



UNIVERSITY OF TWENTE.

Faculty of Electrical Engineering,
Mathematics & Computer Science

Dynamic objects in emission tomography reconstructed with temporal B-splines

Xiaoming P. op de Hoek
M.Sc. Thesis
May, 2018

Assessment Committee:

Prof. Dr. Stephan A. van Gils
Dr. Christoph Brune
Dr. Pranab K. Mandal

Daily Supervisor:

Dr. Christoph Brune

Applied Analysis Group
Department of Applied Mathematics
Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Abstract

In medical imaging the role of 4D dynamic reconstruction methods are fundamental, for example for the study of moving objects, like heart or lung motion, or tracer activity (kinetics). Many existing reconstruction methods in nuclear medicine, e.g. PET, are based on Poisson modelling, however often the objects are static and motion is not taken into account. We make use of a model for inhomogeneous poisson point processes and will study the relation between different methods, that reconstruct data from a given source, in the context of motion.

Contents

1	Introduction	2
1.1	Outline	3
1.2	Problem	4
1.3	Main Goals	6
1.4	Research Context	6
2	Model	9
2.1	Continuous Data-term 1D	10
2.2	Continuous Data-term ND	12
2.3	B-splines	13
2.4	Regularisation Methods	15
3	Numerics	16
4	Results and Discussion	19
4.1	Constant Function	19
4.2	Piecewise Function	24
5	Conclusion	24
6	Appendix	29
6.1	Specific Mathematics	29

1 Introduction

In many medical imaging facilities in hospitals in the world, patients are medically examined for the mapping of brain and heart functions or finding tumors. In the examination it is of great benefit to see how active the biochemical process is and where in the body it is active. These processes in the body can be measured and viewed through the use of Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT), both are part of emission computed tomography (ECT). Since PET and SPECT are able to capture processes in the body, both are functional imaging methods.

To capture a process with ECT, an (radioactive) isotope, the *tracer*, is administered to a patient. Every process that needs to be imaged has its own specific tracer, which is an isotope of an element that is present in the specific process. The tracer has a short half-life, because it is unhealthy to be exposed to radiation for too long. A practical advantage of a short half-life is the possibility of more ECT-scans in a short amount of time.

To be able to make an actual image the gamma rays emitted by the tracer¹ need to be observed with a special camera. Multiple attributes of the gamma ray can be measured, e.g. energy, occurrence time and angle, when observed. These attributes are used to acquire better results, but only occurrence time is assumed to be available.

PET is typically measured in an array of detectors that makes up a ring. The array of detectors in SPECT is typically a plane, with often the freedom that this plane can rotate around the patient, often called an Anger camera. The scanners of Philips Healthcare, [16, 15], for both PET and SPECT seen in Figure 1 show that the array placement is apparent even in the end product.

From the acquired data an image is computed that shows the distribution of the tracers activity and thus makes it possible to examine functions inside the body. The computation can be done analytically, but nowadays iterative methods are favoured, which are presented in this work.

Regardless of the method of computation, a tracer that is present in greater quantities emits more gamma rays. Hence the number of detections per

¹PET and SPECT have tracers with different properties. SPECT-tracers emit gamma rays directly in the decay process, whereas PET-tracers emit positrons, these meet free electrons in the tissue, which causes an annihilation event and gamma rays appear.

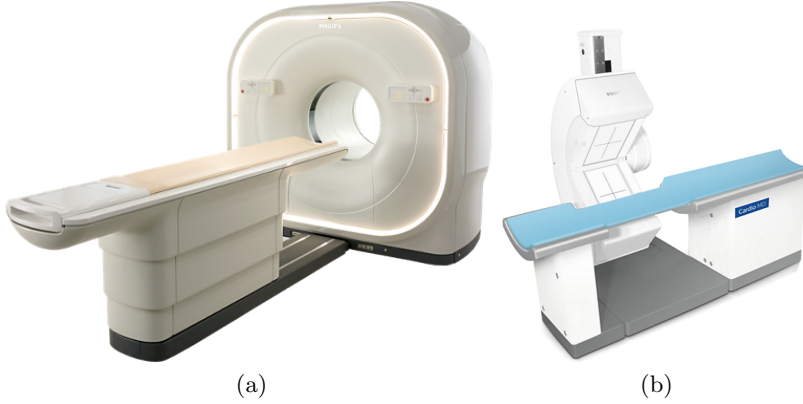


Figure 1: Philips Healthcare: (a) Vereos Digital PET/CT-scanner, it can simultaneously scan PET and CT and overlay the results. (b) CardioMD IV is a SPECT-scanner.

detector is higher, which results in a computed image of better quality. Unfortunately the great quantity of radioactive matter is also harmful. Hence, it is desirable to have low count and good quality.

During the detection of gamma rays from the process in a body it is optimal that the body is static. A static body has fixed location, which means that detected gamma rays are from the same location in the body. This results in an image of better quality, respectively, it represents better the process in the body. In reality the function that needs to be imaged is not part of a static body. Examples of functions with motion are blood flow or lung motion.

Applied Goal

This work is about motion of an object in ECT and the produced image. We focus on the number of detected gamma rays per detector or number of counts per detector. So there is no focus on the details of PET and SPECT or which one is better. For a more comprehensive background on ECT, see [17].

1.1 Outline

This thesis is organised as follows. We start describing our problem in the next section and address our main goal in subsection 1.3. In subsection 1.4

we will address the decisions we made to study our problem. We start to describe the one-dimensional model in section 2 and extend that one-dimensional concept to higher dimensions. The simplification of this higher dimensional problem is done in subsection 2.3. We end that section with the complete discrete model and various methods. The results are shown in section 4 and this work is close in section 5.

1.2 Problem

In medical imaging the role of 4D² is fundamental, because there is a lot of motion inside the patient's body apart from the patient's movements. Examples are the aforementioned heart or lung motion, but also tracer activity (kinetics).

To the best of our knowledge all dynamic ECT studies only investigate the retrieval of time-activity-curves (TAC) and kinetic parameters of physiological processes, in a relatively static object. If the body undergoes motion, this affects the tracer activity at a certain point in space. For visual examples see Figure 2. This multi-dimensional problem can be reduced to multiple one-dimensional problems, due to detectors partitioning space. The role of the one-dimensional subproblem is even more important when there is motion. Over time accelerations occur, because of motion, with shifts in space-time as result. These shifts are preserved as drops of activity, referred to as a *gap* in Figure 2,

The TACs studied in papers that are reviewed in the work by A. Reader [18, R381] are smooth and in some cases just a monotonically decreasing function. Those TACs do not reflect a sense of motion, since there are no gaps. Since the role of movement is fundamental we are particularly interested in the case when we do have a moving object in ECT.

Instantaneous change of the tracer activity exemplifies the shifts, due to motion, in a TAC. This change is not necessarily noticeable, if the time to the next opposite instantaneous change is short. The natural question coming forth would be: how short can an interval between two opposite, instantaneous, changes be to be detected?

Those intervals are the gaps and change the mean number of counts in

²3D space and 1D time.

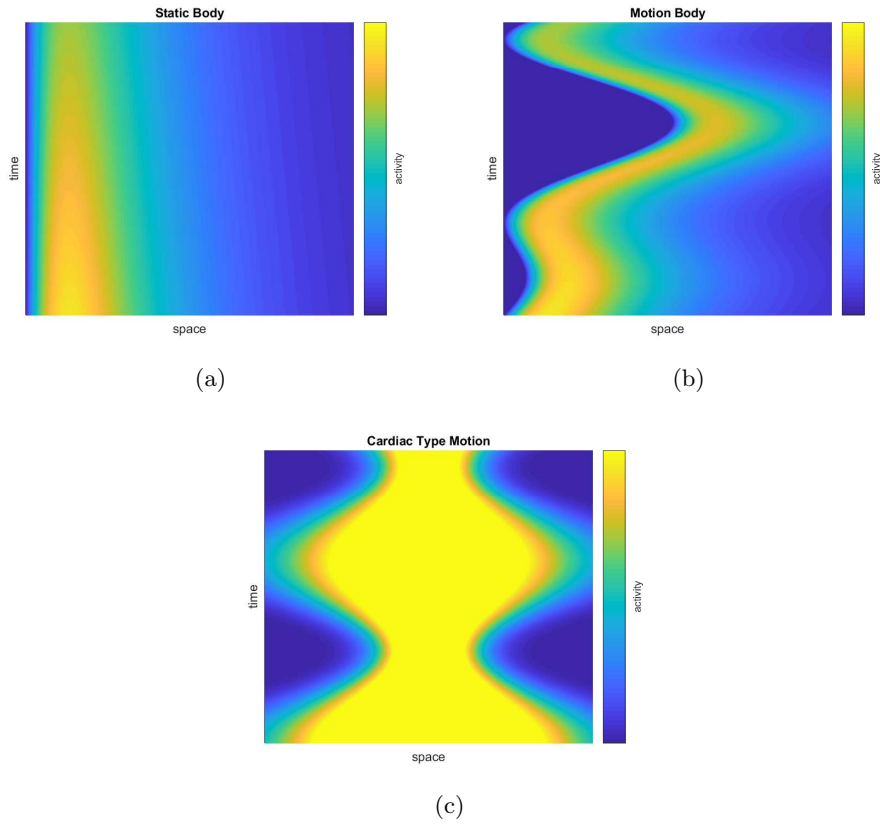


Figure 2: Space-time activity of (a) a static body, (b) a body in motion and (c) motion of the heart, which causes periodic expansion. This is meant to show the difference of a source staying in a point in space and one that is moving.

a detector, which can cause unfaithful reconstructions. Although we know that more counts in a detector yields better quality TACs, it can also be more harmful for the patient. So it is of interest if there is a measurable drop in quality beyond a certain number of counts.

A better understanding of both these mechanics could help medical imaging. It would represent the tracer activity better over time (in a moving body) and minimize the exposure of a patient to radioactive matter.

1.3 Main Goals

To achieve our *Applied Goal* we have goals that are more specific. The main goals of this work is to improve the understanding of the relation between a homogeneous and a non-homogeneous Poisson model in the context of Medical image reconstruction. We do this by comparing results from both models and measure how they perform under different circumstances. Of particular interest is (simulated) motion. The specific goals are elaborated in the upcoming section.

Specific Goals:

- Does there exist an apparent improvement going from low valued SNR to higher valued SNR? Or more general, e.g. can we quantify the word *low* in low SNR?
- How short can an interval between two opposite, instantaneous, changes be to be detected
- Does there exist an optimal ratio between the number of intervals and the number of counts? (i.e. can we find a rule of thumb)

1.4 Research Context

To study our problem we made decisions, for reasons that will be covered in the next paragraph. One decision we made is to interpolate (temporally) over acquired list-mode data. For the interpolation we use uniform B-splines, see subsection 2.3. Due to these decisions we make use of an iterative method to minimize a cost function, see (1), that describes the problem. We will be using a first-order primal-dual algorithm, by Chambolle [4], to minimize our cost function. Note that other authors also suggested this class of algorithms, yet in slightly different forms.

The reason for choosing list-mode data as a way of storing acquisition data is to improve precision and memory usage. ECT acquisition data can be saved in multiple formats. Counts can be binned based on time intervals, i.e. making a histogram, or saved in lists (list-mode). The difference is that binned data only saves the number of counts in an interval, whereas list-mode registers the occurred time of one count (event). So from list-mode it is possible to acquire binned data, but not the other way around, which means list-mode contains more information. In works as early as 1984 [24] it is elaborated how list-mode data could improve ECT. With list-mode not only precision, but also memory usage could be improved, as argued in Barrett in [2].

In list-mode data the occurrence times of an *event* are stored, that event is the detection of a gamma ray, i.e. a photon. In essence this is a multi-dimensional photon counting problem and that is a Poisson process, see [9, ch. 4], [12, ch. 3.9] and [25, p. 69, 93], for more elaborate explanation. However, the classical (analytical) method Filtered Back Projection (FBP), that is used to go from acquisition data to a reconstructed image, can't use this information about the noise.

Favoured nowadays are iterative methods, often in cases of statistical models, because they can deal with the physics of ECT, see [17]. Such as Poisson noise generated by the Poisson process and other various effects that worsen reconstruction, such as scattering and attenuation. List-mode data improves precision, which is our motivation to use a statistical model.

Iterative methods work with a cost function that needs to be maximized or minimized. The commonly used cost function in literature, with the specific noise incorporated into it, is given by,

$$\arg \min_u \int_{\Omega} \int_0^T \mathcal{K}u - f \log(\mathcal{K}u) dt dx + \alpha R(u). \quad (1)$$

Here u is the source, the tracer, that emits gamma rays (photons) per unit time. This is done through the use of $\mathcal{K}$ a (non-)linear compact forward operator, which describes how the measurements are done on u in the specific setup of PET and SPECT. Then f is our observed data. Both $\mathcal{K}$ and f are known, such that we can reconstruct u . The last term $R(u)$ is the regularisation term, that penalizes the values of u to make it smoother for example. We elaborate this in forthcoming sections.

The most used iterative methods to solve variant of (1) are variations of the expectation-maximization (EM) algorithm with and without regularisation term, for one of the early papers see [23]. However EM cannot easily be generalized to a cost function with regularisation, so we use a primal-dual algorithm where this is possible. This primal-dual algorithm is used in only one other work on this topic, see [7], as far as we know. We chose this iterative method, because of its ease of use, adaptability and performance.

To compute an solution from the continuous representation of our problem, see (1), we need to discretise all variables. Multiple works, listed in [18], use a representation for u that has physiological meaning. But we are interested in motion and as described in subsection 1.2 the TAC, u , can have no physiological meaning. So the reason for picking B-splines specifically as discrete representation for u is that B-splines also does not have restrictive physiological meaning.

Furthermore, B-splines, have also been used in older works see [14, 20, 13, 10, 27, 28, 26]. Except Li and Verhaeghe, [10, 26], all these works favour non-uniform B-splines. Verhaeghe even describes succinctly how to jointly optimize the non-uniform knots for B-splines in [28]. A reason for most studies choosing non-uniform B-spline is given in [14], since concentration decay there is less data. Non-uniform B-splines, distribute sparse data more efficiently. This non-uniformity has more freedom in parameter choice, which brings the question how to place the non-uniform mesh. To try to understanding the relationship between the signal-to-noise-ratio (SNR) and the number of basis-function, we do not need extra complexities. Note that high SNR, means a lot of counts. To study those effects it is more intuitive to use uniform B-splines, because solutions don't have a different distribution of B-splines. So the reason of using uniform B-splines is 'fair' comparison between realizations.

As said, none of the aforementioned references, [14, 20, 13, 10, 27, 28, 26], study the effects between SNR in a comprehensive way. That means they are not clear about the number of events that have been used to make a reconstruction. When they do mention the number of events, they do not clarify if that is for one pixel or in the whole plane of pixels. For example in [18] it is said that a scanner detects somewhere between 10^4 and 10^9 number of events depending on the study type and duration. A scanner can mean the whole set-up, but it could also be one detector. Another example is [22]

where the words ‘*extremely low SNRs*’ are used, but not quantified. Hence, addressing this issue is one of our goals, which was previewed in the previous subsection and is summarized in the upcoming subsection.

2 Model

In the general (ideal) image reconstruction we study the operator equation:

$$\mathcal{K}u^\dagger = f^\dagger. \quad (2)$$

Here $\mathcal{K} : X \rightarrow Y$ is a (non-)linear forward operator and has a similar meaning as described in subsection 1.4. It operates on the source $u^\dagger \in X$ and results in the (exact) data $f^\dagger \in Y$. Due to data being changed by noise and modelling effects, we do not acquire $f^\dagger$, but rather a noisy $f^\dagger$. Which can be written in a similar fashion as (2),

$$\begin{aligned} \mathcal{K}u &= f, \\ &= \eta \left(f^\dagger \right). \end{aligned} \quad (3)$$

In this problem we want to resolve $u \in X$, which is the inexact reconstruction of the acquired data $f \in \mathbb{R}^N$ which is typically discrete. The noise, η , is applied to the original signal instead of being multiplicative or additive, because the acquisition data in ECT are, as mentioned earlier photon counts and photon counting is a Poisson process³. So the data is generated with the original signal as the non-negative Poisson intensity. An assumption here is that photon detections are considered independent arrivals, but this does not generally hold, Barrett describes this in [2].

The Poisson intensity can be homogeneous, i.e. constant, and is called a *homogeneous Poisson process* (HPP). The non-constant counterpart is called a *non-homogeneous Poisson process* (NHPP)⁴ for one-dimensional definitions of both one can look in the section 6, for multi-dimensional definitions one can read [25, 11]. The next subsections elaborate on HPP and NHPP and their use. But first we will describe a method to attempt to solve (3).

To solve (3) for u the trivial option is to invert $\mathcal{K}$, but generally speaking it cannot be (continuously) inverted. That is due possible compactness

³Again see [9, ch. 4], [12, ch. 3.9] and [25, p. 69, 93].

⁴Synonymous with inhomogeneous Poisson point process and non-stationary Poisson point process

of $\mathcal{K}$, which in turn means solving (3) for u is ill-posed. We already mentioned in subsection 1.4 that there exist direct (analytical) methods, such as FBP, to solve this problem. But it is advantageous to use iterative methods, since it is possible for extensions which deal with various effects, such as noise models. We do not take into account other effects separately in the model. Although these effects do affect the intensity, we considered them as part of the synthetic source signal.

These iterative methods determine values for u , an example for an iterative method is Newton's method, which uses Taylor expansions to approximate solutions iteratively. Since our generated data suffers from Poisson noise, we can also derive a variational problem specific to this noise. This is done by use of the probability density function of the Poisson process data in a maximum-likelihood (ML) estimation. The use of ML is common, but we take a Bayesian approach and use the maximum-a-posterior (MAP). More specifically the negative log MAP, because of the exponential function in a Poisson distribution. With this Bayesian approach we get the description for a variational problem, which is used to acquire expressions for HPP and NHPP.

$$\begin{aligned}
& \arg \min_{u \in X} -\log(p(u|f)), \\
& \Rightarrow \arg \min_{u \in X} -\log(p(f|u)) - \log(p(u)), \\
& = \arg \min_{u \in X} D(f; \mathcal{K}u) + \alpha R(u).
\end{aligned} \tag{4}$$

The nice thing of MAP is the term $p(u)$ which is the prior probability of u and can also be interpreted as the regularisation of u , since it can be chosen freely. Hence if not explicitly noted, we will be using Gibbs functions $p(u) \sim \exp\{-\alpha R(u)\}$. The data-term $D(f; \mathcal{K}u)$ refers to $-\log(p(f|u))$, the problem specific expression is derived from the stochastic process in the next sections.

2.1 Continuous Data-term 1D

The list-mode data is presumed to be collected for a given end time, in the nuclear-medicine this assumption is referred to as *preset time* (see [2]). Starting with the one-dimensional (temporal $u \in L_1(\mathbb{R})$) problem, where list-mode data $f \in \mathbb{R}^N$, obtained from Poisson intensity $\mathcal{K}u$ are random variables $W_0 < W_1 < W_2 < \dots < W_n < \dots < W_N < T$ denoting the occurrence time of the n -th photon appearance, here $W_0 = 0$, $n \in [0, N] \subset \mathbb{N}$

and N is the total number of events in the time interval $[0, T]$. To be specific, $\Lambda(a, b) = \int_a^b \mathcal{K}u \, dt$ is the *parameter* of the Poisson process and $\mathcal{K}u$ the *intensity*.

For a homogeneous, i.e. constant, $\mathcal{K}u$ parameter $\Lambda(a, b)$ can be computed explicitly. The Poisson distribution that generates the noise,

$$Poisson(n; T\mathcal{K}u) := e^{-T\mathcal{K}u} \frac{(T\mathcal{K}u)^n}{n!}. \quad (5)$$

For the MAP data-term in (4) we need to find out the conditional density function (cdf). Consequently, Proposition 8.2 is used,

$$\begin{aligned} -\log(p(\{w_n\}_n|u)) &= -\log\left(\prod_n p(w_n|W_{n-1} = w_{n-1}, u) p(T|W_N = w_N, u)\right), \\ &= -\log\left(\prod_n \mathcal{K}u(w_n) e^{-(w_n - w_{n-1})\mathcal{K}u} e^{-(T - w_N)\mathcal{K}u}\right), \\ &= T\mathcal{K}u - \sum_n \log(\mathcal{K}u(w_n)), \\ &= T\mathcal{K}u - N \log(\mathcal{K}u(w_n)), \\ &=: D_{hom}(N; \mathcal{K}u). \end{aligned} \quad (6)$$

We made use of the Markov property of a Poisson process, the given cdf of Proposition 8.2 and a variant of the cdf with zero events between the last occurrence time and the preset time T . In the last step the assumption of constant intensity was used.

For the one-dimensional (temporal $u \in L_1(\mathbb{R})$) inhomogeneous problem the steps are fairly similar, but we start with any non-negative function $\mathcal{K}u$ as the Poisson intensity. Thus computing $\Lambda(a, b)$ explicitly is not possible and the major difference, because the integral becomes part of the problem. Assumptions for the random variables and notation are still the same. We also use Proposition 8.2 and a variation with zero events up to the end of the time interval.

$$Poisson(n; \Lambda(0, T)) := e^{-\Lambda(0, T)} \frac{(\Lambda(0, T))^n}{n!}. \quad (7)$$

The expression for the data-term in (4) then reads,

$$\begin{aligned}
-\log(p(\{w_n\}_n|u)) &= -\log\left(\prod_n p(w_n|W_{n-1}=w_{n-1},u)p(T|W_N=w_N,u)\right), \\
&= -\log\left(\prod_n \mathcal{K}u(w_n)e^{-\Lambda(w_{n-1},w_n)}e^{-\Lambda(w_N,T)}\right), \\
&= \Lambda(0,T) - \sum_n \log(\mathcal{K}u(w_n)), \\
&= \int_0^T \mathcal{K}u - f(t)\mathcal{K}u dt, \\
&=: D_{nhom}(\{w_n\}_n; \mathcal{K}u).
\end{aligned} \tag{8}$$

Here $f(t) = \sum_n \delta(t - w_n)$, where $\delta(\cdot)$ is the Dirac-delta function.

It is obvious that (6) results from (8), if $\mathcal{K}u$ is taken constant in (8). That will also be assumed in the multi-dimensional case in the next subsection.

2.2 Continuous Data-term ND

The one dimensional problem can easily be extended to a multi-dimensional problem, e.g. 2D-space and 1D-time, but is too involved for this work. We will only deal with multiple one-dimensional problems as representation of a multi-dimensional problem, which will be derived in subsection 2.3. Therefore precise derivation as done in the previous subsection is not done here. We refer to [25, p. 87] for the derivation of the multi-dimensional ML expression, which is changed in MAP expression. The multi-dimensional NHPP data-term for the variational problem (4),

$$D_{KL}(f|\mathcal{K}u) := \int_{\Omega} \int_0^T \mathcal{K}u - f(x,t) \log(\mathcal{K}u) dt dx. \tag{9}$$

Here $f(x,t) = \sum_{n=1}^{\bar{N}} \delta(x - x_n, t - w_n)$, with $\bar{N} = \sum_k N_k$. Both are an impulse train and just for the occasion x_n is defined as the spatial location and w_n the occurrence time of the n -th event.

If $\mathcal{K}u$ is homogeneous, then the HPP data-term for the variational problem (4) that results from (9), is given by,

$$D_{hom}(N|\mathcal{K}u) := T|\Omega| \mathcal{K}u - N \log(\mathcal{K}u) dt dx. \tag{10}$$

With N the total number of events that occurred up to time T in space Ω . Whereas $|\Omega|$ is the multi-dimensional ‘size’ of the space, e.g area in 2D and volume in 3D.

2.3 B-splines

A simplification that is done to easily use B-splines is to consider the spatial and temporal dimensions to be multiplicative, as is usual in the entire current literature [18] and will be used throughout this work to discretise the continuous expression from subsection 2.1 and subsection 2.2. So first we define the one-dimensional B-spline, which is short for *basic splines*. We do not cover definitions for multi-dimensional B-splines, but these are usually product of multiple one-dimensional B-splines and we assume that too.

Definition 1 ([5, p. 89]). A B-splines basis function of degree p or order $p + 1$,

$$B_{i,0}(t) := \begin{cases} 1, & \text{if } t_i \leq t < t_{i+1}, \\ 0, & \text{otherwise.} \end{cases},$$

$$B_{i,p}(t) = \frac{t - t_i}{t_{i+p} - t_i} B_{i,p-1}(t) + \frac{t_{i+p+1} - t}{t_{i+p+1} - t_{i+1}} B_{i+1,p-1}(t).$$

These basis function have some common properties.

1. The $B_{i,p}$, form a *partition of unity*, hence $\sum_i B_{i,p} = 1$.
2. The $B_{i,p}(t) > 0$ in the interior interval (t_i, t_{i+p+1}) and zero outside this interval. Furthermore a particular given $t_i = t_{i+p+1} \Rightarrow B_{i,p} = 0$. (Support and positivity).

As seen B-splines are built with so called knots, t_i . Knots describe interval endpoints, but also multiplicity of such endpoints by repeating such a point. The interval endpoints are not necessarily uniform, e.g. $|t_i - t_{i+1}| = a, a \in \mathbb{R}, \forall i$. We make a clear distinction between knots with simple multiplicity and those with a higher multiplicity. Knots with simple multiplicity are considered unique and called breakpoints.

We use uniform spaced breakpoints $\{b_1, b_2, \dots, b_{I+1}\}$, with the begin and end breakpoint repeated in the knot vector. I is the total number of intervals in the interval $[b_1, b_{I+1}]$ These are repeated as many times, such that the curve has free endpoints. The number of repetitions depend on the degree of the B-splines. Without free endpoints, the freedom of fitting the solution to data is restricted, which is undesired.

A generic knot vector for a B-spline of degree p (or order $p + 1$) with free end-points has multiplicity $p + 1$ for the begin and end breakpoint.

Such a knot vector is $\{t_1, t_2, \dots, t_{I+1+2p}\}$, with $t_1 = \dots = t_{1+p} = b_1$ and $t_{I+1+p} = \dots = t_{I+1+2p} = b_{I+1}$. This vector is used to build the B-splines of degree p , see Definition 1.

Since B-splines are just polynomial functions a closed form of the integral of a B-spline exists. This definition will be used in a later section.

Definition 2 ([6] and [5, p. 127]).

$$\int_{t_1}^t \sum_{i=1}^n \phi_i B_{i,p}(x) dx = \sum_{i=1}^{s-1} \left(\sum_{j=1}^i \phi_j (t_{i+p+1} - t_j) / (p+1) \right) B_{i,p+1}(t),$$

here $t_1 \leq t \leq t_s$ and $s \geq 2$. Note that the same knot sequence is used in the case of degree p as in the case of degree $p+1$, this is important for getting the right area values. Furthermore t_1 means the left end of the first interval, since it is implied that we use B-splines with variable ends.

So we first define the one-dimensional representation for u ,

$$u(t) = \sum_{i=1}^{I+p} \theta_i B_{i,p}(t). \quad (11)$$

The multi-dimensional $\hat{u}(x, t)$ is discretised with help of the separability of space-time in the following manner,

$$\hat{u}(x, t) = \tilde{u}(x) u(t),$$

reuse (11),

$$\begin{aligned} &= \left(\sum_{k=1}^K \phi_k B_{k,0}(x) \right) \left(\sum_{i=1}^{I+p} \varphi_i B_{i,p}(t) \right). \\ &= \sum_{k,i} \theta_{k,i} B_{k,0}(x) B_{i,p}(t). \end{aligned} \quad (12)$$

Here $\theta_{k,i}$ represent the product between φ_i and ϕ_k . $B_{k,0}(x)$ is the multi-dimensional B-spline of degree 0 and as mention above Definition 1 this represents a product of multiple one-dimensional B-splines.

Inserting (12) into (9) and only computing the spatial integral yields,

$$D_{KL}(f; \mathcal{K}u) = \sum_k \left(\int_0^T \mathcal{K}_k u - f_k(t) \log(\mathcal{K}_k u) dt \right). \quad (13)$$

here $\mathcal{K}_k$ is a temporal functional resulting from integrating the block constant spatial B-splines. It should be straightforward that this is solving k one-dimensional problems like (8). With this reduction all multi-dimensional problems can be tackled, if computers can handle it. The choice for such simple B-splines in space, is that we deal with detectors that partition a region in space. That is a physical limitation, hence higher order spatial B-splines don't make a lot of sense. Especially, since we use synthetic data. Note that in the case $I = 1$ and $p = 0$ for (12) and inserted into (9), then (13) is equivalent to a discretised form of (10) or multiple (6).

We first continue with describing regularisation methods in the next subsection. After that we come back to (12) and (13) to make a system of equation such that we are able to minimize the variational problem.

2.4 Regularisation Methods

The description for the variation problem (4) has an extra term besides the data-term $D(\cdot)$, that is the regularisation term $R(u)$. This penalizes solutions for u , which means that the solution space for u becomes smaller. If the right property is penalized this can give a u with desired properties, but $R(u)$ has to be chosen in such a way that the problem is still convex or well-posed. We see every chosen $R(u)$ term as a different method.

The simplest method is when there is no $R(u)$. Adding the term $f \log(f) - f$ to (9) in the data-terms integral, gives the the Kullback-Leibler functional (14).

$$D_{KL}(f; \mathcal{K}u) = \int_0^T f \log\left(\frac{f}{\mathcal{K}u}\right) - f + \mathcal{K}u \, dt, \quad (14)$$

which is convex, see [3, p. 121] and [19]. Note that the convention $0 \log(0) = 0$ is used. Here f is considered the truth and $\mathcal{K}u$ the approximation. An intuitive explanation is the amount of information lost when $\mathcal{K}u$ is used to approximate f .

It is valid to add the extra term since (9) is minimised for u and not f , so it only adds constants to the problem. This is actually a generalization of the Kullback-Leibler entropy, which is a Bregman distance (divergence) with respect to the Boltzmann entropy, see [3].

We discuss two more regulariser, the first one is a simple temporal derivative

and the second one is the temporal second derivative. All three problems in a column, with a tag.

$$\arg \min_u D_{KL}(f; \mathcal{K}u) + \chi_{\{u \geq 0\}}. \quad (15)$$

$$\arg \min_u D_{KL}(f; \mathcal{K}u) + \frac{\alpha}{2} \|\partial_t u\|_2^2 + \chi_{\{u \geq 0\}}. \quad (16)$$

$$\arg \min_u D_{KL}(f; \mathcal{K}u) + \frac{\alpha}{2} \|\partial_t^2 u\|_2^2 + \chi_{\{u \geq 0\}}. \quad (17)$$

here $+\chi_{\{u \geq 0\}}$ is the characteristic function and the non-negativity constraint for u .

$$\chi_B(x) := \begin{cases} 0, & x \in B, \\ \infty, & x \notin B. \end{cases}$$

The regularisation in (16) enforces C^1 -continuity and (17) the C^2 -continuity also referred to as the curvature of a function. It is common, [13, 26], to use the curvature as regularisation term, hence the reason we use it. In the upcoming section the methods will be fully vectorized.

3 Numerics

We continue with the semi-discrete formulation of our data-term (13). Therefore the operator $\mathcal{K}$ from (3) has implicitly been simplified and we continue to assume $\mathcal{K}_k$ being constant. This means that $\mathcal{K}$ becomes a discretised spatial linear operator.

The assumption of $\mathcal{K}_k$ is done, because we are not sure if the formula for the integral Definition 2 is still valid. It is more efficient to use this formula to compute an exact value for the integral, then to approximate it. Efficiently in the sense of memory, but also implementation plays a huge role. We have not investigated possible extension for Definition 2.

Since $\mathcal{K}$ is now only a spatial operator, $\mathcal{K}_k$ is a constant and the temporal operator is the identity operator. Effectively that means our temporal problem is the simplest class of reconstruction, namely *denoising*. And so for the one-dimensional problem we will be only denoising. To keep it general we build the discrete form of the problem for the multiple dimensions.

We also start by using integration of (13) over the temporal B-splines.

$$\begin{aligned} D_{KL}(f; \mathcal{K}u) &= \sum_k \left(\int_0^T \mathcal{K}_k u - f_k(t) \log(\mathcal{K}_k u) dt \right), \\ &= \sum_k \left(\mathcal{K}_k \int_0^T u dt - \sum_n \log(\mathcal{K}_k u(w_n^k)) \right), \end{aligned}$$

now insert (11) and use Definition 2,

$$\begin{aligned} &= \sum_k \left(\mathcal{K}_k \int_0^T \sum_{i=1}^{I+p} \phi_i B_{i,p}(t) dt - \sum_n \log \left(\mathcal{K}_k \sum_{i=1}^{I+p} \phi_i B_{i,p}(w_n^k) \right) \right), \\ &= \sum_k \left(\int_0^T \sum_{i=1}^{I+p} \theta_{k,i} B_{i,p}(t) dt \right) - \sum_{k,n} \log \left(\sum_{i=1}^{I+p} \theta_{k,i} B_{i,p}(w_n^k) \right). \end{aligned} \tag{18}$$

In the last line $\theta_{k,i} = \mathcal{K}_k \phi_i$ and relates to the θ in (12). Both terms can be put in matrix-vector products, we first show this for the right-most term. We start with an one-dimensional example, e.g. $K = 1$.

$$\begin{aligned} &\sum_{1,n} \log \left(\sum_{i=1}^{I+p} \theta_{1,i} B_{i,p}(w_n^1) \right) \\ &= \sum_{1,n} \log \begin{pmatrix} B_{1,p}(w_1) & B_{2,p}(w_1) & \cdots & B_{I+p,p}(w_1) \\ B_{1,p}(w_2) & B_{2,p}(w_2) & \cdots & B_{I+p,p}(w_2) \\ \vdots & \vdots & & \vdots \\ B_{1,p}(w_N) & B_{2,p}(w_N) & \cdots & B_{I+p,p}(w_N) \end{pmatrix} \begin{pmatrix} \theta_{1,1} \\ \theta_{1,2} \\ \vdots \\ \theta_{1,I+p} \end{pmatrix}, \\ &= \sum \log(\mathcal{B}_{1,p} \boldsymbol{\theta}_{1,p}). \end{aligned} \tag{19}$$

In the last line the 1 means that we chose $K = 1$ as example. Note that the log is element-wise.

Now to go from one to more k terms is somewhat involved, because we kept to the convention of MATLABs column-major. Since we use a lot of MATLAB build-in functions this was a logical step, to reduce work on the code. We start with describing the solution vector $\boldsymbol{\theta}_p$,

$$\boldsymbol{\theta}_p := (\theta_{1,1} \quad \theta_{2,1} \quad \cdots \quad \theta_{K,1} \quad \theta_{K,2} \quad \cdots \quad \theta_{K,2} \quad \cdots \quad \theta_{1,I+p} \quad \cdots \quad \theta_{K,I+p})'$$

Here the accent, $'$, denotes the transpose of the vector. Next up is the spatial transformation matrix $\mathcal{K}$. A temporal slice $\boldsymbol{\theta}_{I+p}$ from $\boldsymbol{\theta}_p$ transformed by this $\mathcal{K}$ is written,

$$\tilde{\boldsymbol{\theta}}_{I+p} = \mathcal{K} \begin{pmatrix} \theta_{1,I+p} & \theta_{2,I+p} & \dots & \theta_{K,I+p} \end{pmatrix}'.$$

If we furthermore define $D_{i,p} = \text{diag}(B_{i,p}(w_1), B_{i,p}(w_2), \dots, B_{i,p}(w_N))$, then the right-most term from (19) can be written as,

$$\begin{aligned} & \sum_{k,n} \log \left(\sum_{i=1}^{I+p} \theta_{k,i} B_{i,p}(w_n^1) \right) \\ &= \sum_{k,n} \log \begin{pmatrix} D_{1,p} & & & \\ & D_{2,p} & & \\ & & \ddots & \\ & & & D_{I+p,p} \end{pmatrix} \begin{pmatrix} \mathcal{K} & 0 & \dots & 0 \\ 0 & \mathcal{K} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mathcal{K} \end{pmatrix} \boldsymbol{\theta}_p, \\ &= \sum \log(\mathcal{B}_p \mathcal{K} \boldsymbol{\theta}_p). \end{aligned} \quad (20)$$

This can be written in different ways, but since not all basis functions will have non-negative values, implementation-wise this is the most optimal for now. All these steps can also be done with the left-most terms in (19) and results in,

$$\sum_k \left(\int_0^T \sum_{i=1}^{I+p} \theta_{k,i} B_{i,p}(t) dt \right) = \bar{\mathcal{B}}_p \mathcal{K} \boldsymbol{\theta}_p.$$

Here $\bar{\mathcal{B}}_p$ is the operator computing the integral. Together with (20) this results in the matrix-vector minimisation problem,

$$\arg \min_{\boldsymbol{\theta}_p} \bar{\mathcal{B}}_p \mathcal{K} \boldsymbol{\theta}_p - \sum \log(\mathcal{B}_p \mathcal{K} \boldsymbol{\theta}_p) + \alpha R(\mathcal{K} \boldsymbol{\theta}_p). \quad (21)$$

It is necessary to define an approximation for the first and second derivative, due to (16) and (17). The numerical derivatives,

$$\begin{aligned} \partial_{t+} u_k &:= \begin{cases} \frac{u_{k,i+1} - u_{k,i}}{h} m, & \text{if } 1 \leq i < I+p, \\ 0, & \text{if } i = I+p. \end{cases} \\ \partial_t^2 u_k &:= \begin{cases} \frac{u_{k,i+2} - 2u_{k,i+1} + u_{k,i}}{h^2} m + u_{k,i}, & \text{if } 1 \leq i < I+p-1, \\ 0, & \text{if } i \geq I+p-1. \end{cases} \end{aligned}$$

Now we define matrix $\mathcal{D}^1$ as the counterpart ∂_{t+} and $\mathcal{D}^2$ for ∂_t^2 . Note that we have to compute our basisfunctions on multiple spots to us that as approximation for the derivative. We do this for a high number of spots, this number is fixed. Finally we can describe the methods (15), (16) and (17) in matrix-vector product that will be minimised with the primal-dual algorithm, see [4]. We use a customized toolbox *Flexbox* [8].

$$\arg \min_{\theta_p} \bar{\mathcal{B}}_p \mathcal{K} \theta_p - \sum \log(\mathcal{B}_p \mathcal{K} \theta_p) + \chi_{\{\theta_p \geq 0\}}. \quad (\text{KL})$$

$$\arg \min_{\theta_p} \bar{\mathcal{B}}_p \mathcal{K} \theta_p - \sum \log(\mathcal{B}_p \mathcal{K} \theta_p) + \frac{\alpha}{2} \sum \|\mathcal{D}^1 \theta_p\|_2^2 + \chi_{\{\theta_p \geq 0\}}. \quad (\text{KLL2})$$

$$\arg \min_{\theta_p} \bar{\mathcal{B}}_p \mathcal{K} \theta_p - \sum \log(\mathcal{B}_p \mathcal{K} \theta_p) + \frac{\alpha}{2} \sum \|\mathcal{D}^2 \theta_p\|_2^2 + \chi_{\{\theta_p \geq 0\}}. \quad (\text{KLL2nd})$$

Notice that I, p, α and $\mathbb{R}(u)$ are all free to pick, giving extra freedom. Would it be a non-uniform B-spline, then not only the length of that knot vector would matter, but the placing of the knots as well. Making it even more complex, which is a whole separate optimisation task and problem to solve. Note that the options for $\mathbb{R}(u)$ are also abundant.

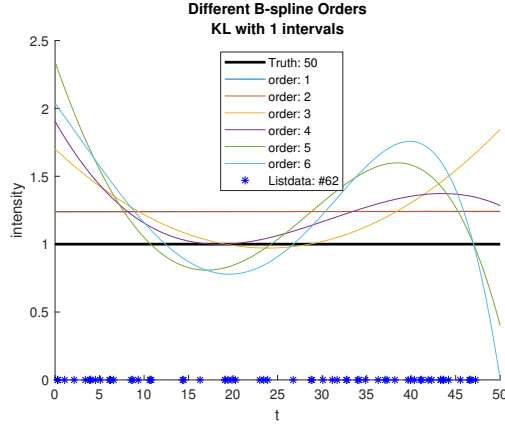
4 Results and Discussion

We start with the easiest function as source, a constant function. This is done as first step as a benchmark, then we investigate the shifted piecewise constant function to simulate the motion with a gap, see Figure 2. The approach to our goals is to look at the number of intervals in relation to the number of events that were registered in the experiment.

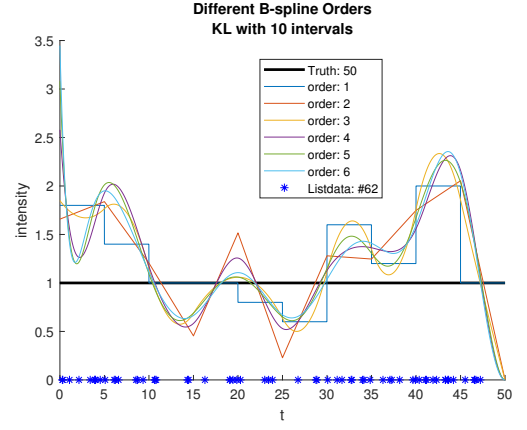
4.1 Constant Function

Every reconstruction in Figure 3 was done with a different number of intervals, but a fixed number of data points. No regularisation was used, hence $\lambda = 0$. The behaviour seen in Figure 3 is as expected for the constant function. The best approximation of the truth constant function is with a constant function, since a B-spline with $p = 0$ (degree) and $I = 1$ (interval). The higher orders look like their boundary basis function, which is not very surprising either.

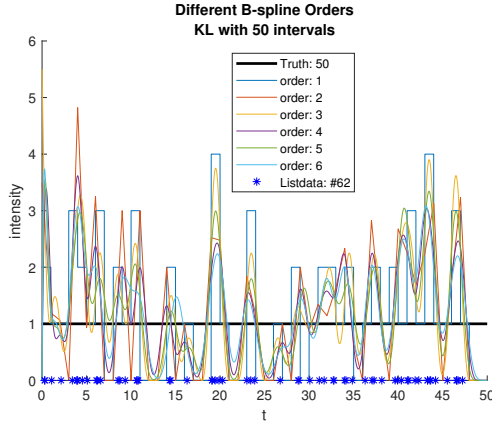
Increasing the number of intervals, i.e. taking smaller steps, for the approximation u to a continuous function would make the approximation more accurate. However, in minimisation problems the (discrete) data is the truth



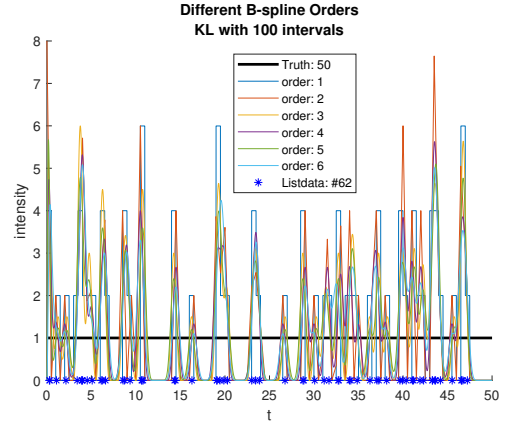
(a)



(b)



(c)



(d)

Figure 3: One realisation of minimising the data-term (KL) with different number of intervals/bins: (a) 1, (b) 10 (c) 50 and (d) 100. Every figure has a B-spline order 1 up to 6. The blue points indicate the location of a data points (63 data-points). The black line is the true intensity with expected number of 50.

that has to be fitted.

If the data is sparse enough or the number of intervals is high enough this means our u can approximate this sparsity with Dirac-delta-like peaks. While this is undesired, it is a good approximation of the data in the minimisation problem, see Figure 3d. Because of the B-spline representation, u has this freedom. That makes the freedom a weakness for a reconstruction u , depending on the number of intervals.

Apparent from Figure 3 is that a higher number of intervals cause oscillations, or rather the intervals without (many) events. This suggest a relationship between the number of events and the number of intervals. Also Furthermore, it is observable in Figure 3a and Figure 3b that the functions of the fourth-order and higher have very similar shapes. That implies that taking a order higher then four for splines is not necessary, because there is almost no distinction.

Very noticeable in Figure 3b is that the mean value of the intensity function is equal to the number of count divided by the time. We did a check for this figure and all intensities have the same behaviour. This is not very surprising, since we only solve the data-term from, (21), derived from a Poisson distribution. Or to put it differently the expected value (mean) is equal to the parameter $\Lambda(0, T) = \beta T$, a homogeneous intensity due to the mean value theorem. It does have something to do with the next figure. Figure 4 shows the relative L_1 -norm between the reconstruction and the truth, with number of events *vs.* number of intervals. The data in Figure 4 looks to suggest that the number 10^4 mentioned in subsection 1.4 is the number of counts per detector. In [17] there is mentioning of 10^4 , because with 10^4 counts, the relative Poisson variation is less then 0.1%. But these are results from a constant function in time that means, those values are nice, on average. With the results from Figure 4 and those references we think the low SNR is somewhere between 0 and 10^4 . Moreover, it is quite clear, at least for a constant function, the more events the better and how fewer intervals the better.

Now some a summary of number of event *vs.* number of interval for the method (KLL2) in The subfigers look very similar for the λ -values that we tried, between 10 and 500. The figure results suggest that 10^4 limit observed in Figure 4 has shifted here. Again it is obvious from the figures that more intervals give worse results. We plotted some instances from the results in Figure 5, on the upper row of Figure 6. The reconstruction are not only peaks, but it is also visible that more intervals doesn't do much, it makes

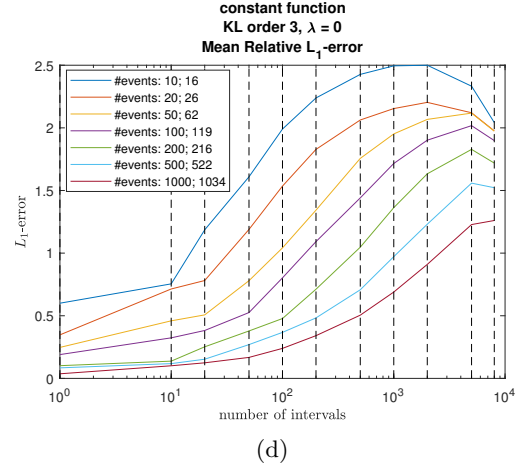
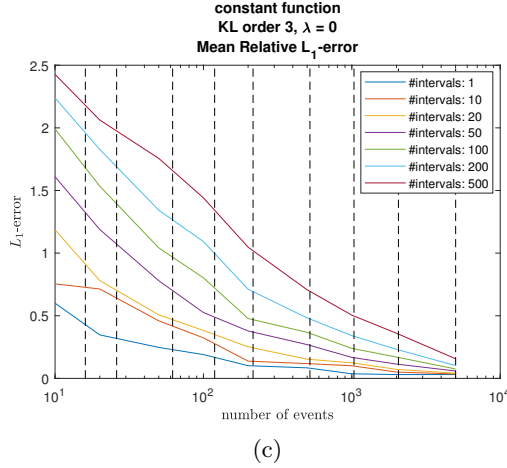
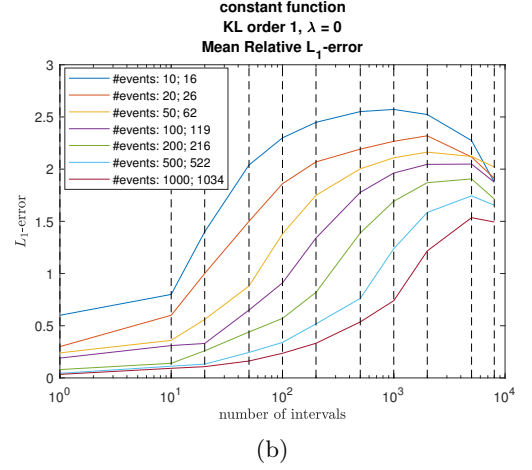
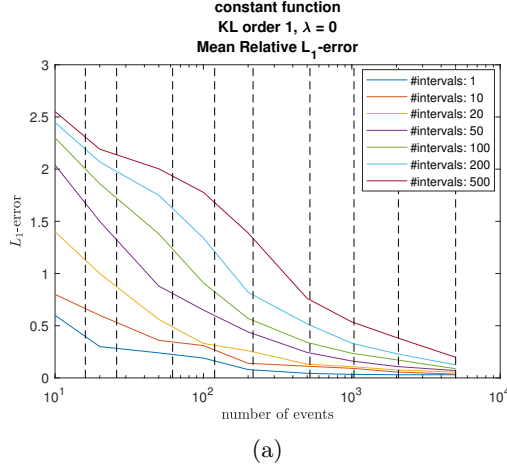
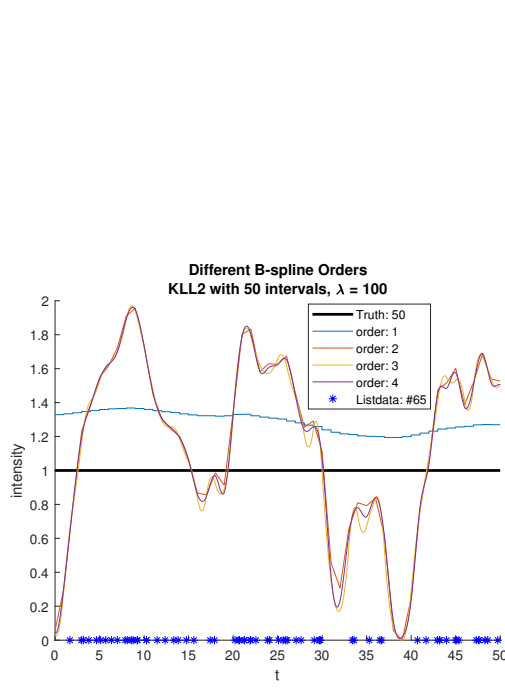
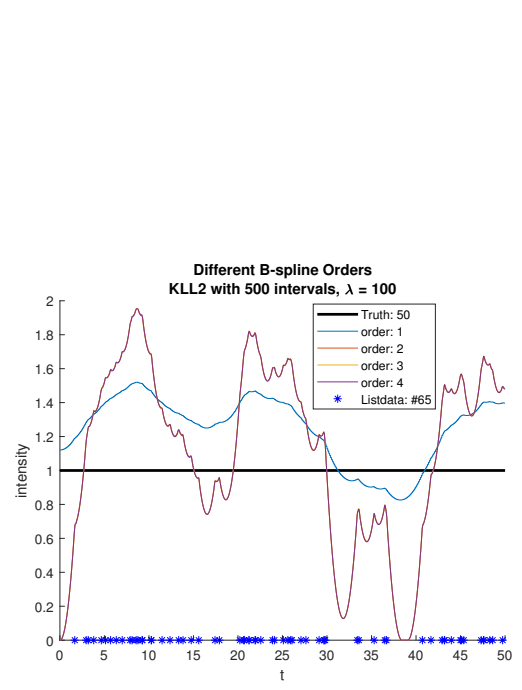


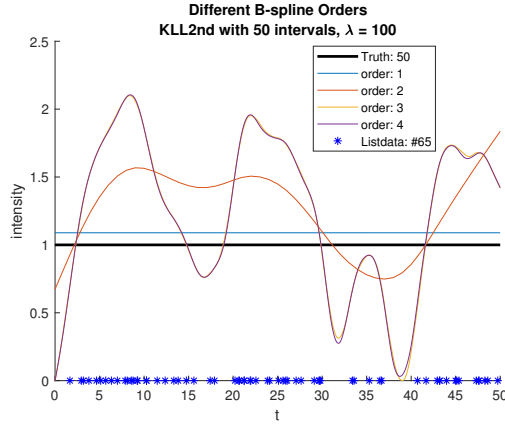
Figure 4: Relative L_1 -error figures, with the constant functions as reference. In (a) the vertical dashed lines are the realized number of events and in (d) the dashed lines represent the actual used number of intervals. In the legend with two separate numeric values, the left value is the analytical expected number and the right value the realized number of events.



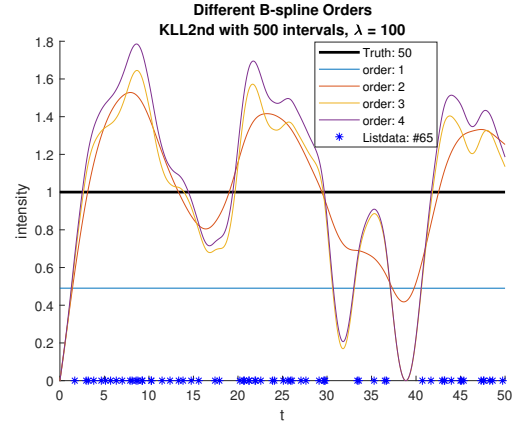
(a)



(b)



(c)



(d)

Figure 5: Top row (KLL2) and bottom row (17).

the intensity look more jagged, but the overall shape is the same. It is also visible that there are zero-trapped intervals, meaning that too many intervals next to each other are zero. Although, we have a solution that is more smooth, it is still noisy (over time, e.g. a flickering film screen). To put it differently C^1 -continuity is not enough to make this signal less noisy.

The bottom row in Figure 6 contains the solutions from (16), which has the C^2 -continuity constraint and the result look already much more smooth, even the zero-trapped intensities. It is notable that even though there is more than one interval, but the order 1 reconstruction looks like it has only one interval. This is due to the C^2 -continuity constraint. Furthermore, the mean value of the intensity function is not equal to the number of event divided by the time, because of regularisation. The snapshots at the bottom row of Figure 5 correspond with the aggregated information in Figure 7. The left column shows the same kind of structure as in Figure 5, but then zoomed in

4.2 Piecewise Function

The shifted functions shows that the B-splines of higher order are actually really good at capturing shifting peaks. Their suitability to these kind of functions is observable in for example Figure 4. In overview Figure 9, where Figure 8 is part of, one can clearly see that the number of intervals must be low for a small error. Another notable result is Figure 9a, since it has a sawtooth error or is somewhat constant. This sawtooth shape could be due to the Nyquist frequency. the function is shifted with 2.5 temporal stepsize and the number of intervals is of length 5. The constant errors are due to this function's analytical expected value to be constant for every gap size.

5 Conclusion

In this paper we discussed dynamic objects and the problem of capturing these with ECT. We modelled this as a one-dimensional problem by looking at a gap between two peaks in time. We believe that a better understanding of the one-dimensional problem, for example with a low number of counts can improve the reconstruction dynamic objects in ECT. To solve the minimisation problem we used the of Poisson processes to derive a variational problem. This was solved with an extended primal-dual toolbox *Flexbox*.

It is now clear for me what low SNR is and that extremely low SNR in

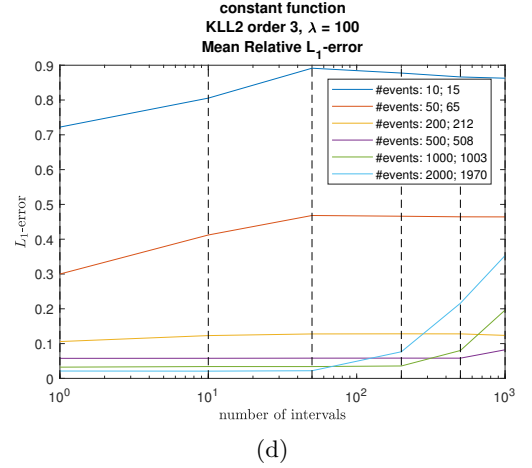
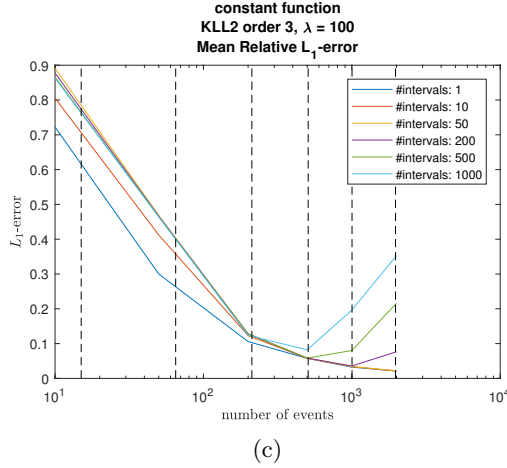
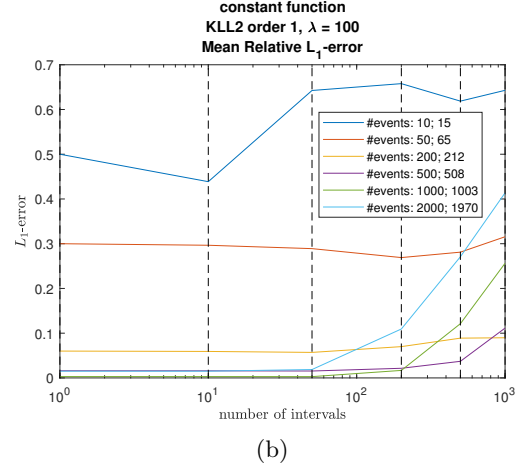
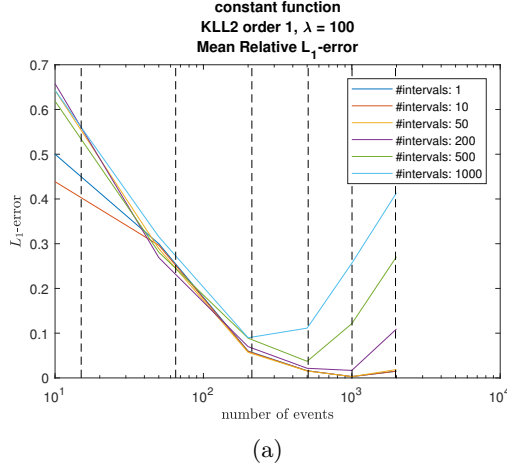


Figure 6: Relative L_1 -error figures, with the constant functions as reference. In (a) the vertical dashed lines are the realized number of events and in (c) the dashed lines represent the actual used number of intervals. In the legend with two separate numeric values, the left value is the analytical expected number and the right value the realized number of events.

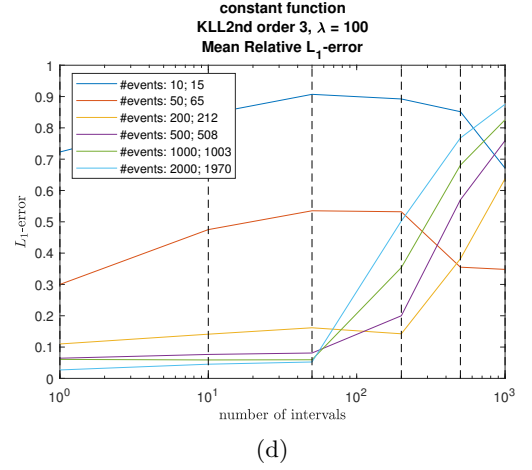
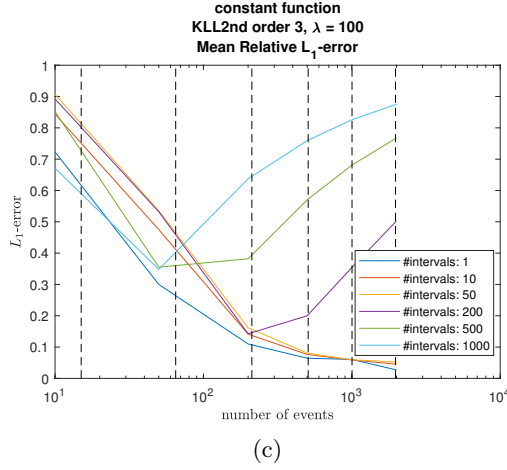
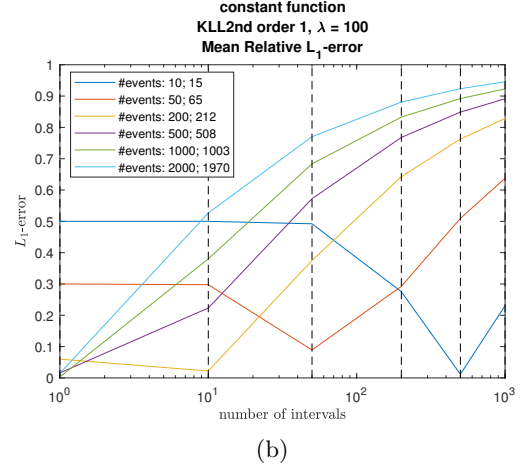
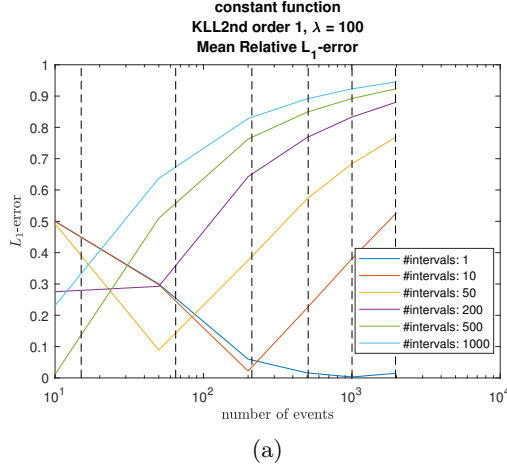
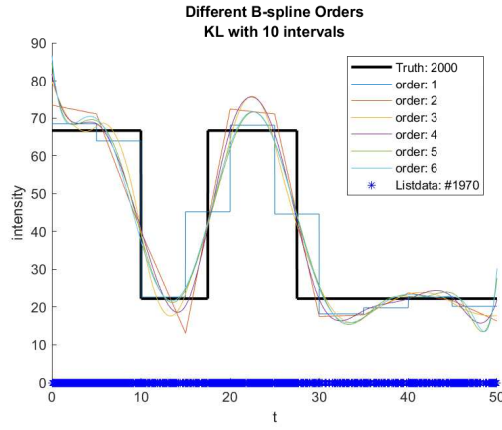
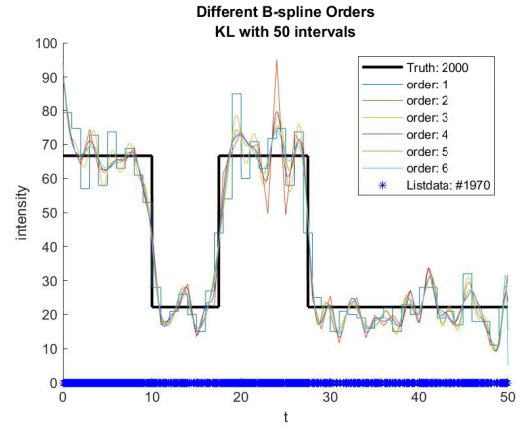


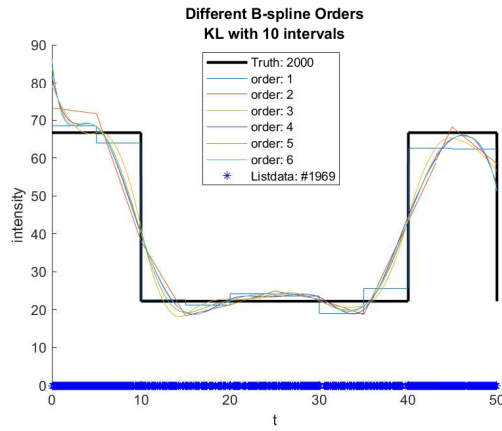
Figure 7: Relative L_1 -error figures, with the constant functions as reference. In (a) the vertical dashed lines are the realized number of events and in (c) the dashed lines represent the actual used number of intervals. In the legend with two separate numeric values, the left value is the analytical expected number and the right value the realized number of events.



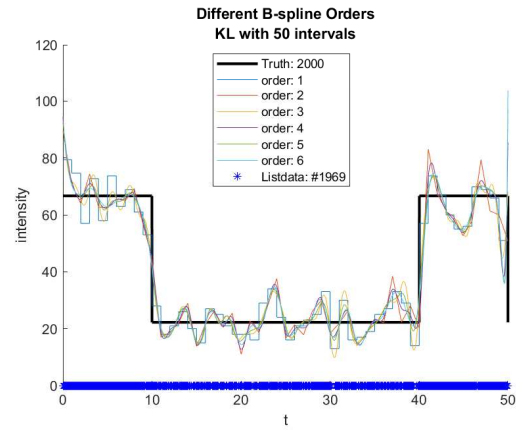
(a)



(b)

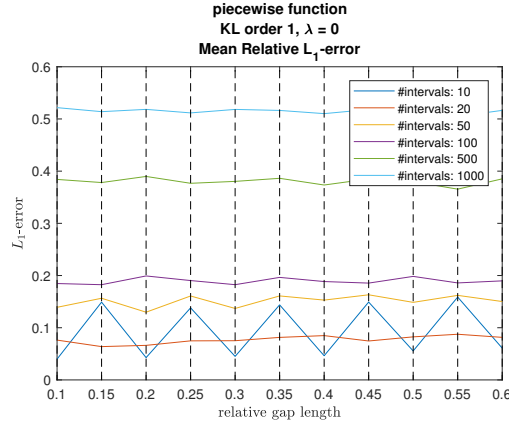


(c)

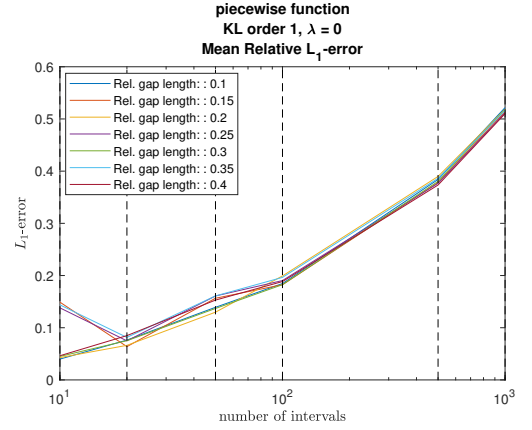


(d)

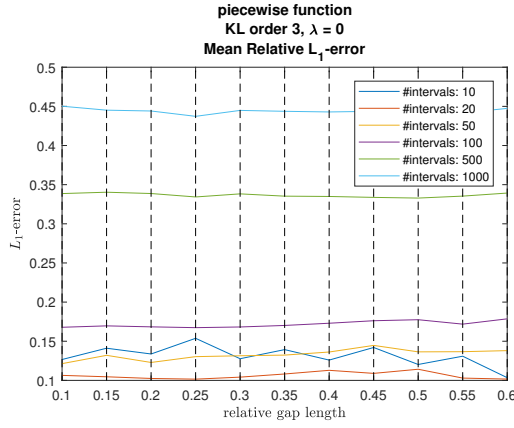
Figure 8: Shifted functions



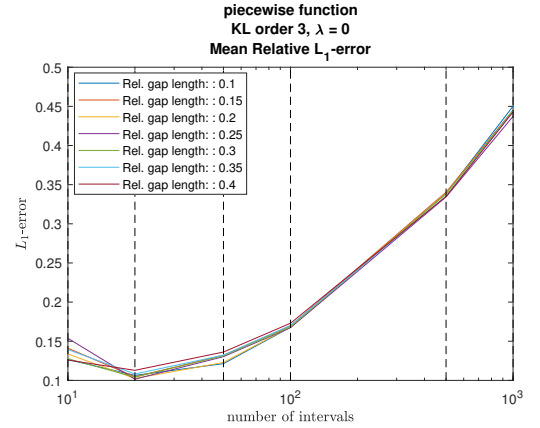
(a)



(b)



(c)



(d)

Figure 9: Relative L_1 -error figures, with the constant functions as reference. In (a) the vertical dashed lines are the realized number of events and in (c) the dashed lines represent the actual used number of intervals. In the legend with two separate numeric values, the left value is the analytical expected number and the right value the realized number of events.

the region $[0, 10^3]$ lies. Furthermore, I could not find any indication of a good ratio between the number of counts and the number of intervals. The best rule is, few intervals (less than 50, with a timespan of 50) and as much counts as possible. The best direction would be non-uniform B-splines although, the placement of knots is a problematic case as well.

6 Appendix

6.1 Specific Mathematics

Definition 3 ([3, p. 33]). (Image). Let $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$ be an image domain. A function $u : \Omega \rightarrow \mathbb{R}$ is called a d -dimensional image if the following conditions are fulfilled,

1. u has a compact support, if Ω is not bounded.
2. $0 \leq u(x) < \infty$, $\forall x \in \Omega$.
3. $\int_{\Omega} u(x) dx \leq \infty$.

This is easily extended for multiple channels (e.g. colour images).

Definition 4 ([3, p. 33]). (Image sequence, video). Let $\Omega_T = \Omega \times [0, T]$ be a space-time cylinder. A function $u : \Omega_T \rightarrow \mathbb{R}$ is called a d -dimensional image sequence if it is a d -dimensional image for every fixed $t \in [0, T]$ and $T > 0$.

Definition 5 ([21, p. 312]). A stochastic process $\{N_t, t \geq 0\}$ is said to be a *counting process* if N_t represents the total number of ‘events’ that occur by time t and satisfies the properties:

1. $N_t \geq 0$.
2. $N_t \in \mathbb{N}$.
3. If $s < t$, then $N_s \leq N_t$.
4. For $s < t$, $N_t - N_s$ equals the number of events that occurred in the interval $(s, t]$,

Definition 6 ([21, p. 313]). The counting process $\{N_t, t \geq 0\}$ is said to be a *homogeneous Poisson process* having rate λ , $\lambda > 0$, if

1. $N_0 = 0$.
2. The process has independent increments.

3. The number of events in any interval of length t is Poisson distributed with mean λt . That is, for all $s, t \geq 0$

$$P(\{N_{t+s} - N_s = n\}) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, n = 0, 1, \dots$$

Note that it follows from the last condition that a Poisson process has stationary increments and also that $\mathbb{E}[N_t] = \lambda t$.

Definition 7 ([21, p. 314]). The counting process $\{N_t, t \geq 0\}$ is said to be a *homogeneous Poisson process* having rate $\lambda > 0$, if

1. $N_0 = 0$.
2. The process has stationary and independent increments.
3. $P(\{N_h = 1\}) = \lambda h + o(h)$.
4. $P(\{N_h \geq 2\}) = o(h)$.

With function $f(\cdot)$ is said to be $o(h)$ if: $\lim_{h \rightarrow 0} f(h)/h = 0$.

Proposition 7.1 ([21, p. 317]). $T_n, n = 1, 2, \dots$, the *interarrival time* between the $(n-1)$ -st and n -th event, are independent identically distributed exponential random variables having mean $1/\lambda$.

It follows that $S_n = \sum_{i=1}^n T_i$, the *waiting time* until the n -th event, has a gamma distribution with parameters n and λ ,

$$p_{S_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, t \geq 0.$$

Definition 8 ([21, p. 339]). The counting process $\{N_t, t \geq 0\}$ is said to be a *nonhomogeneous Poisson process* (NHPP)⁵ with intensity function $\lambda(t)$, $t \geq 0$, if

1. $N_0 = 0$.
2. $\{N_t, t \geq 0\}$ has independent increments.
3. $P(\{N_{t+h} - N_t = 1\}) = \lambda(t)h + o(h)$.
4. $P(\{N_{t+h} - N_t \geq 2\}) = o(h)$.

More specifically if $\Lambda(a, b) = \int_a^b \lambda(t) dt$, then

$$\{N_{t+s} - N_t = n\} \sim \text{Poisson}(n; \Lambda(t, t+s)). \quad (22)$$

⁵Also referred to as *inhomogeneous* or *non-stationary* in literature.

Proposition 8.1 ([21, p. 342]). Let $\{N_t, t \geq 0\}$ and $\{M_t, t \geq 0\}$, be independent NHPP, with respective intensity functions $\lambda(t)$ and $\mu(t)$, and let $N_t^* = N_t + M_t$. Then, the following statements are true.

1. $\{N_t^*, t \geq 0\}$ is a NHPP with intensity function $\lambda(t) + \mu(t)$.
2. Given that an event of the $\{N_t^*\}$ process occurs at time t then, independent of what occurred prior to t , the event at t was from the $\{N_t\}$ process with probability $\frac{\lambda(t)}{\lambda(t) + \mu(t)}$.

Proposition 8.2. For a NHPP with rate $\lambda(t) > 0$, the conditional density function (cdf) of next event's occurrence time w_{n+1} , with knowledge of the past event's occurrence time w_n , is $p(t|W_n = w_n) = \lambda(t) \exp(-\Lambda(w_n, t))$.

Proof. We first define the interarrival times $T_{n+1} = W_{n+1} - W_n$ for $n = 0, 1, \dots$ here $W_0 = 0$.

$$\begin{aligned}
P(W_{n+1} \leq t | W_n = w_n) &= P(T_{n+1} < x | W_n = w_n), \\
&= P(\{N_{w_n+x} - N_{w_n} \geq 1\} | W_n = w_n), \\
&= 1 - P(\{N_{w_n+x} - N_{w_n} = 0\}), \text{ use (22)} \\
&= 1 - \exp\left(-\int_{w_n}^t \lambda_\tau d\tau\right). \tag{23}
\end{aligned}$$

Taking the derivative with respect to t gives the conditional probability density function. $\square$

Theorem 1 ([1]). Let N be a NHPP, suppose $\Lambda(0, t)$ is continuous, and define

$$M_t(\omega) = N_{\tau(t)}(\omega)$$

with $\tau(t) := \inf\{s : \Lambda(0, s) > t\}$ for all $t \geq 0$ and $\omega \in \Omega$. Then $M = \{M_t; t \geq 0\}$ is a homogeneous Poisson process with rate 1.

References

- [1] E. Çinlar and N. J. Sollenberger. Introduction to stochastic processes. Prentice-Hall, New Jersey, 1975.

- [2] Harrison H. Barrett, Timothy White, and Lucas C. Parra. List-mode likelihood. J. Opt. Soc. Am. A, 14(11):2914–2923, Nov 1997.
- [3] Christoph Brune. 4D Imageing in Tomography and Optical Nanoscopy. PhD thesis, Department of Mathematics and Computer Science, University of Münster, 2010.
- [4] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision, 40(1):120–145, May 2011.
- [5] Carl de Boor. A Practical Guide to Splines. Springer-Verlag New York, New York, rev. ed. edition, 2001.
- [6] Carl de Boor, Tom Lyche, and Larry L. Schumaker. On Calculating with B-Splines II. Integration, pages 123–146. Birkhäuser Basel, Basel, 1976.
- [7] Qiaoqiao Ding, Martin Burger, and Xiaoqun Zhang. Dynamic SPECT reconstruction with temporal edge correlation. Inverse Problems, 34(1):014005, jan 2018.
- [8] Hendrik Dirks. A flexible primal-dual toolbox. ArXiv e-prints, March 2016. arXiv:1603.05835.
- [9] Richard P. (Richard Phillips) Feynman, Robert B. (Robert Benjamin) Leighton, and Matthew Linzee. Sands. The Feynman lectures on physics / Vol. III, Quantum mechanics. Addison-Wesley, Reading, MA :, 3rd edition, 1965.
- [10] Quanzheng Li, E. Asma, and R. M. Leahy. A fast fully 4d incremental gradient reconstruction algorithm for list mode pet data. In 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821), pages 555–558 Vol. 1, April 2004.
- [11] M.N.M. van Lieshout. Spatial point patterns, Chapter 3: Point processes. lecture notes Mastermath, 2018.
- [12] Rodney. Loudon. The quantum theory of light. Oxford University Press, Oxford :, 3rd edition, 2000.
- [13] T. E. Nichols, Jinyi Qi, E. Asma, and R. M. Leahy. Spatiotemporal reconstruction of list-mode pet data. IEEE Transactions on Medical Imaging, 21(4):396–404, April 2002.

- [14] Thomas E. Nichols, Jinyi Qi, and Richard M. Leahy. Continuous time dynamic pet imaging using list mode data. In Attila Kuba, Martin Šáamál, and Andrew Todd-Pokropek, editors, Information Processing in Medical Imaging, pages 98–111, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.
- [15] Philips Healthcare. CardioMD IV. [https://images.philips.com/is/image/philipsconsumer/24b25bdc2d904d7eb2f1a873010de1a7?wid=870&hei=490&\\$pnglarge\\$](https://images.philips.com/is/image/philipsconsumer/24b25bdc2d904d7eb2f1a873010de1a7?wid=870&hei=490&$pnglarge$). Web; accessed 30-04-2018.
- [16] Philips Healthcare. Vereos Digital PET/CT. [https://images.philips.com/is/image/philipsconsumer/cf2bf0392b774bfffacc5a77c0160063a?wid=870&hei=490&\\$pnglarge\\$](https://images.philips.com/is/image/philipsconsumer/cf2bf0392b774bfffacc5a77c0160063a?wid=870&hei=490&$pnglarge$). Web; accessed 30-04-2018.
- [17] Jerry L. Prince and Jonathan M. Links. Medical imaging signals and systems. Pearson Prentice Hall, 1st edition, 2006.
- [18] Andrew J Reader and Jeroen Verhaeghe. 4d image reconstruction for emission tomography. Physics in Medicine & Biology, 59(22):R371, 2014.
- [19] Elena Resmerita and Robert S. Anderssen. Joint additive kullback-leibler residual minimization and regularization for linear inverse problems. Mathematical Methods in the Applied Sciences, 30(13):1527–1544, 2007.
- [20] B. W. Reutter, G. T. Gullberg, and R. H. Huesman. Direct least-squares estimation of spatiotemporal distributions from dynamic spect projections using a spatial segmentation and temporal b-splines. IEEE Transactions on Medical Imaging, 19(5):434–450, May 2000.
- [21] Sheldon M. Ross. {CHAPTER} 5 - the exponential distribution and the poisson process. In Sheldon M. Ross, editor, Introduction to Probability Models (Tenth Edition), pages 291 – 370. Academic Press, Boston, tenth edition edition, 2010.
- [22] A. Sawatzky, C. Brune, F. Wubbeling, T. Kosters, K. Schafers, and M. Burger. Accurate em-tv algorithm in pet with low snr. In 2008 IEEE Nuclear Science Symposium Conference Record, pages 5133–5137, Oct 2008.

- [23] L. A. Shepp and Y. Vardi. Maximum Likelihood Reconstruction for Emission Tomography. IEEE Transactions on Medical Imaging, 1(2):113–122, 1982.
- [24] D. L. Snyder. Parameter estimation for dynamic studies in emission-tomography systems having list-mode data. IEEE Transactions on Nuclear Science, 31(2):925–931, April 1984.
- [25] Donald L. Snyder and Michael I. Miller. Random Point Processes in Time and Space. Springer-Verlag New York, 2 edition, 1991.
- [26] J. Verhaeghe, Y. D’Asseler, S. Staelens, S. Vandenberghe, and I. Lemahieu. Reconstruction for gated dynamic cardiac pet imaging using a tensor product spline basis. IEEE Transactions on Nuclear Science, 54(1):80–91, Feb 2007.
- [27] J. Verhaeghe, Y. D’Asseler, S. Vandenberghe, S. Staelens, R. Van de Walle, and I. Lemahieu. Ml reconstruction from dynamic list-mode pet data using temporal splines. In IEEE Symposium Conference Record Nuclear Science 2004., volume 5, pages 3146–3150 Vol. 5, Oct 2004.
- [28] Jeroen Verhaeghe, Yves D’Asseler, Steven Staelens, and Ignace Lemahieu. Optimization of temporal basis functions in dynamic PET imaging. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 569(2):425–428, dec 2006.