Signal separation of dual tracer PET scans

Bachelor's Project Applied Mathematics

July 7, 2023

Student Number:	S4057384
Student:	Maria Ewelina Wasiljew
First Supervisor:	Prof Dr. Juan Peypouquet
Second Supervisor:	Prof Dr. Marco Grzegorczyk

Faculty of Science and Engineering
University of Groningen

Abstract

In this Bachelor's Thesis the topic of signal separation of dual tracer PET scans is discussed in detail. A dual tracer PET scan is a nuclear imaging technique using two radiotracers to depict different physiological processes. All radiotracers used for PET scans emit the same level of energy, as such the scanner cannot distinguish between the incoming signals. For that reason, the use of a signal separation procedure is required. In this paper an Extrapolation Method of signal separation is considered. A full reference tissue model is used to describe the kinetics of a radiotracer. The parameters of the model are approximated with the use of a Gauss-Newton iterative method. Through the work presented in this thesis, the method shown has the necessary conditions to successfully separate signals of a dual tracer PET scan.

Contents

1 Introduction

Medical imagining is a crucial element of modern healthcare. The ability to see the inside of a patient's body, without the need to cut one open, revolutionised medicine. It allowed doctors to form more accurate diagnoses, spot diseases at an earlier stage and track the progress of treatment [1]. There are many methods used in this field, such as an X-ray, MRI or a PET scan. This paper is devoted to modern research on the last-named of these methods.

The acronym of a PET scan stands for *Positron Emitting Tomography*. This nuclear imaging method creates a 3D image of the inside of a body. Depending on the choice of a radiotracer, a drug enhanced with a small amount of a radioactive substance [11] , it can depict different physiological processes, such as metabolic activity or blood flow. Most often it is used within the fields of oncology, neurology and cardiology [20].

Over the past years this method has been improved, from the inclusion of motion correctness to the betterment of image reconstruction. There are many more on-going studies about further advancements of this method [26]. The one of interest for this paper is the use of dual tracers for a PET scan.

The dual tracer scenario implies the use of two radiotracers for a singular PET scan. Currently, if more than one physiological process needs to be studied for a diagnosis, a patient has to go through separate PET scan sessions [26]. Often, they cannot take place on the same day, as the first radiotracer has to stop being radioactive, so that it does not interfere with the results of the next scan. This prolongs the time needed to form a proper diagnosis, increases the medical costs and causes unnecessary stress for the patient. The use of two radiotracers during a single PET scan would solve these problems.

Although it creates an additional challenge. The machine would not be able to differentiate between the incoming signals, as all of the radiotracers used for PET scans produce the same level of energy (511 keV) [7]. Therefore, in order to recover two separate images from one scan, a signal separation method has to be implemented.

In this paper, an Extrapolation Method of signal separation is considered. It was introduced by Koeppe et al. in 2004 as the first method that did not require arterial sampling from a patient [13]. Instead, this approach uses a reference tissue model to predict the behaviour of a radiotracer inside a body.

However, in order to attempt the dual tracer problem, one cannot rely on just mathematical theories. The problem originates from the medical field, which means that it is defined through other terminology, and different components are considered important. Therefore, a basic understanding of the problem through the view of the other discipline is required, before the subject can be reintroduced in mathematical terms.

Another factor that has to be taken into account is the practicality of the results obtained from the application of mathematics. Foremost, it has to be of use to the medical personnel. As such, the results should be outputted in a format that is easy to understand for those with no mathematical background. In this case, the final result of the signal separation procedure should be outputted as two images depicting the activity of each radiotracer.

2 Single Tracer PET scan

From a patient's perspective, a PET scan is a relatively simple procedure. Usually it takes around two hours before a scan is completed. Firstly, the injection of a radiotracer takes place. After that the patient has to wait up to 90 minutes, so that the radiotracer can travel through the body and interact with specific tissues. Only then a scan is performed. The results are then studied by doctors, in order to form a medical diagnosis [2].

On the other hand, from a scientific perspective PET scans are far from simple. To understand this imaging method, one should study different components of it, such as the bio-chemical processes taking place inside the patient's body or technical limitations of the scanner machine. For that reason, most important elements of a PET scan are discussed in this section.

The injection of a radiotracer is a crucial part of the PET scan procedure. Different tracers interact with different types of tissues. For example, F-18 fluorodeoxyglucose collects in areas with a high metabolic activity, such as cancer cells [4], while a different tracer, O-15 water, accurately depicts blood flow [10].

Another difference between radiotracers is their decay rate. It can be seen by comparing their half-lives, being the time needed for half of the radioactive substance to decay. For example, the half-life of a tracer labelled with F-18 is a period of 109 minutes, while for one labelled with O-15 it is only 2.1 minutes [7]. This is an important factor, not only because it decides on the length of a PET scan procedure, but also due to technical difficulties of producing and storing radiotracers.

Due to the significance of radioactive decay, let us take a closer look at the molecular process behind it. Radiotracers used for PET scans are those that go through a beta radioactive decay. As suggested by the meaning of the PET acronym, this means that a positron particle has to be emitted.

The emission comes from a nucleus, which is the center of an atom. The nucleus consists of protons, positively charged particles, and neutrons, particles with neutral charge. When one of the protons turns into a neutron two particles are emitted, one with positive charge called a positron (e^+) and one with neutral charge called an electron neutrino (ν) [7] . This process is depicted in Figure 1.

It can also be described by a chemical equation, where $^A_Z X$ denotes a nucleus with mass number A and atomic number Z. The mass number is the combined amount of neutrons and protons in the nucleus, while the atomic number corresponds to quantity of just neutrons. Finally, X denotes a general chemical symbol. When dealing with a more specific example, it should be replaced by the periodic notation of a considered substance.
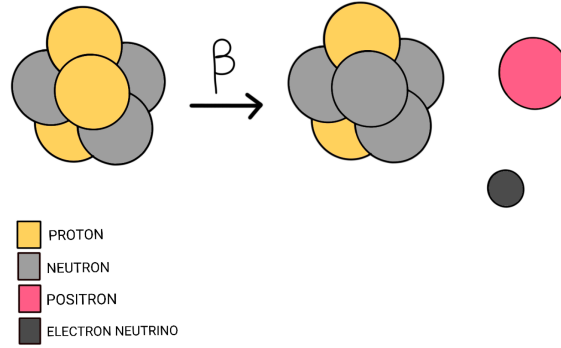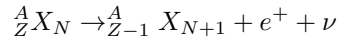
Figure 1: Beta Decay

$$^A_Z X_N \to^A_{Z-1} X_{N+1} + e^+ + \nu$$

Surprisingly, even though a PET scan is named after a positron emission, the scanner machine does not detect that particle. The emitted positron travels around 2-3mm from its source, before it collides with an electron. At the impact, two gamma waves (γ) are created, which travel away from each other at a 180 degree angle [7]. This has been depicted in Figure 2 and described by the following chemical equation. As previously, e^+ denotes a positron, and e^- an electron.
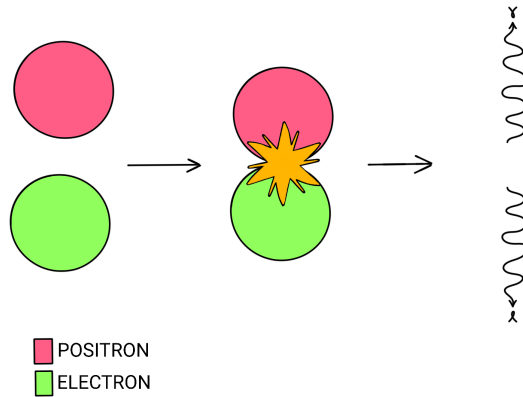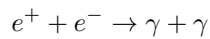

Figure 2: Positron annihilation

$$e^+ + e^- \to \gamma + \gamma$$

The PET scan machine consists of a ring of gamma wave detectors. When two detections take place in a time frame of 12ns they are assumed to come from the same collision [6]. Then, it is known that the emission took place along the line connecting the two detectors. The place of emission can be recovered by considering the time difference of arrivals of each gamma wave.

For example, if both gamma waves were detected at the same time, the point of emission would estimated as the middle of the line. This has been illustrated in the figure below.

Figure 3: PET scanner

The scan is created from multiple voxels, which are 3D pixels. Each voxel corresponds to a specific location and stores data about emissions detected in that area. The voxels are color-coded depending on the level of activity. An example of a PET scan image is visible below, where dark blue denotes very low activity and red the high level of activity.
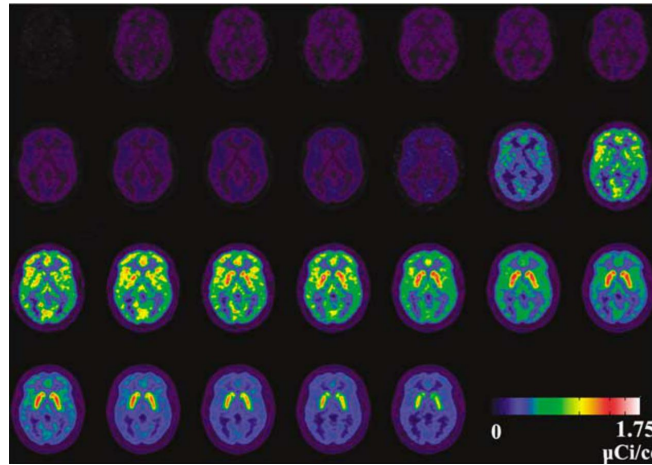

Figure 4: Dual tracer PET scan. Adapted from [12]

One could extend upon the technology behind the PET scan machines or the creation of radio-tracers. Although that information is not needed to comprehend dual tracer PET scans. Hence, a more detailed explanation will be omitted.

3 Dual Tracer PET scan

In terms of the procedure, a dual tracer PET scan does not differ much from a traditional single tracer one. The only noticeable change for a patient is the additional injection. However, the raw data coming from a dual tracer procedure is more difficult to analyse in comparison.

The reason for it, as stated in the Introduction is the fact, that all radiotracers used for PET scans emit the same level of energy, which is 511keV. As a consequence, the scanner cannot distinguish between the signals incoming from different radiotracers. A signal separating method has to be applied to the raw data coming from the scanner, before it can be used for medical analysis.

In order to obtain the best results from the dual tracer scan, the radioactive decay properties of the radiotracers have to be considered. If their half-lives are too similar, and both become active at the same time, the separation of the signals may become very difficult, or even close to impossible [12]. To counter that problem the injections should be performed approximately 30 minutes apart. Additionally, the first radiotracer should have a slow decay rate, while the second one a fast decay rate. Then, if the scan is performed soon after the second injection, only the first radiotracer should be active at the beginning of the procedure. This fact is very useful, as it allows for a more accurate signal separation result.

Furthermore, it is important to note that the signal separation would not have to be applied to all voxels. For example, in some of them only one of the radiotracers may be active, such as in Figure 5 and Figure 6. Another scenario, that does not require signal separation is depicted in Figure 7. There, the first radiotracer stopped emitting a signal before the next became active. The final case, as visible in Figure 8, is that where the signals do overlap. The delay of the second injection decreases the overlapped area, but it does not prevent this scenario from happening.

Therefore, the raw data from a dual tracer PET scan should be studied voxel by voxel, so that the signal separation procedure is applied when needed. The next section is about Extrapolation Method, which is one of the techniques that could be used to separate the signals.
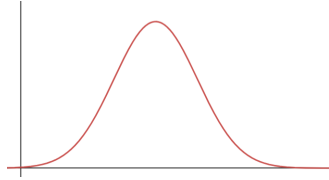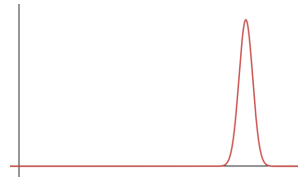

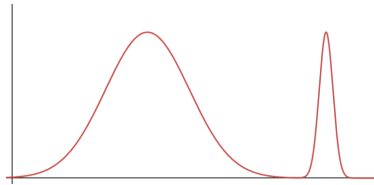Figure 5: Only first tracer is active   Figure 6: Only second tracer is active

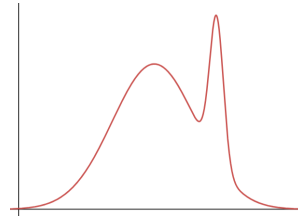

Figure 7: Signals do not overlap   Figure 8: Signals do overlap

4 Extrapolation Method

The Extrapolation Method was originally described in a paper [13] written by Koeppe et al. in 2004. The method depends on extrapolation of the first radiotracer signal. Extrapolation means approximating a function from an initial segment of it. Then, using the obtained estimation, the signal of the second radiotracer can be found.

The method will be illustrated with the help of an example. Let us assume that the activity noted in an arbitrary voxel of a dual tracer PET scan is as visible in Figure 9. In this case signals do overlap, therefore there is a need to use the Extrapolation Method.

As explained in the previous section, due to the delayed second injection, only the first radiotracer is active in the beginning of the scan. Therefore, we can consider the first time segment as known data of the first tracer. Only after a certain amount of time, the second tracer becomes active. This was depicted by a dash line in Figure 10.
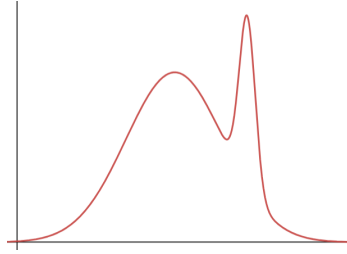
Figure 9: Full signal
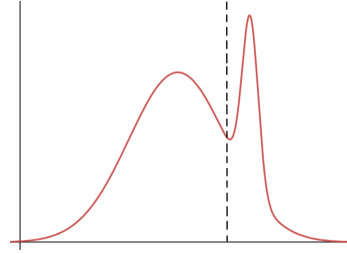
Figure 10: Second tracer becomes active

The data from the first time segment is used to extrapolate the signal of the first radiotracer. The estimation of the first signal is denoted by the blue function in Figure 11.

Then, the signal of the second radiotracer can be estimated by the subtraction of the extrapolated signal from the combined signals. The approximation of the second signal is depicted by the green function in Figure 12.
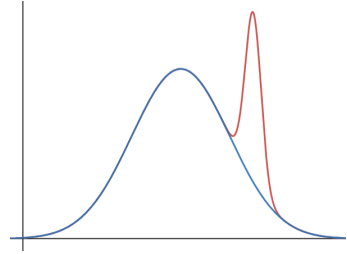
Figure 11: Extrapolation of first signal
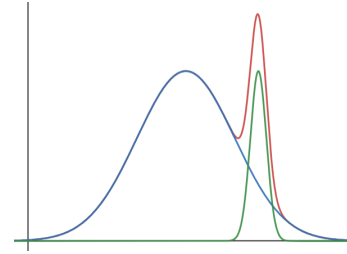
Figure 12: Estimations of both signals

The biggest challenge of using the Extrapolation Method is the estimation of the first signal. In a follow up study of the dual tracer signal separation, conducted by Joshi et al. in 2009 [12], a full reference tissue model was used for the purpose. The next section explains the model in detail.

4.1 Reference Tissue Model

Most approaches to the dual tracer problem require arterial sampling from a patient. It takes the form of arterial cannulation, which allows doctors to track current concentration of the radiotracer in an artery. Additionally a prior study of metabolites in patient's blood is required. This prolongs the PET scan procedure and makes it more stressful for the patient [15].

For that reason, Koeppe et al. [13] introduced an alternative, non-invasive approach, which uses a reference tissue model to estimate the arterial input function.
In medicine, the reference tissue is defined as a region visible on a PET scan, that should be unaffected by the treatment or by the suspected illness. Additionally it should have similar binding properties to that of the region of interest [16].
The effectiveness of the reference tissue model has been proven in a comparison study [14] conducted in 1996, where four different models were considered.

The following full reference tissue model was discussed in detail in [8] and [14]. It is important to note, that this model is limited to a case where both radiotracers used are reversible. A tracer being reversible means that the particles of the tracer do not get trapped inside a tissue for a prolonged period of time [3].

Moreover, as defined in a field of biochemistry, if a tissue has a specific binding towards a radiotracer, we refer to it as a bound molecule. Otherwise, we call it a free tissue molecule.

The diagram of the model is visible in Figure 13. There C_{blood} stands for the concentration of the radiotracer in patients blood system, C_{ref} denotes the concentration in the reference tissue, C_{free} in free molecules and finally C_{bound} is the concentration in bound molecules. The transfer rates between these compartments are denoted by K'_1, K'_2, K_1, K_2, K_3 and K_4.
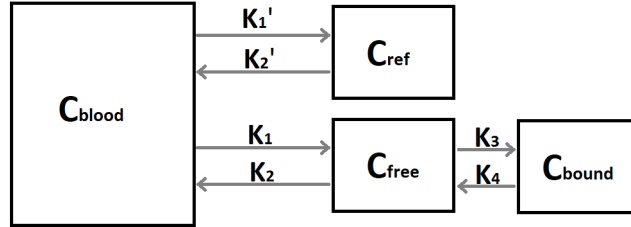

Figure 13: Full Reference Tissue Model

The kinetics of this model can also be described by the following three equations.

$$\frac{d}{dt}C_{ref}(t) = K'_1 C_{blood}(t) - K'_2 C_{ref}(t)$$

$$\frac{d}{dt}C_{free}(t) = K_1 C_{blood}(t) - K_2 C_{free}(t) - K_3 C_{free}(t) + K_4 C_{bound}(t)$$

$$\frac{d}{dt}C_{bound}(t) = K_3 C_{free}(t) - K_4 C_{bound}(t)$$

In practice, it is not possible to accurately measure the concentration of a radiotracer in the free and bound tissues separately. Instead, the sum of these two concentrations, referred to as the total concentration, is to be evaluated [15]. In our case, the region of interest for the concentration is that of an arbitrary voxel i.

Then, the total concentration is denoted by $y_i(t)$ and can be expressed by following equation.

$$y_i(t) = C_{free}(t) + C_{bound}(t)$$

Let us set $R_1 = K_1/K_1'$ and assume that $K_2' = K_2/R_1$.

Under those conditions, according to Lammertsma et al. [14], the relation between the total concentration $y_i(t)$ and the concentration in the reference tissue $C_{ref}(t)$ can be expressed by the following equation, where $\otimes$ denotes convolution.

$$y_i(t) = R_1 \left[C_{ref}(t) + aC_{ref}(t) \otimes e^{-ct} + bC_{ref}(t) \otimes e^{-dt} \right] \tag{1}$$

The parameters a, b, c and d are defined as follows.

$$a = \frac{\left(K_3 + K_4 - \frac{K_2+K_3+K_4+\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2} \right) \left(\frac{K_2+K_3+K_4+\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2} - \frac{K_2}{R_1} \right)}{\sqrt{(K_2+K_3+K_4)^2 - 4K_2K_4}}$$

$$b = \frac{\left(\frac{K_2+K_3+K_4-\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2} - K_3 - K_4 \right) \left(\frac{K_2+K_3+K_4-\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2} - \frac{K_2}{R_1} \right)}{\sqrt{(K_2+K_3+K_4)^2 - 4K_2K_4}}$$

$$c = \frac{K_2 + K_3 + K_4 + \sqrt{(K_2 + K_3 + K_4)^2 - 4K_2K_4}}{2}$$

$$d = \frac{K_2 + K_3 + K_4 - \sqrt{(K_2 + K_3 + K_4)^2 - 4K_2K_4}}{2}$$

The PET scan data for each voxel is saved in a form of a vector consisting of N entries. This is done to reduce the number of computations and data storage required [21]. The vector of total concentration for voxel i is denoted by $\bar{y}_i$.

$$\bar{y}_i = [y_i^1 \ldots y_i^j \ldots y_i^N]^T \tag{2}$$

Each of the entries corresponds to the average concentration over one of the N time frames. An arbitrary time frame $j \in [1, N]$ starts at time t_{start}^j and ends at t_{end}^j. The entry y_i^j is defined in the following manner.

$$y_i^j = \left(\frac{1}{t_{end}^j - t_{start}^j} \right) \int_{t_{start}^j}^{t_{end}^j} y_i(t)dt$$

The data of the concentration in the reference tissue is stored in the same way. Then the following equations represent the vector form of the reference tissue concentration and the formula for the j-th entry. It is important to note, that the values for a reference tissue model are assumed to be known.

$$\bar{C}_{ref} = [C_{ref}^1 \ldots C_{ref}^j \ldots C_{ref}^N]^T \tag{3}$$

$$C_{ref}^j = \left(\frac{1}{t_{end}^j - t_{start}^j}\right) \int_{t_{start}^j}^{t_{end}^j} C_{ref}(t)dt \tag{4}$$

The model equation 1 depends on four parameters. They are stored in the following vector form, denoted by $\bar{\theta}$.

$$\bar{\theta}_i = \left[R_1, K_2, K_3, K_4\right] \tag{5}$$

One of the important values that depends on these parameters is BP_{ND}, which denotes the non-displaceable binding potential [9]. This binding potential is used in reference tissue models and is approximated as $BP_{ND} = \frac{K_3}{K_4}$. As K_3 and K_4 are rates of transfer of the radiotracer between the the free and bound tissues, the binding potential informs us about the ratio of these two values. In the original full reference tissue model, as described in [14], the last entry of the parameter vector $\bar{\theta}_i$ is BP_{ND} instead of K_4. The change to the transfer rate K_4 was made to simplify future calculations.

As vector forms of both, the total concentration $\bar{y}_i$ and the reference tissue concentration $\bar{C}_{ref}$, have been introduced, it is time to discuss the relation between them. This is done by comparing the entries of corresponding time frames with each other. The result of that comparison is visible below. It is important to note, that the relation is similar to that between $y_i(t)$ and $C_{ref}(t)$ expressed in equation 1.

The difference is the addition of a new term ϵ_i^j, which denotes the residual error for voxel i and time frame j. This error is created by the fact that we are using a simplification, in the form of average value of each time segment, instead of a continuous function. The overall error can be minimized by considering smaller time frames.

$$y_i^j = R_1\left[C_{ref}^j(t) + aC_{ref}^j(t) \otimes e^{-ct} + bC_{ref}^j(t) \otimes e^{-dt}\right] + \epsilon_i^j \tag{6}$$

The relation 6 holds for all entries of the corresponding time frames.

We can introduce a vector form of the residual error, which consists of errors for all time frames.

$$\bar{\epsilon}_i = [\epsilon_i^1 \ldots \epsilon_i^j \ldots \epsilon_i^N]^T$$

Then the relation between the vector forms can be denoted as in equation below.

$$\bar{y}_i = f(\bar{C}_{ref}, \bar{\theta}_i) + \bar{\epsilon}_i,$$

where

$$f(\bar{C}_{ref}, \bar{\theta}_i) = \begin{bmatrix} R_1\left(C_{ref}^1(t) + aC_{ref}^1(t) \otimes e^{-ct} + bC_{ref}^1(t) \otimes e^{-dt}\right) \\ \vdots \\ R_1\left(C_{ref}^N(t) + aC_{ref}^N(t) \otimes e^{-ct} + bC_{ref}^N(t) \otimes e^{-dt}\right) \end{bmatrix}. \tag{7}$$

This concludes the theory behind the full reference tissue model.

4.2 Gauss-Newton method

Extrapolation is a necessary step of the signal separation method. With the use of the full reference tissue model and the initial data exclusive to the first radiotracer, one can approximate the signal. This is done by finding the parameters that minimize the difference between the model and the activity noted by the PET scanner. They are stored in a vector form denoted by $\hat{\theta}_i$.

Before fully defining the minimization problem, the variances corresponding to the different time frames have to be normalized. This is done to ensure that all of the N intervals have the same weight in the minimization problem. In other words, if one of the intervals is vastly deviating from the other values, its minimization should not skew the final results in its favour.

The normalization of the variance is done through introduction of a rescaling matrix W, as done by Joshi et al. in [12]. It is applied to the function $f(\bar{C}_{ref}, \bar{\theta}_i)$, defined in equation 7. The only non-zero elements of the matrix W are the diagonal entries, as visible below.

$$W = \begin{bmatrix} \frac{1}{\sigma_1^2} & & & & 0 \\ & \ddots & & & \\ & & \frac{1}{\sigma_j^2} & & \\ & & & \ddots & \\ 0 & & & & \frac{1}{\sigma_N^2} \end{bmatrix} \tag{8}$$

For all $j \in [1, N]$, the parameter σ_j^2 is defined by the following formula, where λ denotes the decay constant of a specific radiotracer and T_j is the midpoint of the time frame j.

$$\sigma_j^2 = \frac{y_i^j e^{\lambda T_j}}{t_{end}^j - t_{start}^j}$$

Therefore, the nonlinear minimization problem, as defined by Joshi et al. in [12], can be written in the following manner.

$$\hat{\theta}_i = \arg\min_{\bar{\theta}_i} ||\bar{y}_i - W f(\bar{C}_{ref}, \bar{\theta}_i)||_2^2 \tag{9}$$

This problem can be written in a more traditional format by introducing a function $g(\bar{\theta}_i)$, that represents the difference between the actual signal, as noted by the PET scanner, and the model values. The minimization problem is about finding a parameter vector $\hat{\theta}_i$ that minimizes this function.

An additional factor of $\frac{1}{2}$ has been included in the equation. It has no effect on the minimization problem defined in equation 9, as a multiplication by a constant does not change the argument $\hat{\theta}_i$ at which the minimum is achieved. The introduction of this factor simplifies future calculations.

Therefore, the nonlinear least squares problem is defined by the following equation.

$$g(\bar{\theta}_i) = \frac{1}{2} ||\bar{y}_i - W f(\bar{C}_{ref}, \bar{\theta}_i)||_2^2 \tag{10}$$

One of the methods used to solve such problems is called the Gauss-Newton method. It is a popular choice of an iterative method due to its relatively simple implementation for nonlinear equations.

In order to use this method a residual $\bar{r}_i$ has to be defined. For this problem it takes the following form.

$$\bar{r}_i = \bar{y}_i - W f(\bar{C}_{ref}, \bar{\theta}_i) \tag{11}$$

12

Let us assume that the residual is stored in the same way as $\bar{y}_i$ and $\bar{C}_{ref}$, which means that it is stored in N segments. Then it takes the following vector form.

$$\bar{r}_i = [r_i^1 \ldots r_i^j \ldots r_i^N]^T \tag{12}$$

Using equations 12, 2, 8 and 7 the residual equation 11 can be written in more detail.

$$\begin{bmatrix} r_i^1 \\ \vdots \\ r_i^j \\ \vdots \\ r_i^N \end{bmatrix} = \begin{bmatrix} y_i^1 \\ \vdots \\ y_i^j \\ \vdots \\ y_i^N \end{bmatrix} - \begin{bmatrix} \frac{1}{\sigma_1^2} & & & & \\ & \ddots & & & \\ & & \frac{1}{\sigma_j^2} & & \\ & & & \ddots & \\ & & & & \frac{1}{\sigma_N^2} \end{bmatrix} \begin{bmatrix} R_1\left(C_{ref}^1(t) + aC_{ref}^1(t) \otimes e^{-ct} + bC_{ref}^1(t) \otimes e^{-dt}\right) \\ \vdots \\ R_1\left(C_{ref}^j(t) + aC_{ref}^j(t) \otimes e^{-ct} + bC_{ref}^j(t) \otimes e^{-dt}\right) \\ \vdots \\ R_1\left(C_{ref}^N(t) + aC_{ref}^N(t) \otimes e^{-ct} + bC_{ref}^N(t) \otimes e^{-dt}\right) \end{bmatrix}$$

From this expression one can find the residual corresponding to each time frame. For $j \in [1, N]$ the residual r_i^j takes the following form.

$$r_i^j = y_i^j - \frac{1}{\sigma_1^2} R_1\left(C_{ref}^j(t) + aC_{ref}^j(t) \otimes e^{-ct} + bC_{ref}^j(t) \otimes e^{-dt}\right)$$

Another term needed for the Gauss Newton method is the Jacobian of the residual. As the considered problem depends on four parameters R_1, K_2, K_3 and K_4, the Jacobian J_{r_i} takes the following form.

$$J_{\bar{r}_i} = \begin{bmatrix} \frac{\partial r_i^1}{\partial R_1} & \frac{\partial r_i^1}{\partial K_2} & \frac{\partial r_i^1}{\partial K_3} & \frac{\partial r_i^1}{\partial K_4} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial r_i^j}{\partial R_1} & \frac{\partial r_i^j}{\partial K_2} & \frac{\partial r_i^j}{\partial K_3} & \frac{\partial r_i^j}{\partial K_4} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial r_i^N}{\partial R_1} & \frac{\partial r_i^N}{\partial K_2} & \frac{\partial r_i^N}{\partial K_3} & \frac{\partial r_i^N}{\partial K_4} \end{bmatrix}$$

Note that the Jacobian can also be written in a shorter form by considering the gradient of the 'time-frame' residuals. This form is used later in the subsection.

$$J_{\bar{r}_i} = \begin{bmatrix} (\nabla r_i^1)^T \\ \vdots \\ (\nabla r_i^j)^T \\ \vdots \\ (\nabla r_i^N)^T \end{bmatrix} \tag{13}$$

Due to the lengthy computations, only the formulas for the entries of the Jacobian are included in this section. The detailed calculations are attachted in the Appendix. The notation used corresponds to that defined in the previous subsection.

Given the similarities of the entries corresponding to different time frames, it suffices to consider the entries of an arbitrary time interval j.

The formula for the first column of the Jacobian entries, namely the partial derivative of the j-th residual with respect to the parameter R_1, where $j \in [1, N]$, is visible below.

$$\frac{\partial r_i^j(t)}{\partial R_1} = -\frac{1}{\sigma_j^2} C_{ref}^j(t) - \frac{1}{\sigma_j^2}\left(a + m\frac{K_2}{R_1}\right)C_{ref}^j(t) \otimes e^{-ct} - \frac{1}{\sigma_j^2}\left(b + n\frac{K_2}{R_1}\right)C_{ref}^j(t) \otimes e^{-dt}$$

The additional parameters m and n were introduced to simplify the equation. They take the following form.

$$m = \frac{K_3 + K_4 - \frac{K_2 + K_3 + K_4 + \sqrt{(K_2 + K_3 + K_4)^2 - 4K_2K_4}}{2}}{\sqrt{(K_2 + K_3 + K_4)^2 - 4K_2K_4}}$$

$$n = \frac{\frac{K_2 + K_3 + K_4 - \sqrt{(K_2 + K_3 + K_4)^2 - 4K_2K_4}}{2} - K_3 - K_4}{\sqrt{(K_2 + K_3 + K_4)^2 - 4K_2K_4}}$$

The formulas for the other three columns of the Jacobian are very similar to each other. In order to simplify the notation further and highlight the difference between the equations, a new constant term l is introduced. It is defined by the expression below.

$$l = \sqrt{K_2^2 + K_3^2 + K_4^2 - 2K_2K_4 + 2K_2K_3 + 2K_3K_4}$$

Then, the full formulas for the entries of the last three columns of the Jacobian are as follows.

$$\frac{\partial r_i^j(t)}{\partial K_2} = -\frac{1}{\sigma_j^2} R_1 a C_{ref}^j(t) \otimes \left(-te^{-ct}\left(\frac{1}{2} + \frac{K_2 + K_3 - K_4}{2l}\right)\right) -$$
$$- \frac{1}{\sigma_j^2} R_1 b C_{ref}^j(t) \otimes \left(-te^{-dt}\left(\frac{1}{2} - \frac{K_2 + K_3 - K_4}{2l}\right)\right)$$

$$\frac{\partial r_i^j(t)}{\partial K_3} = -\frac{1}{\sigma_j^2} R_1 a C_{ref}^j(t) \otimes \left(-te^{-ct}\left(\frac{1}{2} + \frac{K_2 + K_3 + K_4}{2l}\right)\right) -$$
$$- \frac{1}{\sigma_j^2} R_1 b C_{ref}^j(t) \otimes \left(-te^{-dt}\left(\frac{1}{2} - \frac{K_2 + K_3 + K_4}{2l}\right)\right)$$

$$\frac{\partial r_i^j(t)}{\partial K_4} = -\frac{1}{\sigma_j^2} R_1 a C_{ref}^j(t) \otimes \left(-te^{-ct}\left(\frac{1}{2} + \frac{-K_2 + K_3 + K_4}{2l}\right)\right) -$$
$$- \frac{1}{\sigma_j^2} R_1 b C_{ref}^j(t) \otimes \left(-te^{-dt}\left(\frac{1}{2} - \frac{-K_2 + K_3 + K_4}{2l}\right)\right)$$

As the residual term $\bar{r}_i$ and the Jacobian matrix $J_{\bar{r}_i}$ have been defined, the methodology of the Gauss-Newton method can be discussed.

This method is strictly related to the Newton method. The iteration algorithm of the latter is visible below, where θ_k is the estimated parameter vector found at the k-th step.

$$\theta_i^{(k+1)} = \theta_i^{(k)} - \left(\nabla^2 g(\theta_i^{(k)})\right)^{-1} \nabla g(\theta_i^{(k)}) \tag{14}$$

As explained in [22], the Gauss-Newton method is achieved by applying two modifications to the Newton method, namely the approximation of the gradient $\nabla g(\bar{\theta}_i)$ and of the Hessian $\nabla^2 g(\bar{\theta}_i)$.

The gradient of the function $g(\bar{\theta}_i)$ can be written down in the following form [23]. This is done by taking into account the relation between the gradient of a residual corresponding to the j-th time frame and the Jacobian, as visible in equation 13.

$$\nabla g(\theta_i) = \sum_{j=1}^{N} r_i^j \nabla r_i^j = J_{\bar{r}_i}(\theta_i)^T \bar{r}_i \tag{15}$$

The Hessian of the function $g(\bar{\theta}_i)$ can be written down in more detail as follows.

$$\nabla^2 g(\theta_i) = \sum_{j=1}^{N} \nabla r_i^j (\nabla r_i^j)^T + \sum_{j=1}^{N} r_i^j \nabla^2 r_i^j$$

Similarly to the gradient, the above formulation can be simplified with the use of the relation 13. Then the Hessian is as below.

$$\nabla^2 g(\theta_i) = J_{\bar{r}_i}^T J_{\bar{r}_i} + \sum_{j=1}^{N} r_i^j \nabla^2 r_i^j \tag{16}$$

According to [24], the second term of the equation 16 can be omitted. This can be done, as it tends to be of a small weight in comparison to the leading term.

Therefore, the Hessian of the function is approximated in the following way.

$$\nabla^2 g(\theta_i) \approx J_{\bar{r}_i}^T J_{\bar{r}_i} \tag{17}$$

The Gauss-Newton iteration is found by applying the approximations of the gradient 15 and of the Hessian 17 to the Newton method, as defined in equation 14. Then, the following formulation describes the iterative step $k+1$ of the method.

$$\theta_i^{(k+1)} = \theta_i^{(k)} - \left((J_{\bar{r}_i}(\theta_i^{(k)})^T J_{\bar{r}_i}(\theta_i^{(k)})) \right)^{-1} (J_{\bar{r}_i}(\theta_i^{(k)})^T \bar{r}_i \tag{18}$$

The convergence of this method depends on the accuracy of the approximation of the Hessian, which was described through equation 17. As discussed in detail by Wright and Nocedal in [23], the more accurate the approximation, the faster the convergence.

For example, if the following condition holds, then the iterative method results in a rapid local convergence.

$$||[J_{\bar{r}_i}^T J_{\bar{r}_i}(\theta)]^{-1} \sum_{j=1}^{N} r_i^j \nabla^2 r_i^j(\theta)|| \le 1$$

Similarly, the convergence of the method is quadratic, if the approximation of the Hessian, as described in equation 17, is accurate. Hence the equation below has to hold, for the convergence to be of the second order.

$$\sum_{j=1}^{N} r_i^j \nabla^2 r_i^j = 0$$

4.3 Gauss-Newton algorithm

The Gauss-Newton method is used to create algorithms solving nonlinear least squares problems. In this subsection two algorithms are discussed. The first one is a straightforward code using only the iterative step, as denoted in equation 18. The second algorithm is more complex, as it additionally uses a step size, which is meant to increase the convergence towards the solution.

For the minimization problem, originating from the dual tracer signal separation, the output of both algorithms is to be of a vector form $\hat{\theta} = [\hat{R}_1, \hat{K}_2, \hat{K}_3, \hat{K}_4]$.

Moreover, it is important to note, that both algorithms have to be applied voxel by voxel. Hence, without loss of generality, an arbitrary voxel i can be considered to illustrate the process. For that reason, the notation used in the algorithms is kept general.

Since Gauss-Newton is an iterative method, the starting point of an algorithm is defining an initial guess. For this problem, it is to be denoted as $\theta_0 = [R_1^0, K_2^0, K_3^0, K_4^0]$. The closer the initial guess is to the actual value, the less iterations will be needed to reach the solution.

Due to technical limitations, such as the computing power of a device and its storage capacity, finding an 'ideal' solution is close to impossible. Especially as transfer rates between different compartments are rarely integers. For that reason, one has to decide on a level of accuracy of the algorithm. It is to be denoted by ξ. If the computed error for a certain iteration is smaller than it, the iterative algorithm will stop.

There are many ways of defining an error in an iterative algorithm [17]. For example, it can be set equal to the difference between the values of a function at each iteration $g(\theta^{(k+1)}) - g(\theta^{(k)})$, or at the norm of a gradient of the function $||\nabla g(\theta^{(k)})||_2$. Although, as the Gauss-Newton method does not require direct use of the function g or its gradient, the error is set as a norm of the difference of the parameters found during the last two iterations.

$$error = ||\theta^{(k+1)} - \theta^{(k)}|| \tag{19}$$

Furthermore, both algorithms require the use of the residual and the Jacobian, as defined in the subsection 4.1. It is assumed that additional programs that output these functions have been created. They are referred to as $residual(\theta)$ and $Jacobian(\theta)$. The input of these programs is θ, which corresponds to the parameter vector found at each iteration.

With this information, the two algorithms can be discussed. As previously mentioned, the simpler of the two is based solely on the iteration step 18, which was the following.

$$\theta_i^{(k+1)} = \theta_i^{(k)} - \left((J_{\bar{r}_i}(\theta_i^{(k)})^T J_{\bar{r}_i}(\theta_i^{(k)})) \right)^{-1} (J_{\bar{r}_i}(\theta_i^{(k)})^T \bar{r}_i$$

To prevent storing unnecessary data, only the parameters of the two consecutive iterations are to be considered. The parameter vector of the previous iteration is denoted as θ_{old} and that of the current iteration is expressed as θ_{new}. At the end of each iterative step, the value of θ_{old} is replaced by that of the newer parameter vector. Moreover, the first value of the parameter vector is set to equal the initial guess θ_0.

The iterative step is to be repeated until the required accuracy ξ is achieved.

Note, that the error corresponding to the initial guess θ_0 is not known. To ensure that at least one iteration of the Gauss-Newton method is performed, the initial error is set as a number that is larger than the accuracy ξ. For example, in the following algorithm the initial error is set to equal $\xi + 10$.

Then, after each iteration the error is computed as defined in equation 19.

The algorithm of the Gauss-Newton method without a step size is visible below.

Algorithm 1 Simple Gauss Newton algorithm

  Input: θ_0, ξ

  $\theta_{new} = \theta_0$ $\triangleright$ Initial guess

  $error = \xi + 10$ $\triangleright$ Initial error

  while $error > \xi$ **do** $\triangleright$ Iterative step

      $\theta_{old} = \theta_{new}$

      $J = Jacobian(\theta_{old})$

      $r = residual(\theta_{old})$

      $\theta_{new} = \theta_{old} - \left(J^T * J\right)^{-1} * J^T * r$

      $error = ||\theta_{old} - \theta_{new}||$

  end while

This algorithm can be improved by the inclusion of a step size α into the iterative procedure.

Before the step size α can be defined, the direction of descent has to be introduced. This value changes every iteration, where for iteration k it is denoted as $d^{(k)}$. When considering the Gauss-Newton method the direction of descent equals the following formula.

$$d^{(k)} = -\left(\left(J_{\bar{r}_i}(\theta_i^{(k)})^T J_{\bar{r}_i}(\theta_i^{(k)})\right)\right)^{-1} (J_{\bar{r}_i}(\theta_i^{(k)})^T \bar{r}_i$$

The step size α is applied to the direction of descent. Then the modified iterative step of the Gauss-Newton method is as visible below.

$$\theta_i^{(k+1)} = \theta_i^{(k)} + \alpha d^{(k)}$$

There are multiple possibilities for defining the step size α, which vary in their complexity.

One, that could be considered the easiest option, is setting the step size equal to a constant. Then one does not have to update the value at each iteration. An appropriate constant can be found through numerical experimentation. That means running the algorithm multiple times with different constants, to decide on one, that decreases the number of iterations the most. Due to the limited amount of real-life data, simulations are often used. An example of a PET scan simulator is called SMART-PET [18]. This tool requires a single PET scan and CT scan (*Computed tomography*) to create PET scan images with similar features to the original.

A different possibility of a step size is defining it as an exact minimizer of each iteration, as visible in the equation below. Although, the downside of this option is the additional amount of computations required to find the value of each $\alpha^{(k)}$. This would vastly increase the costs of using the Gauss-Newton algorithm.

$$\alpha^{(k)} = arg \min_{\alpha(k)} g(\theta_i^{(k)} + \alpha^{(k)} d^{(k)})$$

Finally, the last option of defining the step size α, that is going to be discussed in this paper, is choosing values that satisfy the Wolfe conditions [25]. This backtracking linear search method consists of the sufficient decrease condition and the curvature condition. They are described below, where $0 < c_1 < c_2 < 1$ are two constants.

$$g(\theta^{(k)} + \alpha^{(k)}d^{(k)}) \leq g(\theta^{(k)}) + c_1\alpha^{(k)}(\nabla g(x^{(k)})^T d^{(k)}$$

$$\nabla g(\theta^{(k)} + \alpha^{(k)}d^{(k)})^T d^{(k)} \geq c_2(\nabla g(\theta^{(k)}))^T d^{(k)}$$

In the following algorithm, the Wolfe conditions are used to define the value of the step size α at each iteration. This choice has been made due to the high efficiency of the method, as discussed in [25].

Let $Wolfe(\theta)$ denote an additional algorithm that finds an appropriate value of the step size α, according to the conditions above. The input of this function is the parameter vector θ, which is found at each iteration.

As the values of the direction of descent d and the step size α have to be calculated at the beginning of each iteration, there is no need to store them. For that reason, their values are updated at each iteration.

Finally, the modified algorithm solving the minimization problem, as defined in 9, is visible below.

Algorithm 2 Gauss Newton algorithm with step size

Input: θ_0, ξ

$\theta_{new} = \theta_0$ ▷ Initial guess
$error = \xi + 10$ ▷ Initial error
while $error > \xi$ **do**   ▷ Iterative step
    $\theta_{old} = \theta_{new}$
    $J = Jacobian(\theta_{old})$
    $r = residual(\theta_{old})$
    $\alpha = Wolfe(\theta_{old})$
    $d = -\left(J^T * J\right)^{-1} * J^T * r$
    $\theta_{new} = \theta_{old} + \alpha d$
    $error = ||\theta_{old} - \theta_{new}||$
end while

The two discussed algorithms, based on the Gauss-Newton method, output the estimated parameter vector $\hat{\theta}_i$. It can be used in combination with the full reference tissue model to approximate the first radiotracer signal, for an arbitrary voxel i.

By subtracting the extrapolated signal from the full signal of the corresponding voxel, the activity of the second radiotracer can be approximated. The procedure has to be repeated for all voxels with overlapping signals. Once it has been completed, two separate PET scan images, corresponding to different radiotracers, can be recreated. Finally, the results of the study can be given to the medical personal to assist them in the process of forming a medical diagnosis.

5 Conclusion

In this paper the signal separation of dual tracer PET scans has been discussed. A dual tracer PET scan refers to the use of two radiotracers during a single scanning procedure. This approach would be used when two different physiological processes have to be studied, in order to form a medical diagnosis. Presently, when such a need arises a patient has to go through separate PET scan sessions. Often they cannot take place on the same day, to ensure that the tracers do not interfere with each others results. Dual tracer PET scans can be considered a more beneficial alternative, as they would shorten the time until the diagnosis, lower the medical costs and not expose the patient to unnecessary stress.

However, the dual tracer approach creates an additional difficulty. All radiotracers used for PET scans emit the same level of energy, due to which the machine cannot distinguish between the two signals. As a result, the medical personnel would not be able to differentiate the activities of the two radiotracer, which is a crucial factor in the process of forming a diagnosis. The delay of the second injection is a preventive measure, which is meant to decrease the overlap of the signals. Despite this, an additional implementation of a signal separation procedure is a necessary element of a dual tracer image reconstruction.

The Extrapolation Method is a signal separation technique introduced by Koeppe et al. in 2004 [13]. For this approach, the scan has to be considered voxel by voxel. The method has to be applied to only those voxels, in which the signals do overlap. In this method, the signal incoming from the first radiotracer is extrapolated with the use of initial data and a reference tissue model. The other signal is found by subtraction of the extrapolated signal from the total signal.

The full reference tissue model describes the kinetics of a radiotracer inside an artery. It is based on the concentration of the tracer and the transfer rates between different compartments. From this model a relation between the total concentration of the radiotracer in the tissues and in the reference region can be derived. In order to extrapolate the first signal, parameters minimizing the difference between the model and the raw data have to be found.

One of the approaches that can be used to estimate the model parameters is called the Gauss-Newton Method. This iterative algorithm can be viewed as a modified Newton Method. Algorithms based on this method can differ in the definition of the error and the inclusion of a step size.

With the use of these components, the radiotracer signals can be separated. Then, two PET scan images can be reconstructed from the original data and given to the medical personnel.

The advantages and limitations of the described signal separation procedure have to be discussed, in order to recognize the areas of possible future improvement.

The first component to be considered is the Extrapolation Method, which was introduced by Koeppe et al. in [13]. The biggest advantage of this method is the lack of arterial sampling, due to the use of a reference tissue model. As mentioned in subsection 4.1, this shortens the time of the procedure and decreases the amount of stress for the patient. On the other hand, the disadvantage of the method is the fact, that the accuracy depends on the limited amount of data exclusive to the first radiotracer. To increase that number, the second injection should be performed as late as possible, which is decided by the decay rate of the first tracer. However, even that measure cannot guarantee a high level of accuracy.

For that reason, alternative signal separation methods ought to be studied, to find one with a higher level of accuracy. An example of such a method is called a Simultaneous Fitting Method. It was introduced by Joshi et al. in a follow up study regarding the signal separation of dual tracer [12]. The method estimates both signals at the same time, and according to the conclusion of the study, it has a higher accuracy than the Extrapolation Method. However, this alternative approach is more difficult to implement due to its complexity. As such, an in depth comparison of both methods should be conducted, to determine which one is more beneficial for the dual tracer problems.

The next component, that has to be discussed, is the full reference tissue model. In 1996 a comparison study of different models describing the kinetics of radiotracers has been published [14]. They were considered in terms of estimating the binding potential parameter BP. It was concluded, that the use of the full reference tissue model is the most expedient due to the high accuracy of the results and easy implementation in comparison to alternative models. Nevertheless the model has a significant limitation, namley it is only valid if the radiotracer considered has reversible binding. A different model has to be used when dealing with irreversible radiotracers. An introduction to such a modified model has been included by Joshi et al. in [12].

Another element that should be studied is the possible chemical reaction between two different types of radiotracers. If both, an irreversible and a reversible tracer, are used during a single PET scan, it could have an effect on their kinetics. Firstly, the research should be conducted from a bio-chemical perspective, to understand the feasible changes on the molecular level. Depending on the results of that study, the reference tissue model may be modified further.

A different perspective that ought to be considered is the computational cost of using the discussed model. The complexity of it may result in a higher amount of computations required, making it more expensive. Because of that, a slightly less accurate model may be preferred by medical institutions. For example, a simplified version of the full reference tissue model has been introduced by Lammertsma and Hume in [15]. According to the conclusion of the paper, the simplified model's accuracy does not differ much from the original, but it is yet to be verified by numerical experimentation. As such, the simplified model should be researched as a possible, cheaper alternative of the full reference tissue model.

The final component of the signal separation procedure, that is to be reviewed, is the Gauss-Newton iterative method. This method is one of the most commonly used for solving nonlinear minimization problems [5]. The reason for that, are the two approximations on which the method is based, namely that of the gradient and the Hessian of the considered function. As the second order partial derivatives do not have to be computed, this method is easier to implement than the Newton method. Another advantage of this method, is the rapid local convergence, which is achieved when the approximation of the Hessian is accurate. Furthermore, if the approximation of the Hessian is exact, the convergence is quadratic. By that reason, the Gauss-Newton iterative method may be considered beneficial to use.

When considering the algorithm of the Gauss-Newton method, a possible difficulty is defining the step size α. Due to the complexity of the problem, finding values that fulfill certain conditions may require multiple computations. That would make the use of the iterative method more expensive. A possible solution to that, would be using a cheaper algorithm at the beginning of the process to find the general area of the solution. When that area is located, the algorithm would be switched to the Gauss-Newton method in order to converge to the solution at a higher rate.

There are many possibilities for a further extension of the study of dual tracers PET scans.

The most straightforward one, is applying the discussed methods to an actual data set. As the signal separation method has to be applied voxel by voxel it would be very time consuming to do it by hand. In order to hasten the process, a more complex algorithm should be developed. Firstly, it should categorize the signal of each voxel into the possible types discussed in Section 3, namely: only one of the tracers is active, both tracers are active but there is no overlap and both tracers are active and signals overlap. In case only one of the radiotracers is active, the program should match the signal to the analogous tracer. If both tracers are active but there is no overlap the algorithm should split the signals instantaneously. Finally, if the voxel is in the overlap category, a signal separation procedure should be implemented. After every voxel has been categorized, the data should be redistributed to two data sets corresponding to each of the radiotracers. Then, separate PET scan images for each tracer can be recreated.

Another possibility for future research is considering a fundamentally different approach of the signal separation of dual tracers. For example, using a Deep Learning technique, which is a type of AI trained to learn patterns from a previous data set and apply it to new problems. This methodology is commonly used in dual tracer research due to its efficiency with nonlinear and complex functions. The difficulty of this method is , that it requires a large training set, which is the data from which the program is supposed to learn. Examples of such research are: [27], [28] and [19].

Finally, once the separation of dual tracers has been studied in detail, the research can be extended to the use of more than two radiotracers during a single PET scan. The signal separation in that scenario is more difficult, as a larger number of model parameters has to be estimated. The main difficulty of the multi tracer problem would be finding a procedure that results in a high level of accuracy.

As discussed, the topic of dual tracers is complex, and the research can be extended in various directions. Although, due to the scope of this paper and limited amount of time, only one of the signal separation methods has been studied in detail. For that reason, this paper should be treated as an introduction to the topic of signal separation of dual tracer PET scans.

Researching this topic makes one realize the significance of creating connections between different disciplines, even when they seem to be unrelated at first glance. In order to consider the problem of signal separation of dual tracers, one needs knowledge not only about mathematics, but also a basic understanding of medical procedures and bio-chemical processes. Similar correlations appear when considering other subjects of modern research. This gives the motivation for conducting interdisciplinary mathematics.

6 Appendix

In this section the calculations of Jacobian entries are discussed. Due to the lengthy computations only the final formulas for these entries are included in section 4.2.

Without loss of generality let us consider voxel i and the j-th time frame interval. Then the expression for the corresponding residual is described as in equation 11.

$$r_i^j(t) = y_i^j(t) - \frac{1}{\sigma_1^2} R_1 \left(C_{ref}^j(t) + a C_{ref}^j(t) \otimes e^{-ct} + b C_{ref}^j(t) \otimes e^{-dt} \right)$$

Parameters a, b, c, d are as defined in section 4.1 and symbol $\otimes$ denotes convolution. By expanding the bracket, the residual equation can be rewritten as follows.

$$r_i^j(t) = y_i^j(t) - \frac{1}{\sigma_j^2} R_1 C_{ref}^j(t) - \frac{1}{\sigma_j^2} R_1 a C_{ref}^j(t) \otimes e^{-ct} - \frac{1}{\sigma_j^2} R_1 b C_{ref}^j(t) \otimes e^{-dt} \qquad (20)$$

In order to find the Jacobian, partial derivatives with respect to entries of the parameter vector $\bar{\theta}_i$, which are R_1, k_2, k_3, k_4, have to be computed. Note that the parameters a, b, c, d depend on the entries of $\bar{\theta}_i$, as such full forms of the variables have to be considered during calculations.

The full forms of the a, b, c, d as defined in section 4.1 are visible below.

$$a = \frac{\left(K_3 + K_4 - \frac{K_2 + K_3 + K_4 + \sqrt{(K_2+K_3+K_4)^2 - 4K_2K_4}}{2} \right) \left(\frac{K_2+K_3+K_4+\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2} - \frac{K_2}{R_1} \right)}{\sqrt{(K_2+K_3+K_4)^2 - 4K_2K_4}}$$

$$b = \frac{\left(\frac{K_2+K_3+K_4-\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2} - K_3 - K_4 \right) \left(\frac{K_2+K_3+K_4-\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2} - \frac{K_2}{R_1} \right)}{\sqrt{(K_2+K_3+K_4)^2 - 4K_2K_4}}$$

$$c = \frac{K_2 + K_3 + K_4 + \sqrt{(K_2 + K_3 + K_4)^2 - 4K_2K_4}}{2}$$

$$d = \frac{K_2 + K_3 + K_4 - \sqrt{(K_2 + K_3 + K_4)^2 - 4K_2K_4}}{2}$$

The detailed calculations behind the Jacobian entries are split over the next four subsections. Each of them concerns partial derivatives with respect to one of the $\bar{\theta}_i$ entries. To simplify the explanations, each of the four terms of the residual equation 20 is considered separately. Only at the end of each subsection, the partial derivatives of all four terms, with respect to the same variable, will be summed, to find the expression for the corresponding Jacobian entry.

6.1 Jacobian entry with respect to R_1

Firstly, let us consider the partial derivative with respect to variable R_1. Note that the first term of equation 20 does not depend on this parameter, as such the following statement holds.

$$\frac{\partial}{\partial R_1} y_i^j(t) = 0$$

Then, the partial derivative of the second term of equation 20 with respect to R_1 is easy to compute as the variable appears only once.

$$\frac{\partial}{\partial R_1}\left(-\frac{1}{\sigma_j^2}R_1 C_{ref}^j(t)\right) = -\frac{1}{\sigma_j^2}C_{ref}^j(t)$$

For the next two terms it is important to note that the a and b depend on the parameter R_1. Therefore one has to use the product rule

$$\frac{d}{dz}\left(v(z)w(z)\right) = v(z)w'(z) + v'(z)w(z),$$

where v and w are two arbitrary functions depending on the variable of interest denoted by z.

For the third term of equation 20, the variable of interest z can be set equal to R_1.

Then set $v(R_1) = -\frac{1}{\sigma_j^2}R_1$ and $w(R_1) = aC_{ref}^j(t) \otimes e^{-ct}$.

Firstly, let us find the expression $v'(R_1)$.

$$\frac{\partial}{\partial R_1}\left(-\frac{1}{\sigma_j^2}R_1\right) = -\frac{1}{\sigma_j^2}$$

Then we have to find the value of $w'(R_1)$, so we need to calculate the following.

$$\frac{\partial}{\partial R_1}\left(aC_{ref}^j(t) \otimes e^{-ct}\right)$$

Note that from the expression above only parameter a depends on the variable R_1. Hence the rest of the expression can be treated as a constant.

$$\frac{\partial}{\partial R_1}\left(aC_{ref}^j(t) \otimes e^{-ct}\right) = \left(\frac{\partial}{\partial R_1}a\right)C_{ref}^j(t) \otimes e^{-ct}$$

Therefore the only partial derivative, which has to be computed to find $w'(R_1)$ is $\frac{\partial a}{\partial R_1}$. The calculations of it are visible below.

$$\frac{\partial a}{\partial R_1} = \frac{\partial}{\partial R_1}\left(\frac{\left(K_3 + K_4 - \frac{K_2+K_3+K_4+\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2}\right)\left(\frac{K_2+K_3+K_4+\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2} - \frac{K_2}{R_1}\right)}{\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}\right)$$

Let m denote a part of expression of a, that does not depend on parameter R_1.

$$m = \frac{K_3 + K_4 - \frac{K_2+K_3+K_4+\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2}}{\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}} \tag{21}$$

Then

$$\frac{\partial a}{\partial R_1} = m\frac{\partial}{\partial R_1}\left(\frac{K_2 + K_3 + K_4 + \sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2} - \frac{K_2}{R_1}\right) =$$
$$= m\frac{\partial}{\partial R_1}\left(-\frac{K_2}{R_1}\right) =$$
$$= m\left(\frac{K_2}{R_1^2}\right)$$

With this, the value of $w'(R_1)$ can be found, and is as following.

$$\frac{\partial}{\partial R_1}\left(aC_{ref}^j(t)\otimes e^{-ct}\right) = m\frac{K_2}{R_1^2}C_{ref}^j(t)\otimes e^{-ct}$$

With these calculations, it can be concluded that the partial derivative of the third term of residual equation 20 with respect to the variable R_1 is equal to the following.

$$\frac{\partial}{\partial R_1}\left(-\frac{1}{\sigma_j^2}R_1aC_{ref}^j(t)\otimes e^{-ct}\right) = -\frac{1}{\sigma_j^2}aC_{ref}^j(t)\otimes e^{-ct} - \frac{1}{\sigma_j^2}R_1m\frac{K_2}{R_1^2}C_{ref}^j(t)\otimes e^{-ct}$$

This can be simplified further.

$$\frac{\partial}{\partial R_1}\left(-\frac{1}{\sigma_j^2}R_1aC_{ref}^j(t)\otimes e^{-ct}\right) = -\frac{1}{\sigma_j^2}\left(a + m\frac{K_2}{R_1}\right)C_{ref}^j(t)\otimes e^{-ct}$$

Now let us consider the partial derivative of the final term of the residual equation 20 with respect to the variable R_1. Therefore we want to find the formula for the following.

$$\frac{\partial}{\partial R_1}\left(-\frac{1}{\sigma_j^2}R_1bC_{ref}^j(t)\otimes e^{-dt}\right)$$

Similarly to previous approach let us split the term into a product of two functions $v(z)$ and $w(z)$, in order to use the product rule. As we are still considering the partial derivatives with respect to R_1, let us set z as that variable. Let $v(R_1) = -\frac{1}{\sigma_j^2}R_1$ and $w(R_1) = bC_{ref}^j(t)\otimes e^{-dt}$.

Note that as $v(R_1)$ has been defined identically to that used for the third term of the residual equation 20, the value of $v'(R_1)$ is also the same.

$$\frac{\partial}{\partial R_1}\left(-\frac{1}{\sigma_j^2}R_1\right) = -\frac{1}{\sigma_j^2}$$

Therefore what is left to compute is $w'(R_1)$, which then takes the following form.

$$\frac{\partial}{\partial R_1}\left(bC_{ref}^j(t)\otimes e^{-dt}\right)$$

Note that only the parameter b depends on the variable R_1, so the rest of the term can be treated as a constant.

$$\frac{\partial b}{\partial R_1} = \frac{\partial}{\partial R_1}\left(\frac{\left(\frac{K_2+K_3+K_4-\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2}-K_3-K_4\right)\left(\frac{K_2+K_3+K_4-\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2}-\frac{K_2}{R_1}\right)}{\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}\right)$$

Let us create a constant n, being a part of formula for b that does not depend on the parameter R_1.

$$n = \frac{\frac{K_2+K_3+K_4-\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2}-K_3-K_4}{\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}} \tag{22}$$

Then partial derivative of parameter b with respect to R_1 can be rewritten in following way.

$$\frac{\partial b}{\partial R_1} = n\frac{\partial}{\partial R_1}\left(\frac{K_2+K_3+K_4-\sqrt{(K_2+K_3+K_4)^2-4K_2K_4}}{2}-\frac{K_2}{R_1}\right) = n\frac{K_2}{R_1^2}$$

Using that, the following equation holds.

$$\frac{\partial}{\partial R_1}\left(bC^j_{ref}(t) \otimes e^{-dt}\right) = n\frac{K_2}{R_1^2}C^j_{ref}(t) \otimes e^{-dt}$$

Therefore it can be concluded that the partial derivative of the last term of the residual equation 20 with respect to the variable R_1 is the following.

$$\frac{\partial}{\partial R_1}\left(-\frac{1}{\sigma_j^2}R_1 bC^j_{ref}(t) \otimes e^{-dt}\right) = -\frac{1}{\sigma_j^2}\left(bC^j_{ref}(t) \otimes e^{-dt}\right) - \frac{1}{\sigma_j^2}R_1 n\frac{K_2}{R_1^2}C^j_{ref}(t) \otimes e^{-dt}$$

It can be simplified in the following manner.

$$\frac{\partial}{\partial R_1}\left(-\frac{1}{\sigma_j^2}R_1 bC^j_{ref}(t) \otimes e^{-dt}\right) = -\frac{1}{\sigma_j^2}\left(b + n\frac{K_2}{R_1}\right)C^j_{ref}(t) \otimes e^{-dt}$$

As all of the partial derivatives of all terms of the residual equation 20 are known, the final formula is the sum of them. Therefore the partial derivative of the residual equation with the respect to the variable R_1 is the formula below, where m, n are as defined in equations 21 and 22.

$$\frac{\partial r_i^j(t)}{\partial R_1} = -\frac{1}{\sigma_j^2}C^j_{ref}(t) - \frac{1}{\sigma_j^2}\left(a + m\frac{K_2}{R_1}\right)C^j_{ref}(t) \otimes e^{-ct} - \frac{1}{\sigma_j^2}\left(b + n\frac{K_2}{R_1}\right)C^j_{ref}(t) \otimes e^{-dt}$$

6.2 Jacobian entry with respect to K_2

To find the next entry of the Jacobian, let us consider the partial derivatives of all four entries of residual equation 20 with respect to variable K_2.

Note that the first term of the equation does not depend on the parameter K_2. As such the following equation holds.

$$\frac{\partial}{\partial K_2}y_i^j(t) = 0 \tag{23}$$

Similarly, the second term of the residual equation does not have any dependencies on that variable. As such the partial derivative of it equals zero.

$$\frac{\partial}{\partial K_2}\left(-\frac{1}{\sigma_j^2}R_1 C^j_{ref}(t)\right) = 0 \tag{24}$$

The other two parameters depend on the variable K_2 nonlinearly.

Firstly, let us consider the partial derivative of the third term of the residual equation.

$$\frac{\partial}{\partial K_2}\left(-\frac{1}{\sigma_j^2}R_1 aC^j_{ref}(t) \otimes e^{-ct}\right)$$

Note that both of the parameters a and c depend on K_2. As such we cannot treat the convolution as a constant. Instead let us consider one of the convolution properties.

When considering convolution of the form $(v \otimes w)$, where both of the functions are differentiable with respect to z the following property holds.

$$\frac{\partial}{\partial z}(v \otimes w) = (\frac{\partial}{\partial z}v) \otimes w = v \otimes (\frac{\partial}{\partial z}w)$$

As both convoluted functions of the considered term are differentiable with respect to K_2, instead of computing the partial derivative of the full term, we can perform the calculation on just the second part of it.

$$\frac{\partial}{\partial K_2}\left(-\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t) \otimes e^{-ct}\right) = -\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t) \otimes \left(\frac{\partial}{\partial K_2}e^{-ct}\right) \tag{25}$$

The following calculations only require the use of a chain rule. Note that the time t does not depend on any parameters.

$$\frac{\partial}{\partial K_2}\left(e^{-ct}\right) = e^{-ct}\frac{\partial}{\partial K_2}(-ct) = -te^{-ct}\frac{\partial c}{\partial K_2} \tag{26}$$

Hence only the partial derivative of c is required.

$$\frac{\partial c}{\partial K_2} = \frac{1}{2} + \frac{(2K_2 - 2K_4 + 2K_3)}{4\sqrt{K_2^2 + K_3^2 + K_4^2 - 2K_2 K_4 + 2K_2 K_3 + 2K_3 K_4}}$$

To simplify the notation, let us set a new constant parameter l equal to the following.

$$l = \sqrt{K_2^2 + K_3^2 + K_4^2 - 2K_2 K_4 + 2K_2 K_3 + 2K_3 K_4}$$

Then the expression for the partial derivative of c becomes the following.

$$\frac{\partial c}{\partial K_2} = \frac{1}{2} + \frac{K_2 + K_3 - K_4}{2l} \tag{27}$$

By combining this results with equations 25 and 26, we reach the following result.

$$\frac{\partial}{\partial K_2}\left(-\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t) \otimes e^{-ct}\right) = -\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t) \otimes \left(-te^{-ct}\left(\frac{1}{2} + \frac{K_2 + K_3 - K_4}{2l}\right)\right) \tag{28}$$

The next step is to compute the partial derivative of the last term of the residual equation 20. The reasoning behind the calculations is similar to that of the third term. Firstly, one has to consider the properties of convolution.

$$\frac{\partial}{\partial K_2}\left(-\frac{1}{\sigma_j^2}R_1 b C_{ref}^j(t) \otimes e^{-dt}\right) = -\frac{1}{\sigma_j^2}R_1 b C_{ref}^j(t) \otimes \left(\frac{\partial}{\partial K_2}e^{-dt}\right) \tag{29}$$

To find the partial derivative of the second part of the term, one has to use the chain rule.

$$\frac{\partial}{\partial K_2}e^{-dt} = -te^{-dt}\frac{\partial d}{\partial K_2}$$

Using the same constant parameter l as defined for the previous term, the partial derivative of parameter d is as written below.

$$\frac{\partial d}{\partial K_2} = \frac{1}{2} - \frac{K_2 + K_3 - K_4}{2l}$$

Notice, that the only difference between this expression, and that written in equation 27 is the change in the sign. This fact simplifies the calculations for the partial derivatives, not only with respect to K_2, but also to K_3 and K_4.

Therefore the partial derivative of the last term from the residual equation 20 with respect to variable K_2 is the following.

$$\frac{\partial}{\partial K_2}\left(-\frac{1}{\sigma_j^2}R_1 b C_{ref}^j(t)\otimes e^{-dt}\right) = -\frac{1}{\sigma_j^2}R_1 b C_{ref}^j(t)\otimes\left(-te^{-dt}\left(\frac{1}{2}-\frac{K_2+K_3-K_4}{2l}\right)\right) \qquad (30)$$

By taking into account equations 23, 24, 28 and 30 the Jacobian entry corresponding to variable K_2 takes the following form.

$$\frac{\partial r_i^j(t)}{\partial K_2} = -\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t)\otimes\left(-te^{-ct}\left(\frac{1}{2}+\frac{K_2+K_3-K_4}{2l}\right)\right) -$$
$$-\frac{1}{\sigma_j^2}R_1 b C_{ref}^j(t)\otimes\left(-te^{-dt}\left(\frac{1}{2}-\frac{K_2+K_3-K_4}{2l}\right)\right)$$

6.3 Jacobian entry with respect to K_3

Let us calculate the formula for the Jacobian entry with respect to variable K_3. To find it one has to take the same steps as for the previous variable.

Note that just as in the previous subsection, the first two terms of the residual equation 20 are zero. It is due to the fact that they do not depend on the variable K_3 in any way.

$$\frac{\partial}{\partial K_3}y_i^j(t) = 0 \qquad (31)$$

$$\frac{\partial}{\partial K_3}\left(-\frac{1}{\sigma_j^2}R_1 C_{ref}^j(t)\right) = 0 \qquad (32)$$

As such the formula for the Jacobian entry depends on the partial derivatives of the latter two terms of the residual equation.

Using the convolution properties as previously, along with the fact that time t does not depend on any variables we have the following.

$$\frac{\partial}{\partial K_3}\left(-\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t)\otimes e^{-ct}\right) = -\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t)\otimes\frac{\partial}{\partial K_3}\left(e^{-ct}\right) =$$
$$= -\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t)\otimes\left(-te^{-ct}\frac{\partial c}{\partial K_3}\right)$$

Using the constant parameter l as introduced in previous subsection, the partial derivative of parameter c with the respect to variable K_3 can be written in following manner.

$$\frac{\partial c}{\partial K_3} = \frac{1}{2}+\frac{K_2+K_3+K_4}{2l}$$

Then the partial derivative of the third term of the residual equation 20 with respect to K_3 is as follows.

$$\frac{\partial}{\partial K_3}\left(-\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t)\otimes e^{-ct}\right) = -\frac{1}{\sigma_j^2}R_1 a C_{ref}^j(t)\otimes\left(-te^{-ct}\left(\frac{1}{2}+\frac{K_2+K_3+K_4}{2l}\right)\right) \qquad (33)$$

27

Using the similarities between the third and fourth terms of the residual equation, and those between parameters c and d the following equations can be found.

$$\frac{\partial}{\partial K_3}\left(-\frac{1}{\sigma_j^2}R_1 bC_{ref}^j(t)\otimes e^{-dt}\right) = -\frac{1}{\sigma_j^2}R_1 bC_{ref}^j(t)\otimes\left(-te^{-dt}\frac{\partial d}{\partial K_3}\right)$$

$$\frac{\partial d}{\partial K_3} = \frac{1}{2} - \frac{K_2 + K_3 + K_4}{2l}$$

Then, the partial derivative of the last term of the residual equation 20 with respect to variable K_4 is the following.

$$\frac{\partial}{\partial K_3}\left(-\frac{1}{\sigma_j^2}R_1 bC_{ref}^j(t)\otimes e^{-dt}\right) = -\frac{1}{\sigma_j^2}R_1 bC_{ref}^j(t)\otimes\left(-te^{-dt}\left(\frac{1}{2}-\frac{K_2+K_3+K_4}{2l}\right)\right) \tag{34}$$

Finally, using equations 31, 32, 33 and 34 the following formula for the Jacobian entry corresponding to variable K_3 is found.

$$\frac{\partial r_i^j(t)}{\partial K_3} = -\frac{1}{\sigma_j^2}R_1 aC_{ref}^j(t)\otimes\left(-te^{-ct}\left(\frac{1}{2}+\frac{K_2+K_3+K_4}{2l}\right)\right) - $$
$$- \frac{1}{\sigma_j^2}R_1 bC_{ref}^j(t)\otimes\left(-te^{-dt}\left(\frac{1}{2}-\frac{K_2+K_3+K_4}{2l}\right)\right)$$

6.4   Jacobian entry with respect to K_4

Let us consider the formula for the Jacobian entry with respect to the last of the variables, namely K_4. Once again, the reasoning behind the computations is similar to those of corresponding to the second and third Jacobian entries.

Just as in previous cases, the partial derivatives of the first two terms of the residual equation 20 are equal to zero, as they do not depend on the variable K_4.

$$\frac{\partial}{\partial K_4}y_i^j(t) = 0 \tag{35}$$

$$\frac{\partial}{\partial K_4}\left(-\frac{1}{\sigma_j^2}R_1 C_{ref}^j(t)\right) = 0 \tag{36}$$

Then the partial derivative of the third term depends on the convolution properties, and the lack of dependency of time t on the parameter.

$$\frac{\partial}{\partial K_4}\left(-\frac{1}{\sigma_j^2}R_1 aC_{ref}^j(t)\otimes e^{-ct}\right) = -\frac{1}{\sigma_j^2}R_1 aC_{ref}^j(t)\otimes\left(-te^{-ct}\frac{\partial c}{\partial K_4}\right)$$

The partial derivative of parameter c with respect to variable K_4 can be written down using the constant parameter l, as defined in subsection 6.2.

$$\frac{\partial c}{\partial K_4} = \frac{1}{2} + \frac{-K_2 + K_3 + K_4}{2l}$$

Then the partial derivative of the third term of the residual expression equals the following.

$$\frac{\partial}{\partial K_4}\left(-\frac{1}{\sigma_j^2}R_1 aC_{ref}^j(t)\otimes e^{-ct}\right) = -\frac{1}{\sigma_j^2}R_1 aC_{ref}^j(t)\otimes\left(-te^{-ct}\left(\frac{1}{2}+\frac{-K_2+K_3+K_4}{2l}\right)\right) \tag{37}$$

28

Due to similarities noted in previous subsections, the partial derivative of the fourth term depends on the two following expressions.

$$\frac{\partial}{\partial K_4}\left(-\frac{1}{\sigma_j^2}R_1bC_{ref}^j(t)\otimes e^{-dt}\right) = -\frac{1}{\sigma_j^2}R_1bC_{ref}^j(t)\otimes\left(-te^{-dt}\frac{\partial d}{\partial K_4}\right)$$

$$\frac{\partial d}{\partial K_4} = \frac{1}{2} - \frac{-K_2+K_3+K_4}{2l}$$

Using that the formula for the partial derivative of the fourth term of the residual equation 20 can be written in the following way.

$$\frac{\partial}{\partial K_4}\left(-\frac{1}{\sigma_j^2}R_1bC_{ref}^j(t)\otimes e^{-dt}\right) = -\frac{1}{\sigma_j^2}R_1bC_{ref}^j(t)\otimes\left(-te^{-dt}\left(\frac{1}{2}-\frac{-K_2+K_3+K_4}{2l}\right)\right) \quad (38)$$

Using equations 35, 36, 37 and 38 the Jacobian entry corresponding to the variable K_4 can be expressed by the following formula.

$$\frac{\partial r_i^j}{\partial K_4} = -\frac{1}{\sigma_j^2}R_1aC_{ref}^j(t)\otimes\left(-te^{-ct}\left(\frac{1}{2}+\frac{-K_2+K_3+K_4}{2l}\right)\right)-$$
$$-\frac{1}{\sigma_j^2}R_1bC_{ref}^j(t)\otimes\left(-te^{-dt}\left(\frac{1}{2}-\frac{-K_2+K_3+K_4}{2l}\right)\right)$$

This concludes the Appendix about the Jacobian calculations.

References

[1] *The Important Role Medical Imaging Plays in Diagnosis and Treatment.* Available online: `https://www.pbmchealth.org/news-events/blog/important-role-medical-imaging-plays-diagnosis-and-treatment`, 2018. Accessed on 03 June 2023.

[2] *Positron emission tomography scan.* Available online: `https://www.mayoclinic.org/tests-procedures/pet-scan/about/pac-20385078`, 2023. Accessed on 24 June 2023.

[3] *Reversible versus irreversible binding.* Available online: `https://trc-p.nl/60/`, 2023. Accessed on 24 June 2023.

[4] M. A. Ashraf and A. Goyal. Fludeoxyglucose (18f). Available online: `https://www.ncbi.nlm.nih.gov/books/NBK557653/`, 2022. Accessed on 03 June 2023.

[5] D. B. Bertsekas. *Nonlinear Programming*, chapter 1.5.1, pages 107–108. Athena Scientific, Belmont, Massachusetts, 2 edition, 1999.

[6] M. Chappell. *Back-Projection Reconstruction—CTand PET/SPECT*, pages 77–86. Springer International Publishing, Cham, 2019.

[7] M. Chappell. *Emission—SPECT/PET*, pages 31–37. Springer International Publishing, Cham, 2019.

[8] V. J. Cunningham, S. P. Hume, G. R. Price, R. G. Ahier, J. E. Cremer, and A. K. P. Jones. Compartmental analysis of diprenorphine binding to opiate receptors in the rat in vivo and its comparison with equilibrium data in vitro. *Journal Of Cerebral Blood Flow and Metabolism*, 11:1–9, 1991.

[9] R. B. Innis, V. J. Cunningham, J. Delforge, M. Fujita, A. Gjedde, R. N. Gunn, J. Holden, S. Houle, S.-C. Huang, M. Ichise, H. Iida, H. Ito, Y. Kimura, R. A. Koeppe, G. M. Knudsen, J. Knuuti, A. A. Lammertsma, M. Laruelle, J. Logan, R. P. Maguire, M. A. Mintun, E. D. Morris, R. Parsey, J. C. Price, M. Slifstein, V. Sossi, T. Suhara, J. R. Votaw, D. F. Wong, and R. E. Carson. Consensus nomenclature for in vivo imaging of reversibly binding radioligands. *Journal of Cerebral Blood Flow & Metabolism*, 27:1533–1539, 2007.

[10] M. Itoh, H. Watabe, M. Miyake, S. Hagisawa, T. Fujiwara, R. Iwata, T. Ido, and T. Nakamura. Chapter 14 - noninvasive determination of arterial input of 15o tracers, using a dual cutaneous -detector set above the radial artery. In R. Myers, V. Cunningham, D. Bailey, and T. Jones, editors, *Quantification of Brain Function Using PET*, page 67. Academic Press, San Diego, 1996.

[11] J. Jones, Y. Mellam, and R. Chieng. Radiopharmaceuticals. Available online: `https://radiopaedia.org/articles/radiopharmaceuticals?lang=us`, 2010. Accessed on 03 June 2023.

[12] A. D. Joshi, R. A. Koeppe, J. A. Fessler, and M. R. Kilbourn. Signal separation and parameter estimation in noninvasive dual-tracer pet scans using reference-region approaches. *Journal of Cerebral Blood Flow Metabolism*, 29:1346–1357, 2009.

[13] R. A. Koeppe, A. Joshi, K. Frey, S. E. Snyder, M. R. Kilbourn, and F. J. Dual-tracer pet studies without arterial sampling. *NeuroImage*, 22:T115–6, 2004.

[14] A. A. Lammertsma, C. J. Bench, S. P. Hume, S. Osman, K. Gunn, D. J. Brooks, and R. S. J. Frackowiak. Comparison of methods for analysis of clinical ^{11}c raclopride studies. *Journal of Cerebral Blood Flow and Metabolism*, 16:42–52, 1996.

[15] A. A. Lammertsma and S. P. Hume. Simplified reference tissue model for pet receptor studies. *Neuroimage*, 4:153–158, 1996.

[16] V. Oikonen. Reference tissue in receptor studies. Available online: `http://www. turkupetcentre.net/petanalysis/model_reference_tissue.html`, 2021. Accessed on 5 July 2023.

[17] J. Peypouquet. Lecture 2: Existence of global minimizers. gradient descent. [slides]. University of Groningen, 2022. Accessed on 25 June 2023.

[18] E. Pfaehler, J. De Jong, and R. e. a. Dierckx. Smart (simulation and reconstruction) pet: an efficient pet simulation-reconstruction tool. *EJNMMI Phys*, 5(16), 2018.

[19] D. Ruan and H. Liu. Separation of a mixture of simultaneous dual-tracer pet signals: A data-driven approach. *IEEE Transactions on Nuclear Science*, 64(9):2588–2597, 2017.

[20] A. Shetty, I. Bickle, P. Qureshi, and et al. Positron emission tomography. Available online: `https://radiopaedia.org/articles/positron-emission-tomography?lang=us`, 2014. Accessed on 03 June 2023.

[21] M. G. Spangler-Bickell, S. A. Hurley, T. W. Deller, F. Jansen, V. Bettinardi, M. Carlson, M. Zeineh, G. Zaharchuk, and A. B. McMillan. Optimizing the frame duration for data-driven rigid motion estimation in brain pet imaging. *Med Phys.*, 48:3031–3041, 2021.

[22] J. N. Stephen and J. Wright. *Numerical Optimization*, chapter 3, pages 44–46. Springer Science+Business Media, 233 Spring Street, New York, NY 10013, USA, 2 edition, 2006.

[23] J. N. Stephen and J. Wright. *Numerical Optimization*, chapter 10, pages 254–257. Springer Science+Business Media, 233 Spring Street, New York, NY 10013, USA, 2 edition, 2006.

[24] J. N. Stephen and J. Wright. *Numerical Optimization*, chapter 10, pages 245–247. Springer Science+Business Media, 233 Spring Street, New York, NY 10013, USA, 2 edition, 2006.

[25] J. N. Stephen and J. Wright. *Numerical Optimization*, chapter 3.1, pages 33–36. Springer Science+Business Media, 233 Spring Street, New York, NY 10013, USA, 2 edition, 2006.

[26] G. Wang, A. Rahmim, and R. N. Gunn. Pet parametric imaging: Past, present, and future. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 4(6):663–675, 2020.

[27] J. Xu and H. Liu. Deep-learning-based separation of a mixture of dual-tracer single-acquisition pet signals with equal half-lives: A simulation study. *IEEE Transactions on radiation and plasma medical sciences*, 3(6):649–659, 2019.

[28] J. Xu and H. Liu. Hybrid-loss guided 3d cnn for dynamic dual-tracer pet reconstruction. *IEEE 16th International Symposium on Biomedical Imaging*, pages 1618–1621, 2019.