrix is the data can be saved and loaded when convenient, which reduces the computation time. However there are disadvantages to this method such as the requirement of a significant amount of storage space depending on the resolution of H, and time needed to calibrate. Also, since the system matrix is specific to the radioactive source that is used in imaging, changing sources requires recalibrating the system.

The system matrix is obtained by combining the Mean Detector Response Function (MDRF) with the Point Spread Function (PSF). The MDRF is obtained by scanning a collimated radioactive source in a two-dimensional grid of 79×79 points on the detector face. The output of the MDRF measurement is used to estimate the position of interaction via a maximum likelihood technique as discussed in Section 2.6.1. The gamma ray detectors at CGRI utilize an array of 3×3 PMT tubes. However, the spreading of light over multiple detectors ensure that the effective detector resolution is much higher. The PSF is measured by moving a radioactive point source through a three-dimensional $21 \times 21 \times 27$ grid of points, which coarsely samples the field of view of the imager. Using the response of the detectors as well as the data of the point source at each voxel, the blur function of the hot source in space for each detector is obtained, which become the columns of the H matrix.

Interpolating H

The object space, or columns of the H matrix, must be interpolated to achieve high resolution. In order to interpolate the object voxels to a higher dimension, first parameterization of the blur function on each detector by a multivariate normal distribution must be performed. A multivariate normal distribution is fully represented by the following coefficients

- A: Amplitude,
- (μ_x, μ_y): Mean,
- (λ_x, λ_y): Major and minor eigenvalues,
- ϕ: Angle of rotation,

with the associated covariance matrix, represented by

$$\mathbf{\Sigma} = \mathbf{R}_{\phi}\mathbf{\Sigma}_0\mathbf{R}_{\phi}^{T}. \tag{3.1}$$

Given these parameters, the Normal distribution is expressed as the following

$$h_d(\mathbf{r}) = \frac{A}{\sqrt{(2\pi)^y|\mathbf{\Sigma}|}} \exp(-\frac{1}{2}(\mathbf{r}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{r}-\boldsymbol{\mu})). \tag{3.2}$$

As discussed in [21] [58], to increase the effective object space resolution of the imager, the response between voxels is obtained by averaging of the Gaussian coefficients from neighboring voxels.

3.1.2 Imaging Equation

The discrete to discrete mapping from a voxelized object $\mathbf{f}$ from $\mathbb{U}$ in $\mathbb{L}_2$ to a projection image $\mathbf{g}$ in $\mathbb{V}$ space expressed in operator form is given by the following expression

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}, \tag{3.3}$$

where $\mathbf{g}$ is an $M \times 1$ vector corresponding to the bins of the detector, $\mathbf{f}$ is an $N \times 1$ vector representing threw object, an $\boldsymbol{n}$ is the noise in the system. As discussed in [6], in a photon-counting detector system such as FastSPECT II, Poisson statistics are used to describe the photon detection process, thus the noise model in the system is also characterized by Poisson statistics.

In FastSPECT II, the object function $f(\mathbf{r})$ is written as a linear combination of the basis functions, or voxels, which can be thought of as small cubes that partition

the field of view.

$$f \approx f_a(\mathbf{r}) = \sum_{n=1}^{N} f_n \phi_n(\mathbf{r}) \tag{3.4}$$

Given the imaging equation from Eq. 3.3, using the system matrix and the measured projection data, trying to estimate the original object would require estimating the $\mathbf{f}$ and $\mathbf{n}$. In the presence of noise, and since the mapping of the object to a discrete image space is generally not a one-to-one process, a modification of Eq. 3.3 is represented by Eq. 3.5, where $\hat{\mathbf{f}}$ is an approximation to the original object.

$$\boldsymbol{g} \approx \boldsymbol{H}\hat{\mathbf{f}} \tag{3.5}$$

where $\hat{\mathbf{f}}$ is an approximation of the object which includes the noise component. An approximation for the object, $\hat{\mathbf{f}}$, can be obtained using iterative methods, which will be briefly discussed below.

The Landweber algorithm in Eq. 3.6 yields an approximation for the object $\hat{\mathbf{f}}$, given the matrix equation in Equation 3.5. Successful implementation of the Landweber algorithm is seen in both SPECT and CT systems [50] [102], and [106]. The Landweber is mentioned here because it will provide a means to estimate parameters from projection data in Sec. 5.2.1.

$$\hat{\mathbf{f}}^{(k+1)} = \hat{\mathbf{f}}^{(k)} + \mathbf{H}^{\dagger}(\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}^{(k)}) \tag{3.6}$$

The MLEM algorithm, is an iterative algorithm that maximizes the likelihood for the data, which in our case is Poisson in nature. The MLEM algorithm is typically utilized at CGRI. Some of the desirable characteristics of MLEM is that the algorithm preserves positivity, while seeking to minimize the Kullback-Leibler distance between

the projected object and the image data.

$$\hat{f}_i^{(k+1)} = \frac{\hat{f}_i^{(k)}}{s_i} \sum_m \frac{g_m}{(\boldsymbol{H}\hat{\boldsymbol{f}}^{(k)})_m} H_{m,i} \tag{3.7}$$

As discussed in [107] [90] [29], iterative reconstruction techniques are accurate since they are able to incorporate system geometry, suffer less from noisy data [65], and have a higher resolution, but come at the cost of extra computational power needed.

CHAPTER 4

SPECT Imaging Systems

4.1 SPECT

Single photon emission computed tomography (SPECT), is a tomographic imaging technique that is used to image the distribution of a gamma-ray emitting isotope. A SPECT imaging systems consists of gamma-ray detectors, and image forming elements such as a pinhole or a parallel-hole collimators that take two-dimensional acquisitions of a three-dimensional object. For a tomographic reconstruction of the object, multiple projections through the object must be obtained at different angles. This requires either the object within the field of view to be rotated, the detector rotated around the object, or multiple detectors as at CGRI.

4.2 FastSPECT II

Four-dimensional Arizona Stationary Single-Photon Emission Computed Tomography II, or FastSPECT II for short, is a SPECT imaging system designed for use with small animals. FastSPECT II has 16 stationary gamma cameras, placed on two rings, with eight cameras on each ring. Each gamma camera has one pinhole. The object is imaged onto the gamma camera with a pinhole of 1mm diameter, placed 1.9 inches from the center of the imaging system axis, providing a magnification of 2.4X. At the center of the imaging system, the field of view is about 40mm along each axis [21].

4.2.1 List-Mode data from FastSPECT II

When an object is imaged, the data from each gamma ray scintillation event is saved in a list-mode data format. The list-mode data, as shown in Fig. 4.1, consists of an array of digitized voltage outputs for each of the 9 PMTs for each recorded event. The header file for each of the detector specifies the number of events recorded and a time stamp. The acquisition electronics used to save the detected scintillation events are discussed in greater detail by Chen [21], and Furenlid [39].

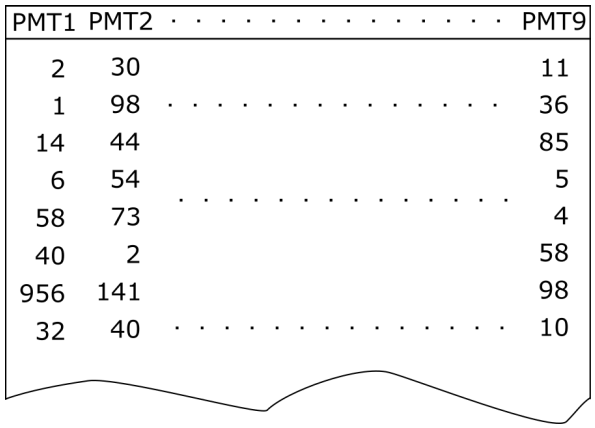


Figure 4.1: Example of the listmode data obtained from a single detector. For each PMT of the detector, and for each event there is an associated voltage.

Each of the 16 detectors in FastSPECT II records a 2-D projection image of the object acquired at a different angle. This eliminates the need to rotate the cameras or the object during the imaging study. In an experiment, the acquired list-mode data is processed, which involves estimating the 2-D position from which a particular event originated from, and creating a histogram of the events for each bin. The format after the list-mode data has been processed is a vector, $\mathbf{g}$ that is $\mathrm{M} \times 1$ elements long. In Sec. 3.1.1, the system mapping for object to image space is being represented by

the sensitivity matrix, hence once the data $\mathbf{g}$ is measured, an estimate for the original object can be obtained using either a least squares or an iterative approach. The drawback to using either of the approaches discussed in Sec. 3.1.1 is the amount of information that is lost due to the data undergoing a binning process.

Using raw list-mode data has several advantages. For instance, the measured data in the raw form is the closest representation of the actual object. When the counts collected for each position is small, the memory footprint of the data is much smaller compared to the case where the projection data is binned. Also, when the data is binned, the original attribute vector is discarded, and the information regarding the individual events is lost. Finally, in the case where the data is binned, the collection time must be complete before binning occurs, and a reconstruction can commence. With list-mode data, an image reconstruction can be initiated in real time as events are being collected [14].

MLEM using Listmode Data

Object reconstruction using list-mode data can be utilized and is well studied and has been implemented at CGRI [16], [17], [53], [61], [72], and [54]. This section serves as a compliment to the MLEM algorithm for binned data from Eq. 3.7 from Sec. 3.1.2.

In contrast to the binned approach, an image reconstruction using list-mode data does not utilize a system matrix. Instead the raw PMT values are measured and saved, and the data is represented as a set of attribute vectors $\{\mathbf{g}_i : i = 1, \cdots, N_{list}\}$, where each attribute vector $\mathbf{g}_i$ gives information about an individual scintillation event. Each individual attribute vector, consists of nine elements, one voltage output for each PMT, or ten elements if the time stamp is included. A probability density

function is created, which relates the object to the measured data, $p(\mathbf{g}_i|\mathbf{f})$. The object is estimated using list-mode attribute data, by utilizing a probability model relating the PMT voltages to the object. The likelihood of measuring $\mathbf{g}$ given the scintillation event i, was emitted from the the object $\mathbf{f}$, is represented by

$$p(\mathbf{g}_i|\mathbf{f}) = \sum_{n=1}^{N_{basis}} p(\mathbf{g}_i|n)p(n|\mathbf{f}) \tag{4.1}$$

where the region associated with the basis function $\phi_n(\mathbf{r})$, will indicated by the integer n. From Eq. 4.1 the density function $p(n|\mathbf{f})$ represents the the probability that the i^{th} event originating from the object $\mathbf{f}$ is associated with the n^{th} voxel. The elements of f_n are proportional to the strength of the signal, or the number of gamma-ray photons being emitted from the particular voxel. Each voxel has an associated sensitivity, which is a probability that a signal from voxel n will be detected. Hence, a representation for the probability of an event being detected originating in voxel n can be written as

$$p(n|\mathbf{f}) = \frac{f_n s_n}{\sum_{n'=1}^{N_{basis}} f_{n'} s_{n'}} \tag{4.2}$$

The density function, $p(\mathbf{g}_i|n)$, is the probability that an event originating from voxel n will produce the attribute vector $\mathbf{g}_i$ associated with the i^{th} event.

Given the measured data, we would like to obtain a representation for the original object. The process involves utilizing the likelihood, or the log-likelihood function over the entire data set $\mathbf{g}_i$

$$l(\hat{\mathbf{f}}|\mathbf{g}_i) = \log p(\mathbf{g}_i|\hat{\mathbf{f}}) \tag{4.3}$$

and maximizing the likelihood function over the unknown coefficients $\hat{\mathbf{f}}$

$$\hat{\mathbf{f}}_{ML} = \underset{\hat{\mathbf{f}}}{\operatorname{argmax}}\, l(\mathbf{f}|\mathbf{g}_i). \tag{4.4}$$

The final list-mode maximum likelihood is shown below, which is derived in detail in [14], [7], and [61].

$$\hat{f}_n^{(k+1)} = \sum_{i=1}^{N_{list}} \frac{pr(\mathbf{g}_i|n)\hat{f}_n^{(k)}}{\sum_{n'=1}^{N_{basis}} pr(\mathbf{g}_i|n')\hat{f}_{n'}^{(k)} s_{n'}} \tag{4.5}$$

CHAPTER 5

Estimation of Activity Curve from Image Data

Early Pharmacokinetic modeling methodologies have evolved from single compartment linear decay relationships of a zero or first order process [93] [98], to multi compartment [57], complex stochastic models [60] [34] that incorporate random system such as variations of heat and how it changes the solubility of the drug within the plasma.

The models to describe the drug kinetics within a system are numerous [100] [99] [75] [27], and PK models are valuable for drug development [87] [96] [89], cancer imaging [104], as well as patient dosage regimens [92] [36] [37].

Furthermore, the applications of imaging in classical PK experiments for both PET [44] [64] [9] an SPECT systems [83] [109], has enabled estimation of the kinetic parameters *indirectly* from reconstructions [31] [32] [91] [23] [59], or *directly* from projection data [62] [108] [47] [49] [55] [80] [22] [63]. In current research, the estimation of the kinetic parameters from experiments is automatically factored in to reconstruction algorithms, in what is referred to as a 4D reconstruction [2] [81] [56] that utilize a likelihood model and estimate parameters via a ML estimation techniques [35].

The interested reader is encouraged to follow up on the following papers to get more information regarding models, and topics related to pharmacokinetics [66] [71] [105] [52].

The goal of this research is a feasibility study to determine the estimability of the kinetic parameters using a novel approach combing an indirect approach involving

estimating the activity curve will be estimated from raw, un-binned list-mode data, and estimating the kinetic parameters from the activity curve. The particular model that will be used in this research is linear vector differential equation describing the kinetics in a dense system [75] [67].

Estimation of the activity curve will from a system using binned and list-mode data, and estimating the kinetic parameters from the approximated activity curve will be covered in Chapter 6.

5.1 Overview: Two Compartment Model Simulation

Using the kinetic equation from Eq. 1.18, an open two compartment model was simulated with known kinetic parameters $\mathbf{K}$. For a particular $\mathbf{K} = [-2,\ 2; 1,\ -3]$ which satisfies the constraints discussed in Sec. 1.4.3, the continuous time activity curves for each compartment are shown in Fig. 5.1. The chosen time range was $t = (1, \cdots, 10)$ seconds, with a bolus input, $\mathbf{I}(t)$, into compartment 1 for total duration of one second.

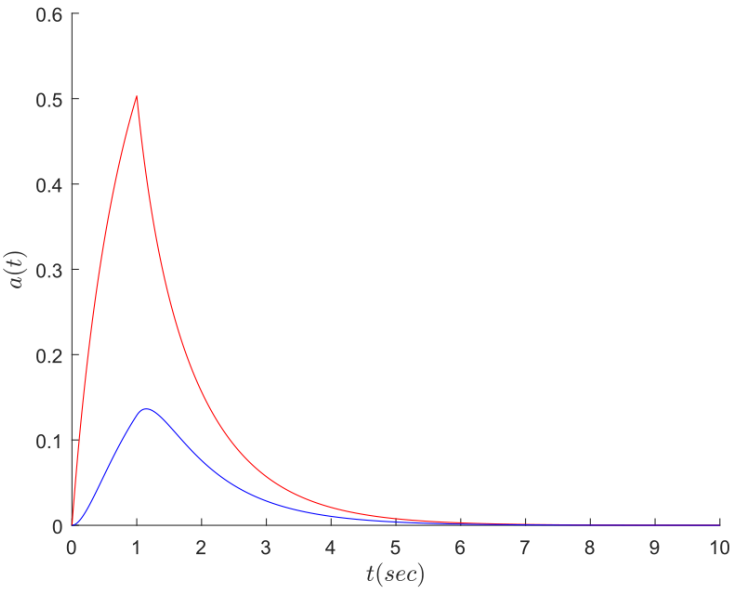


Figure 5.1: Time activity curve for two-compartment model. Red curve represents compartment 1 tracer kinetics, blue curve represents compartment 2 tracer kinetics.

5.1.1 Time Sample Density

For simulation purposes, the continuous time varying activity curves, from Sec. 5.1 needed to be discretized. We are interested in the tracer kinetics after the drug has been completely administered to the system, only samples of the activity curve after $t = t_{input}$ were considered. Choosing the points to sample the activity curve has been explored [84] [74] and the results suggest the optimal time sampling is a function of the model itself. Since the time activity curves in our simulations are exponentially decaying, the time step Δt between sample points of the activity curve will also follow an exponential type function: immediately to the right of $t = t_{input}$, Δt will be smallest, and increase exponentially as $t = t_{max}$ as shown in Fig. 5.2.

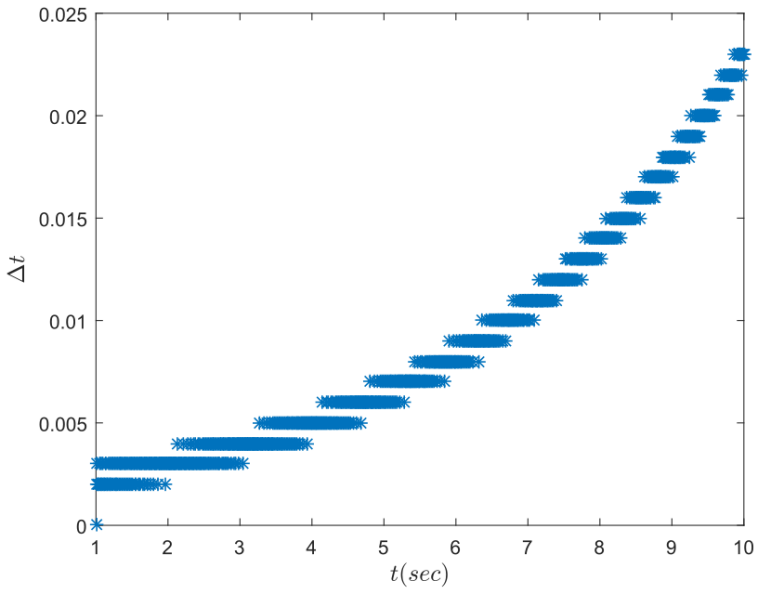


Figure 5.2: Near the input time cutoff $t_{input} = 1$, the activtiy curve is sampled more, the spacing between sample points Δt is smaller, and increases as t increases.

Observe in Fig. 5.3, for thirty time samples of the activity curve, $N_t = 30$, Δt is smaller near $t = t_{input} = 1$, and larger near $t = t_{max} = 10$.

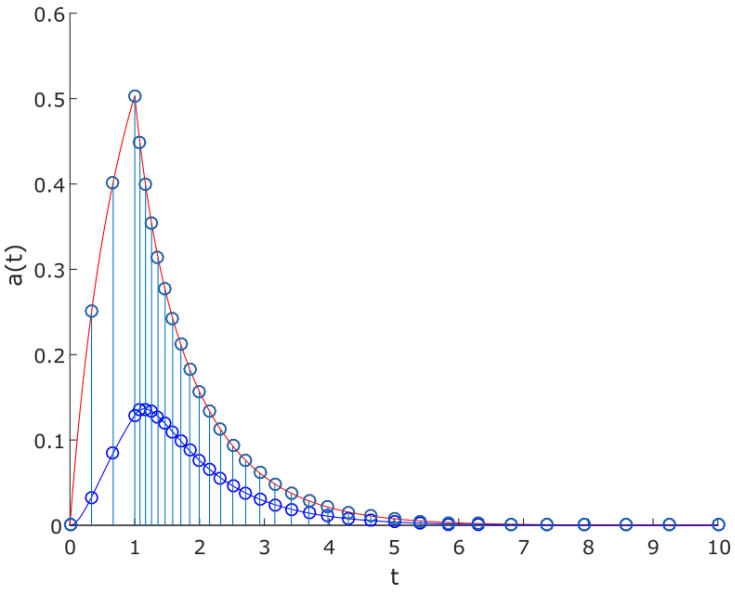


Figure 5.3: Sampling of the time activity curve. Near $t = t_{input}$ the curve is sampled more as opposed to near $t = t_{max}$, where the activity curve is sampled less.

5.2 Estimation of Activity from an Imaging System

5.2.1 Binned Data Approach

Using MOBY, a digital phantom was generated consisting of multiple compartments. Each compartment of the MOBY phantom was fit to discrete time points of the time activity curve such that the summation of all the voxels within a specific compartment support region was equal to the activity $a_l(t)$ in the individual compartment at time t. Fitting the object to the activity curve was necessary for the time evolution of the tracer within the system to be studied.

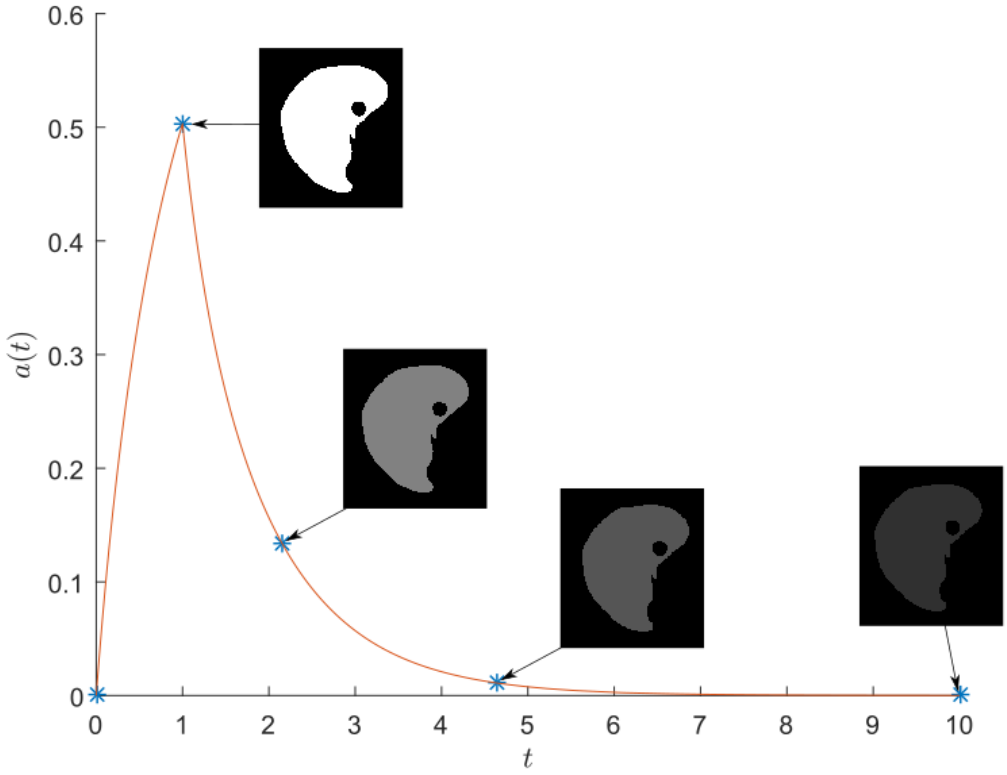


Figure 5.4: Example of object (liver) of interest being fit to multiple discrete time points from the time activity curve.

$$a_l(t) = \sum_{n \in S_l} f_n(t) \tag{5.1}$$

Reconstruction Data

Using the system H matrix, a time varying projection image was obtained following the using the imaging equation in Sec. 3.1.2. From the projection data, a reconstruction of the original object, at each time sample point was obtained from the

projection data using the maximum likelihood estimation maximization algorithm (MLEM). Once the object reconstruction was obtained, using the spatial support region knowledge at each time sample point, the compartmental activity was obtained using Eq. 1.9, and was used to construct an approximation for the time activity curve. In Fig. 5.5, multiple slices of reconstruction data for the MOBY phantom are shown at a fixed time.

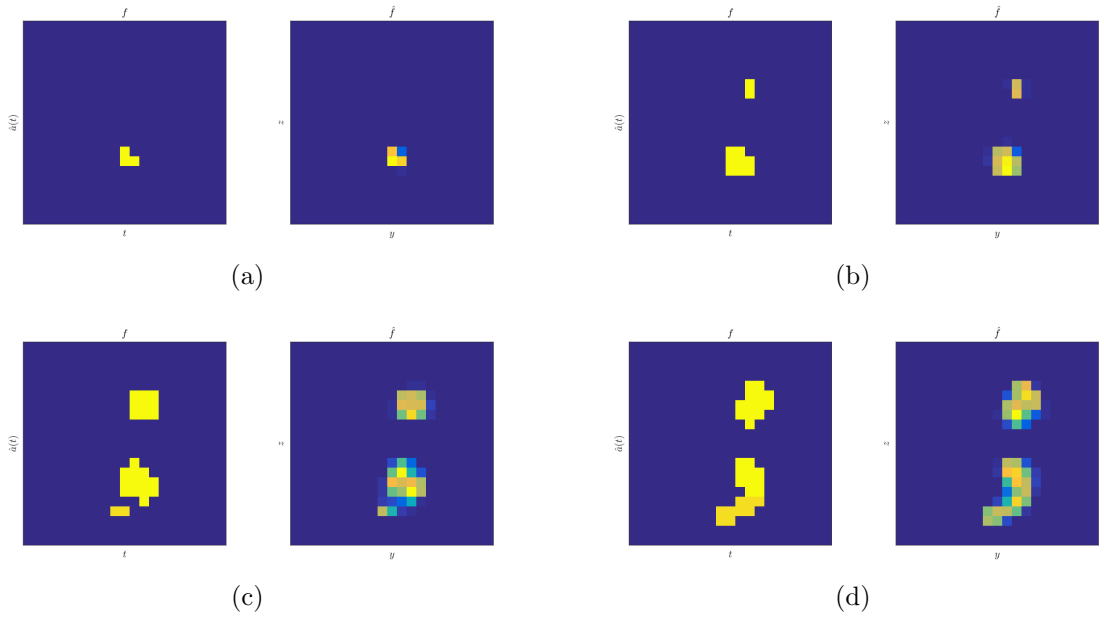
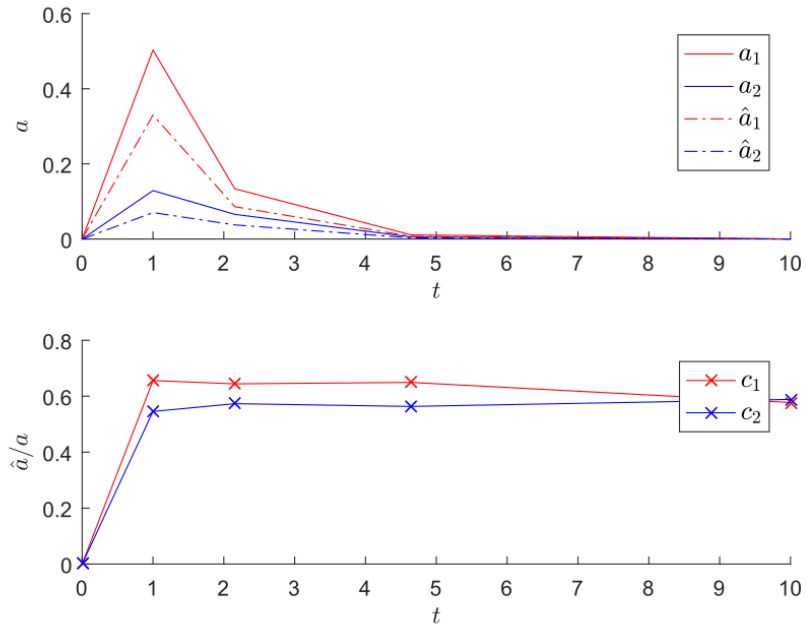
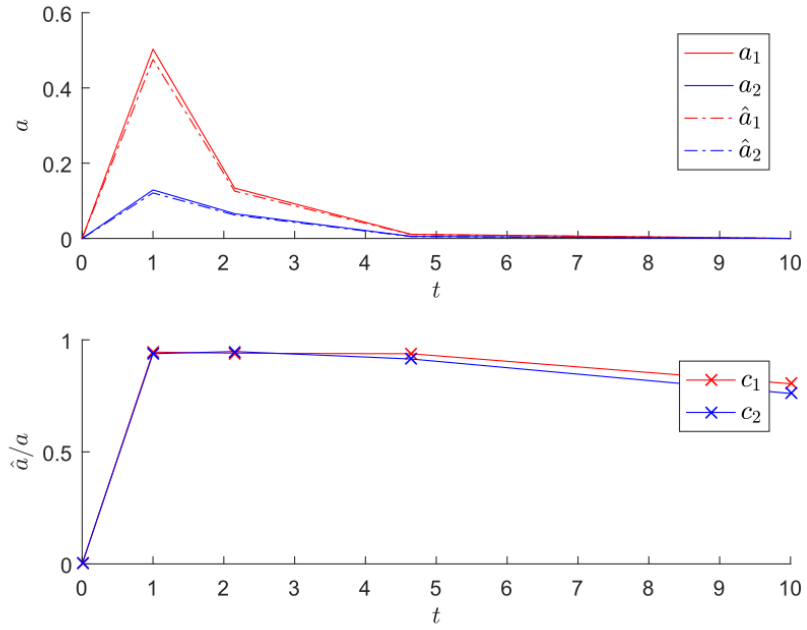


Figure 5.5: Multiple slices of the reconstruction data of the MOBY phantom at a fixed time.

Using knowledge of the support region, the reconstruction algorithm was efficient and converged after only a few iterations. Fig. 5.6 shows the convergence of the activity curves for two cases, Fig. 5.6a when the number of MLEM iteration is small (5), and Fig. 5.6b the number of iterations of the MLEM algorithm is larger (60).



(a)



(b)

Figure 5.6: Convergence of $\hat{\mathbf{a}}(t)$ to $\mathbf{a}(t)$ as the number of iterations of the MLEM algorithm increases from (5) Fig. 5.6a to (60) Fig. 5.6b.

Projection Data

Obtaining an approximation for the activity via projection data is the next step, which proves advantageous for the following reasons.

1. Using the projection data removes the intermediate reconstruction step saving processing time.

2. Using projections data maintains information that would have been further lost during reconstruction due to binning.

In order to estimate the activity curve from the projection data, a mapping similar to that in Eq. 1.8 needs to be constructed, which instead maps the projection data $\mathbf{g}$ to the activity vector $\mathbf{a}$. Let $\mathbf{E}$ denote the mapping from data space to compartment space as shown below.

$$\mathbf{a} = \mathbf{E}\mathbf{g} \tag{5.2}$$

Recall that the object to image relationship in the absence of noise is given by the following expression.

$$\mathbf{g} = \mathbf{H}\mathbf{f} \tag{5.3}$$

Combining Eq. 1.9, 5.2, and 3.5, an equivalent expression for $\mathbf{D}$ can be obtained.

$$\mathbf{a} = \mathbf{D}\mathbf{f} = \mathbf{E}\mathbf{g} = \mathbf{E}\left(\mathbf{H}\mathbf{f}\right) \tag{5.4}$$

$$\mathbf{D} = \mathbf{E}\mathbf{H} \tag{5.5}$$

For the sensitivity matrix $\mathbf{H}$, we assume that each voxel within the field of view has the potential to produce a signal on each of the detectors. If the rank of the sensitivity matrix is such that $R = M$, where M is the total number of pixels on all detectors of FastPSPECT II ($M = 16 * 79 * 79 = 99856$), Eq. 5.5, $\mathbf{H}$ has an associated right-hand inverse, $\mathbf{H_R^{-1}}$, such that $\mathbf{H H_R^{-1}} = \mathbf{I_M}$. Solving for the operator $\mathbf{E}$ can be done in theory, since given both $\mathbf{D}$ and $\mathbf{H}$ are known. Directly solving for $\mathbf{E}$ via inverse methods would not be feasible due to storage constraints.

$$\mathbf{E} = \mathbf{D}\left(\mathbf{H}^\dagger\right)\left(\mathbf{H}\mathbf{H}^\dagger\right)^{-\mathbf{1}} \tag{5.6}$$

Rather than solve for $\mathbf{E}$ directly, an iterative method utilizing the Landweber algorithm from Eq. 3.6, could be used to approximate the compartmental activity directly. Multiplying both sides of the LWA by $\mathbf{D}$ from Sec. 1.4.2, an expression yielding an approximate expression for the compartmental activity from the projection data, $\mathbf{D}$, and $\mathbf{H}$ is obtained,

$$\mathbf{D}\hat{\mathbf{f}}^{(k+1)} = \mathbf{D}\hat{\mathbf{f}}^{(k)} + \mathbf{D}\mathbf{H}^\dagger\left(\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}^{(k)}\right) \tag{5.7a}$$

$$\hat{\mathbf{a}}^{(k+1)} = \hat{\mathbf{a}}^{(k)} + \mathbf{D}\mathbf{H}^\dagger\mathbf{g} - \mathbf{D}\mathbf{H}^\dagger\mathbf{H}\hat{\mathbf{f}}^{(k)} \tag{5.7b}$$

$$\hat{\mathbf{a}}^{(k+1)} = \hat{\mathbf{a}}^{(k)} + \mathbf{D}\mathbf{H}^\dagger\mathbf{g} - \mathbf{D}\mathbf{H}^\dagger\mathbf{H}\tilde{\mathbf{D}}\hat{\mathbf{a}}^{(k)} \tag{5.8}$$

where $\tilde{\mathbf{D}}$ is given by Eq. 1.11, and is shown again below, and N_l indicates the number of non-zero voxels within the l^{th} compartment.

$$\tilde{\mathbf{D}} = \mathbf{D}^T\mathrm{diag}(1/N_1, 1/N_2, \cdots, 1/N_L) \tag{5.9}$$

5.2.2 List-Mode Data Approach

The feasibility of estimating the activity curve is explored in the following sections using a two-compartment model with parameters discussed in Sec. 5.1.

Kernel Density Estimation

Before discussing the methods in which the activity curves were estimated from the raw list-mode data, some basic kernel density estimation must be introduced. This material was gathered from multiple sources and course notes [43] [46] [110] [86] [85] [45] and [70]. Given the measured data, in order to perform a statistical analysis, knowledge of the probability distribution which the data was sampled from is necessary. Density estimation revolves around using the data to estimate the PDF, Given a continuous, real valued random variable $\mathbf{X}$ that can be described in terms of a probability density function $p(x)$, the probability that the random variable falls within a certain interval $[a, b]$ can be expressed as

$$P(a \leq X \leq b) = \int_a^b p(x)dx. \tag{5.10}$$

There are two main types of density estimation techniques: a parametric and a non-parametric approach. The parametric estimation lacks flexibility since the shape of the PDF is assumed to be a known distribution. A non-parametric approach does not assume an initial shape of the distribution, and the shape can be entirely constructed using the measured data only. For the research presented here, a non-parametric approach will be explored. A well known non-parametric estimate of the probability density function is the histogram. Multiple bins of width B partition the

measurement space, and are typically of the form

$$b\left(x,B\right)=\begin{cases}1 & \left(x+B\left(i-1\right)b,x+iB\right)\\ 0 & elswhere\end{cases} \tag{5.11}$$

where the estimate of the PDF in terms of the kernel function is given by

$$\hat{p}\left(x\right)=\frac{1}{nB}\sum_{i}^{n}b_{B}\left(x-x_{i}\right). \tag{5.12}$$

From 5.12 above, the width of the bin, as well as the center location plays an important factor in the shape of the estimated PDF. When the data is binned, only the counts recorded from the object within the bin of width B are saved, and the actual data values are discarded. Furthermore, histogramming the data yields a discrete PDF.

An alternative to the histogram is to replace the bin with a kernel function. Kernel density estimation, also known as the Parzen-Rosenblatt window method, is a convenient non-parametric method to estimate the probability density function of an independent and identically distributed (i.i.d.) random variable. In a KDE approach, rather than a uniform bin, a smooth and continuous kernel of a specific bandwidth σ, is placed at each of the data points. The kernel function $k(u)$ is chosen such that it is normalized over all space, and for convenience, it is positive and radially symmetric. Therefore we have

$$\begin{aligned}\int k(u)du&=1\\ \int uk(u)du&=0.\end{aligned} \tag{5.13}$$

Given the kernel function, $k(x)$, the estimate for the PDF can be written as

$$\hat{p}(x) = \frac{1}{N}\sum_{i=1}^{N} k_\sigma (x - x_i) \quad \text{or} \tag{5.14a}$$

$$\hat{p}_\sigma(x) = \frac{1}{N\sigma}\sum_{i=1}^{N} k\left(\frac{x - x_i}{\sigma}\right) \tag{5.14b}$$

In Eq. 5.14, the choice of the kernel bandwidth is critical. Typically the bandwidth is chosen such that the mean-square error, which is the error between the true and estimated density, is minimized. The MSE is a method to quantify the accuracy of our estimator, and is represented as

$$\begin{gathered} MSE(\hat{p}) = E\left[(p(x) - \hat{p}(x))^2\right] = \\ (E[\hat{p}(x)] - p(x))^2 + E\left[(p(x) - E[\hat{p}(x)])^2\right] \\ Bias^2[\hat{pr}(x)] + Var[\hat{p}(x)] \end{gathered} \tag{5.15}$$

From Eq. 5.15, the bandwidth is highly dependent on the bias and variance of the estimator, which can be observed as the expressions for each are expanded. First the bias will be calculated,

$$E[\hat{p}_\sigma(x)] = \frac{1}{N}\sum_{i=1}^{N} E\left[\frac{1}{\sigma}k\left(\frac{x - x_i}{\sigma}\right)\right] \tag{5.16}$$

and using a substitution of variables $u = \frac{x - x_i}{\sigma}$, the expectation value from the right hand side of Eq. 5.16 can be written as,

$$E\left[\frac{1}{\sigma}k\left(\frac{x - x_i}{\sigma}\right)\right] = \int (k(u)p(x - \sigma u))\, du$$

which can be further simplified using a Taylor series expansion of the density in terms of (σu).

$$= \int k(u) \left(p(x) - \sigma u p^{(1)}(x) + \frac{(\sigma u)^2}{2} p^{(2)}(x) + O(\sigma^2) \right) du$$

Furthermore, after imposing the symmetry of the kernel from Eq. 5.13, while also substituting $\mu_j(k) = \int_\infty u^j k(u) du$, Eq. 5.16 can be simplified to

$$E\left[\hat{p}_\sigma(x)\right] = p(x) + \frac{\sigma^2 \mu_2}{2} p^{(2)}(x) + O(\sigma^2), \tag{5.17}$$

The final expression for the bias of the estimator using the simplified expression for the mean from Eq. 5.17 can be written as

$$Bias\left[\hat{p}_\sigma(x)\right] = E\left[\hat{p}_\sigma(x)\right] - p(x) = \frac{\sigma^2 \mu_2(x)}{2} p^{(2)}(x) + O(\sigma^2). \tag{5.18}$$

The variance of the estimator can be derived in a similar fashion.

$$Var\left[\hat{p}_\sigma(x)\right] = \frac{1}{N} \left(E\left[\frac{1}{\sigma^2} k^2 \left(\frac{x - x_i}{\sigma} \right) \right] - \left(E\left[\frac{1}{\sigma} k \left(\frac{x - x_i}{\sigma} \right) \right] \right)^2 \right) \tag{5.19}$$

After a change of variables, a Taylor expansion, and using the expression obtained for the bias from Eq. 5.18, the variance can be simplified further.

$$\begin{aligned} Var\left[\hat{p}_\sigma(x)\right] &= \frac{1}{N} \left(\frac{1}{\sigma} \int \left(k^2(u) p\left(x - \sigma u\right) \right) du - \left(p(x) + Bias\left(\hat{p}(x)\right) \right) \right) = \\ &= \frac{1}{N} \left(\frac{1}{\sigma} \int \left(k^2(u) \left(p(x) - \sigma u p^{(1)}(x) + O(\sigma) \right) \right) du - \left(p(x) + O\left(\sigma^2\right) \right) \right) \end{aligned} \tag{5.20}$$

We now have

$$Var\left[\hat{p}_\sigma(x)\right] = \frac{p(x)}{N\sigma} R(k) + O\left(\frac{1}{N\sigma} \right), \tag{5.21}$$

where $R(k) = \int k^2(u)du$. Using the expressions for the bias and variance in Eq. 5.18 and 5.21, the mean-squared error in Eq. 5.15 becomes

$$MSE\left(\hat{p}(x)\right) \approx \frac{\sigma^4 \mu_2^2 (p^{(2)}(x))^2}{4} + \frac{1}{N\sigma} R(k) p(x) \tag{5.22}$$

and the mean integrated square error, which is a measure of the global accuracy of the estimator

$$MISE\left(\hat{p}(x)\right) \approx \frac{\sigma^4 \mu_2^2 R(p^{(2)})}{4} + \frac{1}{N\sigma} R(k) \tag{5.23}$$

where $R(p^{(2)}) = \int (p^{(2)}(x))^2 dx$, is the measure of the roughness of the kernel. From Eq. 5.23, it is evident that as the number of samples gets large, $N \to \infty$, and the bandwidth of the kernel gets small, $\sigma \to 0$, the MISE tends to zero. However, the bandwidth must go to zero at a slower rate than the increase in the sample size. Hence, given a set number of samples, there exists an ideal bandwidth that minimizes the MISE, in order to find it, there is a trade-off between the bias and the variance. For our simulations research, a Gaussian kernel will be used since it is smooth, and is easily implemented in code.

$$k_\sigma(x) = \frac{1}{(2\pi)^{(d/2)}} \exp\left(-\frac{1}{2} x^T x\right) \tag{5.24}$$

For a Gaussian kernel, by assuming that the underlying distribution is normal, the optimal bandwidth can be obtained using Eq. 5.25 Silvermann's rule which is given by the following.

$$\sigma = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{1/5} \approx 1.06 \hat{\sigma} n^{-1/5} \tag{5.25}$$

where $\hat{\sigma}$ the standard deviation of the data. In situations where several estimations are necessary on large data sets, subjectively choosing the bandwidth parameter

via an automatic procedure must be performed which makes use of two types of techniques for selecting the bandwidth parameter: a plug in method and classical method, which makes use of various cross-validation techniques.

5.2.3 KDE Applied to LM Data

Using KDE, we can construct a distribution, $p(\mathbf{v}|d,\mathbf{f})$, representing the probability that an object produces a vector of voltage outputs $\mathbf{v}$ on a particular detector d. This is accomplished by placing a kernel at each of the voltage outputs of the list-mode data list. Using KDE allows the data to be represented as a continuous probability density function. Below is the representation of the probability density function, where $G(\mathbf{v})$ is a multivariate version of the Gaussian kernel from Eq. 5.24. The detector index is represented by d, k represents the recorded event, and K_d represents the total number of events collected on the detector.

$$p(\mathbf{v}|d,\mathbf{f}) = \frac{1}{K_d}\sum_{k\in d}^{K_d} G_\sigma(\mathbf{v}-\mathbf{v}_{d,k}) \tag{5.26}$$

From the list-mode data set, we need to estimate the temporal activity curve. Since the object is a function of the activity in each compartment, $\mathbf{f}(\hat{\mathbf{a}})$, we can represent the PDF from Eq. 5.26 in terms of the parameters of interest, $p(\mathbf{v}|d,\mathbf{f}(\hat{\mathbf{a}}))$. The PSF calibration data from FastSPECT II is structured such that for each detector voxel combination, $J_{d,n}$ events that are recorded, and for each individual event j, a vector of voltages $\mathbf{v}_j$ is obtained. Using KDE, we can obtain a continuous probability density function $p(\mathbf{v}|d,n))$ which gives us the most probable voltage output

from each detector-voxel combination.

$$p(\mathbf{v}|d,n) = \frac{1}{J_{d,n}} \sum_{j}^{J_{d,n}} G_{\sigma}(\mathbf{v} - \mathbf{v}_{d,n,j}) \tag{5.27}$$

By multiplying 5.27 by the object we obtain an expression which predicts the voltages that an object with a specific activity will produce, where $f_n(\hat{\mathbf{a}}) = \hat{a}_l/N_l$ is the object as a function of the compartmental activity to be estimated.

$$p\left(\mathbf{v}|d,\mathbf{f}\left(\hat{\mathbf{a}}\right)\right) = \frac{1}{N_d} \sum_{n\in d}^{N_d} f_n\left(\hat{\mathbf{a}}\right) p(\mathbf{v}|d,n) \tag{5.28}$$

Given the expressions for the densities in Eq. 5.26 and 5.28, a least squares method will be used to find the activity parameters.

$$\hat{\mathbf{a}}_{LS} = \underset{\mathbf{a}}{\operatorname{argmin}} \sum_{d=1}^{D} \int \left(\left\| p(\mathbf{v}|d,\mathbf{f}) - p(\mathbf{v}|d,\mathbf{f}\left(\hat{\mathbf{a}}\right)) \right\|^2 \right) d\mathbf{v} \tag{5.29}$$

Derivation of Simplified Equation

In order to use Eq. 5.29 to estimate the compartmental activity, some simplification is required to run as efficiently as possible.
Given a multivariate normal distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$

$$N_{\boldsymbol{x}}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}\right) = \left((2\pi)^{Q} \left|\boldsymbol{\Sigma}\right| \right)^{-1/2} \exp\left(-\frac{1}{2} \left(\boldsymbol{x}-\boldsymbol{\mu}\right)^{T} \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{x}-\boldsymbol{\mu}\right) \right) \tag{5.30}$$

that is normalized over all space,

$$\int_{\mathbb{R}^d} N_{\boldsymbol{x}}\left(\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}}\right) d\boldsymbol{x} = 1 \tag{5.31}$$

the integral product of two multi-variate normal distributions can be simplified to the following

$$\int_{\mathbb{R}^d} N_{\boldsymbol{x}}\left(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1\right) N_{\boldsymbol{x}}\left(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2\right) d\boldsymbol{x} = \int_{\mathbb{R}^d} c_3 N_{\boldsymbol{x}}\left(\boldsymbol{\mu}_3, \boldsymbol{\Sigma}_3\right) d\boldsymbol{x} = c_3 \tag{5.32}$$

where the mean and variance of the combined product is

$$\boldsymbol{\Sigma}_3 = \left(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1}\right)^{-1} \tag{5.33}$$

$$\boldsymbol{\mu}_3 = \left(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1}\right)^{-1} \left(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2\right) \tag{5.34}$$

and the constant is

$$\begin{aligned} c_3 = N_{\boldsymbol{\mu}_1}\left(\boldsymbol{\mu}_2, \left(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2\right)\right) = \\ \left(\left(2\pi\right)^Q \left|\left(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2\right)\right|\right)^{-1/2} \exp\left(-\frac{1}{2}\left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\right)^T \left(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2\right)^{-1} \left(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\right)\right). \end{aligned} \tag{5.35}$$

The PDF from Eq. 5.26 and Eq. 5.28 can be written in a form similar to Eq. 5.30

$$p(\mathbf{v}|d, f) = \frac{1}{K_d} \sum_{k=1}^{K_d} N_{\mathbf{v}}\left(\mathbf{v}_{d,k}, \boldsymbol{\Sigma}_{d,k}\right) \tag{5.36}$$

$$p(\mathbf{v}|d, f(\hat{\mathbf{a}})) = \frac{1}{N} \sum_{n}^{N} f_n(\hat{\mathbf{a}}) \frac{1}{J_{d,n}} \sum_{j}^{J_{d,n}} N_{\mathbf{v}}\left(\mathbf{v}_{d,n,j}, \boldsymbol{\Sigma}_{d,n,j}\right) \tag{5.37}$$

Returning to the least square expression from Eq. 5.29, the goal is to expand the terms within the parenthesis, and simplify such that it could be implemented in an simulation.

$$\int \left(p(\mathbf{v}|d, f) - p(\mathbf{v}|d, f(\hat{\mathbf{a}}))\right)^2 d\mathbf{v} =$$

$$\int \left|p\left(\boldsymbol{v}|d,f\right)\right|^2 d\mathbf{v}+ \tag{5.38a}$$

$$\int \left|p\left(\boldsymbol{v}|d,f\left(\hat{\mathbf{a}}\right)\right)\right|^2 d\mathbf{v}- \tag{5.38b}$$

$$2\int p\left(\boldsymbol{v}|d,f\right)p\left(\boldsymbol{v}|d,f\left(\hat{\mathbf{a}}\right)\right)d\mathbf{v} \tag{5.38c}$$

The expression from Eq. 5.38a becomes

$$\begin{aligned}\frac{1}{K_d^2}\sum_{k\in d}^{K_d}\sum_{k^*\in d}^{K_d}\int\left(N_{\mathbf{v}}\left(\mathbf{v}_{d,k},\boldsymbol{\Sigma}_{d,k}\right)N_{\mathbf{v}}\left(\mathbf{v}_{d,k^*},\boldsymbol{\Sigma}_{d,k^*}\right)\right)d\mathbf{v} = \\ \frac{1}{K_d^2}\sum_{k\in d}^{K_d}\sum_{k^*\in d}^{K_d}c_{d,k,k^*}\int\left(N_{\mathbf{v}}\left(\mathbf{v}_{d,k,k^*},\boldsymbol{\Sigma}_{d,k,k^*}\right)\right)d\mathbf{v} = \\ \frac{1}{K_d^2}\sum_{k\in d}^{K_d}\sum_{k^*\in d}^{K_d}\left(c_{d,k,k^*}\right)\end{aligned} \tag{5.39}$$

where the constant, c_{d,k,k^*}, is

$$c_{d,k,k^*} = \mathcal{N}_{\mathbf{v}_{d,k}}\left(\mathbf{v}_{d,k^*},\left(\boldsymbol{\Sigma}_{d,k}+\boldsymbol{\Sigma}_{d,k^*}\right)\right). \tag{5.40}$$

The expression from Eq. 5.38b becomes

$$\begin{aligned}
&\frac{1}{N^2}\sum_{n}^{N} f_n(\hat{\mathbf{a}})\frac{1}{J_{d,n}}\sum_{n^*}^{N} f_{n^*}(\hat{\mathbf{a}})\frac{1}{J_{d,n^*}} \\
&\int\left(\sum_{j}^{J_{d,n}}\sum_{j^*}^{J_{d,n^*}} N_{\mathbf{v}}\left(\mathbf{v}_{d,n,j},\boldsymbol{\Sigma}_{d,n,j}\right) N_{\mathbf{v}}\left(\mathbf{v}_{d,n^*,j^*},\boldsymbol{\Sigma}_{d,n^*,j^*}\right)\right) d\mathbf{v} = \\
&\frac{1}{N^2}\sum_{n}^{N} f_n(\hat{\mathbf{a}})\frac{1}{J_{d,n}}\sum_{n^*}^{N} f_{n^*}(\hat{\mathbf{a}})\frac{1}{J_{d,n^*}} \\
&\sum_{j}^{J_{d,n}}\sum_{j^*}^{J_{d,n^*}} c_{d,nj,(nj)^*}\int\left(N_{\mathbf{v}_{d,n,j}}\left(\mathbf{v}_{d,nj,(nj)^*},\boldsymbol{\Sigma}_{d,nj,(nj)^*}\right)\right) d\mathbf{v} = \\
&\frac{1}{N^2}\sum_{n}^{N} f_n(\hat{\mathbf{a}})\frac{1}{J_{d,n}}\sum_{n^*}^{N} f_{n^*}(\hat{\mathbf{a}})\frac{1}{J_{d,n^*}}\sum_{j}^{J_{d,n}}\sum_{j^*}^{J_{d,n^*}}\left(c_{d,nj,(nj)^*}\right)
\end{aligned} \tag{5.41}$$

where the constant, $c_{d,nj,(nj)^*}$, is

$$c_{d,nj,(nj)^*} = \mathcal{N}_{\mathbf{v}_{d,n,j}}\left(\mathbf{v}_{d,n^*,j^*},\left(\boldsymbol{\Sigma}_{d,n,j}+\boldsymbol{\Sigma}_{d,n^*,j^*}\right)\right) \tag{5.42}$$

And finally, the expression from Eq. 5.38c becomes

$$\begin{aligned}
&-2\frac{1}{K_d N}\sum_{k\in d}^{K_d}\sum_{n}^{N} f_n(\hat{\mathbf{a}})\frac{1}{J_{d,n}}\sum_{j}^{J_{d,n}}\int\left(N_{\mathbf{v}}\left(\mathbf{v}_{d,k},\boldsymbol{\Sigma}_{d,k}\right) N_{\mathbf{v}}\left(\mathbf{v}_{d,n,j},\boldsymbol{\Sigma}_{d,n,j}\right)\right) d\mathbf{v} = \\
&-2\frac{1}{K_d N}\sum_{k\in d}^{K_d}\sum_{n}^{N} f_n(\hat{\mathbf{a}})\frac{1}{J_{d,n}}\sum_{j}^{J_{d,n}}\int\left(c_{d,k,(nj)}N_{\mathbf{v}_{d,k}}\left(\mathbf{v}_{d,k,nj},\boldsymbol{\Sigma}_{d,k,nj}\right)\right) d\mathbf{v} = \\
&-2\frac{1}{K_d N}\sum_{k\in d}^{K_d}\sum_{n}^{N} f_n(\hat{\mathbf{a}})\frac{1}{J_{d,n}}\sum_{j}^{J_{d,n}}\left(c_{d,k,(nj)}\right)
\end{aligned} \tag{5.43}$$

where the constant $c_{d,k,(nj)}$, is

$$c_{d,k,(nj)} = \mathcal{N}_{\mathbf{v}_{d,k}}\left(\mathbf{v}_{d,n,j}, (\mathbf{\Sigma}_{d,k} + \mathbf{\Sigma}_{d,n,j})\right) \tag{5.44}$$

Combining the final simplified expressions for Eq. 5.39, 5.41, and 5.43, we obtain the following for Eq. 5.29, which is a simplified representation that is easy to implement due to the fact it deals with sums rather than integral products.

$$\begin{aligned}\hat{\mathbf{a}}_{LS} = \underset{\hat{\mathbf{a}}}{\operatorname{argmin}} \sum_{d=1}^{D} \Bigg[\frac{1}{K_d^2} \sum_{k\in d}^{K_d} \sum_{k^*\in d}^{K_d} \left(c_{d,k,k^*}\right) \\ + \frac{1}{N^2} \sum_{n}^{N} f_n(\hat{\mathbf{a}}) \frac{1}{J_{d,n}} \sum_{n^*}^{N} f_{n^*}(\hat{\mathbf{a}}) \frac{1}{J_{d,n^*}} \sum_{j}^{J_{d,n}} \sum_{j^*}^{J_{d,n^*}} \left(c_{d,nj,(nj)^*}\right) \\ - 2 \frac{1}{K_d N} \sum_{k\in d}^{K_d} \sum_{n}^{N} f_n(\hat{\mathbf{a}}) \frac{1}{J_{d,n}} \sum_{j}^{J_{d,n}} \left(c_{d,k,(nj)}\right) \Bigg]\end{aligned} \tag{5.45}$$

Next we calculate the gradient of the argument of Eq. 5.45 in order to use it in our optimization schemes. Let Q represent the terms within the argument. The gradient can be expressed as the following expression.

$$\nabla_{\hat{\mathbf{a}}} Q = \mathbf{J}^T \nabla_{\hat{\mathbf{f}}} Q \tag{5.46}$$

The explicit expressions for gradient of the argument as well as the Jacobian are shown below.

$$\begin{aligned}\left(\nabla_f Q\right)_n = \frac{1}{N^2 J_{dn}} \left[\sum_{n^*}^{N} \frac{f_{n^*}}{J_{dn^*}} \sum_{j} \sum_{j^*} \left(c_{d,nj,(nj)^*}\right) \right] - \\ 2 \frac{1}{K_d N J_{dn}} \sum_{k\in d}^{K_d} \sum_{j}^{J_{d,n}} \left(c_{d,k,(nj)}\right)\end{aligned} \tag{5.47}$$

$$\mathbf{J}(\hat{\mathbf{a}}) = \frac{\partial \mathbf{f}}{\partial \hat{\mathbf{a}}}(\hat{\mathbf{a}}) = \begin{bmatrix} \frac{\partial f_1}{\partial \hat{a}_1} & \cdots & \frac{\partial f_1}{\partial \hat{a}_L} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial \hat{a}_1} & \cdots & \frac{\partial f_N}{\partial \hat{a}_L} \end{bmatrix} \tag{5.48}$$

$$\mathbf{J}(\hat{\mathbf{a}}) = \frac{1}{\left(\sum_l \hat{a}_l N_{S_l}\right)^2} \begin{bmatrix} ((\sum_l \hat{a}_l N_{\hat{a}_l}) - \hat{a}_1 N_{\hat{a}_1}) & -\hat{a}_1 N_{\hat{a}_2} & \cdots \\ -\hat{a}_2 N_{\hat{a}_1} & ((\sum_l \hat{a}_l N_{\hat{a}_l}) - \hat{a}_2 N_{\hat{a}_2}) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{5.49}$$

5.3 Summary Statistics of PSF Calibration Data

Prior to estimating the activity curve from list-mode data, for each detector-voxel combination, the output of each individual PMT was histogrammed, and the mean $\boldsymbol{\mu}$ and standard deviation $\boldsymbol{\sigma}$ of the histogram was obtained. An example of the process to estimate $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ from the list-mode calibration data is shown in Fig. 5.8. Given the geometry of FastSPECT II, certain voxels will collect more counts on a particular detector if it is closer to the detector as indicated in Fig. 5.7. When simulating the PSF calibration data, the number of simulated events from each detector-voxel combination remained a percentage of the number of original events collected at each detector-voxel in the raw PSF calibration data. Knowledge of the sensitivity of the detector, voxels, $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ allowed for the list-mode data to be generated on the fly. An example of simulated PSF calibration is shown in Fig. 5.9.

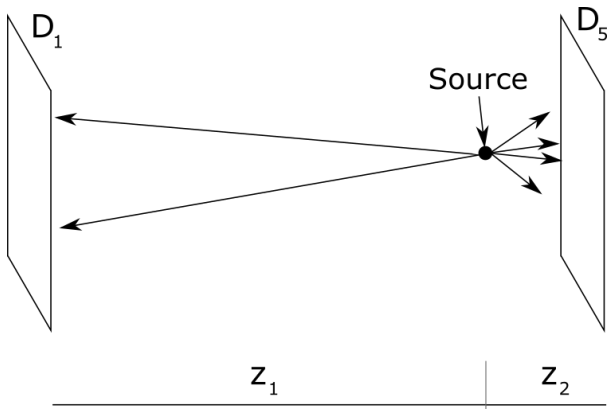


Figure 5.7: The source is closer $z_2 < z_1$ to detector D_5 resulting in more counts collected at D_5 as indicated by the number of arrows.

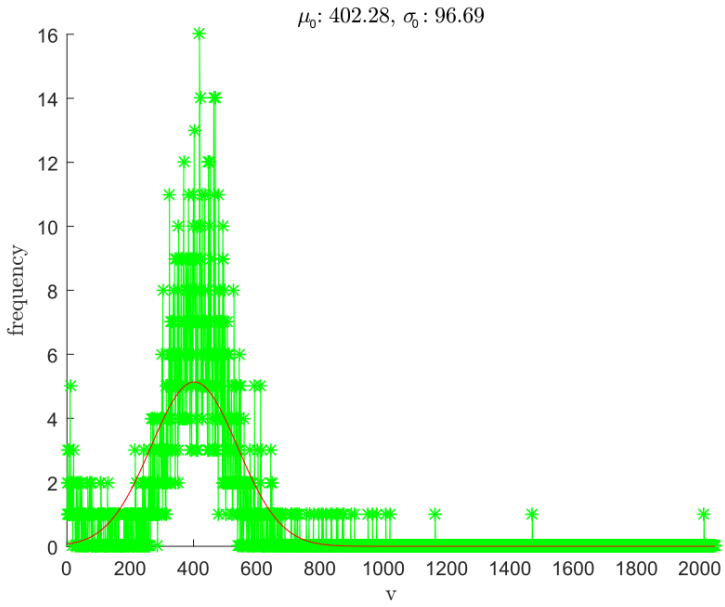


Figure 5.8: Histogrammed voltage output for detector $d = 1$, voxel $n = 1355$, and PMT $PMT = 9$. Using MLE, the mean μ_0 and standard deviation σ_0 of the data was found. The normal distribution corresponding to the estimated μ and σ is represented by the red curve.

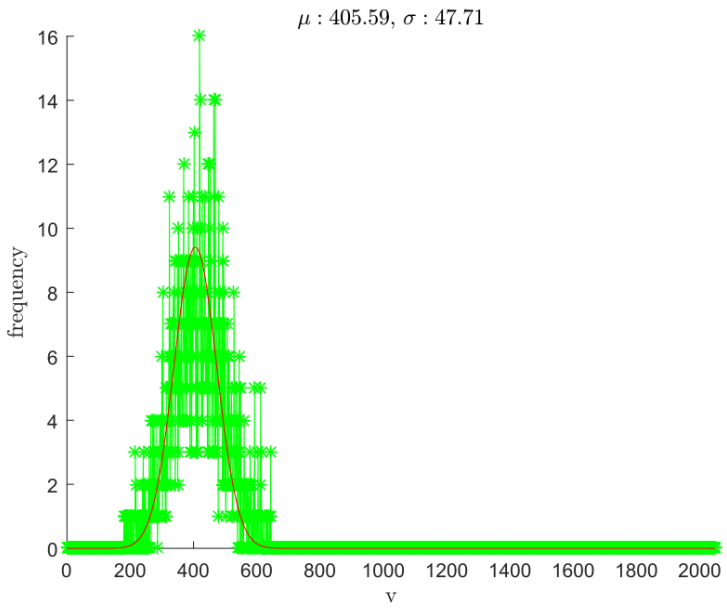


Figure 5.9: Given the μ and σ, PSF calibration data of any size (total number of counts) can be simulated.

5.4 Simulated Object List-Mode Data

Prior to generating the list-mode data for a particular object , the PSF calibration data needed to be split in two parts:

1. The top half of the list was used as the training data set, which is required to build the simulated list-mode object data.

2. The bottom half of the list is the testing data set, which was used to estimate the vector of compartmental activities.

With the raw unfiltered listmode data, a bootstrapping method was used since the PSF calibration data consists of a finite number of counts, which was tremendously

reduced after filtering. Hence, one of the advantages of using the mean and variance of each individual PMT for each detector-voxel combination, discussed in the previous section, to generate data is we are no longer limited by a finite number of counts. Instead a list of any size can be generated and the system performance as a function of Poisson noise can be analyzed since the noise is inversely proportional to the number of counts.

To simulate the object list-mode data using the PSF calibration data requires a few extra steps. The detector sensitivity s_d representing the probability that an event will be detected by detector d, and the voxel sensitivity s_n, representing the probability that an event recorded originated from voxel n, are needed. An example of the detector and voxel sensitivity for the PSF calibration data are shown in Fig. 5.10 and Fig. 5.11. Given that we want to build a list of $N_{events,total}$, the number of events recorded by each detector-voxel combination is given by

$$N_{events,d,n} = P_d P_n N_{events,list}. \tag{5.50}$$

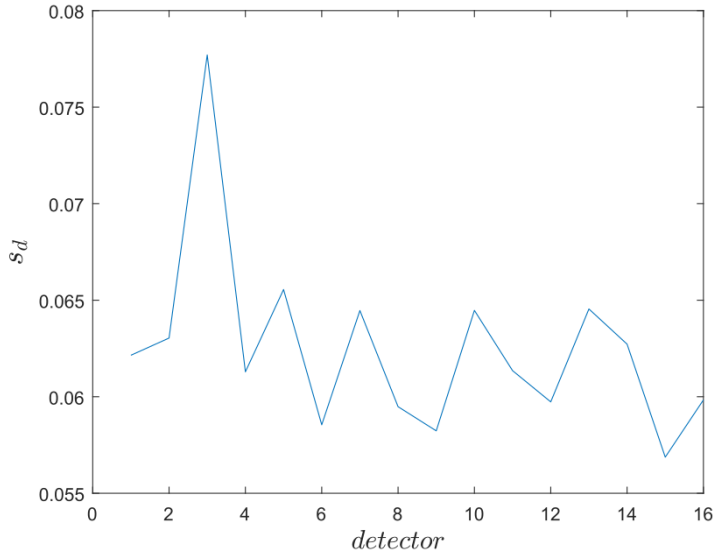


Figure 5.10: Sensitivity of each detector from FastSPECT II.

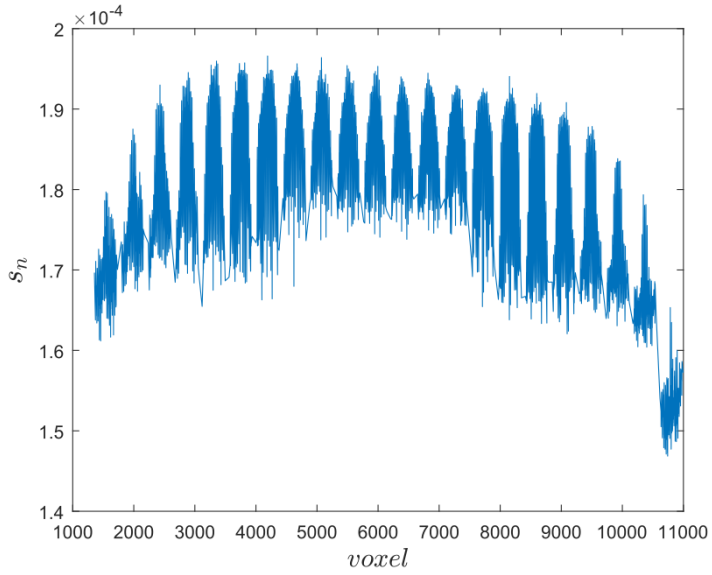


Figure 5.11: Sensitivity for all voxels within the field of view of FastSPECT II.

To build a simulated list-mode data set, multiple events were sampled from detectors and voxels that had the highest sensitivity. Once the object list-mode data is simulated, using KDE, a Gaussian kernel is fit to each of the outputs as shown in Eq. 5.26.

5.4.1 Bandwidth Selection

In multivariate density estimation, typically the bandwidth $\boldsymbol{\sigma}$ is a symmetric positive-definite matrix that controls smoothing of the PDF $p_{\boldsymbol{\sigma}}(\mathbf{v})$. The choice of the bandwidth is the main factor that determines how well the PDF $p(\mathbf{v})$ describes the data. As discussed in the end of Sec. 5.2.2, the bandwidth can either be obtained using the plug-in method via Silvermann's rule of thumb, or cross-validation methods, involving minimizing the integral-square error of the PDF. A symmetric positive definite bandwidth matrix will allow for flexibility in smoothing the data [103], and allow for a more accurate representation of $p(\mathbf{v})$. However, in our simulations we will assume the voltage output between PMT's is independent, and the bandwidth matrix will be a diagonal matrix $\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \cdots, \sigma_{N_{PMT}})$.

Bandwidth of Output Voltages

The underlying physics of the photon counting detector adhere to Poisson statistics. Since the voltage from each PMT is independent, and the mean and the variance of data are equivalent, $\sigma_i \approx v_i$ the bandwidth matrix for the j^{th} recorded event as a function of the detector, voxel and PMT can be represented as

$$\boldsymbol{\Sigma}_{d,n,j} = diag\left(\sigma_{d,n,j,1}, \sigma_{d,n,j,2}, \cdots, \sigma_{d,n,j,N_{PMT}}\right) \tag{5.51}$$

where the variance of each individual PMT according to Eq. 2.13 is written as

$$\sigma_{d,n,j,i} = \frac{v_{d,n,j,i}}{6}. \tag{5.52}$$

A diagonal matrix corresponding to a scaled versions of the PMT outputs is a more accurate representation of the PDF since the voltages with a lower value are given less weight than those of a higher value.

5.5 Results: Activity Curve from LM Data

An approximation for the temporal activity curve estimated from the list-mode data was obtained by finding the solution that minimized Eq. 5.45 as a function of the: 1. sample density, and the 2. noise which is proportional to the number of counts collected.

5.5.1 Experimental Time Sample Density Values

To generate the simulated object list mode data, the original activity curve was sampled $N_t = [4, 7, 9, 11, 13]$.

5.5.2 Experimental Noise Values

In the imaging system, the noise model used was Poisson, and the probability of recording $J_{d,n}$ events on detector d for a finite time interval T is given by

$$P(N) = \frac{\mu^{J_{d,n}} e^{-\mu}}{J_{d,n}} \tag{5.53}$$

The standard deviation of the Poisson distribution is given by

$$\sigma_{J_{d,n}} \approx \sqrt{J_{d,n}} \tag{5.54}$$

As the number of counts $J_{d,n}$ increases, the average of $j_{d,n}$ approaches the mean μ. By taking the ratio of the standard deviation of the Poisson and the number of counts collected, $\%\sigma_{J_d,n} = \sigma_{J_{d,n}}/\sqrt{J_{d,n}}$, as the number of counts increase, $\%\sigma_{J_d,n}$ or percent of standard deviation decreases, which implies a more accurate measurement [41].

For our object, a phantom consisting of $N_{obj} = 14$ voxels, a total of $N_{obj,list} = 279515$ events were collected for all voxels/detectors. Noise in the system was simulated by taking a percentage of the total counts, or $N_{noise} = N_{obj,list} \cdot [.05, .1, .15, .2, .3]$

5.5.3 Experimental Object

The object used in simulations was consisted of two compartment, of equal size and shape, with $N_{c_1} = 7$voxels and $N_{c_2} = 7$voxels.

5.5.4 Attempts to Obtain Activity Curves

Estimates for the activity curve were made at each discrete samples point by finding the best parameters that minimized Eq. 5.45. Using MATLAB, the function `fmincon()` was utilized which seeks to minimize a function of multiple variables. The advantage of using `fmincon()` is the bounds and linear constraints on the parameters we wish to estimate can be specified.

Local, Global Minima

The cost function used in the least squares algorithm Eq. 5.45, is a quadratic equation which theoretically has a global minima. Ideally the global and local minima coincide,

however in simulation, the global minima for each $\hat{\mathbf{a}}(t)$, was located at zero.

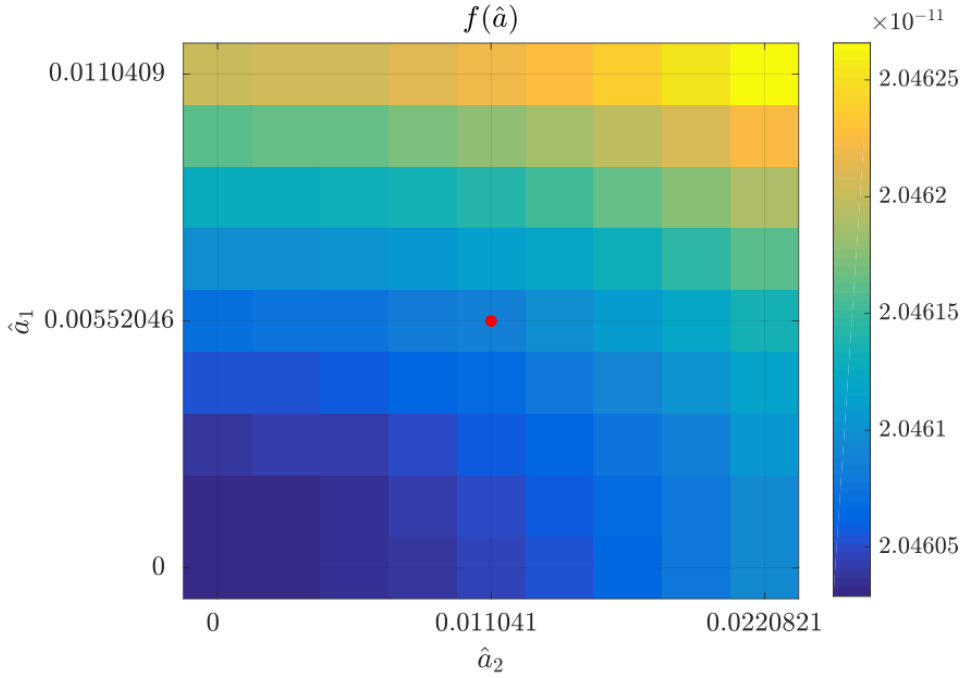


Figure 5.12: Contour plot of the quadratic cost function from Eq. 5.45. The global minima is located at $\hat{\mathbf{a}}(t) = [0, 0]$, whereas the true value for this particular instance, $\mathbf{a}(t) = [.0110, .0055]$ at $t = 4.6420sec$.

The estimate, $\hat{\mathbf{a}}(t)$, was dependent on the initial condition chosen. Multiple points around the global minimum as shown in Fig. 5.13, were tested to determine the locations of local minimums in relation to the true value of the compartmental activity at a fixed time point. Any initial points chosen were solutions to 5.45. In fig 5.12, the surface is approximately flat, as indicated by the scale to the right of the surface plot. In MATLAB, the change in the value of the object function during a step needed to be manually changed to $\Delta_{steptol} < 10^{-11}$, in order for the solver to function properly.

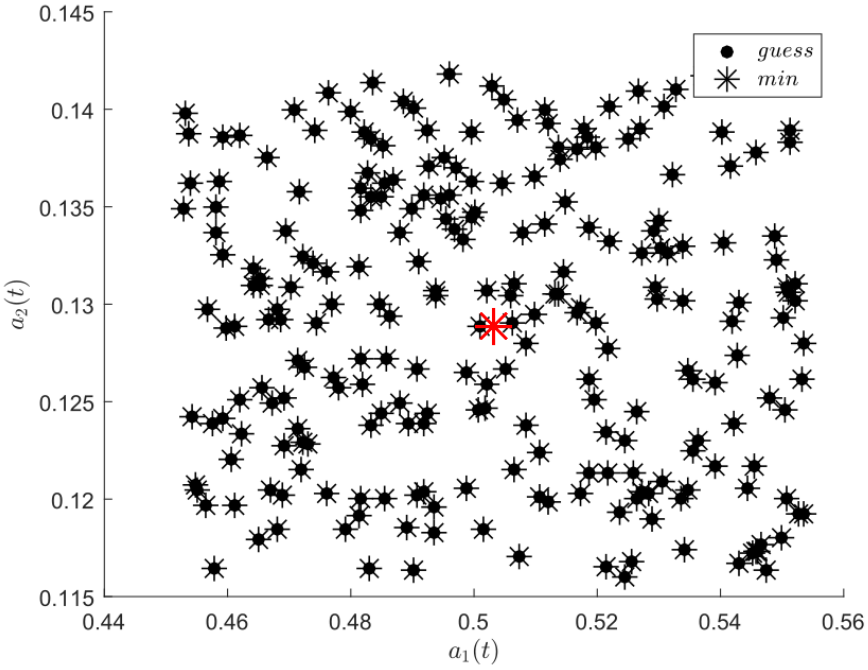


Figure 5.13: Multiple points were sampled near the global minimum to determine the locations of the local minimums $\hat{\mathbf{a}}(t)$ that satisfied Eq. 5.45. The global minimum is indicated by the red asterisk, and the local minimums are indicated by the black asterisk.

Recall in Sec. 5.2.1, Fig. 5.6, the estimates for the activity curves $\hat{\mathbf{a}}(t)$ using a binned approach were possible with less than a 10% error in relation to the actual value $\mathbf{a}(t)$. This was used to constrain the range of possible values $\hat{\mathbf{a}}(t)$ to a small region: $[(\alpha)\hat{\mathbf{a}}_l, (1+\alpha)\hat{\mathbf{a}}_l]$, where α was chosen to be 10%, as shown in Fig. 5.14 .

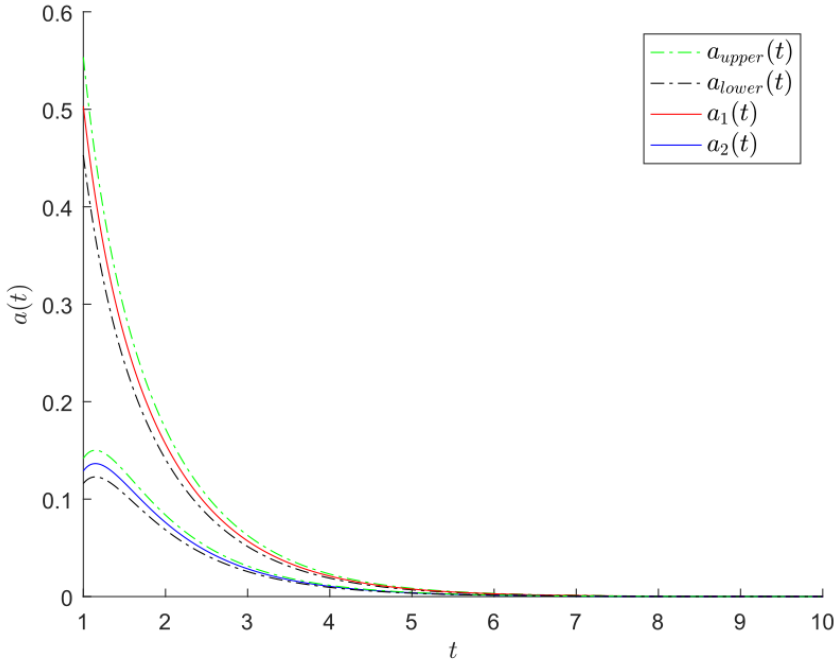
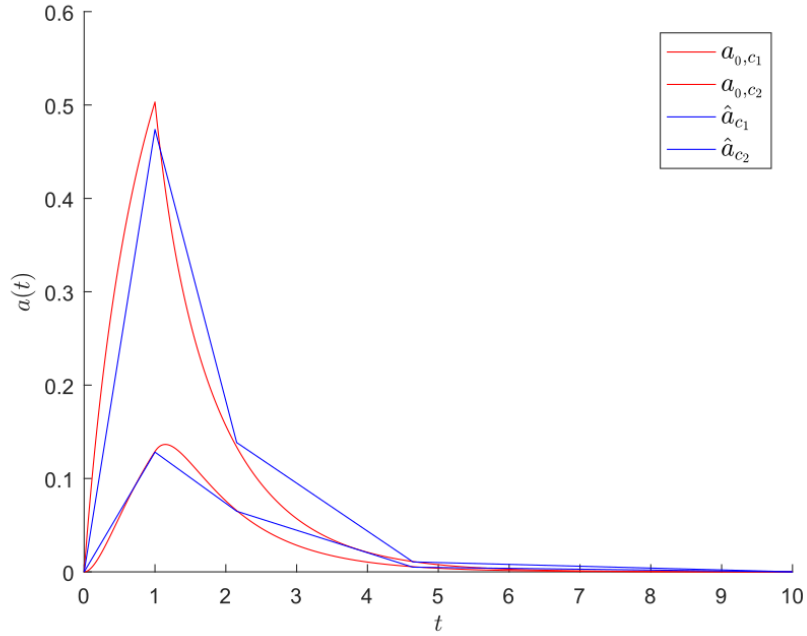
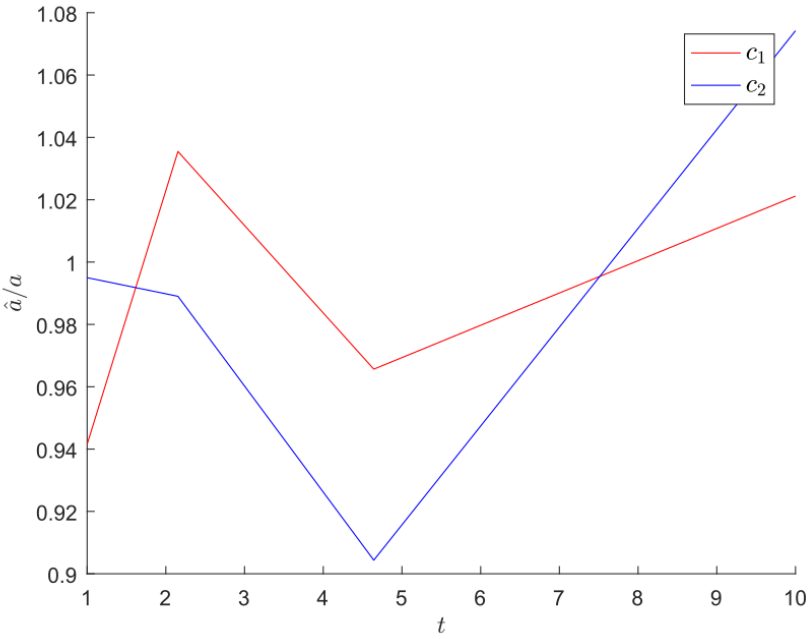


Figure 5.14: Upper (green) and lower bounds (black) for the possible values of $\hat{\mathbf{a}}_l(t)$.

The progression of Fig. 5.15 through Fig. 5.18, show the effect on the estimate of the activity curve, when changing the sample density and noise from the lowest values $N_t = 4$ and $N_{noise} = N_{obj,list}(.05)$ to the highest values, $N_t = 13$ and $N_{noise} = N_{obj,list}(.3)$. Also, in Fig. 5.15 through Fig. 5.18, the estimate for $\hat{\mathbf{a}}(t)$ that satisfied Eq. 5.45, were within the bounds $[(\alpha)\hat{\mathbf{a}}_l, (1+\alpha)\hat{\mathbf{a}}_l]$. However, in order to obtain reasonable estimates for $\hat{\mathbf{a}}(t)$, the range of possible values needed to be specified.

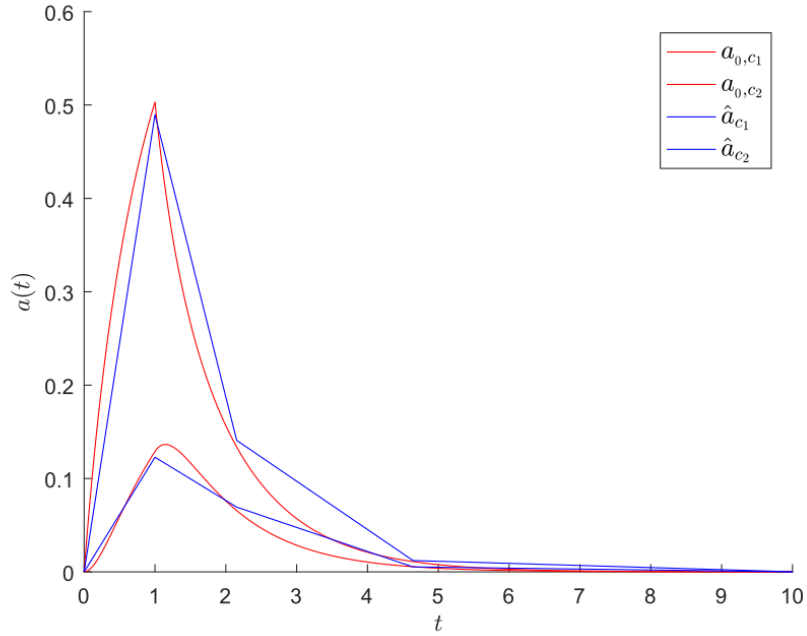


(a) $\mathbf{a}(t)$

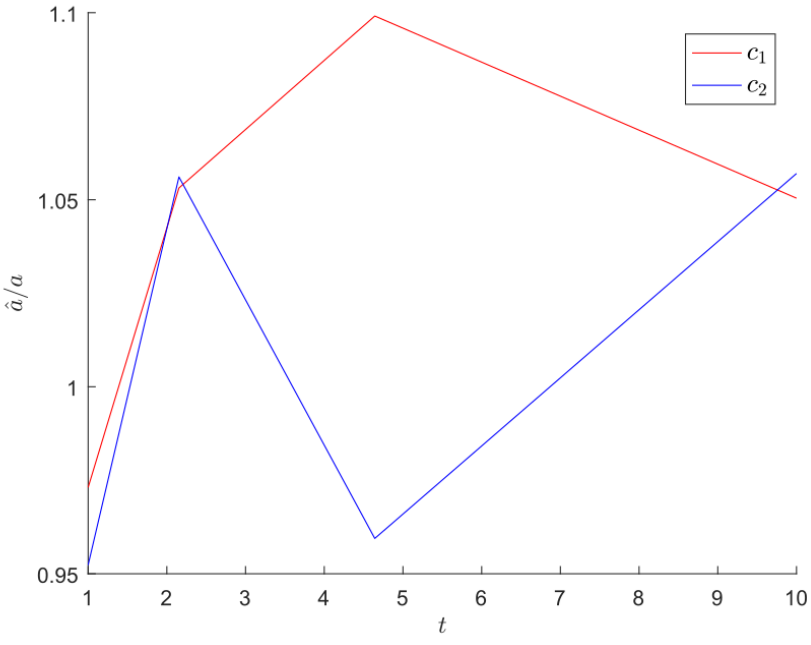


(b) $\frac{\hat{\mathbf{a}}(t)}{\mathbf{a}(t)}$

Figure 5.15: $N_t = 4$, $N_{noise} = N_{obj,list}(.05)$.

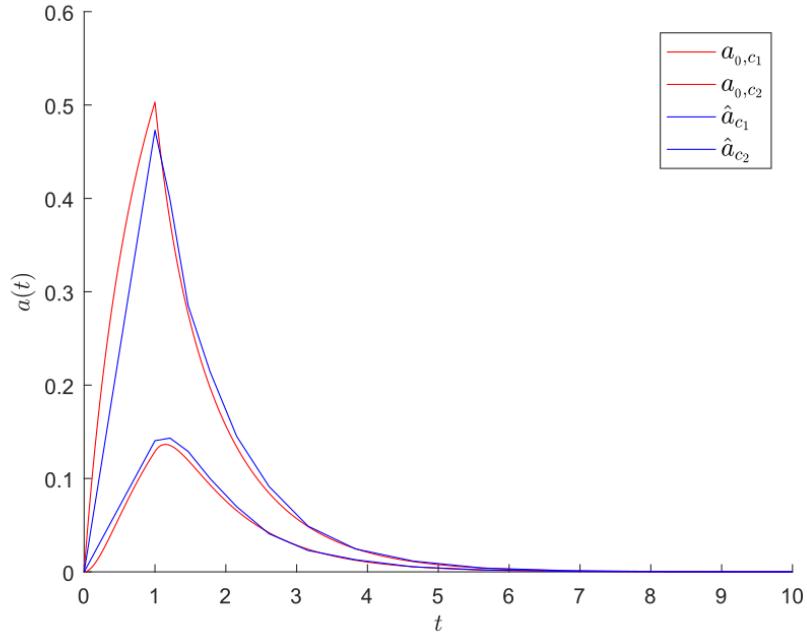


(a) $\mathbf{a}(t)$

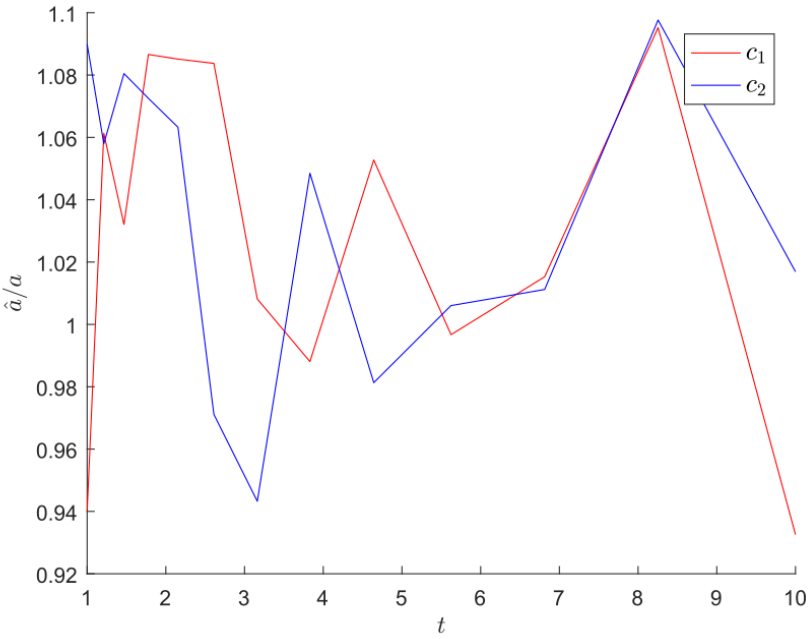


(b) $\frac{\hat{\mathbf{a}}(t)}{\mathbf{a}(t)}$

Figure 5.16: $N_t = 4$, $N_{noise} = N_{obj,list}(.3)$.

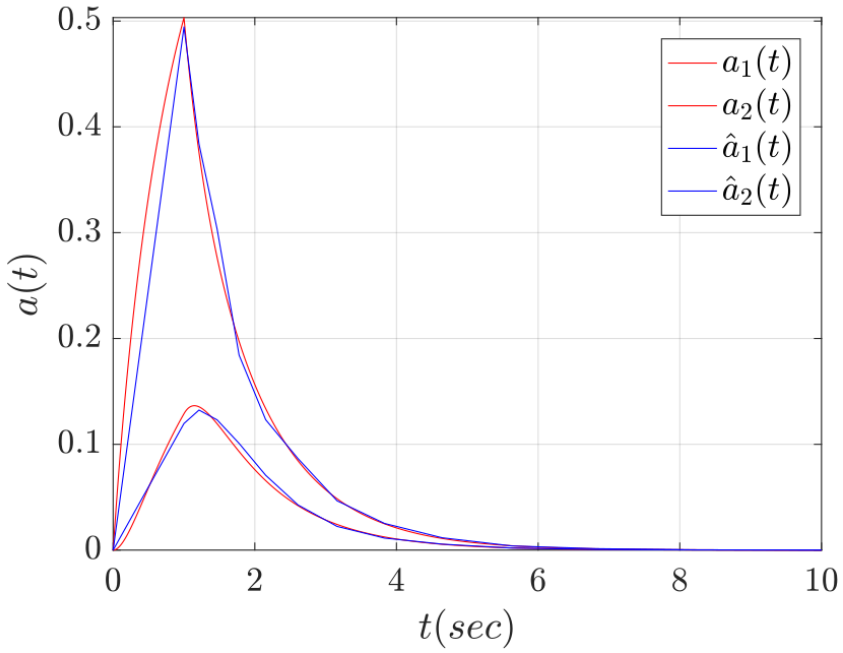


(a) $\mathbf{a}(t)$

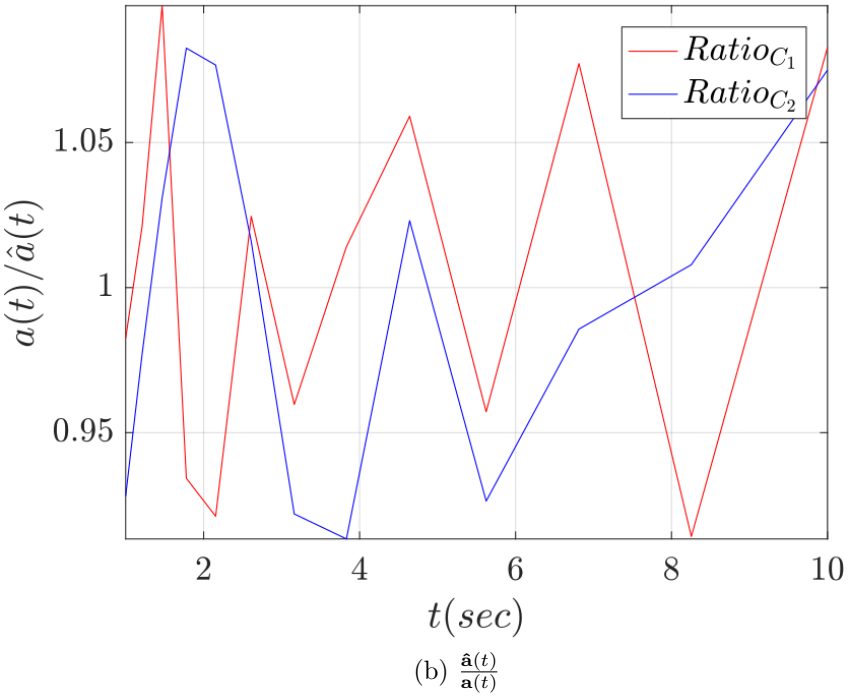


(b) $\frac{\hat{\mathbf{a}}(t)}{\mathbf{a}(t)}$

Figure 5.17: $N_t = 13$, $N_{noise} = N_{obj,list}(.05)$.



(a) $\mathbf{a}(t)$



(b) $\frac{\hat{\mathbf{a}}(t)}{\mathbf{a}(t)}$

Figure 5.18: $N_t = 13$, $N_{noise} = N_{obj,list}(.3)$.

CHAPTER 6

Estimation of Kinetic Parameters from Activity Curves

6.1 Estimating the Kinetic Parameters

The goal of pharmacokinetic modeling is to be able to quantify the kinetic parameters given data from the system using a similar process as shown in Fig. 1.4. In order to better model the uptake, a priori knowledge of the system is incorporated into the model. An example of a priori information that would increase the chances of success in estimating the kinetic parameters:

1. Knowledge of the number of unknowns in the model.
2. Knowledge of the spatial support region of each compartment.

6.1.1 Unknown Parameters

The number of unknown parameters depends on the type of model being used. Using the data from the experiment, the structural identifiability of the system, which deals with whether the unknown kinetic parameters can be uniquely determined, is explored [48] [42] [28]. Changing the dynamics of the system such as the compartments that have partial flow rates between them, or whether leakage occurs at a specific compartment changes the identifiability of the system [40] [68] [4] [27]. One of the most complex situation that will be encountered is a dense mixture model, consisting of L compartments within the system, with reversible transfer and elimination in every compartment, and input $\mathbf{I}(t)$ into one compartment only, is shown

in Fig. 6.1. In this particular example, the system is open since an external input is fed into the system, the kinetic matrix is full rank, and all elements of the kinetic matrix are non-zero. For a system consisting of L-compartment model, there are L^2 unknowns, and of those unknowns, $(2L-1)$ identifiable parameters and $(L-1)^2$ free parameters for a system with only one accessible compartment.

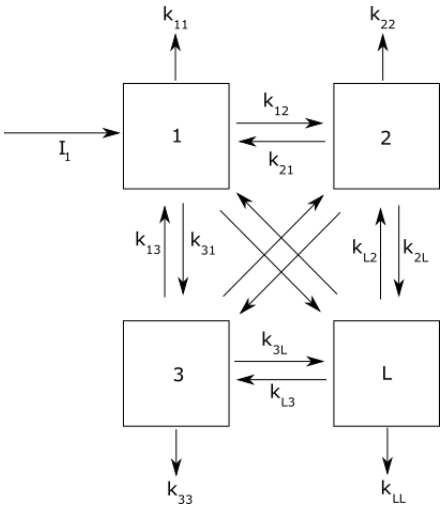


Figure 6.1: PK Model consisting of L compartments. Open dense mixture model: partial flow between each compartment, as well as out of each compartment indicated by $k_{i,j}$, with input $\mathbf{I}(t)$ into a single compartment.

Spatial Support Region

Knowledge of the spatial support region and the ability to estimate the kinetic parameters has been discussed by Clarkson [25] and Cobelli [27], as well as many of the references mentioned in Sec. 6.1.1. In an imaging context, using the support region knowledge is necessary to integrate the spatial component of the activity distribution to obtain only the temporal activity as done in Eq. 1.13. Furthermore, knowledge of the spatial support is imperative in reducing the number of unknowns, which in-

creases the number of identifiable parameters. For instance, with the Mammillary and Catenary models, which are two of the most commonly used models, shown in Fig. 6.2 and Fig. 6.3, the number of kinetic parameters is less than L^2.

Catenary Model

The Catenary model is such that the individual compartments are organized in a interconnecting chain. There exists only flow between adjacent compartments. The kinetic matrix corresponding to this particular model is a tridiagonal matrix consisting $(2L-1)$ non-zeros values, and $(3L-1)$ parameters to be estimated.

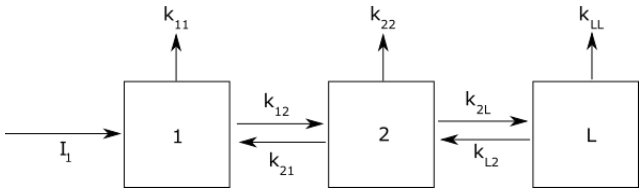


Figure 6.2: Catenary Model. Input $\mathbf{I}(t)$ into the first compartment. Partial flow indicated by $K_{i,j}$ only occurs between adjacent compartments.

Mammillary Model

The Mammillary model is such that the peripheral compartments are all connected to a central compartment, and a reversible rate of flow occurs between the central compartment and each of the individual compartment only. In this model, there exists $(L-1)$ unknown parameters to be estimated.

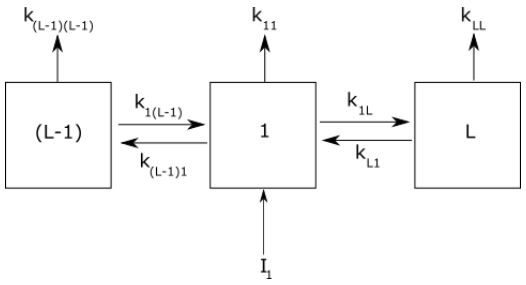


Figure 6.3: Mammillary Model. Input $\mathbf{I}(t)$ into the first compartment. Partial flow indicated by $K_{i,j}$ only occurs between the central compartment (compartment 1).

6.2 Estimating $\mathbf{K}$: Least Squares Approach

To find the unknown kinetic parameters, the methods outlined in [51] were used, involving fitting the data to the rate equations, and fitting the data to the sum of exponentials. A least squares approach was used as the primary fitting method since it is the method typically used in classical pharmacokinetic experiments to estimate the parameters of interest, and it is computationally easier to implement into computer code.

The goal behind the least squares approach is to find the parameters of a model function that best fit the measured data via a linear regression which involves minimizing the sum squared of the residual. In PK experiments the model is typically of the form $\mathbf{y}(t_i) = f(t_i, \hat{\mathbf{K}}) + \mathbf{r}_i$. Hence the residual would be represented as $\mathbf{r}_i = \mathbf{y}(t_i) - f(t_i, \hat{\mathbf{K}})$, and a least squares solution would be one that minimizes the $\mathbb{L}_2$ norm of $\mathbf{r}_i$.

$$\hat{\mathbf{K}}_{LS} = \underset{\mathbf{K}}{\operatorname{argmin}} \sum_{i=1}^{T} \|\mathbf{r}_i\|^2 \tag{6.1}$$

6.2.1 Residuals

In order to estimate the kinetic parameters from the measured activity curves, the following residuals were used

$$1.\ \mathbf{r}_a = \dot{\mathbf{a}}(t) - (\mathbf{K}\mathbf{a}(t) + \mathbf{I}(t))$$

$$2.\ \mathbf{r}_b = \mathbf{a}(t) - \left[\mathbf{K}\int_0^T (\mathbf{a}(s) + \mathbf{I}(s))\, ds\right]$$

$$3.\ \mathbf{r}_c = \mathbf{a}(t) - \mathbf{C}\left[\exp(\mathbf{\Lambda}(t))\right]\left[\text{ones}(1, L)\right]^T$$

where in residual 3, $\mathbf{C}$ is an $L \times L$ matrix, and $\mathbf{\Lambda}$ is $L \times L$ diagonal matrix.

Fitting the Rate Equations

In regards to the expression for residual 1 and residual 2, since time activity curves from each compartment in time are obtained experimentally, numerically finding the derivative or the integral could be done using various methods such as finite-difference approximation (for the derivative) or trapezoidal/Simpson's rule (for the integral). The potential drawback to using these methods comes from irregularities and discontinuities caused by noisy data, which can usually be overcome by utilizing smoothing techniques [20]. The advantage of using residual 1 and residual 2 is that the kinetic parameters can be estimated directly, as opposed to the method from residual 3, which requires an intermediate step, that will be discussed next.

Fitting to Sum of Exponentials

In regards to the expression from residual 3, as derived by Clarkson [26], to estimate the kinetic parameters for $t \geq T_{input}$, an equivalence relationship must be setup

between the $\mathbf{C}$ and the matrix of eigenvectors $\mathbf{V}$ to obtain $\mathbf{K}$. To the right of the input, the activity can be represented as

$$\mathbf{a}(t) = \exp(\mathbf{K}t)\mathbf{a}(t_{input}) \tag{6.2}$$

Recall from Eq. 1.17, $\mathbf{K} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$. Using this relationship, Eq. 6.2 can be written as

$$\mathbf{a}(t) = \mathbf{V}\exp(\mathbf{\Lambda}t)\mathbf{V}^{-1}\mathbf{a}(t_{input}) \tag{6.3}$$

From Eq. 6.2, $\mathbf{a}(t_{input})$ is unknown. By letting $\mathbf{V}^{-1}\mathbf{a}(t_{input}) = \mathbf{u}$, the we can represent $\mathbf{u}$ as

$$\mathbf{u} = \mathbf{D_u}[\text{ones}(1, L)]^T \tag{6.4}$$

where $\mathbf{D_u}$ is a diagonal vector with $\mathbf{u}$ along the diagonal. Since diagonal matrices commute, Eq. 6.3 can be written as

$$\mathbf{a}(t) = \mathbf{V}\mathbf{D_u}\exp(\mathbf{\Lambda}t)[\text{ones}(1, L)]^T \tag{6.5}$$

Comparing residual 3 with Eq. 6.5, we see that $\mathbf{C} = \mathbf{V}\mathbf{D_u}$, and an estimate for $\mathbf{K}$ can be obtained using the results from fitting.

$$\mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1} = \mathbf{V}\mathbf{D_u}\mathbf{\Lambda}\mathbf{V}\mathbf{D_u}^{-1} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} = \mathbf{K} \tag{6.6}$$

6.2.2 Estimate for Kinetic Matrix Using LM Data

The figure of merit used to determine how well the model was able to predict the kinetic parameters was the mean-squared error (MSE).

$$MSE = \frac{1}{N_{k_{i\neq 0}}} \sum_{i=1}^{N_{k_{i\neq 0}}} (\hat{k}_i - k_i)^2 \tag{6.7}$$

MSE Using Binned Data

In Fig. 6.4 through Fig. 6.6 the results of the MSE are shown as a function of noise and sample density when using a binned approach to estimate the kinetic parameters. The function `fmincon()` in MATLAB was used to estimate the Kinetic parameters. The constraints on $\mathbf{K}$ outlined in Sec. 1.4.3, were implemented into the algorithm. Finally, the initial condition for each element k_i of $\mathbf{K}$, was chosen such that $\hat{k_i} = [(\alpha)k_i, (1+\alpha)k_i]$, where $\alpha = 10\%$. The motivation for obtaining the kinetic parameters using a binned approach was to compare the performance of each of the residuals mentioned in Sec. 6.2.1, and to also see if similarities existed between estimating the kinetic parameters using a binned approach vs a list-mode approach.

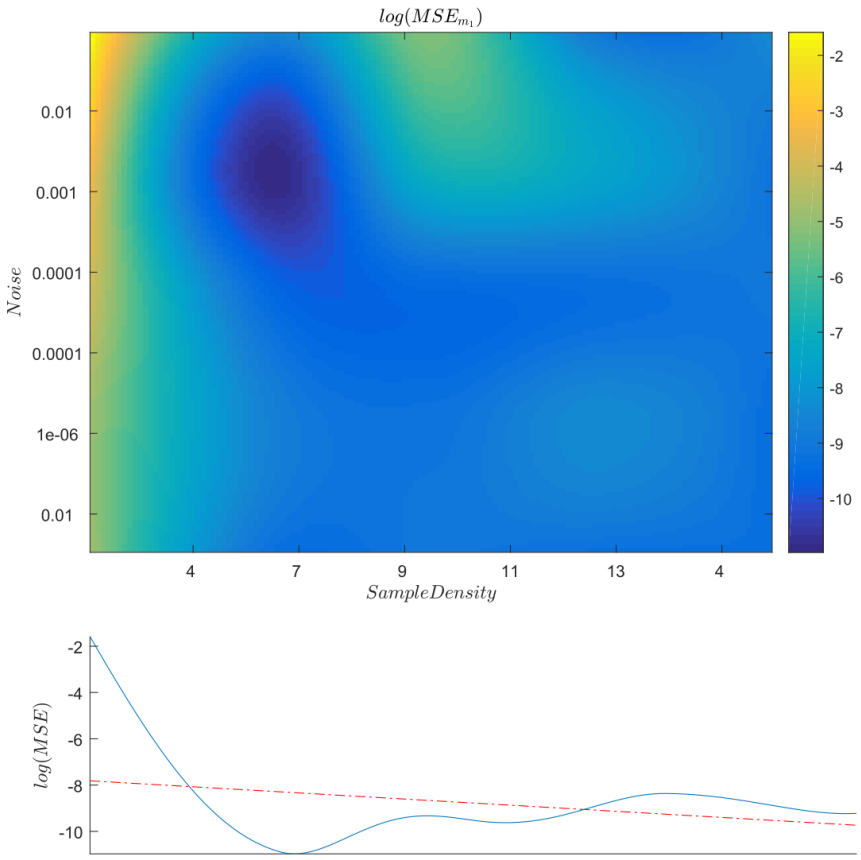


Figure 6.4: Residual 1: $log(MSE)$ of $\hat{\mathbf{K}}$.

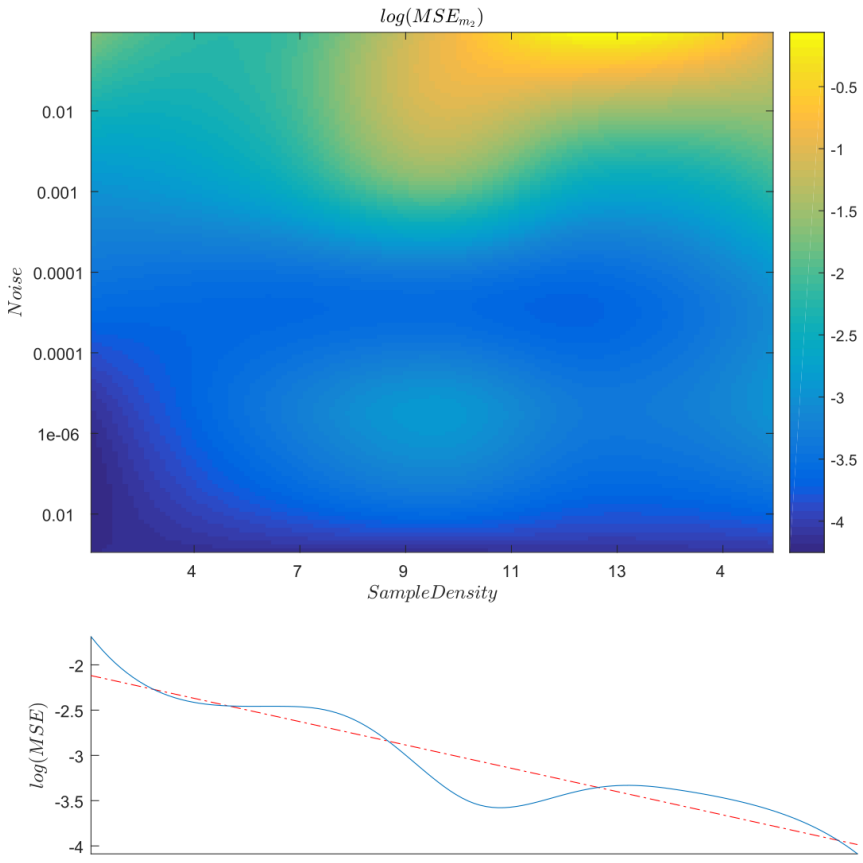


Figure 6.5: Residual 2: $log(MSE)$ of $\hat{\mathbf{K}}$.

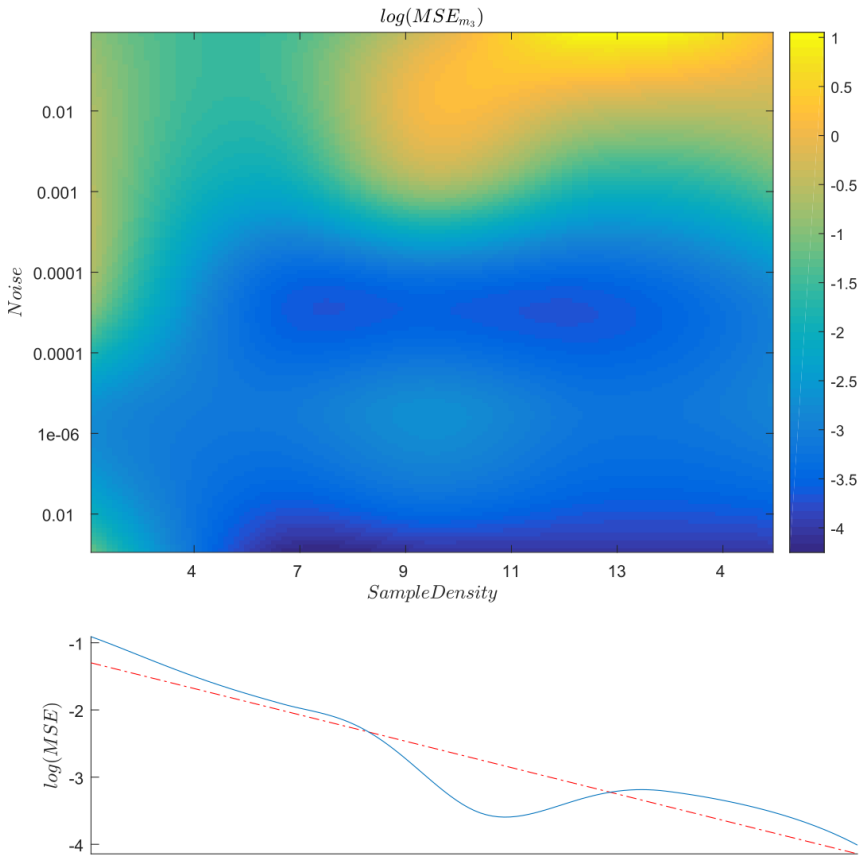


Figure 6.6: Residual 3: $log(MSE)$ of $\hat{\mathbf{K}}$.

MSE Using List-mode Data

In Fig. 6.7 through Fig. 6.9 the results of the MSE as a function of noise and sample density when using a list-mode approach to estimate the kinetic parameters. The function `fmincon()` in MATLAB was used to estimate the kinetic parameters. The

constraints on $\mathbf{K}$ outlined in Sec. 1.4.3, were implemented into the algorithm. Finally, the initial condition for each element k_i of $\mathbf{K}$, was chosen such that $\hat{k_i} = [(\alpha)k_i, (1+\alpha)k_i]$, where $\alpha = 10\%$.

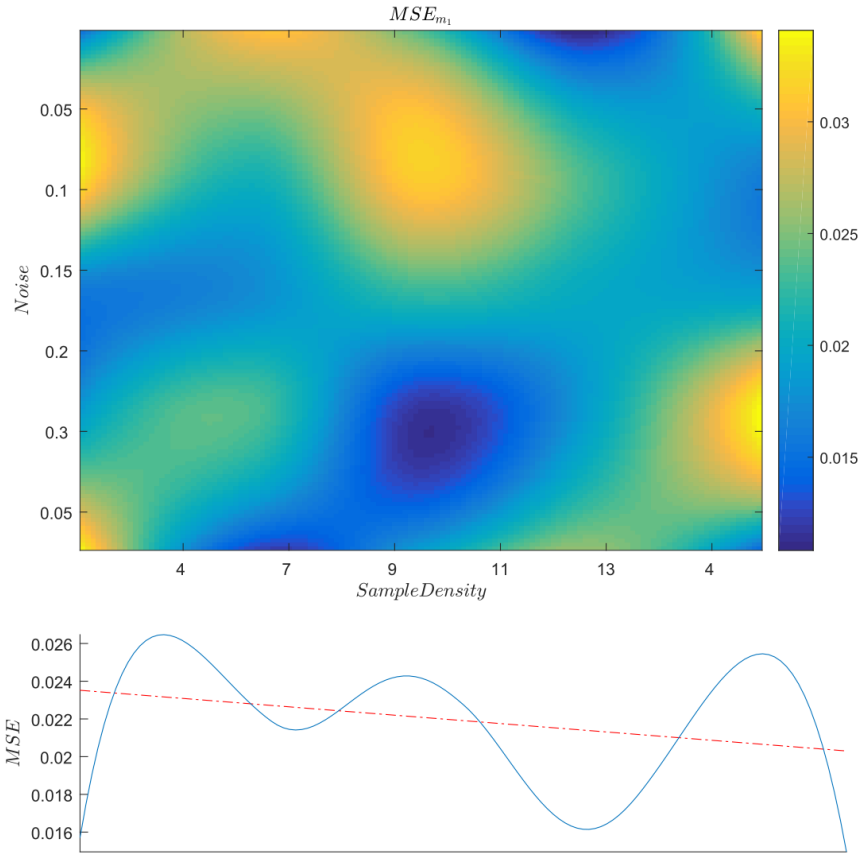


Figure 6.7: Residual 1: MSE of $\hat{\mathbf{K}}$.

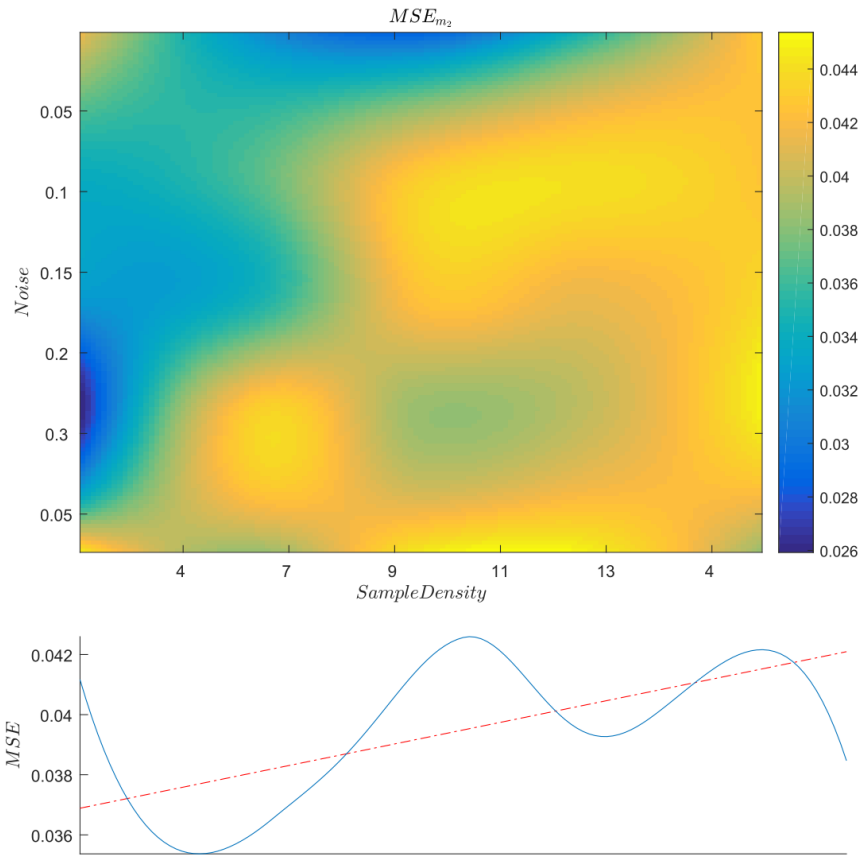


Figure 6.8: Residual 2: MSE of $\hat{\mathbf{K}}$.

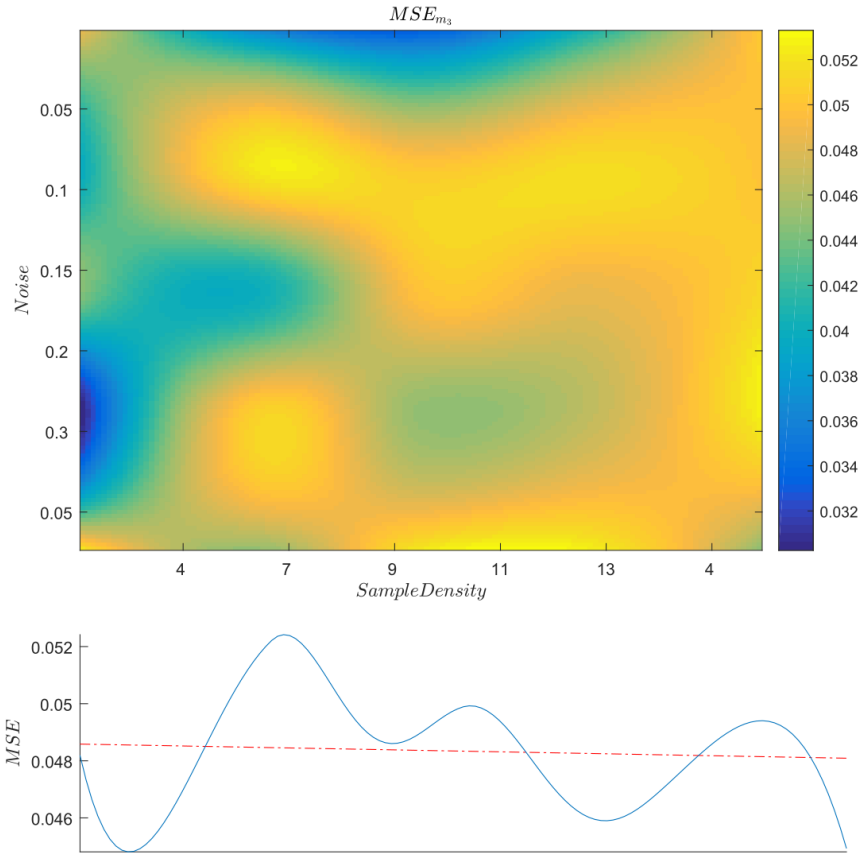


Figure 6.9: Residual 3: MSE of $\hat{\mathbf{K}}$.

CHAPTER 7

Conclusion

7.1 Conclusion

The format of this dissertation is as follows: Chapters 1-4 was introductory material and relevant theory, while Chapters 5 and 6 contain mostly original material. In Chapter 1, a brief overview of the field of classical pharmacokinetics was introduced. This included the relevant equations for modeling the uptake of the drug within a system, as well as an explanation of kinetic parameters, which are key in determining the partial flow rates of a drug between compartments. A key point of this chapter discusses how an imaging approach can be applied to solve classical PK problems. In Chapter 2-4, the process of imaging with gamma rays was discussed, beginning with the interaction of a gamma ray photon with a scintillation material, followed by an overview of different types of gamma detectors used to record the interaction, and finally the necessary elements to form an image. In Chapter 3, the object to image relationship was described mathematically. In Chapter 4, FastSPECT II was introduced, which is the 16 gamma camera imager used to collect the data, and the concept of list-mode data, which is the raw electronic outputs of each of the detector photo-multiplier tubes (PMT), was introduced.

7.2 Summary

The kinetic parameters discussed in Chapter 1 were estimated via an indirect approach from simulated list-mode data obtained from FastSPECT II PSF calibration data. Estimating the parameters via an indirect approach required a two part process: first the activity curves were obtained, and from the curves a best estimate for the kinetic parameters. In chapter 5, estimation of the activity curves from the list-mode data was discussed, which is the novel aspect of this project. In typical indirect approaches, the data from an imager is binned, and the parameters are obtained via segmentation of an object reconstruction. An advantage in using list-mode data was the amount of raw information available. Given the raw list-mode data, the task was to determine whether the kinetic parameters could be estimated. The algorithm to estimate the activity curves, via a least squares approach was presented, and implemented into code. In Chapter 6, three different cost functions were presented in order to estimate the kinetic parameters, and the success of each cost function was determined by using the mean squared error as the figure of merit.

7.3 Results

A surface plot Fig. 5.12 of the cost function from Eq. 5.45 for different values of the compartmental activity showed a global minima located at $[0, 0]$. In order to estimate the compartmental activity from simulated listmode data, the range of possible values had to be tightly constrained over a small region, within $\pm 10\%$ of the true value.

7.4 Future Work

7.4.1 Incorporate MDRF Calibration Data

The approach taken in this research was to disregard the mean detector response (MDRF) calibration data, and only use the PSF calibration data. As discussed in Chapter 2, the MDRF provides a means of measuring the location of the interaction event on each detector face by maximizing the likelihood that an event originated at a particular location. Calibrating the system using MDRF data allows for imperfections in the collection optics, as well as the detectors and system geometries to be accounted for and corrected during the reconstruction process. Furthermore, the advantages of using MDRF data is the ability to filter detected events via a likelihood windowing technique as discussed by Chen [21] or more recently by Chaix [18]. Since an information theoretic learning approach was taken to generate the probability distribution for each element in the list-mode data set, the extraneous events could have been a key source of errors in building the PDF used in simulations. Hence, by utilizing the MDRF, and building a parametric model, as opposed to a non-parametric model, the results could be significantly better.

7.4.2 Estimate **K** via a Direct Approach

Using a direct approach to estimate the kinetic parameters from raw list-mode would decrease the time needed to run the simulation, and potentially increase the accuracy of estimate since the kinetic parameters are estimated directly from the raw PMT data rather than an approximated activity curve.

7.4.3 Object

The phantom used in the experiment was a binary object consisting of only a small number of voxels, with a known support region. In future research, having the object boundary as a parameter that needed to be estimated as well as organs with texture would be a step closer to modeling an actual imaging system.

Processing Time

The original plan for the research was to utilize a MOBY digital phantom to create a realistic object, with organ compartments that filled the entire field of view, but processing large amount of data turned out to be out of reach of the CPU, and the entire project had to be scaled to a much smaller problem. In future work, since this research has shown that it is possible to estimate the activity from list-mode data, multiple compartments will be added, and the estimability of a dense mixture model incorporating textured boundaries, could prove to be useful. Since the total number of events recorded at each detector-voxel pair was different in size, the process could have benefited from a GPU feature called dynamic parallelism which allows individual threads on a device to launch more threads of different kernel resolution without involving the CPU. Most recently Caucci [15] has explored the use of dynamic parallelism in biomedical imaging applications.

REFERENCES

[1] Andersen (2005). Physiologically Based Pharmacokinetic Modeling: Science and Applications.

[2] Angelis, G. I., J. C. Matthews, F. A. Kotasidis, P. J. Markiewicz, W. R. Lionheart, and A. J. Reader (2014). Evaluation of a direct 4D reconstruction method using generalised linear least squares for estimating nonlinear micro-parametric maps. *Ann Nucl Med*, **28**(9), pp. 860–873. ISSN 0914-7187, 1864-6433. doi:10.1007/s12149-014-0881-2.

[3] Atkinson Jr., A. J. (2012). Chapter 2 - Clinical Pharmacokinetics A2 - Markey, Arthur J. AtkinsonShiew-Mei HuangJuan J.L. LertoraSanford P. In *Principles of Clinical Pharmacology (Third Edition)*, pp. 13–26. Academic Press. ISBN 978-0-12-385471-1.

[4] Audoly, S., L. D'Angio, M. P. Saccomani, and C. Cobelli (1998). Global identifiability of linear compartmental models-a computer algebra algorithm. *IEEE Transactions on Biomedical Engineering*, **45**(1), pp. 36–47. ISSN 0018-9294. doi:10.1109/10.650350.

[5] Barrett, H., W. Hunter, B. Miller, S. Moore, Y. Chen, and L. Furenlid (2009). Maximum-Likelihood Methods for Processing Signals From Gamma-Ray Detectors. *IEEE Transactions on Nuclear Science*, **56**(3), pp. 725–735. ISSN 0018-9499. doi:10.1109/TNS.2009.2015308.

[6] Barrett, H. H. and K. J. Myers (2004). *Wiley: Foundations of Image Science*. Wiley series in pure and applied optics. Hoboken, N.J.: Wiley-Interscience.

[7] Barrett, H. H., T. White, and L. C. Parra (1997). List-mode likelihood. *J. Opt. Soc. Am. A-Opt. Image Sci. Vis.*, **14**(11), pp. 2914–2923. ISSN 0740-3232. doi:10.1364/JOSAA.14.002914. WOS:A1997YD31900006.

[8] Beekman, F. J. and B. Vastenhouw (2004). Design and simulation of a high-resolution stationary SPECT system for small animals. *Phys. Med. Biol.*, **49**(19), p. 4579. ISSN 0031-9155. doi:10.1088/0031-9155/49/19/009.

[9] Bentourkia, M. (2005). Kinetic modeling of PET data without blood sampling. *IEEE Transactions on Nuclear Science*, **52**(3), pp. 697–702. ISSN 0018-9499. doi:10.1109/TNS.2005.851442.

[10] Bora, V. J. S. (2015). *Photon Statistics in Scintillation Crystals*. Ph.D. thesis.

[11] Bruyant, P. P. (2002). Analytic and Iterative Reconstruction Algorithms in SPECT. *J Nucl Med*, **43**(10), pp. 1343–1358. ISSN 0161-5505, 2159-662X.

[12] Cao, Z., C. A. Cardi, P. D. Acton, and M. L. Thakur (2007). Simulation of a pinhole-collimator insert for small animal PET using GATE. In *2007 IEEE Nuclear Science Symposium Conference Record*, volume 4, pp. 2933–2936. doi: 10.1109/NSSMIC.2007.4436748.

[13] Carruthers, T. (2016). Pharmacodynamics.

[14] Caucci, L. (2012). *Task Performance with List-Mode Data*. Ph.D. thesis.

[15] Caucci, L. and L. R. Furenlid (2015). GPU programming for biomedical imaging.

[16] Caucci, L., W. C. J. Hunter, L. R. Furenlid, and H. H. Barrett (2010). List-mode MLEM Image Reconstruction from 3D ML Position Estimates. *IEEE Nucl Sci Symp Conf Rec (1997)*, **2010**, pp. 2643–2647. ISSN 1095-7863. doi: 10.1109/NSSMIC.2010.5874269.

[17] Caucci, L., A. K. Jha, L. R. Furenlid, E. W. Clarkson, M. A. Kupinski, and H. H. Barrett (2013). Image Science with Photon-Processing Detectors. *IEEE Nucl Sci Symp Conf Rec (1997)*, **2013**. ISSN 1095-7863.

[18] Chaix, C. (2015). *AdaptiSPECT: a Preclinical Imaging System*. Ph.D. thesis.

[19] Chaix, C., S. Kovalsky, M. A. Kupinski, H. H. Barrett, and L. R. Furenlid (2014). Design and fabrication of a preclinical adaptive SPECT imaging system : AdaptiSPECT.

[20] Chartrand, R. (2011). Numerical Differentiation of Noisy, Nonsmooth Data.

[21] Chen, Y.-C. (2006). *System Calibration and Image Reconstruction for a New Small-Animal SPECT System*. Ph.D. thesis.

[22] Cheng, X., Z. Li, Z. Liu, N. Navab, S.-C. Huang, U. Keller, S. Ziegler, and K. Shi (2015). Direct Parametric Image Reconstruction in Reduced Parameter Space for Rapid Multi-Tracer PET Imaging. *IEEE Transactions on Medical Imaging*, **34**(7), pp. 1498–1512. ISSN 0278-0062. doi: 10.1109/TMI.2015.2403300.

[23] Chiao, P.-C., E. P. Ficaro, F. Dayanikli, W. L. Rogers, and M. Schwaiger (1994). Compartmental Analysis of Technetium-99m-Teboroxime Kinetics Employing Fast Dynamic SPECT at Rest and Stress. *J Nucl Med*, **35**(8), pp. 1265–1273. ISSN 0161-5505, 2159-662X.

[24] Chou, T.-C. (2006). Theoretical Basis, Experimental Design, and Computerized Simulation of Synergism and Antagonism in Drug Combination Studies. *Pharmacol Rev*, **58**(3), pp. 621–681. ISSN , 1521-0081 (Online). doi: 10.1124/pr.58.3.10.

[25] Clarkson, E. and M. A. Kupinski (2009). Global Compartmental Pharmacokinetic Models for Spatiotemporal SPECT and PET Imaging. *SIAM J Imaging Sci*, **2**(1), pp. 203–225. ISSN 1936-4954. doi:10.1137/080715226.

[26] Clarkson, E. W. (2015). BME 524. Kinetics: Identifiable Kinetic Parameters.

[27] Cobelli, C., D. Foster, and G. Toffolo (2002). *Tracer Kinetics in Biomedical Research : From Data to Model*. Springer Science & Business Media, New York. ISBN 978-0-306-46427-0.

[28] Delforge, J. (1977). The problem of structural identifiability of a linear compartmental system: Solved or not? *Mathematical Biosciences*, **36**(1), pp. 119–125. ISSN 0025-5564. doi:10.1016/0025-5564(77)90019-0.

[29] DePuey, E. G., R. Gadiraju, J. Clark, L. Thompson, F. Anstett, and S. C. Shwartz (2008). Ordered subset expectation maximization and wide beam reconstruction half-time gated myocardial perfusion SPECT functional imaging: A comparison to full-time filtered backprojection. *Journal of Nuclear Cardiology*, **15**(4), pp. 547–563. ISSN 1071-3581. doi:10.1016/j.nuclcard.2008.02.035.

[30] Dhillon, S. (2006). *Clinical Pharmacokinetics*. Pharmaceutical Press. ISBN 978-0-85369-571-4.

[31] Di Bella, E., G. Gullberg, A. Barclay, and R. Eisner (1997). Automated region selection for analysis of dynamic cardiac SPECT data. *IEEE Transactions on Nuclear Science*, **44**(3), pp. 1355–1361. ISSN 0018-9499. doi:10.1109/23.597013.

[32] Di Bella, E. V. R., S. G. Ross, D. J. Kadrmas, H. S. Khare, P. E. Christian, S. McJames, and G. T. Gullberg (2001). Compartmental Modeling of Technetium-99mLabeled Teboroxime with Dynamic Single-Photon Emission Computed Tomography. *Invest Radiol*, **36**(3), p. 178. ISSN 0020-9996.

[33] DiPiro, J. T. (2010). *Concepts in Clinical Pharmacokinetics (5th Edition)*. ASHP, Bethesda, MD, USA. ISBN 978-1-58528-338-5.

[34] Ditlevsen, S. and A. d. Gaetano (2014). Stochastic vs. deterministic uptake of dodecanedioic acid by isolated rat livers. *Bull. Math. Biol.*, **67**(3), pp. 547–561. ISSN 0092-8240, 1522-9602. doi:10.1016/j.bulm.2004.09.005.

[35] Donnet, S. and A. Samson (2013). A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic models. *Advanced Drug Delivery Reviews*, **65**(7), pp. 929–939. ISSN 0169-409X. doi:10.1016/j.addr.2013.03.005.

[36] Dresser, G. K., J. D. Spence, and D. G. Bailey (2012). Pharmacokinetic-Pharmacodynamic Consequences and Clinical Relevance of Cytochrome P450 3A4 Inhibition. *Clin Pharmacokinet*, **38**(1), pp. 41–57. ISSN 0312-5963, 1179-1926. doi:10.2165/00003088-200038010-00003.

[37] Durham, S. H., K. B. Garza, and L. S. Eiland (2016). Relationship between vancomycin dosage and serum trough vancomycin concentrations in pediatric patients with cystic fibrosis. *Am. J. Health-Syst. Pharm.*, **73**(13), pp. 969–974. ISSN 1079-2082. doi:10.2146/ajhp150605. WOS:000378661500007.

[38] Furenlid, L., J. Hesterman, and H. Barrett (2005). Real-time data acquisition and maximum-likelihood estimation for gamma cameras. In *Real Time Conference, 2005. 14th IEEE-NPSS*, pp. 4 pp.–. doi:10.1109/RTC.2005.1547506.

[39] Furenlid, L., D. Wilson, Y.-C. Chen, H. Kim, P. Pietraski, M. Crawford, and H. Barrett (2004). FastSPECT II: a second-generation high-resolution dynamic

SPECT imager. *IEEE Transactions on Nuclear Science*, **51**(3), pp. 631–635. ISSN 0018-9499. doi:10.1109/TNS.2004.830975.

[40] Garcia-Sevilla, F., M. Garcia-Moreno, M. Molina-Alarcon, M. J. Garcia-Meseguer, J. M. Villalba, E. Arribas, and R. Varon (2012). Linear compartmental systems. I. kinetic analysis and derivation of their optimized symbolic equations. *J Math Chem*, **50**(6), pp. 1598–1624. ISSN 0259-9791, 1572-8897. doi:10.1007/s10910-012-9991-z.

[41] Gedke, D. (2001). How Counting Statistics Controls Detection Limits and Peak Precision.

[42] Griffiths, D. (1979). Structural Identifiability for Compartmental Models. *Technometrics*, **21**(2), pp. 257–259. ISSN 0040-1706. doi:10.2307/1268525.

[43] Guidoum, A. C. (2013). Kernel Estimator and Bandwidth Selection for Density and its Derivatives.

[44] Gupta, N., P. M. Price, and E. O. Aboagye (2002). PET for in vivo pharmacokinetic and pharmacodynamic measurements. *European Journal of Cancer*, **38**(16), pp. 2094–2107. ISSN 0959-8049. doi:10.1016/S0959-8049(02)00413-6.

[45] Gutierrez-Osuna, R. (2016). CSCE six hundred and sixty six.

[46] Hansen, B. E. (2009). Lecture notes on nonparametrics. *Lecture notes*.

[47] Hebber, E., D. Oldenburg, T. Farnocombe, and A. Celler (1997). Direct estimation of dynamic parameters in SPECT tomography. *IEEE Transactions on Nuclear Science*, **44**(6), pp. 2425–2430. ISSN 0018-9499. doi:10.1109/23.656447.

[48] Hof, J. M. v. d. (1998). Structural identifiability of linear compartmental systems. *IEEE Transactions on Automatic Control*, **43**(6), pp. 800–818. ISSN 0018-9286. doi:10.1109/9.679020.

[49] Huesman, R. H., B. W. Reutter, G. L. Zeng, and G. T. Gullberg (1998). Kinetic parameter estimation from SPECT cone-beam projection measurements. *Phys. Med. Biol.*, **43**(4), p. 973. ISSN 0031-9155. doi:10.1088/0031-9155/43/4/024.

[50] Hwang, D. and G. L. Zeng (2006). One-step Backprojection Algorithm for Computed Tomography. In *2006 IEEE Nuclear Science Symposium Conference Record*, volume 6, pp. 3453–3457. doi:10.1109/NSSMIC.2006.353744.

[51] Jacquez, J. A. and P. Greif (1985). Numerical parameter identifiability and estimability: Integrating identifiability, estimability, and optimal sampling design. *Mathematical Biosciences*, **77**(1), pp. 201–227. ISSN 0025-5564. doi: 10.1016/0025-5564(85)90098-7.

[52] Jang, G. R., R. Z. Harris, and D. T. Lau (2001). Pharmacokinetics and its role in small molecule drug discovery research. *Med. Res. Rev.*, **21**(5), pp. 382–396. ISSN 0198-6325. doi:10.1002/med.1015. WOS:000170474700003.

[53] Jha, A., H. Barrett, E. Clarkson, L. Caucci, and M. A. Kupinski (2013). Analytic Methods for List-Mode Reconstruction.

[54] Jha, A. K., H. H. Barrett, E. C. Frey, E. Clarkson, L. Caucci, and M. A. Kupinski (2015). Singular value decomposition for photon-processing nuclear imaging systems and applications for reconstruction and computing null functions. *Phys. Med. Biol.*, **60**(18), p. 7359. ISSN 0031-9155. doi: 10.1088/0031-9155/60/18/7359.

[55] Kamasak, M., C. Bouman, E. Morris, and K. Sauer (2005). Direct reconstruction of kinetic parameter images from dynamic PET data. *IEEE Transactions on Medical Imaging*, **24**(5), pp. 636–650. ISSN 0278-0062. doi: 10.1109/TMI.2005.845317.

[56] Karakatsanis, N. A., M. E. Casey, M. A. Lodge, A. Rahmim, and H. Zaidi (2016). Whole-body direct 4D parametric PET imaging employing nested generalized Patlak expectationmaximization reconstruction. *Phys. Med. Biol.*, **61**(15), p. 5456. ISSN 0031-9155. doi:10.1088/0031-9155/61/15/5456.

[57] Krger-Thiemer, E. (1977). Pharmacokinetics. In Rossum, P. D. J. M. v. (ed.) *Kinetics of Drug Action*, number 47 in Handbuch der experimentellen Pharmakologie / Handbook of Experimental Pharmacology, pp. 63–123. Springer Berlin Heidelberg. ISBN 978-3-642-66539-4 978-3-642-66537-0.

[58] Kupinski, M. A. and H. H. Barrett (eds.) (2005). *Small-Animal Spect Imaging*. Springer US, Boston, MA. ISBN 978-0-387-25143-1 978-0-387-25294-0.

[59] Lecomte, R., E. Croteau, M. E. Gauthier, M. Archambault, A. Aliaga, J. Rousseau, J. Cadorette, J. D. Leroux, M. D. Lepage, F. Benard, and M. Bentourkia (2004). Cardiac PET imaging of blood flow, metabolism, and function in normal and infarcted rats. *IEEE Transactions on Nuclear Science*, **51**(3), pp. 696–704. ISSN 0018-9499. doi:10.1109/TNS.2004.829608.

[60] Lee, D.-Y., S. W. Sung, S. Y. Lee, and S. Park (2004). Combined DeterministicStochastic Approach for Pharmacokinetic Modeling. *Ind. Eng. Chem. Res.*, **43**(4), pp. 1133–1143. ISSN 0888-5885. doi:10.1021/ie0305364.

[61] Lehovich, A. (2005). *List-mode SPECT reconstruction using empirical likelihood*. Ph.D. thesis.

[62] Limber, M., M. Limber, A. Celler, J. Barney, and J. Borwein (1995). Direct reconstruction of functional parameters for dynamic SPECT. *IEEE Transactions on Nuclear Science*, **42**(4), pp. 1249–1256. ISSN 0018-9499. doi:10.1109/23.467872.

[63] Loeb, R., N. Navab, and S. I. Ziegler (2015). Direct Parametric Reconstruction Using Anatomical Regularization for Simultaneous PET/MRI Data. *IEEE Transactions on Medical Imaging*, **34**(11), pp. 2233–2247. ISSN 0278-0062. doi:10.1109/TMI.2015.2427777.

[64] Logan, J. (2000). Graphical analysis of PET data applied to reversible and irreversible tracers. *Nuclear Medicine and Biology*, **27**(7), pp. 661–670. ISSN 0969-8051. doi:10.1016/S0969-8051(00)00137-2.

[65] Madsen, M. T. (2007). Recent Advances in SPECT Imaging. *J Nucl Med*, **48**(4), pp. 661–673. ISSN 0161-5505, 2159-662X. doi:10.2967/jnumed.106.032680.

[66] Mager, H. and G. Gller (1995). Analysis of Pseudo-Profiles in Organ Pharmacokinetics and Toxicokinetics. *Statist. Med.*, **14**(9), pp. 1009–1024. ISSN 1097-0258. doi:10.1002/sim.4780140920.

[67] Matis, J. H. and T. E. Wehrly (1979). Stochastic Models of Compartmental Systems. *Biometrics*, **35**(1), pp. 199–220. ISSN 0006-341X. doi:10.2307/2529945.

[68] Meshkat, N. and S. Sullivant (2013). Identifiable reparametrizations of linear compartment models. *arXiv:1305.5768 [math]*. ArXiv: 1305.5768.

[69] Mullard, A. (2014). New drugs cost US$2.6 billion to develop. *Nat Rev Drug Discov*, **13**(12), pp. 877–877. ISSN 1474-1776. doi:10.1038/nrd4507.

[70] Narsky, I. and F. C. Porter (2013). Density Estimation. In *Statistical Analysis Techniques in Particle Physics*, pp. 89–120. Wiley-VCH Verlag GmbH & Co. KGaA. ISBN 978-3-527-67732-0.

[71] Nielsen, E. I. and L. E. Friberg (2013). Pharmacokinetic-Pharmacodynamic Modeling of Antibacterial Drugs. *Pharmacol. Rev.*, **65**(3), pp. 1053–1090. ISSN 0031-6997. doi:10.1124/pr.111.005769. WOS:000320986500007.

[72] Parra, L. and H. H. Barrett (1998). List-mode likelihood: EM algorithm and image quality estimation demonstrated on 2-D PET. *IEEE Trans. Med. Imaging*, **17**(2), pp. 228–235. ISSN 0278-0062. doi:10.1109/42.700734. WOS:000074840000009.

[73] Paul Segars, W. and B. M. W. Tsui (2009). MCAT to XCAT: The Evolution of 4-D Computerized Phantoms for Imaging Research. *Proc IEEE Inst Electr Electron Eng*, **97**(12), pp. 1954–1968. ISSN 0018-9219. doi:10.1109/JPROC.2009.2022417.

[74] Pedersen, A. R. (2002). Spurious results in therapeutic drug monitoring research. *Ther. Drug Monit.*, **24**(6), pp. 775–784. ISSN 0163-4356. doi:10.1097/00007691-200212000-00015. WOS:000179353700015.

[75] Pedersen, P. V. (1978). General treatment of linear pharmacokinetics. *J. Pharm. Sci.*, **67**(2), pp. 187–191. ISSN 1520-6017. doi:10.1002/jps.2600670216.

[76] Peterson, T. E. and L. R. Furenlid (2011). SPECT detectors: the Anger Camera and beyond. *Phys. Med. Biol.*, **56**(17), pp. R145–R182. ISSN 0031-9155. doi:10.1088/0031-9155/56/17/R01. WOS:000294786400001.

[77] Purves, D., G. J. Augustine, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O. McNamara, and S. M. Williams (2001). *Receptor Types*. Humana Press.

[78] Qi, J. and R. H. Huesman (2005). Effect of Errors in the System Matrix on MAP Image Reconstruction. *Phys Med Biol*, **50**(14), pp. 3297–3312. ISSN 0031-9155. doi:10.1088/0031-9155/50/14/007.

[79] Rahman, T., M. Tahtali, and M. R. Pickering (2014). An evaluation to design high performance pinhole array detector module for four head SPECT: a simulation study.

[80] Rakvongthai, Y., J. Ouyang, B. Guerin, Q. Li, N. M. Alpert, and G. El Fakhri (2013). Direct reconstruction of cardiac PET kinetic parametric images using a preconditioned conjugate gradient approach. *Med Phys*, **40**(10). ISSN 0094-2405. doi:10.1118/1.4819821.

[81] Reader, A. J. and J. Verhaeghe (2014). 4D image reconstruction for emission tomography. *Phys. Med. Biol.*, **59**(22), p. R371. ISSN 0031-9155. doi:10.1088/0031-9155/59/22/R371.

[82] Reisfeld, B. and A. Mayeno (2012). *Computational Toxicology - Volume I, Volume II*. Humana Press.

[83] Reutter, B., G. Gullberg, and R. Huesman (1998). Kinetic parameter estimation from attenuated SPECT projection measurements. *IEEE Transactions on Nuclear Science*, **45**(6), pp. 3007–3013. ISSN 0018-9499. doi:10.1109/23.737657.

[84] Schumacher, G. E. (1985). Choosing optimal sampling times for therapeutic drug monitoring. *Clin Pharm*, **4**(1), pp. 84–92. ISSN 0278-2677.

[85] Scott, D. W. and S. R. Sain (2005). 9-Multidimensional Density Estimation. *Handbook of statistics*, **24**, pp. 229–261.

[86] Shalizi, C. (2011). 36-402, Undergraduate Advanced Data Analysis (2011).

[87] Sheiner, L. B. and J.-L. Steimer (2000). Pharmacokinetic/Pharmacodynamic Modeling in Drug Development. *Annual Review of Pharmacology and Toxicology*, **40**(1), pp. 67–95. doi:10.1146/annurev.pharmtox.40.1.67.

[88] Shokouhi, S., S. D. Metzler, D. W. Wilson, and T. E. Peterson (2009). Multi-pinhole collimator design for small-object imaging with SiliSPECT: a high-resolution SPECT. *Phys. Med. Biol.*, **54**(2), p. 207. ISSN 0031-9155. doi: 10.1088/0031-9155/54/2/003.

[89] Simpson, J. A., S. Zaloumis, A. M. DeLivera, R. N. Price, and J. M. McCaw (2014). Making the Most of Clinical Data: Reviewing the Role of Pharmacokinetic-Pharmacodynamic Models of Anti-malarial Drugs. *AAPS J.*, **16**(5), pp. 962–974. ISSN 1550-7416. doi:10.1208/s12248-014-9647-y. WOS:000341434100008.

[90] Singh, S., M. K. Kalra, J. Hsieh, P. E. Licato, S. Do, H. H. Pien, and M. A. Blake (2010). Abdominal CT: Comparison of Adaptive Statistical Iterative and Filtered Back Projection Reconstruction Techniques. *Radiology*, **257**(2), pp. 373–383. ISSN 0033-8419. doi:10.1148/radiol.10092212.

[91] Smith, A. M., G. T. Gullberg, P. E. Christian, and F. L. Datz (1994). Kinetic Modeling of Teboroxime Using Dynamic SPECT Imaging of a Canine Model. *J Nucl Med*, **35**(3), pp. 484–495. ISSN 0161-5505, 2159-662X.

[92] Stangier, J., K. Rathgen, H. Stahle, and D. Mazur (2010). Influence of Renal Impairment on the Pharmacokinetics and Pharmacodynamics of Oral Dabigatran Etexilate An Open-Label, Parallel-Group, Single-Centre Study. *Clin. Pharmacokinet.*, **49**(4), pp. 259–268. ISSN 0312-5963. WOS:000276326300004.

[93] Thron, C. D. (1974). Linearity and Superposition in Pharmacokinetics. *Pharmacol Rev*, **26**(1), pp. 3–31. ISSN , 1521-0081 (Online).

[94] Thurber, G. M. and R. Weissleder (2011). A Systems Approach for Tumor Pharmacokinetics. *PLOS ONE*, **6**(9), p. e24696. ISSN 1932-6203. doi:10.1371/journal.pone.0024696.

[95] Tozer, T. N. (1981). Concepts basic to pharmacokinetics. *Pharmacology & Therapeutics*, **12**(1), pp. 109–131. ISSN 0163-7258. doi:10.1016/0163-7258(81)90077-2.

[96] Tuntland, T., B. Ethell, T. Kosaka, F. Blasco, R. X. Zang, M. Jain, T. Gould, and K. Hoffmaster (2014). Implementation of pharmacokinetic and pharmacodynamic strategies in early research phases of drug discovery and development

at Novartis Institute of Biomedical Research. *Front Pharmacol*, **5**. ISSN 1663-9812. doi:10.3389/fphar.2014.00174.

[97] VengPedersen, P., J. A. Widness, L. M. Pereira, C. Peters, R. L. Schmidt, and L. S. Lowe (1995). Kinetic Evaluation of Nonlinear Drug Elimination by a Disposition Decomposition Analysis. Application to the Analysis of the Nonlinear Elimination Kinetics of Erythropoietin in Adult Humans. *Journal of Pharmaceutical Sciences*, **84**(6), pp. 760–767. ISSN 0022-3549. doi:10.1002/jps.2600840619.

[98] Wagner, J. (1968). Pharmacokinetics. *Annual Review of Pharmacology*, **8**, pp. 67–&. doi:10.1146/annurev.pa.08.040168.000435. WOS:A1968A934200004.

[99] Wagner, J. G. (1975). Do you need a pharmacokinetic model, and, if so, which one? *Journal of Pharmacokinetics and Biopharmaceutics*, **3**(6), pp. 457–478. ISSN 0090-466X, 1573-8744. doi:10.1007/BF01059477.

[100] Wagner, J. G. (1994). Some Contributions and Guidance for Future Graduate Students. *Ann Pharmacother*, **28**(7-8), pp. 957–960. ISSN 1060-0280, 1542-6270. doi:10.1177/106002809402800722.

[101] Walker, D. K. (2004). The use of pharmacokinetic and pharmacodynamic data in the assessment of drug safety in early drug development. *Br J Clin Pharmacol*, **58**(6), pp. 601–608. ISSN 0306-5251. doi:10.1111/j.1365-2125.2004.02194.x.

[102] Walrand, S., F. Jamar, and S. Pauwels (2009). Improved solution for ill-posed linear systems using a constrained optimization ruled by a penalty: evaluation in nuclear medicine tomography. *Inverse Problems*, **25**(11), p. 115004. ISSN 0266-5611. doi:10.1088/0266-5611/25/11/115004.

[103] Wand, M. P. and M. C. Jones (1993). Comparison of Smoothing Parameterizations in Bivariate Kernel Density Estimation. *Journal of the American Statistical Association*, **88**(422), pp. 520–528. ISSN 0162-1459. doi:10.1080/01621459.1993.10476303.

[104] Wilks, M. Q., S. M. Knowles, A. M. Wu, and S.-C. Huang (2014). Improved Modeling of In Vivo Kinetics of Slowly Diffusing Radiotracers for Tumor Imag-

ing. *J Nucl Med*, **55**(9), pp. 1539–1544. ISSN 0161-5505, 2159-662X. doi: 10.2967/jnumed.114.140038.

[105] Williams, P. J. and E. I. Ette (2000). The Role of Population Pharmacokinetics in Drug Development in Light of the Food and Drug Administration??s ???Guidance for Industry: Population Pharmacokinetics???:. *Clinical Pharmacokinetics*, **39**(6), pp. 385–395. ISSN 0312-5963. doi:10.2165/00003088-200039060-00001.

[106] Wilson, D. W. and H. H. Barrett (2002). The effects of incorrect modeling on noise and resolution properties of ML-EM images. *IEEE Transactions on Nuclear Science*, **49**(3), pp. 768–773. ISSN 0018-9499. doi:10.1109/TNS.2002.1039561.

[107] Zakavi, S. R., A. Zonoozi, V. D. Kakhki, M. Hajizadeh, M. Momennezhad, and K. Ariana (2006). Image Reconstruction Using Filtered Backprojection and Iterative Method: Effect on Motion Artifacts in Myocardial Perfusion SPECT. *J. Nucl. Med. Technol.*, **34**(4), pp. 220–223. ISSN 0091-4916, 1535-5675.

[108] Zeng, G., G. Gullberg, and R. Huesman (1995). Using linear time-invariant system theory to estimate kinetic parameters directly from projection measurements. *IEEE Transactions on Nuclear Science*, **42**(6), pp. 2339–2346. ISSN 0018-9499. doi:10.1109/23.489438.

[109] Zeng, G. L., G. T. Gullberg, and D. J. Kadrmas (2010). Closed-form kinetic parameter estimation solution to the truncated data problem. *Phys. Med. Biol.*, **55**(24), p. 7453. ISSN 0031-9155. doi:10.1088/0031-9155/55/24/005.

[110] Zucchini, W., A. Berzel, and O. Nenadic (2003). Applied smoothing techniques. *Part I: Kernel Density Estimation*, **15**.